

V. V. S. S. S. Chakravarthy ·
Vikrant Bhateja · Wendy Flores Fuentes ·
Jaume Anguera ·
K. Padma Vasavi *Editors*

Advances in Signal Processing, Embedded Systems and IoT

Proceedings of Seventh ICMEET-2022

Lecture Notes in Electrical Engineering

Volume 992

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
Yong Li, Hunan University, Changsha, Hunan, China
Qilian Liang, Department of Electrical Engineering, University of Texas at Arlington, Arlington, TX, USA
Ferran Martín, Departament d'Enginyeria Electrònica, Universitat Autònoma de Barcelona, Bellaterra, Barcelona, Spain
Tan Cher Ming, College of Engineering, Nanyang Technological University, Singapore, Singapore
Wolfgang Minker, Institute of Information Technology, University of Ulm, Ulm, Germany
Pradeep Misra, Department of Electrical Engineering, Wright State University, Dayton, OH, USA
Sebastian Möller, Quality and Usability Laboratory, TU Berlin, Berlin, Germany
Subhas Mukhopadhyay, School of Engineering and Advanced Technology, Massey University, Palmerston North, Manawatu-Wanganui, New Zealand
Cun-Zheng Ning, Electrical Engineering, Arizona State University, Tempe, AZ, USA
Toyooki Nishida, Graduate School of Informatics, Kyoto University, Kyoto, Japan
Luca Oneto, Department of Informatics, Bioengineering, Robotics and Systems Engineering, University of Genova, Genova, Genova, Italy
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 and 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
Walter Zamboni, DIEM—Università degli studi di Salerno, Fisciano, Salerno, Italy
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. ****

V. V. S. S. S. Chakravarthy · Vikrant Bhateja ·
Wendy Flores Fuentes · Jaime Anguera ·
K. Padma Vasavi
Editors

Advances in Signal Processing, Embedded Systems and IoT

Proceedings of Seventh ICMEET-2022

 Springer

Editors

V. V. S. S. Chakravarthy
Department of Electronics
and Communication Engineering
Raghu Institute of Technology
Visakhapatnam, India

Vikrant Bhateja
Department of Electronics Engineering
Faculty of Engineering and Technology
Veer Bahadur Singh Purvanchal University
Jaunpur, Uttar Pradesh, India

Wendy Flores Fuentes
Autonomous University of Baja California
Mexicali, Baja California, Mexico

Jaume Anguera
Department of Electronics
and Telecommunication
Universitat Ramon Llull
Barcelona, Spain

K. Padma Vasavi
Shri Vishnu Engineering College for
Women (A)
Bhimavaram, Andhra Pradesh, India

ISSN 1876-1100

ISSN 1876-1119 (electronic)

Lecture Notes in Electrical Engineering

ISBN 978-981-19-8864-6

ISBN 978-981-19-8865-3 (eBook)

<https://doi.org/10.1007/978-981-19-8865-3>

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

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

Committee

Chief Patron

Sri K. V. Vishnu Raju, Chairman, SVES

Co-chief Patrons

Sri Ravichandran Rajagopal, Vice Chairman, SVES

Sri K. Aditya Vissam, Secretary, SVES

Patrons

Dr. G. Srinivasa Rao, Principal, SVECW

Dr. P. Srinivasa Raju, Vice Principal, SVECW

Convener

Dr. K. Padma Vasavi, Professor and HOD, ECE, SVECW

General Chairs

Prof. Jaume Anguera, Ph.D., Universitat Ramon Llull, Spain

Dr. Vikrant Bhateja, Department of Electronics Engineering, Faculty of Engineering and Technology, Veer Bahadur Singh Purvanchal University, Jaunpur, Uttar Pradesh, India

Program Chairs

Prof. P. Satish Rama Chowdary, Vice Principal and HoD ECE, Raghu Institute of Technology, Vizag

Dr. V. V. S. S. S. Chakravarthy, Professor, Department of ECE, Raghu Institute of Technology, Vizag

Technical Advisory Committee

Prof. Xin-She Yang (Middlesex University, London, UK)

Celia Shahnaz, Ph.D., SMIEEE, FIEB, Professor, Department of EEE, BUET, Bangladesh

Prof. Jaume Anguera, Universitat Ramon Llull, Spain

Dr. Santosh Pande, Georgia Tech, Atlanta, GA

Dr. Windhya Rankothge, SLIIT, Sri Lanka

Dr. K. M. Prasad, Senior R&D Engineer, ENGINIA RESEARCH Inc., Canada

Dr. Steven Fernandes, Assistant Professor of Computer Science, Creighton University, NE, USA

Dr. G. S. N. Raju, Vice Chancellor, CUTMAP

Dr. N. V. S. N. Sarma, Director, IIIT, Tiruchirappalli

Dr. C. P. Ravi Kumar, Director, Texas Instruments India

Dr. Sudeb Das Gupta, Professor, IIT Roorkee

Dr. Anil Kumar Vuppala, IIIT, Hyderabad

Dr. Raghvendra Kumar Chaudhary, IIT Dhanbad

Dr. R. Nakkeeran, Pondicherry Central University

Dr. D. C. Pande, Former Scientist DRDO, LRDE, Bangalore

Dr. V. Vakula, NIT, Warangal

Dr. Suresh Chandra Satapathy, School of Computer Engineering, KIIT, Bhubaneswar, Odisha.

Dr. P. Rajesh Kumar, Professor and Head, ECE, AUCE, Andhra University

Dr. N. Balaji, Director, IQAC, JNTUK, Kakinada

Dr. B. T. Krishna, Professor and Head, ECE Department, JNTUK, Kakinada

Organizing Chair

Dr. K. Padma Vasavi, Professor and HOD, ECE, SVECW

Organizing Committee

Dr. T. Sudheer Kumar
Dr. S. Hanumantha Rao
Dr. G. R. L. V. N. Srinivasa Raju
Dr. M. Pradeep
Dr. K. Pushpa
Dr. M. Prema Kumar

Technical Program Committee

Dr. T. Sudheer Kumar, Associate Professor, ECE, SVECW
Celia Shahnaz, Ph.D., SMIEEE, FIEB, Professor, Department of EEE, BUET, Bangladesh
Dr. N. Balaji, Director, IQAC, JNTUK, Kakinada
Dr. B. T. Krishna, Professor and Head, ECE, JNTUK, Kakinada
Dr. M. V. Subba Rao, Associate Professor, ECE, SVECW
Mr. T. Sairam Vamsi, Associate Professor, ECE, SVECW
Dr. P. Bramhanandam, Professor, Physics, SVECW
Dr. Gaurav Trivedi, Associate Professor, IIT Guwahati

Publication Committee

Dr. S. Hanumantha Rao, Professor, ECE, SVECW
Dr. N. Prasad, Associate Professor, ECE, SVECW
Dr. R. Sahoo, Assistant Professor, ECE, SVECW

Hospitality Committee

Dr. G. R. L. V. N. Srinivasa Raju, Professor, ECE, SVECW
Mr. P. Ravi Kumar, Associate Professor, ECE, SVECW
Mr. K. S. N. Raju, Assistant Professor, ECE, SVECW

Mr. G. Challaram, Assistant Professor, ECE, SVECW
Dr. R. Viswanadham, Assistant Professor, ECE, SVECW
Mr. D. Ramesh Varma, Assistant Professor, ECE, SVECW

Finance Committee

Dr. M. Pradeep, Associate Professor, ECE, SVECW
Mr. A. Narayana Kiran, Assistant Professor, ECE, SVECW

Media and Publicity Committee

Dr. K. Pushpa, Professor, ECE, SVECW
Dr. M. V. Ganeswara Rao, Associate Professor, ECE, SVECW
Mr. E. R. Praveen Kumar, Assistant Professor, ECE, SVECW
Mr. D. Ramesh Varma, Assistant Professor, ECE, SVECW

Website Committee

Dr. A. Sri Krishna, Associate Professor and HOD, AI, SVECW
Dr. R. Viswanadham, Assistant Professor, ECE, SVECW
Dr. R. Sahoo, Assistant Professor, ECE, SVECW

Registration and Invitation

Dr. M. Prema Kumar, Associate Professor, ECE, SVECW
Ms. T. Pavani, Associate Professor, ECE, SVECW
Mr. A. Narayana Kiran, Assistant Professor, ECE, SVECW
Mr. M. A. S. Mohan Raju, Assistant Professor, ECE, SVECW

Venue

Mr. V. Srinivasa Rao, Associate Professor, ECE, SVECW
Mr. K. Murthy Raju, Associate Professor, ECE, SVECW
Mr. D. Murali Krishna, Associate Professor, ECE, SVECW
Mr. M. Padmanabha Raju, Assistant Professor, ECE, SVECW

Mr. D. Gireesh Kumar, Assistant Professor, ECE, SVECW

Mr. S. Teja, Assistant Professor, ECE, SVECW

Exhibition/Stalls

Dr. S. Hanumantha Rao, Professor, ECE, SVECW

Mr. D. Ramesh Varma, Assistant Professor, ECE, SVECW

Special Sessions

Dr. Rohit Anand, G. B. Pant DSEU Okhla-1 Campus (formerly GBPEC), Government of NCT of Delhi, New Delhi, India

Dr. N. Udaya Kumar, SRKR Engineering College, Bhimavaram

Dr. Ramesh Babu B. S. S. V., Raghu Institute of Technology, Visakhapatnam, India

Dr. Nidhi Sindhwani, Amity Institute of Information Technology, Amity University, Noida, Uttar Pradesh, India

Dr. M. Vamshi Krishna, Dhanekula Institution of Engineering and Technology, Vijayawada

Preface

The book titled *Advances in Signal Processing, Embedded Systems and IoT: Proceedings of Seventh ICMEET-2022* compiles the papers presented at the Seventh International Conference on Microelectronics, Electromagnetics, and Telecommunication (ICMEET-2022). Before this, six versions of the Conference with the title ICMEET were organized consecutively from 2015 to 2021. Every time, the proceedings is published in Lecture Notes in Electrical Engineering Series of Springer and indexed in SCOPUS. The Seventh ICMEET was organized by Shri Vishnu Engineering College for Women (A), Bhimavaram, Andhra Pradesh, India, during July 22–23, 2022. The Conference received 315 manuscripts which were thoroughly reviewed, and 61 papers were selected for presentation at the Conference, maintaining an acceptance ratio below 20%. The Conference was organized in hybrid mode, in compliance with COVID restrictions. The Conference is power packed with a couple of keynotes and two invited lectures. Dr. C. P. Ravi Kumar, Texas Instrument, India, delivered the first keynote address on the first day of the Conference. Following this, the second keynote address is given by Dr. Jaume Anguera, IEEE Fellow, Founder and CTO at Ignion, Spain, on the topic of “Antenna Booster Technology: From Fundamentals to Applications”. The Conference also featured the first invited talk by Dr. Lakshmi Narayana, NERTU, on “Navigation of Natural Flyers and Autonomous Navigation for UAVs and MAVs”. Dr. Celia Shahnaz delivered the other invited talk on “Deep Learning Applications for Medical Image Analysis”.

The Conference aimed to provide the latest research and technological updates in machine learning, artificial intelligence, and other novel techniques to signal processing, telecommunication, Internet of things, VLSI, and embedded Systems. The theme of the Conference covers significant areas of the latest technology like nanotechnology, electronic devices, 5G communication systems, and schemes. The Conference provides an annual International forum for presenting and discussing recent advances in microelectronics, electromagnetics, and telecommunications. It brings together leading international researchers, engineers, and practitioners interested in related technologies, which is the significant interest behind organizing Conference.

We want to thank Chief Patron Sri K. V. Vishnu Raju (Chairman, SVES), Co-chief Patrons Sri Ravichandran Rajagopal (Vice Chairman) and Sri Aditya Vissam (Secretary of SVES), and Patrons Dr. G. Srinivasa Rao (Principal, SVECW) and Dr. P. Srinivasa Raju (Vice Principal, SVECW) for support and commitment to organize the Seventh ICMEET in their campus. We also congratulate all the heads and faculty members of all the departments at SVECW for a successful International Conference on their campus. Special thanks to all the session chairs, co-chairs, track managers, and reviewers for their excellent support. Last but certainly not least, our sincere thanks go to all the authors who submitted papers and all the presenters for their contributions and fruitful discussions that made this Conference a great success.

Visakhapatnam, Andhra Pradesh, India
Jaunpur, Uttar Pradesh, India
Mexicali, Mexico
Barcelona, Spain
Bhimavaram, Andhra Pradesh, India

Dr. V. V. S. S. S. Chakravarthy
Dr. Vikrant Bhateja
Dr. Wendy Flores Fuentes
Dr. Jaume Anguera
Dr. K. Padma Vasavi

Contents

Interference and Cancellation Issues for an Indoor VLC Network	1
Himani Sharma and Rakesh Kumar Jha	
FPGA-Based 128-Bit RISC Processor Using Pipelining	11
T. Subhashini, M. Kamaraju, and K. Babulu	
Analysis and Implementation of a Multi-path and Metal-Stacked 8-Shaped Inductor	27
Keerthi Somraj and P. Akhendra Kumar	
Improved Conversion Gain with High SFDR and Highly Linear RF Mixer Using Inductive Gate Biasing Technique for Low Power WAS and Radio LAN Applications	37
Avvaru Subramanyam and R. V. S. Satyanarayana	
Performance of a MIMO-OFDM-Based Opto-Acoustic Modem for High Data Rate Underwater Wireless Communication (UWC) System	51
CH. Pallavi and G. Sreenivasulu	
A New Hybrid Islanding Detection Scheme Using Discrete Wavelet Transform and Artificial Neural Networks	67
M. Krishna Goriparthi and B. Geethalakshmi	
Design of Speed and Area Efficient Non Linear Carry Select Adder (NLCSLA) Architecture Using XOR Less Adder Module	81
Yamini Devi Ykuntam, Bujjibabu Penumutchi, and Srilakshmi Gubbala	
Design and Analysis of 1 × 4 Corporate Feed Conformal Microstrip Antenna Array for X-band Spaceborne Synthetic Aperture Radar Applications	93
Bala Ankaiah Nunna and Venkata Kishore Kothapudi	
Multilingual Text Recognition System	103
Mansi Maithani, Dibyasha Meher, and Sumita Gupta	

Implementing Symmetric Boundary Condition in Electromagnetic Harmonic Analysis: Two Different Approaches	115
Sreekanth Karanam, Durgarao Kamireddy, and Arup Nandy	
Performance Evaluation of Hardware Trojan Using FPGA	127
Ravikant Khamitkar and R. R. Dube	
Emotion Recognition in Tweets Using Optimized SVM and KNN Classifiers	135
D. N. S. B. Kavitha, P. V. G. D. Prasad Reddy, and K. Venkata Rao	
A Fused LBP Texture Descriptor-Based Image Retrieval System	145
Akbar Khan, Mohammad Hayath Rajvee, B. L. Deekshatulu, and L. Pratap Reddy	
Normal and Alcohol EEG Signals Classification Using Singular Spectrum Analysis	155
Venkata Keshava Krishna Paramkusham and Sachin Taran	
Tropospheric Zenith Delay (TZD) for Microwaves During Severe Weather Events Over a Few Indian Stations	165
A. Narendra Babu, P. S. Brahmanandam, G. Uma, K. Pushpa, K. Srinivas, and A. Praneetha	
Biomedical Implantable Wideband Antenna with Rectangular C-shaped Radiator	173
Pradyut Mohapatra and Sumit Kumar Khandelwal	
Raspberry Pi Alive Human Detection Robot Using PIR Sensor	183
E. V. Krishna Rao, B. Snehitha, J. Visweswara Rao, P. Mamatha, and M. Gowtham Chowdary	
Design of Band Reconfigurable UWB Microstrip Patch Antenna for Cognitive Radio Application	195
Bammidi Deepa, V. Rajyalakshmi, Karri Sindhuja, Chellaboyina Lalitha Devi, Damodara Satya Sushma, Bommali Mothilal, and Thota Rajesh	
Design of Multiband Frequency Reconfigurable Antenna for Wireless Applications	207
Deepa Bammidi, Gayatri Gorle, Venkatesh Sabbiseti, Reshma Gude, Govardhan Lingampalli, and Vinay Kumar Bodepu	
Automatic Modulation Classification Under AWGN and Fading Channels Using Convolutional Neural Network	215
M. Venkata Subbarao, Beeram Keerthana, D. Ramesh Varma, Sudheer Kumar Terlapu, and G. Challa Ram	

Fusion of Panchromatic and Low-Resolution Multispectral Images Using Combination of PCA and Morphological Hat Transformation . . . 227
 Rupa Banerjee, Vikrant Bhateja, Reetika Mishra, Jerry Chun-Wei Lin, and Carlos M. Travieso-Gonzalez

A Compact Circularly Polarized MIMO Diversified Antenna for 5G Mobile Applications 237
 M. Venkateswararao, M. Vamsikrishna, P. Manjusha, B. Suryaprakashrao, G. Sruthi, and B. Raghavkrishna

Metasurface-Enabled Fork-Shaped Antenna for 2.45 GHz ISM Band Wearable Applications 247
 G. Srilatha, G. S. N. Raju, and P. A. Sunny Dayal

Segmentation of Cell Periphery from Blood Smear Images Using Dark Contrast Algorithm and K-Medoid Clustering 255
 Siddharth Verma, Vikrant Bhateja, Sourabh Singh, Sparshi Gupta, Ayush Dogra, and Nguyen Gia Nhu

Groundwater Quality Assessment of Raipur City Using Machine Learning Models 263
 Anushree Shrivastava, Mridu Sahu, and D. C. Jhariya

Performance Analysis of Breast Cancer Data Using Mann–Whitney U Test and Machine Learning 277
 Priyanka Khanna and Mridu Sahu

Skin Lesion Analysis and Classification Techniques 287
 Dilip Kumar Sharma, Anand Singh Jalal, and Bilal Sikander

Dual Band Microstrip Patch Antenna with Annulated Circular Ring . . . 301
 Chirag Arora

Genetically Optimized Quad-Band mm-Wave Microstrip Patch Antenna 309
 Arebu Dejen, Jeevani Jayasinghe, Murad Ridwan, and Jaume Anguera

Blackhole Attack Prediction in Wireless Sensor Networks Using Support Vector Machine 321
 Niharika Panda and M. Supriya

Identifying Exoplanet Candidates with Machine Learning 333
 Chaitanya Virmani, Ria Singhla, Priyanka Gupta, and Hardeo Kumar Thakur

Feature Selection in Corporate Bankruptcy Prediction Using ML Techniques: A Systematic Literature Review 345
 Mohd Raagib Shakeel, Taufeeque Ahmad Siddiqui, and Shahzad Alam

Activity Recognition from Videos Using Semantic Motion Patterns 365
 M. Sivarathinabala and R. Jothi Chitra

Machine and Deep Learning Technique for Depression Detection Using EEG Data 375
 Shikha Tiwari, Kiran Pandey, Vivek Sharma, and Bhupendra Verma

Next-Generation Firewall with Intelligent IPS 387
 Parth Barot, Sharada Valiveti, and Vipul Chudasama

Security Challenges During Handoff Authentication Operation for Wireless Mesh Network 397
 Vanlalhraia and Ajoy Kumar Khan

Implementation of Photon Hash Function on FPGA 407
 W. Heera, K. Bhagyashree, Rohan Patil, Nalini Iyer, and Shraddha Hiremath

Predictive Generation Trajectory Clustering Model for Google Online Learning: OGMM 419
 Vishnu Kumar Mishra, Megha Mishra, Sameera Khan, Sheela Verma, and Ashish Kumar Tamrakar

A Review on Fake News Identification in Online Social Networks 431
 A. B. Athira, S. D. Madhu Kumar, and Anu Mary Chacko

Comparative Assessment of State-of-the-art Image Fusion Techniques for Fusion of Multi-modality Medical Images 439
 Vineeta Singh and Vandana Dixit Kaushik

Audio and Text-Based Emotion Recognition System Using Deep Learning 447
 Palash Thakur, Ronit Shahu, and Vikas Gupta

Land Use Land Cover Segmentation of LISS-III Multispectral Space-Born Image Using Deep Learning 461
 Nirav Desai and Parag Shukla

Particle Swarm Optimization for Web Document Retrieval Based on Term-Document Matrix 475
 Vishal Kumar, Manoj Kumar, Simran Sirohi, Ravi Kumar, and Santosh Kumar

Light-Weight Deep Learning Models for Visual Malware Classification 485
 E. Akshay Kumar and Jothi Rangasamy

IndoorGML Modeling for WiFi-Based Indoor Positioning and Navigation 497
 Manjarini Mallik and Chandreyee Chowdhury

Assessing Damage of Natural Disasters from Satellite Imagery Using a Deep Learning Model 509
 Shubham Tikle, P. Jidesh, and A. Smitha

Fundamental Graphical User Interface Design of an Educational Android Application 519
 Vandana, Kapil Kumar Nagwanshi, Anil Kumar, and Manish Paliwal

Personality Detection Using Signature Analysis 531
 Garapati Jaya Surya Koushik, Kasukurthi Tirudeepak, and Dwijen Rudrapal

Trajectory Tracking and Link Vibration Reduction of Flexible Manipulator in the Presence of Matched Uncertainty and External Disturbances Using Lyapunov-Based Controller 543
 Sanjay Thakur, Ranjit Kumar Barai, and Anagha Bhattacharya

Multimodal Approach for Code-Mixed Speech Sentiment Classification 553
 S. Keshav, G. Jyothish Lal, and B. Premjith

Despeckling of Ultrasound Imagery with Qualitative Filtering Techniques 565
 Satwinder Kaur, Bhawna Goyal, and Ayush Dogra

An Analysis of Different Noise Removal Techniques in Medical Images 579
 Jaspreet Kaur, Bhawna Goyal, and Ayush Dogra

Study and Analysis of Classification Techniques for Specific Plant Growths 591
 Riya Sharma, Rashmi Vashisth, and Nidhi Sindhvani

Comparative Analysis of Optimization Algorithms for Antenna Selection in MIMO Systems 607
 Nidhi Sindhvani, Rohit Anand, G. Nageswara Rao, Sudhir Chauhan, Alka Chaudhary, Ankur Gupta, and Digvijay Pandey

Methodology for Classifying Objects in High-Resolution Optical Images Using Deep Learning Techniques 619
 P. Lalitha Kumari, Santanu Das, B. Kannadasan, Niranjana Sampathila, C. Saravanakumar, Rohit Anand, and Ankur Gupta

PID Control and Estimation of the Attitude Model Applied to Geostationary Satellites 631
 Radhey Shyam Meena, Rongeet Talukdar, B. Kannadasan, G. Nageswara Rao, Kumar A. Shukla, Rohit Anand, and Ankur Gupta

Comparative Analysis of Autoencoders and U-net-Based Image Steganography 643
 Juhi Singh, Akshat Yadav, Ayushi Siddhu, and Shivani Sharma

Challenges in VLSI Design for Efficient Energy Harvesting 657
Sanjay Kumar and Mansi Jhamb

Area Efficient Design of In-Place RFFT Scaling for OFDM Applications 673
A. Padmavathi and G. L. Sumalata

Design and Implementation of Comparator Using GDI Decoder 685
Sangeeta Singh, T. Akankhsa, Y. Vamshi, Shubham Munratiwar,
and M. Venkata Jayanth

High Speed Efficient Three Operand Adder 695
N. Udaya Kumar, K. Bala Sindhuri, S. S. Harsha Varma,
K. Asha Shaini, K. Sri Hari, and K. Sai Sowmya

Author Index 707

About the Editors

V. V. S. S. Chakravarthy is Professor and Dean R&D in the Department of Electronics and Communication Engineering at Raghu Institute of Technology, Visakhapatnam. He is Senior Member of Communication, Signal Processing and Antenna and Propagation Societies of IEEE. He is serving as Vice Chair of IEEE COMSOC/SPS Joint Chapters of IEEE Vizag Bay Section. His research interests include computational intelligence, smart antenna, data modeling, machine learning, and evolutionary computing tools. He has 18 years of teaching. He served as Co-editor to proceedings of third and fifth International Conference on Microelectronics, Electromagnetics and Telecommunications published in Lecture Notes in Electrical Engineering. He is also Life Member of professional bodies like Instrumentation Society of India, International Computer Science and Engineering Society (ICSES), and Soft Computing Research Society. He published more than 40 journal and conference papers along with one chapter which are indexed in SCOPUS and SCI.

Vikrant Bhateja is Associate Professor in Department of Electronics Engineering, Faculty of Engineering and Technology, Veer Bahadur Singh Purvanchal University, Jaunpur, Uttar Pradesh, India. He holds a doctorate in ECE (Bio-Medical Imaging) with a total academic teaching experience of 19+ years with around 190 publications in reputed international conferences, journals and online book chapter contributions; out of which 37 papers are published in SCIE indexed high impact factored journals. One of his paper published in Review of Scientific Instruments (RSI) Journal (under American International Publishers) has been selected as “Editor Choice Paper of the Issue” in 2016. Among the international conference publications, four papers have received “Best Paper Award”. He has been instrumental in chairing/co-chairing around 30 international conferences in India and abroad as Publication/TPC chair and edited 52 book volumes from Springer-Nature as a corresponding/co-editor/author on date. He has delivered nearly 22 keynotes, invited talks in international conferences, ATAL, TEQIP and other AICTE sponsored FDPs and STTPs. He has been Editor-in-Chief of IGI Global—*International Journal of Natural Computing and Research* (IJNCR) an ACM & DBLP indexed journal from 2017–22. He has guest edited

Special Issues in reputed SCIE indexed journals under Springer-Nature and Elsevier. He is Senior Member of IEEE and Life Member of CSI.

Wendy Flores Fuentes received the master's degree in engineering from Technological Institute of Mexicali in 2006, and the Ph.D. degree in science, applied physics, with emphasis on Optoelectronic Scanning Systems for SHM, from Autonomous University of Baja California in June 2014. She has more than 115 publications which includes journal articles in Elsevier, IEEE Emerald and Springer, chapters and books in Intech, IGI global and Springer, proceedings articles in IEEE. She has been Panel Reviewer of Taylor and Francis, IEEE, Elsevier, and EEMJ (Gh. Asachi Technical University of Iasi. Currently, she is Full-time Professor-Researcher at Universidad Autónoma de Baja California, at the Faculty of Engineering.

Jaume Anguera is Founder of and CTO at Ignion. Prior to this, he was Partner and R&D Manager at Fractus, Barcelona, Spain. He is also serving as Associate Professor at Universitat Ramon LLull, Barcelona, Spain. He is IEEE Antennas and Propagation Distinguished Lecturer. He holds more than 150 patents. His biography is listed in Who's Who in the World, Who's Who in Engineering. He is Author of more than 250 scientific widely cited papers and international conferences with citations above 7500, h-index 50, and i10 index of 150 and Author of 6 books. He has participated in more than 21 competitive research projects financed by the Spanish Ministry. He is Author of 6 books, directed more than 100 bachelor and master thesis, and 3 Ph.Ds. He is Inventor of Virtual Antenna™ technology, which enables full functional multi-band wireless connectivity to wireless devices through miniature and off-the-shelf antenna boosters. He has taught more than 20 antenna courses around the world (USA, China, Korea, India, UK, France, Poland, Czech Republic, Tunisia, Spain). With over 21 years of R&D experience, he has developed part of his professional experience with Fractus in South Korea in the design of miniature antennas for large Korean companies such as Samsung and LG. He has received several national and international awards. He is Associate Editor of the *IEEE Open Journal on Antennas and Propagation*, *Electronics Letters*, *International Journal of Electronics and Communications*, and Reviewer in several IEEE and other scientific journals. He is Vice-chair of the working group "Software and Modeling" at EurAAP.

K. Padma Vasavi is Professor and Head, Department of ECE at Shri Vishnu Engineering College for Women. She has 20 years of teaching experience for undergraduate and postgraduate students. She has contributed immensely toward the development of assistive technology products for the physically and mentally challenged people along with elderly persons. She has done her Ph.D. in digital image processing with title "Multi Scale Multi Directional Edge Detection using Statistical Thresholding" from Jawaharlal Nehru Technological University Hyderabad. She pursued her M.Tech. in Digital Systems and Computer Electronics from Jawaharlal Nehru Technological University Hyderabad and B.Tech. in Electronics and Instrumentation Engineering from Kakatiya University, Warangal. She was Principal Investigator of two DST sponsored projects sponsored by Department of Science and Technology.

She has published more than 15 research papers in peer-reviewed and refereed international journals. She has presented more than 15 papers in reputed International Conferences in India and abroad. She is Fellow of IETE and IE and Life Member of ISTE, ISOI, and BMI. She is Mentor for several projects which won International and National Laurels. She has received Appreciation Award from the office of Honorable President of India for the project “Buzzing Band for Hearing Challenged People”. Her areas of interests include digital signal processing, digital image processing, embedded systems, VLSI, pattern recognition and classification, brain–computer interface, and assistive technology.

Interference and Cancellation Issues for an Indoor VLC Network



Himani Sharma and Rakesh Kumar Jha

Abstract With the high escalating demands of the user for uploading or downloading the data, voice call or sending a text, high speed is required with low latency. This requires larger bandwidth and vast frequency spectrum range. Light Fidelity (Li-Fi) is such a technology which caters to achieve high rising demands of the users thereby providing high quality of service (QoS) and quality of experience (QoE). In Li-Fi, the visible light is used for illumination as well as communication. In Li-Fi, data is transferred from the VLC apex points and received by the user cell phones. Li-Fi is a short range communication link formed with the VLC access point for data transmission and user for its reception. In order to fulfill the desired services, light will be emitted and transferred to the desired demanding users. This causes interference and the desired user may or may not get the light beam signals due to interference. The purpose of the proposal is to reduce the interference level occurred due to the light emitted by the array of LED when the demands of the user rise for an indoor VLC environment. Such interferences can be mitigated with the help of some SIC technique. This paper sheds light on the increase in the interference level and how it impacts the overall system performance rate. This is a next step toward 6G technology.

Keywords Light-fidelity (Li-Fi) · Visible light communication (VLC) · Light emitting diodes (LED) · Successive light interference cancellation (SIC)

1 Introduction

With the massive increase in the usage of LED apex points for illumination and new multimedia applications, the rise in the demand for high speed data rate is growing at fast pace. The sprint in the demand for high speed users, good QoS and QoE is escalating every now and then. Users need the required services for accessing online applications, uploading, downloading, chatting voice call, etc. Li-Fi or VLC

H. Sharma (✉) · R. K. Jha
Shri Mata Vaishnu Devi University, (SMVDU), Katra, Jammu and Kashmir, (J&K), India
e-mail: himanisharma931@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
V. V. S. S. Chakravarthy et al. (eds.), *Advances in Signal Processing, Embedded Systems and IoT*, Lecture Notes in Electrical Engineering 992,
https://doi.org/10.1007/978-981-19-8865-3_1

is such a potent at hand in the wireless communication era. Thus, Li-Fi technology which is a short range optical wireless communication link formed between the LED which emits light by the light emitting diodes (LED) and the user at the receiving end where the photodiode converts the changes occurred in the received light into the electric current which is further used for the recovery of the data stream. Li-Fi plays an eminent role in the wireless industry. This illuminating LED which is fitted at every household is used to emit light signal (transmitting source), and the photodiode (receptor) is used to detect the signal and demodulate it into its original signal with the help of optical-to-electrical convertor (OEC). Li-Fi satisfies the user with high data rate requirements and providing safe and secure data transmission and reception. Li-Fi has attracted the attention of both the researchers and the industry due to its fast acceptance in the market.

Li-Fi or VLC is an amalgamation of illumination and communication simultaneously over the line-of-sight (LoS) link. Li-Fi provides the user with continuous or uninterrupted connectivity services because it exhibits large bandwidth. Li-Fi uses visible light for transmission and possesses high bandwidth for indoor area. The LEDs are cost-efficient and highly durable. LEDs are highly energy-efficient, reliable and easy to install. They have faster response. The quality of service improves and this enables high priority applications to run efficiently with lower latency rate. Various other factors such as channel gain must be considered to provide the signals according to the demand of the user. The users enjoy high data rate services with less time delay. If QoS is good, the quality of experience (QoE) automatically becomes better. Better the QoS, better the QoE experienced by the users.

When the user urges for the services, the light beam is concentrated to the desired user based on its demand. LEDs possess no security threat to human life. It is a highly secured scheme. It is highly durable and consumes low power. Therefore, much power is saved. The major threat to this green communication network is interference from the neighboring LED which adversely affects the whole system performance rate. Thus, some energy-efficient technique is required in order to reduce this interference effect which will lead to high QoS and QoE. Our proposal work is based on the mitigating the interferences and boosting up the overall system performance rate of the system.

2 Research Proposal

Contributions: The research proposal aims at designing of an optimal energy-efficient architecture for high quality of communication with high speed user data rate and reduced complexity as depicted in Fig 1. Various modeling techniques are represented in [1], for indoor, outdoor, underwater and underground environment. Research is carried out for designing of an energy-efficient system, thereby reducing excess power consumption. In [2], a novel sequential load balancing technique with reinforcement learning (RL)-based access point (AP) is discussed which reduces the system complexity. In [3], a novel handover approach has been proposed where the

handover for a particular category is done, keeping in mind various attributes such as velocity of the user and quality of the channel in order to take the decision of handover. In [4], the impact of blockages and shadow formation is discussed. The experimental demonstration for a non-orthogonal multiple access (NOMA) Visible Light Communication (VLC) system with the use of Non-Hermit Symmetry (NHS) IFFT/FFT and other parameters such as transmitting distance and network coverage area are discussed in [5]. In [6], an optical ray tracing approach is studied in order to reduce the path loss in different link length for indoor laser-based VLC system which is done by defining the right positions for the transmitter and the receiver lowers to a point where the collection efficiency will be highest. The design for multiple VLC points for a hybrid system is discussed in [7], which aims in enhancing the data rate and improved capacity with respect to the desired demand as per the user. With the proposed SIC scheme, the interference from the neighboring array of LED can be reduced to a large extent. This is gaining a lot of attention from the researchers and industry as it is a road toward next generation (6G). In the proposed approach, as the user walks away from the LoS link, the light signals available to the user are lessened. The user may not receive the desired services as now it is not under the LoS (line-of-sight) link and FoV (field of view). Thus, there is an urgent need for an energy-efficient architecture for interference mitigation. With the proposed approach, there will be less complexity as the light beam is directed toward the demanding user according to its requirement, i.e., if a user wishes to upload or download a video or tends to do a video call, that user will be given utmost preference (here, A1). After the demands are fulfilled, the light emission for the said user is shut and the light signals do not interfere anymore to its neighboring cells with the help of SIC technique. Now, the demands for the user requiring light signals for voice or text will be given preference according to its need, considering the channel conditions.

In conventional VLC apex system, the users experience interference from its neighboring VLC apex points. This interference keeps on enumerating in traditional VLC APs and light from distinct VLC APs will intersect with one another. This hastens up the interference level and fall in the system performance. With our proposed approach, the mitigation of interferences is done with the help of SIC technique, and then further allocating the light signals to the desired users, with successive interference cancellation algorithm, the interferences are then successively cancelled out after satisfying the user based on its desired applications and their channel conditions.

3 System Model

The proposed system model drafted in Fig. 1 provides high speed data rate, high efficiency and low interference. The interference mitigation technique is done by SIC and then allocating the light signals to its desired users for attainment of high QoS and good QoE. With our proposed interference cancellation technique, the neighboring interferences are eliminated. This results in high QoS experienced by the users and

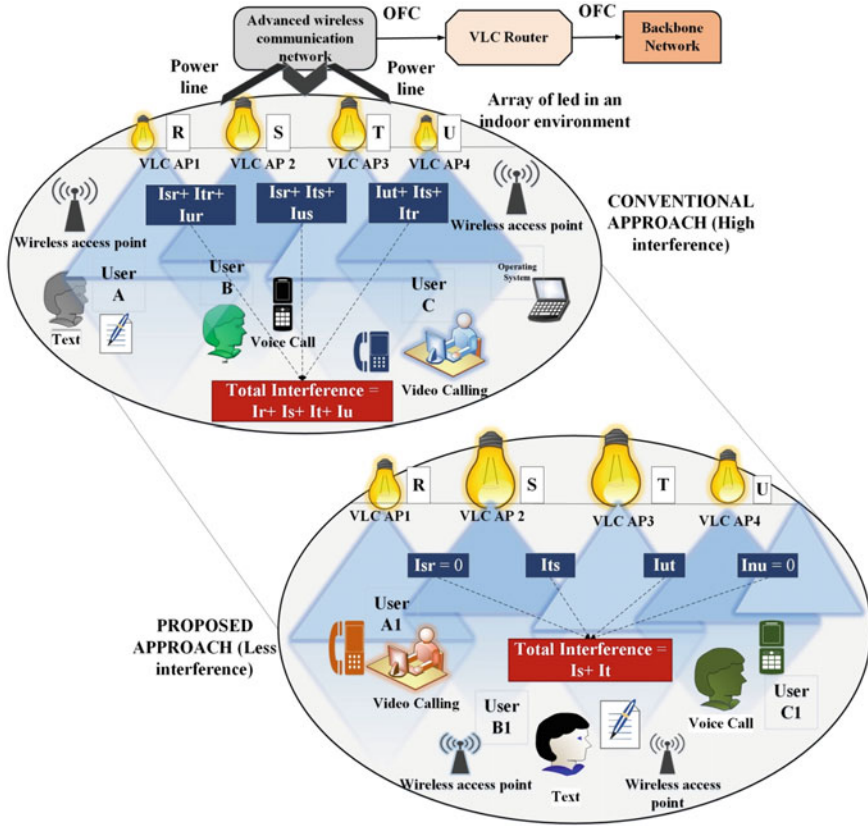


Fig. 1 VLC model for conventional and proposed approach

low complexity. The VLC system model with traditional and proposed approach is depicted in Fig. 1 as shown. The proposed system model helps in mitigating the interferences with the aid of Successive Interference Cancellation (SIC) scheme and allocation method.

The system model is deployed for an indoor room VLC system where the users are taken in a real-time scenario. Here, we have assumed three users A, B and C in case of conventional scheme and the users are at a certain distance to one another. The number of users may rise or fall for a general scenario. The transmitting power coming from the array of light emitting diode (LEDs) is fixed. The transmitter and the receptor must be in LoS for a proper communication to take place. The LEDs are connected to the VLC router and the backbone network with the help of advanced intelligent wireless network through optical fiber cable (OFC). The optimal channel conditions provide a better communication link for transmission and reception of signals. The reflectivity from various surfaces such as glass and wall is ignored here.

In conventional approach, LED transmits continuously the light signals whether a user demands for it or not leading to much more power loss and energy wastage. In this approach, the transmission and reception go on continuously. In conventional system, user experiences interference from various neighboring VLC access points (APs), and this interference keeps on adding and light from these VLC apex points will intersect with each other. The user will experience interrupted signal. This leads to low quality of experience (QoE) by the user.

We have considered three users A , B and C and A_1 , B_1 and C_1 for the conventional and proposed approach, respectively. R (VLC AP₁), S (VLC AP₂), T (VLC AP₃) and U (VLC AP₄) are the apex points emitting light signals. The users in conventional and proposed approach require different applications as per need. In conventional approach, the subscribers will be interrupted by neighboring VLC APs. For example, here, R (VLC AP₁) will face interferences from S , T and U ($I_{sr} + I_{tr} + I_{ur}$). Similarly, S (VLC AP₂) will get interference from R , T and U ($I_{rs} + I_{ts} + I_{us}$). Thus, we see a constant addition in the interference level in the conventional approach. Interference and the SINR are inversely proportional to one another. Therefore, lesser the interference level, the better is the signal-to-interference-noise ratio (SINR) and vice versa. The addition in the total interference level lowers the overall system performance rate.

With the proposed approach, the interferences are mitigated to a large extent after fulfilling the users demand based on the channel conditions. The optimal channel selection is done considering the good signal strength. Here, h_1 is considered as the best channel path taken for the light signals to travel with. In our real-time scenario, firstly, the users are detected which are in need of higher signal strength and need services such as uploading or downloading or doing a video call and after that the channel conditions are seen. If the channel condition is apt, the light signals emitted and given to the desired user urging for video call and then the light beam is given to the users requiring services for voice call and text. Thus, we can say the services are given on priority basis to avoid excessive power wastage.

We will see a sudden drop in the interference level and high rise in the QoS experienced by the users. Interference tends to rise as the distance increases. Therefore, the users and the VLC apex points must be in FoV and in line of sight (LoS) to each other. Proper evaluation of channel conditions must be done to allot the signal strength to the desired users. The link which provides best channel condition for communication will be given first preference.

In case of the proposed approach, the interference will be eliminated and cancelled out from the total interference with SIC technique. With the proposed interference cancellation scheme, decoding of the signals and then subtracting the unwanted interference is done. All the VLC apex points would transmit light signals to the user which are in LoS and the optical-to-electrical convertor converts the optical signal into the electric data stream. After serving the desired user which is in line of sight and bearing good channel condition, this particular signal will be eliminated from the total signal strength. This reduces the interference level. This increases the data rate and provides good QoS and high capacity. In conventional approach, the interference factor is high which tends to lessen with the proposed approach (Fig. 2).

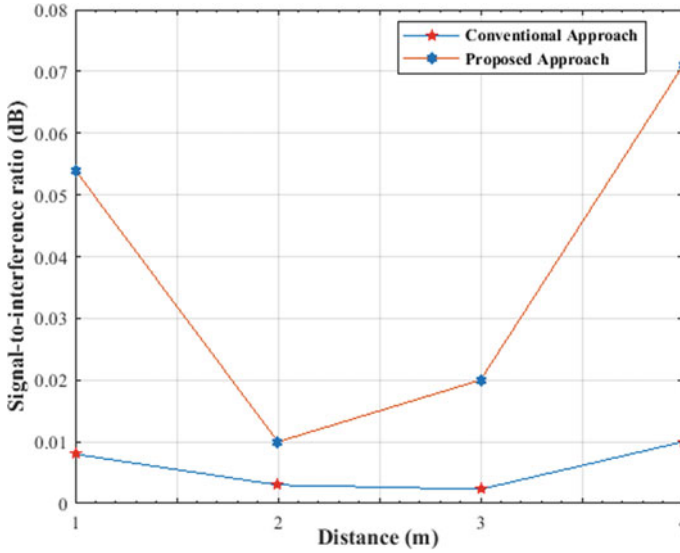


Fig. 2 SINR variation for conventional and proposed approach

The impact of the proposed approach will be seen with lowering of the interference concentration from the respective adjacent VLC APs. Since in the conventional approach, user experiences interference from the neighboring VLC APs and there will be fall in the SINR and low system capacity. But, with the proposed interference mitigation approach and SIC technique, there will be improvement in the SINR values and the capacity also increases. Thus, overall system performance rate gets enhanced. The energy and the spectral efficiency of the user get enhanced. The conventional system provides low data rate and high latency rate due to large interferences from adjacent VLC APs. Thus, we see fall in the overall throughput of the system. This is eliminated with the help of proposed interference mitigation technique.

The main agenda is to enhance the throughput and the performance level of the system. If the user moves away from the light source, the user may not receive the light signals so efficiently as the ones where transmitter and the receiver are perpendicular to each other. With the said proposed approach, the complexity is reduced. This leads to low power exhaustion. The user will enjoy high quality of service. The proper analysis has been shown below. The statistics collected in the table clearly show the rise in the SINR and the capacity with the proposed approach. The energy efficiency rises and the overall system performance rate gets enhanced.

4 Results and Analysis

With regard to the deployment scenario of the users for an indoor environment, the VLC APs and the users are at a distance of [1–4] m. The minimum number of users considered here is four for the real-time scenario. The count of the users may be high or low. The power emitted by the VLC APs is taken as 30 W. This optical wireless communication takes place with transmitter (here LED) and receiver (photodiode) being perpendicular to one another; that is, they must share LoS link in order to gain high signal strength of light signals. The probability of latency or delay signal strength is reduced in LoS communication. Thus, we see a high rise in the throughput level (as depicted in Fig 4) and the overall performance rate of the system. The neighboring or the adjacent VLC APs interference would be cancelled with the help of Successive Interference Cancellation (SIC) approach (Fig. 3).

In our system model, the results are formulated based on the simulations done on MATLAB for the conventional and proposed approach. This section illustrates the cancellation of the unwanted interference coming from neighboring VLCAPs in order to enhance the overall performance rate of the system. This paper shows the increased signal-to-interference-noise ratio (SINR) as it can be seen in the graph with the proposed scheme. The proper results and simulations have been performed.

The simulated results from Fig. 2 shows the SINR variation with respect to the user placed at different distances for an indoor VLC. It could be seen with each passing distance that the SINR increases for the proposed approach in contrast to the conventional approach. The graph depicts increased throughput rate, lower interference level and high SINR with the proposed approach. The increased throughput rate

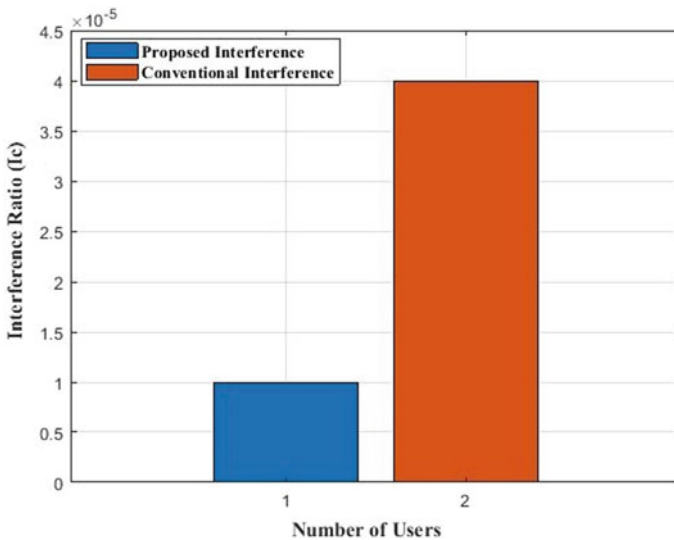
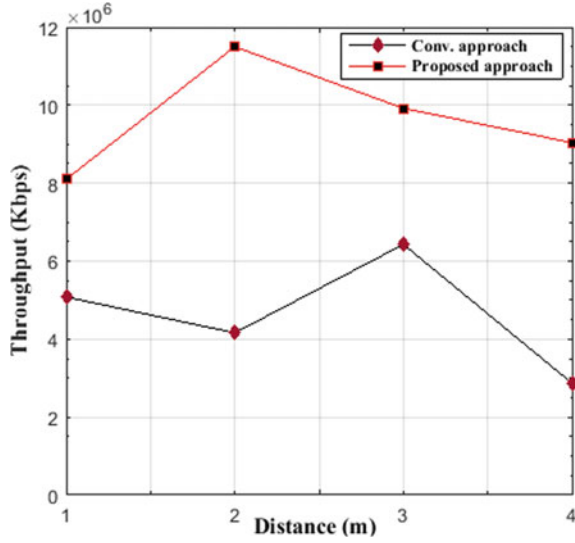


Fig. 3 Interference variation ratio for conventional and proposed approach

Fig. 4 Throughput versus distance (m)



for the proposed VLC indoor system is illustrated in Fig. 4, where 4 VLC AP and 4 users are deployed at 90° to each other so that they share LoS link. The distance is varied at each instant. The demands of the user are given priority based on the application demanded whether a person wants to upload or download, perform an audio call or simply text.

This priority-based approach is achieved by analyzing the channel conditions. If the channel conditions are favorable, the user's demand of desired application can be fulfilled without any delay. Thus, user enjoys high rate data services without any issue. There will be high-level interference issue from the neighboring VLC APs as the VLC APs are in close proximity to each other as shown in Fig. 1. Hence, the complexity of the system rises. Thus, the SINR and the throughput level of the system tend to fall as shown in Fig. 2. In contrast, there is an instant rise in the graph for the proposed scheme with respect to SINR and throughput as shown in Figs. 2 and 4, respectively. As shown in Fig.3, for the conventional approach, the interference keeps on adding due to adjacent VLC APs continuous transmission. Therefore, the graph sees a fall in the system performance rate for the same said approach.

When the interference is mitigated or gets reduced with the light cancellation SIC technique, the graph count sees a sharp increase in context to the overall system performance rate and the system capacity rises. With this approach, the VLC APs points concentrates its energy or emit light signals to the intended user and after the demands for the necessary services are fulfilled, and the interference is eliminated or gets subtracted from the total signal strength, thereby reducing the interference level and enhancing the throughput rate with a lead as depicted in Fig. 4.

The capacity values will vary as the SINR varies which is given by the below formula:

$$\Gamma_i = \frac{B}{2} \log_2(1 + \text{SINR}_i), \text{ bps} \quad (1)$$

As the SINR varies, the capacity also changes. The hike in the SINR performance rate is shown in the graph above. For better capacity or throughput rate, the energy efficiency and the SINR must be high. The proposed approach is able to show variation in the throughput level and increased SINR due to the least amount of interference level. Thus, we can say lower the interference, better the performance rate of the system and the user enjoys high QoS. With this approach, a lot of power is saved thereby lowering the power consumption. Similarly, Fig. 4 clearly describes the overall increase in the throughput level with the movement of the user as per distance. Thus, we can say that with lowering of the interference the system efficiency rate enhances.

5 Conclusion

To meet the high end requirements by the users for high data rate and intensified applications, we are moving toward this new technique of eliminating the interference coming from adjacent VLC APs via mitigation technique. The mitigation of the interference and the enhancement of the overall system performance rate are done with the help of the proposed approach. The proposed scheme fulfills the high data rate demand by the users and increases the overall throughput of the system by lowering the interference level. This paper focuses on the analysis done for the conventional and the proposed approach. The difference between the two approaches is shown via simulation done on MATLAB. With the increase in the system's capacity, the subscriber gets high data rate and experience high quality of services. In near future, this indoor type of optical wireless communication may extend to outdoor communication too. A lot of research is going by the researchers on the VLC and its high tech use in the future for enabling optical communication to each household, labs, universities, etc. The use of array of light via VLC APs apex points reduces the power wastage and thus saves a lot of power consumption. This can be a great step toward 6G network.

References

1. Sharma H, Jha RK (2022) VLC enabled hybrid wireless network for B5G/6G communications. *Wireless Personal Commun* 1–31
2. Ahmad R, Srivastava A (2022) Sequential load balancing for link aggregation enabled heterogeneous LiFi WiFi network. *IEEE Open J Vehicular Technol* 3:138–148
3. Ma G, Parthiban R, Karmakar N (2022) An adaptive handover scheme for hybrid LiFi and WiFi networks. *IEEE Access* 10:18955–18965
4. Tang T, Shang T, Li Q (2020) Impact of multiple shadows on visible light communication channel. *IEEE Commun Lett*

5. Adnan A, Liu Y, Chow CW, Yeh CH (2020) Demonstration of non-hermitian symmetry (NHS) IFFT/FFT size efficient OFDM non-orthogonal multiple access (NOMA) for visible light communication. *IEEE Photonics J* 12(3):1–5
6. Ahmad F, Ramachandrapura S, Manattayil J, Raghunathan V (2020) Path-loss optimized indoor laser-based visible light communication system for variable link length gigabit-class communication. *IEEE Photonics J* 12(4):1–12
7. Valencia-Estrada JC, Béchadergue B, García-Márquez J (2020) Full field radiant flux distribution of multiple tilted flat lambertian light sources. *IEEE Open J Commun Soc* 1:927–942

FPGA-Based 128-Bit RISC Processor Using Pipelining



T. Subhashini, M. Kamaraju, and K. Babulu

Abstract The main aim is to implement 128-bit RISC processor using pipelining techniques through FPGA with the help of von Neumann architecture. With the increase in the use of the FPGA in various embedded applications, there is a need to support processor designs on FPGA. The type of processor proposed is a soft processor with a simple instruction set which can be modified according to use because of the reconfigurable nature of FPGA. The type of architecture implemented is von Neumann. Prominent feature of the processor is pipelining which improves the performance considerably such that one instruction is executed per clock cycle. Due to the increase in innovations in the development of processors, the increasing popularity of open source projects like RISC-V ISA (Instruction Set Architecture), there is a need to also rapidly understand these designs and also upgrade them which can easily be performed on FPGA with trade off in speeds and size as compared to commercial ASIC processors, and hence, we are motivated to understand these systems. In this paper, a 128-bit RISC processor is implemented using FPGA pipelining.

Keywords RISC—reduced instruction set computer · FPGA—field programmable gate array · ISA—instruction set architecture · ASIC—application specific integrated circuit

T. Subhashini
JNTUK, Kakinada, India

M. Kamaraju (✉)
Department of ECE, Seshadri Rao Gudlavalleru Engineering College, Gudlavalleru, India
e-mail: profmkr@gmail.com

K. Babulu
Department of ECE, JNTUK, Kakinada, India

1 Introduction

The two primary techniques within the traditional computation for the execution of algorithms are: ASIC (Application Specific Integrated Circuit) and Microprocessors [1, 2]. ASICs are wont to perform some specific application or are often manufactured as per customer needs. The most drawback of this microchip is its static architecture. The architecture remains same throughout its life cycle. Hence, the second technique, i.e., Microprocessors, came into picture. They execute instructions to perform the task of computation. The architecture of this technique is not static since it can alter the system functionalities, through software coding, without touching the hardware. But the disadvantage with this system is that every time processor has to fetch the instructions stored in the memos and executes only after the prior understanding of what operation it must do with those instructions. Hence, this results overhead within the execution. So reconfigurable computing came into existence which connects both hardware and software by making architecture of the system dynamic and reducing the execution overhead. The performance of a machine becomes determined by way of three key elements: clock cycles per instruction (CPI) [3], clock cycle time and instruction count. The instruction set architecture and therefore the compilers required for a given program. However, each of the clock cycle time and therefore the big variety of clock cycles in step with preparation are often determined through the processor implementation. The development of the control unit and data path for single cycle implementation scheme of the RISC instruction set. This paper emphasizes on architecture of the proposed RISC [4] processor, succeeding with examining each building blocks of the architecture, instruction formats and kinds of instructions. Finally, we conclude it with checking power dissipated by the proposed processor and its simulation results.

2 Literature Survey

In outlining a chip, there are a couple of parameters that have got to be considered. A number of these parameters include: speed review, throughput, number of bits that the microchip manages directly, number of instructions the chip can execute and different contemplations that contribute for the execution of the chip. Factors that incorporate multifaceted nature, possibility for usage, configuration structure and capacity to be actualized within the accessible apparatuses were likewise considered. So to deliver and to understand the contemplations for these parameters, the accompanying investigations were looked into and examined.

2.1 Pipelined Versus Non-pipelined

Normally, pipeline-implemented processors are faster than non-pipelined processors. In non-pipelined processors [5], one instruction is executed per clock cycle, but in pipeline-implemented processor every part of it is kept busy. Many instructions are executed per clock cycle in pipelining technique by dividing one instruction into many number of instructions. Pipelining increases the CPU throughput but increases the latency thanks to overhead of the pipelining process itself. Hardware or software implementation of pipeline technique is feasible. Pipeline may be a series of stages where some work is completed at each stage and work is not completed until it is skilled all the stages.

2.2 Architecture Comparison

RISC versus CISC: RISC is often abbreviated as Reduced Instruction Set Computing, whereas Complex Instruction Set Computing is an abbreviation of CISC. CISC had sizable amount of instructions wherein some instructions for special tasks were used infrequently. Compilers became more prevalent. RISC is meant to perform smaller number of sorts of computer instruction so it is ready to operate at a better speed than CISC. RISC has less number of instructions. It could also be 128 or less and hence few addressing modes. Access instructions were limited to LOAD and STORE. All the operations were performed within the registers of the CPU. The foremost important feature of RISC compared thereupon of CISC is single cycle execution, i.e., done by overlapping fetch, decode and execute phases of two or three instructions referred to as pipelining.

3 Existing and Proposed System

Earlier works based on without using pipelining technique, i.e., non-pipelined RISC processor. This non-pipelined processor has many cons compared to the pipelined system. When a non-pipelining system is taken into account, the most processes like decoding, fetching, execution and writing into memory are executed during a single step, and a single instruction is executed at a time. In case of a non-pipelined system, CPU scheduler selects the instruction or command from the pool of instructions in wait state, when execution is free. Execution takes longer or more number of cycles comparatively.

The proposed system uses FPGA pipelining with 128-bit RISC processor. Pipelining [6] is the concept of gathering the instructions from the processor through a pipeline. Pipelining can also be defined as a gaggle of knowledge processing units that are arranged serial in order that the output from one element becomes the input of

the next or next element. During this technique, multiple instructions are overlapped during execution. It is mainly wont to create the instructions during a processor such that the processes run in concurrent way. Thus, pipelining [7] may be a simple process which adds the new tasks frequently by removing the finished ones. Many instructions are executed parallel. The efficiency of the system depends upon the performance of CPU scheduler and the execution time is extremely low, and therefore, the execution is completed during a less number of cycles.

RISC processor is implemented using von Neumann architecture. This architecture has the advantages of one shared memory for programs and data, one bus is employed for access, an ALU and a program control unit. Control unit [8] fetches the info and instruction during a similar manner from memory. The control unit's design and development became cheaper and faster. Similar way is used to urge the info from I/O devices and memory.

4 128-Bit RISC CPU Architecture

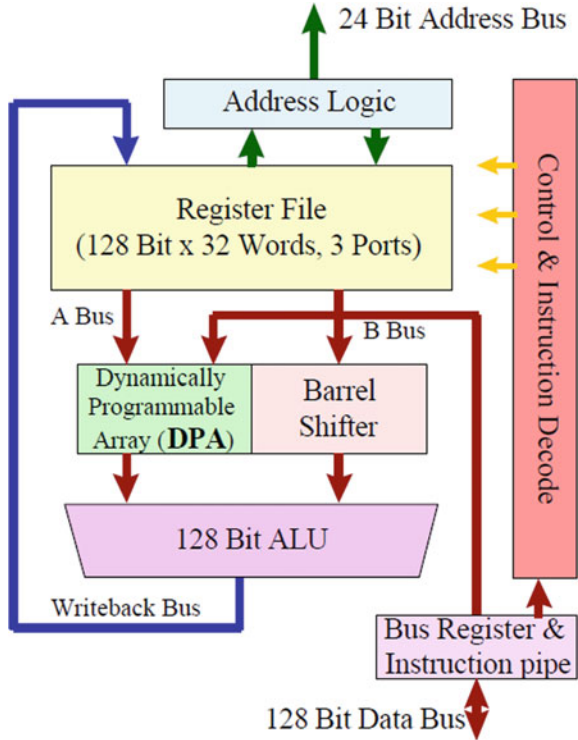
4.1 Introduction

Reduced Instruction Set Computer commonly known as RISC is a sort of architecture-based processor. "Architecture" resembles the way a processor is designed and implemented in the both hardware and software that is nearest to the silicon on which the processor runs. Hardware architecture requires code that breaks down instructions into 0 s and 1 s that the machine can identify—known as machine language. Processor architectures are entirely dissimilar to others, and therefore, the ISA software might be a replica of that. The difference between both can be determined by the way in which the tasks are completed, like handling of registers, interrupts, addressing of memory, external I/O and so forth. The proposed RISC processor is for 128-bit inputs, and the ALU of the processor contains 128-bits in order to perform the operations on the processor. The processor is implemented by using a code snippet developed on the Xilinx tool. Mainly, the processor contains control unit, CPU execution unit and ALU.

Figure 1 shows the 128-bit RISC CPU Architecture. It consists of address logic, register file, DPA, barrel shifter, 128-bit ALU, control and instruction decoder and bus register with instruction pipe. Address logic is just a reference address used to access the memory location of the CPU. Register file is basically an array of processors. They are used to set data between memory and the functional units on the processor chip. DPA's size changes as per the requirement of the processor. Barrel shifter is used to shift the operations in powers of 2. ALU is used to perform arithmetic and logical operations. Control and instruction decoder is used to generate control signals and decode the instruction.

Figure 2 shows the pipelined architecture. In the proposed system, we are using pipelining technique. Pipelining is used for the faster execution of the system. In

Fig. 1. 128-bit RISC CPU architecture



pipelining, all instructions gets overlapped. Pipelining is a process of arranging hardware components in CPU (Central processing unit) so it can increase the performance of the processor. Pipelining architecture also used for storing and executing instructions in a sequenced manner. Pipelining technique has multiple stages and these stages are interlinked with one another in a pipelining manner that is pipe-like structure. Pipelining results in faster ALU [9]. These pipelined architectures are executed with higher frequencies. And in this pipelining technique, we are using von Neumann architecture. Von Neumann architecture has same memory for both data and instructions.

4.2 RISC Instruction Set

All the processor instructions can be classified into three groups in terms of instruction encoding.

1. R-type (Register).
2. I-type (Immediate).
3. J-type (Jump).

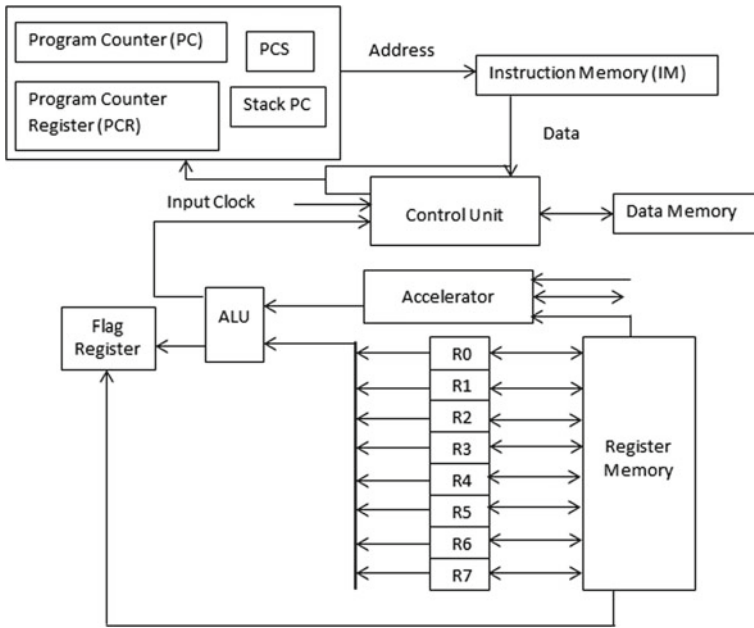


Fig. 2 Pipelined architecture

In the instruction encoding, 32 bits of instructions are divided into several fields of fixed widths. All instructions may not use all the fields. Since the relative positions of some of the fields are same across instructions, decoding these instructions is very simple.

4.2.1 R-type

In *R*-type, an instruction can use up to three register operands: two source and one destination. In addition, for shift instructions, the number of bits to shift can also be specified. *R*-type instruction set is shown in Fig. 3

4.2.2 I-type

I-type instruction set is shown in Fig. 4. *I*-type contains a 16-bit immediate data field. It supports one source and one destination register.

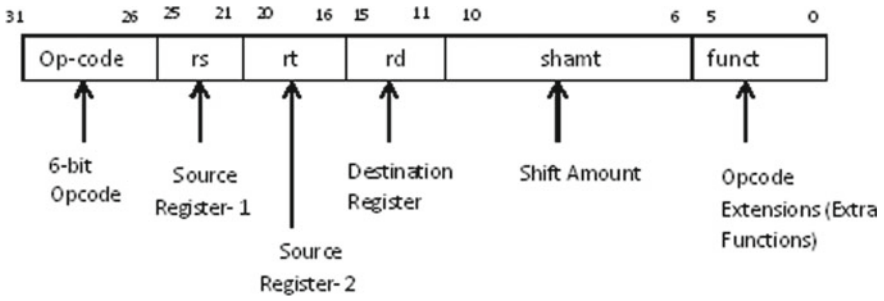


Fig. 3 R-type instruction set

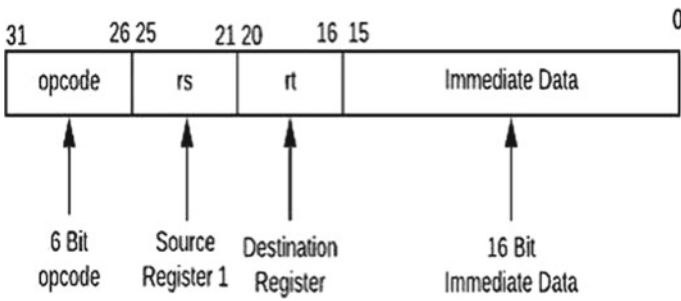


Fig. 4 I-type instruction set

4.2.3 J-type

J-type contains a 26-bit jump address field. It is extended to 28 bits by padding two 0s on the right. Figure 5 shows J-type instruction set.

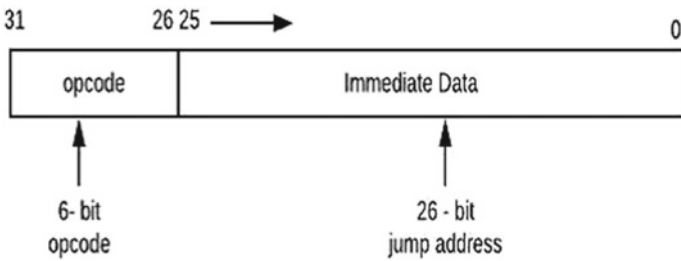


Fig. 5 J-type instruction set

4.3 RISC Pipelining Stages

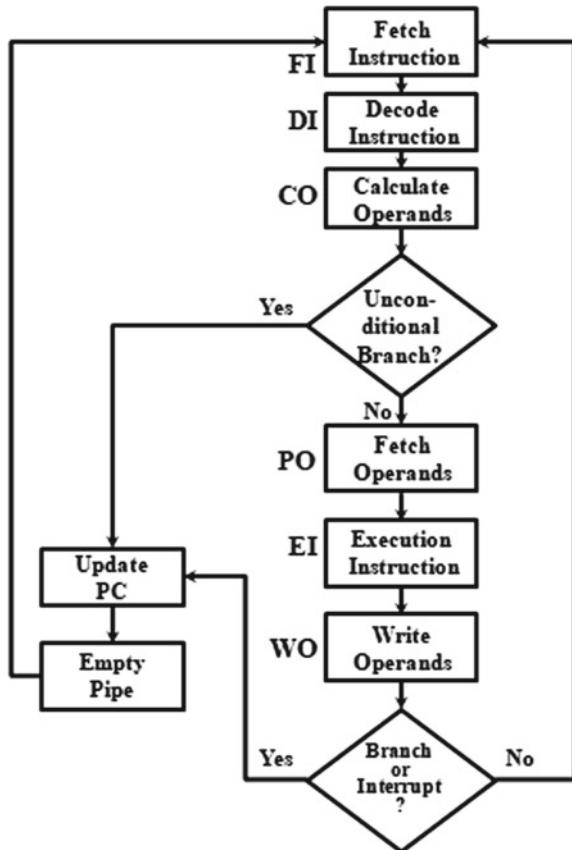
This processor consists of multiple pipelined stages. Flow chart of pipelining stages is shown in Fig. 6. The stages are classified as: instruction fetch (IF) stage, instruction decode (ID) stage, execution (EX) stage, memory write back stage and register write back stage. Each stage is separated by a register buffer for pipelining. Consecutive stages are fed with two separate clocks and alternate with same clocks. These two clocks are non-overlapping, half cycle separated clocks to ensure certain guard band.

Taking an instruction example: **ADD R1, R2, R3.**

4.3.1 IF Stage: (Instruction Fetch)

Figure 7 shows the IF stage. As the name suggests, in this stage, the instruction is fetched from the instruction memory. This is done using a special register called

Fig. 6 Flowchart of pipelining stages



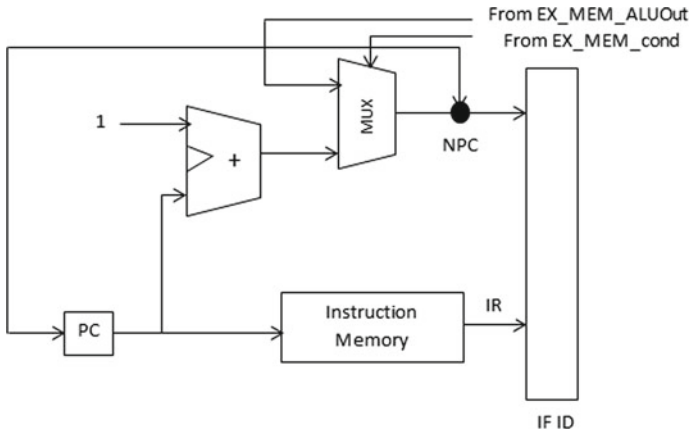


Fig. 7 IF stage

program counter (PC). Every instruction is 128 bit wide and every memory word is also 128 bits; therefore, each memory word has a unique address. The PC stores the value of this address. The value pointer by program counter in the memory is the address of instruction; this instruction is given to instruction register ($IR \leftarrow Mem[PC]$). The value of new program counter register is also incremented by 1, i.e., ($NPC \leftarrow PC + 1$).

4.3.2 ID Stage: (Instruction Decode)

ID stage is shown in Fig. 8. The instruction is fetched from the instruction register or decoded. The opcode field is 8 bits. First source operand R_s and second source operand R_t are present. The last stage is 64 bit immediate data field. In case of jump instruction, the data field is 120 bits. The register operands along with data field from the instruction are read in parallel. This is possible because these fields are in fixed location in the instruction format. This stage also uses sign extension where the immediate data is sign extended to make it 128 bits. In this stage, the value of first source register R_2 is given to register A and second source register R_3 is given to B from the register bank ($A \leftarrow Reg[rs1]; B \leftarrow Reg[rs2]$).

4.3.3 EX Stage: (Execution Stage)

This is the stage where actual computation on given data is performed, depending on the type of instruction which is determined from the ID stage. Among NPC, A, B and IMM, two are selected and their value is given to the ALU. Figure 9 shows the EX stage.

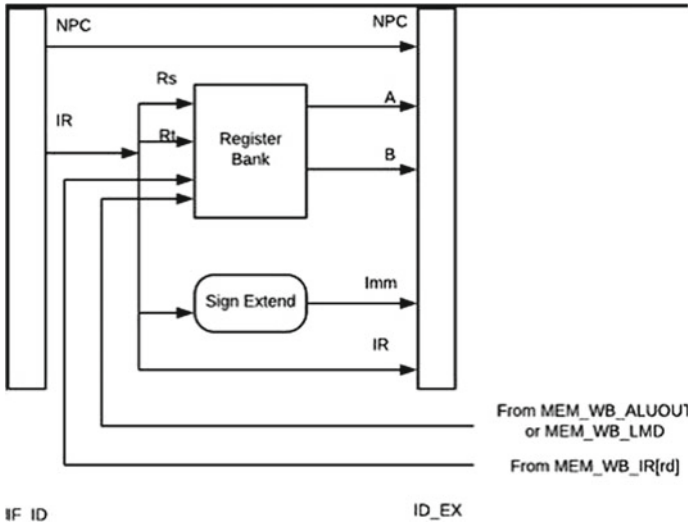


Fig. 8 ID stage

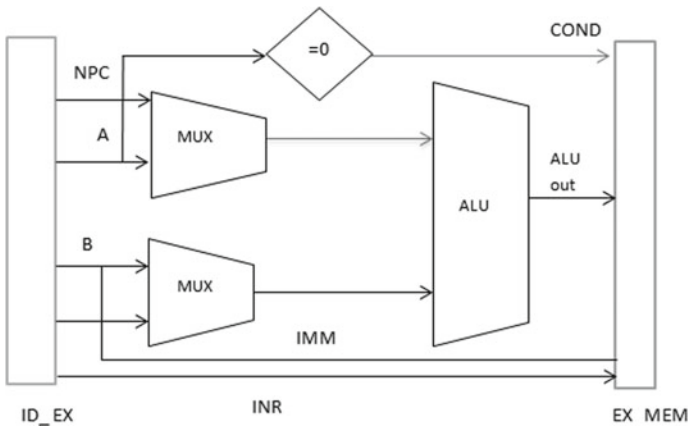


Fig. 9 EX stage

In the ALU (Arithmetic and Logic Unit) block, depending on the type of instruction either arithmetic, logical, shift, comparison operation is performed. For branch type instructions, value of register A is compared to zero using a comparator and branch condition is accordingly checked. Since the type of instruction is ADD (for example), which has a specific opcode value, addition is performed in the ALU and the result is stored in the ALU output register ($ALU\ output \leftarrow A + B$).

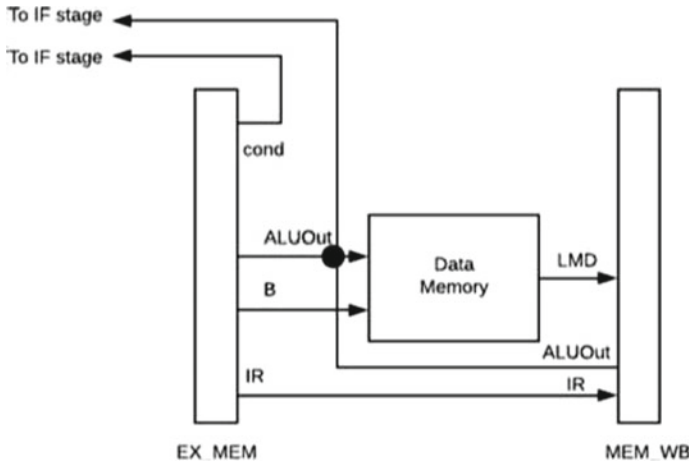


Fig. 10 MA stage

4.3.4 MA Stage: (Memory Access Stage)

Memory access stage is shown in Fig. 10. This is the memory write back stage where result obtained from EX stage is written at a specific memory location or read from it. This stage is used only in case of load and store instructions. In case of other instructions, this stage is simply ignored. For storing instructions, memory data is loaded in register LMD.

4.3.5 WB Stage: (Write Back Stage)

In the final stage of the processor, data generated by the ALU is written back to the register bank in case of register type instructions. Here, there are two possibilities: the data can either come from the memory in case of LOAD type instruction which is loaded into the register LMD or it can come from the EX stage; hence, accordingly, data is selected from the two and given to the register bank. The address of the register at which data is to be stored is specified in the instruction encoding. Figure 11 shows write back stage.

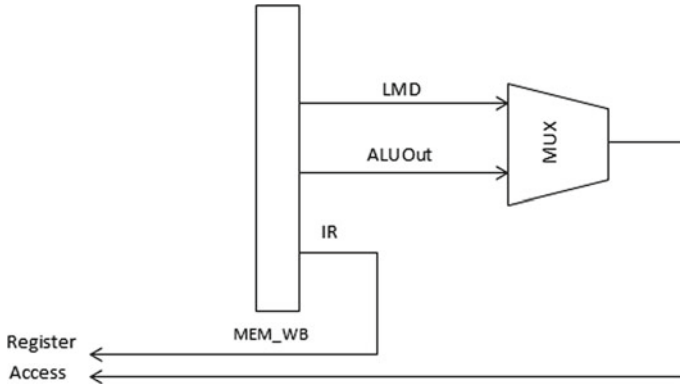


Fig. 11 WB stage

5 Results

5.1 Simulation Results

Figure 12 shows the RTL Schematic Diagram of 128-bit RISC processor implemented using FPGA Pipelining. It contains address lines, status, enable pin, input-outputs, ALU and clk signal. Each of this has internal block diagram that is shown in Fig. 13.

Figure 13 shows the Technology Schematic Internal Diagram of 128-bit RISC processor implemented. It consists of control unit, CPU, execution unit and ALU. This internally consists of mux, latches, flip flops, adders and the blocks.

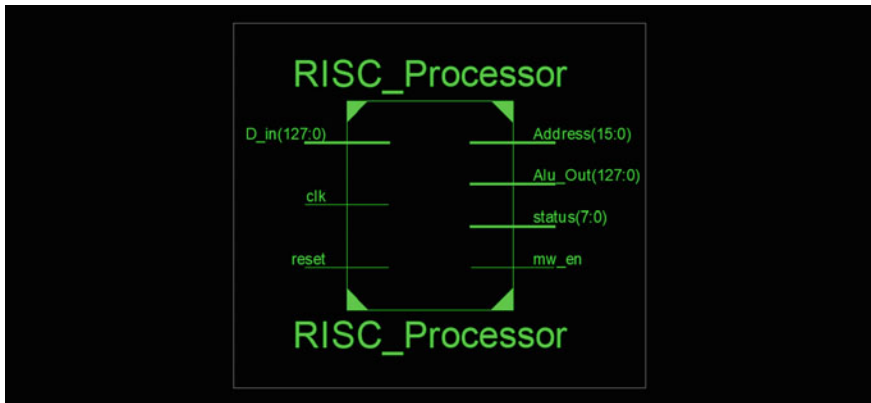


Fig. 12 RTL schematic diagram of 128-bit RISC processor

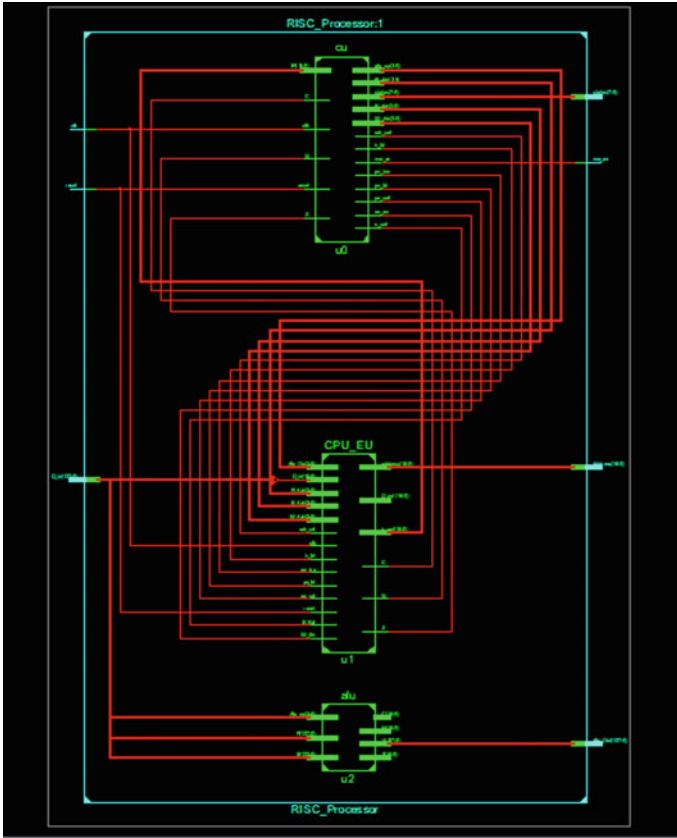


Fig. 13 Technology schematic internal diagram of 128-bit RISC processor

Figure 14 is the output waveforms of the 128-bit RISC processor implemented. The ALU is varied between 124 and 128 bits. Here, the ALU has 128 bits to perform operations on the inputs. Inputs are enabled and outputs are verified accordingly.

5.2 Analysis Data

Figure 15 is the detailed delay analysis of the 128-bit RISC processor implemented. The total delay obtained here is 11.493 ns.

Figure 16 shows the detailed power analysis report for the 128-bit RISC processor implemented. The total power dissipated is 0.362 W. Figure 17 shows the number of registers, flip flops, latches, number of LUTs, etc....



Fig. 14 Output waveforms

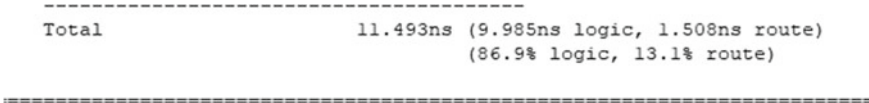


Fig. 15 Delay analysis

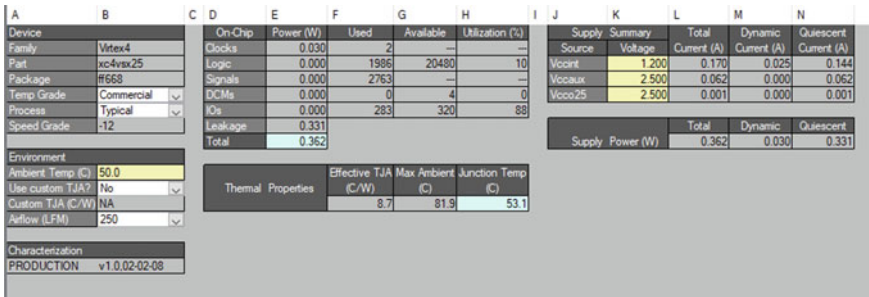


Fig. 16 Power analysis

6 Conclusion

The designed 128-bit RISC processor incorporated with pipelining technique for performance enhancement, simulated and implemented the same on Xilinx Virtex series FPGAs. The complete instruction set was tested individually, and at last, a code snippet is tested which consisted of all types of instructions. Pipelining technique yields better results in terms of speed of execution. It consumes less power of 0.362, operated at a frequency of 87 MHz. The implemented system has the uplift

Device Utilization Summary				
Logic Utilization	Used	Available	Utilization	Note(s)
Total Number Slice Registers	207	20,480	1%	
Number used as Flip Flops	168			
Number used as Latches	39			
Number of 4 input LUTs	2,183	20,480	10%	
Number of occupied Slices	1,369	10,240	13%	
Number of Slices containing only related logic	1,369	1,369	100%	
Number of Slices containing unrelated logic	0	1,369	0%	
Total Number of 4 input LUTs	2,444	20,480	11%	
Number used as logic	2,183			
Number used as a route-thru	261			
Number of bonded IOBs	283	320	88%	
IOB Latches	9			
Number of BUFG/BUFGCTRLs	2	32	6%	
Number used as BUFGs	2			
Average Fanout of Non-Clock Nets	2.72			

Fig. 17 Summary report

of low dissipation of power, occupies lesser area and achieves faster concurrent programming execution.

References

1. Raj V, Patil R, Patil A, Vishwakarma V, Hemnani P (2019) 32-bit processor design on FPGA. In: Electronics and telecommunication. Mumbai University
2. Michael KG, Biro LL, Jackson DB (1998) Power considerations in the design of the alpha 21264 microprocessor. In: Proceeding of the 35th design automation conference (DAC'98), vol 5, no 6, pp 726–731 [5.4][6.9]
3. Wu Q, Pedram M, Wu X (2000) Clock-gating and its application to low power design of sequential circuits. *IEEE Trans Circ Syst-1: Fundamental Theory Appl* 47(103):415–420 [4.4][2.38][6.8]
4. Indira P, Kamaraju M, Dwivedi VV (2019) Design and analysis of a 32-bit pipelined MIPS RISC processor. *Int J VLSI Des Commun Syst (VLSICS)* 10(5)
5. Hakenes R, Yiannos M (2000) A novel low—power microprocessor architecture. In: Proceeding of the Int'l conference on computer design 2000 IEEE, pp 141–146 [6.12]
6. Vijaykumar J, Nagaraju B, Swapna C, Ramanujappa T (2014) Design and development of FPGA based low power pipelined 64-bit RISC processor with double precession floating point unit. *Int Conf Commun Sig Proc*
7. Olanrewaju RF, Fajingbesi FE, Junaid SB, Alahudin R, Anwar F, Pampori BR (2017) Design and implementation of a five stage pipelining architecture simulator for RiSC-16 instruction set. *Indian J Sci Technol* 10(3):1–9
8. Haris L, Jorg H, Wayne W (2005) Approximate arithmetic coding for bus transition reduction in low power designs. *IEEE Trans VLSI Syst* 13(6):696–707
9. David ED, Vijaykrishnan N, Irwin MJ (2002) A clock power model to evaluate impact of architectural and technology optimizations. In: IEEE transaction on very large scale integration (VLSI) systems, vol 10, no 6, pp 844–855 [4.3a][6.19]

Analysis and Implementation of a Multi-path and Metal-Stacked 8-Shaped Inductor



Keerthi Somraj and P. Akhendra Kumar

Abstract The enhancements in technology, have reduced the area of the transceiver, increased the frequency of operation which in turn is increasing the power requirement, and the distance between the elements on the chip has to be highly optimized. It is in this scenario that an 8-shaped on-chip inductor can be used in a transceiver circuit. This paper presents a compact multi-path and metal-stacked 8-shaped inductor which offers better inductance as compared to the single-path, multi-path, and metal-stacked inductors. The inductance of the proposed inductor is observed as 1.4 nH which is nearly 20% more than the inductance of a single-path metal-stacked inductor.

Keywords Transceiver area · Multi-path · Metal-stacking · 8-shaped inductor

1 Introduction

The rapid growth of technology in the formulation of an RF transceiver has led to the continuous miniaturization of CMOS building blocks. This miniaturization which led to a reduction in the overall cost of the system is a challenge in the case of designing the circuits. For RF circuits, inductors play a major role in transmitting/storing energy and also occupy most of the circuit's surface. To overcome this problem integrated planar components have been introduced. The conductive coils are inserted on a magnetic circuit which mostly is made of ferrite to design the inductive components.

More often than not the spiral inductors provide better Q -factor and inductance but at the cost of increased area. Consequently, stacking of the metal layers which was first introduced for GaAs technology and later adopted for the CMOS technology has gained importance due to its compact area. The stacked structures generally exhibit

K. Somraj (✉) · P. Akhendra Kumar

Department of ECE, Faculty of Science and Technology, IFHE University, Hyderabad, India
e-mail: keerthisomraj@gmail.com

P. Akhendra Kumar

e-mail: akhendra.p@ifheindia.org

a single resonance frequency and thus can be modeled with the help of lumped RLC tank network [1]. Regrettably, this has affected the quality factor and self-resonant factor with issues such as the proximity effect, skin effect, and inter-wire capacitance. Due to these losses, the Q -factor of integrated inductors has been degraded.

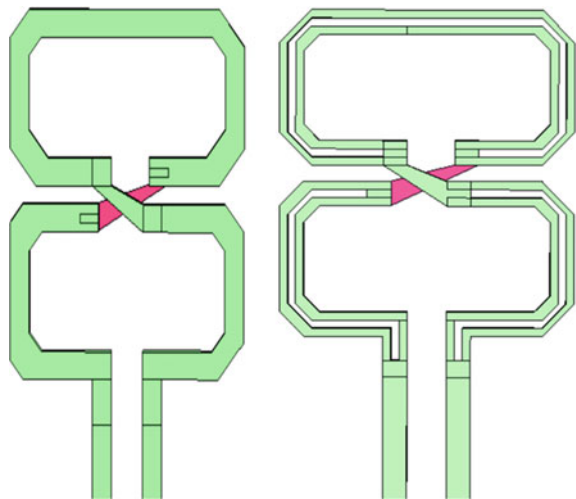
Another technique to improve this Q -factor is to increase the metal thickness or the width of the coil. This can reduce the DC resistance in the coil, but it doesn't offer any advantages at the higher frequencies due to current crowding on the surface of the coil [2]. Additionally, optimizing the width of the coil also affects the losses due to the skin effect and eddy currents which are present due to the current loop present inside the coil. Hence multi-path inductors are proposed to reduce the eddy currents and the resistive losses which are reduced due to the non-uniform distribution of the current [3].

The applications of on-chip inductors vary from tuned amplifiers and mixers with high frequencies to low noise amplifiers (LNA) and voltage controlled oscillators (VCO). In most of these applications, the inductor is used to maximize the gain of the electronic circuit.

This paper describes an 8-shaped inductor that is less susceptible to external magnetic fields when compared to spiral structures. Figure 1 represents a typical 8-shaped inductor that comprises two symmetric loops. These loops lead to providing equal magnitude distribution with opposite polarities in the structure which leads to a reduction of the coupling effects caused by the Electromagnetic Compatibility (EMC) which results in lesser noise generated in the inductor or the electronic circuit [4].

The DC self-inductance in general is relatively proportional to the metal length. As we can see here from the Fig. 1 the metal length of an 8-shaped inductor is relatively high compared to a circular structure, it offers better self-inductance [5]. Due to the presence of symmetrical and reverse magnetic fields present in the two

Fig. 1 Single-path and multi-path 8-shaped inductor layout



loops of this structure, the mutual inductance is negligible. As suggested in [6], the usage of the 8-shaped structures has led to an increase in the inductance value by a range of 25–70%.

The rest of this paper is divided as follows: Sect. 2 provides a description of the 8-shaped inductor, Sect. 3 analyzes the performance of the inductor, and this paper is concluded in Sect. 4.

2 Literature Survey

The implementation of an on-chip inductor necessitates a multitude of trade-offs that depend on the lateral and vertical geometries of the inductor. The lateral geometry is specified based on

1. number of turns, n
2. spacing between the edges, s
3. metal width, w
4. Number of sides, and
5. average diameter d_{avg} , determined from the outer and inner diameters [7].

However, the vertical geometry is dependent on the cross-section of the layout. The inductance of the inductor can be calculated from the above lateral parameters, but the parasitic capacitance and resistance are determined using both the lateral and vertical parameters [8].

The best compromise to obtain better quality factor, electrical performance, and high integration density is when the length D_x and width D_y of the inductor are equal or D_x/D_y is equal to 1 [7]. The length of the metal track is also higher in the 8-shaped layout when compared to the circular layout. The self-inductance of the inductor is proportional to the length of the metal wire, and since it is higher in the former case, it offers better inductance compared to the latter (Fig. 2).

The self-resonating frequency (SRF) is an important parameter in the inductor configurations and is given by

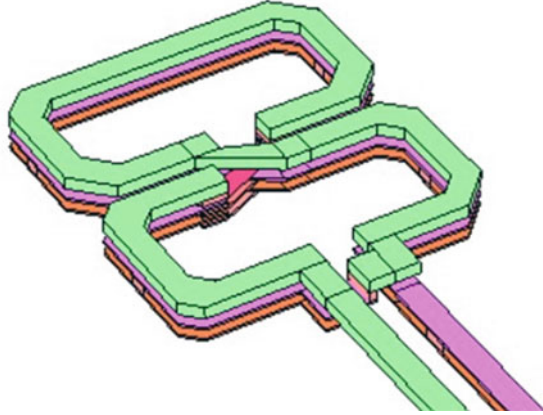
$$\text{SRF} = \frac{1}{2\pi\sqrt{LC}}, \quad (1)$$

where L = inductance and C = capacitance.

It is used to determine the frequency at which the inductor starts behaving as a capacitor. The total capacitance of this inductor is dependent on the length of the metal wires as well as the footprint of the metal track on the substrate. Apart from this, there is also some capacitance due to the presence of the crossing of layers between the two loops [9].

The quality factor of the inductor is also affected by the crossing as it increases resistance, which leads to a decrease in the value of Q .

Fig. 2 Single-path metal-stacked 8-shaped inductor layout



$$Q = \frac{L}{R} \quad (2)$$

Here,

ω is the angular frequency,

L —Inductance of the coil,

R —Resistance of the coil.

In general, the Q value for an O -shaped inductor and a single turn 8-shaped inductor is approximately equal. But, in the case of two-turn or three-turn structures, there is a significant reduction in the peak Q value as the bottom layers are less conductive when compared to the top metal layer. This is a limiting factor that restrains the usage of multi-turn inductor configurations.

To overcome this restraint, metal stacking is used. Here, multiple layers of metal are stacked together to obtain lower resistance. However, when the lower metal layers are stacked, it is seen that the parasitic capacitance between the inductor and the substrate has been increased since the distance between the substrate and the lower metal layer has been decreased. It has been observed that the metal stacking between the two top metal layers has improved the Q -factor considerably [10].

As observed in equation (2), one method to improve Q is by reducing R , which in turn is dependent on the frequency of operation. The increase in frequency leads to an increase in the crowding of current on the surface of the metal coil. In general, for a CMOS process, the width of the metal is much greater than the metal thickness, which leads to the distribution of current only on the two ends of the coil. This creates a non-uniform current in the coil and leads to increased resistance. The effective cross-sectional area is also reduced due to this non-uniformity which induces proximity effect in the configuration, which is worse in the case of multi-turn inductors.

As mentioned in [11], an effective method to reduce this current crowding and non-uniform distribution is to use multi-path configurations which can also reduce the resistance. This method is effective in the case of 8-shaped inductors but cannot

be applicable for spiral structures as they have smaller inner loop lengths compared to outer loops.

3 Analysis of Multi-path and Metal-stacked 8-shaped Inductor

Figure 3 shows the proposed multi-path and metal-stacked 8-shaped inductor which has two paths and three layers of metal stacking with outer dimensions of $74 \times 90 \mu\text{m}$, respectively. The top layer is implemented with metal 6 to incur low losses, and the bottom layer is implemented using metal 2. The two metal layers below the top layer are interconnected. In this way, for the same given area, the inductance value obtained is more than twice the inductance observed in the case of multi-path only.

The maximum Q -factor attained after the simulation has been observed to be 6.4 at 14 GHz, (see Fig. 4), and the inductance has been observed as 1.4 nH which can be observed from Fig. 5.

This proposed inductor has further been compared to a single path, multi-path, single path metal stacked, as well as with spiral inductor configurations of the same dimensions.

In the above Fig. 6 which depicts the relation between the frequency and quality factor, we can observe the SRF of the single-path and multi-stack inductor described in Fig. 2, in comparison with the proposed inductor layout described in Fig. 3, at different frequencies. The comparison of the inductance values for the above-mentioned inductors can be observed from the below Fig. 7.

The above-mentioned configurations and their values can be seen in Table 1.

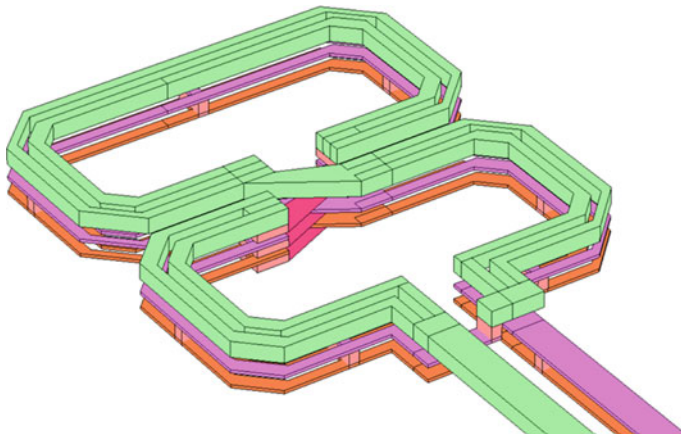


Fig. 3 Multi-path and metal-stacked 8-shaped inductor layout

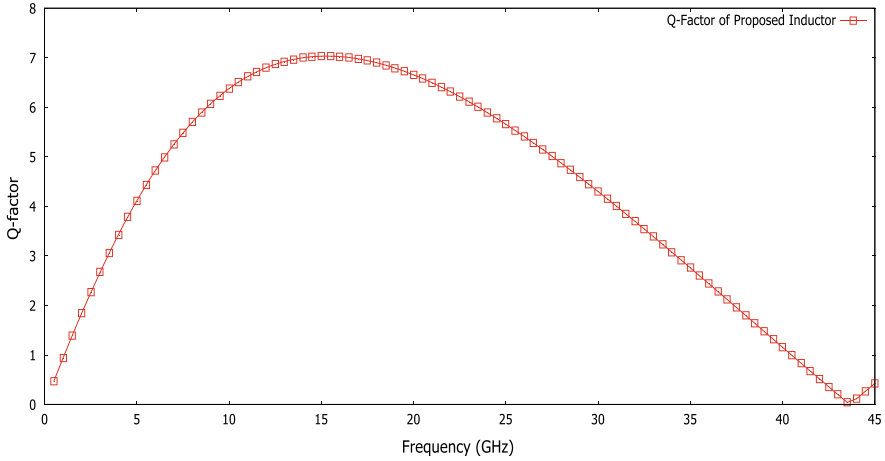


Fig. 4 Q -factor attained for the proposed inductor

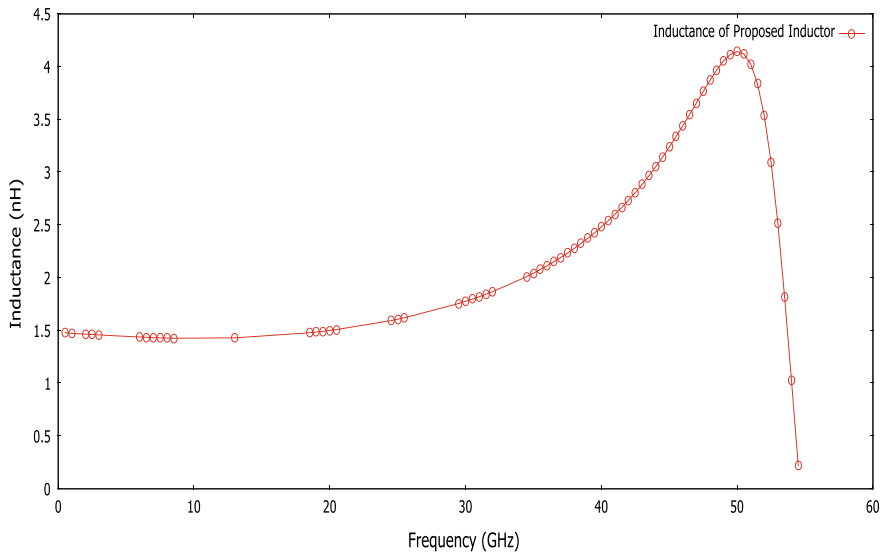


Fig. 5 Inductance attained for the proposed inductor

Table 1 provides a comparison of the values obtained after simulation for the different configurations. Here, we can see that though the spiral inductor as shown in Fig. 6 has the highest Q value, it has the lowest inductance of all the above configurations.

This proposed inductor though offers a lower Q -factor compared to the other configurations but offers higher inductance compared to all of them.

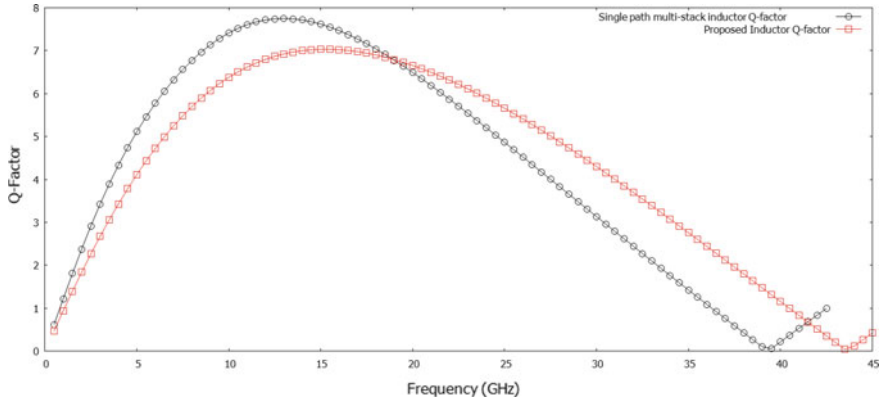


Fig. 6 Comparison of single-path and multi-path inductor’s Q -factor values

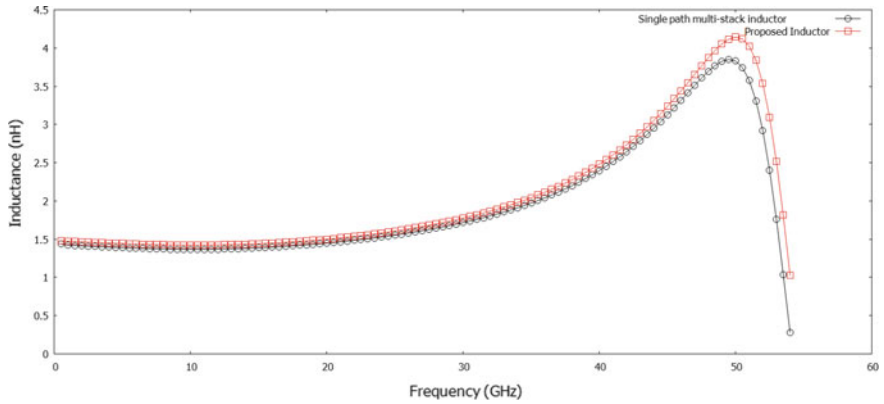


Fig. 7 Comparison of inductance

Table 1 Comparison of parameters for various inductor configurations

Inductor configuration	Maximum Q -factor	Inductance (nH)	SRF (GHz)
Single-path	14.45 at 22.5 GHz	0.2	118.5
Single-path and metal stacked	7.75 at 13 GHz	1.2	39.5
Multi-path	9.2 at 28.5 GHz	0.2	115.5
Spiral	18.39 at 20 GHz	0.134	117.5
This work	6.4 at 14 GHz	1.41	39

Table 2 Comparison of proposed inductor with existing inductors

Ref No.	Dimension	Quality factor	Inductance
4	$466 \times 800 \mu\text{m}$	13	3.5 nH
11	$200 \times 200 \mu\text{m}$	15 at 10 GHz and 17 at 13 GHz	0.5 nH
10	–	20.6	584 pH
9	$360 \times 400 \mu\text{m}$	22	906 pH
This work	$74 \times 90 \mu\text{m}$	6.4	1.41 nH

Table 2 describes a comparison of the proposed 8-shaped inductor with already existing inductors. The area occupied by the proposed inductor is very less compared to the previous inductors and also offers much better inductance.

4 Conclusion

In this paper, we have proposed a high inductance two-path 3-layered metal-stacked 8-shaped inductor, and its Q -factor and inductance have been simulated and are observed to be 6.4 at 14 GHz and 1.41 nH. Apart from this, various configurations of the 8-shaped inductor for the same dimensions as the proposed inductor, i.e., $74 \times 90 \mu\text{m}$ have also been designed, and their simulated values are also compared with the proposed inductor. In addition to this, a spiral inductor of the same dimensions has also been designed and compared with the proposed configuration. Although, the single-path, metal-stacked inductor produces approximately similar inductance but a difference of nearly 20%, makes the proposed inductor better. Also, the losses due to eddy currents and non-uniform current distribution are considerable in case of the single-path, metal-stacked inductor. Though the metal stacking improves the Q -factor, it also increases the proximity effect between the top and bottom metal layers, thereby increasing the resistance. In the future work, we would be using this inductor in designing a low noise amplifier circuit for the RF transceiver.

References

1. Zolfaghari A, Chan A, Razavi B (2001) Stacked inductors and transformers in CMOS technology. *IEEE J Solid-State Circ* 36(4):620–628. <https://doi.org/10.1109/4.913740>
2. Vanukuru VNR, Chakravorty A (2014) Design of novel high-Q multipath parallel-stacked inductor. *IEEE Trans Electron Dev* 61(11):3905–3909. <https://doi.org/10.1109/TED.2014.2359497>
3. Murali B, Rao NB (2022) Design of a novel high Q multi-layer inductor for VCO applications. *Silicon*. <https://doi.org/10.1007/s12633-022-01750-3>
4. Sun Y, Deng W, Chi B (2020) A FoM of -191 dB, 4.4-GHz LC-VCO integrating an 8-shaped inductor with an orthogonal-coupled tail-filtering inductor. In: 2020 IEEE international

- symposium on circuits and systems (ISCAS), pp 1–4. <https://doi.org/10.1109/ISCAS45731.2020.9180559>
5. Le Guillou Y, Fahs B (2012) 8-shaped inductor
 6. Tesson O (2008) High quality monolithic 8-shaped inductors for silicon RF IC design. In: 2008 IEEE topical meeting on silicon monolithic integrated circuits in RF systems, pp 94–97. <https://doi.org/10.1109/SMIC.2008.30>
 7. Mohan SS, del Mar Hershenson M, Boyd SP, Lee TH (1999) Simple accurate expressions for planar spiral inductances. *IEEE J Solid-State Circ* 34(10):1419–1424. <https://doi.org/10.1109/4.792620>
 8. Akhendra Kumar P, Bheema Rao N (2017) Fractal series stacked inductor for radio frequency integrated circuit applications. *Electron Lett* 53:1387–1388. <https://doi.org/10.1049/el.2017.2623>
 9. Fahs Bassem, Gamand P, Berland Corinne (2010) Low-phase-noise LC-VCO using high-Q 8-shaped inductor. *Electron Lett* 46:140–141. <https://doi.org/10.1049/el.2010.2323>
 10. Zou W, Zou X, Ren D, Zhang K, Liu D, Ren Z (2019) 2.49–4.91 GHz wideband VCO with optimised 8-shaped inductor. *Electron Lett* 55:55–57. <https://doi.org/10.1049/el.2018.6012>
 11. Zou W et al (2016) Experimental investigation of multi-path and metal-stacking structure for 8-shape on-chip inductors on standard CMOS. *Electron Lett* 52:1998–1999

Improved Conversion Gain with High SFDR and Highly Linear RF Mixer Using Inductive Gate Biasing Technique for Low Power WAS and Radio LAN Applications



Avvaru Subramanyam  and R. V. S. Satyanarayana

Abstract Wireless access systems (WASs) are ubiquitous, fulfilling various technological needs in our day to day life. In order to meet the latest requirements in terms of power, size and cost, they impose very stringent specifications w.r.t. the performance of various subsystems in them. Inductive gate biasing technique (IGBT) is proposed for performance enhancement of active mixer, a subsystem of RF front end of any superheterodyne receiver, to cater the need of low power WAS including Radio LAN (RLAN) applications. It features a double balanced down conversion mixer topology with inductive source degeneration and active loads along with single ended to differential converters (SE2DCs) for both RF and LO inputs of the mixer. Post layout simulation (PLS) results account for a conversion gain (CG) of 13.62, 93.58 dB spurious free dynamic range (SFDR), 39.10 dBm third order input intercept point (IIP₃) and a single side band-noise figure (SSB-NF) of 12.09 dB at 3.49 mW of D.C power consumption in 180 nm CMOS technology with 1.8 V supply voltage for 5.15 GHz RF and 100 MHz IF.

Keywords Mixer · SE2DC · IGBT · RLAN · CG · IIP₃ · SFDR · SSB-NF

1 Introduction

Adoption of latest WLAN (802.11ax) needs to be encouraged in public Wi-Fi projects for greater use of relatively less congested 5 GHz band [1] in India for applications [2] in Fig. 1. IP-based technologies shaped necessity for unified regulatory regime for multimedia services across communication platforms not only to promote adoption of newer technologies including 5G but also to reduce cost to the end user [3]. Table 1 summarizes the regulatory requirements for applications like WAS and Radio LAN. Also referred as frequency translator, a mixer converts a signal from one frequency to

A. Subramanyam (✉) · R. V. S. Satyanarayana
Department of ECE, SVU College of Engineering, Sri Venkateswara University, Tirupati, Andhra Pradesh, India
e-mail: subramanyam.avvaru@gmail.com

another by a definite conversion gain or loss, *CG* or *CL* which is the ratio of amplitudes of the output to the input, shown in Fig. 2a. When the difference frequency, used as the *IF*, the mixer is called downconverter [4]. A single ended mixer (*SEM*) having a single mixing element is basic with simplicity as its merit while *RF* loss due to external *RF* or *LO* coupling being its setback. Two *SEM*s are combined in single balanced mixer for port isolation [5]. A double balanced mixer has four mixing elements leading to improved linearity [6]. As frequency translation devices, role of mixers is indispensable in communication [7], found in entertainment equipment, test instruments, communication gear, counter measure systems and radar units, etc. [8], categorized by *CG* or *CL*, in addition to their noise and distortion contribution [9].

Noise determines system requirements w.r.t. the smallest input signal power, characterized by noise figure [10]. Third order intermodulation distortion, *IM*₃, is the ratio of amplitudes of the third order intermodulation product, *ID*₃, and the fundamental output component, *ID*₁. 1-dB compression point is the level of input which roots the small-signal linear gain to descent by 1 dB. Dynamic range of an *RF* system is

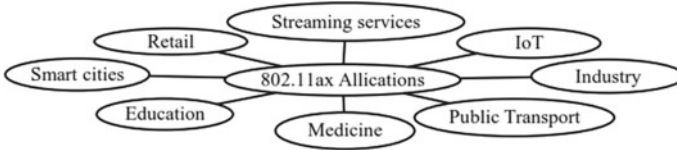


Fig. 1 Low power WAS and RLAN applications specific to 802.11ax (Wi-Fi 6) standard

Table 1 Summary of regulatory requirements in India

Regulation code: NFAP2011/2018-IND 67/29-GSR # 46E/1048E	
Band (MHz)	5150–5350, 5725–5875
License Type	Unlicensed
Max. mean EIRP	200 mW (23dBm)
Max. mean EIRP density	10 mW/MHz
Band usage	Low power WAS and Radio LAN

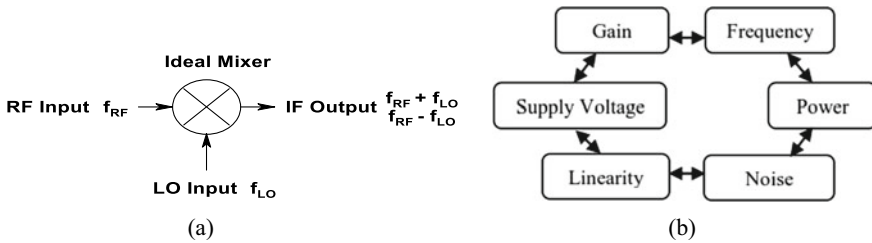


Fig. 2 a Typical mixing process of an ideal mixer. b RF design hexagon

referred by its *SFDR*. *RF* circuits deal with trade-offs in Fig. 2b, to lower the noise of a front-end block (say, amplifier), we must sacrifice linearity or consume more power [11].

2 Literature Survey

Balanced *RF* performance is offered by downconversion mixers w.r.t. *CG*, linearity and *NF*, under the common supply rail, heaps transconductance and switching quad stages and load resistors, one over the other, maneuvering a higher V_{DD} [12]. To reduce the transistor count beneath the supply rail, *IF* output stage and *LO* switches are folded out [13, 14]. Bias current for *LO* and *RF* stages is partially shared causing an additional power consumption [15]. Substitute is the current-bleeding active mixer, which lifts *CG* and *NF* [16]. *nMOS* current-bleeders shield *RF* and *LO* ports for isolation. Without any excess power consumption due to inductive gate bias on the current-bleeding transistors to pump up the transconductance current, improved *CG* achieved in [17]. But, this degrades *NF*. Instead of heterodyne systems, in receivers of direct conversion type, the silicon area can be reduced [18]. But, they suffer from DC-offsets [19]. Linearity gets degraded owing to noise impairments due to leakage triggered by the D.C offset which affects system's dynamic range [20]. By lowering IM_3 , linearity of active mixers or *LNAs* can be increased using derivative superposition (*DS*), modified, and enhanced *DS* methods [21–23]. Current bleeding/reuse techniques require multitude of biasing circuits which increases the complexity and size. In feed forward distortion cancelation, a feed forward path using a shunt-feedback power-hungry auxiliary amplifier used in [24]. Active circuit in the feed forward path generates IM_3 component whose magnitude is same, but, with opposite phase [25], causing gain and *NF* penalties. Improved linearity was achieved by a criterion that integrated techniques like active/passive loads and current bleeds, passive source degeneration that reported a moderate *SFDR*, a little improved *CG* at a little expense of *NF* [26]. For *RFIC* designs, differential approach gained preference due to increased dynamic range, agreeable rejection of parasitics and its invulnerability to common-mode noises [27], however, at the cost of *CG* and *NF*. In view of these works, there is a need to address mixer topology/technique to mitigate the trade-offs and is the motive of this work. In this paper, Sect. 3 describes design of conventional double balanced mixer, proposes *IGBT* and coalesces it with active load, source degeneration and *SE2DCs* to elude transformer baluns. Section 4 reports post layout simulation results and discussion followed by comparison of results in Sect. 5 and conclusion at the end.

3 Design Methodology

3.1 Conventional Double Balanced Mixer (CDBM)

Current switching is the key in operating the mixer shown in Fig. 3a, where direction changing current is offered to the load resistors depending on LO stage transistors (M_3 – M_6) state. M_1 – M_2 forms RF stage and thus referred as stacked stage mixer topology [28]. Equations (1–4) govern the design of $CDBM$; the design parameters and device sizing are presented in Table 2. g_m/I_d refers to the transistor's inversion level. For values in the 5–8 V/A range, *strong inversion* occurs and *weak inversion* for 20–25 V/A . A trade-off among intrinsic gain, speed (f_T) and power efficiency can be used to find the inversion level. Higher speed is exhibited by transistor in strong inversion compared to that in weak inversion which results in lower intrinsic gain. *Moderate inversion* provides good trade-off among performance metrics and is used here.

$$P_{dc} = V_{DD} \cdot I_{SS}; CG \approx \frac{2}{\pi} g_m R_L; \quad (1-4)$$

$$I_{ds} = \frac{g_m^2}{2\mu C_{ox}} \frac{W}{L}; g_m = \frac{2I_{ds}}{(V_{gs} - V_t)},$$

where μ is the surface mobility, C_{ox} is the oxide capacitance, V_t is the threshold voltage and V_{gs} is the gate to source voltage of respective $MOSFET$ s in the design [29].

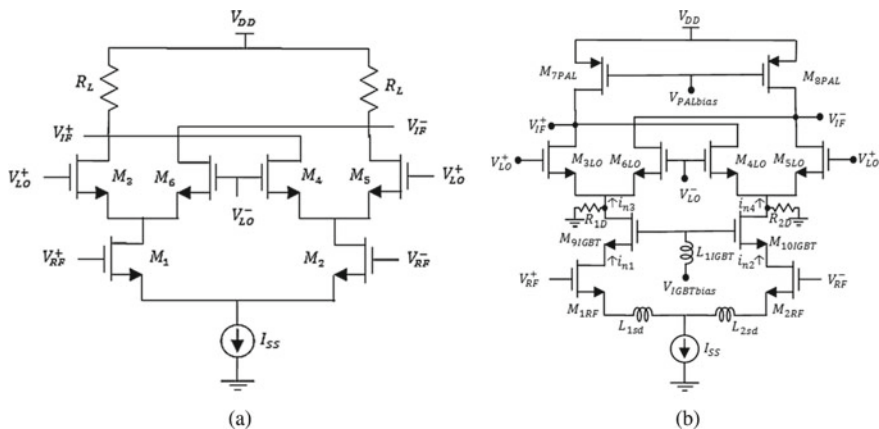


Fig. 3 a Conventional mixer with resistive loads. b Proposed mixer using IGBT

Table 2 Design parameters and device sizing of the mixers

Design parameter/device size	Value
Technology/process node	0.18 μm
Supply voltage (V_{DD}), current source (I_{SS})	1.8 V, 1.94 mA
D.C power (P_{dc}), expected conversion gain (CG)	3.5 mW, 10 dB
Overall transconductance (g_m), load resistance (R_L)	32.33 m Ω , 50 Ω
g_m/I_{ds} ratio	16.66 Ω/A
RF and LO stage transistor widths (W_{RF}, W_{LO})	240, 120 μm
RF and LO transistor overdrive voltages (V_{odRF}, V_{odLO})	60, 120 mV
pMOS active load transistor width (M_{7-8PAL})	5.6 μm
Gate biasing transistor width ($M_{9-10IGBT}$)	26 μm
Gate biasing inductance (L_{1IGBTs})	2.26 nH
Source degenerative inductor (L_{1-2SD})	0.8 nH
Bias voltage for pMOS active load ($V_{PALbias}$)	0 V
Bias voltage for gate biasing transistor ($V_{IGBTbias}$)	0.36 V
Inductance of SE2DC for RF and LO inputs (L_{1-2RF}, L_{1-2LO})	1.54 nH, 1.57 nH
Capacitance of SE2DC for RF and LO inputs) (C_{1-2RF}, C_{1-2LO})	618 pF, 630 pF
RF and LO input frequencies (f_{RF}, f_{LO})	5.15, 5.05 GHz
IF output frequency (f_{IF})	100 MHz
Input RF and LO powers (P_{RF}, P_{LO})	-30, 0 dBm

3.2 Proposed Double Balanced Mixer (PDBM)

PDBM uses IGBT to improve CG and linearity based on the techniques in [26], especially linearity. Resistive loads are replaced by pMOS active loads (M_{7-8PAL}) to improve CG as they are less likely affected by the flicker noise [30] and inductive source degeneration (through L_{1-2sd}) improves linearity. L_{1IGBT} and $M_{9-10IGBT}$ form inductive gate bias network. CG can be improved using IGBT, analyzed using Fig. 4.

Small-signal current flowing through M_{1RF} be i_{n1} ,

$$i_{n1} = g_{m1RF} \frac{V_{RF}}{2} \sin(\omega_{RF}t). \quad (5)$$

Small-signal voltage w.r.t. the source of M_{9IGBT} ,

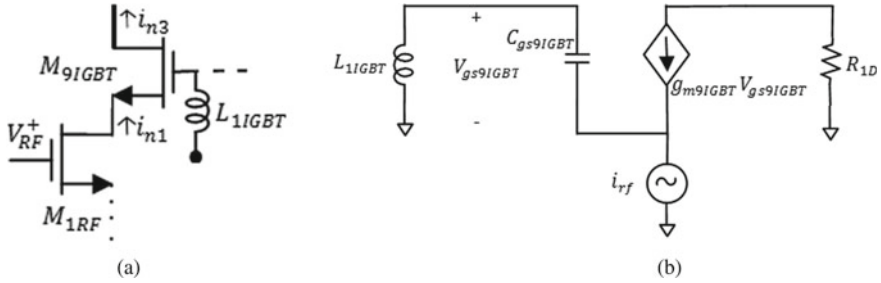


Fig. 4 **a** Small-signal analysis of a portion of the PDBM. **b** Equivalent circuit

$$V_{n1} = \frac{g_{m1RF}}{g_{m9IGBT}} \frac{V_{RF}}{2} \sin(\omega_{RF}t). \quad (6)$$

Gate to source voltage of transistor M_{9IGBT} ,

$$V_{gs9IGBT} = \frac{-\left(\frac{1}{j\omega_{RF}C_{gs9IGBT}}\right)V_{n1}}{\frac{1}{j\omega_{RF}C_{gs9IGBT}} + j\omega_{RF}L_{1IGBT}}, \quad (7)$$

(6) in (7) results in,

$$V_{gs9IGBT} = \frac{-\left(\frac{g_{m1RF}}{g_{m9IGBT}} \frac{V_{RF}}{2} \sin(\omega_{RF}t)\right)}{1 - \omega_{RF}^2 C_{gs9IGBT} L_{1IGBT}}. \quad (8)$$

Thus, when $0 < (1 - \omega_{RF}^2 C_{gs9IGBT} L_{1IGBT}) < 1$, then $V_{gs9IGBT}$ gets enhanced by L_{1IGBT} . The small-signal current, i_{n3} will be

$$i_{n3} = \frac{-g_{m9IGBT} \left(\frac{g_{m1RF}}{g_{m9IGBT}} \frac{V_{RF}}{2} \sin(\omega_{RF}t)\right)}{1 - \omega_{RF}^2 C_{gs9IGBT} L_{1IGBT}} = \frac{-g_{m1RF} \frac{V_{RF}}{2} \sin(\omega_{RF}t)}{1 - \omega_{RF}^2 C_{gs9IGBT} L_{1IGBT}}. \quad (9)$$

Similarly, the out phased small-signal current, i_{n4} , can be deduced. The differential mixer output current is thus given by

$$\begin{aligned} i_{IF} &= (i_{n4} - i_{n3})Sg(\omega_{LO}t) \\ &= \left(\frac{g_{m2RF} \frac{V_{RF}}{2} \sin(\omega_{RF}t)}{1 - \omega_{RF}^2 C_{gs10IGBT} L_{1IGBT}} \frac{-g_{m1RF} \frac{V_{RF}}{2} \sin(\omega_{RF}t)}{1 - \omega_{RF}^2 C_{gs9IGBT} L_{1IGBT}} \right) \\ &\quad \left(\frac{4}{\pi} \sum_{n=1,3,5,\dots}^{\infty} \frac{1}{n} \sin\left(\frac{n\pi}{T_{LO}}\right) \right) \end{aligned} \quad (10)$$

Due to symmetry in mixer stages and assuming identical device geometries,

$$i_{IF} = \frac{2}{\pi} g_{mRF} \frac{V_{RF}}{1 - \omega_{RF}^2 C_{gs} g_{IGBT} L_{1IGBT}} [\cos(\omega_{RF} - \omega_{LO})t - \cos(\omega_{RF} + \omega_{LO})t + \dots] \quad (11)$$

The desired down conversion CG of the proposed mixer will be

$$CG_p \approx \frac{2}{\pi} g_{mRF} \frac{1}{(1 - \omega_{RF}^2 C_{gs} g_{IGBT} L_{1IGBT})}, \quad (12)$$

$$Z_{in}(j\omega) = \frac{1}{j\omega_{RF} C_{gsRF}} + Z_{sdg} + \frac{\omega_T Z_{sdg}}{j\omega_{RF}} + Z_{IGBT}, \quad (13)$$

where $\omega_T = \frac{g_{mRF}}{C_{gsRF}}$; $Z_{sdg} = \omega_T L_{sdg} + \left(\frac{1}{j\omega_{RF} C_{gsRF}} + j\omega_{RF} L_{sdg} \right)$;

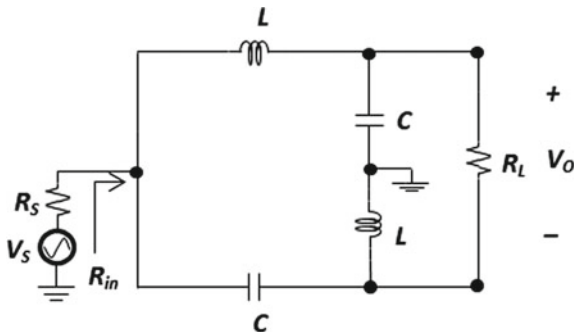
$$Z_{IGBT} = \frac{1}{g_{mIGBT} + j\omega_{RF} C_{gsIGBT}} (1 - \omega_{RF}^2 C_{gs} g_{IGBT} L_{1IGBT}).$$

From Eq. (12), as long as $0 < (1 - \omega_{RF}^2 C_{gs} g_{IGBT} L_{1IGBT}) < 1$ holds well, increased CG is expected. In $RFIC$ design, inductors are used for source degeneration to enhance linearity as noise is not added to the circuit by ideal inductor and it reduces required V_{DD} [31]. Z_{in} , input impedance at the gate of RF stage with $IGBT$ and source degeneration impedance is given in Eq. (13). Z_{sdg} is the input impedance w.r.t. the source, Z_{IGBT} is the input impedance due to $IGBT$ transistors. Series inductance L_{sdg} produces a frequency independent positive real part and a series LC resonant network. Due to the integration of L_{1IGBT} at gates of biasing transistors which causes a negative impedance (Z_{in}) referring to its source terminal, stability concern arises. However, as far as $(1 - \omega_{RF}^2 C_{gs} g_{IGBT} L_{1IGBT}) > 0$, the mixer will be stable. Balun (balanced to unbalanced) does single ended to differential conversion. It is acclaimed to use differential inputs when RF /electromagnetic interference exist, and noise is generally a problem. Thus, LO and RF inputs are converted into differential before applying them to the $PDBM$ using $SE2DC$ in Fig. 5. It is valid over a bandwidth centered at frequency f , where $f = \omega_0/2\pi$. In spite of being lumped LC structure, it appeared in $RFIC$ s and for narrow band applications using discrete components on $PCBs$ and is a common low cost approach [32].

Differential voltage across the load is V_o ,

$$V_o = \frac{R_L}{jX} V_{in} \Rightarrow \frac{V_o(\omega_0)}{V_i(\omega_0)} \cong \sqrt{\frac{R_L/R_{in}}{\pi/2}}. \quad (14)$$

Fig. 5 *SE2DC* designed for differential *RF* and *LO* inputs



4 Results and Discussion

PLS results of the *SE2DC* for *RF* at 5.15 GHz show return loss ≈ -9.5 dB for the ports, with forward voltage gain ≈ -3.5 dB for the differential outputs with phase difference $\approx 180^\circ$; similarly for *LO* at 5.05 GHz. The results of *CDBM* and *PDBM* with *SE2DC*s for *CG* (red and blue curves) and *SSB-NF* (light green and light blue curves) are shown in Fig. 6a, for various *LO* powers. As noise is added to the signal, mixer's *S/N* degrades, and hence its *NF*. In Fig. 6b, **m3** and **m4** marks 1-dB gain compression for *CDBM* (blue curve) at -14.7 dBm *RF* power and the same by **m5** and **m6** at -10.45 dBm for *PDBM* (blue curve), thus, displays improved linearity and *CG*. Figure 7 shows *IMD*₃ products; lower and upper *IMD* terms, centered around the *IF* at 100 MHz are of same amplitude. Monte-Carlo simulation (*MCS*) is performed for 200 (*N*) trials to assess how much deviation the mixers go through for device mismatches, shown in Fig. 8a. *MCS* results for *CDBM* and *PDBM*'s mean *CG* values are 9.69 dB and 14.04 dB, respectively, which match with the *PLS* values. Noise power spectral density of the *PDBM* is shown in Fig. 8b and is combined with *TOI* simulation results to calculate the *SFDR*. Thus, to compute *SFDR*, the results from *IMD* and noise floor simulations are needed. *LO*-to-*IF*, *RF*-to-*IF* and *LO*-to-*RF* port leakages vary significantly with *LO* power, as shown in Fig. 8c.

5 Comparison of Results

In order to appraise the performance of the mixers pedantically, a bench mark figure of merit (*FoM*) [33] characterized by Eq. (15) which includes the key parameters like gain, noise figure, third order intercept, input power and frequency along with D.C power, given by *FoM*. Improvements in *CG*, third order intercept and *SFDR* along with the deterioration in noise figure as well as increased size of the mixer are portrayed in Fig. 9. *PDBM* is integrated with *RF* and *LO SE2DC*s, and its performance stands out with the recently reported mixers of similar kind as shown in Table 3.

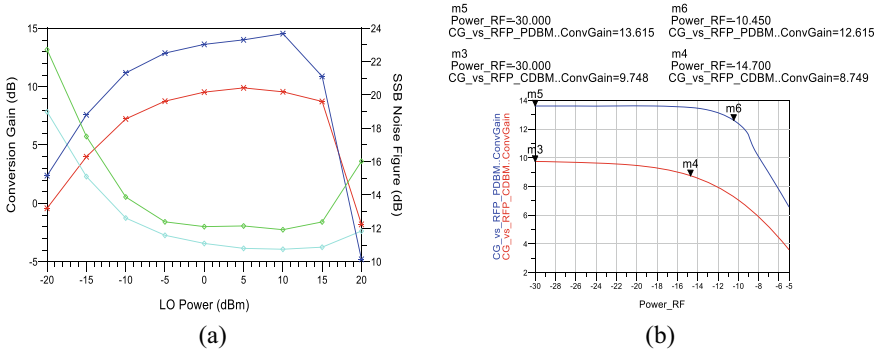


Fig. 6 a CG and SSB-*NF* vs. *LO* power. b 1-dB gain compression of *CDBM* and *PDBM*

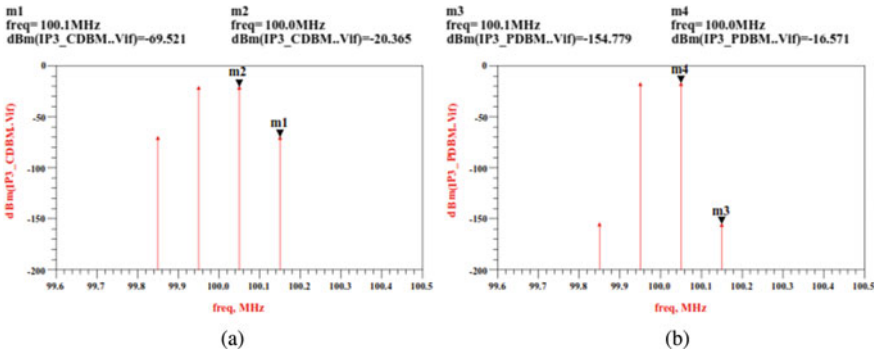


Fig. 7 Third order *IMD* products of a conventional mixer, and b proposed mixer

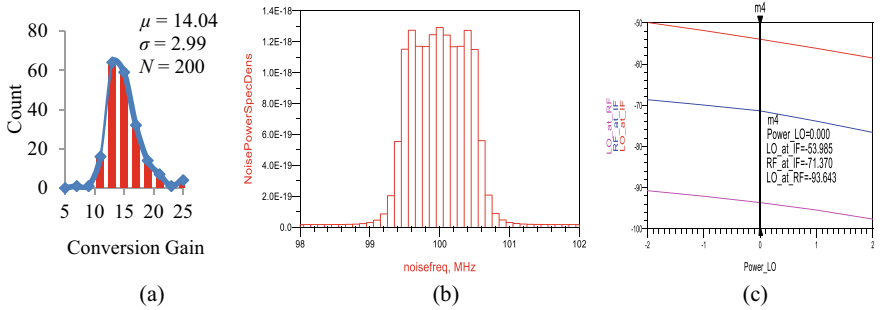


Fig. 8 *PDBM* a MCS for CG. b Noise power spectral density at output. c Port isolation

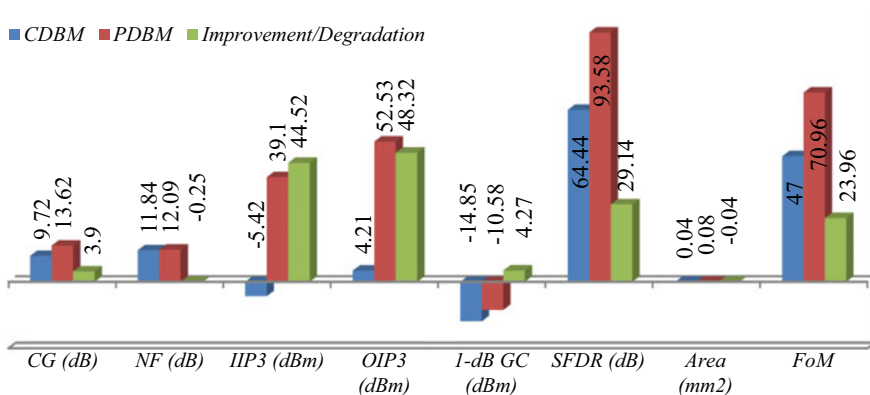


Fig. 9 Comparison of various performance metrics of CDBM and PDBM

$$FoM = 10 \log \left[\frac{10^{(CG - 2NF_{SSB} + IIP3 - 10 - P_{Lo})/20} f_{RF}/1 \text{ KHz}}{P_{d.c}/1 \text{ mW}} \right] \quad (15)$$

Layout of the *SE2DC* for *RF* input is shown in Fig. 10a, which occupies an area $\approx 0.065 \text{ mm}^2$. Layouts of the conventional and proposed mixer cores are shown in Fig. 10b and c, which occupy an area $\approx 0.04 \text{ mm}^2$ and 0.08 mm^2 , respectively. The increased size of the *PDBM* is mainly due to the use of inductors for conversion gain as well as linearity enhancements.

6 Conclusion

Inductive gate biasing technique (*IGBT*) is proposed for performance enhancement of differential double balanced active mixer for low power *WAS* and Radio *LAN* applications. *SE2DCs* are used to obtain differential mixer inputs, *RF* at 5.15 GHz and *LO* at 5.05 GHz. A conversion gains of 13.62, 93.58 dB *SFDR*, 39.10 dBm *IIP₃* and *SSB*-noise figure of 12.09 dB are achieved at a D.C power consumption of 3.49 mW on 180 nm *CMOS* technology. It exhibits an improvement of 3.9 dB in *CG*, 29.14 dB in *SFDR*, 44.52 dBm in *IIP₃* and thus leads to a 23.96 dB improvement in *FoM*, at an expense of 0.25 dB degradation in *NF* and occupies double the mixer core area (0.08 mm^2) compared to the conventional mixer. Overall performance of the proposed mixer competes well with the recently reported double balanced mixers.

Table 3 Comparison of performance parameters with related works in the recent past

Parameters	[17] (Meas)	[20] (PLS)	[23] (Meas)	[26] (Sim)	[34] (PLS)	[35] (Meas)	[36] (PLS)	This work (PLS)
<i>Tech.</i> (nm)	130	180	65	130	180	180	180	180
V_{DD} (V)	1.2	1	1	1.3	1	1.8	1.6	1.8
f_{RF} (GHz)	2.4 ^x	2.4	2.4	5.825	0.4–6	1–6	0.5–10	5.15
<i>CG</i> (dB)	22	11.4	12.5	10.2	15.1 [^]	13–10	18.5 [^]	13.6
<i>NF</i> (dB)	7.2 [*]	8.5 [*]	10.5 [*]	9.95	13.4 [*]	12–18 [*]	8.8	12.0
<i>IIP₃</i> (dBm)	16	26.36	7.6	3.08	-3 [^]	-4.5	-4	39.1
<i>1-dB GC</i> (dBm)	-20.4	n/a	-2.5	-6.6	-13	n/a	-14	-10.5
<i>SFDR</i> (dB)	n/a	91.5	n/a	67.4	n/a	n/a	n/a	93.5
<i>Area</i> (mm ²)	1.16 ⁺	0.33	0.22	n/a	0.25	0.54	0.27	0.08
<i>Power</i> (mW)	3.15	0.26	1.2	2.5	3.1	3.45	12	3.49
<i>FoM</i>	62.6	67.5	54.5	55.3	43.5	38.6	41.6	70.9
<i>Technique</i>	<i>T1</i>	<i>T2</i>	<i>T3</i>	<i>T4</i>	<i>T5</i>	<i>T6</i>	<i>T7</i>	<i>T8</i>

T1 Inductive source degeneration, integrated transformer-based gate, forward body bias
T2 Ultra low power *RF* mixer using complimentary pre-distortion technique
T3 Double balanced *CMOS* active mixer with enhanced derivative superposition method
T4 Active load and current bleed with inductive source degeneration mixer
T5 Complementary current mirror-based bulk-driven *CMOS* mixer
T6 Complementary and self-biased fully differential *CMOS* analog mixer
T7 Active *CMOS* sub-harmonic mixer with enhanced transconductance
T8 Active load, inductive gate biasing, inductive source degeneration mixer
^x Single ended ^{*} $DSB_{NF} \approx SSB_{NF} - 3\text{ dB}$ ⁺ with test buffer [^] peak value
Meas. Measured *PLS.* Post layout simulation *Sim.* Simulation *n/a.* not available

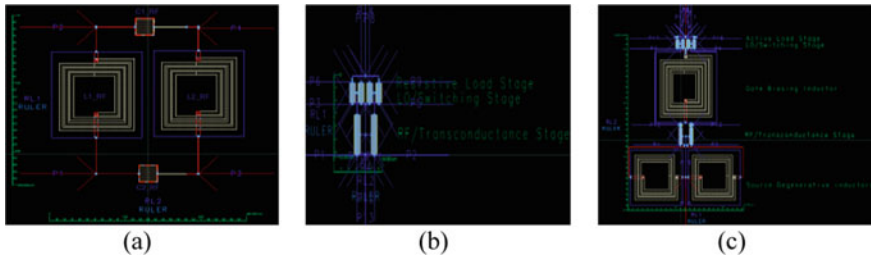


Fig. 10 Core layouts. **a** *LC* Balun for differential *RF* inputs. **b** *CDBM*. **c** *PDBM*

References

1. Response to TRAI consultation paper on proliferation of broadband through public Wi-Fi networks. <https://traai.gov.in>
2. Wireless planning and coordination wing, national frequency allocation plan-2011, and 2018, Ministry of Communications, Department of Telecommunications, Government of India
3. Wi-Fi Alliance, <https://www.wi-fi.org>

4. Zumbahlen H (2006) Basic linear design. Analog Devices
5. Razavi B (2013) RF microelectronics, 2nd edn. Pearson Education India
6. Leung B (2011) VLSI for wireless communication, 2nd edn. Springer
7. Oxley TH (2002) 50 years development of the microwave mixer for heterodyne receivers. *IEEE Trans Microw Theor Tech* 50(3)
8. Gilbert B (1997) The MICROMIXER: a highly linear variant of the gilbert mixer using a bisymmetric class-AB input stage. *IEEE J Solid-State Circ* 32(9)
9. Mini-Circuits (2015) How to select a mixer. File: AN00010.doc
10. Agilent EEs of EDA (2008) Overview on mixer simulation with agilent's ADS, mixer simulation with HP advanced design system. Technical Note
11. Gray PR, Meyer RG (1993) Analysis and design of analog integrated circuits, chapter 10, 3rd edn. Wiley, New York
12. Sulivan PJ, Xavier BA, Ku WH (1997) Low voltage performance of a microwave CMOS Gilbert cell mixer. *IEEE J Solid-State Circ* 32(7):1151–1155
13. Vidjokovic V, Tang J, Leeuwenburgh A, Roermund AH (2005) A low-voltage folded-switching mixer in 0.18 μm CMOS. *IEEE J Solid-State Circ* 40(6)
14. Huang MF, Kuo CJ, Lee SY (2006) A 5.25-GHz CMOS foldedcascode even-harmonic mixer for low-voltage applications. *IEEE Trans Microw Theor Techn* 54(2)
15. Cruz H, Lee L (2015) A 1.3 mW low-IF, current-reuse, and current-bleeding RF front-end for the MICS band with sensitivity of 97 dBm. *IEEE Trans Circ Syst I, Reg Papers* 62(6):1627–1636
16. Hsieh HH, Lu LH (2007) Design of ultra-low-voltage RF frontends with complementary current-reused architectures. *IEEE Trans Microw Theor Techn* 55(7)
17. Tan R, Mak M (2017) A 0.35-V 520- μW 2.4-GHz current-bleeding mixer with inductive-gate and forward-body bias, achieving >13-dB conversion gain and >55-dB port-to-port isolation. *IEEE Trans Microw Theor Tech* 65(4)
18. Behzad R (1997) Design considerations for direct-conversion receivers. *IEEE Trans Circ Syst II: Analog Dig Sig Proc* 44:428–435
19. Abidi Asad A (1995) Direct-conversion radio transceivers for digital communications. *IEEE J Solid-State Circ* 30:1399–1410
20. Gladson B, Praveen S (2019) A 261-mW ultra-low power RF mixer with 26-dBm IIP3 using complementary pre-distortion technique for IEEE 802.15.4 applications. *Int J Electron Commun (AEÜ)* 107:70–82
21. Ganesan S, Sanchez-Sinencio E, Silva-Martinez J (2006) A highly linear low-noise amplifier. *IEEE Trans Microw Theor Techn* 54(12):4079–4085
22. Aparin v, Larson LE (2005) Modified derivative superposition method for linearizing FET low-noise amplifiers. *IEEE Trans Microw Theor Techn* 53(2)
23. Kashani A, Yavari M (2021) A +7.6 dBm IIP3 2.4-GHz double-balanced mixer with 10.5 dB NF in 65-nm CMOS. *IEEE Trans Circ Syst II Exp Briefs*
24. Li C, Lai H, Liu H (2014) A feed forward noise and distortion cancellation technique for CMOS broadband LNA-mixer. *IEEE Asian Solid-State Circ Conf*
25. Li H, Saavedra CH (2019) Linearization of active downconversion mixers at the IF using feedforward cancellation. *IEEE Trans Circ Syst I* 66(4)
26. SRVS, Avvaru S (2020) Design and optimization of double balanced gilbert cell mixer in 130 nm CMOS process. *Solid State Electron* 2:129–139
27. Gomez FRI, De Leon MTG (2019) A study of zero-if double-balanced mixer for wimax receivers. *Heliyon* 5:e01741
28. NacEachern LA, Manku T (1998) A charge-injection method for Gilbert cell biasing, conference proceedings. In: *IEEE Canadian conference on electrical and computer engineering*
29. Piccinni G, Talarico C, Avitabile G, Coviello G (2019) Innovative strategy for mixer design optimization based on gm/ID methodology. *Electronics* 8:954
30. Kumar SD (2020) A 2.4 GHz double balanced downconversion mixer with improved conversion gain in 180 nm technology. *Microsyst Technol* 26:1721–1731
31. Pimenta TC, Crepaldi PC, Ferreira LHC, Moreno RL, Zoccal LB (2011) Main RF structures-current trends and challenges in RFID. IntechOpen Limited

32. Impedance matching, Lumped LC Balun, LC Balun calculator, analysis, design equations and examples. <https://analog.intgckts.com/impedance-matching/lumped-lc-balun>
33. Lee T, LShen C (2013) A current bleeding CMOS mixer featuring LO amplification based on current-reused topology. *Circ Syst* 4(1):58–66
34. Wan Q, Xu D, Hui Z, Jun D (2018) A complementary current mirror-based bulk-driven down-conversion mixer for wideband applications. *Circ Syst Sig Proc*
35. Bhatt D, Mukherjee J, Redouté J-M (2017) A self-biased mixer in 0.18 μm CMOS for an ultra-wideband receiver. *IEEE Trans Microw Theor Tech*
36. Neda S, Abdolreza N (2017) Design of an active CMOS subharmonic mixer with enhanced transconductance. *Int J Electron Comm (AEÜ)* 73:98–104

Performance of a MIMO-OFDM-Based Opto-Acoustic Modem for High Data Rate Underwater Wireless Communication (UWC) System



CH. Pallavi and G. Sreenivasulu

Abstract This paper fundamentally bases on the arrangement of MIMO-OFDM-based opto-acoustic modem using MATLAB-Simulink to additionally foster the data speed of an optical-acoustic correspondence in underwater wireless communication (UWC) system. Orthogonal frequency-division multiplexing (OFDM) and multiple input multiple output (MIMO) advancement for the most part used in far-off associations to further develop the data rates, protection from multipath blurring, and an augmentation in the channel's spatial multiplexing and spatial diversity gain, etc. The OFDM design segregates the available bandwidth information into a number of overlying sub-channels. This multipath spread on the channel's removes the inter symbol interference (ISI), and hence, the available bandwidth limit can be improved. Therefore, the OFDM methodology was directed towards to reduce the ideal subcarriers in the channel conveyed as the bit error ratio or rate (BER) or (Eb/No). Using the MATLAB-Simulink tool, the opto-acoustic modem was designed for the transmission and reception of optical and acoustic (EM) signals, respectively. From the simulation results, it is clear that the proposed method was very accurate and efficient. The BER of acoustic (EM) and optical signals were calculated successfully. Thus, by combining the MIMO and OFDM strategy in single opto-acoustic modem will grow the spectral efficiency, gain, and bandwidth of the channel and furthermore to work on the exhibition of submerged remote correspondence framework with single opto-acoustic modem through optical and acoustic (EM) signals. Therefore, the simulation results showed that, there is an improvement in the performance of an opto-acoustic modem that is, reduction in BER.

Keywords Underwater wireless communication (UWC) · Optical signal · Acoustic (EM) signal and bit error ratio or rate (BER)

CH. Pallavi (✉) · G. Sreenivasulu

Department of ECE, S V University College of Engineering, S V University, Tirupati, Andhra Pradesh, India

e-mail: pallavich.svu@gmail.com

1 Introduction

In the current scenario, the UWC increases rapidly due to a lot of applications in far-off controller in seaward and smoke stations, smog monitoring, weather nursing in eco-friendly environment, military applications, recorded scientific data collection at the sea-bottom areas and unmanned subaquatic vehicles, diver communications, recording of the deep-sea level for finding the things and objects, etc. Acoustic, electromagnetic (RF), optical, and fibre optics are the communication channels in the UWC. Every channel has its own advantages and disadvantages depends on data rate capacity, frequency, and distance, respectively. In seawater, electromagnetic (RF) and optical signals have the poor propagation that is in UWC, acoustic signals are only the feasible solution for high range. The speed of acoustic spread in water, is around 1500 m/s, contrasted and that of electromagnetic and optical waves, is one more restricting component for productive correspondence and systems administration [1, 2]. Acoustic frameworks are capable for significant distance correspondence. Be that as it may, acoustic correspondence has limits. They have exceptionally low information rates for checking applications and because of speed of sound in water, it has an immense dormancy which implies a critical time delay. To adapt up to the constraints of acoustic correspondence, we utilize optical correspondence whose frequency lies in the apparent area. Optical correspondence is known for high information rate and low idleness.

A novel and diverse beneath water wireless communications are presented by relating to wired and wireless communications available in environment, with associated cultured communication strategies to enhance comparatively small transmission rates including short range. Truly, the submarine atmosphere includes several characteristic features which makes distinct from atmospheric radio propagation in comparative use of traditional communications. Beneath water influenced communication system by salt concentration, humidity, pressure, light intensity, air effecting the waves are represented a few [3, 4]. The performance analysis of an UWC system is depicted in Fig. 1 [5]. A unique gateway furnished with many antennas, and an arbitrary number of individual bases are buried from their endpoint(s) uses a zero-compelling strategy by employing the space division multiple access technique. The bases are sited in underneath water, and the main mean to interface them with their endpoint(s) over a door, which turns as amplify and forward electric transfer [5].

In the blue-green region, light waves circulate properly with low attenuation, scattering, and inadequate distances of several metres [6]. When acoustic signals circulate with heavy gravity, then it can simply transportable in many kilometres, however, cover a more extended distance at lesser recurrence rate. Compared to extra-terrestrial communication, the auditory communication has less bandwidth. Currently, auditory modems work with limited bandwidth as kHz. Even though beneath water, acoustic (auditory) communication is best over some thousand kilometres [7]. In demand to overwhelmed the complications in beneath water acoustic communication, use under (beneath) water optical communication system, which overwhelms the boundaries of auditory communication but the optical communication is right for short distances

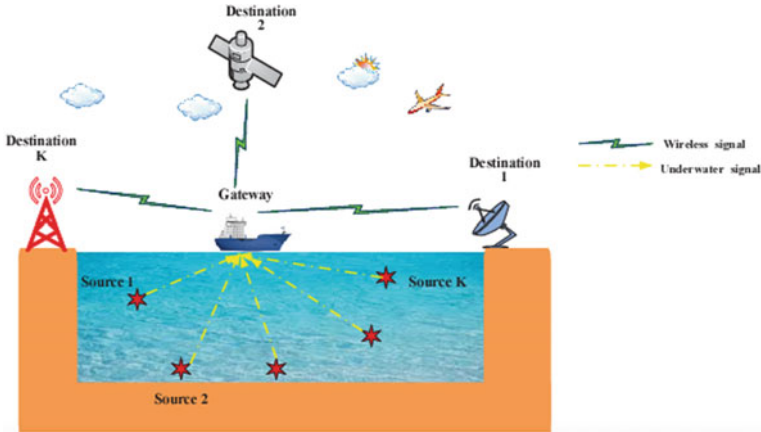


Fig. 1 Performance analysis of UWC system [5]

with highest data rates. When merging the both acoustic and optical communication into a solitary modem, it overwhelms the problems of individuals. The design of an opto-acoustic modem based on MIMO-OFDM using MATLAB-Simulink tool was recommended. For current wireless communication systems, large information rate is an utmost leading constraint. In the past decades, the ideology of OFDM modulation has its important. Therefore, in recent communications systems, these practices are widely used. Radiocommunication networking, data communication, digital radio receiver, and television are some of the applications of OFDM. One of the most important applications of a parallel-data communication scheme is the OFDM. So that, the effect of multipath fading has been reduced [8, 9].

The technique to use of several aerials at the sender and receiver sections to expand communication performance is the MIMO. The smart antenna system uses this scheme. Now this scheme should be considered for radiocommunications like 4G, 5G, and 6G also, for it offers substantial growths in data output and linkage without other bandwidth or improved diffused power [9, 10]. To achieve this goal, the aerials must fully transmit and evenly distribute the power distribution done by the aerials to realize an array gain with expanded bandwidth efficiency. Therefore, MIMO plays a significant role in the current radiocommunication system with these properties. The mixture of OFDM system with MIMO scheme [9, 11] was viewed by means of an auspicious resolution for radiocommunication systems with enhanced signalling rates. Therefore, the baseband signal processing uses the MIMO-OFDM with less number of deviations. This typical model is used to simulate the bit error rate (BER) performance of an opto-acoustic modem with MIMO-OFDM by various modulation techniques such as QAM, QPSK, and MSK.

2 Related Works

A few related previous works in UWC using OFDM, acoustic, and optical modems are described in this section. The techniques and the result of their works are underlined. At long last, every one of the strategies is utilized to work on the constraints of these procedures and in the light of these methods that will be focused in the exploration work. Submerged acoustic/optical correspondence is one of the most difficult assignments for transmission, and it is restricted by the accompanying boundaries are (i) restricted transfer speed on the grounds that the sign lessening increments by expanding distance and recurrence, (ii) time variation multipath spreading, (iii) signalling rate, (iv) BER, (v) SNR, and (vi) reduced auditory signals in water and power also. Restricted data transmission is a significant issue in underneath water acoustic/optical (UWAO) channels, and acoustic/optical waves in UWAO climate are caught up in high frequencies. Thus, the offered transmission capacity is limited to kHz. Consequently, utilizing techniques that can further develop the signalling rate as well as data transfer capacity is vital. One such effective method is to combine the multi-transporter OFDM approach with MIMO, which increases bandwidth efficiency by distributing information equally among a variety of transmitters and recipients [2, 12].

Alessandro [13] explored transmission plot for OFDM. The adaptive communication structure is characterized by connecting their exhibition through static communication arrangement. To expand the throughput output by using a well-known mathematical technique. The signal-to-noise ratio (SNR) of the subcarriers in the substituting limit was defined by the above calculation method. Therefore, there is an improvement in BER calculation with better throughput outcome [2]. Marwa [14] examined the channel coding procedures through additive white Gaussian noise (AWGN) and hence, improvement in OFDM framework by utilizing this channel model. Reed Solomon coding, convolutional coding, concatenated coding (by consolidating Reed Solomon with convolutional), and inter leaved linked coding methods are some of the examples [2].

In paper [15], due to numerous dissemination singularities, it is necessary to comprehend the viability and dependability of high signalling rate subaquatic optical links in order to improve the underwater optical communication system performance. In paper [16], presented a development of acoustic modem is the optical modem with a greater speed, less power, and with small complicated frameworks are required. A low cost and low power acoustic modem has been designed in paper [17] to have better short-range underwater communication. In paper [18], the optical and acoustic signals were combined in a single integrated modem probable for UWC by means of OFDM. In what way the long-distance communication was possible with a low power dissipation, large bandwidth, and greater speed which was discussed (M. Lenin Kumar et al., 2019).

3 OFDM, MIMO System, and MIMO-OFDM

The key perception in OFDM is the symmetry of the sub-transporters. In the OFDM, the “symmetrical” part of displays that there is a specific scientific rapport among the frequencies of the transporters in the framework. Assembling of the transporters in an OFDM signal is a basic requirement. Thus, the distinct carrier signals are joined by the sidebands and there is no adjacent channel interference. Therefore, the transporters must be scientifically symmetrical to each other. The design of a transporter is a multiple of $1/T_s$, then the transporters are unique, where T_s is the symbol duration and the spectrum of OFDM with five symmetrical frequencies as shown in Fig. 2 (Shadrach Kukuchuku et al., 2018).

The symmetry among the transporters can be defined by the OFDM signal and is characterized by utilizing Fourier transformation techniques. The OFDM framework communicates countless narrowband transporters, which are firmly dispersed. Note that at the focal recurrence of each sub-channel, there is no crosstalk from other sub-channels. The symmetry permits synchronous transmission on a great deal of sub-transporters with a tight recurrence space without obstruction from one another (Shadrach Kukuchuku et al., 2018).

Effect of Recurrence Offset in OFDM:

The existing broadband among a set of impertinent overlying subcarriers is splitted by OFDM. Thus, the presence of a transporter recurrence offset can present extreme twisting in an OFDM framework, as it brings about a deficiency of symmetry among the subcarriers. Hence, the presence of a transporter recurrence offset introduces inter-symbol interference (ISI) in the OFDM system. Therefore, the impact of frequency offset on the performance of the OFDM system is defined by considering a recurrence offset Δf such that,

$$\epsilon = \Delta f \div BW, \tag{1}$$

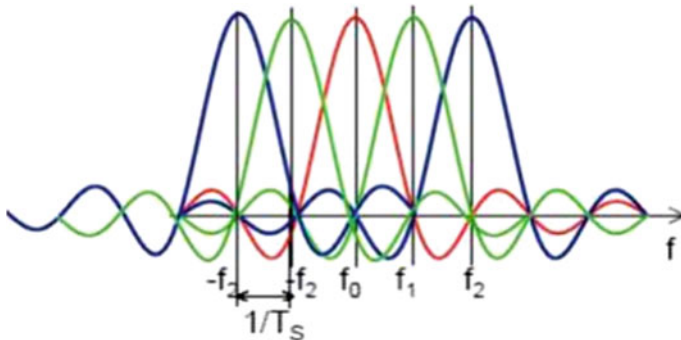
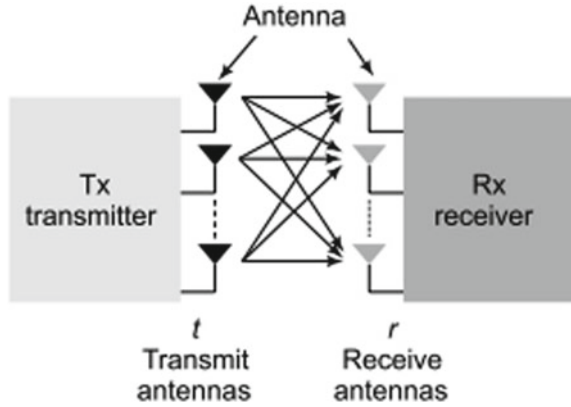


Fig. 2 Spectrum of OFDM with five symmetrical frequencies [19]

Fig. 3 MIMO system [20]



where ϵ denotes the normalized frequency offset, normalized with respect to the subcarrier bandwidth (BW) [20].

MIMO System: MIMO system consists of many antennas at the Tx. and Rx. Sectors. A MIMO system with multiple antennas is depicted in Fig. 3. The strength of the signal can be improved by employing the many antennas in the MIMO system [20].

MIMO-OFDM: The combination of the MIMO and OFDM enhances the transmitting capability of the several antennas with different signals and provides more reliable and secure communications at large signalling speeds, higher data rates, greatest spectral efficiency or large bandwidth efficiency, minimum BER, and low latency.

Like OFDM, the frequency-selective MIMO channels are converted into multiple parallel flat fading MIMO channels using MIMO-OFDM. Therefore, it significantly resolves the baseband signal processing by simply reducing the complex MIMO equalizer. The schematic representation of a MIMO-OFDM Tx. and Rx. sections is illustrated in Fig. 4a and b for broadband wireless communication systems [20].

4 Design of a MIMO-OFDM-Based Opto-Acoustic Modem

The design of a MIMO-OFDM-based opto-acoustic modem using MATLAB-Simulink for large signalling rate UWC system is illustrated in Fig. 5. This developed model consists various blocks that govern the essential features of the opto-acoustic modem using MIMO-OFDM technique. The transmitter section components are described in the uppermost row, and the receiver section components are described in the lowermost row of the opto-acoustic modem. In this model, simultaneously can send optical signal and acoustic (EM) signal as an input by using toggle switch called variant sink. When $V = 1$, it selects acoustic signal and it selects light (optical) signal when $V = 2$, respectively. The bit error ratio of acoustic (EM) and light (optical)

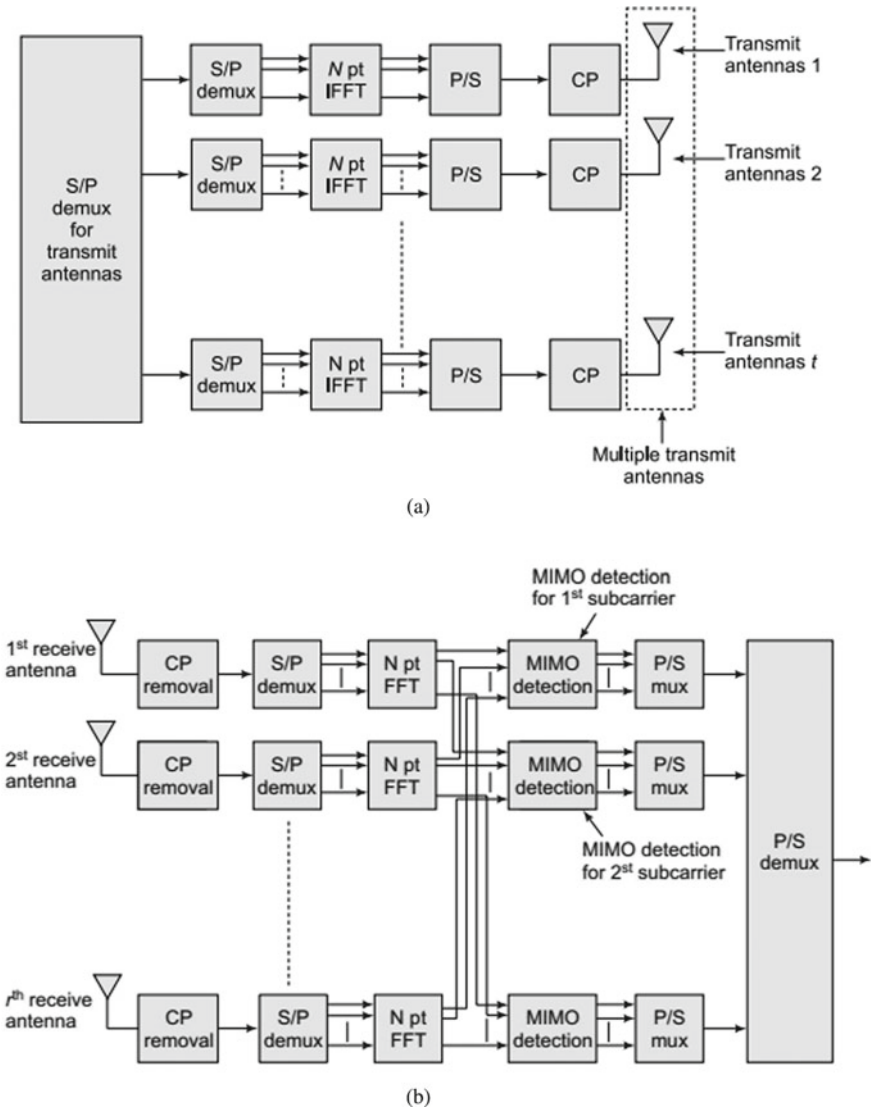


Fig. 4 a MIMO-OFDM Tx and b MIMO-OFDM Rx [20]

signals were calculated successfully. Therefore, by using this adaptive modulation technique called MIMO-OFDM improves the opto-acoustic modem performance by increasing the speed of propagation, signalling rate for UWC system with minimum BER and good latency. The function of each block in the designed opto-acoustic modem using MATLAB-Simulink was explained below.

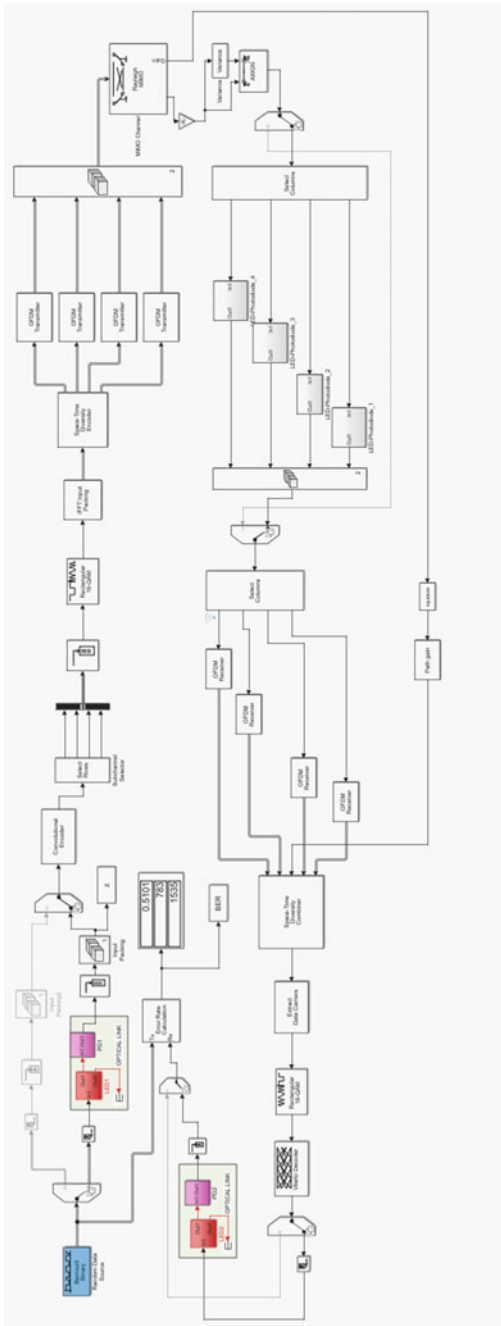


Fig. 5 MIMO-OFDM-based opto-acoustic modem using MATLAB-Simulink

Bernoulli Binary Generator: The Bernoulli binary generator block creates arbitrary binary numbers with the use of a Bernoulli distribution. This block is used to produce random data bits and to simulate digital communication systems and attain the performance system of measurement such as bit error rate (BER).

Variation Sink and Variation Source: Variation sink is a closure switch that initiates single of its optional picks at the output to accept the input. It consists dual or extra ports in which one input and two or more output ports, respectively. An optional controller is linked to every output port. It chooses either EM path or optical path [21]. The variation source offers discrepancy on the source of a signal. Blocks associated to the input ports explain different choices, and during simulation, only one input port must be active. Blocks related to the inactive ports will be removed from the simulation. The variation control governs the input port or if any is active. It selects either EM signal or optical signal [21].

Unbuffer and Buffer: Disables the output buffering when it is redirected from the input. That is, inputs are row-wise, every matrix row alters the independent variable into the output. The volume of the structure accepts the inputs is normally less than the actual volume. The rectangular buffer constantly executes outline-based measurement. The buffer reforms the information in every segment of the contribution to yield a result with a variable frame size. Buffering a signal to a larger frame size induces an output with a low frame than the input. Buffering a signal to a smaller frame size induces an output with a nearer frame than the input. When the input and output frames are same, then this block provisions prompted subsets.

Vector Concatenation Block: It defines a series of input signals of similar data type can be used to generate the output signal. The vector concatenate and matrix concatenate blocks concatenate the input signals to produce a nonvirtual output signal whose basic elements are exist in the memory. In the Simulink library, these blocks have changed formations of the identical block. The change is built on the set of their methods, which governs the block in vector or multidimensional array chain mode.

LED-Photodiode: LED causes light with the help of a charge carriers while photodiode produces current due to incident photons. In a nutshell, LED translates electric energy into light energy but photodiode alters light energy into electrical energy. LED is a light source, used to emit the light and changes the light intensity, when signal flows through it. Photodiode is a device when it is exposed to light, induces an electrical current.

Convolution Encoder: A stream of binary input vectors is converted into a stream of binary output vectors by using this encoder. It is closely related to generator polynomials, and it receives inputs that vary in length during simulation process. If the encoder has 'n' input bit streams, then it gives 2^n possible input bit streams. Similarly, when the encoder receives 'm' output bit streams, then it has 2^m possible output bit streams. To state the convolutional encoder, use the "Trellis structure" parameter.

Sub-channel selector/Input Packing: The translation of a single user data type or multiple user data type in to a single valid data type was done by the process called input packing.

Bus Creator: The bus creator combines a set of input elements into a bus and can connect any type of element to the input ports and with the other buses. The bus contains unique elements. In default, each element of the bus gets the name of the element coupled to the bus creator block.

Rectangular QAM Modulator: It is a one of the digital modulation schemes used in communication system for transmitting the data or message. QAM modulation changes the amplitude, or power level, of two signals. QAM modulation effectively transmits an analogue data into digital data or information and also increases the available bandwidth.

IFFT Input Packing: Packing data to perform IFFT.

Space Time Diversity Encoder (Frame Conversion): Space–time code offers both the diversity and coding gain while using several transmit antennas to enhance the radiocommunication system with efficient.

OSTBC Encoder: It is a class of linear STBC code and encodes the input message using an orthogonal space–time block code (OSTBC). It is a most popular and widely used STBC code with full spatial diversity. It is used to define the rate of two sender antennas as 1, for 1/2 or 3/4 uses three and four transmit antennas [21].

OFDM Transmitter: Figure 6 shows the simulated block diagram of OFDM transmitter and the working of each block is described as follows:

Reorder: The multidimensional elements of an input signal can be defined by select or reorder and used to guide the input port with each element.

IFFT: It executes IFFT operation.

Add Cyclic Prefix: The robustness of an OFDM signal can be added by using the cyclic prefix and also reduces the inter symbol interference.

Reshape: Reshape the data and translate parallel data in to serial data.

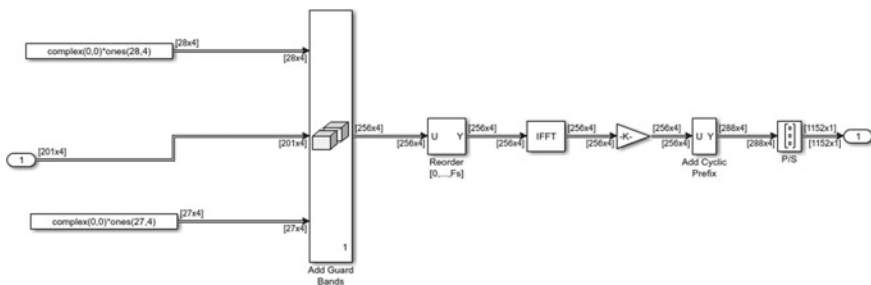


Fig. 6 OFDM transmitter



Fig. 7 OFDM receiver

Combine four Tx. signals for feed into MIMO Channel: Matrix concatenation is the use of square brackets to join existing matrices together. The way of generating a matrix is called concatenation.

MIMO Channel: Referring to the MIMO channel, when a data to be sent through a channel using more than one aerial and uses multiple aeriels while emerging from the channel. The channel capacity of a MIMO channel can be improved by increasing the number of transmitting aeriels and receiving aeriels, respectively.

Gain and Variance: It is used to set the gain and variance of a AWGN channel.

AWGN Channel: It adds the white noise with a constant amplitude of spectral bandwidth to the given input signal. The input signal may be a real or complex. It supports multichannel processing.

Multiport Selector: It picks the 4 output ports for 4 transmitted signals into one. The extraction of various subsets of rows or columns from the given input matrix with size $M \times N$ and circulates a new submatrix into a discrete output.

LED-Photodiode: Implements the transmission and reception through light in optical path.

Matrix Concatenation: Once again concatenate 4 different transmitted signals into one.

Squeeze: The unity dimensions (1D) of a multidimensional input signal can be removed by using squeeze method. For suppose, a $3 \times 1 \times 2$ signal should be converted into a 3×2 signal. The unmovable signals are 1D and 2D.

Path gain: Multiport selection of 4 transmitted signals, serial to parallel conversion, removal of cyclic prefix, DC components, etc.

OFDM Receiver: Figure 7 shows the simulated block diagram of OFDM receiver.

Reshape: Converts the serial data to parallel data.

Remove/Reorder: Used reorder block.

FFT: Implements FFT operation.

Frame Conversion: Converts sample-based signal into frames.

Removal of Guard bands: Select and reorder the data.

STBC Combiner: It performs frequency evaluation, removal of DC component, etc.

OSTBC Combiner: This is same as that of OSTBC block used in transmitter section and performs reverse operation to that.

Extract Data Carriers: Used to select data and remove pilots and parallel to serial conversion.

Rectangular QAM Demodulator: It is similar to the rectangular QAM modulator, but it performs reverse modulation.

Deconvolutional Encoder (Viterbi Decoder): The Viterbi algorithm is used to decrypt the encoded input data and decrypts input ciphers to yield digital symbols. A trellis limit defines the poly2trellis function [21].

Error Rate Calculation: It is used to calculate the error rate ratio of the given input signals like acoustic, optical or RF signals, etc., which calculates only the error rate or ratio but not the amplitude.

5 MATLAB Simulation Results

The MIMO-OFDM-based opto-acoustic modem was designed and simulated using the MATLAB-Simulink. The performance evaluation of a MIMO-OFDM-based UWC system was examined by transmitting the acoustic and optical signals through a toggle switch. Figures 8 and 9 show the error rate plot for both acoustic (EM) and optical (light) signals, respectively. These constraints were implemented to learn and assist the origin of MATLAB environment modulation. Therefore, this performance metrics shows that the better BER behaviour and good SNR. The graphical representation of an error rate or ratio plot for both acoustic and optical signals was done by using the MATLAB simulation environment. When $V = 1$, the acoustic (EM) signal is selected and the error rate (BER) achieved is 0.49 and for $V = 2$, optical (light) signal is selected, the error rate achieved is 0.51. The acoustic (EM) and optical (light) signals transmit the number of possible bits from 0 to 1535. Therefore, from the simulation results, it is cleared that there is an improvement in the UWC system performance with the use of MIMO-OFDM-based opto-acoustic modem is a good choice in the most challenging underwater communication channels.

r

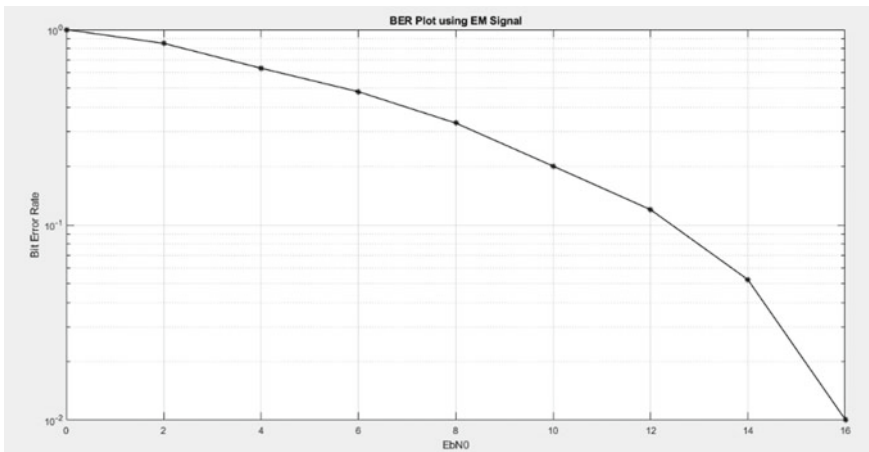


Fig. 8 Error rate plot for acoustic (EM) signal

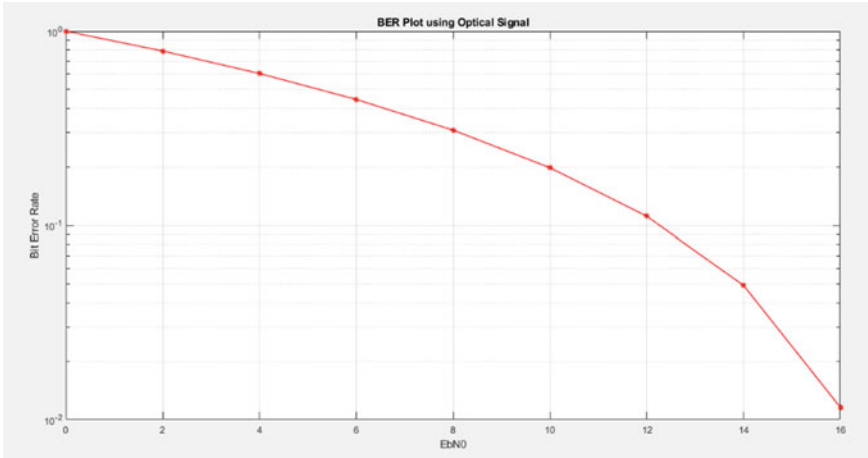


Fig. 9 Error rate plot for optical (light) signal

6 Conclusion

The propagation of acoustic (EM)/optical (light) communication in underwater through MIMO-OFDM is a relatively new scheme, and hence, it requires different digital modulation schemes with new procedures. To enumerate all these techniques with good characteristics and parameters like signalling rate, error rate (BER), bandwidth efficiency and range, etc. An efficient and effective channel scheme in underwater was still hard to find in high-speed communication environment. The design of a MIMO-OFDM-based opto-acoustic modem was proposed here, and both optical (light) and acoustic (EM) signals were communicated through a toggle switch. Both optical (light) and acoustic (EM) signals have unique channels for transmission and reception. In MATLAB, there are no audio files to give the speech signal as an input to acoustic channel, so EM signal was given as an input to acoustic channel and for optical channel, light signal was given as an input, respectively. Both optical/acoustic signals were simultaneously transmitted to an opto-acoustic modem by using a toggle switch provided in the design. In this modem, both optical and acoustic signals were communicated through a single modem based on MIMO-OFDM technique. In the previous developed modems, optical signals and acoustic signals transmitted through individual modems. That is, one modem used for acoustic signals and another modem for optical signals separately. Therefore, it requires two separate modems for communication. So, it reduces the design complexity and cost of the modem and also increases the speed of the underwater wireless communication system through a single modem. Using the MATLAB platform, the functionality of the designed modem was simulated and the BER performance of the optical and acoustic signals were calculated. Hence, it was clearly observed from the simulated results that, there is a greater improvement in the signalling rate for the both optical and acoustic (EM) signals and this method was found to be the best. The designed opto-acoustic

modem proved that there is an enhancement in the performance in terms of error ratio, channel bandwidth efficiency, large signalling rate and speed of frequency, etc. This paper gives the comprehensive research on UWC system through a single modem by sending acoustic/optical signals and also aims to provide innovative designs that would help in the improvement of future UWC system.

References

1. Shengli Z, Zhaohui W (2014) OFDM for underwater acoustic communications. Wiley, Chichester, England. <https://doi.org/10.1002/9781118693865.ch2>
2. Shadrach Kukuchuku et al., (2018) Improved underwater wireless communication system using OFDM technique. *Am J Eng Res (AJER)* 7(09):82–89
3. Lurton X (2010) *An introduction to underwater acoustics: principles and applications*, 2nd edn. Springer
4. Lanzagorta M (2012) *Underwater communication*, ser. synthesis lectures on communications, vol 5. Morgan & Claypool Publishers
5. Alfitouri A, Hamdi KA, Hadi A, Alkhaled M (2017) Performance analysis of underwater-wireless communication with zero forcing precoder. In: 2017 IEEE international black sea conference on communications and networking (BlackSeaCom). <https://doi.org/10.1109/BlackSeaCom.2017.8277661>
6. Dayal A (2016) Nonlinear doppler warp correction for acoustic OFDM
7. Farhana E (2012) Performance evaluation of orthogonal frequency division multiplexing (OFDM) based wireless communication system with implementation of least mean square equalization technique
8. Koringaand RD, Patel TS (2012) Design and analysis of bit error rate performance of simulink based DSSS-OFDM model for wireless communication. *Int J Eng Res Technol (IJERT)* 1(3)
9. Ananda Murthi LS, Bhagya R (2013) Design and performance analysis of Mimo-Ofdm for WLAN standard. *Int J Electron Commun Eng (IJECE)* 2(3):77–84. ISSN 2278-9901, © IASET
10. van Zelstand A, Schenk TCW (2004) Implementation of a MIMO OFDM-based wireless LAN system. *IEEE Trans Sig Proc* 52(2)
11. Schenk TCW, Dolmans G, Modonesi I (2004) Throughput of a MIMO OFDM based WLAN system. In: *Proceeding symposium IEEE Benelux chapter on communications and vehicular technology*
12. Muquet B, Wang Z, Giannakis G (2002) Cyclic prefix or zero padding for wireless multicarrier transmissions
13. Alessandro T (2013) Efficient OFDM channel estimation via an information criterion in IEEE transactions on wireless communications, vol 22, Issue. Print ISSN 1536-2276
14. Li B, Zhou S, et al (2008) Scalable OFDM design for underwater acoustic communications. In: ICASSP, IEEE international conference on acoustics, speech and signal processing—proceedings (March), pp 5304–5307. <https://doi.org/10.1109/ICASSP.2008.4518857>
15. Kaushal H, Kaddoum G (2016) Underwater optical wireless communication. *IEEE Access* 4:1518–1547. <https://doi.org/10.1109/ACCESS.2016.2552538>
16. Gayathri CB, Singh D, Durga Satya Dhanusha M, Narendara Raj V (2015) Design of high speed underwater optical communication using on-off keying algorithm. *IEEE ICCSP 2015 conference*, pp 1355–1360
17. Wills J, Ye W, Heidemann J (2006) Low-power acoustic modem for dense underwater sensor networks. In: *Wireless underwater communications'06*, pp 79–85
18. Lenin Kumar M, Janaki Rani M (2019) A design of novel hybrid opto-acoustic modem for underwater communication. *Int J Innov Technol Explor Eng (IJITEE)* 8(8). ISSN: 2278-3075

19. Wu Z, Li X (2015) An improved underwater acoustic network localization Algorithm. J Commun 12(3):77–83
20. Jagannatham AK, Principles of modern wireless communication systems theory and practice. Mc Graw Hill Education (India) Private Limited New Delhi
21. www.mathworks.com

A New Hybrid Islanding Detection Scheme Using Discrete Wavelet Transform and Artificial Neural Networks



M. Krishna Goriparthi and B. Geethalakshmi

Abstract In today's world, sources of renewable energy (RES) using PV arrays are the most extensively used. When RES is connected to the grid, there will be some issues owing to the unexpected circuit breakers connected to the grid trip, which creates islanding. This islanding condition should be detected within two seconds, as per IEEE standards. This paper presents frequency disturbance triggered hybrid islanding detection using artificial neural network (ANN) and discrete wavelet transformation (DWT). The ANN model is trained by these WT features and this approach calculates the detection time for various loading and non-islanding conditions. DWT analysis is performed up to level 4 in this work which is fed to an ANN model to predict islanding detection time. Simulation of frequency triggered hybrid islanding detection approach is implemented on the MATLAB 2018b platform. Python 3.9.5 is used for the discrete wavelet transform and ANN.

Keywords ANN · Current injection · Frequency disturbance · DWT · Islanding detection

1 Introduction

Integration of distributed generators into electrical grid will raise serious concerns about the stability and safe operation of the grid [1]. The nature of DG to power local loads even though the electric utility is unable to supply power is called islanding. The deviations in electrical parameters will cause significant effects on electrical appliances in the islanded segment connected to it. Moreover, according to IEEE 1547-2003 standards, detection of islanding condition has to be done within 2 s of its occurrence. Islanding detection schemes are broadly divided as active, passive, and hybrid. Operation of passive islanding methods depends on the information available

M. K. Goriparthi (✉) · B. Geethalakshmi
Department of Electrical and Electronics Engineering, Puducherry Technological University
Puducherry, Puducherry, India
e-mail: goriparthymuralikrishna@gmail.com

on the DG side which helps to detect islanding conditions. On the other hand, these schemes are affected by broad non-detection zone (NDZ). Whereas, active islanding schemes are based on periodic perturbation signal transmission between grid and DG to determine islanding detection and these techniques have less NDZ. However, active methods are complex [2]. Hybrid islanding detection techniques are far more effective than other islanding detection techniques. But these are complex and costlier.

2 Test System and Methodology

The test system details are given in Table 1 [3].

It generally consists of grid which is connected to DG with a parallel RLC load using a synchronous phase-locked loop (PLL) as shown in Fig. 1. To operate DG in constant power control mode, switching pulses are required for VSC which are generated with a pulse width modulation (PWM). Discrete PLL blocks are used to measure the frequency deviation at DG (f_{DG}) and grid side (f_G). If frequency deviation (f_d) exceeds threshold value (Th_1), then output pulse will be sent to the PWM controller to initiate disturbance current injection using the current controller

$$|f_G - f_{DG}| > Th_1. \quad (1)$$

A current with a low frequency of disruption (i_d^*) will be introduced after occurrence of disturbance event [4]. It is 1% of inverter DG current, and the introduction will be done through the d -axis for a period of time current controller with 0.3 s as represented in Fig.2.

$$i_d^* = ki_d \cos(\omega_d t), \quad (2)$$

Table 1 List of parameters

S. No.	Variable	Values
1	PV array	P_{max} : 100 kW at 1000 W/m ² sun irradiance.
2	Grid voltage	125 kV distribution feeder + 120 kV equivalent transmission system
3	DC-DC boost converter	Step up 273–500 V DC.
4	3-phase voltage source converter	Convert link voltage of 500–260 V AC at unity power factor.
5	Sampling time	100 microseconds
6	Capacitor bank	10 KVAR
7	Step up transformer	100 kVA 260 V/25 kV

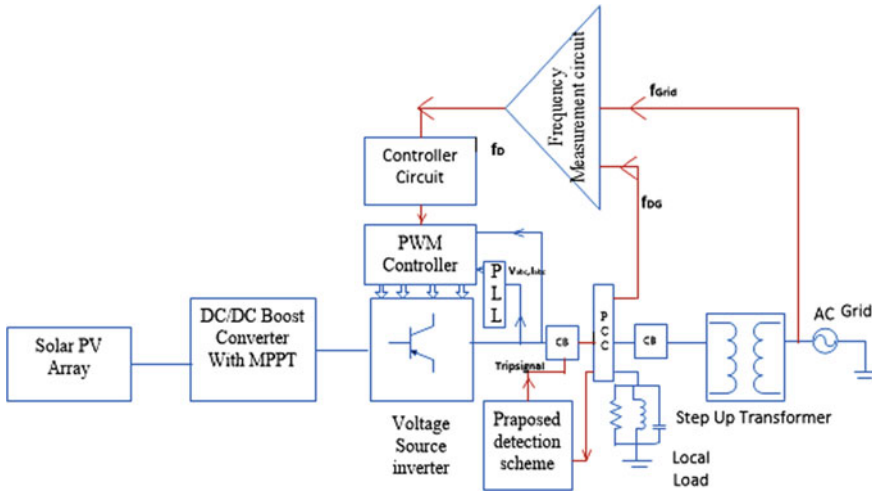


Fig. 1 Test system

where $i_d = d$ -axis current; $\omega = \text{constant}$ its value lies between 0.01 and 0.03; $\omega = \text{disturbance current frequency}$.

Fault detection parameters are identified by a passive islanding parameter called rate of change of phase angle between positive sequence voltage and current (RCPABPSVAC). Due to the difficulty of analyzing DG voltage and current signals in abc frame, these quantities will be converted into dq reference frame. Circuit breakers are placed at both DG and grid sides to simulate islanding conditions and also to isolate DG at the instant of islanding. DWT analysis is done up to level 4 to generate various wavelet coefficients. ANN model algorithm is utilized to determine fault detection time for various fault conditions under grid-connected mode [5].

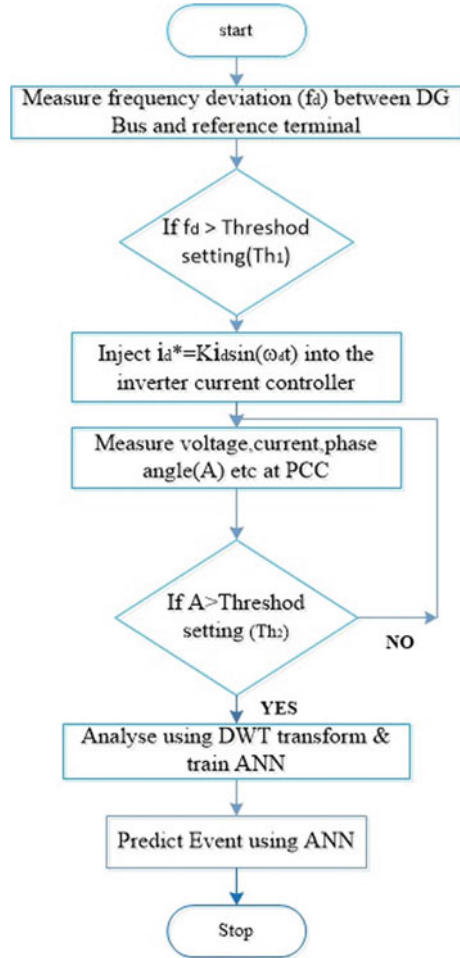
3 Transformation of Discrete Wavelets

Signal and speech processing applications are best served by wavelet transforms and ANN. A wavelet transform consists of the location and scale of two basic components. A wavelet transform consists of a succession of wavelet functions of various sizes. The HPF and LPF can be used to create lower resolution components. This decomposition process is repeated until all of the components have been created. The objective signal (S) is sent through the HPF and LPF. The wavelet transform is separated into two types, continuous (CWT) and discrete (DWT).

A signal's CWT can be written as

$$(v, x, y) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} v(t) \psi^* \left(\frac{t-y}{x} \right) \quad (3)$$

Fig. 2 Flow chart of proposed method



In this equation, x is the scale (dilation) constant, while y is the translation (time shift) constant, and the mother wavelet is denoted by ψ^*

A signal's DWT can be written as

$$(v, x, y) = \frac{1}{\sqrt{x_0^m}} \sum_k v(k)^* \left(\frac{n - kx_0^m}{x_0^m} \right) \tag{4}$$

The integer variables x_0^m and kx_0^m are used to replace the x and y terms in Eq. 1. The HPF preserves signal properties in the wavelet function, whereas the LPF represents the coarser information of the approximation signal. In this procedure, the Daubechies level 4 filter is used by DWT analysis ($d=4$) [6].

4 DWT and ANN Analysis

In general, PyWT. dwt applies to all islanding data at a single level. During a single-level discrete wavelet transform, Haar or $d-1$ is used. For multilayer discrete wavelet transforms up to level 4, the Pywt.wavedec function is utilized. The selection of hidden layers has a big impact on the ANN's output accuracy prediction. 100 epochs are used during the ANN model of momentum rate 1msec/step. ReLU, Sigmoid is used as transfer functions of ANN input, hidden, and output layers. ANN models take into account the WT indices in order to detect faults more accurately, based on the differences reflected therein. The data set consists of 25001 samples. In the cross-validation technique, the simulated divided as training data of 20000 (70% of samples) samples and testing data of 5001 samples (30% of samples). During DWT analysis, these samples are finally reduced to 1563 samples of which training data has 1250 samples and testing data has 313 samples at level 4. So with the decrease in the number of samples, islanding detection rate increases. The objective of the learning process is to achieve the smallest mean square error possible. The hidden layer neurons were changed till the mean square error (MSE) was reduced [7].

5 Results and Discussion

After the occurrence of islanding was manufactured at 0.2 s in all of the case studies below by tripping the circuit breaker on the grid side and the frequency of the micro-grid begins to vary from grid frequency after the islanding instant. As a result, starting at $t=0.3525$ s, the difference in frequency parameter begins to diverge from zero, as seen in Fig. 3a. However, because the d-axis voltage is a local measurement, it begins to change at $t=0.2$ s, as illustrated in Fig. 3b. As the fd approaches Th1 (0.33), disturbance current is injected through the inverter current controller for a brief period of 0.3 s. During this period, if the RCPABPSVAC value surpasses threshold value Th2(0.02 pu) islanding will be identified as shown in Fig. 3c. As a result, trip signals are generated for various power mismatch situations as shown in Fig. 3d. The proposed methodology is implemented using MATLAB/Simulink. Analysis of discrete wavelet transform is performed Python programming is used up to level 4, and the ANN methodology is employed for numerous case studies to predict islanding. Figures. 3a–d, 4a–d, 5a–d, 6a–d and 7a–d represent frequency deviation, d-axis voltage, RCPABPSVAC, trip signal generated at DG side circuit breaker. Similarly, Figs. 3f–i, 4f–i, 5f–i, 6f–i and 7f–i represent Daubechies wavelet approximate and detailed coefficient up to level 4 for various fault conditions.

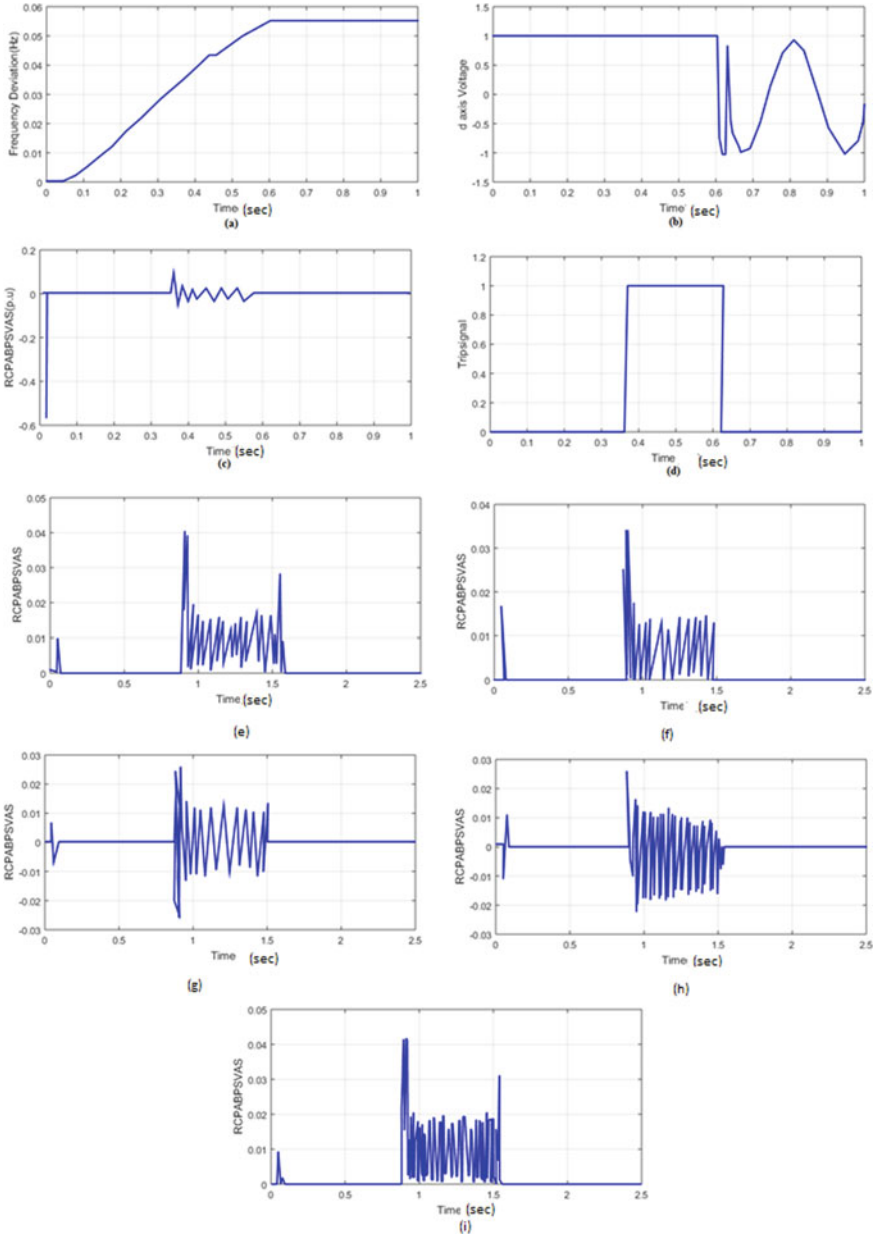


Fig. 3 a-i System performance under normal loading conditions

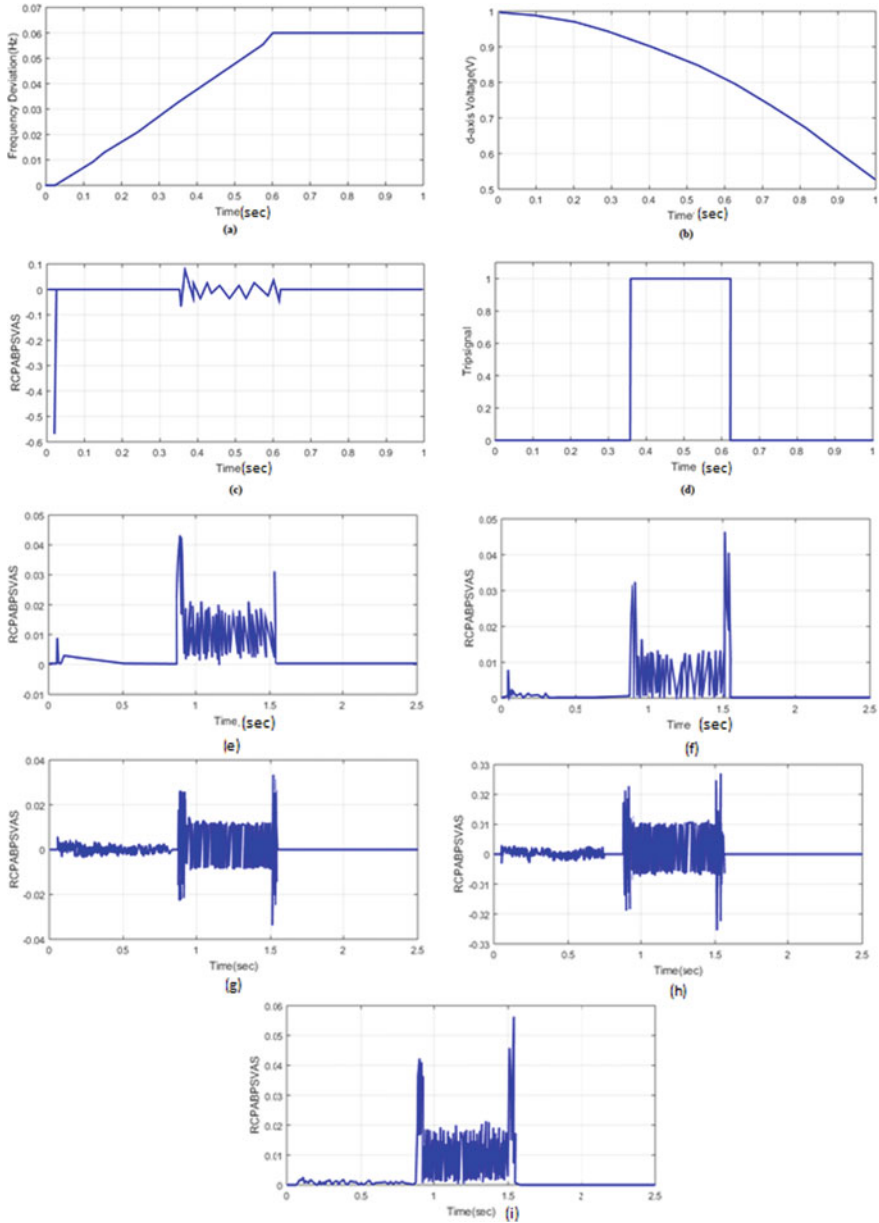


Fig. 4 a–i System performance under balanced loading conditions

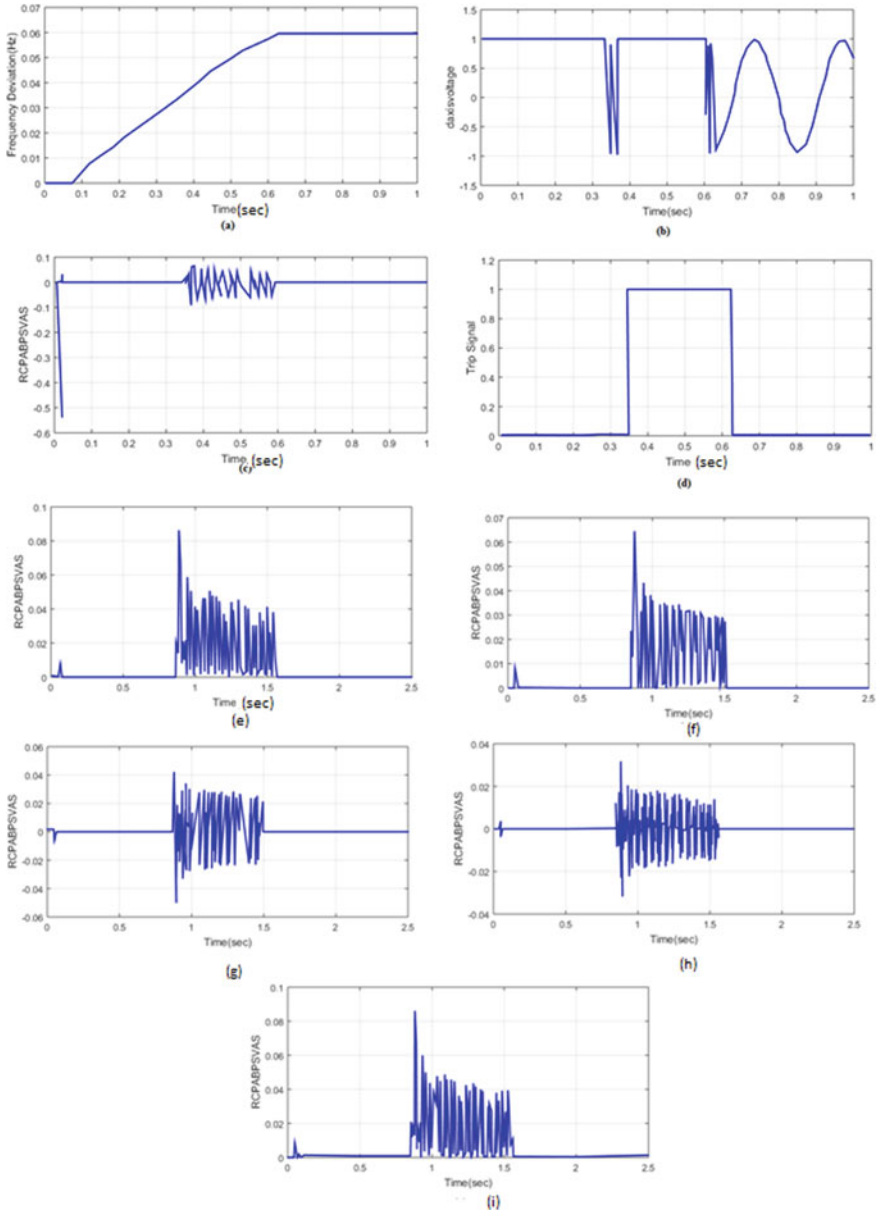


Fig. 5 a-i System performance under over loading conditions

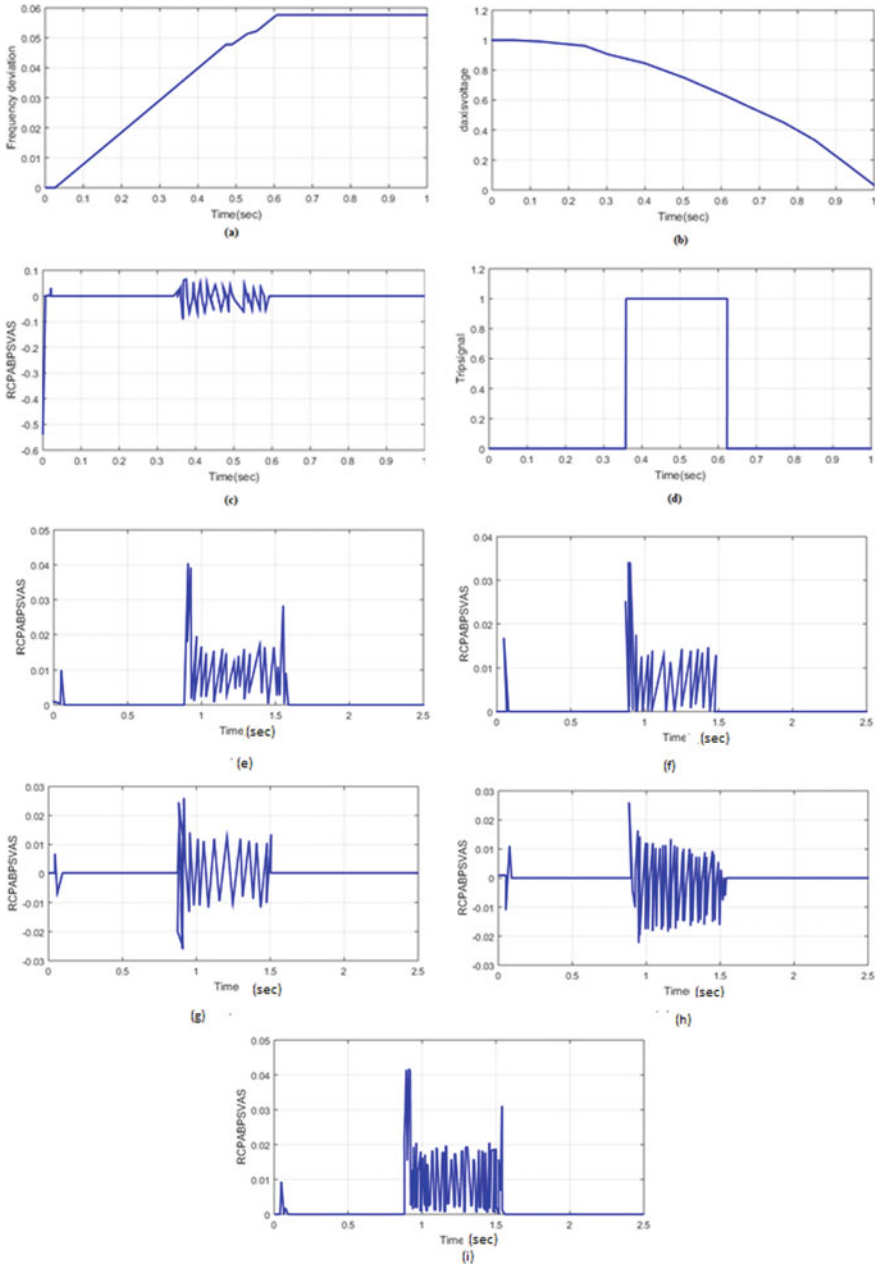


Fig. 6 a–i System performance under inductive load switching

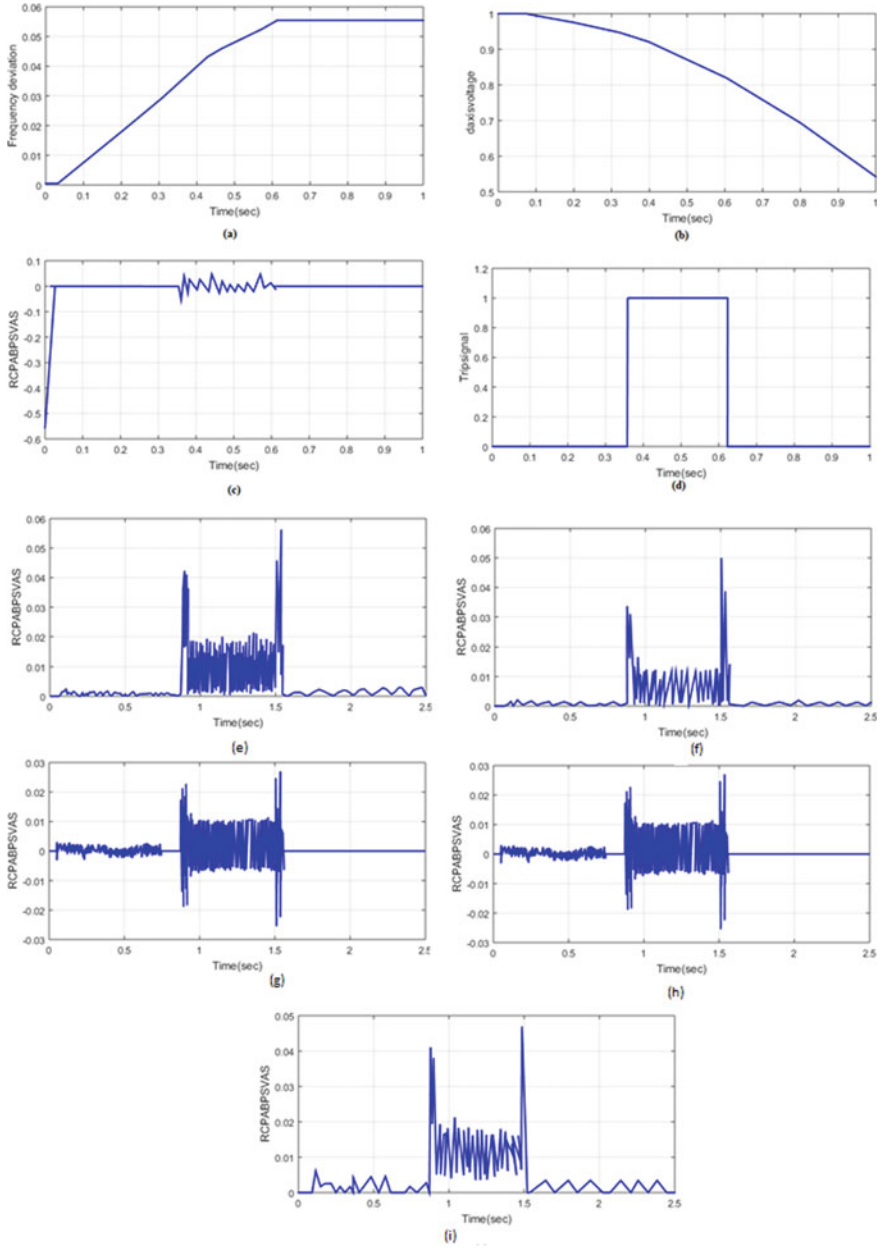


Fig. 7 a-i System performance under capacitor switching

5.1 Conditions of Normal Loading

See Fig. 3.

5.2 Conditions of Balanced Loading

See Fig. 4.

5.3 Conditions of Over Loading

See Fig. 5.

5.4 Switching Conditions of Inductive Load(50HP) (Non-islanding)

See Fig. 6.

5.5 Switching Condition of Capacitor (10 KVA) (Non-islanding)

See Fig. 7.

6 Comparative Assessment

In this paper, DWT-ANN-based frequency disturbance triggered passive islanding detection scheme gradually decreased at level 4 below 18 m s due to a reduction in the number of samples. MSE also determined for various case studies, lower the value of MSE better will be the accuracy. As shown in Table 3, the results are based on the minimum square error for each case study. The injected current in the current controller has not produced sinusoidal d-axis voltage during both balanced and non-islanding conditions which will affect islanding detection time. So, this method detects islanding even under balanced load conditions also. Table 2 indicates the proposed method has better detection time compared to existing procedures. The ranges of F1-score, recall, and precision values are 0.9703, 1.9424 at 70 KW

Table 2 Suggested approach against existing procedures

Approach	Islanding estimation time (m s)
Rate of change of frequency	500
Positive sequence voltage and current	100
Active rate of change of frequency	200
Regulator voltage over reactive power	300
Proposed technique	Less than 20

Table 3 Suggested approach at various conditions

S. No.	Case studies	Normal detection time (m s)	MSE (at level4)	Accuracy testing data (%)	WT-ANN detection time(msec) (at level4)
1	70 KW Load	60.4	0.0415	99.85	14.9
2	100 KW Load	116.7	0.05750	94.24	16.98
3	120 KW Load	159	0.04472	97.52	16.54
4	Inductive load switching	105.6	0.057509	94.24	17.86
5	Capacitor switching	117.6	0.0575	94.24	17.89

load. F1-score ranges at 100KW load include recall; precision is stated as 0.9703, 1, 0.9744. The ranges of F1-score, recall, and precision value for 120 KW load are 0.9834, 0.9908, and 0.9569, respectively. During non-islanding, F1-score, recall, and precision are 0.89422, 1.0, 0.9744.

7 Conclusion

The research provides a negative sequence islanding detection technique based on WT-ANN that utilizes mean square error to detect islanding conditions for various wavelet coefficients. The energy content as well as the standard deviation are measured and fed into the ANN model. It measures islanding detection time with an accuracy of 88–92% for varied loads. In balanced conditions with 0% NDZ, the suggested islanding technique can also detect islanding and overload conditions in a short amount of time. Advanced ANN approaches such as multiple regressions, recurrent neural networks, and others can be used to implement the recommended approach for detecting islands without sacrificing accuracy in the future.

References

1. Manohar M, Sheetal C, Pravat Kumar R (2020) Taxonomy islanding detection techniques for distributed generation in microgrid. *Renew Energy Focus* 31:9–30
2. Ahmed G, Ahmed Bilal A, Abdel-Rahman A (2018) Comparative study of passive and active islanding detection methods for PV grid-connected systems. *Sustainability* 10(6):1–15
3. Krishna MG, Geetha Lakshmi B (2021) Balanced islanding detection of integrated DG with phase angle between voltage and current. *Indonesian J Electric Eng Comput Sci* 23(1):32–40
4. Phanindra Ganivada K, Premalata J (2020) Frequency disturbance triggered d-axis current injection scheme for islanding detection. In: *IEEE transactions on smart grid*, pp 1949–3053
5. Basanta P, Rajendra S, Ajay Kumar J (2019) Islanding detection in distributed generation integrated Thimi—Sallaghari distribution feeder using wavelet transform and artificial neural network. *J Instit Eng* 15(2):55–61
6. Paiva S, Ribeiro R, Alves D, Costa, Rocha T (2020) A wavelet-based hybrid islanding detection system applied for distributed generators interconnected to AC microgrids. *Int J Electric Power Energy Syst* 121:106–132
7. Krishna MG, Geetha Lakshmi B (2021) Balanced islanding detection of distributed generator using discrete wavelet transform and artificial neural network. *Int J Eng Trends Tech* 69(10):57–71

Design of Speed and Area Efficient Non Linear Carry Select Adder (NLCSLA) Architecture Using XOR Less Adder Module



Yamini Devi Ykuntam, Bujjibabu Penumutchi, and Srilakshmi Gubbala

Abstract In any arithmetic processor, adder acts as a primary element for performing various computational operations. In order to have good speed of operation for the processor, the adder used in it should be efficient in speed of operation. Among different adder architectures, Carry Select Adder (CSLA) is having good speed of operation which can be used in processors and various digital systems. The reason behind the speed of CSLA is its architecture that consists of twin cascaded full adders which is nothing but Ripple Carry Adder (RCA). These twin RCAs execute addition operation in parallel by assuming input carry as '0' for one RCA whereas the other RCA input carry is assumed as '1'. The drawback of CSLA is more area due to usage of twin RCAs. A low area CSLA is proposed which has BEC instead of RCA with input carry as '1' with the cost of more delay than the existing architecture. To achieve a CSLA architecture having low area without affecting the speed parameter, a novel architecture of CSLA is proposed in this paper using XOR less adder module also called as mirror adder. The proposed CSLA architecture is compared with the existing CSLA architectures in area and speed parameters.

Keywords NLCSLA · RCA · BEC · Mirror adder (MA)

Y. D. Ykuntam (✉) · B. Penumutchi
Department of ECE, Aditya Engineering College, Surampalem, Andhra Pradesh, India
e-mail: yaminieie@gmail.com

B. Penumutchi
e-mail: bujjibabu_penumuchi@aec.edu.in

S. Gubbala
Department of ECE, Vishnu Institute of Technology, Bhimavaram, Andhra Pradesh, India
e-mail: srilakshmi.g@vishnu.edu.in

Y. D. Ykuntam · B. Penumutchi · S. Gubbala
Jawaharlal Nehru Technological University Kakinada, East Godavari District, Kakinada, India

1 Introduction

As the electronics field is advancing quickly, requirement of digital systems with high speed, small area and low power consumption is increasing. The main design entities which are to be considered while designing a digital circuit are: area, which must be less; speed, which should be more; and power consumption, which should be less [1]. The major area of research in VLSI domain urges for circuits having more speed and reduction in area. In any system or processor, computational unit plays an important role. Any type of computations performed by the computational unit involves addition operation which should be fast enough to have an efficient performance computational unit.

An adder module performs the addition operation whose performance should be good in order to have good operation speed. In order to satisfy the VLSI circuit design criteria of less area and power consumption with high speed operation, the computational unit must have an adder architecture satisfying all the mentioned criteria.

In literature survey, number of adder architectures are studied like Ripple Carry Adder (RCA) in which number of 1-bit full adders are connected in series, Carry Look ahead Adder (CLA) in which addition is performed with the help of generate and propagate functions and many other architectures (Carry Save Adder (CSA), Carry Increment Adder (CIA) and Carry Select Adder (CSLA) [2–6].

Majorly among all adder architectures, the main problem is speed of addition. The speed of addition in most adder architectures is limited by carry generation and its propagation. The addition of each bit in the given data depends on the carry input obtained from the previous data bit addition operation. So the next level data bit addition is delayed until it gets the carry input from the previous stage data bits addition operation.

Among all the adder architectures studied, Non linear CSLA (NLCSLA) is having less delay in addition operation but its area is more as it consists of two sets of RCA [7]. These RCAs in NLCSLA perform operation by assuming one as carry input equal to '1' and the other assumes carry input as '0' [8–12]. A 2-to-1 multiplexer is used to choose the final sum output from the two RCA outputs. The select input to multiplexer is nothing but the carry output from previous stage of addition operation. In order to minimize the area of NLCSLA, the circuit is designed using Binary to Excess one Converter (BEC) [13].

But this NLCSLA with BEC architecture delay increased when compared to existing structure even though its area is less. To attain a NLCSLA with minimum delay and reduced area, a new architecture is proposed in this paper. The proposed architecture is modification of NLCSLA with BEC architecture which consists of XOR less adder [14–16] module as a replacement for RCA module which assumes carry input as '0' for performing addition. The proposed architecture and existing architectures of NLCSLA are contrasted in area and delay.

In brief, the paper starts with the explanation about existing and proposed adder architectures. The next section evaluates the adder architectures in terms of area and

delay calculations. Later on, the adder architectures are contrasted with the help of synthesized results. As a final point, the paper is concluded with explaining the future scope of the work.

2 Non Linear Carry Select Adder

There are two types of CSLA architectures—Linear CSLA (LCSLA) and Non Linear CSLA (NLCSLA). Among both, NLCSLA is having good operation speed. The NLCSLA is also called as Square Root CSLA (SQRT CSLA) which is considered in this paper for architecture modification in order to obtain an efficient CSLA architecture design.

2.1 Existing NLCSLA

With the design of CSLA architecture, the carry data transmission difficulty is conquered to most extent. Its architecture is planned in such a way that the addition operation is performed parallel by considering carry data inputs both as ‘1’ and ‘0’ and sum output is chosen using the precious stage carry data output. A NLCSLA of 16-bit size is shown in Fig. 1 in which different stages with different input sizes are present. Each stage of adder architecture consists of twin RCAs and multiplexer (MUX).

In twin RCAs, one performs addition assuming carry input as ‘0’ and the other performs addition assuming carry input as ‘1’. The sum and carry outputs from the twin RCAs are applied as inputs to multiplexer which selects the final sum and carry output of the particular stage. Depending on the control signal, any one RCA’s, sum and carry output is selected by the MUX. This architecture is having good speed of operation when compared to other adder structures but its area is more due to the presence of twin RCAs. At the same time, power consumption will be more due to more area.

2.2 NLCSLA Using BEC

To prevail over the drawbacks of NLCSLA, a modified architecture for it is proposed using Binary to Excess one Converter (BEC). In modified architecture of NLCSLA, BEC is used in place of RCA with carry data input as ‘1’. The area occupied by BEC is less when compared to RCA. So with the usage of BEC in the modified adder architecture will result in adder architecture with reduced area and power consumption. A BEC of $n + 1$ bit size can be used instead of an RCA of n -bit size where n is input size for the respective RCA or BEC.

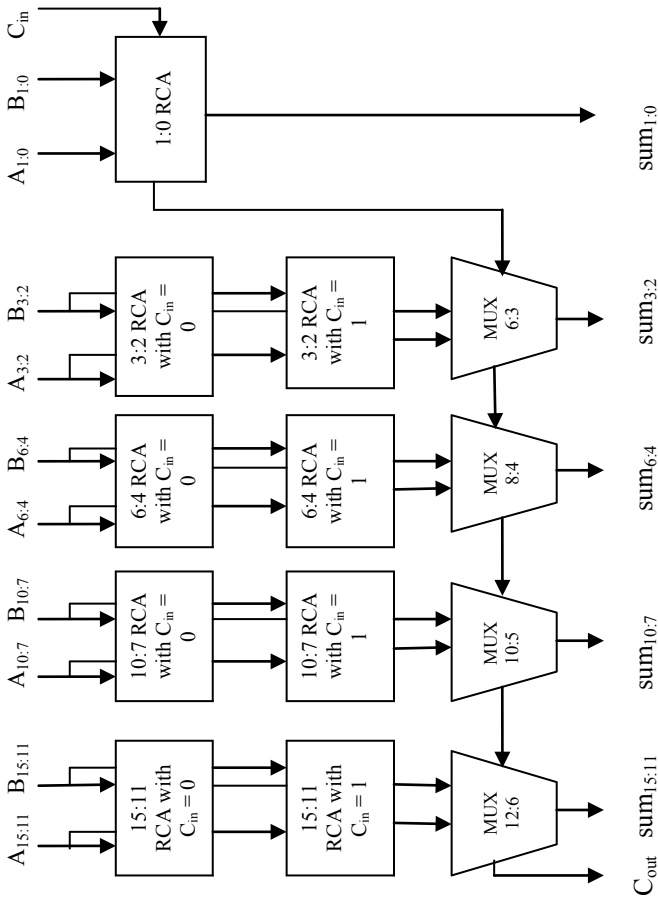


Fig. 1 NLCSLA of 16-bit size

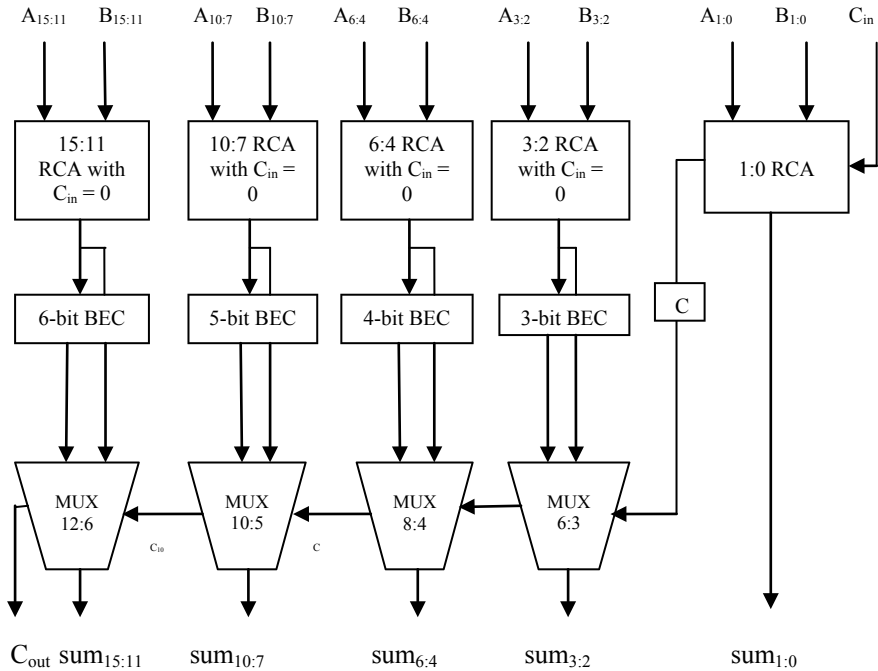


Fig. 2 NLCSLA using BEC of 16-bit size

The operation of the adder remains the same, i.e., the MUX will select either output of RCA with carry data input as ‘0’ or the BEC output which depends on the carry data output resulting from the previous stage addition operation. If the carry data output is ‘0’, the MUX selects RCA output and if it is ‘1’, then MUX selects BEC output. A NLCSLA using BEC of 16-bit size is shown in Fig. 2. Although the area of the adder architecture is reduced, the delay in adder operation increased.

2.3 NLCSLA Using XOR Less Adder Module (MA)

The NLCSLA using BEC architecture has less area but at a cost of increased delay. To have architecture with minimum area and power consumption, a novel NLCSLA is proposed in this paper. An XOR less adder module is used in this novel adder architecture as an alternative for RCA with carry data input as ‘0’. In an adder module, an XOR gate has more number of transistors which increase area also the delay as the number of transistors in critical path increases. An XOR less adder module is also called as Mirror Adder (MA) which is shown in Fig. 3.

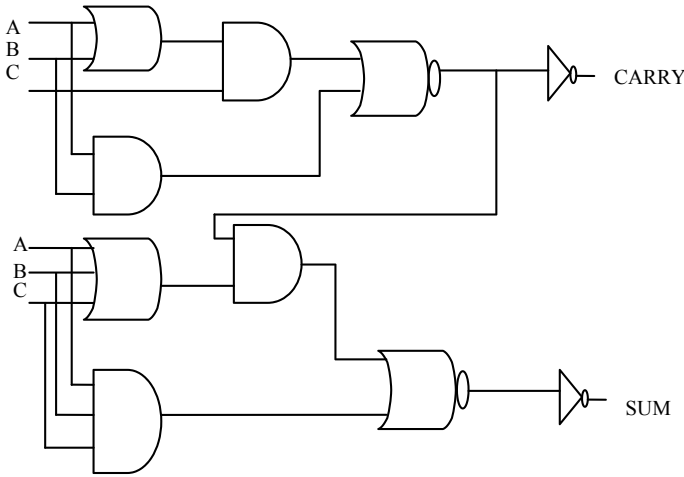


Fig. 3 XOR less adder circuit

Further minimization of carry delay in MA circuit can be done by removing inverter at the carry output. The removing of inverter at carry output is called elimination of inverter in the path of carry output which results in great reduction of carry delay. Two properties were used in elimination of inverter scheme. Full adders were applied with inverted inputs to get an inverted output which is the first property. The second property is that MA generates the inverted carry output and then inverts it to obtain final carry output.

Figure 4 shows a NLCSLA using MA of 16-bit size in which RCA with carry data input as '0' is replaced by MA. The modified adder architecture with MA also have BEC and MUX. The operation of the architecture remains the same as explained in previous Sects. 2.1 and 2.2.

3 Assessment of Delay and Area(Theoretical)

3.1 Assessment of Delay and Area for Modules in NLCSLA

For assessing the area only basic gates are considered, i.e., AND, OR and Inverter. The NLCSLA architecture consists of RCA and MUX. Full adders and a half adder connected in series forms an RCA. With the help of two XOR gates, one OR gate and two AND gates, a 1-bit full adder structure can be designed. Two AND gates, one OR gate and two inverters are required for designing a 1-bit XOR gate, i.e., XOR is designed using 5 basic gates. In total, 13 gates are required to design a 1-bit full adder. One XOR gate and one AND gate are required for a half adder design. The gate count required for design of half adder is 6. A 2:1 MUX design call for one OR

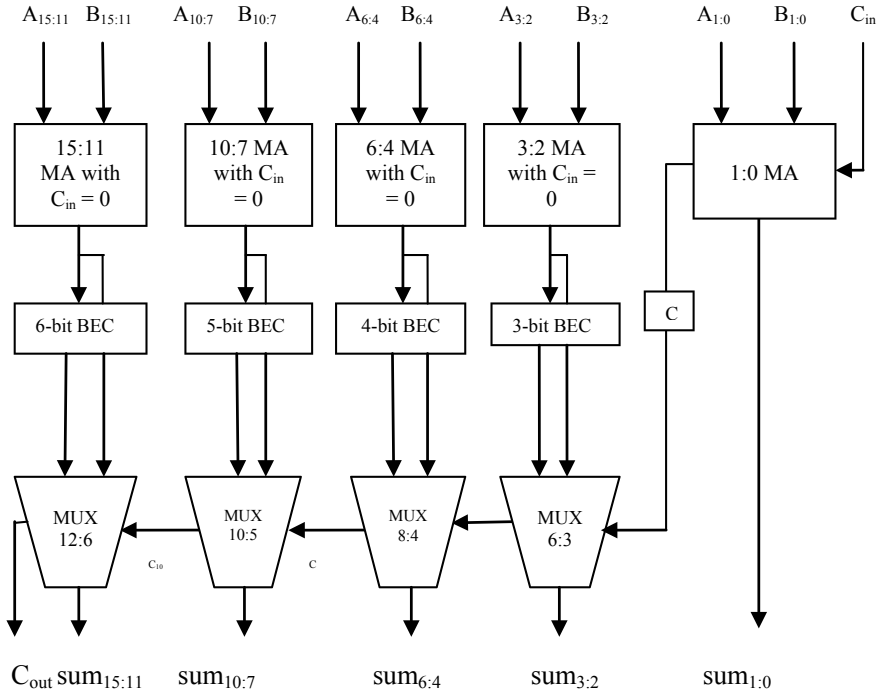


Fig. 4 NLCSLA using MA of 16-bit size

gate, one inverter and two AND gates which gives the total gate count as 4. Each gate constitutes 1 unit of delay and area.

The critical path which is the longest path in the circuit offers the highest delay which is considered as the delay of entire circuit. The number of gates in critical path helps in assessing the delay of the path, i.e., by summing the gates delay in the path. The total number of gates in a module helps in area assessment (Table 1).

Table 1 Assessment of area and delay for modules in NLCSLA

Module name	Area assessment	Delay assessment
XOR	5	3
2:1 mux	4	3
Half adder	6	3
Full adder	13	6

3.2 Assessment of Delay and Area for Existing NLCSLA

The NLCSLA can be divided into five blocks where block 1 consists of only RCA and remaining blocks consist of twin RCAs, MUX which can be seen in Fig. 1. Let us assess the area for block 2 which is designed using twin RCAs of 2-bit size and three 2:1 MUX. In twin RCAs, one RCA assumes carry input as '0' which can be designed using one half adder and one full adder and the other RCA assumes carry input as '1' which can be designed using two full adders.

Using the methodology for assessment of area discussed in 3.1, gate count for group 2 is 57 which is the total number of gates in 3 full adders ($3 \times 13 = 39$, where 13 is gate count of one full adder), one half adder ($1 \times 6 = 6$, where 6 is gate count of one half adder) and three 2:1 MUXs ($3 \times 4 = 12$, where 4 is gate count of one 2:1 MUX). In the same way, the area is assessed for remaining blocks. The delay for the architecture is assessed by summing all gate delays in the critical path.

3.3 Assessment of Delay and Area for NLCSLA Using BEC

The area and delay assessment for this adder architecture is also performed in the same way as explained in Sect. 3.2. From Fig. 2, it is clear that the NLCSLA using BEC also has five blocks. RCA assumes carry input as '0' of 2-bit size, BEC of 3-bit size and three 2:1 MUX are the modules of block 2 in the adder architecture.

Gate count of block 2 is 43 which is the total number of gates in one full adder ($1 \times 13 = 13$, where 13 is gate count of one full adder), one half adder ($1 \times 6 = 6$, where 6 is gate count of one half adder), three 2:1 MUXs ($3 \times 4 = 12$, where 4 is gate count of one 2:1 MUX), one AND gate, one NOT gate and two XOR gates ($2 \times 5 = 10$, where 5 is gate count of one XOR gate). The area assessment for remaining blocks in the architecture can be performed in same way where as the delay for the architecture is assessed by summing all gate delays in the critical path.

3.4 Assessment of Delay and Area for NLCSLA Using MA

The area and delay assessment for this adder architecture is also performed in the same way as explained in Sect. 3.2. From Fig. 4, it is clear that the NLCSLA using MA also has five blocks. RCA designed using MA assuming carry input as '0' of 2-bit size, BEC of 3-bit size and three 2:1 MUX are the modules of block 2 in the adder architecture.

Gate count of block 2 is 41 which is the total number of gates in one mirror adder ($1 \times 11 = 11$, where 11 is gate count of one mirror adder), one half adder ($1 \times 6 = 6$, where 6 is gate count of one half adder), three 2:1 MUXs ($3 \times 4 = 12$, where 4 is gate count of one 2:1 MUX), one AND gate, one NOT gate and two XOR gates (2×5

= 10, where 5 is gate count of one XOR gate). The area assessment for remaining blocks in the architecture can be performed in the same way whereas the delay for the architecture is assessed by summing all gate delays in the critical path.

4 Assessment of Delay and Area (Practical)

All the three adder architectures (two existing and one proposed), are designed with the help of VHDL for different input size. Then the designed architectures are functionally verified and synthesized by using Xilinx ISE Navigator of version 14.2. A simulator called ISIM is integrated in Xilinx ISE which helps to perform functional simulation of designed adder architectures for functional verification. Later on the adder architectures are synthesized to obtain the area which is specified in number of LUTs and delay which is specified in Nano seconds.

Table 2 interprets the area and delay for all the three designed adder architectures. From Table 2, it can be said that the delay for all the three adders are the same for lower input size, i.e., for 8-bit input size. But as the input size increases, the delay of the proposed adder, i.e., NLCSLA with MA decreases when compared to the existing adders, i.e., NLCSLA and NLCSLA using BEC. At the same time, the area of proposed adder is increased but by a small amount.

5 Conclusion

The NLCSLA, NLCSLA using BEC and NLCSLA using MA are designed for different input sizes (8, 16, 32 and 64-bit) with the help of VHDL in this paper. Also the designed adder architectures are functionally verified by performing functional simulation. Further, the designs are synthesized for area and delay reports which are tabulated in Table 2. The proposed adder, NLCSLA using MA had less delay with a little bit penalty of more area when compared to other two existing adder architectures. But still, the proposed adder area is less than the existing NLCSLA. As the delay of proposed adder is less also having less area than the existing NLCSLA, it can be used in the design of high speed VLSI and other circuits. As a future scope, the proposed adders can be designed for higher-order input size and can be used to build different multipliers and filters.

Table 2 Area and delay assessment (practical) among three designed adder architectures

Size of input	Type of adder	Area in No. of LUTs	Delay in ns
8-bit	Existing NLCSLA	21	12.38
	NLCSLA using BEC	19	12.38
	NLCSLA using MA	19	12.38
16-bit	Existing NLCSLA	46	18.46
	NLCSLA using BEC	40	19.65
	NLCSLA using MA	46	17.45
32-bit	Existing NLCSLA	105	25.22
	NLCSLA using BEC	93	26.48
	NLCSLA using MA	99	23.88
64-bit	Existing NLCSLA	229	31.66
	NLCSLA using BEC	203	32.50
	NLCSLA using MA	211	30.05

References

1. Rabaey JM (2001) Digital integrated circuits—a design perspective. Prentice-Hall, Upper Saddle River, NJ
2. SaiKumar M, Samundiswary DP (2013) Design and performance analysis of various adders using verilog. *Int J Comput Sci Mob Comput* 2(9):128–138
3. Jayanthi AN, Ravichandran CS (2013) Comparison of performance of high speed VLSI adders. In: 2013 international conference on current trends in engineering and technology (ICCTET). IEEE, pp 99–104
4. Ramkumar B, Kittur HM, Kannan PM (2010) ASIC implementation of modified faster carry save adder. *Eur J Sci Res* 42(1):53–58
5. Kavya KVSSSS, Penumutchi B, Nandan D (2021) Analysis on high-performance full adders. In: Next Generation Information Processing System. Springer, Singapore, pp 122–131
6. Siva Charan Kumar V, Bujjibabu P (2017) Design and implementation of radix-8 based 32-bit pipelined multiplier by using CLA. In: 2017 2nd international conference on communication and electronics systems (ICCES). IEEE, pp 866–871
7. Bedrij OJ (1962) Carry-select adder. *IRE Trans Electron Comput* 340–344
8. Kim Y, Kim L-S (2001) 64-bit carry-select adder with reduced area. *Electron Lett* 37(10):614–615

9. Ceiang TY, Hsiao MJ (1998) Carry-select adder using single ripple carry adder. *Electron Lett* 34(22):2101–2103
10. He Y, Chang CH, Gu J (2005) An area efficient 64-bit square root carry-select adder for low power applications. *Proc IEEE Int Symp Circ Syst* 4:4082–4085
11. Weste N, Eshragian K, *Principles of CMOS VLSI designs: a system perspective*, 2nd edn. Addison-Wesley, pp 1985–1993
12. Morinaka H, Makino H, Nakase Y, et al (1995) A 64 bit carry look-ahead CMOS adder using modified carry select. *Cz/stoin Integr Circ Conf* 585–588
13. Ramkumar B, Kittur HM (2012) Low-power and area-efficient carry select adder. *IEEE Trans Very Large Scale Integr (VLSI) Syst* 20(2)
14. Ykuntam YD, Nageswara Rao MV, Locharla GR (2013) Design of 32-bit carry select adder with reduced area. *Int J Comput Appl* 75(2):47–51
15. Ykuntam YD, Prasad SH (2021) A modified high speed and less area BCD adder architecture using Mirror adder. In: 2021 2nd international conference on smart electronics and communication (ICOSEC). IEEE, pp 624–627
16. Ykuntam YD, Rajan Babu M (2019) A novel architecture of high-speed and area-efficient wallace tree multiplier using square root carry select adder with mirror adder. In: *Innovations in electronics and communication engineering*. Springer, Singapore, pp 319–326

Design and Analysis of 1×4 Corporate Feed Conformal Microstrip Antenna Array for X-band Spaceborne Synthetic Aperture Radar Applications



Bala Ankaiah Nunna and Venkata Kishore Kothapudi

Abstract Synthetic aperture radar (SAR) collects data about distant objects using electromagnetic waves. It has measuring capabilities without requiring physical contact, making it ideal for remote sensing applications. In the last decade, space-based synthetic aperture radars for earth remote sensing have been created that can operate in all weather and day-and-night circumstances. The antenna works as the interface between the outer surface and the internals of the radar, is a vital component of synthetic aperture radar. The new technique is presented for simulating conformal microstrip patch antenna arrays. A single, 1×2 and 1×4 corporate feed microstrip antenna arrays using inset feeding technique operating at 9.65 GHz is designed. The characteristics of the antenna array are simulated and also examined. It includes reflection coefficient, VSWR, gain and side lobe level. The improved cross-range resolution of spaceborne Synthetic Aperture Radar is aided by the increased bandwidth and gain. The gain achieved improves cross-range coverage while lowering aerodynamic drag. The proposed designs are simulated using electromagnetic software CST Microwave Studio. The performance characteristics of the single conformal microstrip patch antenna, 1×2 and 1×4 conformal arrays are compared with planar arrays.

Keywords Conformal antenna · Spaceborne · Synthetic aperture radar · Microstrip patch antenna · Corporate feed · X-band

B. A. Nunna (✉) · V. K. Kothapudi

Center of Excellence Advanced RF Microwave and Wireless Communications, Department of Electronics and Communication Engineering, Vignan's Foundation for Science, Technology, and Research (VFSTR), Vadlamudi, Guntur District, Andhra Pradesh 522213, India
e-mail: balaankaiah@ieee.org

V. K. Kothapudi

e-mail: v.k.kothapudi@ieee.org

1 Introduction

Synthetic Aperture Radar (SAR), initially created in the 1950s to improve the resolution of military reconnaissance radar, has quickly matured as a remote sensing technology for a wide range of civilian uses. Although the majority of currently operational spaceborne SAR operates in the L, C and X-bands, there is an increasing need for observations at higher frequencies. The design of aircraft antennas is difficult because it must meet a number of criteria in order to give the optimal aerodynamic force [1]. Synthetic aperture radar in space is a versatile sensor that may be used for earth monitoring in any weather. The majority of space-based SAR systems is based on large-satellite platforms and utilizes phase-arrays or mechanical steering, resulting in high costs, high power consumption and restricted performance [2]. A conformal antenna is one that is meant to follow a certain shape, such as a flat curving antenna installed on or incorporated in a curved surface. The impact of mutual coupling is explored using a four-element conformal antenna array on a cylindrical surface [3, 4]. Aerodynamic drag is increased by planar antennas, which reduces fuel efficiency. Conformal antennas address this issue by adapting to the contour of the item to which they are mounted. In X-band, the antenna size is small [5]. The cross-range resolution can be improved by using a wider transmission bandwidth and a reception band [6]. Microstrip antennas are preferred and deemed extremely favorable for meeting the aforementioned standards [7, 8]. Microstrip antennas have a low profile, can adapt to planar and conformal surfaces, easy and inexpensive to fabricate using contemporary PCB technology. These antennas can be mount on aircrafts, satellites, missiles and also on cell phones [9].

In [10], conformal antenna, configured with enhanced bandwidth for aircraft applications, is presented. The design is achieved 10.2% of bandwidth and gain of 9 dB at 5.2 GHz for improving the synthetic aperture radar images. This paper presents the omnidirectional antenna array on cylinder. A 1×4 antenna array is conformed on cylinder surface with a diameter of 90 mm. This design obtains the VSWR of 1.16 at the operating frequency of 2.34 GHz [11]. In [12–14], 3×3 series-fed planar array for X-band, 1×3 series-fed linear array for X-band and 10 element Linear array for K-band is presented respectively. In this paper [15], design of 4 element and 8 element conformal antennas on cylinder surface with different curvatures has been presented. The proposed curvatures are 30° and 40° at 10 GHz frequency. In [16], conformal antennas that are placed on cylinder surfaces has been studied. The antennas' elements of 4, 8, 16, and 32 are considered with different inter element spacing and different radius of curvatures. In this paper [4], performance analysis of microstrip conformal antenna array is presented with different curvature at 2.4 GHz with the gain of 6 dB. In [17], conformal antenna is designed for Ku-band. A gain of 13.65 dB with 16.7° 3-dB beam width is achieved. In this paper [18], a conformal antenna is designed on cylindrical surface. The peak gain is 23.5 dB, and 4° beam deflection is achieved. In [19], a rectangular microstrip antenna is conformed on spherical surface. The achieved spherical microstrip antenna gain is 7.2 dBi and side lobe level is -11.6 dB. In this paper [20], design of high gain microstrip antenna

is designed for X-band RADAR applications. In [21], Planar and conformal 2×2 microstrip patch antenna is designed for C-band avionics applications.

2 Conformal Antenna Array Design

This section describes the design process for conformal antenna arrays, with distinct configurations. Microstrip antennas have become more attractive for space applications. They are being employed for both military and commercial purposes. In this paper, a single conformal antenna, 1×2 corporate feed, 1×4 corporate feed inset fed conformal arrays has been designed, analyzed and compared with planar arrays for spaceborne SAR applications. These arrays are designed at X-band (9.65 GHz). The simulations were carried out in CST MWS software. The antenna spacing is taken as 0.7λ in arrays. In this work, the loss tangent $\tan \delta = 0.0009$ is considered, material substrate with relative permittivity $\epsilon_r = 2.2$ is chosen for all designs, with RT/Duroid-5880 substrate [22].

2.1 A Single Conformal Antenna

A single conformal antenna is designed at the frequency of 9.65 GHz. Weather monitoring, air traffic management, defense tracking and vehicle speed detection are all done with X-band radar and phased arrays at civic, military and government institutions. The inset feeding approach for microstrip antennas is straightforward to use and understand, as the inset gap and inset length determine the antenna's behavior. The single microstrip patch antenna is designed as planar antenna and then bent on a cylinder surface with the radius of 10 mm to make conformal antenna. The top view of the single planar antenna conformed on cylinder is presented in Fig. 1a, b. The optimized dimension of the single conformal antenna is given in Table 1.

The return loss $|S_{11}|$ for the single conformal antenna is presented in Fig. 1c. The return loss is -31.24 dB. The surface current distribution is shown in Fig. 1d. Figures 1e, f demonstrate the simulated radiation pattern of E-plane and H-plane. Figure 1g presents simulated 3D far-field gain. The simulated gain is 7.38 dBi and SLL is -13.2 dB for the single conformal antenna design.

1×2 Corporate Feed Conformal Antenna Array. There is no requirement for matching the impedance in a series-fed system, and the signal travels a shorter distance than in a corporate-fed system. These networks are more difficult to design because the amplitude and phase of the impedance must be precisely estimated. In a corporate-fed network, all elements have the same electrical distance from the source to the radiating element, and hence the same signal phase and amplitude. The most common feeding approach for fabricating antenna arrays is a corporate feed. The

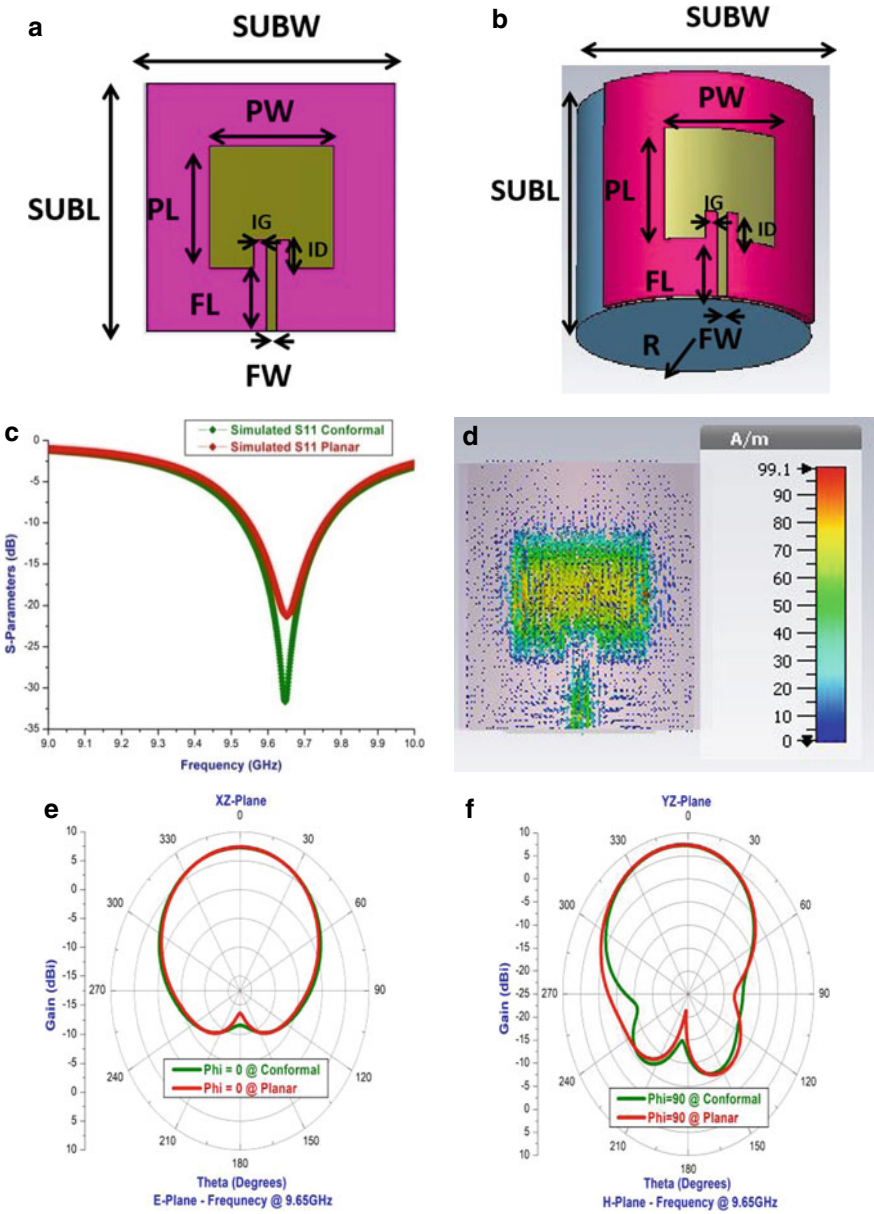


Fig. 1 **a** Single Planar antenna. **b** Single conformal antenna. **c** Simulated return loss. **d** Simulated surface current distribution. **e** Radiation pattern (E-Plane). **f** Radiation pattern (H-Plane). **g** Simulated far field gain

Fig. 1 (continued)

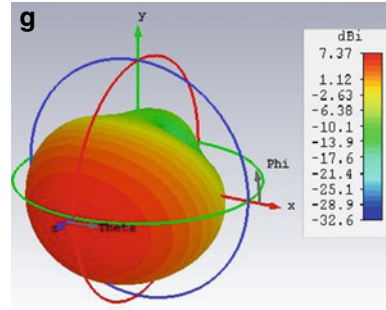


Table 1 A single conformal antenna dimensions

Parameter	Value (mm)
PL	10
PW	10
FL	5.5
FW	1.0
IG	1.0
ID	2.4
SUBL	20
SUBW	20
R	10

incident power is separated and distributed evenly to the different antenna components in this situation. This approach allows for greater control over each element’s feed, making it excellent for scanning phased arrays and multi beam arrays. The corporate feed network is used to provide power splits of $2n$ (i.e., $n = 2, 4$, etc.). To avoid mutual coupling effects, a gap of 0.7 is employed between two subsequent elements of the array antenna. The distance between two patches is 22 mm. The patch components are coupled using a microstrip line and a T-junction power divider in this configuration. The width of 50Ω line feed is 2.8 mm and 100Ω line feed is 0.716 mm.

The top view of 1×2 corporate feed planar and conformal antenna array is presented in Fig. 2a, b. The cylinder radius is considered as 20 mm for the 1×2 corporate feed conformal antenna array. The 1×2 and 1×4 corporate feed conformal antenna arrays’ optimized dimensions are given in Table 2.

The simulated return loss $|S_{11}|$ for the 1×2 corporate feed conformal array is presented in Fig. 2c. The obtained return loss is -21.21 dB. The surface current distribution is shown in Fig. 2d. Figure 2e, f present the simulated radiation pattern of E-plane and H-plane. Figure 2g shows simulated 3D far-field gain. The simulated gain is 8.38 dBi and SLL is -16.6 dB for the 1×2 corporate feed conformal antenna array.

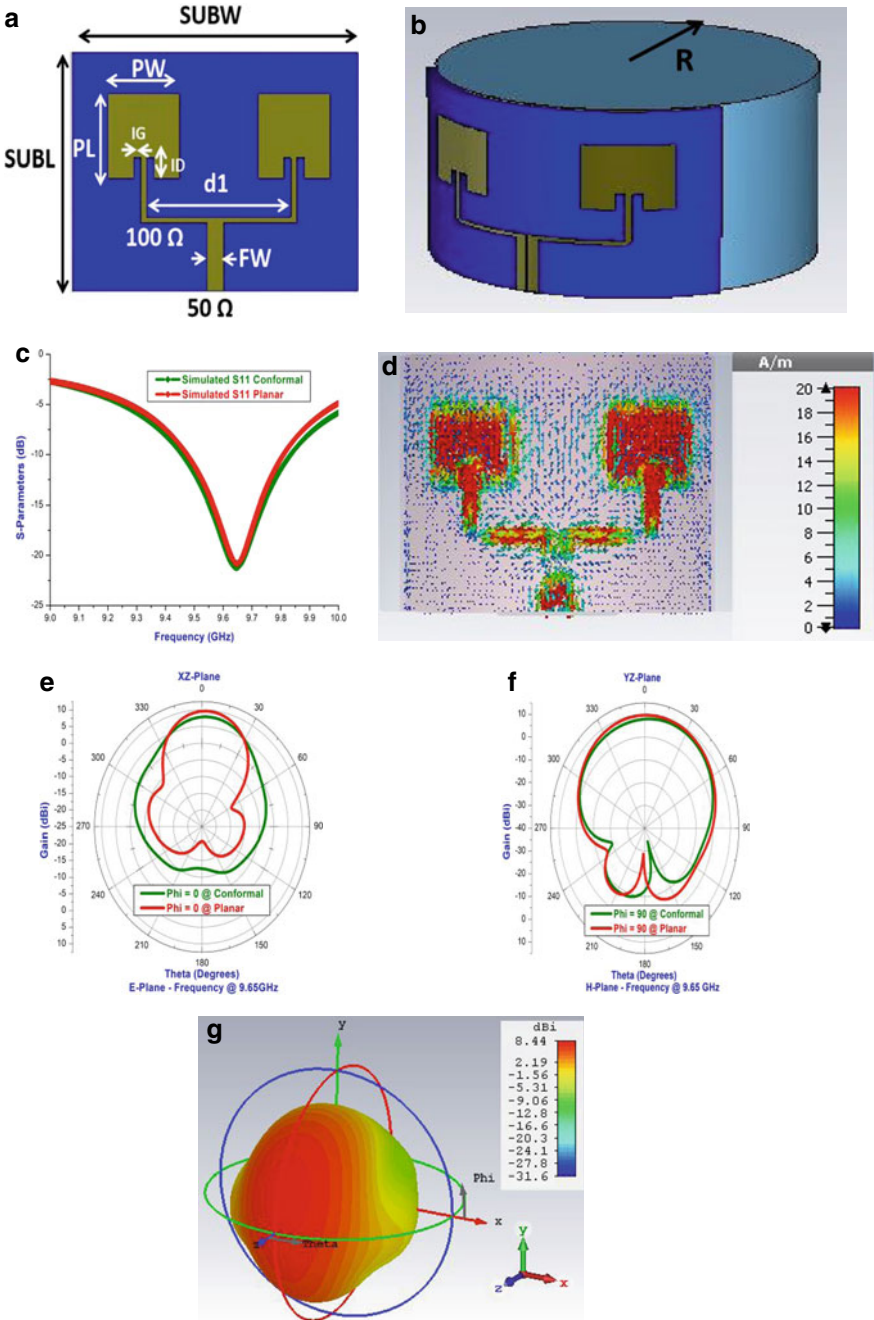


Fig. 2 **a** 1×2 planar array antenna. **b**. 1×2 conformal array antenna. **c** Simulated Return loss. **d** Simulated Surface current distribution. **e** Radiation Pattern (E-Plane) **f** Radiation Pattern (H-Plane). **g** Simulated far field gain

Table 2 Optimized dimensions of the 1×2 and 1×4 conformal array antenna

Parameter	Value (mm)
PL	10
PW	10
IG	1.0
ID	2.4
FL	5.5
FW	1.0
d1	22
d2	22
d	44
R for 1×2 array	20
R for 1×4 array	80
SUBL for 1×2 array	30
SUBW for 1×2 array	40
SUBL for 1×4 array	50
SUBW for 1×4 array	85

1×4 Corporate feed Conformal Antenna Array. The top view of the 1×4 corporate feed planar and conformal antenna array is shown in Fig. 3a, b. The cylinder radius is considered as 80 mm for the 1×4 conformal antenna array. The simulated return loss $|S_{11}|$ for the 1×4 corporate feed conformal antenna array is presented in Fig. 3c. The obtained return loss is -41.80 dB. The surface current distribution is shown in Fig. 3d. Figure 3e, f show the simulated radiation pattern of E-plane and H-plane. Figure 3g, h show simulated 3D far-field gain and front-to-back ratio over frequency. The simulated gain is 10.5 dBi and SLL is -18.9 dB for the 1×4 series fed conformal antenna array.

3 Conclusion

This study describes the design and implementation of a conformal microstrip patch antenna array on a conducting cylindrical surface. The X-band (9.65 GHz) conformal antenna arrays for spaceborne synthetic aperture radar applications are presented in this research article. Table 3 shows the performance of a single conformal antenna, 1×2 corporate feed conformal antenna array and 1×4 corporate feed conformal antenna array. The gain and directivity obtained for the 1×4 conformal array compared to the 1×2 conformal array is better. When the array bends, the performance characteristics deteriorate even further. A maximum gain of 10.5 dBi with -18.9 SLL was achieved for 1×4 conformal array in X-band. Telemetry, satellites,

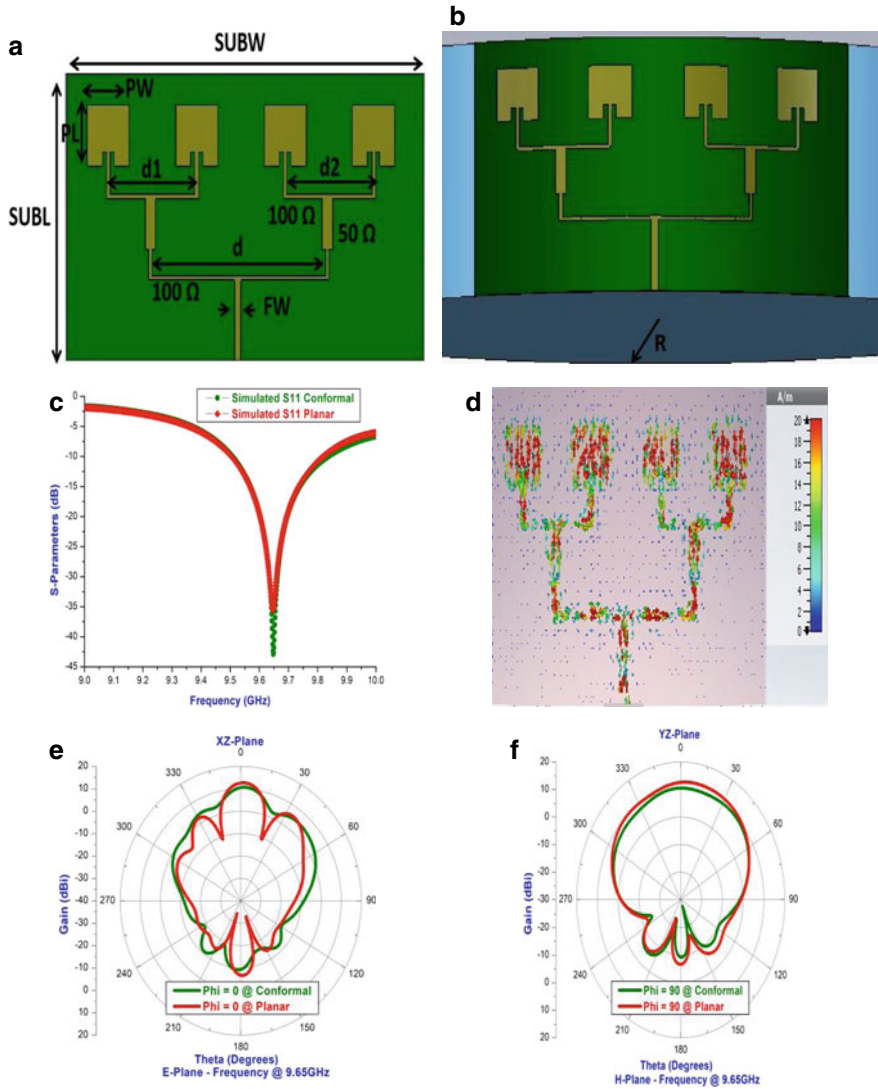


Fig. 3 **a** 1×4 planar array antenna. **b** 1×4 conformal array antenna. **c** Simulated return loss. **d** Simulated surface current distribution. **e** Radiation pattern (E-Plane). **f** Radiation pattern (H-Plane). **g** Simulated far field gain. **h** Front-to-back ratio over frequency

aeroplanes and medical operations are all possible uses for the suggested architecture. This research will pave the way for the creation of printed antennas on flexible substrates.

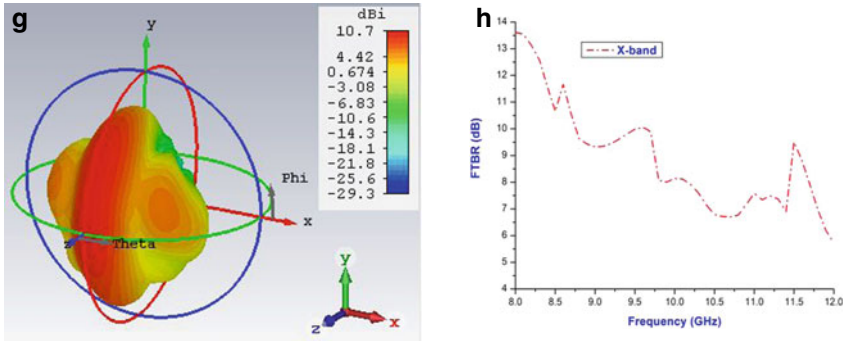


Fig. 3 (continued)

Table 3 Performance analysis of conformal antenna arrays

Parameter	Single		1 × 2 array		1 × 4 array	
	Planar	Conformal	Planar	Conformal	Planar	Conformal
Return loss (dB)	-21.24	-31.24	-20.80	-21.21	-34.69	-41.80
Bandwidth (GHz)	9.55–9.75	9.53–9.76	9.47–9.80	9.45–9.83	9.50–9.81	9.50–9.83
VSWR	1.18	1.05	1.20	1.19	1.03	1.01
Gain (dBi)	7.48	7.38	9.64	8.38	12.8	10.5
SLL (dB)	-13.0	-13.2	-15.1	-16.6	-18.1	-18.9

References

1. Measuring Ocean Currents - 1st Edition. <https://www.elsevier.com/books/measuring-ocean-currents/joseph/978-0-12-415990-7>
2. Gal T Design of a C-band conformal phased-array antenna for airborne synthetic aperture radar 90
3. Chauhan B (2016) Four element conformal array antenna for X-band and Ku-band applications 14:10
4. Kumari P, Chauhan B, Vijay S (2016) Performance analysis of microstrip conformal antenna array and effect of mutual coupling for different curvature. IJCA 135:30–35
5. Frasier SJ, Camps AJ (2001) Dual-beam interferometry for ocean surface current vector mapping. IEEE Trans Geosci. Remote Sens. 39:401–414
6. Kissinger D (2012) Millimeter-wave receiver concepts for 77 GHz automotive radar in silicon-germanium technology. Springer Science & Business Media
7. Josefsson L, Persson P (2006) Conformal array antenna theory and design: Josefsson/conformal array antenna theory and design. Wiley. <https://doi.org/10.1002/047178012X>
8. Bancroft R (2009) Microstrip and printed antenna design. The Institution of Engineering and Technology (2009)
9. Balanis CA (2005) Antenna theory: analysis and design. Wiley (2005)
10. Monica J, Jothilakshmi P (2020) A design of bandwidth-enhanced conformal antenna for aircraft applications. IETE J Res 1–13. <https://doi.org/10.1080/03772063.2020.1829507>
11. Patil S, Ar M, Sudhendra C, Ts R, Biswas D Design and Implementation of a Conformal Omnidirectional Microstrip Antenna Array on Cylindrical Surface 2, 5

12. Kothapudi VK, Kumar V (2018) A 6-port two-dimensional 3×3 series-fed planar array antenna for dual-polarized X-band airborne synthetic aperture radar applications. *IEEE Access* 6:12001–12007
13. Kothapudi VK, Kumar V (2019) Vertical polarized 1×3 series-fed linear array with gain and front-to-back ratio enhancement for airborne SAR-X applications. *PIER M* 80:169–179
14. Kothapudi VK, Kumar V (2019) SFCFOS uniform and chebyshev amplitude distribution linear array antenna for K-band applications. *J Electromagn Eng Sci* 19:64–70
15. Shankar SG, Beenamole KS (2015) A comparative study on arc conformal arrays for different curvatures 4
16. Beenamole KS, Shankar G (2016) Studies on conformal antenna arrays placed on cylindrical curved surfaces 3
17. Bhowmik LM et al. (2013) Design and analysis of conformal ku-band microstrip patch antenna arrays. In: 2013 IEEE international symposium on phased array systems and technology. IEEE, pp 815–820. <https://doi.org/10.1109/ARRAY.2013.6731932>
18. Bai T et al (2021) Design of cylindrical conformal array antenna based on microstrip patch unit. *ACES* 36:1008–1014
19. Sharma A, Dev Gupta S (2015) Design and analysis of rectangular microstrip patch antenna conformal on spherical surface. In: 2015 international conference on signal processing and communication (ICSC). IEEE, pp 366–369. <https://doi.org/10.1109/ICSPCom.2015.7150678>
20. Paul LC, Hasan MdI, Azim R, Islam MdR, Islam MT (2020) Design of high gain microstrip array antenna and beam steering for X Band RADAR application. In: 2020 Joint 9th international conference on informatics, electronics and vision (ICIEV) and 2020 4th international conference on imaging, vision and pattern recognition (icIVPR), IEEE, pp 1–7. <https://doi.org/10.1109/ICIEVicIVPR48672.2020.9306519>
21. Design of Planar and Conformal Microstrip Patch Antenna for Avionics Applications. *IJITEE* 8:4363–4365
22. Kothapudi VK, Kumar V (2018) Compact 1×2 and 2×2 Dual polarized series-fed antenna array for X-band airborne synthetic aperture radar applications. *J Electromagn Eng Sci* 18:117–128

Multilingual Text Recognition System



Mansi Maithani, Dibyasha Meher, and Sumita Gupta

Abstract Most of us take our reading and writing abilities for granted until the onset of visual impairment. After all, literacy is a key to personal independence and access to information. Tasks that require little thinking on our part, such as reading labels and business cards, can prove to be very challenging for visually impaired people. The proposed approach suggests the use of OCR techniques to extract information from printed and electronic documents and convert this information into a voice output for the benefit of the visually impaired. This paper considers the functioning of two of the common OCR engines, Tesseract OCR and EasyOCR. We have used these tools for the text recognition of Indian languages, Bengali and Tamil along with English text. We conclude this paper with a comparative study of these two tools by considering business cards as an input. From business cards we have tried to extract all legible data by using Tesseract and EasyOCR and compared these tools based on their accuracy parameter.

Keywords Optical character recognition (OCR) · Open source · Tesseract · EasyOCR · Multilingual text recognition

1 Introduction

OCR (Optical Character Recognition) is a method of converting text into images, scanned or printed text images into content that can be edited for additional processing [1]. This technology makes it possible for text to be automatically recognised by a machine. It's as if the human body's eye and mind combined. An eye can see or view the text from the visuals, but the brain processes the information in a different way as well as reads the retrieved text by eye and interprets it. A few difficulties can be solved by developing a computerised OCR system. It could be tough for the computer to

M. Maithani (✉) · D. Meher · S. Gupta
Department of Computer Science and Engineering, Amity School of Engineering and
Technology, Amity University, Noida, India
e-mail: mansi0356@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
V. V. S. S. Chakravarthy et al. (eds.), *Advances in Signal Processing, Embedded
Systems and IoT*, Lecture Notes in Electrical Engineering 992,
https://doi.org/10.1007/978-981-19-8865-3_9

103

tell the difference, for example between the number ‘0’ and the letter ‘o’ Second, it’s possible that text extraction will be extremely difficult and challenging. In the year 1955 the first OCR commercial system was deployed. OCR was utilised to input a sales report into a computer. After that, the OCR approach has shown to be extremely useful wherein documents present physically in the office are being converted to digital format. There are numerous applications wherein OCR is being used. In this paper we discuss about business card character recognition. This paper considers the functioning of two of the common OCR engines, Tesseract OCR and EasyOCR. Experimental analysis for the results of the working of these two tools is given for business cards in English, as well as for business cards with multilingual text. We have used these tools for the text recognition of Indian languages, Bengali and Tamil along with English text. Section 2 discusses existing techniques used to perform OCR and their performance and drawbacks. Section 3 presents our approach to the problem, and Sect. 4 contains the methodology applied by us. Section 5 discusses the experiment results, and Sect. 6 concludes the article with our findings. We conclude this paper by comparative study of these two tools by considering business cards as an input.

2 Text Recognition Systems

The OCR technology was introduced back in the early 1990s. Since then, OCR has undergone various changes but it still remains one of the breakthroughs of the digitized world. Some of the common use-cases of OCR technology are forms processing, e.g. bills, receipts, cheque processing etc. The performance of OCR models draws on multilayer artificial neural networks. For computer vision, the most common types are recurrent neural networks (RNN) or more precisely long short-term memory (LSTM), and convolutional neural networks (CNN). Some of the common open-source OCR engines are Tesseract, OCROpus, Calamari OCR, Kraken, Keras OCR and EasyOCR. Kraken and Calamari OCR are both dissidents of OCROpus where Kraken runs on Linux and OSX only and Calamari OCR is a CLI-only framework which has also been derived from Kraken [2]. Kraken currently only contains the generalized English language model and a Syriac language model and for Calamari OCR, French is the primary alternate language to English.

For this experiment, we have chosen two OCR frameworks to estimate the performance of text recognition. These are Tesseract OCR and EasyOCR by JaidedAI [3, 4]. The framework is written in Python and supports multiple languages. The main criteria for choosing these OCR engines were their support for character recognition on Indian languages, open-source code and flexible APIs. In this experiment, we test these engines on business cards.

Adaptive recognition: Tesseract searches pixels, letters, phrases, and sentences for templates. It employs a two-step process known as adaptive recognition. It necessitates one data step for character identification, followed by a second stage to fill in any missing letters with letters that match the word or phrase context.

Tesseract reached maturity with version 3.x, when it began to support a wide range of picture formats and introduced a huge number of scripts (languages). Traditional computer vision algorithms are used in Tesseract 3.x. In many areas of Computer Vision, Deep Learning- based algorithms have outperformed standard machine learning techniques by a large margin in terms of accuracy in recent years. One of the most well-known examples is handwriting recognition. It was just a matter of time before Tesseract got its hands on a deep learning- based identification engine as well. Tesseract 4 has additional deep-learning based capability with LSTM network (a type of Recurrent Neural Network) based OCR engine which is focused on the line recognition but also provisions the legacy Tesseract OCR engine of Tesseract 3 which works by recognising character designs. Figure 1 shows the architecture of the Tesseract framework.

The other OCR engine that we will be using is EasyOCR. It is also an open-source OCR like Tesseract. EasyOCR is built with Python and Pytorch deep learning library. It is a python module that allows you to transform a picture into text. The CRAFT method is used for detection, and the CRNN model is used for recognition. Feature extraction, sequence labelling (LSTM), and decoding (CTC) are the three basic components. EasyOCR has few software prerequisites and can be utilised directly through its API.

CRAFT (Character-Region Awareness For Text detection)—CRAFT is used for scene text detection, which is a task that involves detecting and labelling text sections with bounding boxes, in a complicated background. CRAFT was proposed in 2019 and its main objective is to localize the individual character regions and link the detected characters to a text instance. The backbone of CRAFT is VGG-16, which is a fully convolutional network architecture. To put it another way, VGG16 is the feature extraction architecture that is utilised to encode the network’s input into a feature representation [5] CRAFT is multilingual i.e. it can detect text written in any script. Figure 2 shows the architecture of the EasyOCR framework.

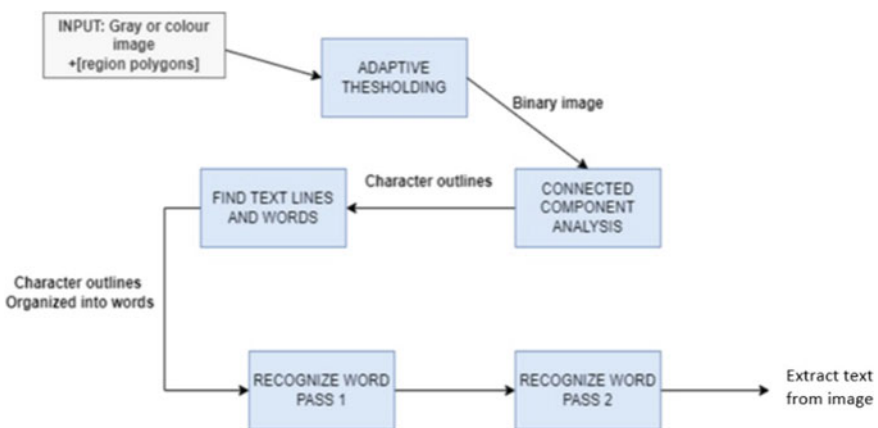


Fig. 1 Architecture of tesseract

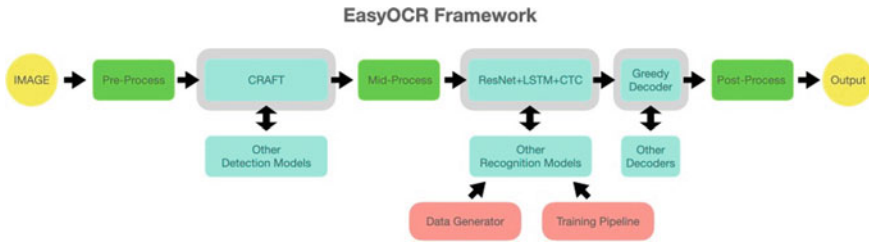


Fig. 2 Architecture of EasyOCR

In this research paper, the main focus is on Tesseract and EasyOCR due to their support for multiple languages and ease of use [6]. The Tesseract framework works on a wide range of FOSS and proprietary interfaces and GUIs and supports 116 languages with room for adaptation [7]. Tesseract is considered one of the most precise OCR frameworks [8]. EasyOCR is also a repository supporting more than 80 languages. It supports many popular script types, including Latin, Cyrillic, Chinese and Arabic. It is by far the most user-friendly approach to implement OCR. It has been developed by the JaidedAI company. In this research paper, we use these OCR engines on Indian languages—Bengali and Tamil, with their respective scripts.

3 Multilingual Text

In this multilingual text recognition system, we are able to detect text in combination of two languages—Bengali and English as well as Tamil and English. For this we have used both Tesseract and EasyOCR and used the standard trained datasets provided with these engines, for the respective language. EasyOCR can read multiple languages at the same time but they have to be compatible with each other. English is compatible with all languages. Languages that share most of the character (e.g. Latin script) with each other are compatible. To use Tesseract to do OCR on multilingual documents, we place all language data files in the same folder so that they can be accessed under the same path, then pass their associated language codes using the ‘+’ operator. We have used business cards in Bengali and Tamil as our inputs for this purpose. Figure 3 shows the proposed approach for the Multilingual Text Recognition System.

3.1 Bengali Script

There are several well-known languages around the world, such as English, Chinese, Arabic, Japanese, Bengali, and so on. Bengali is the world’s fifth most spoken native language and the world’s seventh most spoken language overall [9]. Many letters,

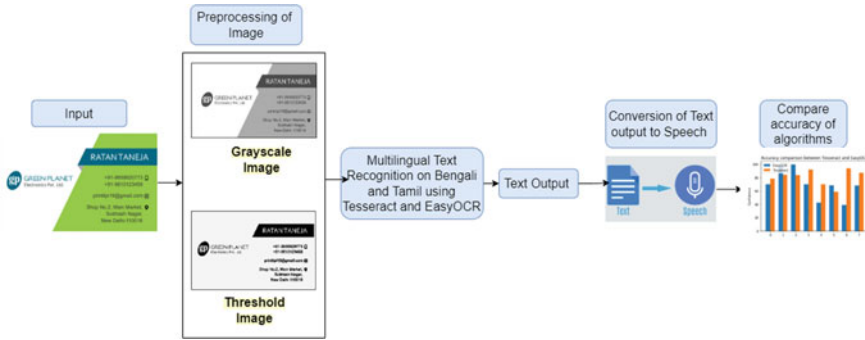


Fig. 3 Architecture of proposed text recogniser system

textbooks, and other documents are examples of many types of documents. Novels, government documents, historical documents, newspapers, and other sources of information periodicals, data input forms, and other materials are available in Bengali that needs to be upgraded. Many well- known pieces of Bengali literature have to be computerised and archived as well.

The Bengali language has far more complex characters than any other language due to its structure. It comprises of 11 vowels (Shoroborno), 39 consonants (Benjon-borno), and two or more characters blending to create compound (Juktoborno) characters in the Bengali language.

As a result, developing an OCR that identifies Bengali characters is more difficult than the usual method from other languages. As a result, since the middle of the 1980s, numerous researchers have been working on constructing an OCR system for recognising Bengali characters. Since then, Bengali OCR has become a huge topic of study, with numerous academics proposing cutting-edge methods [10].

The following items make up the overall contribution:

- We introduce an end-to-end word recognition system for the Bengali language as part of our total contribution. To the best of our knowledge, this research project comprises and investigates end-to-end strategy in Bengali OCR.
- We test the performance of the end-to-end technique using the well-known Bengali dataset BanglaWriting.
- In addition, the recovered features (scanned from left to right of a word image) are fed into the tesseract character recognition model and the easy-OCR model at the same time.
- Figure 4 shown below is an example of a Bengali business card used in our experiment.

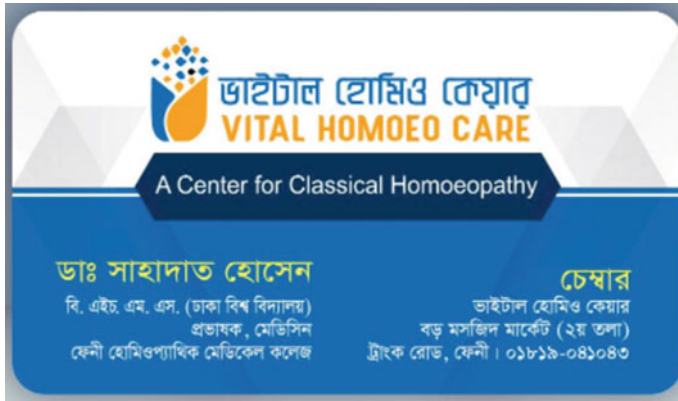


Fig. 4 Bengali business card

3.2 Tamil Script

There are 12 vowels, 18 consonants, and one unique character in Tamil (known as Ayudha Ezhuthu). Grantha Letters are five additional consonants borrowed from Sanskrit and English to express north Indian and English words/syllables [11]. As a result, the script features 36 distinct fundamental letters: 12 vowels, 18 consonants, 1 Ayudha Ezhuthu, and 5 Granthas. Figure 5 depicts the basic characteristics.

The combination of 12 vowels and 18 consonants yields 216 compound characters, bringing the total number of characters to 247 [216 + 12 + 18 + 1]. Figure 6 shows an example of compound character generation for a consonant.

அ	ஆ	இ	ஈ	உ	ஊ	எ	ஏ	ஐ	ஒ	ஔ	ஔள	ஃ	ஐ	ஷ	ஹ	ஸ	க்ஷ
a	ā	i	ī	u	ū	e	ē	ai	o	ō	au	aytam	j	ṣ	h	s	ks
க	ச	ந	ன	த	ந	ப	ம	ய	ர	ல	ள	ற	வ	ழ	ள		
k	c	ṅ	ṅ	t	n	t	n	p	m	y	r	l	l	r	v	z	n

Fig. 5 Vowels, consonants, Ayutha Ezhuthu and Grantha letters of Tamil Script

க்+அ	க்+ஆ	க்+இ	க்+ஈ	க்+உ	க்+ஊ	க்+எ	க்+ஏ	க்+ஐ	க்+ஒ	க்+ஔ	க்+ஔள
க	கா	கி	கீ	கு	கூ	கெ	கே	கை	கொ	கோ	கௌ

Fig. 6 Compound character formation for a consonant with set of vowels



Fig. 7 Input images

4 Methodology

In this experiment, we chose pytesseract as the Python wrapper for Tesseract. The version of Tesseract being used is 5.0.1. The dataset fed to the algorithm is tessdata/eng.traineddata for the English OCR. The Language Abbreviation instructs the OCR engine the language to seek for during OCR, and the Language Data Path should have the corresponding language's data file. This file contains all of the information that were utilised to train the OCR engine in the first place tessdata is the standard model that only works with Tesseract 4.0.0 and above. Similarly, the dataset used for Bengali and Tamil are the standard traineddata files—ben.traineddata, tam.traineddata—stored in the tessdata folder.

4.1 Load Image

The images shown in Fig. 7 have been uploaded from the internet. The output quality and accuracy are determined by the image quality that was loaded. It should be clear [12].

4.2 Process Image

The image has been processed in the second stage. It determines whether or not the image is clear. The output may not be clear and correct if the image is fuzzy or not taken properly.

Pre-processing is a crucial step before moving on to feature extraction since it ensures that the results are appropriate for the subsequent phases. The proportion of each stage that is completed successfully provides the OCR rate of success.

A. Character Recognition Performance Affecting Variables

The accuracy with which characters are recognised using OCR is influenced by a number of factors. The sharpness of the scan and the clarity of the scanned photograph [13], published document category (photocopied or new technology printer), paper performance, and semantic nuances are all factors to take into account. Uneven illumination and watermarks are two issues that affect the correctness of OCR.

B. The Importance of Character Recognition Pre-processing phase

The pre-processing stage is required to achieve a higher text recognition pace. Using structural and methodical pre-processing set of rules, the text recognition technique becomes more strongly built [13] with noise elimination, image enhancement, picture thresholding, page and text segmentation, text normalisation, and morphological operations all being used.

C. Pre-processing methods

Binary/grey pictures are used in the majority of OCR applications. Without conducting the pre-processing stage, the photos may have watermarks and/or a non-uniform backdrop, making recognition harder. To accomplish this, you'll need to take a few steps. The first stage, known as the image enhancement technique, is to alter the contrast or remove noise from the image. Following that, thresholding is used to remove page segmentation to distinguish the images from the text, accompanied by watermarked image and/or noise. Text segmentation to individual character separation is the next phase, followed by morphological processing. If the preprocessed image has eroded areas in the characters, morphological processing is required to add pixels.

4.2.1 Grayscale Image

For numerous reasons, greyscale photos have higher OCR accuracy than bitonal ones:

- Image processing such as deskewing, despeckling, and other techniques. Greyscale image processing provides for more efficient image processing, resulting in page versions with less speckling, cleaner backgrounds, and sharper text characters and artwork. Greyscale, in particular, is expected to provide much better deskewing since images may be rotated without generating breaks in the letter shapes, which could confuse the OCR application [14]
- Binarization algorithms that are specialised. Scanners are designed to be viewed by humans. Higher-contrast photos, on the other hand, operate better with OCR applications. As a result, OCR accuracy will be improved by using specialised binarization algorithms that are optimised for OCR.
- Further details. Inherently, greyscale scans carry more information that OCR systems can employ to process text.

Figure 8 shows the converted grayscale images for the corresponding input images shown in Fig. 7.



Fig. 8 Greyscale image

4.2.2 Adaptive Thresholding

Not every document in the office is a perfect replica printed on pure white paper. Adaptive thresholding can help with the following types of documents:

- The originals were hand-drawn with very tiny strokes. The tiny strokes combine with the scanner’s low resolution to produce pixels that easily vanish when thresholded [13].
- Copies that have been poorly duplicated. Text that has been lightened can have the same impact as thin strokes. The gaps between strokes may darken as a result of the darkening.
- Text on a coloured or halftoned backdrop, or beneath a coffee stain.

Thresholding can easily eliminate such text. Even though the text is preserved, allowing halftone dots to pass through the OCR machine might cause chaos. Figure 9 shows the converted threshold images for the corresponding input images shown in Fig. 7.



Fig. 9 Threshold image

4.3 Extract Text from Image

In a database, each character in a given language is saved. The image’s characters are compared to the characters in the database. The character will be displayed on the output if a match is detected. The same procedure will be followed for each character. If a match is not found then the pointer will move onto the next character and this process will occur until a space character or an empty character is found [12].

5 Implementation, Result and Analysis

The two OCR engines were tested on a variety of business cards, with varying degrees of output accuracy. The results of experiments are presented in Tables 1, 2 and 3.

Table 1 shows the accuracy of Tesseract and EasyOCR on English text, specifically on business cards. Here, we have taken the confidence as the average of confidence of each line in a business card.

Figure 10 shows a business card and its line-wise accuracy using both EasyOCR and Tesseract.

Table 1 Probability of correct character recognition(confidence) for a business card in English

File name	Tesseract OCR	EasyOCR	Better OCR engine
	<i>English</i>		
visitingcard3.png	0.81	0.69	Tesseract
businesscard2.jpg	0.76	0.65	Tesseract
visitingcard2.png	0.70	0.82	EasyOCR
visitingcard.png	0.86	0.80	Tesseract
card.png	0.90	0.80	Tesseract

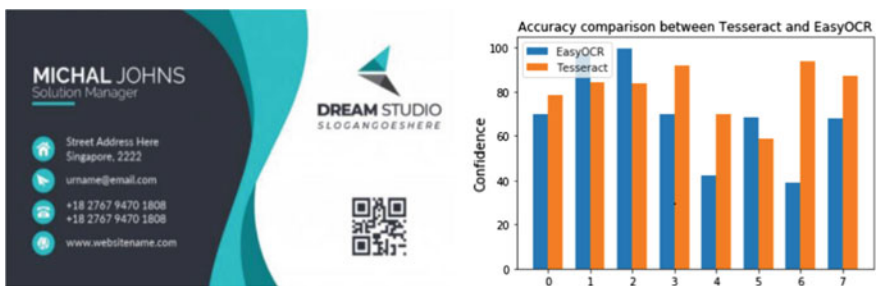


Fig. 10 Business card and its line-wise accuracy using both Tesseract and EasyOCR

Table 2 Probability of correct character recognition of business cards with Bengali script

File name	Tesseract OCR	EasyOCR	Better OCR engine
	<i>Bengali</i>		
bengali.png	0.46	0.47	EasyOCR
bengali2.png	0.36	0.58	EasyOCR
bengali3.png	0.30		EasyOCR

Table 3 Probability of correct character recognition of business cards with Tamil script

File name	Tesseract OCR	EasyOCR	Better OCR engine
	<i>Tamil</i>		
tamil1.png	0.62	0.75	EasyOCR
tamil2.png	0.51	0.66	EasyOCR
tamil3.png	0.20	0.63	EasyOCR

Table 2 shows the probability of correct character recognition for a business card in the Bengali language. Table 3 shows the probability of correct character recognition for images with Tamil text.

In this experiment, we have found that Tesseract usually performs better in comparison to EasyOCR when it comes to business cards in the English language. Although, in our experiment, Tesseract shows output with more accuracy(in case of English script), it fails to recognise text on business cards with complex or irregular backgrounds whereas EasyOCR remains consistent in its output irrespective of background. If the background is too noisy, then Tesseract might fail on that kind of output. For the Bengali and Tamil scripts, EasyOCR performs far better and is able to extract the entire text. Tesseract, in comparison, lacks accuracy and only recognises some parts of the text while failing to extract some other text.

6 Conclusion

In this paper, we set up the experiment and made a comparative analysis of the performance of the most common OCRs (Tesseract OCR and EasyOCR). The results obtained in the above sections are obtained by extracting text from business cards. The algorithms face greater difficulty in identifying words with font variations, slanting characters and abnormal writing patterns. It is also observed that similar shaped characters and joined characters face difficulty in being recognised properly. The results show that Tesseract works better on English script and EasyOCR works far better on Bengali and Tamil multilingual text that we have considered in this paper. It is also observed that when processing images with more content such as outlines or other text, it may change the outcome as EasyOCR works better with noise and

extracts the entire text. As we are interested in extracting text from business cards, we have considered both tools for serving this specific purpose.

References

1. What is OCR and OCR Technology, 2020. <https://pdf.abbyy.com/learning-center/what-is-ocr>
2. The OCRopus Open Source OCR System Thomas M. Breuel DFKI and U. Kaiserslautern Kaiserslautern, Germany Sections
3. Tesseract Open-Source OCR: <http://code.google.com/p/tesseract-ocr>
4. EasyOCR, 2021. <https://www.jaided.ai/easyocr>
5. Object Detection with Deep Learning: A Review Zhong-Qiu Zhao, Member, IEEE, Peng Zheng, Shou-tao Xu, and Xindong Wu
6. Smelyakov K, Chupryna A, Darahan D, Midina S (2021) Effectiveness of modern text recognition solutions and tools for common data sources. Kharkiv National University of Radio Electronics, 14 Nauky Ave., Kharkiv, 61166, Ukraine
7. OCR4all - An Open-Source Tool Providing a (Semi-)Automatic OCR Workflow for Historical Printings Christian Reul 1 , Dennis Christ 1 , Alexander Hartelt 1 , Nico Balbach 1 , Maximilian Wehner 1 , Uwe Springmann 2 , Christoph Wick 1 , Christine Grundig 3 , Andreas Büttner 4 , and Frank Puppe 1
8. Patel C, Patel A, Patel D (2012) Optical character recognition by open source OCR tool tesseract: a case study. *Int. J. Comput. Appl.* 55(10) (0975–8887)
9. Safir FB, Ohi AQ, Mridha MF, Monowar MM, Abdul Hamid Md. End-to-end optical character recognition for Bengali handwritten words
10. Sahare P, Dhok SB (2018) Multilingual character segmentation and recognition schemes for Indian document images. In: *IEEE Access*, vol 6, pp 10603–10617
11. uTHCD: A New Benchmarking for Tamil Handwritten OCR Nousath Shaf i and Faizal Hajamohidden
12. Chawla M, Jain R, Nagrath P (2020) Implementation of tesseract algorithm to extract text from different images. In: *International conference on innovative computing and communication (ICICC 2020)*
13. Karthick K, Ravindrakumar KB, Francis R, Ilankannan S (2019) Steps involved in text recognition and recent research in OCR; A Study. *Int J Recent Technol Eng (IJRTE)* 8(1). ISSN: 2277-3878
14. Akhil S (2016) An overview of tesseract OCR engine. Department of Computer Science and Engineering National Institute of Technology, Calicut Monsoon-2016

Implementing Symmetric Boundary Condition in Electromagnetic Harmonic Analysis: Two Different Approaches



Sreekanth Karanam, Durgarao Kamireddy, and Arup Nandy

Abstract In electromagnetic analysis, we can identify a specific problem that is symmetric about a plane. Due to the symmetry, we can use half of the domain for FEM analysis with proper symmetric boundary condition on the symmetry face. In this work, we have proposed and compared two different ways to apply this symmetric boundary condition. Both the methods result in significant reduction in the computational cost maintaining same accuracy.

Keywords FEM · Electromagnetics · Harmonic analysis · Symmetric boundary condition

1 Introduction

In the field of computational electromagnetics, the importance of application of the finite element method (FEM) is well known, and still it is an interesting research area. This field of research has a wide range of applications in antenna radiations, waveguide transmissions, cavity resonant problems, etc. [1, 2]. FEM is used to solve eigenvalue problems by using both nodal elements [3, 4] and edge elements [5–12]. It is also used in harmonic analysis [13–16], and transient analysis [17] of electromagnetic radiation and scattering in both the interior and exterior domains. To apply the FEM in the numerical analysis, domain has to be truncated at a certain distance to convert to a finite domain, and then proper absorbing boundary conditions (ABCs) has to be imposed on the truncated surface. These ABCs include both first order such as Sommerfeld radiation condition [18] and higher-order ABCs reported in [13, 15, 16]

S. Karanam (✉) · D. Kamireddy · A. Nandy
Department of Mechanical Engineering, Indian Institute of Technology Guwahati, Guwahati
781039, India
e-mail: skaranam@iitg.ac.in

D. Kamireddy
e-mail: durga176103010@iitg.ac.in

A. Nandy
e-mail: arupn@iitg.ac.in

to solve various electric field problems. In [19, 20], radiation condition at truncated boundary is applied with the finite element method along with the boundary integral method (BI). Here, internal domains and truncated domains (external domains) are discretized with nodal finite elements and edge finite elements, respectively. But in [4, 21–27], nodal finite elements are used for both interior and exterior domains. While using the nodal elements, the main difficulty exists in modelling objects with sharp corners and edges in predicting the singular field values correctly [18, 21, 24]. Here, in the present work, we want to simulate numerically the symmetric boundary condition in solving the scattering and radiation fields problems.

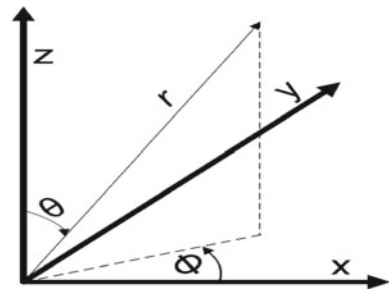
The rest of the paper is arranged in the following manner. In Sect. 2, we have discussed the existence of symmetry from physical perspective citing field conditions on the symmetric face. Mathematical formulations are presented in Sect. 3, followed by numerical examples in Sect. 4. Finally, we have concluded our findings in Sect. 5.

2 Existence of Symmetry in Electromagnetic Scattering by Conducting Sphere: Physical Interpretation with Field Conditions

Consider an electromagnetic plane wave $E_{\text{inc}} = E_0 e^{-ikz} e_x$ travelling along z -axis which is incident on a conducting sphere of radius a . The wave vector is only in x -direction; hence, the incident wave is symmetric about xz -plane. Also, the conducting sphere and its surrounding air is symmetric about xz -plane. Hence, the resulting scattering field also will be symmetric about xz -plane. Due to the existing symmetry, for the scattering field, $E_y(y = 0)$ should be zero, and following relations should hold.

$$\begin{aligned} E_x(\bar{x}, \bar{y}, \bar{z}) &= E_x(\bar{x}, -\bar{y}, \bar{z}), & E_x(r, \theta, \phi) &= E_x(r, \theta, -\phi), \\ E_y(\bar{x}, \bar{y}, \bar{z}) &= -E_y(\bar{x}, -\bar{y}, \bar{z}), & E_y(r, \theta, \phi) &= -E_y(r, \theta, -\phi), \\ E_z(\bar{x}, \bar{y}, \bar{z}) &= E_z(\bar{x}, -\bar{y}, \bar{z}), & E_z(r, \theta, \phi) &= E_z(r, \theta, -\phi). \end{aligned}$$

Fig. 1 Spherical coordinates system



Where Cartesian and spherical coordinates are as shown in Fig. 1. We have cross-checked that all the above conditions are satisfied by the analytical scattered field [28].

3 Mathematical Formulation

For electromagnetic wave propagation, the Maxwell's governing equation for electric field \mathbf{E} can be given as [29].

$$\frac{\epsilon_r}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} + \mu_0 \frac{\partial \mathbf{j}}{\partial t} + \nabla \times \left(\frac{1}{\mu_r} \nabla \times \mathbf{E} \right) = \mathbf{0}. \quad (1)$$

where $\epsilon_r := \epsilon/\epsilon_0$ is the relative permittivity and $\mu_r := \mu/\mu_0$ is the relative permeability. ϵ_0 and μ_0 are permittivity and permeability for vacuum respectively and $c = 1/\sqrt{\epsilon_0\mu_0}$ is the speed of light.

3.1 Harmonic Analysis

For harmonic analysis Eq. (1) can be written as

$$\nabla \times \left(\frac{1}{\mu_r} \nabla \times \mathbf{E} \right) - \frac{k^2}{\mu} \mathbf{E} = -i\omega \mathbf{j}, \quad (2)$$

\mathbf{j} is the current density, $i = \sqrt{-1}$, μ is the magnetic permeability, $k = k_0\sqrt{\mu_r\epsilon_r}$ is the wave number of the medium, $k_0 = \omega/c$ is the wave number of vacuum, and ω is the excitation frequency. In the absence of charges, the electric field is subject to the condition

$$\nabla \cdot (\epsilon \mathbf{E}) = 0. \quad (3)$$

3.2 Potential Formulation

In Potential formulation, \mathbf{E} is replaced with $\mathbf{A} + \nabla\psi$, where ψ and \mathbf{A} represent scalar and vector potentials, respectively. Then,

$$\mathbf{E} = \mathbf{A} + \nabla\psi, \quad (4a)$$

$$\mathbf{H} = \frac{i}{\mu\omega} \nabla \times \mathbf{A}. \quad (4b)$$

In terms of potentials, the governing differential Eq. (2) and the constraint Eq. (3) can be expressed as [14]

$$\nabla \times \left(\frac{1}{\mu} \nabla \times \mathbf{A} \right) - \frac{k^2}{\mu} (\mathbf{A} + \nabla \psi) = -i\omega \mathbf{j}, \quad (5a)$$

$$\nabla \cdot (\epsilon \mathbf{A}) + \nabla \cdot (\epsilon \nabla \psi) = 0. \quad (5b)$$

3.3 Finite Element Formulation

We can now write the FEM equation for the above two Eqs. 5a and 5b as [14]

$$\begin{bmatrix} \mathbf{K}_{AA} & \mathbf{K}_{A\psi} \\ \mathbf{K}_{\psi A} & \mathbf{K}_{\psi\psi} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{A}} \\ \hat{\psi} \end{bmatrix} = \begin{bmatrix} \mathbf{F}_A \\ \mathbf{F}_\psi \end{bmatrix}, \quad (6)$$

where the terms \mathbf{K}_{AA} , $\mathbf{K}_{A\psi}$, $\mathbf{K}_{\psi A}$ and $\mathbf{K}_{\psi\psi}$ can be found in [14]. After solving for $\hat{\mathbf{A}}$ and $\hat{\psi}$ values, \mathbf{E} and \mathbf{H} are obtained using Eq. 4 as

$$\mathbf{E} = \mathbf{N} \hat{\mathbf{A}} + \mathbf{B}_\psi \hat{\psi},$$

$$\mathbf{H} = \frac{i}{\mu\omega} \mathbf{B} \hat{\mathbf{A}}.$$

3.4 Implementation of Symmetric Boundary Condition Using Lagrange Multiplier Method (Method-1)

Consider the condition,

$$\int_{\Gamma} \lambda (\mathbf{E} \cdot \mathbf{n}) d\Gamma = \mathbf{0}, \quad (7)$$

where λ is Lagrange multiplier. From potential formulation, Eq. 7 can be written as

$$\int_{\Gamma} \lambda [(\mathbf{A} + \nabla \psi) \cdot \mathbf{n}] d\Gamma = \mathbf{0}, \quad (8)$$

Applying variational operator to the above Eq. 8, we get as

$$\int_{\Gamma} (\lambda_{\delta} (\mathbf{A} \cdot \mathbf{n} + \nabla \psi \cdot \mathbf{n}) + \lambda (\mathbf{A}_{\delta} \cdot \mathbf{n} + \nabla \psi_{\delta} \cdot \mathbf{n})) d\Gamma = \mathbf{0}. \quad (9)$$

We can discretize λ and λ_δ as $\lambda = N_\lambda \hat{\lambda}$ and $\lambda_\delta = N_\lambda \hat{\lambda}_\delta$ where

$$N_\lambda = [N_1 \ N_2 \ N_3 \ \dots]. \quad (10)$$

Other terms $\mathbf{A} \cdot \mathbf{n}$, $\nabla \psi \cdot \mathbf{n}$, $\mathbf{A}_\delta \cdot \mathbf{n}$, and $\nabla \psi_\delta \cdot \mathbf{n}$ can be discretized as given in [14]. Different terms in Eq. 9 can be expressed as

$$\begin{aligned} \lambda_\delta (\mathbf{A} \cdot \mathbf{n}) &= \hat{\lambda}_\delta^T N_\lambda^T \mathbf{n}^T N \hat{\mathbf{A}}, & \lambda (\mathbf{A}_\delta \cdot \mathbf{n}) &= \hat{\mathbf{A}}_\delta^T N^T \mathbf{n} N_\lambda \hat{\lambda}, \\ \lambda_\delta (\nabla \psi \cdot \mathbf{n}) &= \hat{\lambda}_\delta^T N_\lambda^T \mathbf{n}^T \mathbf{B}_\psi \hat{\psi}, & \lambda (\nabla \psi_\delta \cdot \mathbf{n}) &= \hat{\psi}_\delta^T \mathbf{B}_\psi^T \mathbf{n} N_\lambda \hat{\lambda}. \end{aligned}$$

After imposing symmetric boundary condition, FEM equation in potential formulation can be written as [14]

$$\begin{bmatrix} \mathbf{K}_{AA} & \mathbf{K}_{A\psi} & \mathbf{K}_{A\lambda} \\ \mathbf{K}_{\psi A} & \mathbf{K}_{\psi\psi} & \mathbf{K}_{\psi\lambda} \\ \mathbf{K}_{\lambda A} & \mathbf{K}_{\lambda\psi} & \mathbf{K}_{\lambda\lambda} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{A}} \\ \hat{\psi} \\ \hat{\lambda} \end{bmatrix} = \begin{bmatrix} \mathbf{F}_A \\ \mathbf{F}_\psi \\ \mathbf{0} \end{bmatrix}, \quad (11)$$

where

$$\begin{aligned} \mathbf{K}_{A\lambda} &= \int_{\Gamma} N^T \mathbf{n} N_\lambda d\Gamma, & \mathbf{K}_{\psi\lambda} &= \int_{\Gamma} \mathbf{B}_\psi^T \mathbf{n} N_\lambda d\Gamma, \\ \mathbf{K}_{\lambda A} &= \int_{\Gamma} N_\lambda^T \mathbf{n}^T N d\Gamma, & \mathbf{K}_{\lambda\psi} &= \int_{\Gamma} N_\lambda^T \mathbf{n}^T \mathbf{B}_\psi d\Gamma, \\ \mathbf{K}_{\lambda\lambda} &= \mathbf{0}, & \mathbf{F}_A &= -i\omega \left(\int_{\Gamma_h} N^T \bar{\mathbf{H}} d\Gamma + \int_{\Omega} N^T \mathbf{j} d\Omega \right), & \mathbf{F}_\psi &= \mathbf{0}. \end{aligned}$$

3.5 Implementation of Symmetric Boundary Condition Using a Thin Patch (Method-2)

A cylindrical patch of very small thickness t is added to the half annular sphere from radius a (radius of the conducting sphere) to radius R_∞ (radius of the truncation boundary), as shown in Fig. 2. The thickness t of the patch should be very small as compared to the other dimensions. In this method, the symmetric boundary condition ($\mathbf{E} \cdot \mathbf{n} = 0$) is applied in two parts. ($\mathbf{A} \cdot \mathbf{n} = 0$) is imposed explicitly by keeping the normal component of \mathbf{A} , i.e. A_y to be zero for all the nodes on the outer blue surface (Fig. 2) of the cylindrical patch. Also, we impose $\psi = 0$ at all the nodes inside the volume of the cylindrical patch to satisfy $\nabla \psi \cdot \mathbf{n} = 0$ on the symmetry face. As the

Fig. 2 Sphere with a thin cylindrical patch with thickness, t

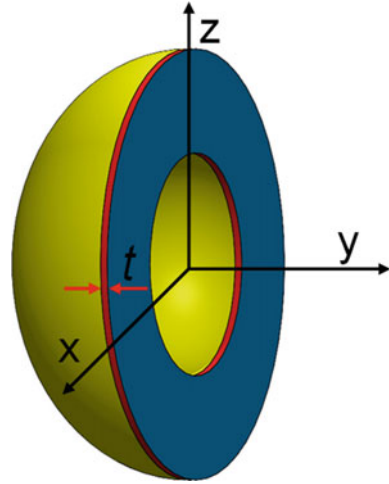


Table 1 Mesh and equation details for Method-1, Method-2 and full domain

	Method-1	Method-2	Full sphere [14]
Mesh size ($r \times \theta \times \phi$)	$8 \times 6 \times 6$	$8 \times 6 \times 6 +$ $8 \times 12 \times 1$	$8 \times 6 \times 12$
Number of equations	9585	11,697	16,698

patch is very thin, it will satisfy $\mathbf{E} \cdot \mathbf{n} = 0$ on the symmetry face without disturbing the scattered fields.

4 Numerical Example

We have demonstrated the performance and effectiveness of the Method-1 and Method-2 solving the problem of electromagnetic scattering by a conducting sphere. One electromagnetic plane wave $\mathbf{E}_{\text{inc}} = E_0 e^{-ikz} \mathbf{e}_x$ is incident on one conducting sphere of radius a . We truncate our numerical domain at radius R_∞ ; on this truncation boundary, we apply Sommerfeld absorbing boundary condition. Hence, we find the scattered electric field in the hollow spherical domain of inner and outer radii of a and R_∞ , respectively. On inner spherical surface at radius a , perfectly conducting boundary condition is applied as described in [14].

As discussed in Sect. 2, the problem is symmetric about xz plane. Hence, we consider half of the hollow sphere as our computational domain. On the symmetry face (xz), we apply suitable boundary condition as discussed in Sects. 3.4 and 3.5. The half annular spherical domain is meshed with $8 \times 6 \times 6$ ($r \times \theta \times \phi$) elements of 27-node hexahedral (b27) and 18-node wedge (w18) elements.

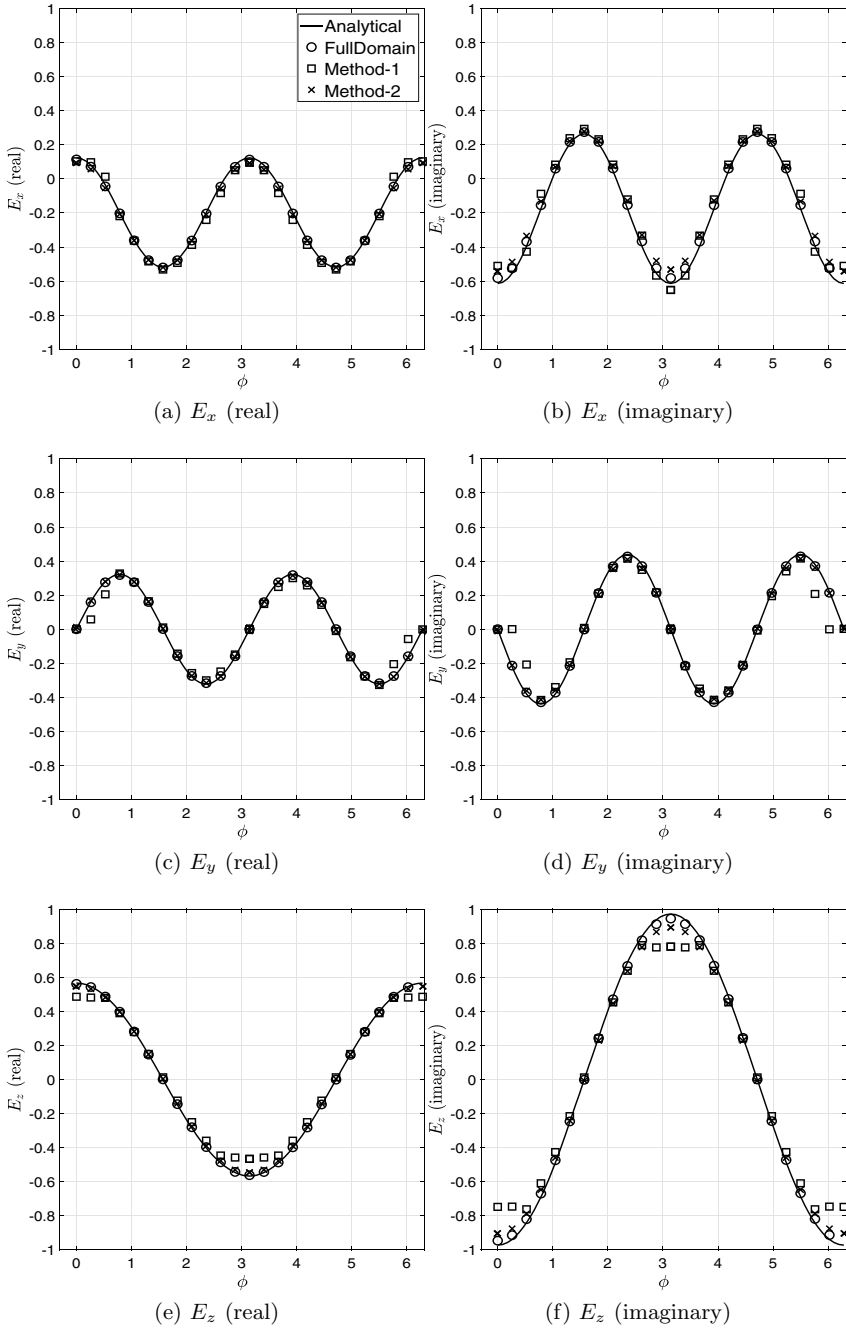


Fig. 3 Near field variation in the electric field along ϕ for $k_0r = 1.25, \theta = \pi/4, k_0a = 1, k_0R_\infty = 5$ for the scattering from a conducting sphere

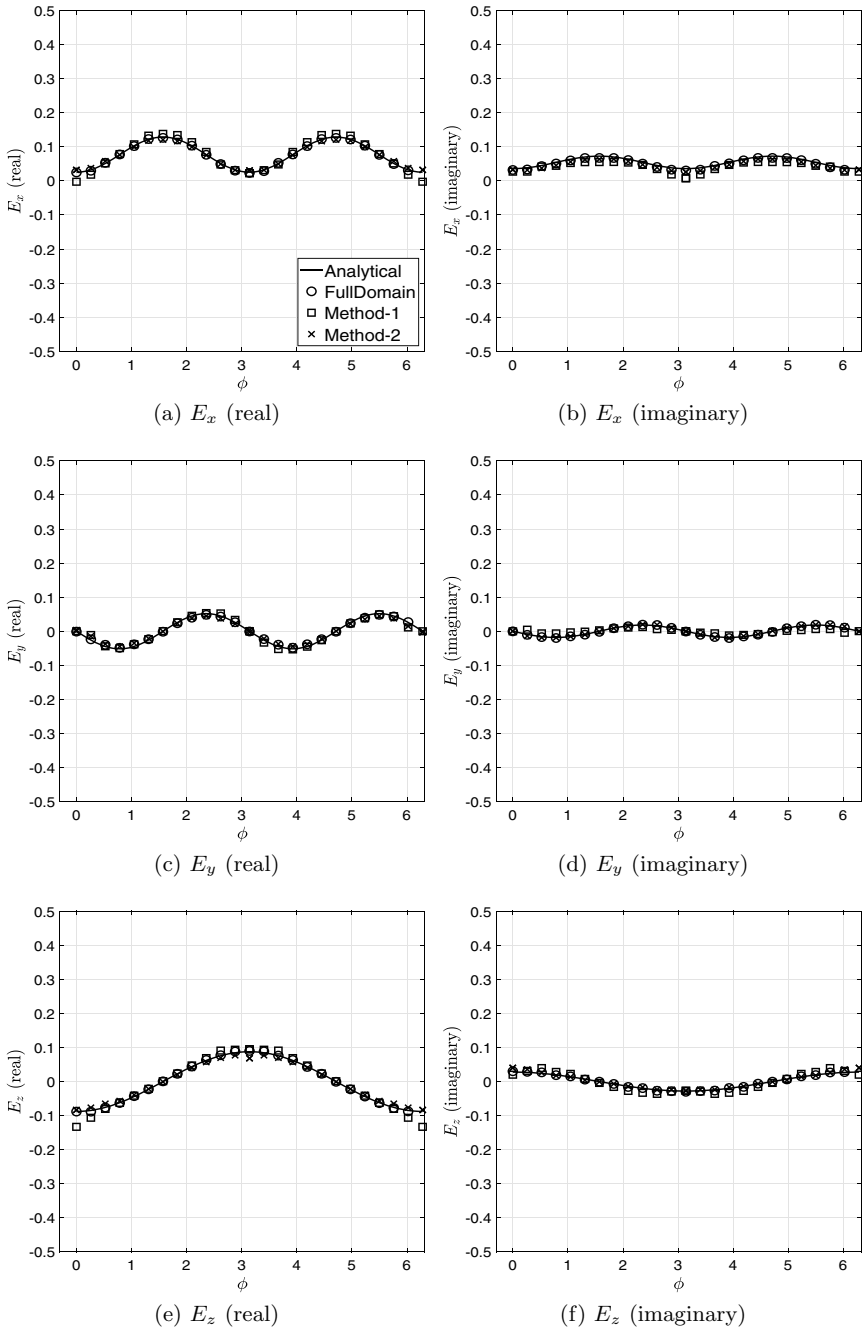


Fig. 4 Far field variation in the electric field along ϕ for $k_0r = 4.75, \theta = \pi/4, k_0a = 1, k_0R_\infty = 5$ for the scattering from a conducting sphere

In Method-1, on the symmetry face $\mathbf{E} \cdot \mathbf{n} = \mathbf{0}$ is implemented as described in Sect. 3.4. Then we have to solve 9585 equations. In Method-2, additional cylindrical patch has $8 \times 12 \times 1$ ($r \times \theta \times t$) elements of 27-node hexahedron (b27). Here, as discussed in Sect. 3.5, we have to solve 11,697 equations. Table 1 summarizes the mesh details in both methods along with the full domain.

We have taken $E_0 = 1$, $k_0 a = 1$, $k_0 R_\infty = 5$. The scattered electric field components are compared along ϕ at two different locations: (a) at near field, i.e. at $k_0 r = 1.25$ and (b) at far field with $k_0 r = 4.75$) for $\theta = \frac{\pi}{4}$. Figure 3 and Fig. 4 represent real and imaginary components of all three components of the scattered field \mathbf{E} at near field and far field respectively.

In all the plots, we have compared the results of Method-1 and Method-2 with the analytical benchmarks [14], and the result of the full domain. For both the Method-1 and Method-2, we directly solve for $0 \leq \phi \leq \pi$, then because of symmetry of the problem, we have $\mathbf{E}(2\pi - \phi) = \mathbf{E}(\phi)$ for the remaining half domain.

We can notice that the results from both Method-1 and Method-2 are closely matching with the analytical benchmarks. But the total numbers of equations are 42% and 30% less for Method-1 and Method-2, respectively (see Table 1) as compared to the full domain. As computational cost is directly co-related with the total no. of equations; hence, both the methods are computationally very efficient.

5 Conclusions

The present work imposes symmetric boundary condition on the symmetry face in two different approaches while simulating electromagnetic harmonic analysis. In one approach (Method-1), symmetric boundary condition is applied using Lagrange multiplier method on the symmetry surface. In the other method (Method-2), a very thin cylindrical patch is added on the symmetry surface. The symmetric boundary conditions are imposed explicitly with proper conditions on the potentials in this additional layer. Here, no additional computational treatment is required. Both methods have considerably less computational cost than the full domain, but Method-2 involves more number of equations than Method-1 for same level of meshing. But Method-2 produces more accurate result than the Method-1. For the future work, the same methods can be extended in transient electromagnetic analysis.

Acknowledgements The authors gratefully acknowledge the support from SERB, DST under the project IMP/2019/000276, and VSSC, ISRO through MoU No.: ISRO:2020:MOU:NO: 480.

References

1. Jin JM (2015) The finite element method in electromagnetics. Wiley (2015)

2. Volakis JL, Volakis JL, Chatterjee A, Kempel LC (1998) Finite element method for electromagnetics. Universities Press
3. Nandy AK (2016) Robust finite element strategies for structures, acoustics, electromagnetics and magneto-hydrodynamics. Ph.D. thesis, Indian Institute of Science Bangalore, Department of Mechanical Engineering, IISc, Bangalore, India. <https://etd.iisc.ac.in/handle/2005/2913>
4. Jog CS, Nandy A (2014) Mixed finite elements for electromagnetic analysis. *Comput Math Appl* 68(8):887–902
5. Boffi D, Farina M, Gastaldi L (2001) On the approximation of maxwell’s eigenproblem in general 2d domains. *Comput Struct* 79:1089–1096
6. Boffi D, Fernandes P, Gastaldi L, Perugia I (2001) Computational models of electromagnetic resonators: analysis of edge element approximation. *SIAM J Numer Anal* 36(4):1264–1290
7. Boffi D, Brezzi F, Fortin M (2013) Mixed finite element methods and applications. Springer
8. Kamireddy D, Nandy A (2020) Combination of triangular and quadrilateral edge element for the eigenvalue analysis of electromagnetic wave propagation. *Eur J Molecular Clin Med* 7(11):1656–1663 (2020). https://ejmcm.com/article_5696.html
9. Kamireddy D, Nandy A (2021) Creating edge element from four node quadrilateral element. *IOP Conf Ser: Mater Sci Eng* 1080(1):012015. <https://doi.org/10.1088/1757-899x/1080/1/012015>
10. Kamireddy D, Nandy A (2022) A novel conversion technique from nodal to edge finite element data structure for electromagnetic analysis. *Comput Ass Methods Eng Sci* 28(4):291–319
11. Kamireddy D, Chavan SM, Nandy A (2022) Comparative performance of novel nodal-to-edge finite elements over conventional nodal element for electromagnetic analysis. Preprint available at <https://arxiv.org/abs/2203.14522>
12. Kamireddy D, Chavan SM, Nandy A (2022) Electromagnetic eigen analysis: performance comparison of four node and four edge quadrilateral elements with the effect of distortion. *Mater Today: Proc.* <https://www.sciencedirect.com/science/article/pii/S2214785322037506>
13. Stupfel B (1997) Numerical implementation of second-and third-order conformal absorbing boundary conditions for the vector-wave equation. *IEEE Trans Antennas Propag* 45(3):487–492
14. Nandy A, Jog CS (2016) An amplitude finite element formulation for electromagnetic radiation and scattering. *Comput Math Appl* 71(7):1364–1391
15. Joly P, Mercier B (1989) A new second order absorbing boundary condition for Maxwell’s equations in dimension 3. *Institut national de recherche en informatique et en automatique* (1989)
16. Assous F, Sonnendrücker E (2010) Joly-mercier boundary condition for the finite element solution of 3d maxwell equations. *Math Comp Modell* 51(7–8):935–943
17. Nandy A, Jog CS (2018) Conservation properties of the trapezoidal rule for linear transient electromagnetics. *J Adv Math Comput Sci* 26(4):1–26 (2018). <http://www.sciencedomain.org/abstract/23334>
18. Chatterjee A, Jin JM, Volakis JL (1993) Edge-based finite elements and vector ABCS applied to 3-d scattering. *IEEE Trans Antennas Prop* 41(2):221–226
19. Sertel K, Volakis JL (2004) Multilevel fast multipole method solution of volume integral equations using parametric geometry modeling. *IEEE Trans Antennas Propag* 52(7):1686–1692
20. García-Castillo LE, Gómez-Revuelto I, de Adana FS, Salazar-Palma M (2005) A finite element method for the analysis of radiation and scattering of electromagnetic waves on complex environments. *Comput Methods Appl Mech Eng* 194(2–5):637–655
21. Boyse WE, Seidl AA (1994) A hybrid finite element method for 3-d scattering using nodal and edge elements. *IEEE Trans Antennas Propag* 42(10):1436–1442
22. Tang J, Paulsen KD, Haider SA (1998) Perfectly matched layer mesh terminations for nodal-based finite-element methods in electromagnetic scattering. *IEEE Trans Antennas Propag* 46(4):507–516
23. Bardi I, Biro O, Preis K (1991) Finite element scheme for 3d cavities without spurious modes. *IEEE Trans Magn* 27(5):4036–4039

24. Bardi I, Biro O, Dyczij-Edlinger R, Preis K, Richter KR (1994) On the treatment of sharp corners in the fem analysis of high frequency problems. *IEEE Trans Magn* 30(5):3108–3111
25. Costabel M, Dauge M (2002) Weighted regularization of Maxwell equations in polyhedral domains. *Numer Math* 93(2):239–277
26. Otin R (2010) Regularized Maxwell equations and nodal finite elements for electromagnetic field computations. *Electromagnetics* 30:190–204
27. Otin R (2013) Ermes: a nodal-based finite element code for electromagnetic simulations in frequency domain. *Comput Phys Commun* 184(11):2588–2595
28. Stratton JA (2007) *Electromagnetic theory*, vol 33. Wiley
29. Griffiths DJ (1999) *Introduction to electrodynamics*, 2nd edn. Prentice Hall, New Jersey

Performance Evaluation of Hardware Trojan Using FPGA



Ravikant Khamitkar  and R. R. Dube 

Abstract In this technologically advancing world, different things are connected together using various integrated systems to collect the data for processing and wise decision-making. These systems are developed using various hardware and software components using different VLSI design and technology with focus on quality of integrated circuits (ICs). There exists security concern during design and development of these ICs in the form of trigger-based hardware that can lead to malicious behavior or destruction of system. This malevolent hardware in ICs is known as Hardware Trojan (HT). Various researchers throughout the world have proposed different methods to detect the Hardware Trojans and to mitigate their effect. There are majorly two perspectives considered for HT detection: one is using layout images examination with physical constraint validation of ICs and the other is during design phase of ICs using hardware description language and assertion checkers to verify and validate the design. In this paper, the implementation of HT in FPGA based System on Chip is carried out. The hardware Trojan is implemented in three different systems and their comparative analysis is carried using the evaluation parameters. These various approaches used to implement are studied in this paper therein labeling the security threats and requirements to deliver the commendation for future research in this field.

Keywords Hardware Trojan · SoC · FPGA

1 Introduction

In development of ICs during last decade, there exists a security concern of malicious modification in hardware, which is also known as Hardware Trojan (HT). This has become the major security concern in the IC development industry. Integrated Circuits having HTs can cause leakage of data, failure in work or cause other devastating consequences. Due to this concern, HT has been a problem of concern and

R. Khamitkar (✉) · R. R. Dube
Walchand Institute of Technology, Solapur, Maharashtra 413006, India
e-mail: khamitkar.ravikant@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
V. V. S. S. Chakravarthy et al. (eds.), *Advances in Signal Processing, Embedded Systems and IoT*, Lecture Notes in Electrical Engineering 992,
https://doi.org/10.1007/978-981-19-8865-3_11

127

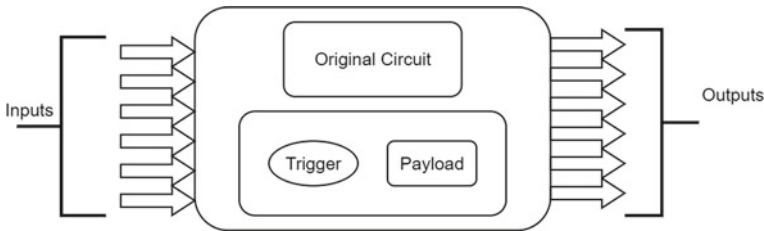


Fig. 1 Hardware Trojan circuit

the various researchers from industry academic are working on detection and mitigating of it [1–4]. Figure 1 shows the basic Hardware Trojan diagram which has the triggering mechanism and when it receives the trigger, the payload is given as output. This payload can be secret key/data, change in sequence of execution order or deadlock condition, etc.

The HT insertion can be done at user logic (glue logic) level, control level or System on Chip level [5]. The first with glue logic level insertion can be achieved easily using simple implementation of user logic in hardware description language and the HT can be implemented in the same hardware description language. The insertion and detection are easy in user logic-based implementation. The control level or bus level hardware Trojan implementation can be achieved using circuit implemented using HDL but the triggered output of payload can be modification in bus operation or control signal operations. Using various standard comparison of output, it is possible to detect the presence of HT in the IC. The SoC level implementation is little complex to implement and more complex to detect. The insertion of covert hardware can be done using hardware description language but its triggering mechanism can be controlled through either hardware input or software-based instruction execution.

To insert hardware Trojan in IC, it is important to choose the proper integrated circuits to masquerade the malicious hardware. For the study of this HT insertion, the Advanced Encryption Standard is considered as basic hardware in integrated circuit. It is implemented using hardware description language and Xilinx Vivado tool Set; the hardware Field Programmable Gate Array (FPGA) based kit of Xilinx Arty-S7-50 is used for implementation of the same. The various implementation parameters are compared to get the insights of HT insertion into the system.

2 Literature Survey

There is lot of research still going on for the hardware Trojan insertion and detection in the IC. Out of which some of the recent and relevant resources and contributions are discussed in this part.

Table 1 Various HT implementations

Work	Base circuit	Hardware used	Triggering mechanism and payload
[6]	RC4	FPGA	Sequence of bit-stream and secure key payload
[7]	ARM7	ASIC	External pin and register contents
[8]	Floating point multiplier	FPGA	Input bits and Specific result
[9]	General encryption	FPGA	Sequence of bits and specific result
[10]	Ring oscillator	FPGA	Ring oscillator sequence number and specific output signal

Fotios Kounelis et al. have implemented the hardware Trojan insertion in RC4 algorithm and checked the result of implementation using parameters like device utilization, operating frequency, throughput, etc. [6]

Yumin Hou et al. have implanted a framework for the Analog HT detection using microprocessor trustworthiness by design and fabricating the ARM7 based system on chip and manufacturing the IC of it with HT [7].

Nikhila S et al. have implemented floating point multiplier and its hardware Trojan models using FPGA implementation and checked the result of implementation using hardware kit programming [8].

Devu Manikantan Shila et al. have implemented HT in FPGA device and discussed about the detection mechanisms by security and threat analysis of general encryption model [9].

Kaige Qu et al. have implemented HT in circuit having ring oscillator and the network-based hardware Trojan which is used to disrupt the operation of oscillator and give the malicious output [10].

Here, from Table 1, it can be seen that the different methods and use case are used for studying the impact of HTs.

3 Methodology

The Crypto cores can be implemented using various Hardware description languages like VHDL or Verilog. These crypto cores are programmed inside the FPGA and tests are conducted for detection of HTs. The cores can be designed and implemented directly using HDL (Fig. 2).

The implementation have crypto cores with and without HT and the performance evaluation is carried out on the basis of various parameter like area, power, operating frequency, etc. [11, 12].

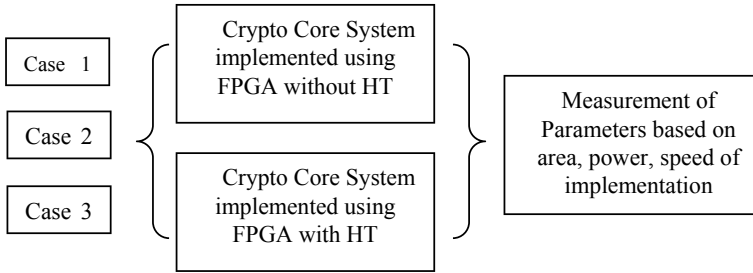


Fig. 2 HT detection in simple crypto core

4 System Design

4.1 Hardware Design with HT

The integration of Hardware Trojan in the system can be done at three levels as discussed earlier.

1. Glue Logic level
2. Control level
3. SoC level

Designing Glue Logic based HT insertion and detection is easy to implement, hence it is not considered in the scope of this paper.

The Control level HT system is designed by integrating the crypto core with Hardware Trojan. Basic crypto core is implemented for security of data built around Advanced Encryption Standard (AES). This implemented system takes input plain text and gives out cipher text based on the secure key stored inside the system. This is implemented using FPGA. The hardware Trojan is integrated in this system which when triggered using external trigger input gives out the secure key as payload output (Fig. 3).

The system is implemented using VHDL programming language in which each of the part inside system is written as independent module and it is integrated in single file. The external pin can be used as trigger or input plain text with particular sequence can also take as trigger for the circuit. Once circuit receives trigger, the hardware Trojan gets activated and gives out secure key as payload. In normal operations, the cipher text will be the output and when trigger is received by HT, the output will be secure key.

The SoC level HT system is designed by integrating the crypto core as IP and part of Microblaze based System on Chip inside Xilinx FPGA. The crypto core communicates with Microblaze soft core and gives out the normal output as per the input. If input is plain text then it gives out the cipher text, and if input is cipher text then it decrypts it and gives out the plain text as output. When the external trigger is received, then it gives out the cipher key or secure key as output instead of cipher text

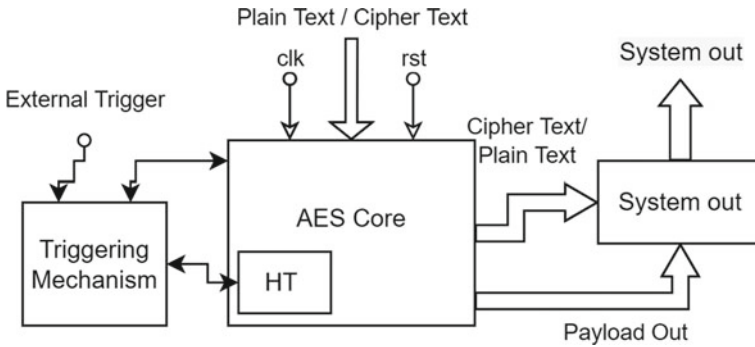


Fig. 3 Crypto core and hardware Trojan in control level integration

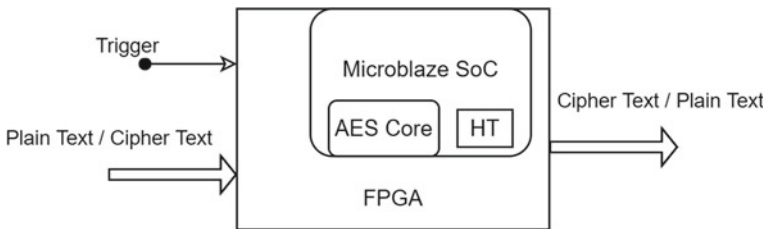


Fig. 4 Crypto core and hardware Trojan in SoC level integration

or plain text. This HT can also get triggered with special sequence detection from plain or cipher text input and gives out the secure key as payload.

From Fig. 4, it can be seen that the SoC based Hardware Trojan integration is done for crypto core. The system is provided the external trigger as well as it can be triggered using special sequence from Plain text input as well.

4.2 Software Design with HT

The system designed with Control level HT integration needs only hardware description language for implementation and the results of simulation and hardware implementation are discussed in the next sessions.

The system designed with SoC level HT integration need Hardware description language for hardware core and the software part to work as program inside the embedded hardware is written using Embedded C language. The SoC level and Control level HT have the following common software flow of working with Hardware Trojan (Fig. 5).

From the flow diagram, it can be seen that the hardware Trojan gets triggered either by external trigger input or special pattern in the plain text input.

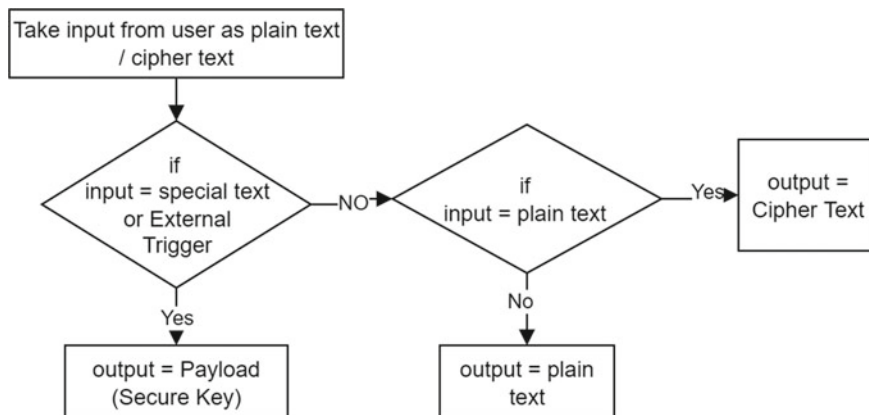


Fig. 5 Software flow of hardware Trojan in SoC level integration

5 Results and Discussion

The Hardware Trojan insertion in the Encryption core is evaluated on the basis of device utilization, operating frequency, Power dissipation.

From Table 2, it can be seen that the device utilization increases with insertion of hardware Trojan. The number of Look up Tables increased from 10 to 22 in number, the Flip flops increased from 31 to 59, number of slices from 16 to 30, so overall it can be seen that the device hardware utilization is increased.

From Table 3, it can be seen that the static power dissipation remains the same in hardware Trojan insertion, the Logic power increases from 0.217 to 0.476 W. Also it can be seen that the signals' power consumption is increased from 0.672 to 1.103 W. From this, it can be concluded that the power consumption increases when the hardware Trojans are inserted in the circuits.

In Fig. 6, the output waveform showing the payload delivery with reception of trigger signal is shown. It can be seen that at 400 ns, the trigger signal is made high and then the output final (32 bits) and payload (32 bits) give out the cipher (secret) 128 bits key as output in 4 words of 32 bits each.

Table 2 Device utilization in hardware Trojan insertion

Resource	Without HT		With HT	
	Utilization	Available	Utilization	Available
LUT	10	32,600	22	32,600
FF	31	65,200	59	65,200
Slices	16	16,300	30	16,300
IO	33	210	66	210
BUFG	1	32	1	32

Table 3 Power utilization in hardware Trojan insertion

	Without HT	With HT
Type	Power (W)	Power (W)
PL static	0.485	0.485
Logic	0.217	0.476
Signals	0.672	1.103



Fig. 6 Waveform of payload activating

6 HT Challenges and Conclusion

From different work review, it can be seen that the HT detection need very well ordered approach for security of Hardware. Most of the previous researches were based on image processing of implemented system comparing it with golden ICs. This approach is complex as refabricating the IC is both time-consuming and costly affair. The other approach is based on the design stage to compare the electrical parameters of standard IC developed and the one manufactured through third-party developers. This parameter comparison can produce good result in HT detection as seen in this paper. The further this approach can be taken ahead for HT detection in SoC based systems, also some Machine Learning approaches can be incorporated to give better performance in HT detection.

References

1. Bhunia S, Hsiao MS, Banga M, Narasimhan S (2014) Hardware Trojan attacks: threat analysis and countermeasures. Proc IEEE 102(8):1229–1247. <https://doi.org/10.1109/JPROC.2014.2334493>
2. Tehranipoor M, Koushanfar F (2010) A survey of hardware Trojan taxonomy and detection. IEEE Des Test Comput 27(1):10–25. <https://doi.org/10.1109/MDT.2010.7>
3. Chakraborty RS, Narasimhan S, Bhunia S (2009) Hardware Trojan: threats and emerging solutions. In: 2009 IEEE international high level design validation and test workshop, San Francisco, CA, 2009, pp 166–171. <https://doi.org/10.1109/HLDVT.2009.5340158>
4. Wu TF, Ganesan K, Hu YA, Wong H-SP, Wong S, Mitra S (2016) TPAD: hardware Trojan prevention and detection for trusted integrated circuits. IEEE Trans Comput Aided Des Integr Circ Syst 35(4):521–534. <https://doi.org/10.1109/TCAD.2015.2474373>
5. Khamitkar R, Dube RR (2022) A survey on using machine learning to counter hardware Trojan challenges. In: Senjyu T, Mahalle PN, Perumal T, Joshi A (eds) ICT with intelligent applications. Smart innovation, systems and technologies, vol 248. Springer, Singapore. https://doi.org/10.1007/978-981-16-4177-0_53

6. Kounelis F, Sklavos N, Kitsos P (2017) Run-time effect by inserting hardware Trojans, in combinational circuits. *Euromicro Conf Dig Syst Des (DSD) 2017*:287–290. <https://doi.org/10.1109/DSD.2017.27>
7. Hou Y, He H, Shamsi K, Jin Y, Wu D, Wu H (2019) On-chip analog Trojan detection framework for microprocessor trust worthiness. *IEEE Trans Comput Aided Des Integr Circuits Syst* 38(10):1820–1830. <https://doi.org/10.1109/TCAD.2018.2864246>
8. Nikhila S, Yamuna B, Balasubramanian K, Mishra D (2019) FPGA based implementation of a floating point multiplier and its hardware Trojan models. In: 2019 IEEE 16th India council international conference (INDICON), pp 1–4. <https://doi.org/10.1109/INDICON47234.2019.9030341>
9. Shila DM, Venugopal V (2014) Design, implementation and security analysis of hardware Trojan threats in FPGA. *IEEE Int Conf Commun (ICC) 2014*:719–724. <https://doi.org/10.1109/ICC.2014.6883404>
10. Qu K, Wu L, Zhang X (2015) A novel detection algorithm for ring oscillator network based hardware Trojan detection with tactful FPGA implementation. In: 2015 11th international conference on computational intelligence and security (CIS), pp 299–302. <https://doi.org/10.1109/CIS.2015.80>
11. Huang Z, Wang Q, Chen Y, Jiang X (2020) A Survey on machine learning against hardware Trojan attacks: recent advances and challenges. *IEEE Access* 8:10796–10826. <https://doi.org/10.1109/ACCESS.2020.2965016>
12. Xue M, Gu C, Liu W, Yu S, O'Neill M (2020) Ten years of hardware Trojans: a survey from the attacker's perspective. *IET Compute Dig Techniques* 2020

Emotion Recognition in Tweets Using Optimized SVM and KNN Classifiers



D. N. S. B. Kavitha, P. V. G. D. Prasad Reddy, and K. Venkata Rao

Abstract Social media is reflected as a rich source of information because it is widely used by people to share their emotions, views, and opinions. Therefore, this information should be mined qualitatively for the purpose of Emotion Recognition (ER). Various other applications associated with ER are public mood detection, identifying psychological (depression/anxiety) disorders, online product sales prediction, etc. This work exploits optimized SVM (Support Vector Machine) and K-Nearest Neighbor (KNN) classifiers for recognition of emotions in textual Twitter messages. The multiple feature extraction model combined with optimal Binary Grasshopper Optimization Algorithm (BGOA)-based feature selection (FS) enhances the computation process. The meta-heuristic algorithm-based FS improves the classification accuracy with minimal computation time and speed up the classifier for recognizing the multiclass emotions. The performance measures are done at various training rates. The simulations depict the superiority of KNN and SVM classifiers over traditional Naive Bayes approach.

Keywords Emotion recognition · Text analysis · Feature selection · BGOA · Bayesian optimization

1 Introduction

Human ER (Emotion Recognition) plays a significant role to build the interactive relationship. Emotions are the reflections of hand, body gestures, speech, and face expressions. The process of understanding and extracting the emotions shows greater importance between the machine and human interaction. Recognizing the emotions can induce us to take specific actions and influence the decision. [1]. SA recognizes the sentiment represented in a text, whereas Opinion Mining collects and examines

D. N. S. B. Kavitha (✉) · P. V. G. D. Prasad Reddy · K. Venkata Rao
Department of Computer Science and Systems Engineering, Andhra University, Visakhapatnam, India
e-mail: kavi.moki@gmail.com

people's views about a specific thing. SA is classified into three levels: file, paragraph, and aspect levels. Paper-level SA aims to classify an opinion text as a positive or negative emotion or opinion. It considers the entire document to be a fundamental information unit. Sentence-level SA tries to categorize the emotion conveyed in each phrase. The first step is to figure out whether the phrase is objective or subjective. In this case, sentence-level SA will look at whether the sentence is positive or negative.

According to Wilson et al. [2], emotions are not inherently subjective. However, because sentences are essentially short text, there is no fundamental gap between document-level and sentence-level categorization [3]. Text categorization just at paragraph or sentence level will not provide the necessary depth for opinions on all parts of the entity, which would be required for a number of applications. Therefore, we must proceed to the aspect level to acquire these specifics. Aspect-level SA aims to classify sentiment in reference to particular characteristics of entities. The first step is to identify the entities and their properties. Opinions may differ on separate characteristics of the same object, such as this statement: "This phone's voice quality is terrible, but its battery life is long." This survey focuses on the initial two categories of SA.

Reviews are the most important sources of information. These reviews are significant to entrepreneurs because they can base their strategic decisions on a study of user views about their items. The vast majority of the feedback comes from review sites. SA can be employed not only in consumer reviews but also in stock markets [4, 5], news articles [6, 7], or political debates. In political arguments, for example, we could learn about people's attitudes toward specific presidential nominees or political parties. Political ideologies can also be used to foretell election outcomes.

The rest of the article is arranged as follows: The second section presents the literature review. The third section describes the methodology. In Sect. 4, simulations and discussions are provided. The paper concluded with Sect. 5.

2 Literature Survey

Social media and blogging websites are considered valuable sources of information since users openly share and discuss their thoughts on a specific issue. They are also employed as data sources in the SA process. Many uses and enhancements to SA algorithms have been proposed recently. Some of them are as follows.

Kumar et al. [8] presented a new structure to recognize the Twitter emotions. This framework is the integration of data preprocessing, feature extraction (BOW, Unigram + POS, Unigram), and classification (Multinomial NB). Only textual contents were used to analyze the Tumblelogs (microblogging) emotions. The best text class was attained using 4-way classification model with 82.25% (unigram) accuracy. The data was gathered from sentiment140. However, the proposed multinomial NB classifier with 5-way classification model achieved lower accuracy of 69%.

Yang et al. [9] introduced a new ensemble approach called Dynamic weighted Attention (DA) with multi-channel CNN (DACNN) for recognizing the emotions in

social media. The combination of attention and multi-channel CNN was utilized to adjust the weights automatically and efficiently enhance the recognition outcome. The higher-quality FVs (feature vectors) were obtained with the XLNet technology. Three different datasets namely Twitter Emotion Corpus (TEC), SemEval (SE)-2018, and Cleaned Balanced Emotional Tweet (CBET) were used for the experimentation. Thus, the proposed DACNN model gained better scores than the standard models.

Stojanovski et al. [10] developed a DL model called Deep Neural Network (DNN) for identifying the emotions and sentiments in tweets. The textual features were extracted using CNN and the classifier categorizes the emotions (Fear, Sadness, Anger, Joy, Surprise, Love, and Thankfulness) and sentiments (positive, neutral, negative). SE dataset was used for sentiment evaluation and tweets related with FIFA 2014 world cup was used to identify the emotions. Hence, the accuracy (58.84%) was achieved with the DNN model in recognizing the emotions.

Chatterjee et al. [11] proposed a novel DL-based method SS-BED (Sentiment Semantic (SS)-Based Emotion Detector) for the detection of various emotions like Sad, Happy, and Angry in the textual dialogues. The DL approach has been embedded with SVM, NB, and DT classifiers to detect various emotions efficiently. The sentiment and semantic-based representations are combined in this research to obtain more accuracy in the detection of emotions. The model was trained by a huge data with various ways of stating feelings and these emotions were collected using the semi-automated approaches. Real-world dialogue datasets were used in this research for the emotion detection. Table 1 shows the extensive literature review.

Many of the supervised learning algorithms achieved an accuracy of only 70–80% even at 80% training rate. Many researchers have done the analysis with single dataset and at a constant training rate of 80%, and it is identified there is lack of information about the consistency of classifier with less training rate (50–60%). It is also observed that many approaches are facing the overfitting problem because of redundant features.

Based on these motivations, this research work is carried out on investigations into the sentiment classification of political tweets about an individual on Twitter using SVM and KNN approaches under different training rates.

3 Methodology

Figure 1 depicts the framework of the proposed emotion classifier. The procedure is divided into two stages. Phase I is concerned with data preparation, feature extraction, and BGOA-based feature selection.

Preprocessing comprises operations such as removing numerals, stemming words, labeling parts of speech, removing commas and punctuation marks, and noise removal. Following preprocessing, there is feature extraction, which includes extraction of bag-of-words (BOW), term-frequency (TF), and term-frequency-inverse document frequency (TF-IDF). After feature extraction, the BGOA is used to choose features. After feature extraction, the BGOA is used to choose features.

Table 1 Review of SA techniques

Ref. No.	Classifier	Research direction	Performance	Advantages	Disadvantages
[12]	NB and semantic DT (decision tree) classifier	Sentimental analysis (SA)-Amazon reviews	Acc (for 3 categories) = 71.7%	Solves the general problem which arise among negative and positive classification	It can only be utilized for data assessments written in English
[13]	NB, SVM	SA on political Twitter data	SVM (precision = 0.86, recall = 0.84, <i>F</i> -measure = 0.87) NB (precision = 0.925, recall = 0.921, <i>F</i> -measure = 0.92)	Arabic language is used in this research for SA	Difficult to attain higher accuracy of classification using Arabic language as it is not a case-sensitive language
[14]	Automatic SVM	Automatic analyzing of electoral tweets	Acc = 56.84%	Aids in comprehending how public opinion is formed, tracking public opinion and polarization	Fail to differentiate among several different kinds of purpose
[15]	Logistic regression (LR) algorithm	Twitter SA for stock market movements (SMM) prediction	Acc = 69.09%	Emergence of SA can judge the type of sentiment present in the tweet	Execution process takes much time
[16]	Modified <i>N</i> -gram	Identify top companies in a particular industry	Acc = 82%	Used to improve MI (marketing intelligence)	-

In Phase II, KNN and SVM classifiers are implemented for SA. The experiments are carried out at various training rates, with various datasets, and with different SVM kernel functions and with different distance functions in KNN.

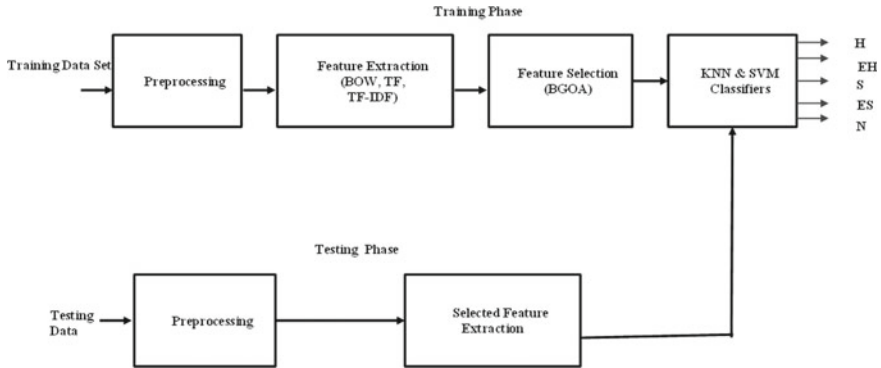


Fig. 1 Framework of SA

Table 2 KNN classifiers

Type of KNN	Distance function	Distance(d)
Fine KNN	Euclidean distance	$\sqrt{\sum_{i=1}^n (p_i - q_i)^2}$
Cubic KNN	Cubic distance/Minkowski Dist. (r = 3)	$(\sum_{i=1}^m p_i - q_i ^2)^{\frac{1}{r}}$

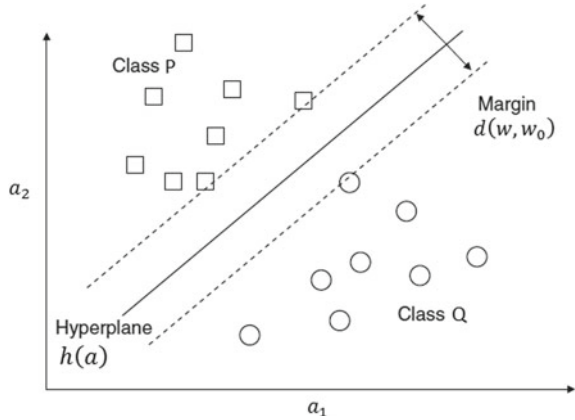
3.1 KNN Classifiers

KNN classifier stores all existing cases and categorizes new cases using a similarity metric. It calculates the distance among a data record and its references. It examines the k-closest data records, with the class label in the group serving as the anticipated class. In this paper, we have implemented two optimized KNN classifiers, namely Fine KNN and Cubic KNN [17–19]. For optimization of the model, hyperparameters such as distance metric, number of neighbors, distance weight, and standardize data are considered. Bayesian optimization approach is used for optimizing the model. Table 2 represents the proposed KNN classifiers.

3.2 SVM Classifiers

The SVM classifier has the ability in classifying high dimensional and noisy data. It is a supervised method that uses a portion of training instances to classify the data. With the use of training data, the SVM classifier constructs a feature space. Then, it helps to determine a hyperplane that divides the plane into two parts, where each half contains only one class [20, 21]. To construct the hyperplane SVM follows two

Fig. 2 Principle of linear SVM



principles, and they are selecting the best hyperplane for classification, and it should be the maximum distance between two supporting planes shown in Fig. 2.

Linear SVM Classifier:

In linear SVM, a linear kernel is used to discriminate between classes and is provided as,

$$F(a, w) = a^T w$$

Here a represents the raw feature vector and w represents the weight vector.

The linear hyperplane with a constant w_0 is

$$h(a) = a^T w + w_0$$

The decision taken by SVM is defined as

$$M = \begin{cases} M(P), & \text{if } h(a) \geq 0 \\ M(Q), & \text{if } h(a) < 0 \end{cases}$$

If P and Q are two classes, the class is identified as a member (M) of P if and only if $a^T w + w_0 \geq 0$, otherwise, it is recognized as Q .

Non-linear SVM Classifier

In this work, non-linear kernels such as polynomial and Gaussian are used for classification of two modulation classes, and they are given as

$$K(P, Q) = (\gamma \cdot P^T Q + r)^d, \gamma > 0,$$

where d is the degree of polynomial and it represents the polynomial kernel

$$K(P, Q) = \exp(-\|P - Q\|^2 / 2\sigma^2)$$

here σ depends on the number of features and it represents the Gaussian kernel.

For optimization of the model, hyperparameters such as standardize data, kernel function, kernel scale, multiclass method, and box constraint level are considered. Bayesian optimization approach is used for optimizing the model. The performance analysis of these optimized SVM and KNN classifiers is discussed in Sect. 4.

4 Results and Discussions

The experiments in this paper were carried out using Twitter datasets from the 2016 and 2020 US Presidential Elections. The US 2016 presidential election Twitter dataset contains 66,075 tweets classified as Extremely Happy, Extremely Sad, Happy, Sad, and Neutral Emotions. The Twitter dataset for the US 2020 election contains 365,000 tweets. Table 3 represents the dataset information of different classes.

Tables 4 and 5 represent the summary of different performance matrices of Fine KNN and Cubic KNN classifiers at different learning rates. From these, it is depicted that the accuracy of SA using KNN classifiers is varying between 83.8 and 89.4%. Among two distance functions Euclidean distance function is superior for SA.

Tables 6 and 7 represent the summary of different performance matrices of linear SVM and non-linear SVM with Gaussian kernel classifiers at different learning rates. From the tables, it is depicted that the accuracy of SA using SVM classifiers is varying

Table 3 Training and testing dataset information at 80% training rate

Sentiment	No. of training samples	No. of testing samples
Happy	10,380	2595
Sad	10,520	2630
Extremely happy	10,544	2636
Extremely sad	10,632	2658
Neutral	10,784	2696
Total	52,860	13,215

Table 4 Performance of cubic KNN

Training rate	Accuracy	Precision	Recall	F1 score
80	87.3	0.872	0.874	0.873
70	86.5	0.867	0.863	0.865
60	84.3	0.845	0.841	0.843
50	83.8	0.841	0.836	0.838

Table 5 Performance of fine KNN

Training rate	Accuracy	Precision	Recall	F1 score
80	89.4	0.893	0.895	0.894
70	88.3	0.886	0.881	0.883
60	87.1	0.872	0.870	0.871
50	85.9	0.861	0.857	0.859

between 90.1 and 94.8%. Among two kernel functions Gaussian kernel function is superior for SA.

Table 8 represents the performance comparison of ML classifiers. It is observed that proposed non-linear SVM optimized with Bayesian optimization attained ~ 5.5% accuracy higher than KNN and 4.5% higher than linear SVM classifier.

Table 6 Performance of linear SVM

Training rate	Accuracy	Precision	Recall	F1 score
80	90.1	0.907	0.90	0.902
70	89.6	0.893	0.89	0.894
60	87.8	0.876	0.875	0.876
50	84.3	0.846	0.841	0.843

Table 7 Performance of non-linear SVM with Gaussian Kernel

Training rate	Accuracy	Precision	Recall	F1 score
80	94.8	0.95	0.948	0.949
70	93.7	0.932	0.938	0.936
60	92.3	0.924	0.925	0.924
50	91.7	0.918	0.915	0.916

Table 8 Performance comparison of SVM and KNN classifiers

Method	Accuracy (%) at training rate			
	80%	70%	60%	50%
Cubic KNN	87.3	86.5	84.3	83.8
Fine KNN	89.4	88.3	87.1	85.9
Linear SVM	90.1	89.6	87.8	84.3
Non-linear SVM	94.8	93.7	92.3	91.7

5 Conclusion

In this paper, KNN and SVM classifiers are employed to classify the tweets based on its sentiment. The classifiers performance was evaluated using various training rates and datasets. The performance metrics of the proposed KNN Classifier with Minkowski (cubic) distance function are accuracy (87.3%), precision (0.872), recall (0.874), and *F1* score (0.873). Further, the analysis is carried out with the KNN classifier with Euclidean distance function, and it achieved the accuracy (89.4%), precision (0.893), recall (0.895), and *F1* score (0.894). It is observed that there is an improvement of 2% accuracy with KNN classifier with Euclidean distance function, and it also observed that KNN classifier is in poor to identify the natural sentiment accurately. Further, the analysis is carried out with the SVM classifiers with linear and Gaussian kernel functions. SVM with linear kernel achieved an accuracy of 90.1%, precision (0.907), recall (0.90), and *F1* score (0.902). Similarly, SVM with Gaussian kernel attained an accuracy (94.8%), precision (0.95), recall (0.948), and *F1* score (0.949). It is observed from the analysis Gaussian kernel is superior to liner kernel for sentiment analysis

References

1. Mikalai T, Themis P (2012) Survey on mining subjective data on the web. *Data Min Knowl Discov* 24:478–514
2. Wilson T, Wiebe J, Hoffman P (2005) Recognizing contextual polarity in phrase-level sentiment analysis. In: *Proceedings of HLT/EMNLP*
3. Liu B (2012) Sentiment analysis and opinion mining. *Synth Lect Human Lang Technol*
4. Liang-Chih Y, Jheng-Long W, Pei-Chann C, An-Shou CH (2013) Using a contextual entropy model to expand emotion words and their intensity for the sentiment classification of stock market news. *Knowl-Based Syst* 41:89–97
5. Hagenau M, Liebmann M, Neumann D (2013) Automated news reading: stock price prediction based on financial news using context-capturing features. *Decis Supp Syst*
6. X, Tao, Qinke P, Yinzhaoh C (2012) Identifying the semantic orientation of terms using S-HAL for sentiment analysis. *Knowl-Based Syst* 35:279–289
7. Isa M, Piek V (2012) A lexicon model for deep sentiment analysis and opinion mining applications. *Decis Support Syst* 53:680–688
8. Kumar ER, Rama Rao AKVSN, Nayak SR (2020) Emotional level classification and prediction of Tweets in Twitter. In: *Emotion and information processing*. Springer, Cham, pp 161–169
9. Yang C-T, Chen Y-L (2020) DACNN: dynamic weighted attention with multi-channel convolutional neural network for emotion recognition. In: *2020 21st IEEE international conference on mobile data management (MDM)*. IEEE, pp 316–321
10. Stojanovski D, Strezoski G, Madjarov G, Dimitrovski I, Chorbev I (2018) Deep neural network architecture for sentiment analysis and emotion identification of Twitter messages. *Multimedia Tools Appl* 77(24):32213–32242
11. Chatterjee A, Gupta U, Chinnakotla MK, Srikanth R, Galley M, Agrawal P (2019) Understanding emotions in text using deep learning and big data. *Comput Human Behav* 93:309–317
12. Hassan MK, Shakthi SP, Sasikala R (2017) Sentimental analysis of Amazon reviews using naïve bayes on laptop products with Mongo DB and R. In: *Materials science and engineering conference series*, vol 263, no 4

13. Elghazaly T, Mahmoud A, Hefny HA (2016) Political sentiment analysis using twitter data. In: Proceedings of the international conference on internet of things and cloud computing ACM
14. Mohammad M, Saif XZ, Kiritchenko S, Martin J (2015) Sentiment, emotion, purpose, and style in electoral tweets. *Inf Proc Manag* 51(4):480–499
15. Pagolu VS, Reddy KN, Panda G, Majhi B (2016) Sentiment analysis of Twitter data for predicting stock market movements. In: 2016 international conference on signal processing, communication, power and embedded system (SCOPEs), IEEE, pp 1345–1350
16. He W, Wu H, Yan G, Akula V, Shen J (2015) A novel social media competitive analytics framework with sentiment benchmarks. *Inf Manag* 52(7):801–812
17. Subbarao MV, Terlapu SK, Chowdary PSR (2022) Emotion recognition using BiLSTM classifier. In: 2022 international conference on computing, communication and power technology (IC3P). IEEE, pp 195–198
18. Venkata Subbarao M, Terlapu SK, Geethika N, Harika KD (2022) Speech emotion recognition using K-nearest neighbor classifiers. *Recent Adv Artif Intell Data Eng* 1386
19. Subbarao MV, Samundiswary P (2019) K-nearest neighbors based automatic modulation classifier for next generation adaptive radio systems. *Int J Secur Appl* 13:41–50
20. Subbarao MV, Padavala AK, Harika KD (2022) Performance analysis of speech command recognition using support vector machine classifiers. In: Communication and control for robotic systems. Smart innovation systems and technologies, vol 229
21. Venkata Subbarao M, Pravallika C, Ramesh Varma D, Prema Kumar M (2022) Power quality event classification using wavelets, decision trees and SVM classifiers. In: Saini HS, Singh RK, Tariq Beg M, Mulaveesala R, Mahmood MR (eds) Innovations in electronics and communication engineering. Lecture notes in networks and systems, vol 355. Springer, Singapore. https://doi.org/10.1007/978-981-16-8512-5_27

A Fused LBP Texture Descriptor-Based Image Retrieval System



Akbar Khan, Mohammad Hayath Rajvee, B. L. Deekshatulu,
and L. Pratap Reddy

Abstract Texture analysis is critical in a variety of computer vision applications, including object recognition, defect detection on surfaces, pattern recognition, and medical picture analysis. The purpose of this research is to offer a novel method for content-based texture picture classification that is based on the discrete wavelet transformation and several texture properties. Three approaches (LBP, DWT, and Tamura) are combined to build an efficient hybrid function vector capable of extracting the finest texture information. The study extracts LBP and Tamura features in two methods, via wavelet transform and fusion, to create an effective hybrid texture feature vector. Experiments on the Brodatz and MIT-VisTex databases demonstrate that the proposed approach is more precise than a single feature texture algorithm and also than a combination of Tamura texture features and wavelet transform features. Additionally, the technique that employs an SVM classifier achieves a higher level of accuracy, up to 99%.

Keywords SVM · LBP · CBIR · Texture descriptor · Visual patterns · Tumara features

A. Khan (✉)

Nimra College of Engineering and Technology, Vijayawada, A.P, India
e-mail: sarak123in@yahoo.com

M. H. Rajvee (✉)

PBR Visvodaya Institute of Technology and Science, Kavali, A.P, India
e-mail: razwe2003@gmail.com

B. L. Deekshatulu

IDRBT, RBI, Government of India, Hyderabad, India

L. Pratap Reddy

JNTUH, Hyderabad, India

1 Introduction

The expansion of the Internet and technology has resulted in a massive increase in the number of digital photos. It is vital to efficiently store and retrieve these photos. Content-based image retrieval (CBIR) is the most advanced area of image processing and computer vision. Shape, shading, and texture are often used image highlights. At the moment, determining the discriminative picture highlights is a more difficult job, as images of distinct classes have a variety of different sorts of highlights. Appropriately, there is still a dearth of effective and adaptable picture inclusions, which is especially beneficial when the dataset contains a variety of image formats. Frequently, descriptors are evaluated using the spatial data surface, form, shadings, and so on, and global descriptors are utilized for image recovery. The usage of neighboring descriptors has increased in recent years because they remain compatible with comparative qualities, with the caveat that neighborhood descriptors are extracted from picture districts rather than the entire picture [1]. Haralick et al. [2] presented a co-occurrence network representation of surface components in the mid-1970s. This approach examined the surface's dim level spatial ward. Tamura et al. [3] explored surface depiction from a variety of perspectives and presented a computer method for estimating six visual qualities, including directionality, contrast, coarseness, line likeness, consistency, and unpleasantness. The wavelet transform was suggested for surface representation in the mid-1990s. Smith and Chang [4, 5] visualized the surface by separating mean and fluctuation highlights from wavelet sub-bands. Thyagarajan et al. [6] used wavelet transforms in combination with KL extension and kohenon guides to explore the surface., Manjunath et al. [7] and Kundu et al. integrated wavelet change and co-occurrence frameworks to exploit the benefits of both insights- and change-based surface investigations. Mama and Ojala [8] examined the explanations for surface images using a range of wavelet surface representations, including symmetrical and bi-symmetrical wavelet changes, tree-organized wavelet changes, and Gabor wavelet changes. At first, the surface range was employed to denote the separation between surfaces [9]. The utility of the surface range approach is determined by the extraction of adjacent surface data for each pixel and the representation of the textural component of a computerized picture as a range.

Similarly, Ojala et al. [10] proposed the formally dressed nearby parallel examples (LBP) method for dealing with extricating revolution and histogram evening out invariant highlights, which was expanded upon by Hemachandran [11] by defining subsidiary-based neighborhood twofold examples and applying them to the use of face arrangement. The conventional LBP approach is straightforward and proficient, since it considers the consistent instances in photographs that will serve as nearby highlights of a photograph. In [12], a noise-invariant local ternary pattern (LTP) was presented that operates in three states and is capable of identifying and extracting differences between the center pixel and its neighborhood. Raghuwanshi et al. [13] have suggested a system for picture recovery based on half breed highlight vectors. Local twofold example, color second, and automatic division measurements are all taken separately. Each of these highlights was condensed for the purpose of

arranging a half and half component vector utilizing the standardizing cycle. Additionally, support vector machines (SVMs) have been used to determine layout. Over the last several years, extensive study has been conducted on describing texture patterns, as this is a particularly difficult and time-consuming operation in the field of image processing. Due of its low computing complexity and great selectivity, the LBP texture description has attracted considerable interest [14]. Due to these qualities, LBP has been effectively applied in a variety of applications, including pattern recognition, image analysis, and the CBIR systems. The discrete wavelet transform is utilized as the second texture predictor in this case [15]. Due of the multi-orientation and multi-resolutional qualities of these functions, this wavelet transform retrieves local features [16]. Raghuwanshi et al. [13] suggested a CBIR technique for extracting the finest features of texture by combining (LBP and DWT) into an efficient hybrid feature vector. On the MIT-VisTex and Brodatz benchmark datasets, the performance of SVM and ELM machine learning classifiers is evaluated. The primary goal of this research is to enhance the efficiency of texture feature-based image retrieval systems by developing a novel and advanced hybrid CBIR system that combines well-known texture extraction techniques, such as discrete wavelet transform, local binary pattern (LBP), and Tamura features. The texture descriptors utilized accurately reflect the texture while also accounting for rotation and scale variations that may impact the appearance of the images.

2 Related Work

The texture of a picture is one of the most crucial low-level characteristics. It organizes the visual patterns depicted in the pictures spatially. The literature examines numerous CBIR systems employing a variety of texture characteristics such as the DWT, LBP, Gabor transform, gray level co-occurrence matrix (GLCM), and Fadaei [17]. A texture retrieval method based on adaptive tartlet transformations that provides exceptionally precise texture information was proposed. The optimal mix of tetrominoes was chosen to optimize the image's shape at all levels [18]. In a unique method [19], local tetra patterns were described for image indexing and retrieval in CBIR systems. To reduce the information content of pictures, a novel approach called local derivative radial pattern (LDRP) was presented [20]. It employs multi-level coding in place of binary coding in several directions, which reduces information loss while increasing accuracy somewhat. In [21], the Tamura feature extraction method was applied, which determined the degree of similarity between images based on the fuzzy humming distance. Kokare et al. [22] developed a method for extracting fresh and supplementary information from a query image using LBP maps. The descriptors are then computed using the fractal dimension. Texture and color are utilized by the hybrid CBIR system described by [23] authors. Using wavelet decomposition, the texture features and classification of texture classes with variation in the high-frequency coefficients were retrieved. This paper presents a novel and effective CBIR strategy based on a combination of LBP and Tamura

features obtained using DWT and an SVM classifier. This exhaustive method is used to two traditional datasets, i.e., Brodatz and MIT-VisTex outperformed other state-of-the-art approaches and the results produced. The LBP, Tamura features, DWT, and SVM are the widely used methods for feature extraction.

3 Proposed Framework

In Fig. 1, a novel CBIR method based on integrating the characteristics of two prominent methodologies, LBP and Tamura, to extract finer texture information is shown, while the DWT supports several resolutions and orientations. It is capable of extracting information about the image's shape at a larger scale, hence enhancing the retrieval rate. The features of the database photos at various resolutions and scales are retrieved using both (LBP and Tamura) approaches and integrated throughout the normalization phase. The system's steps are listed below.

3.1 System Framework Algorithm

Part1: Feature vector construction.

Input: An image from the database.

Output: Feature vector. The steps followed are:

- (1) If the image is colorful, extract it from the image database and convert it to gray scale.
- (2) Use DWT to generate sub-band pictures.
- (3) Compute the histogram using LBP for the estimated coefficient picture.
- (4) Compute the histogram for all sub-band pictures using Tamura features.
- (5) Using max–min normalization, concatenate the two histograms generated in steps 2–4 to create the feature vector

Part 2: Retrieval of images using LBP and Tamura characteristics.

Utilize a database query to retrieve a picture. Output: Images that were retrieved using the similarity measure.

- (6) Input the query picture.
- (7) Repeat steps 2–5 in part 1 to extract the query image's feature vector.

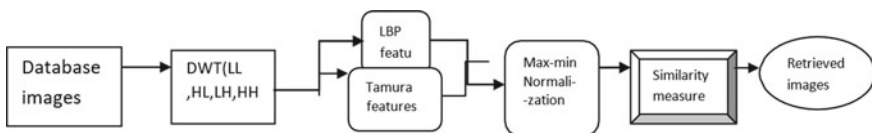


Fig. 1 Block diagram of CBIR system

- (8) Using several similarity metrics, compute the similarity index of the query image vector to each database picture.
- (9) To obtain the set of related photos, sort the similarity indices from highest to lowest.
- (10) Utilize analytics to assess performance.

3.2 Similarity Measure

Along with feature vector estimation, the similarity test is critical for picture recovery in a content-based image recovery approach. After the feature vectors have been measured, i.e., after the dissimilarity between the pictures has been calculated, this similarity measure computes the distance between the feature vector of the query image and the feature vector of each image in the database. With an emphasis on indexing, this computation is carried out, and indices are organized as a collection of recovered photos with smaller measurements. The D1 distance metric is used to determine similarity matches.

4 Experimental Results and Analysis

The proposed image retrieval technique is evaluated in terms of precision and recall. The proposed method's superiority was proven by comparing it to numerous existing texture patterns for picture retrieval using the previously mentioned assessment criteria. Each database received many images depending on a query image. Each database image has been used as a query image once.

The experiments test the most critical CBIR system parameters that are measured as equations, precision, and recall (1) as well as (2).

$$\text{Precision}(P) = \frac{\text{Total number of relevant images retrieved from the database}}{\text{Total no of images in the da}(N)} \quad (1)$$

$$\text{Recall}(R) = \frac{\text{Total number of relevant images retrieved from the database}}{\text{Total no of images in the dataset}} \quad (2)$$

For each category, the average accuracy and recall values may be calculated using Eqs. 3 and 4.

$$P_{\text{avg}} = 1/L \sum_{r=1}^L P^r \quad (3)$$

$$R_{\text{avg}} = 1/L \sum_{r=1}^L R^r \quad (4)$$

Using Eqs. 5 and 6, we can calculate the total precision and total recall for our experiment.

$$P_{\text{total}} = \sum_{k^c=1}^{k-1} P_{\text{avg}}(c) \quad (5)$$

$$R_{\text{total}} = \sum_{k^c=1}^{k-1} R_{\text{avg}}(c) \quad (6)$$

Two reference datasets, VisTex and Brodatz are utilized for the experiments. In every trial, the LBP and DWT techniques are combined to measure the parameters accuracy and recall.

Experiment 1.

The Brodatz texture database, which contains 13 distinct groups of pictures such as raffia, grass, brick, and bark at a scale of 512×512 , are rotated in the very first experiment, with each class consisting of seven images with orientations (0, 30, 60, 90, 120, 150, and 200). Each picture is then partitioned into 16 smaller photos, resulting in a total of 1456 images, each 128 by 128 in size. Figure 2 illustrates a selection of images from this database. The top 25, 35, 45, 55, and 65 retrievals are returned by selecting a query image from the whole database. The average precision and recall graphs for the combination of DWT and LBP, as well as for SVM, demonstrate that the proposed framework outperforms the basic combined approach of both texture approaches.

For each input image, local binary pattern is applied and the retrieved images are shown in Figs. 2 and 3.

Experiment 2

In the second experiment, 40 distinct texture pictures with a resolution of 512×512 pixels were picked from the MIT-VisTex collection. Following that, each image was segmented into 16 bits, each 128 by 128 in size, and the final database had 640

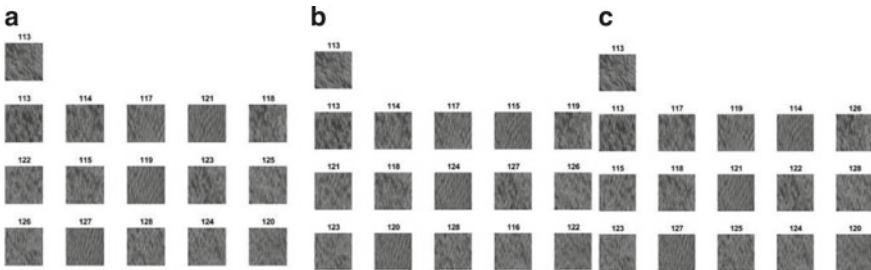


Fig. 2 Retrieved top 15 images from Brodatz database by proposed methods. **a** LBP **b** Wavelbp + Tamurac LBP + Tamura

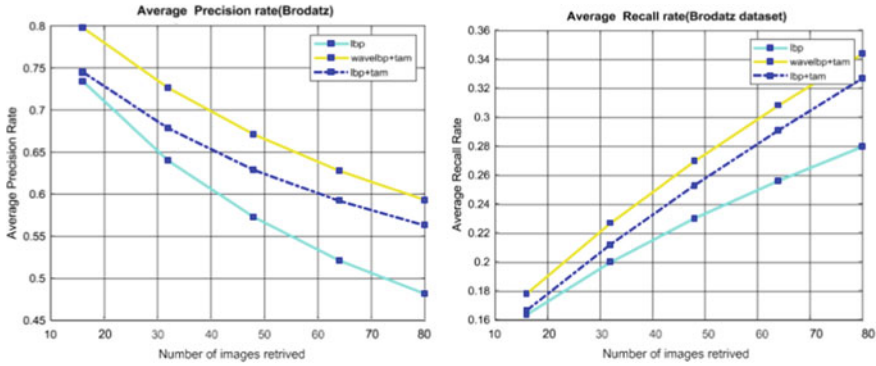


Fig. 3 Average precision and average recall rate graphs

photos. As with the previous experiment, all images are taken as the query image and the top 16, 32, 48, 64, and 80 images are obtained. Figure 5 illustrates representative images from this collection. The combination of LBP and Tamura is used to examine average accuracy and recall using DWT and the specified methodologies. In Fig. 4, comparison graphs for precision and recall at various retrievals are studied. According to the produced graphs, the suggested approach combined with the SVM classifier also produces the best results when compared to the other methods in this dataset.

To demonstrate our proposed work’s uniqueness in terms of precision, Table 1 compares it to other recent state-of-the-art approaches for retrieving the top 25 and 16 photos from the Brodatz and MIT-VisTex datasets, respectively.

As shown in Table 1, the average precision rate (APR) values produced by our proposed method are much greater than those obtained by other current texture descriptor techniques in CBIR systems. The average precision of MIT-VisTex is virtually identical to that of the other three approaches. However, in the instance of

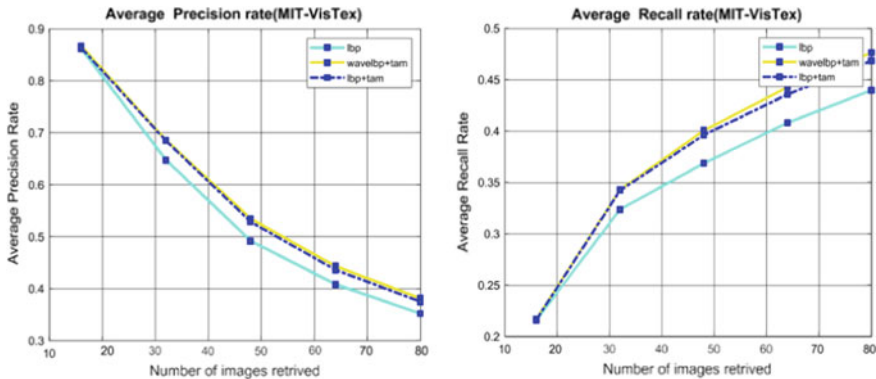


Fig. 4 Average precision and average recall rate graphs—comparison of the suggested methods’ performance to that of other existing approaches

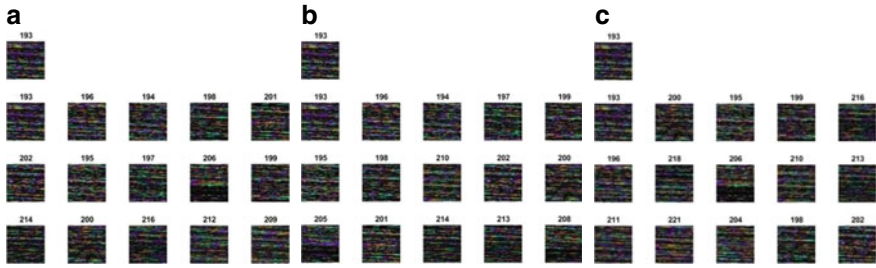


Fig. 5 Retrieved top 15 images from MIT-VisTex database by proposed methods. **a** LBP **b** Wavelbp + Tamurac **LBP + Tamura**

Table 1 Average precision comparison of suggested approaches with alternative retrieval counts of 25 and 16 for the Brodatz and VisTex databases

Database	Parameter/feature	LBP	Wavelbp + Tamura	LBP + Tamura
Brodatz	Average precision	73.4	79.9	74.5
Brodatz	Average recall	16.4	17.8	16.6
MIT-VisTex	Average precision	86.2	86.6	86.3
MIT-VisTex	Average recall	21.5	21.7	21.6

Table 2 Comparison of all the methods

Features with SVM classifier	Accuracy for Brodatz dataset	Accuracy for MIT-VisTex dataset
Wavelbp + Tamura + SVM	100	99.8
DWT + Tamura + SVM	98.73	97.6

the Brodatz database, the suggested technique achieves a greater average accuracy. As shown in Table 2, the SVM classifier has the best accuracy for both techniques and datasets.

5 Conclusion and Future Scope

For texture-based images, a hybrid CBIR system is presented to improve the system’s overall precision and accuracy. In this built framework, the combination of traits and categories is crucial. To acquire the finest texture details, two features LBP and Tamura were extracted and combined using DWT to create an effective hybrid feature vector. In comparison with SVM classifier, the CBIR method performs better when the d1 distance is used. The performance of SVM classifiers is compared in this

section. This novel overall system is evaluated on two benchmark datasets, MIT-VisTex and Brodatz, and Euclidean distance is utilized to calculate image similarity. The collected findings demonstrate that the suggested system outperforms previous texture-based approaches using an SVM classifier.

References

1. Zhao R, Grosky WI (2000) From features to semantics: some preliminary results. In: Multi-media and Expo, 2000. ICME 2000. 2000 IEEE international conference on, 2(c), vol 2, pp 679–682
2. Haralick RM, Shanmuga K, Dinstein (1973) Textural features for image classification. *IEEE Trans Syst Man Cyber SMC*3:610–621
3. Tamura H, Mori S, Yamawaki T (1978) Textural feature scor responding to visual perception. *IEEE T Syst Man Cyber* 8(6):460–472
4. Smithand JR, Chang S-F (1997) visually searching the web for content. *IEEE Multimedia Mag* 4(3):12–20. [Columbia U. CU/CTR Technical Report 459-96-25]
5. Smith JR, Chang SF (1994) Transform features for texture classification and discrimination in large image databases. In: *Proceeding IEEE international conference on image proceeding 1994*
6. Thyagarajan KS, Nguyen T, Persons C (1994) A maximum likelihood approach to texture classification using wavelet transform. In: *Proceeding IEEE international conference on image proceeding*
7. Manjunath BS, Ma WY (1996) Texture features for browsing and retrieval of image data. *IEEE Trans Pattern Anal Mach* 18(8):837–842
8. Ojala T, Pietikäinen M, Mäenpää T (2002) Multiresolution gray-scale and rotation in variant texture classification with local binary patterns. *IEEE Trans Pattern Anal Mach Intell* 24(7):971–987
9. Huang G, Member S, Zhou H, Ding X, Zhang R (2012) Extreme learning machine for regression and multiclass classification 42(2):513–529
10. Ricardo A, Joaci J, Sá DM (2017) Neuro computing LBP maps for improving fractal based texture classification 266:1–7
11. Hemachandran K, Paul A, Singha M (2012) Content-based image retrieval using the combination of the fast wavelet transformation and the colour histogram. *IET Image Proc* 6(9):1221–1226
12. Wang H, Feng L, Zhang J, Liu Y (2016) Semantic discriminative metric learning for image similarity measurement. *IEEE Trans Multimedia* 18(8):1579–1589
13. Raghuwanshi G, Tyagi V (2015) Texture image retrieval using adaptive tetrolet transforms. *Dig Sig Proc* 1(3):1–8
14. Dhingraa S, Bansalb P (2019) An intelligent multi-resolution and rotational invariant texture descriptor for image retrieval systems. *ADCAIJ: Adv Distrib Comput Artif Intell J Regular Issue* 8(2):33–49
15. Naghashi V (2018) Optik co-occurrence of adjacent sparse local ternary patterns: a featured script or for texture and face image retrieval. *Optik—Int J Light Electron Optics* 157(877–889):2018
16. Murala S, Maheshwari RP, Bala subramanian R (2012) Local tetra patterns: a new feature descriptor for content-based image retrieval
17. Fadaei S, Amirfattahi R, Ahmadzadeh MR (2017) Local derivative radial patterns: a new texture descriptor for content-based image retrieval. *Sig Proc* 137:274–286
18. Sreena PH, George DS (2013) Content based image retrieval system with fuzzified texture similarity measurement, (ICCC) 80–85

19. Liao S, Law MWK, Chung ACS (2009) Dominant local binary pattern for texture classification. *IEEE Trans Image Proc* 18(5):1107–1118
20. Alaei F, Alaei A, Pal U, Blumenstein M (2018) ACPT USCR. *Expert systems with applications*. <https://doi.org/10.1016/j.eswa.2018.12.007>
21. Jain AK, Vailaya A (1995) Image retrieval using color and shape. *Pattern Recognition* 29(1):233–1244
22. Kokare M, Biswas PK, Chatterji BN (2005) Complex wavelet filters 35(6):1168–1178 [26]; Pham M (2017) Color texture image retrieval based on local extrema features and riemannian distance
23. Das R, Dash JK, Mukhopadhyay S (2013) Rotation invariant textural feature extraction for image retrieval using eigen value analysis of intensity gradients and multi-resolution analysis 46:3256–3267

Normal and Alcohol EEG Signals Classification Using Singular Spectrum Analysis



Venkata Keshava Krishna Paramkusham and Sachin Taran

Abstract Consumption of alcohol alters the functioning process of nervous system by disturbing the neuron process, which leads to the behavioral changes in a human life. An automatic identification of alcoholics can address these issues. Electroencephalogram (EEG) is a widely used tool for monitoring the brain activities. In this study, singular spectrum analysis (SSA) and machine learning-based algorithm are proposed for the automatic detection of normal and alcohol EEG signals. Kruskal Wallis test is performed as a part of a statistical study and the features which satisfy $p < 0.05$ are considered in the classification. Initially, multiple SSA-based features are extracted out of which the inter-quartile range and wavelength provide the best classification performance with an optimizable support vector machine classifier. The achieved classification accuracy is 94.2%.

Keywords Electroencephalogram · Singular spectrum analysis · Support vector machine · Optimizable classifier

1 Introduction

Consumption of alcohol has far reached socio-cultural and economic consequences for the drinkers. An increase in traffic accidents, machine-related mishaps, and violence are some results of drinking. Regular intake of alcohol damages the organs and DNA of human beings. Drinking leads to the third position in diseases. Alcoholism-related genetic diseases include weaker and less organized theta rhythms [1]. Drinking is the third major cause of disease as per the world health organization [2]. Furthermore, alcoholism increases as per disability-adjusted life-years (DALY) which estimates that 5.1% of global diseases are related to alcohol consumption [3]. Furthermore, the rising incidence of malignancies linked to alcohol adds to the seriousness of the situation [4]. These negative consequences highlight the importance

V. K. K. Paramkusham (✉) · S. Taran
Delhi Technological University, New Delhi, India
e-mail: pvkk00007@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
V. V. S. S. Chakravarthy et al. (eds.), *Advances in Signal Processing, Embedded Systems and IoT*, Lecture Notes in Electrical Engineering 992,
https://doi.org/10.1007/978-981-19-8865-3_14

155

of enhanced researches aiming at price efficient and early alcohol misuse monitoring and diagnosis [5]. The non-invasive nature of electroencephalogram (EEG)-based techniques makes it more viable for real-time diagnosis of alcoholics [6].

Several studies are conducted for the identification of alcohol EEG data. A variational mode and empirical mode decompositions (EMD) with least square support vector machine (LS-SVM) and K -nearest neighbor (KNN) algorithms are employed for the identification of alcohol EEG signals [7]. A wavelet filter bank-based approach is suggested for the identification of the alcohol EEG signals [8]. Kolmogorov–Smirnov test-based features are explored with Adaboost k -means algorithms for identifying alcohol EEG signals [9]. The EMD-based feature extraction with ensemble subspace KNN-based classification is explored to identify alcohol EEG signals [10]. A test to check the ability of parametric spectrum and coherence estimators and phase synchrony processor, which identifies the variations in the scalp while eyes remained open in both alcohol and normal EEG data is explored [11]. A fast Fourier transform and autoregression modeling are explored for the identification of alcohol EEG signals with discriminant analysis (DA) [12]. A tunable- Q wavelet transform (TQWT) with ensemble classifiers is explored for the identification of apnea events [13]. For the automatic identification of alcoholics, the nonlinear features of computer-aided diagnostics are examined with SVM [14]. Higher-order spectrum cumulants and other nonlinear features are extracted to assess alcohol-related alterations in EEG data, various machine learning algorithms are employed to classify alcohol EEG data [15]. The classification of alcohol EEG is explored by wavelet packet decomposition with KNN is presented [16]. The automatic identification of alcohol EEG signals is done by using a time–frequency image-based technique is suggested [17]. The diagnosis of alcoholics is explored by the TQWT with SVM [18]. The correlation analysis is adopted for the statistical analysis, and KNN is explored to identify alcohol EEG data [19]. EMD with extreme machine learning and SVM algorithms is explored for the analysis of alcohol EEG signals using EEG rhythms which is proposed [20]. The identification of alcohol EEG data is explored by using wavelet transform with extreme learning machine (ELM) [21]. The power spectral density (PSD) is explored to identify the changes in alcohol EEG signals which is presented [22]. Principle component analysis and singular value decomposition-based extracted features are employed to classify the alcohol EEG data with KNN is explored [23]. PSD of the haar mother wavelet-based features is explored to classify the alcohol EEG data with SVM [24]. The automated identification of alcohol EEG signals is explored using a dual-tree complex WT and SVM classifier [25].

This work suggested singular spectrum analysis (SSA) and an optimizable classifier-based algorithm to classify the alcohol EEG data. SSA-based features were extracted and further driven to multiple classifiers for the classification.

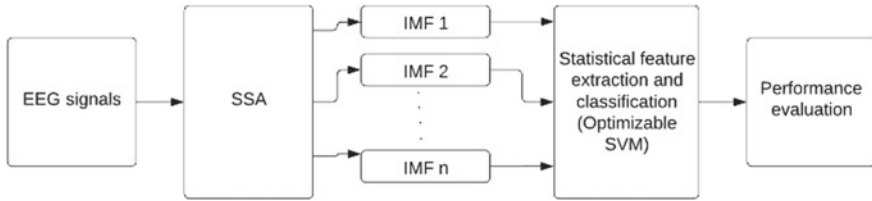


Fig. 1 Block diagram of proposed methodology

2 Methodology

2.1 Dataset

The validation of the suggested methodology is done by using the dataset which contains EEG data of alcoholic and normal person, respectively. The dataset is provided online [26]. The signals are acquired using 64 electrodes which has a sampling frequency of 256 Hz. A total 120 EEG signals of each class are used with 2048 sample lengths to distinguish normal and alcoholic people. Figure 1 shows the data flow block diagram of the proposed work.

2.2 Singular Spectrum Analysis

The SSA is a nonparametric technique that is frequently accessed to analyze meteorological and geophysical data series [27]. SSA is made up of two steps: embedding and reconstruction. During the embedding process, the EEG data vector S is mapped with a multivariate data matrix with arranging the K number of delayed vectors with a size D , which in-turn yields to a trajectory matrix (TM) T of size $A \times B$ is given as [27],

$$T_{A \times B} = \begin{bmatrix} t(1) & t(2) & \dots & \dots & t(K) \\ t(2) & t(3) & \dots & \dots & t(K+1) \\ \vdots & \vdots & \dots & \dots & \vdots \\ t(D) & t(D+1) & \dots & \dots & t(N) \end{bmatrix}, \quad (1)$$

here $K = N - D + 1$ and D stands for window length and can be defined for $D > \frac{f_s}{f}$, where f_s stands for sampling frequency and frequency of the signal is represented by f . The TM is further decomposed into M number of TM, $T_1, T_2, T_3, \dots, T_M$, in the last stage of SSA.

2.3 Feature Extraction and Classification

This work uses IQR and wavelength (WAVELEN) as features. The difference of third quartile to the first quartile is IQR [28]. IQR can be calculated by obtaining the center point of both the upper and lower half of the data. Wavelength is defined as the EEG waveform's cumulative length across a time segment [29].

In this work, SVM is explored to solve two class classification problem. The goal of the SVM algorithm is to find the optimal separating hyperplane, which separates the samples while also maximizing the distance between the two classes [30].

3 Results and Discussion

For the identification of alcohol and normal EEG data, a technique based on SSA is presented. The data comprises 53 features with two classes (i.e., normal and alcohol). Each class has 16 subbands (SBs) with a feature set of 1×120 for both alcohol and normal, respectively. The KW test is used to perform a discriminative analysis of derived features. The features having smaller probabilistic ($p < 0.05$) values are used for the classification, indicating that they have high discriminative power for inter class EEG signals. KW test results are shown in Table 1. The SBs 2, 4, 5, 6, 7, 8, 9, 10, 11, 12, 16 have lower p value as shown in Table 1 which are considered for the classification. For alcohol EEG classification, the subband-wise characteristics are fed to various machine learning tools. A 10-folds cross-validation method is used at the classifier stage. The features are tested with all the variants of the SVM algorithm. Table 2 presents the achieved accuracy of the proposed features using the SVM classifier. Optimizable SVM gives a better accuracy of 94.2% when compared to all other variants of SVM.

In Fig. 2, minimum error classification (MEC) plot depicts the classification error rate at the classification stage. The square block indicates the efficient point hyper

Table 1 For different SB's, the KW test p values of IQR and WAVELEN features

Features\SBs	SB1	SB2	SB3	SB4	SB5	SB6	Sb7	SB8
IQR	0.0293	2.88×10^{-8}	0.1033	2.20×10^{-3}	0.0635	2.20×10^{-3}	0.0046	2.90×10^{-3}
WAVELEN	0.5936	7.85×10^{-12}	0.0065	2.51×10^{-10}	0.0008	4.31×10^{-10}	3.34×10^{-6}	2.42×10^{-7}
Features\SBs	SB9	SB10	SB11	SB12	SB13	SB14	SB15	SB16
IQR	9.37×10^{-6}	1.12×10^{-5}	8.66×10^{-5}	1.53×10^{-2}	6.87×10^{-2}	9.23×10^{-1}	1.64×10^{-1}	1.22×10^{-1}
WAVELEN	4.48×10^{-12}	8.74×10^{-10}	7×10^{-10}	1.90×10^{-5}	3.20×10^{-3}	3.80×10^{-1}	1.22×10^{-1}	2.31×10^{-2}

Table 2 Accuracy of the proposed features using SVM classifier variants. Here, *F* is fine, *M* is medium, *C* is coarse, *L* is linear, *Q* is quadratic, and Cu is cubic

SVM classifier variants	<i>L</i> -SVM	<i>Q</i> -SVM	Cu-SVM	<i>F</i> -Gaussian SVM	<i>M</i> -Gaussian SVM	<i>C</i> -Gaussian SVM	Optimizable SVM
Accuracy	90.0%	89.2%	90.0%	83.8%	93.8%	89.2%	94.2%

parameter, and circle shows the minimum error hyper parameter. The receiver operating characteristics curve (ROC) of the used features is shown in Fig. 3. The quality of the used features can be determined by the ROC using the threshold values which ranging [0, 1]. The area under the curve (AUC) achieved is 0.97 for optimizable SVM model.

The confusion matrix (CM) of proposed features is shown in Fig. 4. The predicted class is represented by the rows of the CM, while the target class is represented by the columns. The other performance measures such as accuracy (AC), sensitivity (S_{en}), specificity (S_{pe}), *F1*-score (F_1), and Matthew’s correlation coefficient (M_{CC}) are also computed for the validation of proposed features.

The performance metrics for the proposed feature are presented in Table 3. Table 3 explains that the suggested method’s misclassification rate of 5.83% is much lower, indicating that both classes of EEG signals are correctly classified. The S_{en} of 91.67% and S_{pe} of 96.67% are almost near to the optimum values of S_{en} and S_{pe} . Precision and recall are defined by the *F1*-score composite. The *F1*-score for the suggested technique is 0.94, which is closer to its maximum value. The classification performance was evaluated using M_{CC} . The M_{CC} value attained by the proposed method is 88.44%, which is closer to the ideal value.

Fig. 2 MEC plot of optimizable SVM of the used features

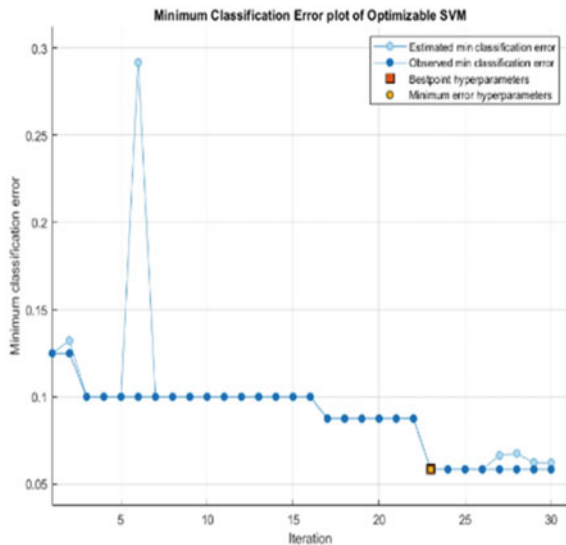


Fig. 3 ROC plot of optimizable SVM for used features

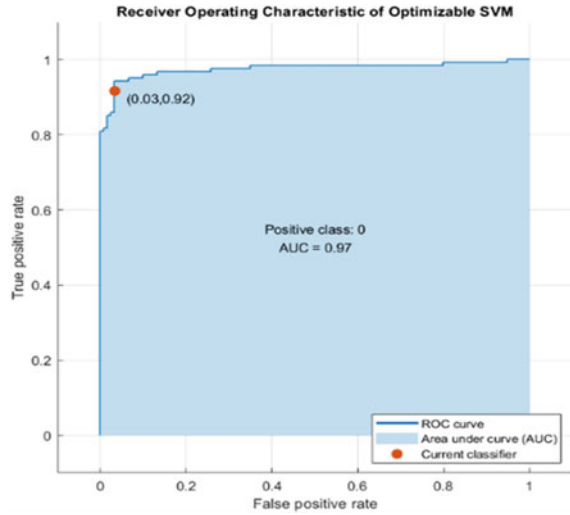


Fig. 4 CM of optimizable SVM for the proposed features

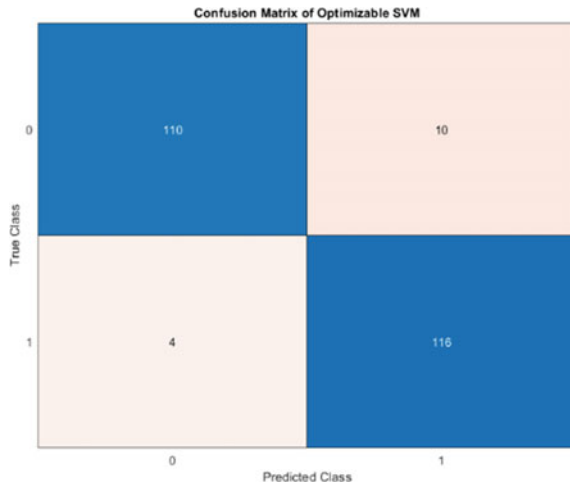


Table 3 Suggested classification model’s performance metrics

Performance measure (ideal value)	Classification performance
AC (100%)	94.17
Error (0%)	5.83
S_{en} (100%)	91.67
S_{pe} (100%)	96.67
Precision (100%)	96.49
F_1 (1)	0.94
M_{CC} (100%)	88.44
Kappa (1)	0.88

Table 4 Suggested method's performance summary in comparison with existing techniques

Author	Features	Classifier	Classification performance (AC)
Faust et al. [5]	PSD, peak amplitude, and frequency	ROC	0.822 (AUC)
Ehlers et al. [14]	Correlation dimension	DA	88%
Acharya et al. [20]	Nonlinear features	SVM (poly kernel)	76.19%
Faust et al. [21]	Higher-order spectra cumulants	Fuzzy sugeno classifier	92.4%
Proposed method	IQR, wavelength	Optimizable SVM	94.2%

The utility of the suggested strategy is now demonstrated by a comparison of performance utilizing the same dataset methods as shown in Table 4. The techniques are explored in terms of the classifier and accuracy that have been achieved. Faust et al. have achieved a significant ROC value of 0.822 using nine features and using fuzzy sugeno classifier (FSC) managed to achieve an accuracy of 92.4% [5, 21]. Ehlers et al. have used DA for the classification and achieved an accuracy of 88% [14]. Acharya et al. have managed to achieve an accuracy of 76.19% using SVM [20]. This table shows that present approaches use a variety of characteristics for the classification of alcohol and normal EEG data, which raises the classification complexity of an online classification system. The performance metrics S_{en} , S_{pe} , AC, and M_{CC} obtained by the suggested technique are 91.67, 96.67, 94.17%, and 88.44%, respectively. The obtained classification performance demonstrates the method's dependability and resilience in classifying alcohol and normal EEG signals.

4 Conclusion

This work presents the classification of the normal and alcohol EEG signals by using SSA-based features and SVM. The KW test is used to perform the statistical analysis. Based on the KW test results, the features are selected. The features used here are IQR and WAVELEN. These features are tested on optimize SVM for the selected subbands. The obtained classification accuracy is 94.17%, sensitivity is 91.67%, specificity is 96.67%, and precision is 96.49%.

The suggested SSA-based feature can be explore for motor imagery tasks EEG signals classification, sleep apnea detection, etc.

References

1. Rangaswamy M, Jones KA, Porjesz B et al (2007) Delta and theta oscillations as risk markers in adolescent offspring of alcoholics. *Int J Psychophysiol* 63(1):3–15
2. World Health Organization (2011) Alcohol fact sheet number 349, 2011, inverse Gaussian parameters and adaptive boosting. *Neurocomputing*
3. Organization WH et al (2004) Global status report on alcohol
4. Druesne-Pecollo N, Tehard B, Mallet Y, Gerber M, Norat T, Hercberg S, Latino-Martel P (2009) Alcohol and genetic polymorphisms: effect on risk of alcohol-related cancer. *Lancet Oncol* 10(2):173–180
5. Faust O, Acharya R, Allen AR, Lin C (2008) Analysis of eeg signals during epileptic and alcoholic states using ar modeling techniques. *IRBM* 29(1):44–52
6. Taran S, Bajaj V, Sharma D, Siuly S, Sengur A (2017) Features based on analytic IMF for classifying motor imagery EEG signals in BCI applications. *Measurement*
7. Salankar N, Qaisar SM, Pławiak P, Tadeusiewicz R, Hammad M (2022) EEG based alcoholism detection by oscillatory modes decomposition second order difference plots and machine learning. *Biocyber Biomed Eng* 42(1):173–186, ISSN 0208-5216
8. Sharma M, Deb D, Acharya UR (2018) A novel three-band orthogonal wavelet filter bank method for an automated identification of alcoholic EEG signals. *Appl Intell* 48:1368–1378
9. Diykh M, Abdulla S, Oudah AY, Marhoon HA, Siuly S (2021) A novel alcoholic EEG signals classification approach based on adaboost k-means coupled with statistical model. In: Siuly S, Wang H, Chen L, Guo Y, Xing C (eds) *Health information science*. Springer International Publishing, Cham, pp 82–92
10. Bavkar S, Iyer B, Deosarkar S (2021) Optimal EEG channels selection for alcoholism screening using EMD domain statistical features and harmony search algorithm. *Biocyber Biomed Eng* 41:83–96
11. Tcheslavski GV, Gonen FF (2012) Alcoholism-related alterations in spectrum, coherence, and phase synchrony of topical electroencephalogram 42(4):0–401
12. Ehlers CL, Havstad J, Prichard D, Theiler J (1998) Low doses of ethanol reduce evidence for nonlinear structure in brain activity. *J Neurosci* 18(18):7474–7486
13. Taran S, Bajaj V, Sinha GR, Polat K (2021) Detection of sleep apnea events using electroencephalogram signals. *Appl Acoustics* 181:108137. ISSN 0003-682X
14. Acharya UR, Sree SV, Chattopadhyay S et al (2012) Automated diagnosis of normal and alcoholic EEG signals. *Int J Neural Syst* 22(03):1250011
15. Faust O, Yanti R, Wenwei Y (2013) Automated detection of alcohol related changes in electroencephalograph signals. *J Med Imaging Health Inf* 3(2):333–339
16. Faust O, Yu W, Kadri NA (2013) Computer-based identification of normal and alcoholic eeg signals using wavelet packets and energy measures. *J Mech Med Biol* 13(3):1350033
17. Bajaj V, Guo Y, Sengur A et al (2017) A hybrid method based on time–frequency images for classification of alcohol and control EEG signals. *Neural Comput Applic* 28:3717–3723
18. Patidar S, Pachori RB, Upadhyay A, Rajendra Acharya U (2016) An integrated alcoholic index using tunable-Q wavelet transform based features extracted from EEG signals for diagnosis of alcoholism. *Appl Soft Comput* S1568494616305713
19. Padma Shri TK, Sriaram N (2016) Spectral entropy feature subset selection using SEPCOR to detect alcoholic impact on gamma sub band visual event related potentials of multichannel electroencephalograms (EEG). *Applied Soft Comput* S1568494616301971
20. Taran S, Bajaj V (2018) Rhythm-based identification of alcohol EEG signals. *IET Sci Measur Technol* 12(3):343–349
21. Malar E, Gauthaam M (2020) Wavelet analysis of EEG for the identification of alcoholics using probabilistic classifiers and neural networks. *Int J Intell Sustain Comput* 1(1):3
22. Di W, Zhihua C, Ruifang F, Guangyu L, Tian L (2010) [IEEE 2010 3rd IEEE international conference on computer science and information technology (ICCSIT 2010)—Chengdu, China (2010.07.9–2010.07.11)] 2010 3rd international conference on computer science and information technology—study on human brain after consuming alcohol based on EEG signal, pp 406–409

23. Yazdani A, Kamaledin Setarehdan S (2007) [IEEE 2007 9th international symposium on signal processing and its applications (ISSPA)—Sharjah, United Arab Emirates (2007.2.12–2007.2.15)] 2007 9th international symposium on signal processing and its applications—classification of EEG signals correlated with alcohol abusers, pp 1–4
24. Kousarrizi N, Reza M, Asadi Ghanbari A, Gharaviri A, Teshnehlab M, Aliyari M (2009) [IEEE 2009 3rd international conference on bioinformatics and biomedical engineering (iCBBE)—Beijing, China (2009.06.11–2009.06.13)] 2009 3rd international conference on bioinformatics and biomedical engineering—classification of alcoholics and non-alcoholics via EEG using SVM and neural networks, pp 1–4
25. Sharma M, Sharma P, Pachori RB, Rajendra Acharya U (2018) Dual-tree complex wavelet transform-based features for automated alcoholism identification. *Int J Fuzzy Syst*
26. American Electroencephalographic Association 1990 (2007) Standard electrode position nomenclature. Available at http://kdd.ics.uci.edu/databases/eeg/eeg_data.html
27. Golyandina N, Nekrutkin V, Zhigljavsky A (2001) Analysis of time series structure: SSA and related techniques, ser. Monographs on statistics and applied probability. Chapman and Hall, Boca Raton, Florida, vol 90
28. Taran S, Sharma PC, Bajaj V (2020) Automatic sleep stages classification using optimize flexible analytic wavelet transform. *Knowl-Based Syst* 192:105367. ISSN 0950-7051
29. Phinyomark A, Phukpattaranont P, Limsakul C (2012) Feature reduction and selection for EMG signal classification. *Exp syst Appl* 39(8):7420–7431
30. Li S, Zhou W, Yuan Q, Geng S, Cai D (2013) Feature extraction and recognition of ICTAL EEG using EMD and SVM. *Comput Biol Med* 43(7):807–816

Tropospheric Zenith Delay (TZD) for Microwaves During Severe Weather Events Over a Few Indian Stations



A. Narendra Babu, P. S. Brahmanandam, G. Uma, K. Pushpa, K. Srinivas,
and A. Praneetha

Abstract To have an accurate signal in navigational systems such as GPS, it must be known how much the radio signal gets delayed when it propagates from the satellite to the receiver. In general, the delay is mostly caused by refractivity of the troposphere when a radio signal propagates in the neutral atmosphere and such refractivity is primarily due to the presence of dry gases and water vapor present over there. Though the troposphere delay or troposphere zenith delay (TZD) can be expected based on the earlier datasets, however, during severe weather events (heavy rainfalls and hail storms), the temporal and spatial variability of water vapor distribution is so dynamic that it is impossible to predict. Therefore, further case studies are needed to assess TZD under various severe weather events over tropical stations where the water vapor distribution is highly dynamic. This study presents tropospheric delay (both wet and dry) calculated during severe weather events over three Indian cities. A comparison of troposphere delay during severe weather conditions with normal days is made. Diurnal variation of ZHD and ZWD using ECMWF data is also presented.

A. Narendra Babu

Department of ECE, Lakkireddy Bali Reddy College of Engineering, Mylavaram, Affiliated to JNTUK, Mylavaram, India

P. S. Brahmanandam (✉) · G. Uma · K. Pushpa

Dept. of Physics, Shri Vishnu Engineering College for Women (A), Vishnupur, Bhimavaram, India
e-mail: dranandpotula@svecw.edu.in

G. Uma

e-mail: drumagouthu@svecw.edu.in

K. Pushpa

e-mail: pushpak@svecw.edu.in

K. Srinivas

Vishnu Institute of Technology (A), Vishnupur, Bhimavaram, India
e-mail: viceprincipal@vishnu.edu.in

A. Praneetha

Department of CSE, Lakkireddy Bali Reddy College of Engineering, Mylavaram, Affiliated to JNTUK, Mylavaram, India

Keywords Neutral atmosphere · Radiowave bending · Severe weather events · Dry and wet delays

1 Introduction

The Earth’s tropospheric delay is a crucial error resource while doing analysis of space geodetic techniques (which operate at microwave frequencies), including global navigation satellite systems (GNSS) and very long baseline interferometry (VLBI) [1]. The troposphere is that the lowest part of the Earth’s atmosphere up to about 80 km altitude, which consists of a mix of several gases with different proportions. The propagation of the microwave signal through this layer depends on the temperature, pressure, and water vapor [2] present over. Figure 1 shows how a typical radio signal gets refracted (bending) in neutral atmosphere and therefore the curvature in actual path is due to the presence of refractive gradients in neutral atmosphere.

The troposphere delay is the average delay, which can be expressed as

$$TZD = ZHD + ZWD$$

where ZHD is the zenith hydrostatic delay, which contributes 90% of TZD and ZWD is the zenith wet delay, which contributes around 2–20% of TZD that affects the propagation of radio signals by the movements of water vapor molecules and is largely related to the water vapor concentration.

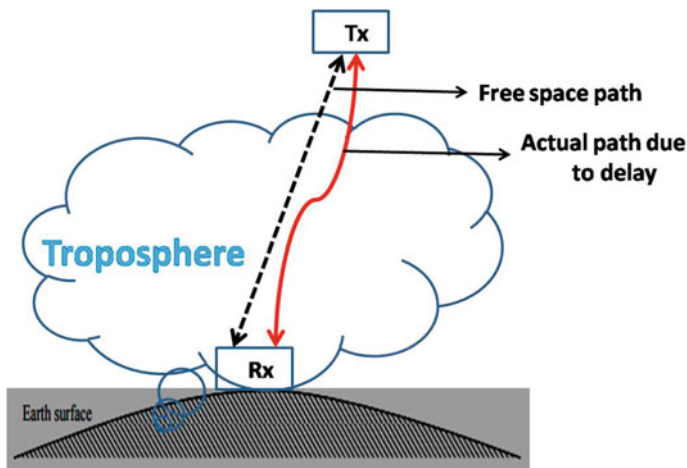


Fig. 1 Depicting the delay caused to a radio signal by troposphere constituents, which leads to a bending known as troposphere delay

The organization of this paper is that the literature survey is included in Sect. 2. Section 3 will have adopted methodology to calculate both troposphere delays, while results and associated discussion are included in Sect. 4. Section contains conclusion which follows acknowledgements.

2 Literature Survey

By effectively using meteorological parameters, several parameter troposphere delay models have been proposed such as Hopfield model, Saastamoinen model, and Black model [3–5]. On the other hand, several empirical troposphere delay models were proposed with the aid of the National Centers for Environmental Prediction (NCEP) and ERA5 reanalysis produced by the ECMWF [6–8].

3 Methodology

Both ZHD and ZWD are calculated using the formula proposed by Saastamoinen [3]

According to Saastamoinen approach, ZHD can be defined as

$$\text{ZHD} = \frac{0.002277 * P}{1 - 0.00266 * \cos(2\phi) - 0.00028 * H}$$

where P is the surface pressure while ϕ and H are the latitude and height of the station.

Similarly, ZWD is calculated as follows

$$\text{ZWD} = 0.002277 * ((1255/T + 0.05) * e)$$

where T is temperature in Kelvin and e is the saturation water vapor pressure (hPa).

4 Results and Discussion

We have collected balloon-borne radiosonde data from Wyoming University, USA (<https://weather.uwyo.edu/upperair/sounding.html>). A quality check of radiosonde data has been done before being used in the analysis, and MATLAB software is used to draw two-dimensional plots presented in this research. We have used balloon-borne radiosonde data (Wyoming University data base), including pressure (hPa), temperature (Kelvin), and saturation water vapor pressure (hPa) to compute troposphere delay, both ZHD and ZWD. The station coordinates and computed delays

during severe and normal days are presented in Table 1. There is a significant difference in delays between days with severe weather and normal days, which implies that the troposphere delays are more during severe weather days.

It is known that hydrometeors during severe weather days would produce heavy rainfalls or snow, fog, and dust storms which are usually connected with a large amount of clouds. As a result, various particles during severe weather days could induce up to 3% of the delay compared to that caused by dry air (which leads to troposphere hydrostatic delay) and water vapor (which leads to troposphere hydrostatic delay) [9]. It can, therefore, be concluded that trans-troposphere signals would encounter profound bending during severe weather days.

One of the limiting parameters with the radiosonde instrument is the availability of continuous data [10]. According to the World's Meteorological Organization (WMO), radiosonde ascents will be planned twice daily, one at 0600 AM local time and another one at 0600 PM local time. So, it is not possible to show the diurnal variation of ZHD and ZWD using radiosonde data. Alternatively, we have relied on ERA5 to plot the diurnal variations of ZHD and ZWD. ERA5 is the fifth-generation ECMWF [11] reanalysis data for the global climate and weather.

Figure 2 shows the diurnal variation of ZHD and ZWD over Hyderabad in universal time (UT). The magnitudes associated with ZHD (ZWD) are found to be between 2.133 m and 2.134 (0.0222 and 0.0228). Both delays show a significant diurnal variation. For example, ZHD shows a peak at 0600 LT (Local Time = UT + 0530 h) and 1800 LT, and again at approximately 0000 h of the next days, while a minimum is noticed at 1200 LT. Similarly, ZWD shows similar diurnal variation, but with a few exceptions.

Figure 3 shows the diurnal variation of ZHD and ZWD over Delhi in universal time (UT). The magnitudes associated with ZHD (ZWD) are found to be between 2.218 m and 2.2193 (0.038 and 0.039). Both delays show a significant diurnal variation. For example, ZHD shows a peak at 0600 LT (Local Time = UT + 0530 h) and 1800 LT, and again at approximately 0000 h of the next days, while a minimum is noticed at 1200 LT. Similarly, ZWD shows similar diurnal variation, but with a few exceptions.

Figure 4 shows the diurnal variation of ZHD and ZWD over Chennai in universal time (UT). The magnitudes associated with ZHD (ZWD) are found to be between 2.280 m and 2.281 (0.0347 and 0.0352). Both delays show a significant diurnal variation. For example, ZHD shows a peak at 0600 LT (Local Time = UT + 0530 h) and 1800 LT, and again at approximately 0000 h of the next days, while a minimum is noticed at 1200 LT. Similarly, ZWD shows similar diurnal variation, but with a few exceptions.

The radio signal from a GPS satellite (located at ~ 22,500 km) in its travel from the GPS satellite altitude to the surface of Earth's surface will suffer further losses due to bending at the ionospheric altitudes, cycle slips (which are the losses of signal lock in the GPS receiver tracking), ionospheric irregularities [12], satellite elevation angles, and troposphere delays. It is worth mentioning here that Earth's ionosphere irregularities can be small-scale disturbances in the ionosphere, which disrupt radio waves by causing amplitude and phase scintillations [13]. Most losses at the ionospheric altitudes are known (based on the earlier values), and even those losses could

Table 1 Station coordinates and delays during severe weather and normal days

Station	Geo. Lat	Geo. Long	MSL (m)	Severe weather day			Normal day		
				ZHD (m)	ZWD (m)	TZD (m)	ZHD (m)	ZWD (m)	TZD (m)
Hyderabad	17.45	78.46	542	2.4	0.21	2.61	2.2	0.20	2.40
Delhi	28.30	77.10	237	2.4	0.25	2.65	2.1	0.25	2.36
Chennai	13.00	80.18	16	2.5	0.29	2.79	2.3	0.26	2.56

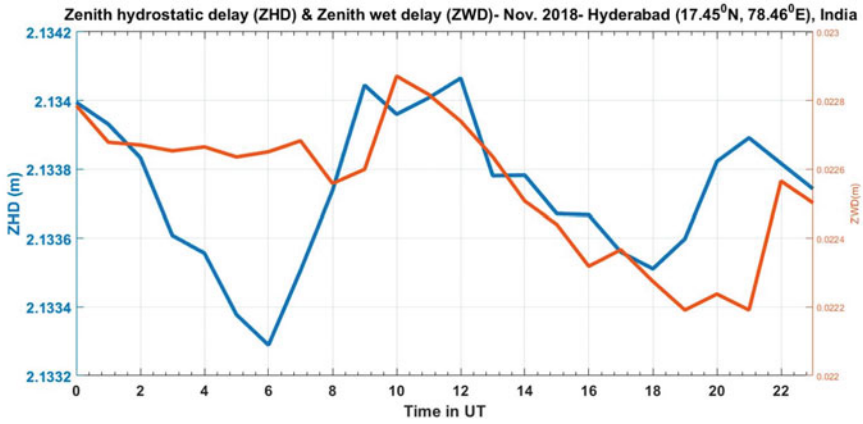


Fig. 2 Average diurnal variation of ZHD and ZWD over Hyderabad in November 2018

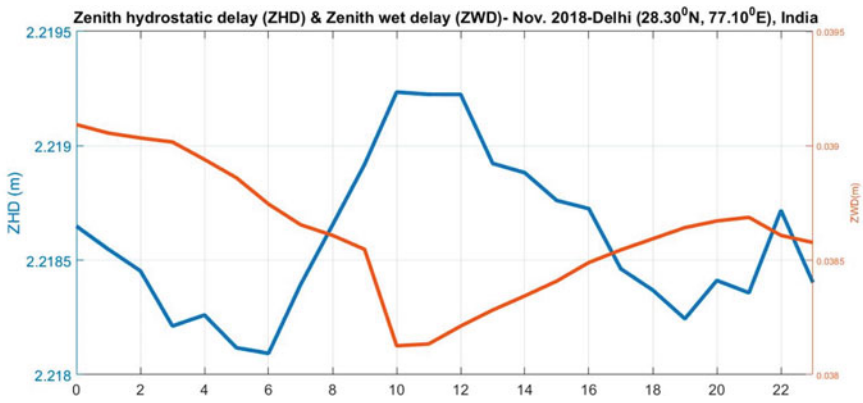


Fig. 3 Average diurnal variation of ZHD and ZWD over Delhi in November 2018

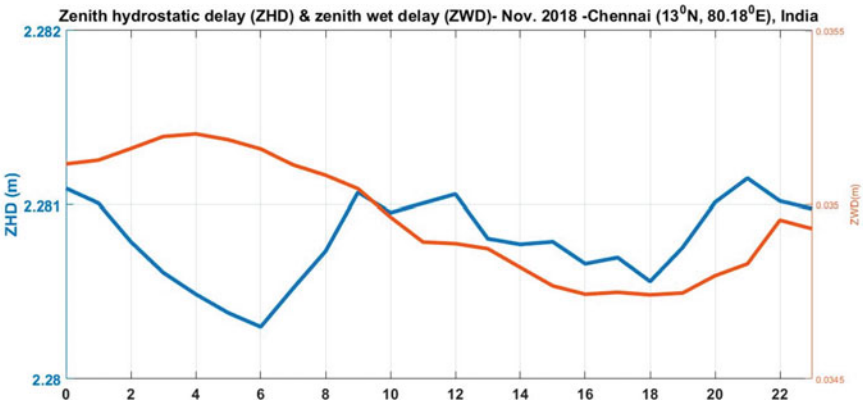


Fig. 4 Average diurnal variation of ZHD and ZWD over Chennai November 2018

be modeled accurately. On the other hand, the error due to dry components at troposphere altitudes could also be predicted so accurately. But the error due to wet parts cannot be modeled due to its highly dynamic nature. In this context, the present study assumes great significance as we have attempted to understand the troposphere delay during severe weather days and the diurnal variation at relatively different stations. It is, therefore, possible to know the impact of the wet component on troposphere delay over a few typical Indian stations with different environments.

5 Conclusion

The salient conclusions of the present study are

- (a) Comparative to arid stations (Hyderabad and Delhi), the coastal station (Chennai) witnessed higher magnitudes of troposphere delays, which could be due to higher moisture content over the coastal station.
- (b) Both troposphere delays show distinct diurnal variations.
- (c) For example, the diurnal variations of troposphere delay show a peak at 0600 AM and 0600 PM, whereas it shows minimum magnitudes at 1200 PM, which is primarily due to the fact that the rise in temperature would lower the moisture content.
- (d) Severe weather days show higher troposphere delays compared to delays during normal days, as expected.

Acknowledgements Dr. P. S. Brahmanandam (corresponding author) expresses his gratitude toward the Management of SVECW (A), Bhimavaram, India, for their logistic help that enabled him to carry out this research.

References

1. Nilsson T, Böhm J, Wijaya D, Tresch A, Nafisi V, Schuh H (2013) Path delays in the neutral atmosphere. In: Böhm J, Schuh H (eds) *Atmospheric effects in space geodesy SE-3*. Springer Berlin Heidelberg, pp 73–136
2. Askne J, Nordius H (1987) Estimation of tropospheric delay for microwaves from surface weather data. *Radio Sci* 22:379–386
3. Hopfield HS (1969) Two-quartic tropospheric refractivity profile for correcting satellite data. *J Geophys Res Space Phys* 74:4487–4499
4. Saastamoinen J (1972) Atmospheric correction for the troposphere and stratosphere in radio ranging satellites. In: Henriksen SW, Mancini A, Chovitz BH (eds) *The use of artificial satellites for geodesy*. <https://doi.org/10.1029/GM015p0247>
5. Black HD, Emsner A (1984) Correcting satellite Doppler data for tropospheric effects. *J Geophys Res Space Phys* 89:2616–2626
6. Li W, Yuan Y, Ou J, Li H, Li Z (2012) A new global zenith tropospheric delay model IGGtrop for GNSS applications. *Chin Sci Bull* 57:2132–2139

7. Mateus P, Catalão J, Mendes VB, Nico G (2020) An ERA5-based hourly global pressure and temperature (HGPT) model. *Remote Sens* 12:1098
8. Mateus P, Mendes VB, Plecha SM (2021) HGPT2: an ERA5-based global model to estimate relative humidity. *Remote Sens* 13:2179
9. Solheim FS, Vivekanandan J, Ware RH, Rocken C (1999) Propagation delays induced in GPS signals by dry air, water vapor, hydrometeors, and other particulates. *J Geophys Res* 104:9663–9670
10. Brahmanandam PS et al (2010) Observations of equatorial Kelvin wave modes in FORMOSAT-3/COSMIC GPS RO temperature profiles. *Terr Atmos Ocean Sci* 21:829–840
11. Dee DP et al (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q J R Meteorol Soc* 137:553–597. <https://doi.org/10.1002/qj.828>
12. Brahmanandam PS, Uma G, Rajababu A, Tejaswani AB, Chu Y-H, Srinivasa Rao A (2014) Typical comparisons between plasma bubble and blob distributions at different altitudes over the Indian sector using a unique combination of satellite-based observations—a case study. *Int J Remote Sens* 35(16):6173–6189. <https://doi.org/10.1080/01431161.2014.950765>
13. Uma G, Brahmanandam PS, Chu YH (2016) A long-term study on the deletion criterion of questionable electron density profiles caused by ionospheric irregularities—COSMIC radio occultation technique. *Adv Space Res* 57(12):2452–2463

Biomedical Implantable Wideband Antenna with Rectangular C-shaped Radiator



Pradyut Mohapatra and Sumit Kumar Khandelwal

Abstract Implantable antenna research has gotten a lot of attention recently in the discipline of biomedical engineering. An implantable antenna operating in MedRadio, MICS, WMTS, and ISM bands for biotelemetry applications has been presented in this paper. The antenna of dimensions $20 \times 30 \times 1.6 \text{ mm}^3$ has a rectangular C-shaped radiator patch, and an inverted rectangular C-shaped ground plane and CPW feeding are provided. The UWB antenna is placed inside a $70 \times 70 \times 70 \text{ mm}^3$ container filled with pork tissue for simulation of the real body environment. The maximum specific SAR value of 1.597 W/Kg is obtained at 401 MHz, which is considered safe by the IEEE standard safety guidelines. The antenna is designed and simulated with the help of ANSYS Electronics Desktop.

Keywords Biomedical · Implantable antenna · Specific absorption rate (SAR) · Wideband

1 Introduction

Biomedical telemetry has gotten a lot of attention as the healthcare industry has advanced. Implantable Medical Devices (IMDs) serve a critical role in healthcare monitoring via wireless telemetry, by enabling communication with an external unit and transferring data obtained by sensors inside the human body [1, 2]. These devices enable the real-time communication of a patient's physiological data to an external unit, such as glucose level, temperature, cardiac pulse, and so on. Patients' vital signs can now be monitored remotely, without the need for typical hospital checkups or follow-up routine checkups, owing to this technology [3, 4]. Implantable and wearable biotelemetry devices are two types of biotelemetry devices. The implantable

P. Mohapatra (✉) · S. K. Khandelwal
Delhi Technological University, Delhi, India
e-mail: pradyut10@yahoo.com

S. K. Khandelwal
e-mail: sumit.khandelwal@dtu.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
V. V. S. S. Chakravarthy et al. (eds.), *Advances in Signal Processing, Embedded Systems and IoT*, Lecture Notes in Electrical Engineering 992,
https://doi.org/10.1007/978-981-19-8865-3_16

173

antenna has several drawbacks. To have robust and continuous performance, various variables such as low power consumption, miniaturization, patient safety, lower operating frequency band, and multi-band or wideband operation must be considered for implantable sensors [5]. The characteristic that impacts the implant's overall performance is the selection of antenna structure in the design of the implantable sensor, which is a challenging task.

The implanted sensory units are inserted inside the human body and monitor heart rate, respiratory rate, blood pressure, temperature, etc., and transmit the data, which they have collected, to an external unit. Implantable antennas usually operate at the Medical Device Radio band (MedRadio) (401–406 MHz), Medical Implantable Communication Service (MICS) (402–405 MHz), Wireless Medical Telemetry Service (WMTS) (1.427–1.432 GHz), or Industrial, Scientific, and Medical (ISM) (433–434 MHz and 2.4–2.48 GHz) bands [6, 7]. To ensure the biocompatibility of implantable antenna, biocompatible encapsulation is used. The materials used for this purpose are polyetheretherketone (PEEK), Silastic MDX-4210 Elastomer, and Zirconia [8]. Due to its electromagnetic characteristics, Zirconia ($\epsilon_r = 29$) is a preferable alternative for bioencapsulation. It has a very low loss tangent and a very high permittivity value, which significantly reduces power loss by confining the near field of the antenna inside the capsulation.

In paper [9], an implantable circular polarized (CP) antenna has been illustrated which operates in the 2.40 GHz ISM band. To successfully minimize the antenna size and expand its CP bandwidth, the authors have used notched ring slots and shorting pin methodology. The percentage bandwidth based on simulation is 21.5% with a peak gain of -33 dBi. In [10], a compact dual-antenna setup with improved isolation is investigated for 2.40 GHz ISM band biotelemetry applications. The proposed antenna has fractional bandwidth of 8.5% and a gain of -28.3 dBi. Two spiral coupled resonators are employed to obtain broader bandwidth at lower bands in an implantable antenna [11]. This dual-band antenna has dimensions of $16.5 \times 16.5 \times 2.54$ mm³ and a gain of -30 dBi. In paper [6], a quad-band PIFA antenna is designed for biomedical purposes, which operates at MedRadio, ISM, and WMTS bands with a peak gain of -23 dBi. An antenna consisting of a monopole radiator with a ground plane (C-shaped) is proposed in [12]. The coupling between the ground plane and the sigma-shaped radiator enabled an additional resonant mode. To accomplish perfect impedance matching, a coplanar waveguide (CPW) feeding line is used. The antenna operates in the MedRadio band with a gain of -27.8 dBi.

Here in this paper, we have presented an innovative design to achieve the wideband nature for the implantable antenna to operate in MedRadio, WMTS, MICS, and both ISM bands. A rectangular C-shaped radiator and an inverted rectangular C-shaped ground plane have been considered for the antenna design. FR4 epoxy ($\epsilon_r = 4.4$) and Roger's RO3010 ($\epsilon_r = 10.2$) are considered for the substrate and superstrate of this antenna, respectively. To ensure biocompatibility, Zirconia coating can be used for bioencapsulation. The gain of an implantable antenna is extremely low due to the lossy nature of human body tissues. In our work, the gain is better than the papers discussed in the previous paragraph. Section 1 gives a brief introduction to biomedical implantable antenna and discusses several other works presented in this field and our

project. The antenna is modeled and simulated on ANSYS's simulation software which is illustrated in Sect. 2. Experimental results obtained from simulation and effects on antenna frequency bands by changing certain parameters are discussed in Sect. 3. This paper is then concluded in Sect. 4 followed by references.

2 Antenna Design

Here, we aim that our proposed antenna operates at all the biomedical frequency bands, i.e., MedRadio, MICS, WMTS, and ISM bands. With the help of ANSYS Electronics Desktop simulation software, an implantable antenna is designed which operates at the frequency bands mentioned above. For the simulation setup, multiple frequencies—401 MHz, 433 MHz, 1.427 GHz, and 2.4 GHz are selected. The antenna comprises a rectangular C-shaped radiator patch and an inverted rectangular C-shaped ground plane in a single layer, as shown in Fig. 1. A Roger's RO3010 superstrate and an FR4 substrate are used, as shown in Fig. 2. To attain perfect impedance matching, the CPW technique is used [13, 14]. In this paper, the $20 \times 30 \times 1.6 \text{ mm}^3$ antenna is positioned at the center of a $70 \times 70 \times 70 \text{ mm}^3$ container which is filled up with pork tissue to present an ideal body environment. The primary idea of antenna design is inspired by paper [12], where a monopole radiator and the ground plane are placed in a single layer. By increasing bends and current path in the radiator, the antenna works at lower frequencies. We aim to make the antenna work at every medical frequency band and to ensure that radiator has a smaller length since metal can be harmful to the body. Therefore, the antenna radiator is a very simple structure and the copper used is quite less. Table 1 displays the measurements of the proposed antenna structure.

3 Results

Gain, radiation pattern, efficiency, and SAR are all factors to be addressed while developing an implantable antenna. The designed antenna is simulated in HFSS software, and the results obtained are plotted. It is observed from Fig. 3 that the antenna behaves as an ultra-wideband (UWB) antenna as it resonates from 320 MHz to 3 GHz and beyond, so it covers the MICS band, MedRadio band, 433 MHz ISM band, WMTS band, and 2.4 GHz ISM band. As shown in Fig. 4, it is evident that peak gain is -14.6 dBi at 401 MHz, which is better in comparison with other antennas discussed in Sect. 1. Peak realized gain at 433 MHz, 1.427 GHz, and 2.4 GHz are -13.4 dBi , -9.2 dBi , and -10.8 dBi , respectively. Figures 5a and b depict the 2D radiation patterns.

The amount of electromagnetic radiation absorbed by human body tissue is described as the Specific Absorption Rate (SAR). For 1 g of human tissue, the average SAR must be less than 1.6 W/kg as given in IEEE C95.1-1999 guidelines.

Fig. 1 Front view of the antenna

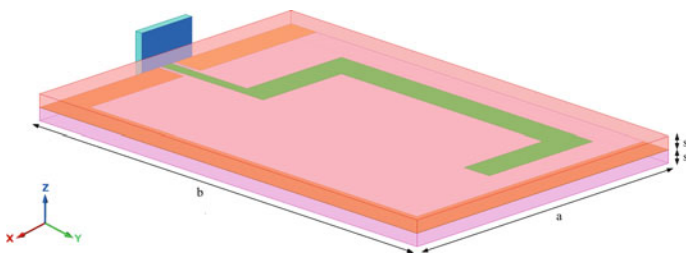
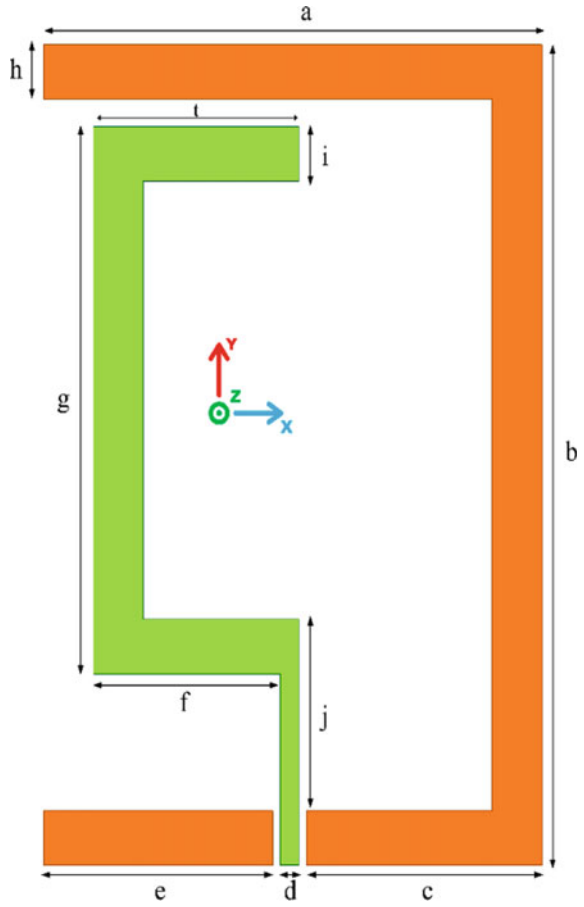


Fig. 2 Dimetric view of the antenna

Table 1 Dimensions of the designed antenna

S. No.	Parameters	Value (mm)
1	<i>a</i>	20
2	<i>b</i>	30
3	<i>c</i>	9.45
4	<i>d</i>	0.75
5	<i>e</i>	9.2
6	<i>f</i>	7.5
7	<i>g</i>	20
8	<i>h</i>	2
9	<i>i</i>	2
10	<i>j</i>	7
11	<i>s</i>	0.8
12	<i>t</i>	8.25

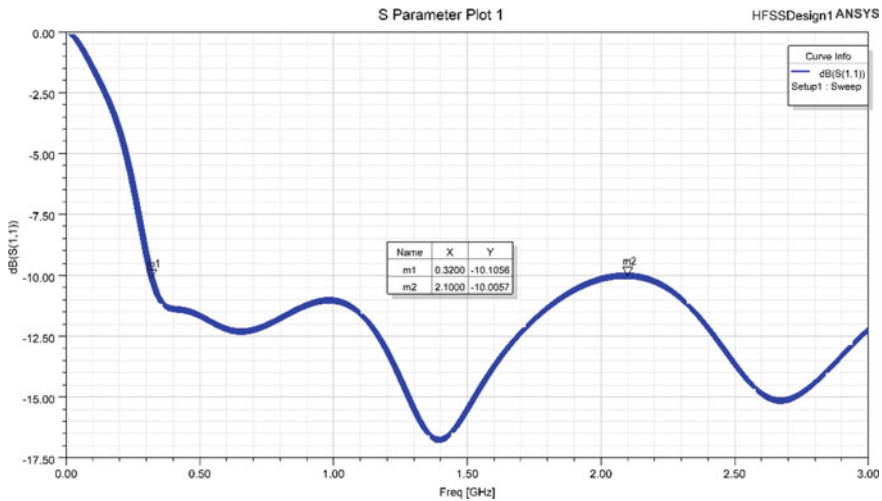


Fig. 3 Return loss versus frequency plot

It is observed from Fig. 6 that by providing an input power of 9.85 m W at 401 MHz satisfactory average SAR value is obtained for 1 g of pork tissue. Peak input power of 10.3, 19.5 and 18.2 m W are required at 433 MHz, 1.427 GHz, and 2.4 GHz, respectively, to obtain SAR less than 1.6 W/kg. The radiation efficiency versus frequency graph is plotted in Fig. 7.

We have observed in Fig. 8 that for values of parameter ‘a’ to be 0, 5, 10, and 15 mm, the antenna is not operating at all the frequency bands. Also, as depicted in Fig. 9 by changing the parametric values of ‘t’ to 0 and 4 mm, the antenna does not work at all frequency bands that we intend to.

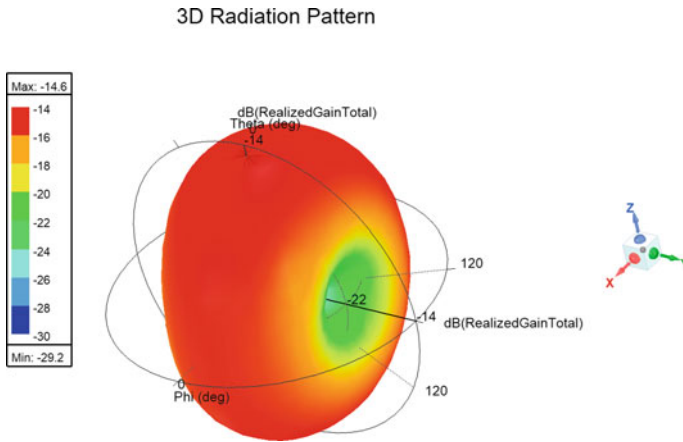


Fig. 4 3D radiation pattern

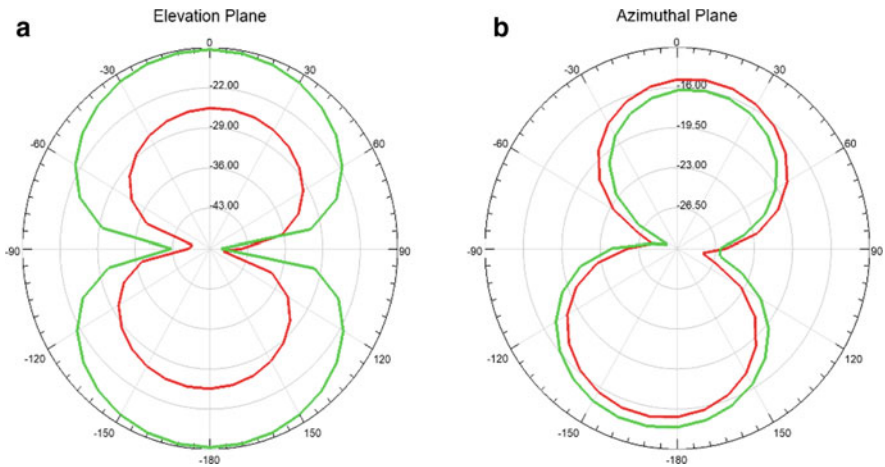


Fig. 5 a Elevation plane, and b Azimuthal plane

4 Conclusion

We have proposed a single layer, wide band, compact, and low SAR implantable antenna for biotelemetry applications. The UWB antenna comprises a rectangular C-shaped radiator and an inverted rectangular C-shaped ground plane. The antenna works at MedRadio, MICS, WMTS, and ISM (433 MHz and 2.4 GHz) bands with a wider bandwidth. At 401 MHz, a maximum gain of -14.6 dBi is obtained with a low SAR of 1.5997 W/Kg with an input power of 9.85 m W. By changing two parameters namely 'a' and 't,' the effect of parametric change on the antenna has been studied. In the future, the antenna can be fabricated and tested, and comparisons

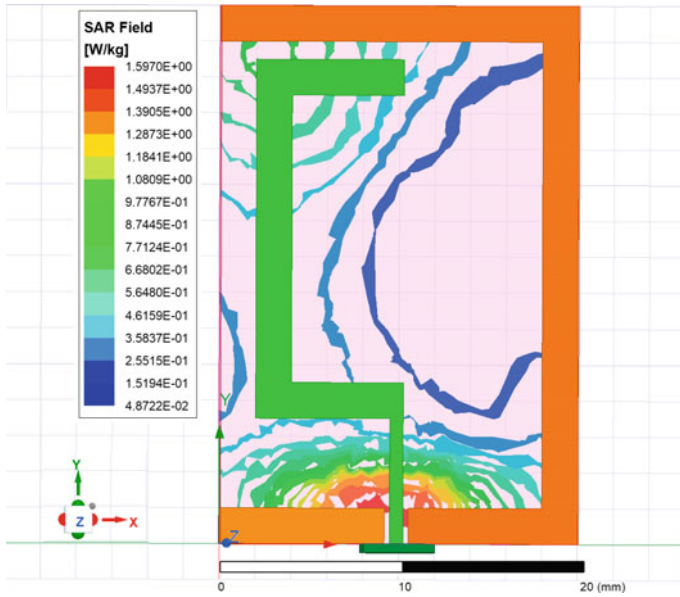


Fig. 6 SAR

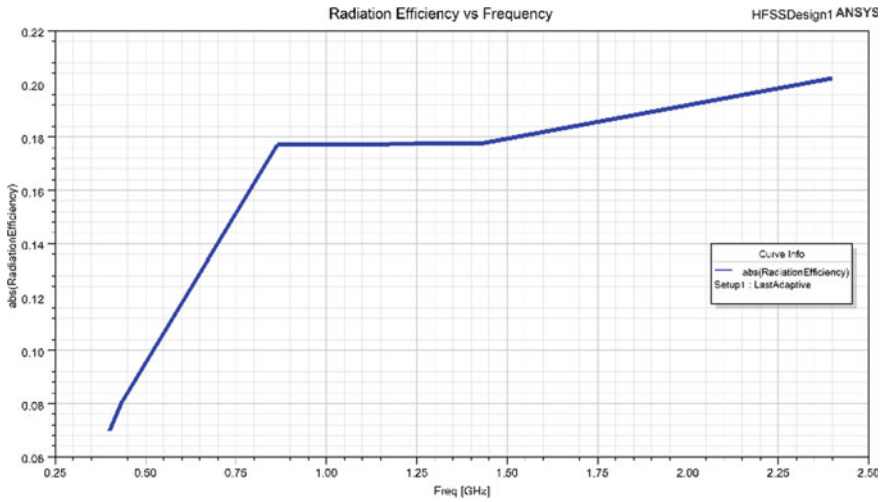


Fig. 7 Radiation efficiency versus frequency plot

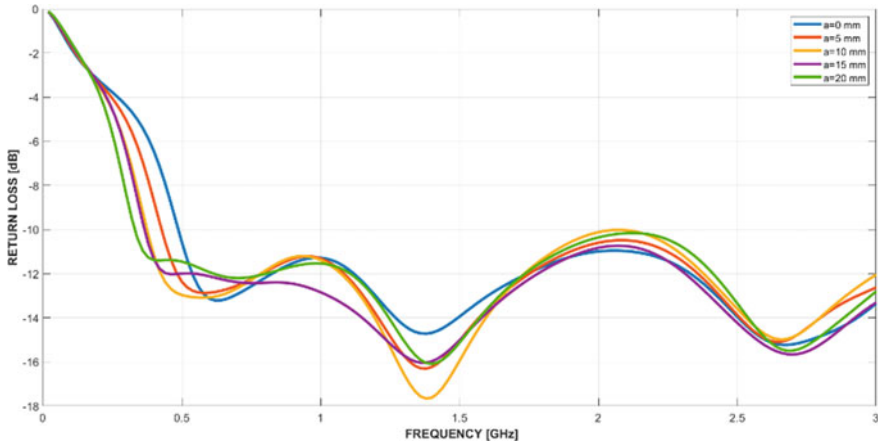


Fig. 8 S_{11} (dB) versus frequency (GHz) plot by changing 'a'

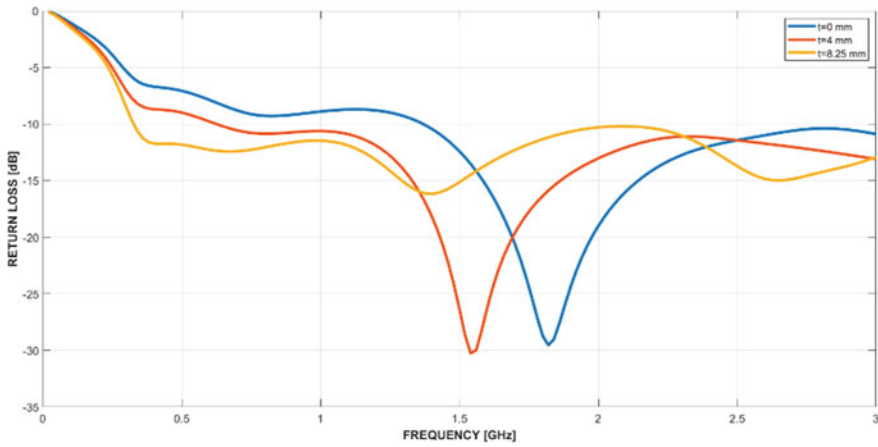


Fig. 9 S_{11} (dB) versus frequency (GHz) plot by changing 't'

can be made with the simulated results. Also, efforts can be made to improve the gain of the implantable antenna.


References

1. Kumar S, Shanmuganatham T (2020) Scalp—implantable antenna for biomedical applications. In: 2020 URSI regional conference on radio science (URSI-RCRS)
2. Xin Oh Z, Ho Yeap K, Shen Voon C, Chun Lai K, Chiong Teh P (2020) A multiband antenna for biomedical telemetry and treatments. J Phys Conf Series 1502:012013

3. Sudam Patil K, Rufus E (2022) Microwave antennas suggested for biomedical implantation. *Antenna Syst, IntechOpen*
4. Feng Y, Li Z, Qi L, Shen W, Li G (2022) A compact and miniaturized implantable antenna for ISM band in wireless cardiac pacemaker system. *Sci Rep* 12
5. Soliman M, Chowdhury M, Khandakar A, Islam M, Qiblawey Y, Musharavati F, Zal Nezhad E (2021) Review on medical implantable antenna technology and imminent research challenges. *Sensors* 21:3163
6. Wu C, Chien T, Yang C, Luo C (2012) Design of novel S-shaped quad-band antenna for MedRadio/WMTS/ISM implantable biotelemetry applications. *Int J Antennas Propag* 1–12
7. Malik N, Sant P, Ajmal T, Ur-Rehman M (2021) Implantable antennas for bio-medical applications. *IEEE J Electromagnet RF Microw Med Biol* 5:84–96
8. Kiourti A, Nikita K (2014) Implantable antennas: a tutorial on design, fabrication, and in vitro/in vivo testing. *IEEE Microw Mag* 15:77–91
9. Xia Z, Li H, Lee Z, Xiao S, Shao W, Ding X, Yang X (2020) A wideband circularly polarized implantable patch antenna for ISM band biomedical applications. *IEEE Trans Anten Propag* 68:2399–2404
10. Singh M, Ghosh J, Ghosh S, Sarkhel A (2021) Miniaturized dual-antenna system for implantable biotelemetry application. *IEEE Antennas Wirel Propag Lett* 20:1394–1398
11. Liu C, Guo Y, Xiao S (2012) Compact dual-band antenna for implantable devices. *IEEE Anten Wirel Propag Lett* 11:1508–1511
12. Tsai C, Chen K, Yang C (2015) Implantable wideband low-SAR antenna with C-shaped coupled ground. *IEEE Antennas Wirel Propag Lett* 14:1594–1597
13. Narmadha G, Malathi M, Kumar S, Shanmuganantham T, Deivasigamani S (2022) Performance of implantable antenna at ISM band characteristics for biomedical base. *ICT Exp* 8:198–201
14. Dubey R, Gupta V, Meshram M (2021) Implantable slot antenna for biomedical application. *Int J Adv Microw Technol (IJAMT)* 6:264–268

Raspberry Pi Alive Human Detection Robot Using PIR Sensor



E. V. Krishna Rao , B. Snehitha, J. Visweswara Rao, P. Mamatha, and M. Gowtham Chowdary

Abstract Modern technology has opened the way for superstructure buildings and houses, thus enlarging the chance of losing life during supernatural events and human made disasters. As reported by Urban Search and Rescue group, the probability of saving the life of a victim is treating them in the first 48 h of the rescue operation. In the destructed weather environment, sometimes the human beings are struck at various places. In order to identify the human's certain approach is recommended operating with the help of a robot by using sensor technology. It is proposed to develop the passive infrared sensor (PIR)-based robot to check whether the human being is alive or not using camera. A Bluetooth module is used to operate the robot. The ultrasonic sensor manages the path of the robot making it as autonomous. The group of sensors set up on Raspberry Pi such as ultrasonic sensor and PIR sensor to detect the livelihood of the human being. The video captured by the camera module is processed through machine learning by using HARR cascade algorithm. When the person is detected, the PIR sensor is triggered to check whether the person is alive or not, and then buzzer activates. When the person identified as alive, then the location along with the image is shared through GPS module to the receiver through mail.

1 Introduction

One of the most damaging events that occurred during the natural process of Earth is natural disasters such as floods, earthquakes, volcanic eruptions, tsunamis, and other metamorphism processes. Many people have been affected by such events. Because of elevated buildings and other man-made infrastructure, the loss of life during disasters is large in urban and industrial locations. On average every year, these disasters cause the death of nearly 60,000 people, resulting in 0.1% of all deaths

E. V. Krishna Rao (✉) · B. Snehitha · J. Visweswara Rao · P. Mamatha · M. Gowtham Chowdary
Lakireddy Bali Reddy College of Engineering (Autonomous), Mylavaram 521230, India
e-mail: krishnaraoede@yahoo.co.in

E. V. Krishna Rao
Jawaharlal Nehru Technological University Kakinada, Kakinada, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
V. V. S. S. Chakravarthy et al. (eds.), *Advances in Signal Processing, Embedded Systems and IoT*, Lecture Notes in Electrical Engineering 992,
https://doi.org/10.1007/978-981-19-8865-3_17

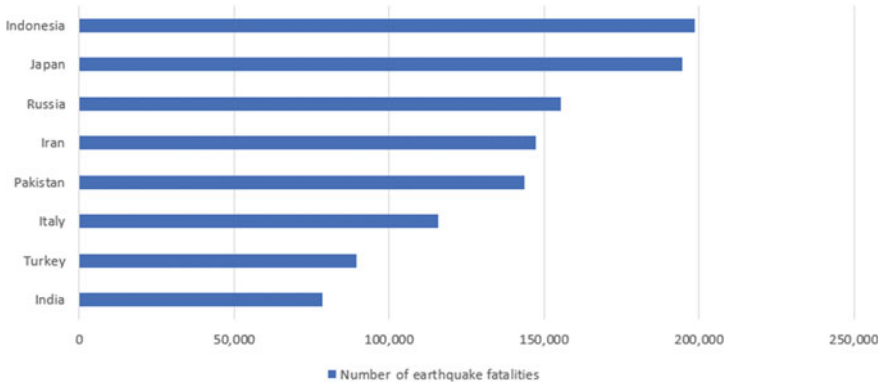


Fig. 1 Number of earthquake fatalities in 2021

worldwide [1]. Natural disasters claim the lives of a large number of individuals each year. As shown in Fig. 1, the number of earthquake fatalities around the world in 2021.

Generally, rescue humans (team) can't dive into a few areas of the earthquake-affected areas. This is done usually by the means of a humans and the skilled dog, regularly at very risky and volatile state, the rescuer might also end up as a sufferer and need to be rescued again [1]. Some of the victims lose their lives due to no longer being treated within time. An evaluation of information states that one-fourth of humans die because of no longer getting treatment at the proper time. Hence to make a rescue operation more secure and easier, a self-sustaining robotic has been proposed which discovers alive humans and which uses wireless communication with the rescue team.

Previously, equipped robots have to deal with a plenty of issues such as

- Communication cost is high as a consequence of transmitting a large number of images to the receiver.
- Over 25% of the communications between the wireless robot and the control unit were excessively high and thus useless. This eventually resulted in the robots and the operator losing contact, as the robots are semi-autonomous and are dependent on humans for their further movement. As a result, the functionality of the robot stops.
- The location of the person detected by the robot cannot be known by the operator, and the search and rescue operation will be pointless if the location of the alive person is not certain.

2 Literature Survey

Mario Di Castro's journal "CERNTAURO: A Modular Architecture for Robotic Inspection and Telemanipulation in Harsh and Semi-Structured Environments" [2], the motive of the project is to reduce the risk for humans in the unplanned and dangerous tasks in the industries or nuclear. The human-robot interface (HRI) is connected to the leader robot through a virtual private network (VPN) using Point-to-Point Tunneling Protocol (PPTP) for safe and secure communication. It is used for controlling or assigning various tasks to different robots according to the functionality of the tasks by analyzing the assigned task using the same human-robot interface. These robots are prepared to replace humans in the work environment this cannot save the life of the human.

Jorge Pena Querela's journal "Collaborative Multi-Robot Search and Rescue: Planning, Coordination, Perception, and Active Vision" [3], the methodology mainly focuses on semi-autonomous robots for search and rescue operations in saving the victims during calamities. We equip the robot with various components and sensors for multipurpose operations; this is particularly used for military applications and communicating the information to the authorized personnel and through a mobile ad hoc network (MANET) for wireless communication between robots. It is difficult to operate these robots in extreme weather conditions.

Giacomo Lunghi's journal "Multimodal Human-Robot Interface for Accessible Remote Robotic Interventions in Hazardous Environment" [4] suggested a technique for rescuing alive humans in a destructed environment within time. Here, radiation and temperature sensors are used to detect alive humans. It uses high-resolution instrument module for camera rotation compensation including tele-proprioception for finding the position of the person. The robot developed here semi-autonomous robot and needs a person for its operation.

Alvin Joseph's paper "Design of Human Detection Robot for Natural Calamity Rescue Operation" [5], the methodology used in creating an autonomous robot that reduces human deaths during natural hazards. Here, the robot model uses a radar sensor through which radio energy signals are sent. These radio signals are detected by the human body and will reflect which is received by the receiver. Automatically, the buzzer turns on indicating the alive human being. The problem here is the range is short.

Mojaharul Islam's paper "Search and Rescue System for Alive Human Detection by Semi-autonomous Mobile Rescue Robot" [6], an approach to creating a human detection robot for Urban Search And Rescue (USAR) is made by using joystick control which will facilitate user to drive the system easily. A system with two levels of human sensing is implemented. If joysticks being not particularly robust and can break easily if too much force is used on them.

Therefore, a system is proposed which uses an ultrasonic sensor that makes the robot autonomous along with a PIR sensor and a camera module to identify the alive human being. The system will detect radiation from the human and transmits a signal

to the control section in addition to the location of the robot. The robot will reply to the commands that have been received from the instructor.

3 ML In Search and Rescue Operations

The recent developments in image processing or recognition involve the detection of a person or an object by using a trained data set. Today, machine learning occupies a part of image processing. The currently available ML algorithms are supervised, unsupervised, and reinforcement learning. Analyses of the image which is collected by the camera module was done by using one of the machine learning (ML) algorithms and other deterministic processes. There has been a lot of talk about ML in the military defense along with machine learning applied to Search And Rescue (SAR) has the potential to save the lives of victims.

4 System Architecture

This section describes the architecture behind the alive human detection robot.

Raspberry Pi is a relatively low-cost, small-sized, and acts like a portable computer device with a little module that allows anyone to experiment with computing and learn how to program in uncomplicated languages like Python. It comes with 1 GB RAM, Quad Core 1.2ghz Broadcom BCM2837 64bit CPU. It acts as a barrier between sensors. The Hardware components such as sensors and motors are wired to this device, and this acts as a processing unit. As Raspberry Pi comes with 802.11 Wireless Local Area Network (LAN) Adapter, it can be connected to the Internet in a single click. The processing unit (Raspberry Pi) has 26 General Purpose Input Output (GPIO) which can be connected to 26 components.

In a specified range, a PIR sensor, as shown in Fig. 2, is used to identify animal or human motion. It consists of a pyroelectric sensor that can detect and recognize various degrees of infrared energy levels. The PIR sensor does not emit any energy into the environment. To activate the alarm, it needs to receive infrared radiation from the alive body present. Any object that has a temperature will emit infrared radiation. The human body's surface temperature ranges from 36 to 27 °C, with the majority of its radiant energy focused in the wavelength range of 8–12 microns. These degrees are sensed by PIR sensor, and the alive person is identified.

Transmitter which emits the ultrasonic sounds and receiver which receives the reflected sound waves are the major components that lie in the ultrasonic sensor. This sensor is usually suited to sense proximity this electric device is used to measure the distance of a target by measuring the time taken by the sound to travel between emission and receiving them back at receiver the distance is calculated by the formula $Distance (D) = 1/2 \text{ Time taken } (T) * \text{ speed of the signal } (C)$.

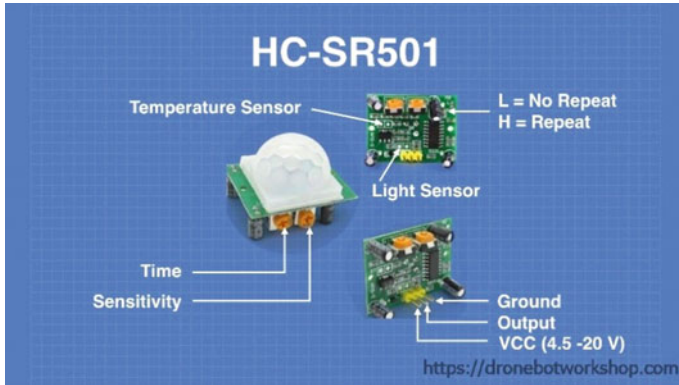


Fig. 2 Pin configuration of passive infrared sensor

The L293D board is an integrated chip (IC); it is a driver module used to drive the DC and step motor both the sides, and also the speed of the motor is controlled. This driver is capable of driving either four motors in same direction or two motors in both the directions. This motor driver module is mostly suited for the robotics.

5 Proposed Methodology

The Harr cascade algorithm is used to detect upper body using Raspberry Pi-based on PIR sensor and the block diagram of alive human detection robot with the components attached is shown in Fig. 3. The following steps are followed to detect alive human being, i.e., (a) computing through Raspberry Pi, (b) motion of robot, and (c) detecting alive human.

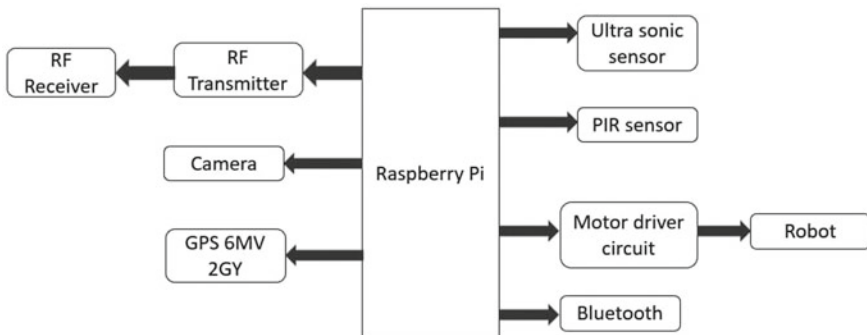


Fig. 3 Block diagram of alive human detection robot

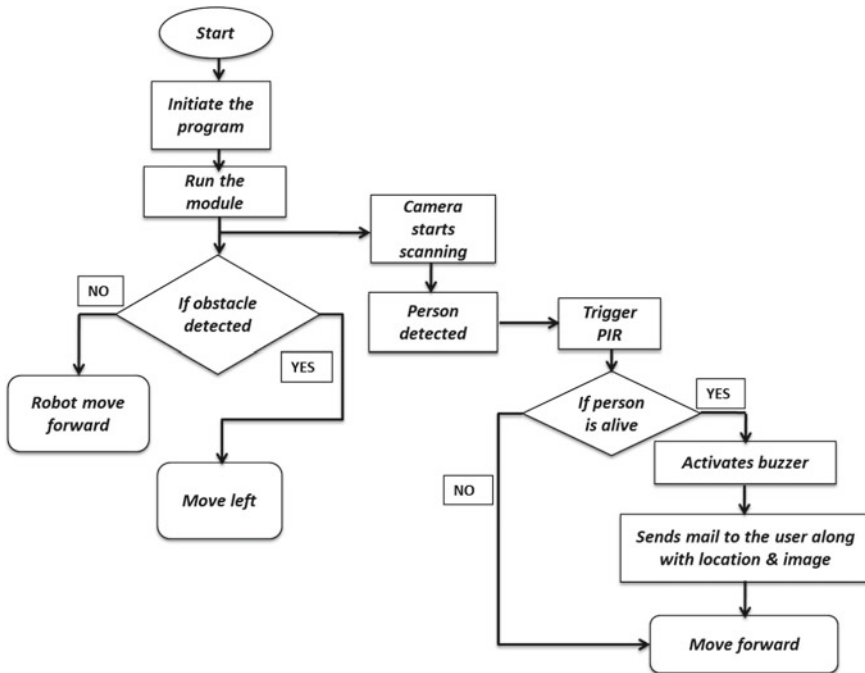
(a) **Computing through Raspberry Pi**

Raspberry Pi is a low-cost computer, just the thing is to connect a monitor, and it has standard keyboard and mouse ports. We can run many programming languages like ‘Scratch, Python, Java, JavaScript, etc.’ The difference between computer and Raspberry Pi is size, storage and memory, and connectivity. And Raspberry Pi has no screen and 1 GB is the average RAM of raspberry pi, we need to use an external SD card to store the data into system. The specifications of Raspberry Pi are Quad Core 1.2 GHz, Broadcom BCM2837, 64bit CPU, and 1 GB RAM.

(b) **Motion of robot**

For the movement of the robot, in this paper, the rover contains four wheels, two are placed at front side of the rover and rest is at back side of the rover. The front two wheels are connected the DC motor for the movement [1]. To make this robot autonomous, we are using an ultrasonic sensor. If there is any obstacle occurs on the way of robot, it turns left side to escape from the obstacle. The direction of the robot can be changed as per the instructions given in the code. The motion of the robot is under control of Raspberry Pi.

(c) **Detecting alive human**



After giving the power supply to the robot, Raspberry Pi activates all the components which are connected to the Raspberry Pi. The camera module is always on while

the robot is moving; if it detects any human being face or upper body of human being, then the PIR sensor will be activated. Detection of a human face or upper body of the human will be done using a Harr cascade classifier. A Harr cascade classifier is an object detection program using machine learning. A Harr classifier has few features such as line features, edge features and four-rectangular features, and these are used to detect the human face. PIR sensor is used as a motion sensor. It contains two pairs of sensors; those two are placed beside each other. If there is any differential between the two sensor signals, then the PIR sensor will detect it as human, then the buzzer will be triggered [5]. After detecting an alive human being, it will send a mail that contains an image of that person and coordinates of that particular location and location link will be sent. By using that location link and location coordinates, rescue operation will be easy.

6 Implementation and Working

Figure 4 depicts the structural flow chart of alive human detection robot. As soon as the power is turned on, the program is compiled, and the ultrasonic sensor begins to work, allowing the robot to move independently by detecting impediments. Simultaneously, the camera module activates to identify the presence of a body and capture an image [7–9]. The camera module scans for a body and activates the PIR sensor, which checks whether or not the person is alive. If the person is alive, the light at the receiver end begins to blink, and the mail is delivered with an attached image of the victim and the location. Rescuers rush to the location sent by the robot to save the victim.

The execution code is stored on a Raspberry Pi, which is a small computer that manages the functionality of all of the robot's components and sensors. The body is detected by three parameters: lower body detection, upper body detection, and full body detection, where the lower body, upper body, or full body of the person must be matched with the outline structured in the algorithm which helps in detecting the body by camera module, and PIR sensor detects the victim's condition by absorbing the thermal radiations emitted by the body, not only outlines and sensors, but also the instructions that must be carried out if no victim can be found in the current area. Clients would obtain desired results utilizing the device, which is a portable tool; as a result, the technique was simple to learn and use. This framework works really well.

7 Result

Figure 5 shows the hardware kit that is implemented using Raspberry Pi, ultrasonic sensor, PIR sensor, and camera module. Raspberry Pi can be considered as a single board computer that works on LINUX operating system. PIR sensor is

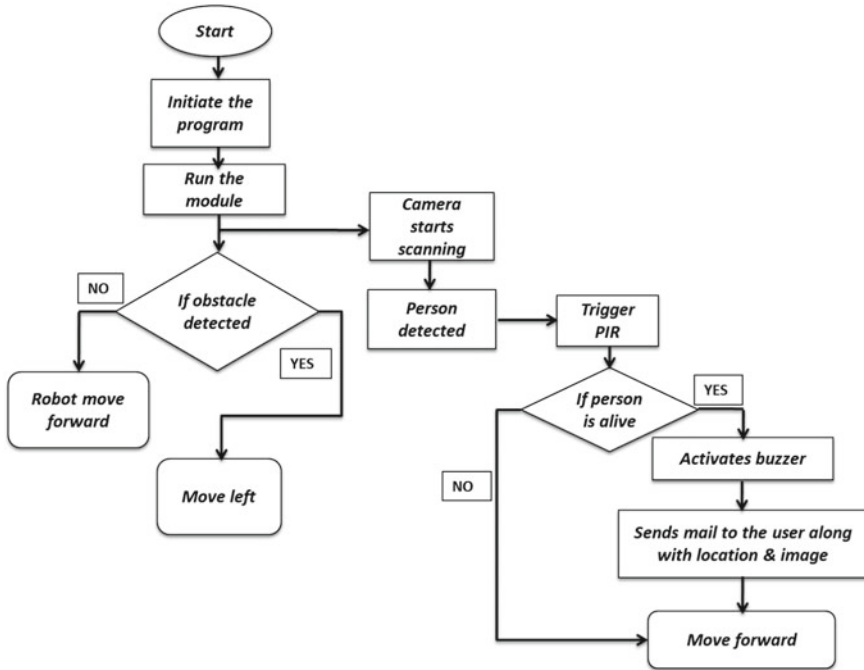


Fig. 4 Flow chart of alive human detection

passive infrared sensors use a pair of pyro electric sensors to detect heat energy in the surrounding. Ultrasonic sensors work by sending out a sound wave at a frequency above the range of human hearing. GPS modules contain tiny processors and antennas that directly receive data sent by satellites through dedicated RF frequencies. The Bluetooth modules can transmit and receives the data wirelessly. DC motors are the motors when kept in a magnetic field, a current-carrying conductor gains torque and develops a tendency to move. Wheels are used to the movement of the robot (Fig. 5).

After giving the power supply for the Raspberry Pi, the code starts running, and the robot moves forward. The screen on the laptop shows the distance traveled by the robot. When the person is detected by the camera module, it activates PIR sensor along with the text “Person Detected” and “Scanning PIR Sensor” is appeared on the screen.

The above figure shows the image is captured by camera module and is processed with the help of machine learning by using Harr cascade algorithm. When the person is captured by the camera sends the information to the user. PIR sensor scans whether the person is alive or not. If the detected person is alive automatically, the buzzer activates along with the text “Person is alive” will be appeared on the screen. Later, GPS 6MV 2GY module is used to access the coordinates of the location. In the end, the detected person image and coordinates of the location is received by the user through the mail.

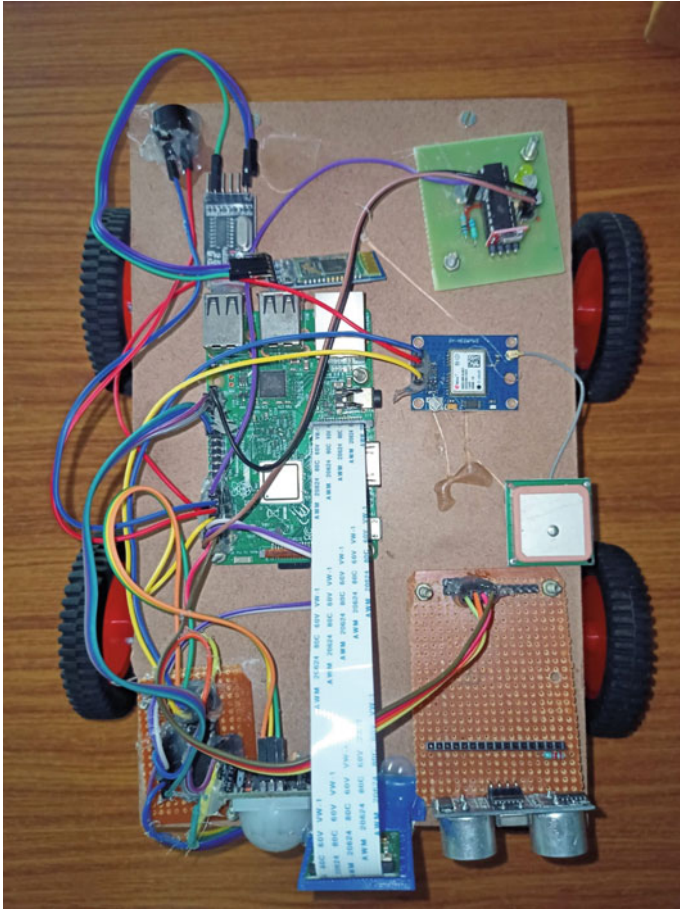


Fig. 5 Hardware kit

8 Conclusion

In this paper, it is developed PIR sensor-based robot to detect victim persons at the place where natural disasters are occurred by using camera module with HARR cascade algorithm. This sensor checks whether the person is alive or not. If the person detected is alive, then the buzzer activates, and LED is ON at the receiver side. At the end, when the person detected is alive, the GPS locates the coordinates of the person to be rescued, and the location along with the image of the person is received through mail at the receiver side.

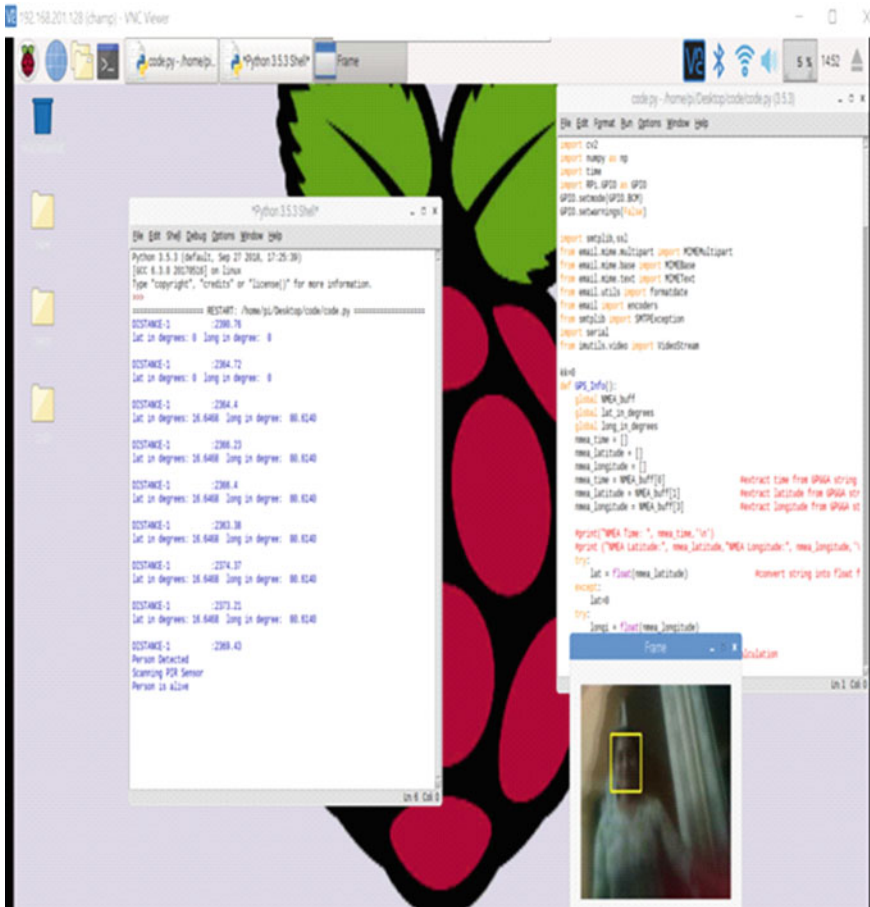


Fig. 6 Results showing on raspberry screen

References

1. Pawde A, Shaheen K, Nazreen K, Jameer K, Salman P, Alive human detector robot for rescue operation, no 39, pp 5–7. ISSN: 2320-8163
2. Di Castro M, Ferre M, Masi A (2018) CERNTAURO: a modular architecture for robotic inspection and telemanipulation in harsh and semi-structured environments. IEEE Access 6:37506–37522. <https://doi.org/10.1109/ACCESS.2018.2849572>
3. Queraltó JP, Taipalmaa J, Collaborative—robot search and rescue: planning, coordination, perception, and active vision. IEEE Access 8:191617. <https://doi.org/10.1109/ACCESS.2020.3030190>
4. Lunghi G, Marin R, Di Castro M, Masi A, Sanz PJ (2019) Multimodal human-robot interface for accessible remote robotic interventions in hazardous environments. IEEE Access 7:127290–127319. <https://doi.org/10.1109/ACCESS.2019.2939493>
5. Joseph A, Parmar R, Bagyaveerewaran V (2019) Design of human detection robot for natural calamity rescue operation. In: 2019 innovations in power and advanced computing technologies

- (i-PACT), pp 1–5. <https://doi.org/10.1109/i-PACT44901.2019.8959959>
6. Uddin Z, Islam M (2016) Search and rescue system for alive human detection by semi-autonomous mobile rescue robot. In: 2016 international conference on innovations in science, engineering and technology (ICISSET), pp 1–5. <https://doi.org/10.1109/ICISSET.2016.7856489>
 7. Satheesh Kumar V, Vasanthkumar C (2013) Alive human detection robot using image acquisition and processing 2(11) ISSN No 2277-8160
 8. Miyama S, Imai M, Anzai Y (2003) Rescue robot under disaster situation: position acquisition with Omni-directional sensor. In: Proceedings 2003 IEEE/RSJ international conference on intelligent robots and systems (IROS 2003) (Cat. No.03CH37453), vol 3, pp 3132–3137. <https://doi.org/10.1109/IROS.2003.1249638>
 9. Shwetha R, Chethan HK (2014) Automatic and manual controlled alive human detection robot during disaster management. Int J Technol Res Eng 1(11) ISSN (Online): 2347-4718

Design of Band Reconfigurable UWB Microstrip Patch Antenna for Cognitive Radio Application



Bammidi Deepa, V. Rajyalakshmi, Karri Sindhuja, Chellaboyina Lalitha Devi, Damodara Satya Sushma, Bommali Mothilal, and Thota Rajesh

Abstract The contemporary communication antennas require ultra-wideband (UWB) spectrum coverage for cost-effectiveness and efficient operation. Reconfigurability is the requirement of present communication systems. The project proposes broadband sensing ultra-wideband cognitive radio (CR) antenna in the frequency range of 3.75–12.75 GHz with a broad radiation pattern. The structure of octagonal-shaped patch antenna is proposed with UWB characteristics. The rectangular coplanar ground plane structure can be modified by incorporating two symmetrical horizontal T-shaped structures to increase the impedance bandwidth. Two PIN diodes are implanted to achieve frequency re-configuration with impedance matching at resonant frequencies. FCC UWB characteristics can be obtained by making suitable changes in the geometry of the ground plane. Four cases of switching of two PIN diodes give S11 minima at different operating bandwidths in the UWB range. The switches must be incorporated to convert a narrowband antenna into a broadband spectrum antenna. The proposed antenna can be used for cognitive radio applications in C-band and X-band.

Keywords Reconfigurable antenna · Cognitive radio antenna · Octagonal-shaped antenna · Non-conventional CPW feed

1 Introduction

As wireless communication technology advances, the demand for numerous wireless services in a single device has grown dramatically, and traditional antennas are no longer capable of meeting this new wireless communication system's needs. To meet the demands, an antenna that can change its direction features depending on the needs is being built. One such antenna is a reconfigurable antenna. Antennas with reconfigurable behavior can modify their behavior according to their needs.

B. Deepa (✉) · V. Rajyalakshmi · K. Sindhuja · C. L. Devi · D. S. Sushma · B. Mothilal · T. Rajesh
Anil Neerukonda Institute of Technology and Sciences, Visakhapatnam, India
e-mail: deepa.ece@anits.edu.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
V. V. S. S. Chakravarthy et al. (eds.), *Advances in Signal Processing, Embedded Systems and IoT*, Lecture Notes in Electrical Engineering 992,
https://doi.org/10.1007/978-981-19-8865-3_18

Frequency reconfigurable antennas enable frequency tuning and efficient spectrum utilization across specified frequency bands [1, 2].

As an intelligent system, cognitive radio (CR) technology can dynamically allow a given spectrum to be used additional users based on the major users' performance [3]. The study's goal is to produce a reconfigurable antenna by looking into the creation of a patch antenna with an octagonal shape, which hasn't been examined in depth for the development of CR antennas in the past [4–9].

Octagonal printed patch architecture is combined with a nontraditional coplanar waveguide (CPW) feed [10–13]. This geometry, which has been modified from the typical rectangular coplanar ground plane design, incorporates sleeve-like structures are symmetrical horizontal T-shaped sleeve-like structures that can enhance impedance bandwidth in printed antennas. By acting as an extra parasitic element that creates resonance, the sleeve can be utilized to make ultra-wideband (UWB) antennas. A frequency configurable ultra-wideband antenna based on a FR-4 substrate is demonstrated in this paper.

The suggested antenna may be modified via PIN diode switches to emit on a variety of bands while keeping a small footprint and high gain. The proposed antenna employs two PIN diode switches. The following is a breakdown of the paper's structure: The suggested switchable multiband antenna's design method and geometry are explained in each section and described the simulated analysis and observed consequences, bringing this investigation to a close.

2 Antenna Design

The proposed multiband frequency reconfigurable antenna's basic geometry and design theory are presented in this section. In the simulation, the suggested antenna's frequency and pattern are re-configured by utilizing lumped RLC components, i.e., the PIN diode is incorporated into the fabrication using the ON/OFF condition of the diodes, the antenna that was constructed may work on a variety of frequencies.

2.1 Design Procedure

The octagonal-shaped antenna recommended is 40 mm × 40 mm in dimension. The substrate is FR-4 and measures 1.6 mm thick with a loss tangent of 0.02 and a dielectric constant of 4.4. Figure 1a, b depicts the top and bottom perspectives of the suggested antenna, respectively. The S11 properties of an octagonal patch are investigated, but the matching is poor. The antenna can function over the FCC UWB spectrum if the structure is further adjusted, as illustrated in Fig. 1, to enhance bandwidth while simultaneously improving matching. In order to acquire an ultra-wide bandwidth, the ground plane geometry was altered to resemble a sleeve-like structure. To improve impedance matching, the ground structure is implemented by step structure. To get the desired frequency re-configuration, a T-pattern is placed on

the left and right portions of the ground. The antenna's S11 characteristics are not quite right, so they are tweaked a little further. The ground structure is further shown in Fig. 1 as a defective ground structure (DGS) improving impedance matching and bandwidth to allow the antenna to operate in the FCC ultra-wideband. HFSS optimization is performed to improve the results, and four S11 minima are notable as shown in Fig. 2, in the S11 features of the suggested antenna. To modify frequency and pattern, the suggested antenna uses a PIN diode. Table 1 provides a detailed breakdown of the planned antenna's dimensions.

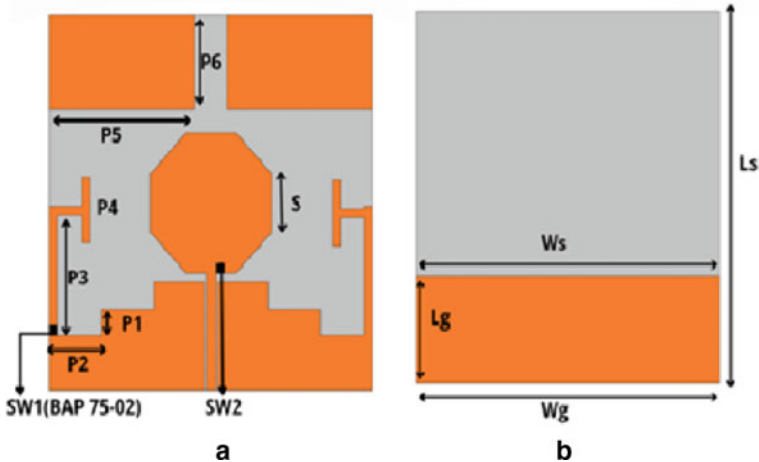


Fig. 1 a Proposed antenna—top view (orange color—copper, black color—PIN diodes, gray color—substrate) and b the proposed antenna—bottom view (orange color—copper) DGS

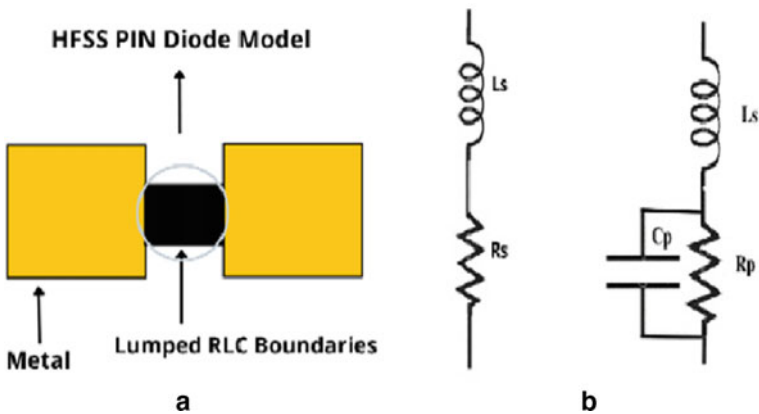


Fig. 2 a HFSS PIN diode model and b PIN diode equivalent circuit schematic in ON ($R = 1.9 \Omega$, $L = 1.5 \text{ nH}$ and OFF ($R = 1.9 \text{ k}\Omega$, $C = 250 \text{ fF}$, $L = 1.5 \text{ nH}$)

Table 1 Optimized parameters of the proposed antenna

Element	Parameter	Value (mm)
Substrate	W_s	40
	L_s	40
	h	1.6
Patch	S	6.2
	P_1	5
	P_2	8
	P_3	13.8
	P_4	7×1
	P_5	18
	P_6	10
Ground	L_g	11.5
	W_g	40
Feed	F	1.2

2.2 Frequency Re-configurability and Switching Techniques

The reconfigurable frequency is accomplished by turning the ON and OFF modes of each PIN diode in the suggested antenna. Because in the radio frequency (RF) band, they act as a variable resistor. Diodes with two pins (BAP 75-02) are utilized for switching. At their respective insertion sites, these PIN diodes can be both open and short-circuiting by modifying the antenna's effective resonant length, and as a result, reconfiguring the antenna's working frequency. Figure 2b demonstrates the PIN diode switch's proper circuits for both the OFF and ON states.

2.2.1 Antenna Design and Layout

The length and width of rectangular path antenna are calculated from below equations, where c is the velocity of light and ϵ_r is the dielectric constant of substrate.

1. Calculation of the width (W): $W = \frac{c}{2f_o\sqrt{\frac{\epsilon_r+1}{2}}}$
2. Calculation of effective dielectric (ϵ_{re}): $\epsilon_{re} = \frac{\epsilon_r+1}{2} + \frac{\epsilon_r-1}{2}\sqrt{1 + \frac{12h}{w}}$
3. Calculation of the effective length (L_{eff}): $L_{eff} = L + 2\Delta L$
4. Calculation of the length extension (Δ): $\Delta L = \frac{h}{\sqrt{\epsilon_r}}$
5. Calculation of actual length of patch (L): $L = \frac{c}{2f_r\sqrt{\epsilon_{re}}} - 2\Delta L$
6. Calculation of the ground plane dimensions (L_g and W_g): $L_g = L + 6h$; $W_g = W + 6$.

In Fig. 2b, it can be seen, it is merely the circuit of RL series with a low value resistor (" R_s ") and an inductor (" L_s ") in its ON state. It has an inductor " L_s " in parallel

with a high value resistor “ R_p ” and a capacitor “ C_p ” in the OFF state, comparable to an RLC circuit. There are three operational modes on the antenna in issue, each with its unique set of resonance frequencies. In Case 1 (SW_1 to $SW_2 = ON$), the antenna is capable of operating at frequencies of 2.2–8.1 GHz. The suggested antenna resonates at 3.7–8.3 GHz in Case 2 at SW_1 is ON, SW_2 is OFF. The antenna covers two separate bands of 2.2–7.3 and 8.1–11 GHz while in Case 3 at SW_1 is OFF, SW_2 is ON.

3 Results

Only one reconfigurable antenna can work in a variety of bands (broad and narrow) and switching scenarios, by altering the switching states of the PIN diodes for the cognitive radio system in numerous time frames. The S11 features for each situation are depicted in Fig. 3. In each sample, Table 2 lists the frequencies correspond to the S11 minima’s gain characteristics were examined. For instance, I have both SW_1 and SW_2 switched on. Two frequency bands, running from 2.2 to 8.1 GHz, can be shown in Fig. 3. In case 2, SW_1 is on and SW_2 is off, as shown in Fig. 4, the frequency band is 3.7–8.3 GHz. In the previous case, all of the switches were in the conducting state which is case 3 (SW_1 OFF, SW_2 ON). This results across a range of frequencies from 2.2 to 11 GHz. Figures 3, 4 and 5 depict the operating frequency in each of the three scenarios. It is worth noting that it uses the S-band (2–4 GHz), C-band (4–8 GHz) at the same time X-band (8–12 GHz).

When both SW_1 and SW_2 are in ON condition, the antenna resonates at 2.2 GHz with an S11 of -14.8 dB and a simulated bandwidth of 106 MHz, resonates at 3.7 GHz with an S11 of -23.2 dB and a simulated bandwidth of 437 MHz, resonates at 5.3 GHz with an S11 of -19.7 dB and a simulated bandwidth of 1120 MHz, resonates at 8.1 GHz with an S11 of -18.3 dB, and a simulated bandwidth of

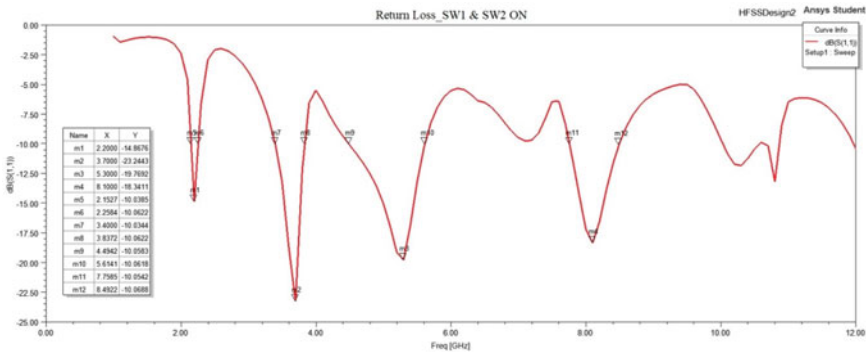


Fig. 3 Simulated S11 characteristics curve of a PIN diode in the first case (SW_1 and SW_2 are both turned ON)

Table 2 Comparison of performance of designed antenna under different switching cases of PIN diodes

Cases	Operating frequency (GHz)	The frequency at S11 minima corresponds to the gain
1. SW ₁ and SW ₂ ON	2.2–8.1	(a) 1.3 dB at 2.2 GHz (b) 2.5 dB at 3.7 GHz (c) 1.8 dB at 5.3 GHz (d) 3.1 dB at 8.1 GHz
2. SW ₁ ON SW ₂ OFF	3.7–8.3	(a) 2 dB at 3.7 GHz (b) 2 dB at 5.3 GHz (c) 3.3 dB at 8.3 GHz
3. SW ₁ OFF SW ₂ ON	2.2–7.3 8.1–11	(a) 2 dB at 2.2 GHz (b) 4.2 dB at 3.6 GHz (c) 2.4 dB at 5.5 GHz (d) 2.06 dB at 6.7 GHz (e) 2.9 dB at 7.3 GHz (f) 3.3 dB at 8.1 GHz (g) 5.4 dB at 10.4 GHz

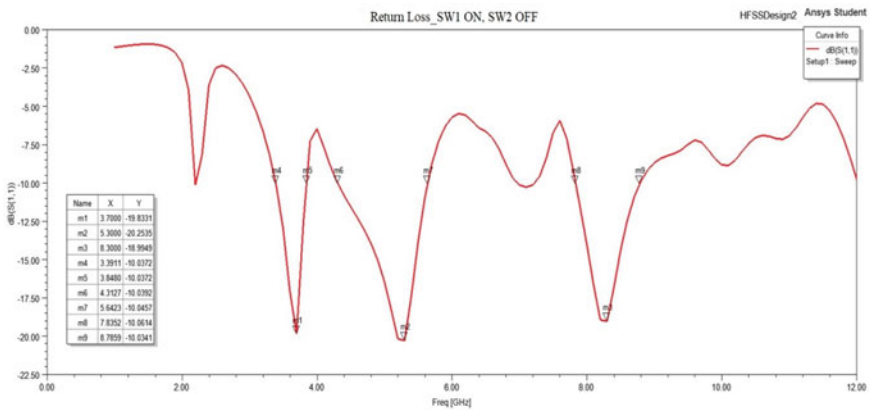


Fig. 4 Simulated S11 characteristics curve of a PIN diode in the second case (SW₁ ON and SW₂ OFF)

734 MHz. When SW₁ is ON and SW₂ is in OFF condition, the antenna resonates at 3.7 GHz with an S11 of -19.8 dB and a simulated bandwidth of 449 MHz, resonates at 5.3 GHz with an S11 of -20.2 dB and a simulated bandwidth of 1330 MHz, resonates at 8.3 GHz with an S11 of -18.9 dB and a simulated bandwidth of 950 MHz. Figure 3 shows the simulated results of the antenna when both diodes are ON. When both SW₁ is OFF and SW₂ is in ON condition, the antenna resonates at 2.2 GHz with an S11 of -19.1 dB and a simulated bandwidth of 139 MHz, resonates at 3.6 GHz with an S11 of -19.6 dB and a simulated bandwidth of 498 MHz, resonates at 5.5 GHz with an S11 of -20.4 dB and a simulated bandwidth of

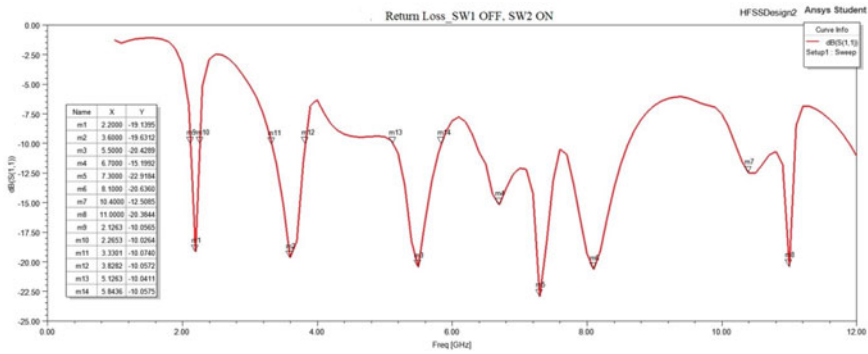


Fig. 5 Simulated S11 characteristics curve of a PIN diode in the third case (SW₁ OFF and SW₂ ON)

717 MHz, resonates at 7.3 GHz with an S11 of -22.9 dB, resonates at 8.1 GHz with an S11 of -20.6 dB.

For all resonant bands, there is a voltage standing waves ratio (VSWR) of less than 2, indicating that an antenna and the feed line connecting to it are matched. Figure 6 shows the VSWR for all switching modes and operational frequency ranges.

For all scenarios, the predicted surface current distribution has been investigated at various frequencies, with some results shown in Fig. 7. Higher-order current modes are activated at higher frequencies, causing the surface current density on the radiator to become unevenly distributed. It means that changing the switching conditions can affect the surface current distribution, causing adjustment in the antenna’s reflection coefficient.

An antenna’s simulated polar plot is 2.2 GHz, 3.6 GHz, 5.5 GHz, 6.7 GHz, 7.3 GHz, 8.1 GHz and 10.4 GHz at case 3 shown in Fig. 8 and its corresponding simulated peak gains of 2 dB, 4.2 dB, 2.4 dB, 2.06 dB, 2.9 dB, 3.3 dB and 10.4 dB,

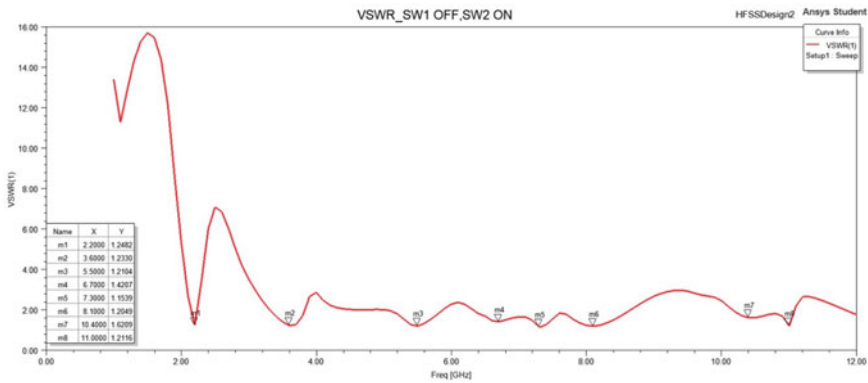


Fig. 6 Simulated VSWR characteristics when Sw₁ OFF and Sw₂ ON

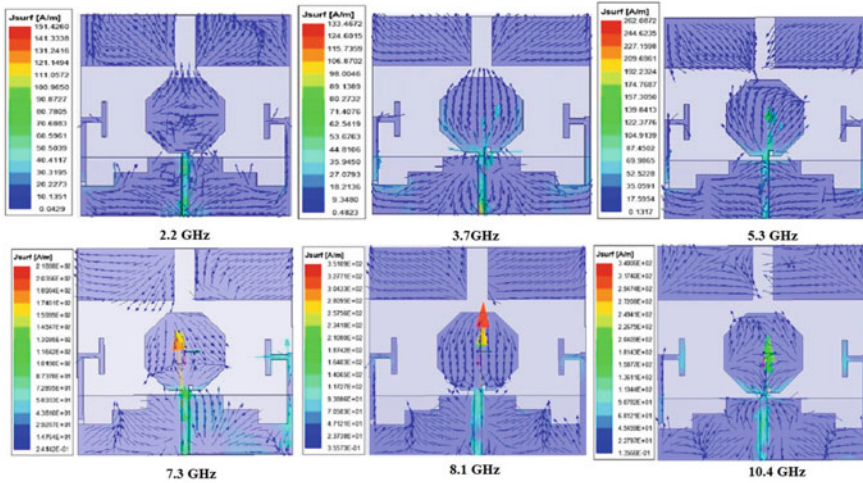


Fig. 7 Proposed model’s surface current distribution of different resonating frequencies at three different cases

respectively. Figure 9 shows the radiation pattern of the antenna under different resonant frequencies, when Sw_1 OFF and Sw_2 ON.

4 Conclusion

The designed antenna is a single-port reconfigurable antenna with a operating bandwidth of 2.2–11 GHz that is appropriate for cognitive radio applications. Its re-configuration is determined by the switching conditions of the diode. Two BAP 75-02 PIN diodes are employed between the sleeves of the patch to achieve the frequency re-configuration. The switches are arranged so that the narrow band communication antenna covers nearly the whole broadband spectrum. In this proposed antenna, three switching scenarios are investigated, and it is discovered that the diverse switching conditions cover practically the full band. Furthermore, the designed antenna can be used as a narrowband communication antenna in the ultra-wideband when (SW_1 ON and SW_2 ON) with operating frequencies of 2.2–8.1 GHz (SW_1 ON and SW_2 OFF) with operating frequencies of 3.7–8.3 GHz and (SW_1 OFF and SW_2 ON) with operating frequencies of 2.2–7.3 and 8.1–11 GHz. The increased use of mobile and radio applications is in need use of wide spectrum coverage and that problem can be addressed by this cognitive radio antenna. As a result, this antenna can be utilized for cognitive radio in the S-band (2–4), C-band (4–8) and the X-band (8–12).

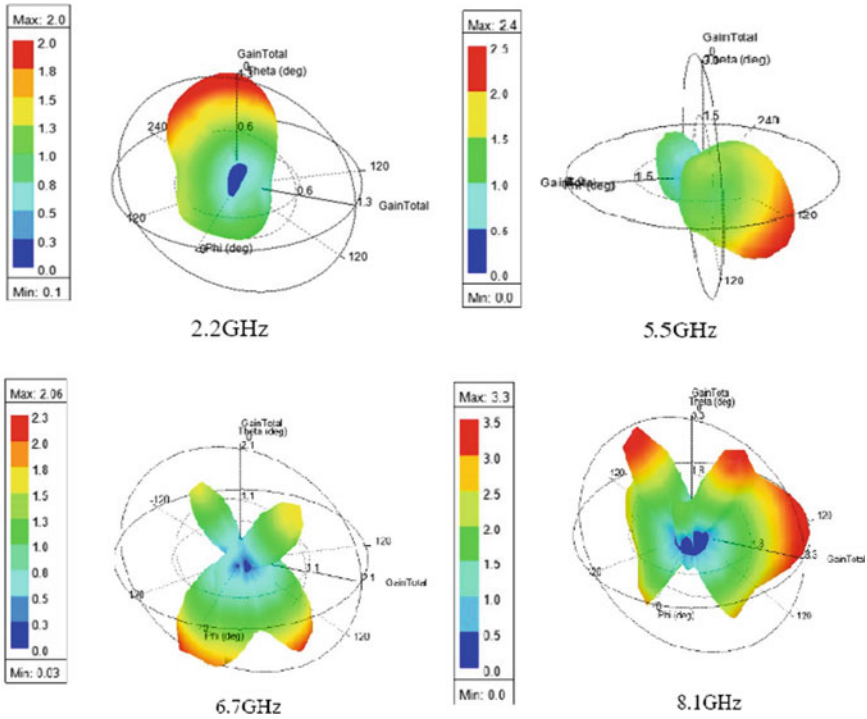


Fig. 8 Proposed model's simulated gain plot of different resonating frequencies at case 3 (SW₁ OFF and SW₂ ON)

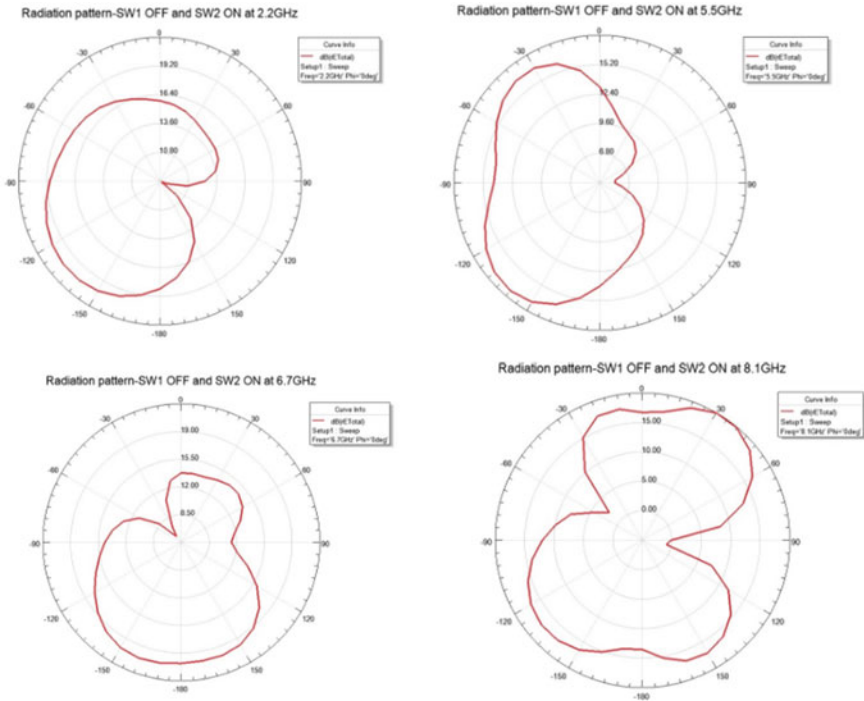


Fig. 9 Proposed model's simulated radiation pattern of different resonating frequencies at case 3 (SW₁ OFF and SW₂ ON)

References

1. Balanis CA (2016) Antenna theory analysis and design, vol TK7871.6.B354. Wiley, USA
2. Grag R (2000) Microstrip antenna design handbook, vol TK7871.6.M515. Artech House, Boston
3. Goswami A, Bhattacharya A, Dasgupta B (2020) Reconfigurable hexagon shaped printed antenna for cognitive radio application. *Int J RF Microwave Comput Aided Eng.* <https://doi.org/10.1002/mmce.22514>
4. Tawk Y, Constantine J, Christodoulou C (2016) Antenna design for cognitive radio. Artech House, Norwood
5. Majid HA, Rahim MKA, Hamid MR, Ismail MF (2014) Frequency and pattern reconfigurable slot antenna. *IEEE Trans Antennas Propag* 62:5339–5343
6. Tawk Y, Christodoulou CG (2009) A new reconfigurable antenna design for cognitive radio. *IEEE Antennas Wirel Propag Lett* 8:1378–1381
7. Khraisat YSH, Al-Tawalbeh N, Abdel-Hafez IA et al (2013) A hybrid reconfiguration antenna design for cognitive radio system combination of sensing and reconfigurable antennas. *J Electromagn Anal Appl* 5:328–332
8. Wu T, Li RL, Eom SY et al (2010) Switchable quad-band antennas for cognitive radio base station applications. *IEEE Trans Antennas Propag* 58(5):1468–1476
9. Al-Husseini M, Tawk Y, Christodoulou CG et al (2010) A reconfigurable cognitive radio antenna design. Paper presented at: IEEE antennas and propagation society international symposium, Toronto, Canada, pp 1–4

10. Ebrahimi E, Kelly JR, Hall PS (2011) Integrated wide narrowband antenna for multi-standard radio. *IEEE Trans Antennas Propag* 59(7):2628–2635. Abdulhameed AA, Alnahwi FM, Swadi HL et al (2019) A compact cognitive radio UWB/reconfigurable antenna system with controllable communicating antenna bandwidth. *Aust J Electr Electron Eng* 16:1–11
11. Bharadwaj SS, Sipal D, Yadav D, Koul SK (2020) A compact tri-band frequency reconfigurable antenna for TE/Wi-Fi/ITS application. *Prog Electromagn Res* 91:59–67. Chaouche YB, Messaoudene I, Benmabrouk I, Nedil M, Bouttout F (2018) Compact coplanar waveguide-fed reconfigurable fractal antenna for switchable multiband systems. *IET Microwaves Antennas Propag* 13:1–8
12. Asadallah FA, Costantine J, Tawk Y (2018) A multiband compact reconfigurable PIFA based on nested slots. *IEEE Antennas Wirel Propag Lett* 17:331–334
13. Ghaffar A, Li XJ, Seet BC, Awan WA, Hussain N (2019) Compact multiband frequency reconfigurable antenna for 5G communications. In: *Proceedings of the 29th international telecommunication networks and applications conference (ITNAC)*, Auckland, New Zealand, 27–29 Nov 2019. IEEE, Piscataway, NJ, USA, pp 1–3

Design of Multiband Frequency Reconfigurable Antenna for Wireless Applications



Deepa Bammidi, Gayatri Gorle, Venkatesh Sabbiseti, Reshma Gude, Govardhan Lingampalli, and Vinay Kumar Bodepu

Abstract To achieve multi-frequency or multiple radiation patterns, reconfigurable antennas are preferable in terms of size and complexity. In this design, a reconfigurable antenna that resonates in multibands for wireless applications, such as WLAN, Wi-Fi, LTE is designed. This design features a slotted wineglass-shaped radiating patch with stubs and a rectangular ring-type slot having two diodes (D_1 , D_2) on the perfect electric ground. The dielectric material used to design the proposed antenna is the Neltec substrate. The outer dimension of the antenna is $30 \text{ mm} \times 30 \text{ mm} \times 0.762 \text{ mm}$. Reconfigurability is achieved by changing the electrical equivalent circuit of p-i-n diodes (diode ON/OFF) which results in four different modes of operations. When both D_1 and D_2 are in the OFF state, the proposed antenna offers multiband operations at 2.52 GHz (S-band), 4.27 GHz (C-band), and 8.88 GHz, 10.80 GHz, 11.78 GHz (X-band), 14.04 GHz (Ku-band). Similarly, the antenna resonates at 1.43 GHz (L-band), 4.89 GHz, 6.48 GHz (C-band), 8.89 GHz (X-band), and 13.71 GHz (Ku-band) in the OFF–ON case. Meanwhile for the ON–OFF case at 3.95 GHz (S-band), 6.18 GHz (C-band), 8.77 GHz, 11.88 GHz (X-band), 14.23 GHz (Ku-band). Whereas when both the diodes are ON, the antenna resonates at 1.71 GHz (L-band), 5.04 GHz, and 6.17 GHz (C-band).

Keywords Reconfigurable antenna · Multiband · Multi-mode operation · Wireless applications · p-i-n diodes

1 Introduction

Nowadays, the wireless technologies, such as 5G, LTE, have been widely used. Thus, the technological expansion of modern communication highly demands wireless devices. So, it raises a requirement of antenna capable of working at various frequency bands. Instead of using multiple antennas operating at various frequencies,

D. Bammidi (✉) · G. Gorle · V. Sabbiseti · R. Gude · G. Lingampalli · V. K. Bodepu
Anil Neerukonda Institute of Technology and Sciences, Visakhapatnam, India
e-mail: deepa.ece@anits.edu.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
V. V. S. S. Chakravarthy et al. (eds.), *Advances in Signal Processing, Embedded Systems and IoT*, Lecture Notes in Electrical Engineering 992,
https://doi.org/10.1007/978-981-19-8865-3_19

207

it is desirable to use a frequency reconfigurable antenna that improves the performance. Hence, the demand for multi-mode reconfigurable antennas has increased over years to meet the requirements of modern devices. Multiband antennas are the ideal option since they reduce the number of antennas that must be installed in the communication system and can meet the needs of various wireless communication systems as discussed by Wi et al. [1]. So to accomplish this task, the best method is to design a Multiband Reconfigurable Antenna. Reconfigurability is the ability of an Antenna to change its characteristics such as radiation pattern, resonant frequency, in a controlled manner [2]. These are low-profile, lightweight, and low-power handling antennas and hence can be used in mobile, aerospace applications. It has a low fabrication cost and dual frequency can be easily made as discussed by Balanis [3]. Many literary works say that using T-shaped, inverted T-shaped, two E-shaped patch and radiating slots makes the antenna operate in multibands was discussed by Pawar et al. [4–6]. The use of a crescent-shaped resonator in a small new Co-Planar Waveguide (CPW) fed printed reconfigurable fractal antenna has been researched by Chaouche et al. [7, 8]. Alternatively, reconfigurable operation is controlled using fiber optic cables, optical switches, varactor diodes, and p-i-n diodes. This is discussed by Abraham and Shekhar [9]. p-i-n diodes are used to build a frequency reconfigurable tri-band antenna for LTE, Wi-Fi, and ITS as discussed by Bharadwaj et al. [10]. But the low-cost reconfigurability function is achieved by using p-i-n diodes at specific locations. Few multiband switchable frequency responses are achieved by changing the states of p-i-n diodes [11].

This paper aimed at designing a compact antenna which is capable of working at various frequency bands. The proposed design consists of a ground plane featuring a rectangular ring-type slot having two diodes and a slotted wineglass-shaped radiating patch with stubs on the substrate. Initially, a rectangular patch is designed and then the edges are truncated to form a triangular patch then two slots are inserted on it, and finally, two stubs are placed at the two edges of the base of the triangular patch.

2 Antenna Geometry Description

The design layout of the proposed antenna is illustrated in Fig. 1. The antenna is fabricated on a Neltec substrate having dielectric constant $\epsilon_r = 3.2$, loss tangent = 0.0024, and the overall dimension of the antenna is 30 mm \times 30 mm \times 0.762 mm. The antenna consists of a triangular patch (LP, WP) with two rectangular cuts (LPC, WPC) on it along with two stubs (L1, W1) at each edge of the base of the triangle and also a rectangular ring-type slot (LR, WR) containing two p-i-n diodes (MA4SPS552) on the ground plane. These diodes are responsible for the reconfigurability of the proposed antenna (Table 1).

Switching these p-i-n diodes ON/OFF results in 4 different modes with different surface current distributions which causes the antenna to resonate at multiple frequencies.

Electrical equivalent circuit of the diode,

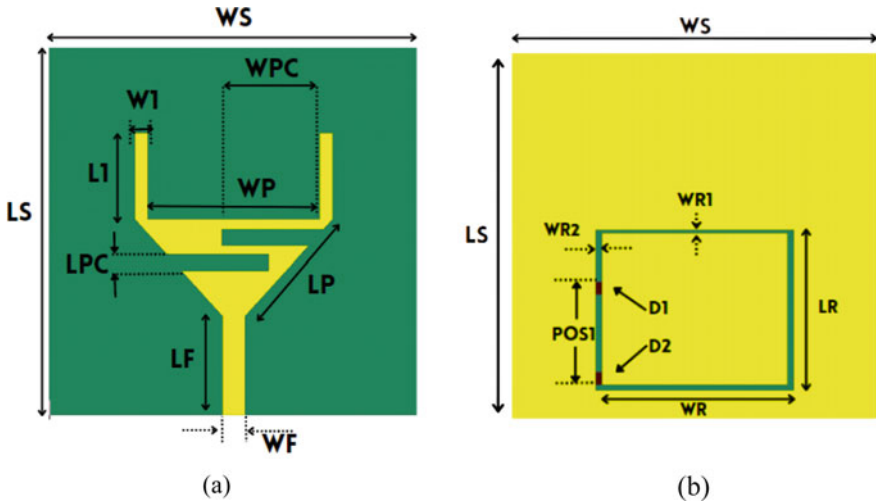


Fig. 1 Schematic a top and b bottom view of the proposed antenna

Table 1 Parameters of the proposed antenna

Parameter	Dimension (mm)	Parameter	Dimension (mm)
LS	30	WPC	8.3
WS	30	L1	1
LP	10.53	W1	7
WP	16	LR	16
LF	1.8	WR	13
WF	8	WR1	0.3
LPC	1.35	WR2	0.5

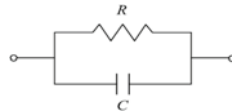
LS = length of the substrate, WS = width of the substrate, WP = triangular patch base length, LP = triangular patch slant length, LF = length of the feedline, WF = width of the feedline, LPC = rectangular cut length, WPC = rectangular cut width, L1 = stub length, W1 = stub width, LR = length of rectangular ring-type slot, WR = width of rectangular ring-type slot, WR1, WR2 = thickness of ring-type slot

Diode ON



$$R = 1.7 \Omega$$

Diode OFF



$$R = 40 \text{ k}\Omega \text{ and } C = 0.06\text{pF}$$

3 Results and Discussions

Various characteristics and parametric quantities such as S -parameter, surface current distribution, gain, and radiation pattern are measured to assess the proposed antenna's performance.

Figure 2 depicts the S -Parameter (S_{11}) for the different modes of operation of the proposed antenna. The simulated results in Fig. 2a illustrates that the antenna resonates in multibands at 2.52 GHz, 4.27 GHz, 8.88 GHz, 10.80 GHz, 11.78 GHz, 14.04 GHz, when both the diodes are in OFF state, whereas Fig. 2b shows that the antenna covers a wide range of bands at 1.43 GHz, 4.89 GHz, 6.48 GHz, 8.89 GHz, 13.71 GHz, when D_1 —OFF and D_2 —ON. Similarly, Fig. 2c, d operate at 3.95 GHz, 6.18 GHz, 8.77 GHz, 11.88 GHz, 14.23 GHz and 1.71 GHz, 5.04 GHz, 6.17 GHz for the cases D_1 —ON, D_2 —OFF and both diodes in ON state, respectively (Table 2).

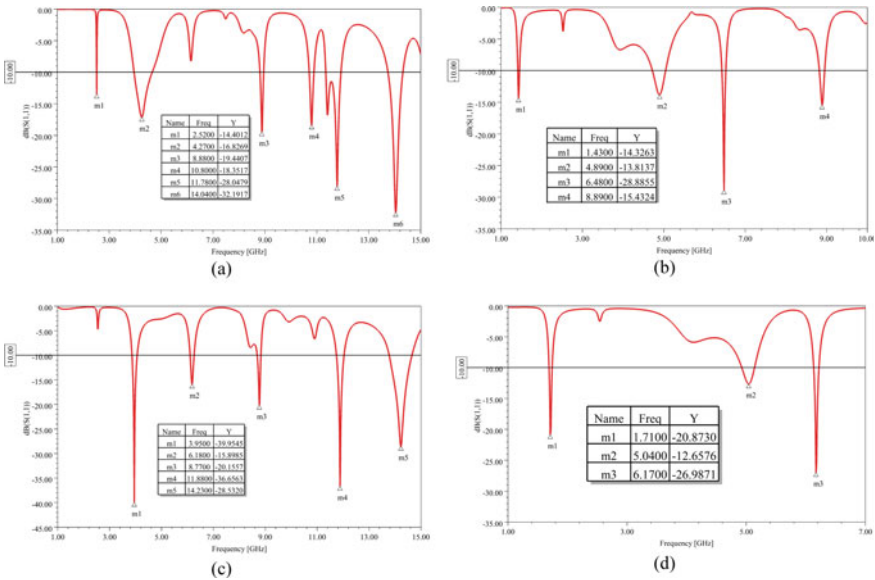


Fig. 2 Frequency versus S_{11} parameter plot for different modes: **a** both diodes OFF, **b** D_1 —OFF and D_2 —ON, **c** D_1 —ON and D_2 —OFF, **d** both diodes ON

Table 2 Various operating modes of the proposed antenna

Operating mode	Diode state (D_1, D_2)	Resonant frequency (GHz)	Operating bands
Mode 1	OFF, OFF	2.52, 4.27, 8.88, 10.80, 11.78, 14.04	S, C, X, Ku
Mode 2	OFF, ON	1.43, 4.89, 6.48, 8.89, 13.71	L, C, X, Ku
Mode 3	ON, OFF	3.95, 6.18, 8.77, 11.88, 14.23	S, C, X, Ku
Mode 4	ON, ON	1.71, 5.04, 6.17	L, C

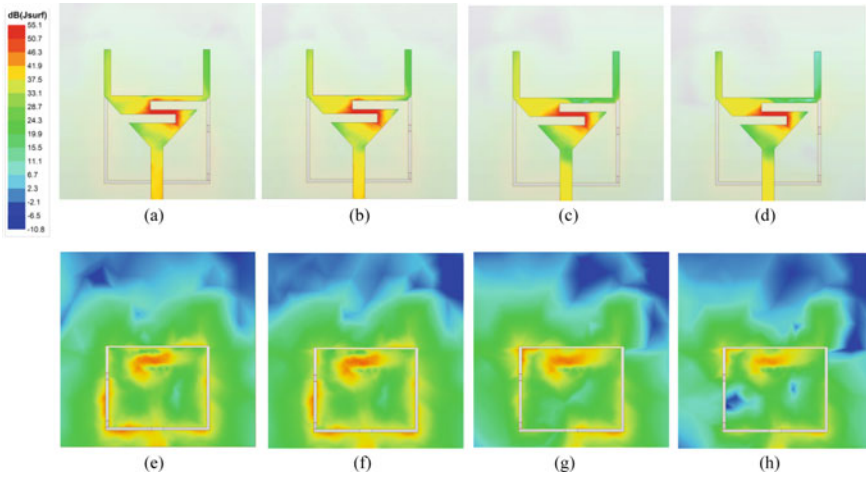


Fig. 3 Surface current distribution for different operation modes of the proposed antenna

Figure 3a, e shows the current distribution in top and bottom view of the proposed antenna, when both the diodes are OFF (mode 1). For D_1 OFF and D_2 ON (mode 2), the current distribution is observed as shown in Fig. 3b, f. When D_1 ON and D_2 OFF (mode 3), the current distribution is observed as shown in Fig. 3c, g. When both the diodes are ON (mode 4), the current distribution is observed as shown in Fig. 3d, h.

At $\theta = 0^\circ$, a maximum of 3.4 dB and a minimum of 2.9 dB gain is observed, when both the diodes are OFF and both the diodes are ON, respectively. For the mode D_1 —OFF, D_2 —ON and D_1 —ON, D_2 —OFF a gain of 3.1 dB and 3.2 dB is observed, respectively (Fig. 4).

The radiation pattern of an antenna is the graphical representation of the radiation properties as a function of space. That is, it tells how the antenna radiates/receives energy out into/from the space. For all the four cases omnidirectional radiation pattern is observed which is depicted in Fig. 5a–d.

4 Conclusion

In this paper, a multiband frequency reconfigurable antenna for wireless applications such as WLAN, Wi-Fi, LTE is designed. The proposed antenna can operate in multiple bands namely L, S, C, X, Ku—bands by adjusting the switching state of the two p-i-n diodes and even in each state without any modifications in patch. The designed antenna has the potential to be employed in multi-mode and wireless communication systems due to its simple design and compact size ($30 \times 30 \times 0.762 \text{ mm}^3$) and its good S-parameter values, gain, radiation properties for multiple

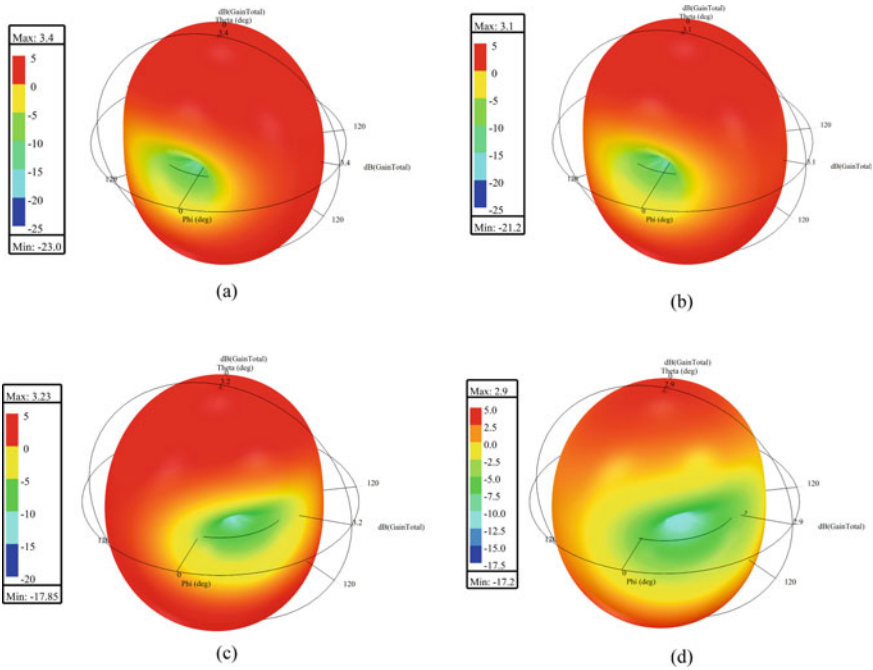


Fig. 4 Gain plot of the antenna for different operating modes: **a** both diodes OFF, **b** D_1 —OFF and D_2 —ON, **c** D_1 —ON and D_2 —OFF **d** both diodes ON

operating bands. When both the diodes are OFF, the antenna offers multiband operations at 2.52 GHz (S-band), 4.27 GHz (C-band), 8.88 GHz, 10.80 GHz, 11.78 GHz (X-band), 14.04 GHz (Ku-band). Similarly, antenna resonates at 1.43 GHz (L-band), 4.89 GHz, 6.48 GHz (C-band), 8.89 GHz (X-band), 13.71 GHz (Ku-band) when D_1 OFF, D_2 ON case. Meanwhile for D_1 ON, D_2 OFF case at 3.95 GHz (S-band), 6.18 GHz (C-band), 8.77 GHz, 11.88 GHz (X-band), 14.23 GHz (Ku-band). Whereas when both the diodes are ON, antenna resonates at 1.71 GHz (L-band), 5.04 GHz, 6.17 GHz (C-band). The proposed antenna can also be used in Bluetooth and ZigBee applications.

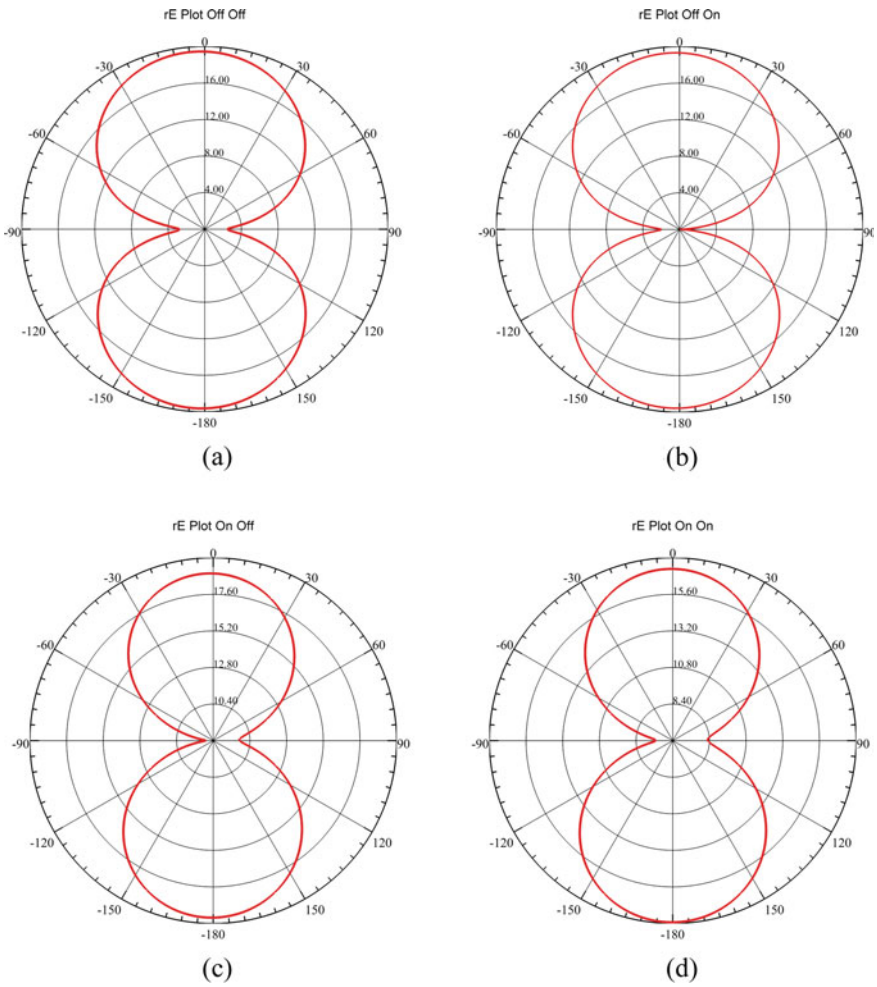


Fig. 5 Radiation pattern of the antenna at different operating modes: **a** both diodes OFF, **b** D_1 —OFF and D_2 —ON, **c** D_1 —ON and D_2 —OFF **d** both diodes ON

References

1. Wi H, Choi M, Mun B, Yoon Y, Lee H, Lee B (2015) A compact frequency reconfigurable antenna for LTE mobile handset applications. *Int J Antenna Propag* 2015, Article ID 764949
2. Awan WA, Naqvi SI, Ali WAE, Hussain N, Iqbal A, Tran HH, Alibakhshikenari M, Limiti E (2021) Design and realization of a frequency reconfigurable antenna with wide, dual, and single-band operations for compact sized wireless applications. *Electronics* 10:1321
3. Balanis CA. *Antenna theory analysis and design*, 3rd edn. Wiley
4. Pawar SS, Jadhav JB (2019) Design of multiband slot antenna for WLAN. *SSRG Int J Electron Commun Eng* 6(1):8–15

5. Saraswat RK, Kumar M (2021) Design and implementation of a multiband metamaterial—loaded reconfigurable antenna for wireless applications. *Int J Antennas Propag* 2021, Article ID 3888563
6. Wang L, Yu J, Xie T, Bi K (2021) A novel multiband fractal antenna for wireless application. *Int J Antennas Propag* 2021, Article ID 9926753
7. Hussain N, Ghaffar A, Naqvi SI, Iftikhar A (2022) A conformal frequency reconfigurable antenna with multiband and wideband characteristics. *Sensors* 22:2601
8. Chaouche YB (2018) Compact coplanar waveguide-fed reconfigurable fractal antenna for switchable multiband systems. *IET Microwaves Antennas Propag*
9. Abraham M, Shekhar H (2020) Social spider optimized design configuration of multiband reconfigurable antenna for 5G applications. *Wirel Pers Commun* 115:1161–1175
10. Bharadwaj SS, Sipal D, Yadav D, Koul SK (2020) A compact tri-band frequency reconfigurable antenna for LTE/Wi-Fi/ITS applications. *Progr Electromagnet Res M* 91:59–67
11. Jenath Sathikbasha M, Nagarajan V (2020) Design of multiband frequency reconfigurable antenna with defected ground structure for wireless applications. *Wirel Pers Commun* 113:867–892

Automatic Modulation Classification Under AWGN and Fading Channels Using Convolutional Neural Network



M. Venkata Subbarao, Beeram Keerthana, D. Ramesh Varma, Sudheer Kumar Terlapu, and G. Challa Ram

Abstract The advancement of 5G technology raises the need for wireless communication, increasing the number of customers who use it. The wireless communication band will become too crowded to support the increasing number of users in the future, resulting in a crowded spectrum. Furthermore, existing wireless technology has security flaws, limited capacity, and poor service quality. The implementation of cognitive radio (CR) technology is intended to improve the efficiency of the existing system and meet the bandwidth requirements of expanding users. Automatic modulation classification (AMC) is critical in CR technology. The objective of this research paper is to develop a modulation classifier by using a convolutional neural network classifier (CNN). The proposed classifier's performance is evaluated at various SNRs and training rates in AWGN and fading channels. The study revealed that the model is more efficient than existing models in identifying 11 digital and analog modulation classes with less computational complexity and higher accuracy.

Keywords AMC · CR · CNN · AWGN · Fading channels

1 Introduction

CR technology is a type of wireless technology in which the transmitter discovers vacant spectrum channels and uses them while avoiding occupied ones. This technology optimizes the frequency spectrum's utilization. Further, without any signal interference, the quality of service and spectrum efficiency are increased. In general, wireless spectrum is a finite resource that must be shared by both licensed and unlicensed users. Licensees will pay to get a dedicated frequency band for data transmission. All unlicensed users are given a tiny band that they can use for free. Some of the bands assigned to licensed users (military applications) are underutilized and stay unoccupied for the majority of the time. These idle channels are used by CR

M. Venkata Subbarao (✉) · B. Keerthana · D. Ramesh Varma · S. K. Terlapu · G. Challa Ram
Department of ECE, Shri Vishnu Engineering College for Women, Bhimavaram, India
e-mail: mandava.decs@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
V. V. S. S. Chakravarthy et al. (eds.), *Advances in Signal Processing, Embedded Systems and IoT*, Lecture Notes in Electrical Engineering 992,
https://doi.org/10.1007/978-981-19-8865-3_20

215

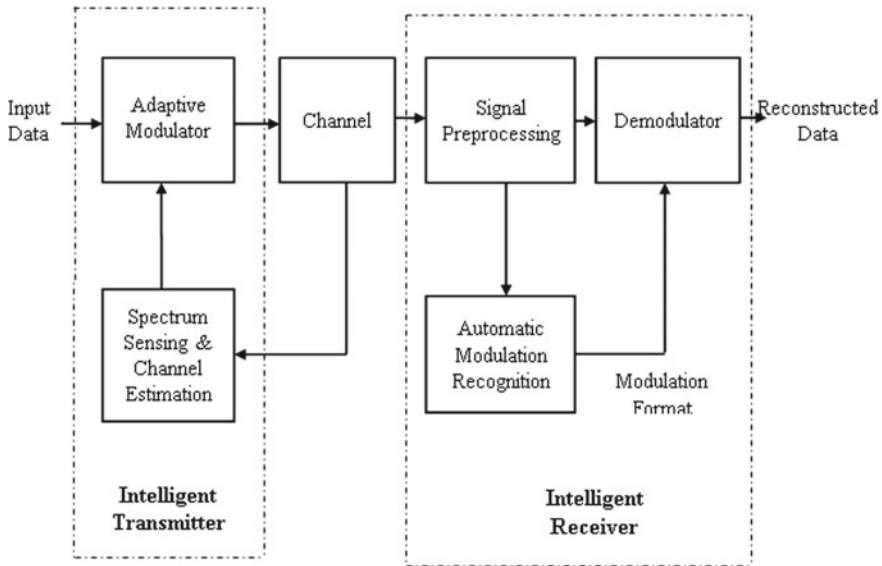


Fig. 1 Adaptive modulation system

technology to meet the bandwidth requirements of expanding licensed users. This can be accomplished through the use of spectrum sensing and channel estimation. The adaptive modulation system using CR technology is depicted in Fig. 1.

The modulation information of the signal is unknown to the demodulator because we are employing an adaptive modulator, which changes the modulation approach based on channel conditions [1]. As a result, we use AMC to assist the demodulator in determining the modulation type. AMC takes place between signal preprocessing and demodulation. AMC methods can be divided into three types: decision-theoretic approaches, feature-based approaches, and advanced techniques [2]. The statistical features of the received signal are used in the decision-theoretic method. They are divided into two types: the likelihood-based (LB) method and the distribution test approach. The feature-based (FB) approach is carried out in two steps: (a) feature extraction and (b) classification. Spectral-based features, transform features, statistical features, and other aspects of the incoming signal are extracted during the feature extraction stage [3]. These characteristics are used by the appropriate machine learning algorithm to classify various modulation kinds. Deep learning technology is being introduced in advanced ways. Deep learning algorithms are utilized to increase accuracy while decreasing AMC's computational complexity.

In this research paper, literature survey and research gap are discussed in the Sect. 2. Section 3 presents the proposed framework. The discussion of results is revealed in Sect. 4 and finally, the conclusions are stated in Sect. 5.

2 Related Work

In [4], Nandi and Azzouz modulation classification is performed with the help of features derived from instantaneous amplitude, phase, and frequency using the decision-theoretic algorithm. Around 12 classes are classified by the classifier with 94% accuracy at 15 dB SNR. Mühlhaus et al. [5] proposed an AMC for spatially multiplexed MIMO systems using fourth-order cumulants, and the classification is based on the likelihood ratio test. Juan-ping et al. [6] published AMR of digital communication signals using statistical and spectral feature set with an artificial neural network (ANN) classifier. The proposed classifier can classify up to 8 digital modulation classes under the AWGN channel. In [7], Hassan et al. focus on classifying various orders of digital modulation types including ASK, PSK, MSK, FSK, and QAM using wavelet transform and neural networks. Shen and Gao [8] proposed an AMR of digital signals based on locality preserved projection and the classifier is KNN. The classifier classifies digital modulation signals at SNR of 10 dB. In [9], Park et al. proposed an AMR system of digital signals using wavelet features and SVM. In the paper [10], Zhang et al. proposed an advanced technique like a hybrid parallel neural network for AMC.

From the study, we can understand that the decision-theoretic approaches suffer from computational complexity and the selection of the right feature set is the major limitation in feature-based approaches. Most of the authors had considered limited modulation techniques for classification. The majority of authors presented their analyses in AWGN channels, however only a few studies provide analysis in fading channels. Hence, this paper focuses on implementing a robust AMC classifier using a CNN and analyzing the performance of the classifier under AWGN and fading channels at different training rates and noisy conditions.

3 Methodology

The data set consists of eleven distinct analogue and digitally modulated signals under AWGN and Fading circumstances with varying SNR values. Later, these 1D signals are converted into 2D spectrograms using time-frequency transforms. Short-Time Fourier Transform (STFT) is used here. The spectrogram images are given as input to the CNN. Here, CNN plays a crucial role by extracting features and classifying the signals. CNN is trained at different training rates ranging from 50 to 90% with a step of 10%, and the CNN performance at different SNRs is analyzed.

Figure 2 represents the framework of the proposed method. This paper considered 7 digital modulation techniques (64QAM, 16QAM, CPFSK, GFSK, QPSK, BPSK, and 8PSK) and 4 analog modulation techniques (B-FM, DSB-AM, SSB-AM, PAM4). The dataset consists of 11 modulation techniques with 10,000 frames/class. Each frame consists of 1024 samples and the sample rate is 200 kHz. The center frequency of digital modulation classes is 902 MHz and analog modulation classes

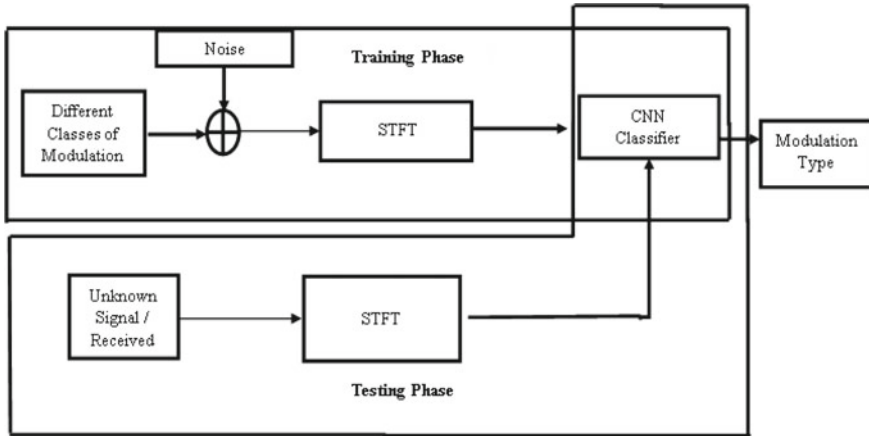


Fig. 2 Framework of the proposed approach

is 100 MHz. The modulated signals are passed through AWGN and fading channels. Both Rayleigh and Rician channels are applied under fading channels. Later, ID signals are converted into spectrograms by using Short-Time Fourier Transform (STFT). These images are fed into the input layer of CNN. Six convolution layers of CNN extract features from received signals. Each convolution layer is followed by a batch normalization layer, rectified linear unit (ReLU), and max-pooling layer. In the last convolution layer, the max-pooling layer is replaced with an average pooling layer. For the classification of signals, one fully connected layer and SoftMax activation were used. Figure 3 represents the proposed CNN classifier for AMC.

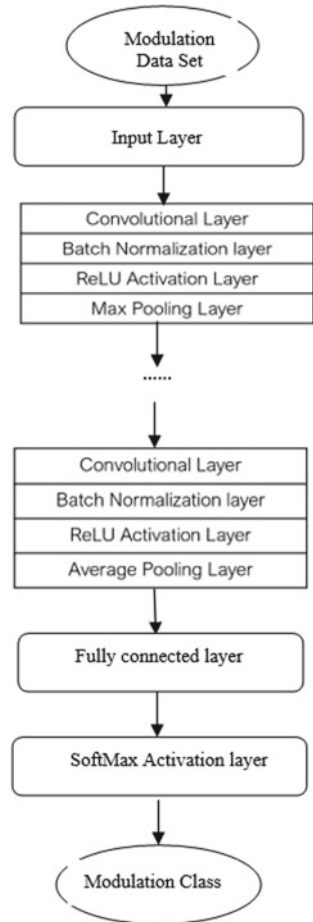
4 Results and Performance Characteristics of Protective Garment Antennas

The CNN model is first trained to classify analog modulation classes (4), followed by digital modulation classes (7), and finally trained to categorize both analog and digital modulation classes (11). In every case, the performance of CNN was examined under AWGN, Rayleigh, and Rician channels.

4.1 Classification of Analog Modulation Signals

The classifier is trained to classify 4 analog modulation classes at different training rates. Figures 4, 5, and 6 represent the performance of CNN at different training rates for AMC of analog signals under AWGN, Rician, and Rayleigh channels,

Fig. 3 Proposed CNN for AMC



respectively. The maximum classification accuracy is obtained at 90% training rate and 30 dB SNR. The maximum classification accuracy in the AWGN channel is 98%, 97.55% in the Rician channel, and 99.57 in the Rayleigh channel. Even at low SNRs, the classification accuracy is around 70%. The classification accuracy was improved with increasing SNRs.

Figure 7 represents the performance of three channels for analog signals at varying SNRs. Rayleigh has the highest classification accuracy followed by AWGN and Rician.

Fig. 4 Performance of CNN in AWGN channel

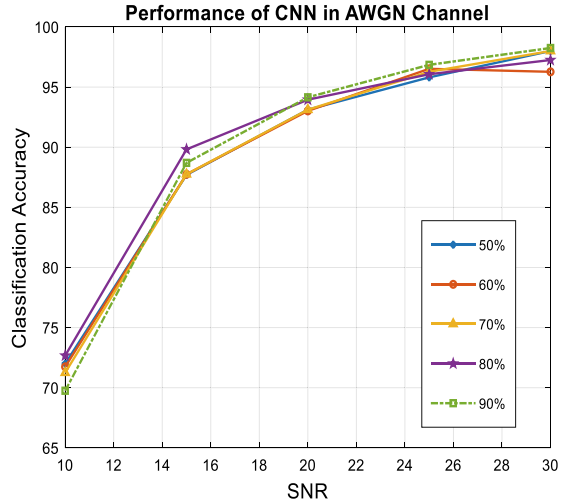
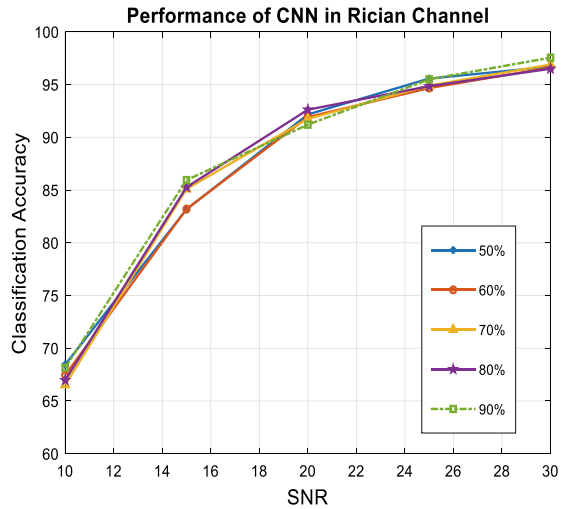


Fig. 5 Performance of CNN in Rician channel



4.2 Classification of Digital Modulation Signals

The classifier is trained to classify 7 digital modulation signals. The CNN is trained from 50 to 90% training rate. At each training rate, the performance is analyzed at different SNR values. The same procedure is carried out for three channels. Figures 8, 9, and 10 represent the performance of CNN for digital classes under AWGN, Rician, and Rayleigh channels at different training rates. Maximum classification accuracy is 97.46% in the AWGN channel, 94.91% in the Rician channel, and 94.14% in Rayleigh.

Fig. 6 Performance of CNN in Rayleigh channel

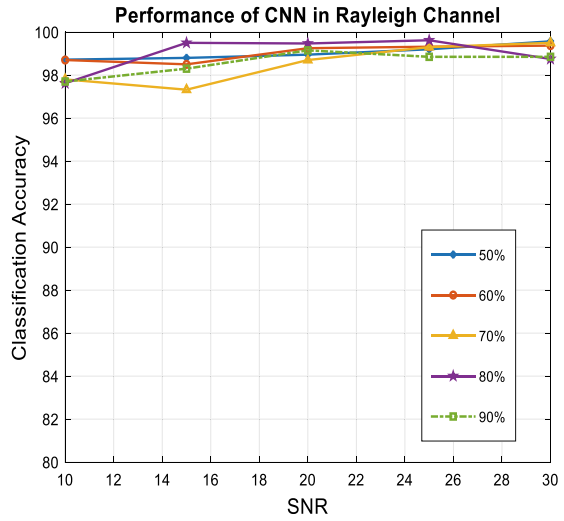


Fig. 7 Performance of CNN under 3 channels

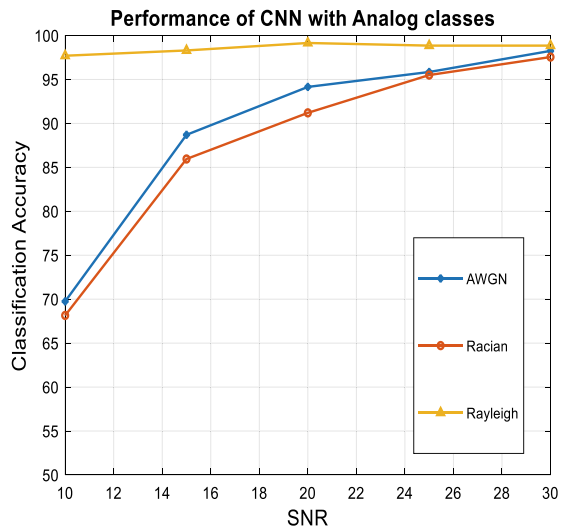


Figure 11 shows the behavior of CNN for digital classes under AWGN and fading channels. The plot is between varying SNRs and classification accuracy at 90% training rate. AWGN has the highest accuracy, followed by Rician and Rayleigh.

Fig. 8 Performance of CNN in AWGN

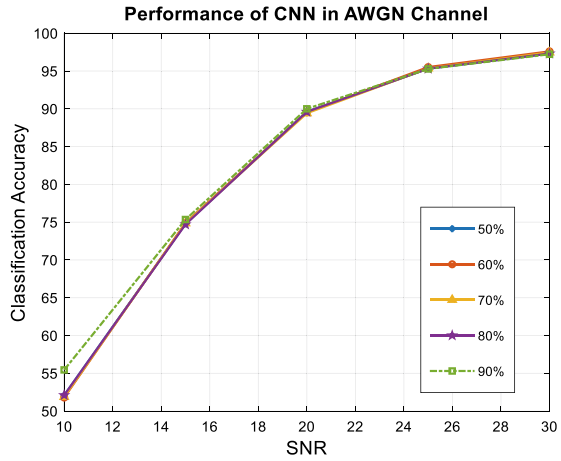
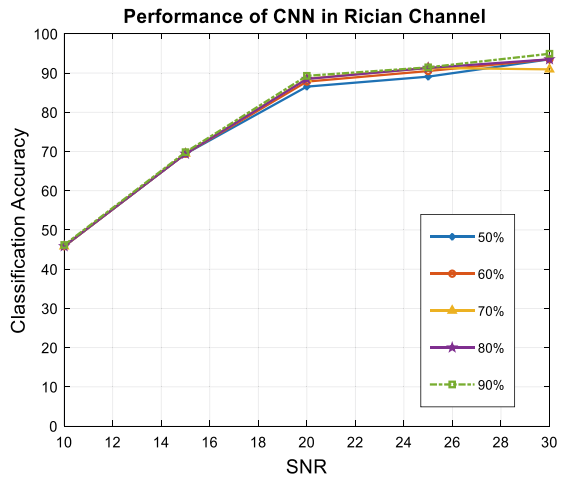


Fig. 9 Performance of CNN in Rician



4.3 Classification of Analog and Digital Modulation Signals

Finally, the classifier is trained to categorize 11 modulation classes, which cover both analog and digital modulation classes. Figures 12, 13, and 14 represent the classification accuracy at different SNRs under AWGN and fading channels. Three paths with varying path gains were taken to create fading channels. Maximum accuracy of 99.72% is attained in the AWGN channel, 92.75% in the Rician channel, and 95.32% in the Rayleigh channel.

Even with 11 classes, the classifier had 99% of accuracy in the AWGN channel. Figure 15 demonstrates the performance of the model with 11 classes under varied SNR values. The plot illustrates that the CNN has good performance in AWGN channel followed by Rayleigh and Rician.

Fig. 10 Performance of CNN in Rayleigh

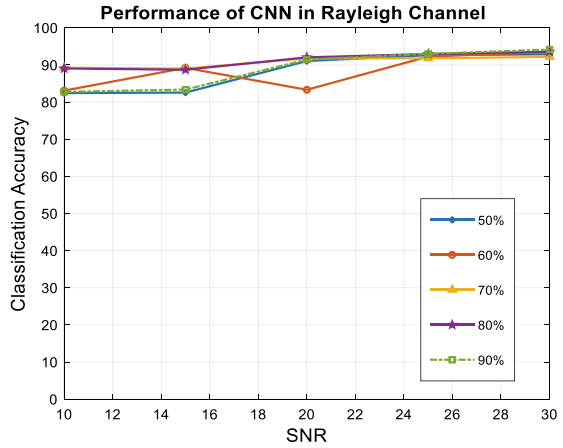


Fig. 11 Performance of CNN in 3 channels

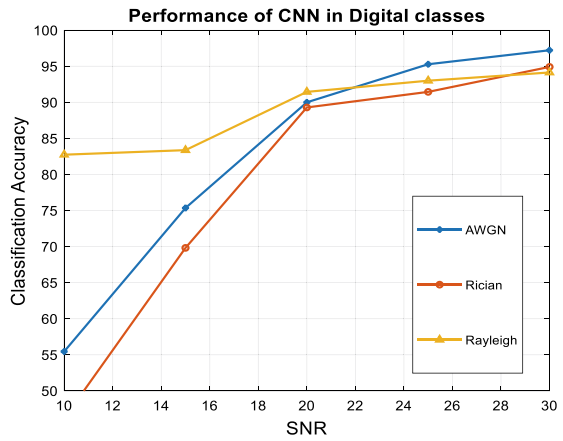


Fig. 12 Performance in AWGN for 11 classes

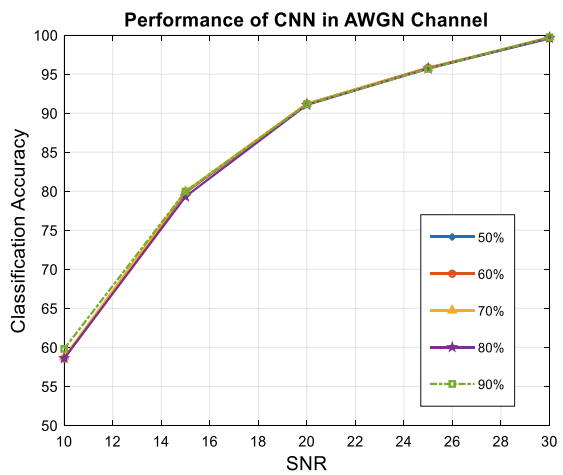


Fig. 13 Performance of in Rician for 11 classes

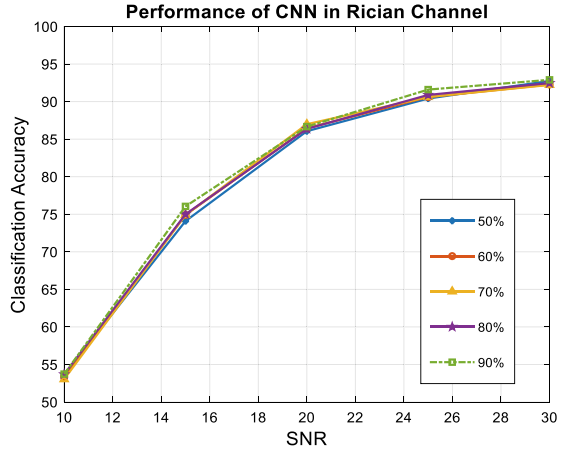


Fig. 14 Performance in Rayleigh for 11 classes

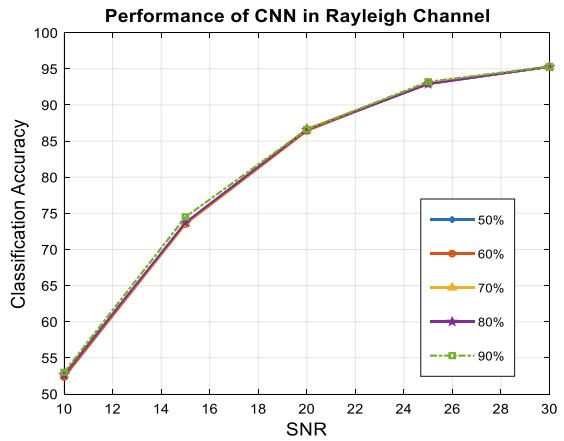
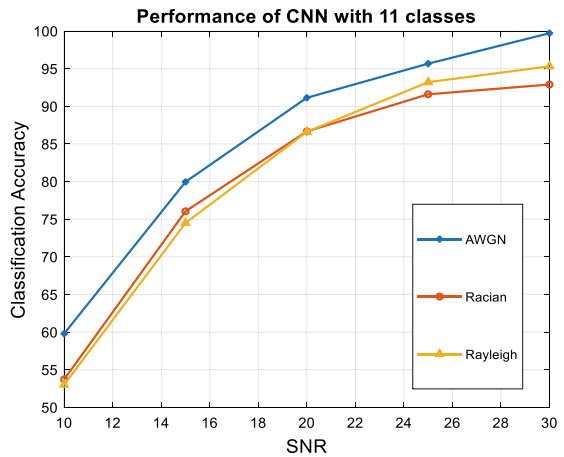


Fig. 15 Overall performance of CNN



5 Conclusion

In this paper, an AMC classifier using a CNN is developed. It can classify both analog and digital modulation classes. The classifier is trained to work in both AWGN and fading channels. The training rates were varied from 50 to 90% with a 10% step, and the SNR was varied from 10 to 30 dB with a 5 dB step. Initially, the classifier is trained with 4 analog modulated classes and the performance of the network is evaluated under AWGN, Rician, and Rayleigh channels. Later performance is evaluated with 7 digitally modulated classes, and finally, classification of 11 classes of both analog and digital modulation classes is examined. The model can classify both analog and digital modulation classes with great efficiency. The results reveal that the accuracy of the model is superior than existing strategies, which consist of limited modulation classes. As the feature extraction component and classification part were conducted in a single phase, the computational complexity is reduced.

References

1. Subbarao MV, Samundiswary P (2019) K-nearest neighbors based automatic modulation classifier for next generation adaptive radio systems. *Int J Secur Appl* 13:41–50
2. Subbarao MV, Samundiswary P (2021) Automatic modulation classification using cumulants and ensemble classifiers. In: Kalya S, Kulkarni M, Shivaprakasha KS (eds) *Advances in VLSI, signal processing, power electronics, IoT, communication and embedded systems. Lecture notes in electrical engineering*, vol 752. Springer, Singapore
3. Venkata Subbarao M, Samundiswary P (2020) Performance analysis of modulation recognition in multipath fading channels using pattern recognition classifiers. *Wirel Pers Commun* 115:129–151. <https://doi.org/10.1007/s11277-020-07564-z>
4. Nandi K, Azzouz EE (1998) Algorithms for automatic modulation recognition of communication signals. *IEEE Trans Commun* 46(4):431–436
5. Mühlhaus MS, Oner M, Dobre OA, Jondral FK (2013) A low complexity modulation classification algorithm for MIMO systems. *IEEE Commun Lett* 17(10):1881–1884. <https://doi.org/10.1109/LCOMM.2013.091113.130975>
6. Juan-ping W, Ying-zheng H, Jin-mei Z, Hua-kui W (2010) Automatic modulation recognition of digital communication signals. In: *First international conference on pervasive computing, signal processing and applications*
7. Hassan K, Dayoub I, Hamouda W, Berbineau M (2009) Automatic modulation recognition using wavelet transform and neural network. In: *2009 9th International conference on intelligent transport systems telecommunications (ITST)*, pp 234–238
8. Shen W, Gao Q (2014) Automatic digital modulation recognition based on locality preserved projection. In: *2014 International conference on wireless communication and sensor network, Wuhan, China*, pp 348–352. <https://doi.org/10.1109/WCSN.2014.78>
9. Park S, Choi J-H, Nah S-P, Jang W, Kim DY (2008) Automatic modulation recognition of digital signals using wavelet features and SVM. In: *2008 10th International conference on advanced communication technology*, pp 387–390
10. Zhang R, Yin Z, Wu Z, Zhou S (2021) A novel automatic modulation classification method using attention mechanism and hybrid parallel neural network. *Appl Sci* 11:1327

Fusion of Panchromatic and Low-Resolution Multispectral Images Using Combination of PCA and Morphological Hat Transformation



Rupa Banerjee, Vikrant Bhateja, Reetika Mishra, Jerry Chun-Wei Lin, and Carlos M. Travieso-Gonzalez

Abstract In past few years, satellite imagery has made significant advances in the sectors of climate, farming, and defense. As seen in the Low-Resolution Multi-Spectral (LRMS) and Panchromatic (PAN) images, satellites acquire data in two domains: spectral resolution and spatial resolution. The goal of remote sensing image fusion is to merge the complimentary information from the PAN and LRMS images. The same has been accomplished in this study using Morphological Hat Transformation and Discrete Wavelet Transformation. Decomposition of LRMS and Pan images into approximation (cA) and detail (cD) coefficients is used in the proposed fusion method. To extract the first principal component, PCA is applied to the cA coefficients of both images, and MHT is applied to both cD coefficients at the same time. The end outcome is a fusion of the results after applying Inverse Wavelet Transformation (IDWT). The performance evaluation of the final-fused image is presented in the later sections.

Keywords DWT · Image fusion · Morphological hat transform · MS image · PAN image

R. Banerjee · V. Bhateja (✉) · R. Mishra

Department of Electronics and Communication Engineering, Shri Ramswaroop Memorial College of Engineering and Management (SRMCEM), Faizabad Road, Lucknow, UP 226028, India
e-mail: bhateja.vikrant@gmail.com

Department of Electronics Engineering, Faculty of Engineering and Technology, Veer Bahadur Singh Purvanchal University, Shahganj Road, Jaunpur, Uttar Pradesh, India

J. C.-W. Lin

Department of Computer Science, Electrical Engineering and Mathematical Sciences, Western Norway University of Applied Sciences, Bergen, Norway
e-mail: jerrylin@ieee.org

C. M. Travieso-Gonzalez

Institute for Technological Development and Innovation in Communication (IDeTIC), University of Las Palmas de Gran Canaria (ULPGC), Campus Universitario de Tafira, sn. Telecomunicacion. Pabellon B. D-111, 35017 Las Palmas de Gran Canaria, Spain
e-mail: carlos.travieso@ulgpc.es

1 Introduction

Modern imaging sensors have emerged as a result of the rapid growth in technology in recent years [1]. For a precise distinction of land covers, a high spectral resolution is required. The multi-spectral and panchromatic images are the two types of images that satellites may collect. High spectral resolution but low spatial resolution can be found in multi-spectral pictures. The panchromatic images, on the other hand, have excellent spatial resolution but poor spectral resolution [2].

Image fusion is a method for fusing data from numerous sources with additional images to enhance human vision [3]. C. A. Rishikeshan proposed an automated mathematical morphology-driven system for water body extraction from remotely sensed photos. The suggested technique dynamically selected the filter size (referred to as the structural element (SE) in MM), allowing one to preserve the most accurate results from different sets of input images with variable spatial resolutions. In order to retain edges and shapes, the paper used mathematical morphological (MM) techniques for feature extraction in image processing. In the approach, there was only one band used for the computation. The amount of processing needed for the suggested MM technique depended on the size of the structuring element used and the total number of pixels in the input images. This ensured that the MM approach would work as intended within a few minutes or less [4]. Sui and Kim [5] proposed a study into the use of wavelet transform-based multimedia image processing technologies. The paper's research item was an image stored using multimedia information storage technology. The wavelet transform was used to divide the image into the benefits of low-frequency and high-frequency characteristics, and a model for multimedia processing technology based on the wavelet transform was developed. The cubic b-spline wavelet was used as the wavelet basis function across four layers of wavelet decomposition, and the accuracy rate was 89.08%. An image fusion algorithm based on Principal Component Analysis (PCA) and Curvelet Transformation was introduced by Wu et al. [6]. The work proposed pre-processing of the LRMS and PAN image using PCA and Morphological Hat Transformation, respectively. It employs PCNN algorithm for segmentation purposes. The experimental results of their study illustrated better fusion results than those of traditional fusion algorithms. Hyperspectral image visualization with edge-preserving filtering and Principal Component Analysis proposed by Kang et al. [7] was an edge—preserving, filtering approach, and averaging-based image fusion. Compared to other popular hyperspectral image visualization techniques, experiments show that the suggested method gives sharper details and makes the land coverings in the landscape easier to differentiate. PCA, DWT, and the Morphological Hat Transformation are used in this paper. The source PAN and LRMS images are decomposed with Discrete Wavelet Transformation. The low-frequency coefficients are pre-processed by PCA and Morphological operations are applied on the high-frequency coefficients to increase the contrast of the image. The following parts contain the remainder of this paper: The proposed approach is described in Sect. 2, the experimental findings are presented in Sect. 3, and the analysis is concluded in Sect. 4.

2 Proposed Fusion Approach of Low-Resolution Multispectral and Panchromatic Images

Panchromatic (PAN) and Low-Resolution Multispectral (LRMS) images are fused in this approach of multimodal image fusion. The proposed methodology does the same by using Discrete Wavelet Transformation (DWT) to divide both images into approximation coefficients (cA) and detail coefficients (cD). The approximation coefficients are also pre-processed using Principal Component Analysis (PCA), and the detail coefficients of the images are transformed using Morphological Hat Transformation (MHT). The MHT outcomes are fused together using the maximum fusion rule, whereas the PCA values are fused together using the Average Fusion Rule. As seen in Fig. 1, the final-fused image is created by combining both components using the Inverse Wavelet Transformation (IDWT). The techniques adopted are detailed in subsections below.

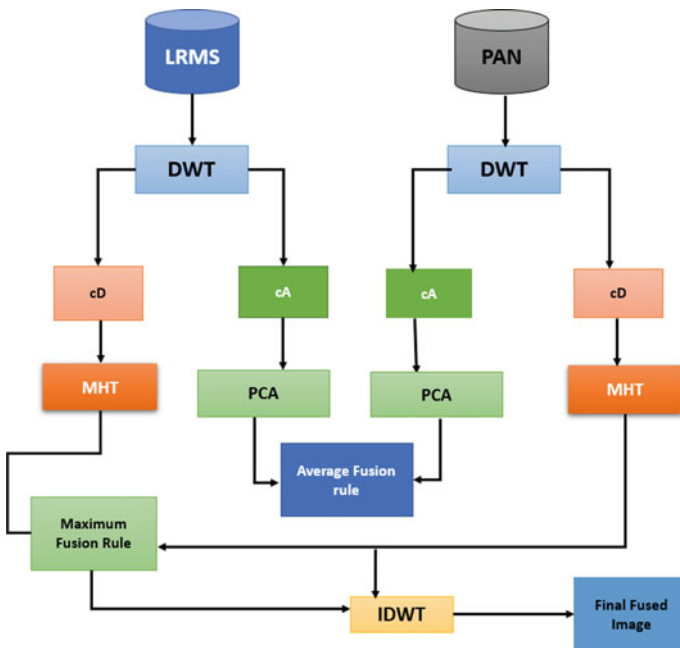


Fig. 1 Block diagram of the proposed fusion approach of low-resolution multispectral and panchromatic images

2.1 Wavelets

Wavelet theory is a generalization of Fourier theory that was presented as a replacement for the short-time Fourier transform (STFT) in image processing and has since been widely used to give a multi-resolution decomposition of an image resulting in non-redundant picture representation. Wavelets, which are defined as functions formed by the translation and dilation of the mother wavelet, are the foundations of the Wavelet theory [8].

2.1.1 Discrete Wavelet Transformation

Discrete Wavelet Transformation (DWT) is an efficient multi-scale transformation to decompose the source images into its approximation and detail coefficients [9]. For a single dimension, the DWT can be defined as in the following equations [10]

$$W_{\Psi}(j, k) = 1/\sqrt{M} \sum_x f(x) \Phi_{j,k}(x), \quad (1)$$

$$W_{\Psi}(j, k) = 1/\sqrt{M} \sum_x f(x) \Psi_{j,k}(x), \quad (2)$$

where $\Psi_{j,k}$ is the function of discrete variable and $\Phi_{j,k}$ is a part of the group of expansion functions that the scaling function $\Phi(x)$ yields.

2.1.2 Inverse Discrete Wavelet Transformation

The Inverse Discrete Wavelet Transformation (IDWT) is employed in reconstructing the image from its decomposed coefficients and is defined as in Eq. (3) [10].

$$f(x) = 1/\sqrt{M} \sum W_{\Phi}(j_0, k) \Phi_{j_0,k}(x) + 1/\sqrt{M_x} \sum_{j=j_0}^{\infty} \sum_k W_{\Psi}(j, k) \Phi_{j,k}(x), \quad (3)$$

where $f(x)$, $\Phi_{j_0,k}(x)$ are functions of discrete variable and $j \geq j_0$

2.2 Principal Component Analysis (PCA)

The Principal Component Analysis is used to reduce the dimensions of an image by applying some transformations. It converts the linearly correlated variables of an image into a set of uncorrelated variables. It is a set of procedures that includes computing the source image's covariance matrix, then computing the eigen vectors

of that covariance matrix to find the Principal Components (PCs) as stated in Eq. (4) [5, 11]. It is noted that the first component with the highest eigen value is the principal component containing the maximum information of the image.

$$\lambda_i v_i = \text{cov}(x)v_i, \quad (4)$$

where $i = 1, 2, \dots, d$, and v_i is an eigenvector of $\text{cov}(x)$.

$$P_1 = \frac{v(1)}{\Sigma v} \quad (5)$$

where P_1 is the first principal component.

The outputs of both the images after the application of PCA are fused together using Average Fusion Rule which can be given as in Eq. (5).

2.3 Morphological Hat Transformation (MHT)

Morphological Hat transform is a shape-based contrast enhancement technique that processes pictures. It takes the source image and a structuring element as inputs and processes the source image into a specific shape and size, producing an output image that is the same size as the source image. Dilation and erosion are the two major morphological activities. In an image, erosion eliminates the pixels of an object, whereas dilation adds the pixels on the object boundaries. Dilation and erosion provide the opening and closing morphological rules that make up the Top-hat and Bottom-Hat morphological operations [12]. Dilation followed by erosion is the definition of opening, whereas erosion followed by dilation is the definition of closing [12]. Bottom-hat transform has been used on the cD coefficients.

Bottom-Hat operation

Bottom-hat operation is defined as subtracting the original image from the closing image [13], as shown in Eq. (6) [12].

$$\text{BH}(x, y) = f_{\text{CL}}(x, y) - f(x, y), \quad (6)$$

where $\text{BH}(x, y)$ is the Bottom-hat transform, $f_{\text{CL}}(x, y)$ is the closing of $f(x, y)$ by a structuring element s as given in Eq. (7) [12]. In the proposed approach, the shape of the structuring element is set to disk and $f(x, y)$ is the original image

$$f_{\text{CL}}(x, y) = f.s. \quad (7)$$

2.4 Fusion Rules

The approximation coefficients after the application of PCA are fused by the Average Fusion rule as defined in Eq. (8) [14], and the detail coefficients after performing MHT are fused together by Maximum Fusion Rule as given in Eq. (9) [13].

$$f(i, j) = (L(i, j) + P(i, j))/2, \quad (8)$$

$$f(i, j) = \max|L(i, j) + P(i, j)|, \quad (9)$$

where $f(i, j)$ is the pixel intensity of fused image and $L(i, j)$ and $P(i, j)$ are the pixel intensities of the images of approximation coefficients of LRMS and Pan image, respectively.

Algorithm #1: Procedural Steps for Image Fusion

Begin

Step 1: Read the source LRMS (I_1) and PAN (I_2) images.

Step 2: Perform DWT on (I_1) and (I_2) to decompose each into approximation and detail coefficients as in Eq. (1).

Step 3: Compute eigen vector of the $\text{cov}(I_1)$ using Eq. (4).

Step 4: $I_3 \leftarrow$ Extract the Principal Component using Eq. (5).

Step 5: Set structuring element as disk shape for Morphological Hat Transform on I_2 .

Step 6: $I_4 \leftarrow$ Compute Bottom-Hat transform on (I_2) using formula in Eq. (6).

Step 7: (I_5) \leftarrow Fuse the first Principal Components by Average Fusion Rule as in Eq. (8).

Step 8: $I_6 \leftarrow$ Fuse the results of bottom-hat MHT by maximum fusion rule as in Eq. (9).

Step 9: $I_7 \leftarrow$ Perform IDWT on the fused coefficients to obtain the fused image using formula in Eq. (3).

End

3 Experimental Results

3.1 Dataset

The two datasets used in the experiment were World-View II and GF-2. LRMS images have spatial resolutions of 1.8 m and 3.2 m, respectively, whereas panchromatic images have spatial resolutions of 0.5 m and 0.8 m [15]. Two images were chosen

at random for testing and simulation purposes. The two test images, as well as the resulting fused image, are shown in the next section. To characterize the fusion performance, entropy and standard deviation are measured to analyze the generated image quantitatively. The standard deviation evaluates the contrast of the image, while the entropy evaluates the quantity of information in the image [16]. A high entropy as well as a high standard deviation indicates greater fusion.

3.2 Simulation Results

The algorithm of the proposed methodology was tested on two random images of the dataset. The software used for simulation was MATLAB r2022a on a i5 processor with 8 GB of memory and 2 GB of GPU system. The fused images were compared with the source images on fusion parameters to prove its effectiveness. The results are illustrated in the cases below.

The source LRMS picture (a), the source PAN image (b), and the fused image are shown in Fig. 2c. The fused image has taken up the color of the source LRMS image and the clarity of the PAN image. Tables 1 and 2 shows the quantitative experimental data of both images.

The experimental results clearly show the fused image's intensity as well as better contrast. While attempting to acquire the saturation of the LRMS image, the fused image focuses on improvised spatial information by enhancing the entropy and standard deviation values from each of its source images.

4 Conclusion

This paper presents MHT and PCA in the Wavelet domain to provide a method for remote sensing image fusion, DWT allows for simultaneous time and frequency domain analysis, resulting in a multi-resolution representation of the image signal. The PCA approach increases the spatial content of the cA coefficients of the image and bottom-hat morphological transform improves the image's edge details. Entropy is a measure of information content, and the fused image has a higher value than the PAN and LRMS images, while a higher Standard Deviation suggests that the fused image has superior contrast. Overall, the observed experimental results justify the fusion procedure used and proves that the fusion is efficient.

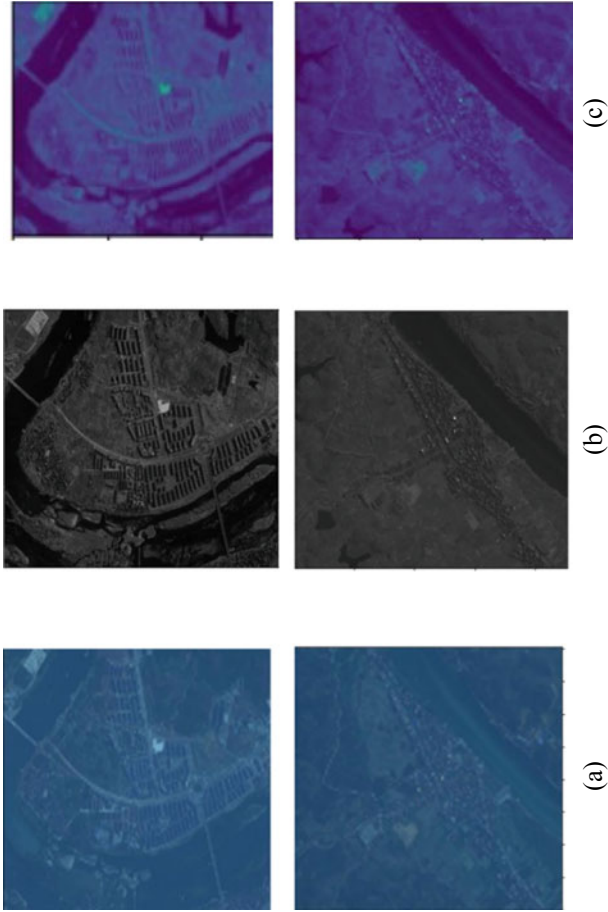


Fig. 2 Experimental results of images: **a** source LRMS image, **b** source PAN image, **c** fused image

Table 1 Qualitative assessment of the fused image I

Metric	LRMS image	PAN image	Fused image
Entropy	7.3412	6.2958	10.141
Standard deviation	21.6020	15.0088	25.63

Table 2 Qualitative assessment of the fused image II

Metric	LRMS image	PAN image	Fused image
Entropy	6.758	5.112	11.110
Standard deviation	27.6476	16.192	28.754

References

1. Moin A, Bhateja V, Srivastava A (2016) Multispectral medical image fusion using PCA in wavelet domain. In: ICTCS' 16, vol 87, pp 1–4
2. Ghassemian H (2016) A review of remote sensing image fusion methods. *Inf Fusion* 32:75–89
3. Singhal A, Bhateja V, Singh A, Satapathy SC (2018) Visible-infrared image fusion using anisotropic diffusion. In: Intelligent computing and information and communication. Advances in intelligent system and computing, vol 6723, pp 525–532
4. Rishikeshan CA, Ramesh H (2018) An automated mathematical morphology driven algorithm for water body extraction from remote sensed images. *ISPRS J Photogramm Remote Sens* 146:11–21
5. Sui K, Kim HG (2019) Research on application of multimedia image processing technology based on wavelet transform. *EURASIP J Image Video Process* 24
6. Wu Z, Huang Y, Zhang K (2018) Remote sensing image fusion method based on PCA and Curvelet transform. *J Indian Soc Remote Sens* 46:687–695
7. Kang X, Duan P, Li S (2020) Hyperspectral image visualization with edge-preserving filtering and principal component analysis. *Inf Fusion* 57:130–143
8. Naidu VPS, Raul JR (2008) Pixel-level image fusion using wavelets and principal component analysis. *Defence Sci J* 58:3
9. Dikshit A, Bhateja V, Rai A (2022) Discrete wavelet transform based fusion of mammogram images for contrast improvement. In: Advances in micro-electronics, embedded and IOT. Lecture notes in electrical engineering, vol. 38, pp 203–210
10. Ravichandran D, Nimmatoori R, Ahaamad MG (2016) Mathematical representation of 1D, 2D, 3D wavelet transform for image coding. *Int J Comput Theory Eng* 5:2319–2526
11. Kurita T (2020) Principal component analysis (PCA). In: Ikeuchi K (eds) Computer vision. https://doi.org/10.1007/978-3-030-03243-2_649-1
12. Bhateja V, Nigam M, Bhadauria AS, Arya A, Zhang ED (2019) Human visual system based optimized mathematical morphology approach for enhancement of brain MR images. *J Ambient Intell Human Comput*
13. Kumar V, Choudhary T (2019) Real-time recognition of malignant skin lesions using ensemble modeling. *J Sci J Res* 78:148–154
14. Prakash O, Park MC, Khare A, Jeong M, Gwak J (2019) Multiscale fusion of multimodal medical images using lifting scheme based biorthogonal wavelet transform. *Optik* 182:995–1014
15. Ma J, Yu W, Chen C, Liang P, Guo X, Jiang J (2020) Pan-GAN: an unsupervised pan-sharpening method for remote sensing image fusion. *Inf Fusion* 62:110–120
16. Bhateja V, Misra M, Urooj S (2019) Quantitative metrics for mammographic image quality assessment. In: Non-linear filters for mammogram enhancement. Studies in computational intelligence, vol 861, pp 87–93

A Compact Circularly Polarized MIMO Diversified Antenna for 5G Mobile Applications



M. Venkateswararao, M. Vamsikrishna, P. Manjusha, B. Suryaprakashrao, G. Sruthi, and B. Raghavkrishna

Abstract In this article, a circular ring-structured MIMO antenna is designed and analysed for 5G applications. The proposed antenna design using ANSYS electronic desktop by using FR-4 material as substrate. The antenna has two complementary split-ring resonators which are etched from the patch. Using the principle of monopole structures and strong isolation characteristics the proposed model has been developed. Proposed model works up to 40 GHz at different resonant bands and covers modern commercial applications WIFI (2.4 GHz), WLAN (5.2 GHz), WiMAX (3.5 GHz), X-band applications (above 8 GHz), 5G and millimetre wave applications. The results show proposed antenna a good candidate for future ready commercial and 5G millimetre wave applications.

Keywords MIMO · CSRR · Isolation reduction

1 Introduction

In the present scenario, with the rapid evolution of wireless communication, the use of antenna services and the usage of wireless systems have been increased briskly. Due to this, there is an immense demand for wireless communication systems and there are several challenges that will affect the antenna design. Abundant antenna solutions which can be used in these systems were developed and progress has been made accordingly with the development. The wideband applications have been increased. But by using these applications, the data rate has been increased. The best solution one can furnish is the usage of a MIMO antenna. The MIMO uses multiple transmitters to send the signal on the same frequency to multiple receivers.

M. Venkateswararao (✉)

Department of ECE, PACE Institute of Engineering and Sciences(Autonomous), Ongole, AP, India

e-mail: drmvrdiet@gmail.com

M. Vamsikrishna · P. Manjusha · B. Suryaprakashrao · G. Sruthi · B. Raghavkrishna

Department of ECE, Dhanekula Institute of Engineering & Technology, Vijayawada, AP, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

237

V. V. S. S. Chakravarthy et al. (eds.), *Advances in Signal Processing, Embedded*

Systems and IoT, Lecture Notes in Electrical Engineering 992,

https://doi.org/10.1007/978-981-19-8865-3_22

MIMO system can also be used under high scattering conditions even when the signal bounces around the environment. When we use multiple antennas, systems are placed in near proximity which can lead to mutual coupling, the mutual coupling is a major factor due to which system capacity and performance can dwindle, we can overcome this by using a MIMO antenna along with this it also provides high isolation. The MIMO uses different types of elements in its structure like metamaterials, defected ground structures and different parasitic elements.

Lately, it has been announced that wireless communication can make use of 28 and 38 GHz frequency bands, as the 5G application requires wide bandwidth to provide a high data rate which is provided by millimetre-wave applications. By using the 5G antenna, we observe shorter wavelengths and high path loss in space. So, antennas of high gain should overcome the space losses and different forms of fading, we can overcome this by using a MIMO antenna as discussed earlier the MIMO antenna can be used to decrease different losses and improve isolation. So, by the fusion of high gain and MIMO configuration, we get a unique antenna for the mm-wave applications by this method we can miniaturize the size of the antenna and decrease the cost of the antenna. Yang et al. [1] designed a MIMO antenna that is suitable for wideband applications by including the use of vias due to which the bandwidth is improved and the usage of different slits in the design provided high isolation. With the use of asymmetric antennas along with different ground planes, we can reduce electromagnetic coupling [2] and we can achieve good polarization diversity and isolation by placing the antenna properly. In [2], stubs are used due to which the frequency band is enhanced.

A Y-shaped branch structure has been introduced between the radiating element in order to get high isolation of about 21 dB and in the design circular slot and rectangular slots on the patch are used which produces two different notch bands [3]. In [4], CPW feeding is used through which the bandwidth has been increased from 3 to 16 GHz and by cup-shaped branches are added high up the ground plate the band-notch characteristics have been increased and good isolation has been achieved. The UWB-MIMO antenna is designed by using a shared structure [5], by incorporating a metal layer to work as a radiator along with it each radiator in the ground plane, and in order to improve the bandwidth, a dual L-shaped structure is used for feeding. In [5], they also made use of fork-shaped slots by which the mutual coupling is decreased along with the increase in the isolation. In [6], the unavoidable mutual coupling has been reduced by using different techniques like Modified Serpentine Structures which act like decoupling units and by using the electronic bandgap is further used to improve the impedance bandwidth.

Quasi Self Complementary method is used on the antenna [7] in order to improve impedance bandwidth. This design contains a semi-elliptical patch whose feeding is done by a tapered microstrip line which can be used for isolation enhancement and different diversity techniques are utilized for decreasing the mutual coupling. In [8] author proposed an antenna that can be used in mobile handset, this design consists of two tri-branch planar inverted-F antennae along with a T-stub used as a ground using this design the isolation has been increased. An all elliptical antenna is proposed in [9] in order to provide adequate impedance matching and different decoupling structures

have been used to provide an efficiency of about 94%. To provide compactness (IB-CEBG cells are used in [10] in order to obtain notches at interfering frequencies and to provide compactness), EBG and spiral-shaped defects are introduced these structures along with individual monopoles are used to reduce mutual coupling and improve the path length and have perfect impedance matching.

2 MIMO Antenna Design and Its Analysis

2.1 Dual-Port Monopole Antenna Analysis

In order to provide high frequencies and omni-directional patterns, a microstrip fed PMA is designed in this work. The proposed MIMO antenna is fabricated on FR-4-Substrate with the dimensions of $a \times b \times 1.6$ having a relative permittivity of 4.4. The substrate is ascribed by 2 radiating circular patches which are incorporated with splitting resonators that provide a fair amount of capacitance. The stepped impedance microstrip lines are used for feeding the circular patches. The proposed work has been analysed iteration-wise (Prior to 5 iterations). In iteration I, a single circular patch antenna is attached with a stepped impedance feedline which results in working at 26 GHz frequency in addition to that, in iteration 2, the antenna is changed to MIMO structure by attaching another port to the similar antenna with a distance of 20.7 mm. In iteration 2, the MIMO antennas work at working a 23 GHz frequency or band and ground is designed on the bottom side with a U-shaped slot which is used to reduce mutual coupling and in order to provide impedance matching. F-Shaped stubs are used in iteration III which works at a frequency 24 GHz and circular rings are embedded along with F-Shaped stubs in iteration IV along with it to improve the efficiency and consistently maintain a less Specific absorption rate (SAR) four electronic bandgap cells with different dimensions are used in iteration IV which results in working at 23 GHz. Cylindrical-shaped stubs are used in iteration V which are placed on the EBG structures and provide a good electrical path and SSR is incorporated on the circular patch that provides an exceptional magnetic susceptibility and capacitance.

The proposed work is also suitable to furnish an additional decoupling path between adjacent elements by diverting the surface current with the help of F-shaped stubs that are embedded with circular rings in the ground plane of the Monopole antenna. The proposed antenna resonant frequency can be calculated by using the below formula (Figs. 1 and 2).

$$f_{\text{res}} = \frac{c}{(2 \times R \times \sqrt{\epsilon_r})}$$

The final dimensions of the proposed antenna have been represented in the diagram and stated in Table 1 (Fig. 3).

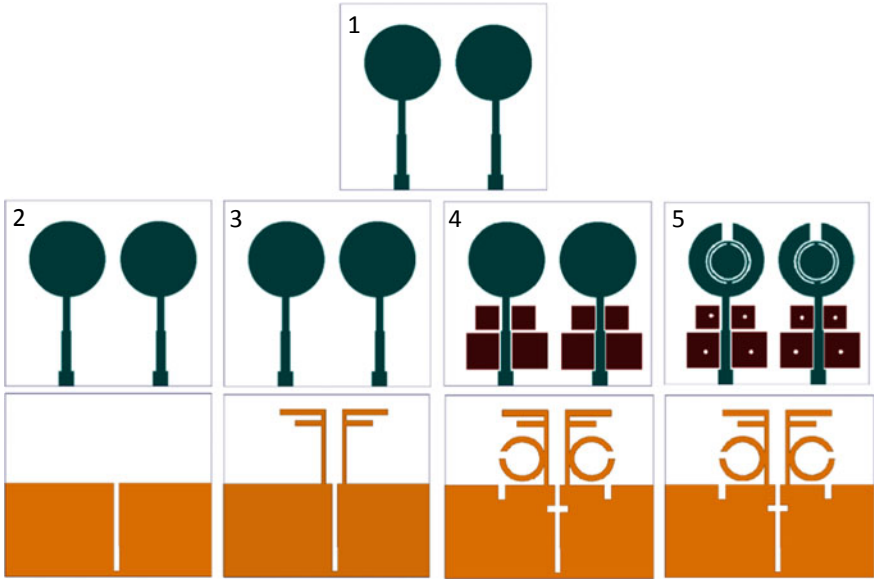
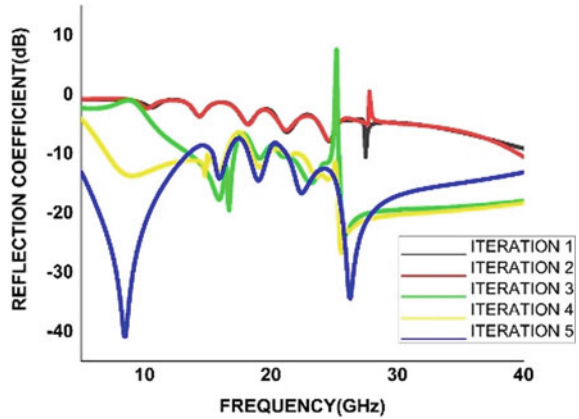


Fig. 1 Dual-port Monopole antenna iterations

Fig. 2 Reflection coefficients at different iterations



2.2 Axial Ratio Generation

By combining various slotted structures and choosing the elements of the antenna deliberately in the perfect geometry, we can attain perfect circular polarization. As the circular polarization mainly depends on the elements which are embedded on the antenna we should very carefully select them. In the proposed model, different SSR structures along which EBG and F-shaped stubs are examined concerning the configuration of circular polarization. Due to the presence of U-shaped slot in the

Table 1 Dimensions of proposed antenna

Parameters	L_s	W_s	L_1	L_2	L_3	L_4
Units (in mm)	49	54.4	9.5	10.5	10.6	9
Parameters	L_5	W_1	W_2	W_3	W_4	W_5
Units (in mm)	6.1	8.5	6.1	3.9	3.9	1.8
Parameters	W_6	W_7	R_1	R_2	R_3	R_4
Units (in mm)	20.7	14.2	1.5	4	3.5	8
Parameters	G	D_9	Lg_1	Lg_2	Lg_3	Lg_4
Units (in mm)	4.6	8.5	24	24.5	1.5	1.5
Parameters	LG_5	LG_6	LG_7	LG_8	LG_9	LG_{10}
Units (in mm)	16.9	5.2	1.9	4	20.5	1
Parameters	Lg_{11}	Wg_1	Wg_2	Wg_3	Wg_4	Wg_5
Units (in mm)	22.5	25.3	8.8	12.3	1.9	2
Parameters	Wg_6	Wg_7	D_1	D_2		
Units (in mm)	11	7	4	6		

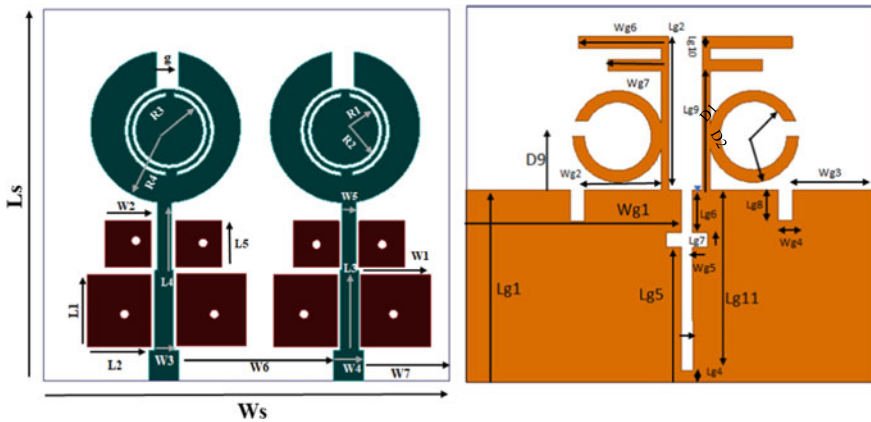


Fig. 3 Dimensions of proposed antenna

antenna geometry in the iteration 2 axial ratio band is present in the iteration 3 ascribed to slot F-shaped stubs are embedded which lead to another axial ratio band. The F-shaped stubs are further modified by encapsulating SSR to the F-shaped stubs leading to greater improvement in the current distributions leading to change in the value of CP.

3 Results and Discussions

3.1 Envelope Correlation Coefficient and Diversity Gain

The ECC immensely required in order to measure the amount of correlation is between the antenna which is excited to the antennas which are not excited they can be measured by using the mathematical equation

$$\rho = \frac{|s_{11} * s_{12} + s_{21} * s_{22}|^2}{(1 - (|s_{11}|^2 + |s_{21}|^2))(1 - (|s_{22}|^2 + |s_{12}|^2))}$$

The above equation can be used in order to measure the value of ECC with the aid of S-parameters but the above equation can only be used if the current distributions are equal in the antennas and they should be lossless. Typically, the value of ECC is zero for antennas that are uncorrelated but practically the value of ECC is less than 5 dB. This work targets achieving ECC < 3 dB. The diversity gain of an antenna is one of the important parameters due to the different diversity schemes the signal to noise ratio is increased which is called as diversity gain.

In Fig. 5b, the mutual coupling has been analysed at different frequencies the mutual coupling is used to measure the electromagnetic interaction between different antennas. If one port of an antenna provides LHCP and the other antenna provides RHCP then the antenna are said to be independent to each other practically the mutual coupling is below - 20 dB. The proposed antenna has a mutual coupling of less than - 15 dB which makes it a good candidate to work at the wide band of frequencies. Figure 4 depicts the axial ratio plot working at 9.9 GHz it shows a omnidirectional patterns at both LHCP and RHCP Planes.

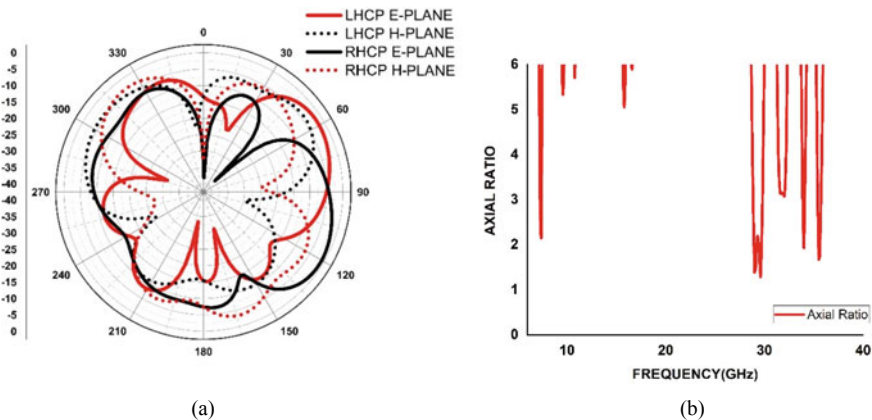


Fig. 4 **a** Axial ratio plot at a frequency of 9.5 (GHz). **b** Axial ratio at different frequencies

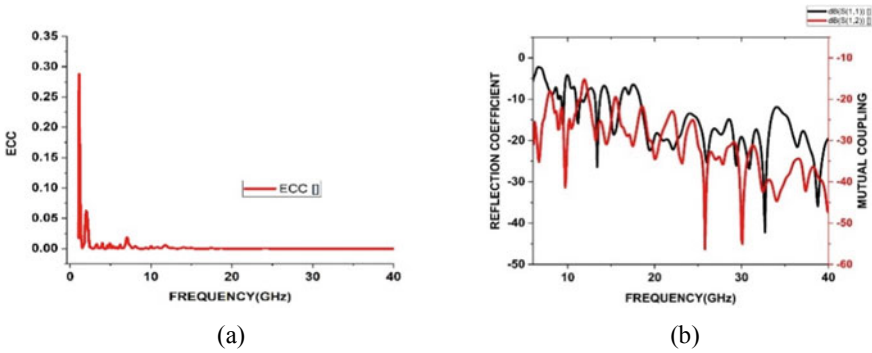


Fig. 5 Simulated results. **a** Frequency vs ECC **b** frequency versus reflection coefficient and mutual coupling

3.2 E-Field Distributions

The electric field distributions at different ports and at different frequencies are observed in the Fig. 6. At 9.5 GHz frequency, when antenna at port 1 is radiated the antenna at port 2 is not affected which can be observed in Fig. 6a. At 9.5 GHz frequency, when antenna at port 2 is radiated the antenna at port 1 is not affected which can be observed in Fig. 6b this shows that the antenna obeys the diversity principle. The same principle is observed at higher frequency in Fig. 6c we can observe how the antenna behaves at 32.5 GHz frequency, when antenna at port 1 is radiated the antenna at port 2 is not affected in the same way when antenna at port 2 is radiated the antenna at port 1 is not affected which can be observed in the Fig. 6d. From the Fig. 6e, we can observe how the antenna behaves when two ports of the antenna are radiating at 9.5 GHz and the Fig. 6f we can observe how the antenna behaves when two ports of the antenna are radiating at 32.5 GHz.

3.3 Current Distributions

See Fig. 7.

3.4 Radiation Patterns

The radiation patterns of the proposed antenna are observed in the Fig. 6 at different frequencies of 9.5 GHz and 32.9 GHz at 9.5 GHz the antenna provides the dipole structure in the H-plane and omni-directional pattern in the plane as

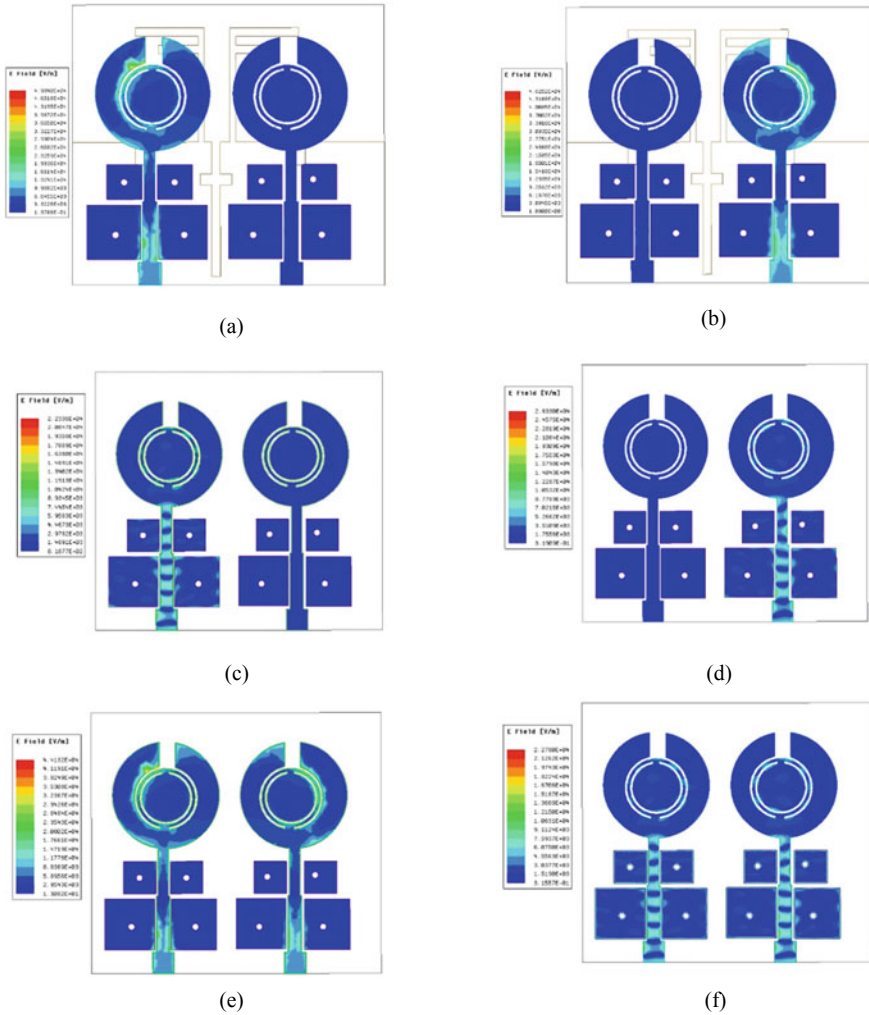


Fig. 6 E-field distribution at **a** 9.5 GHz with radiation at port 1 **b** 9.5 GHz with radiation at port 2 **c** 32.5 with radiation at port 1 **d** 32.5 with radiation at port 2 **e** 9.5 GHz with radiation at two ports **f** 32.5 with radiation at two ports

shown in Fig. 8a whereas at high frequencies like 32.9 GHz the antenna provides Quasi-omni-directional pattern in both the planes as shown in Fig. 8b.

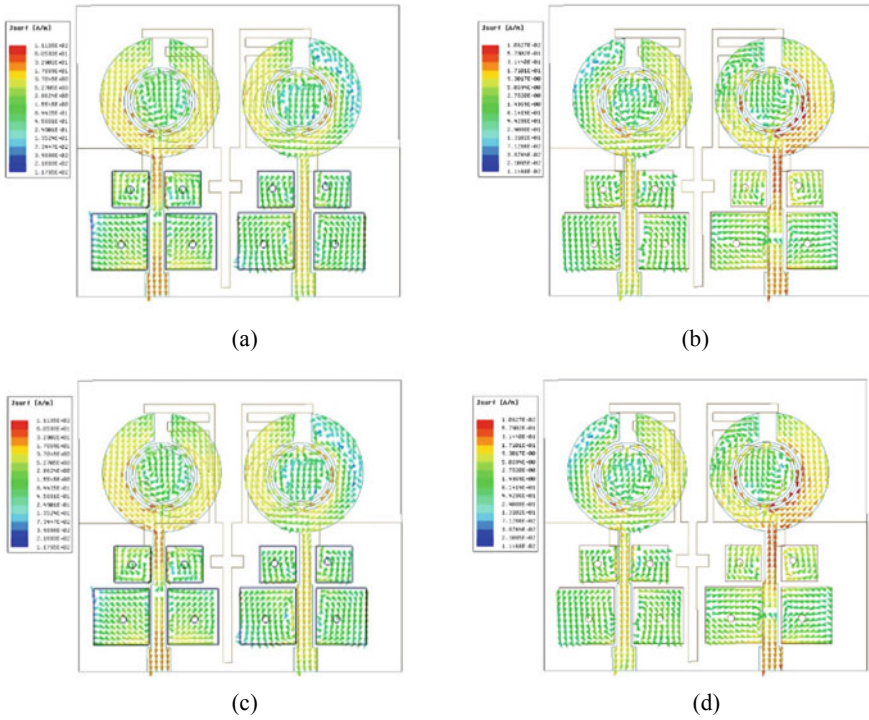


Fig. 7 Current distributions at **a** 9.5 GHz with radiation at port 1 **b** 9.5 GHz with radiation at port 2 **c** 32.5 with radiation at port 1 **d** 32.5 with radiation at port 2

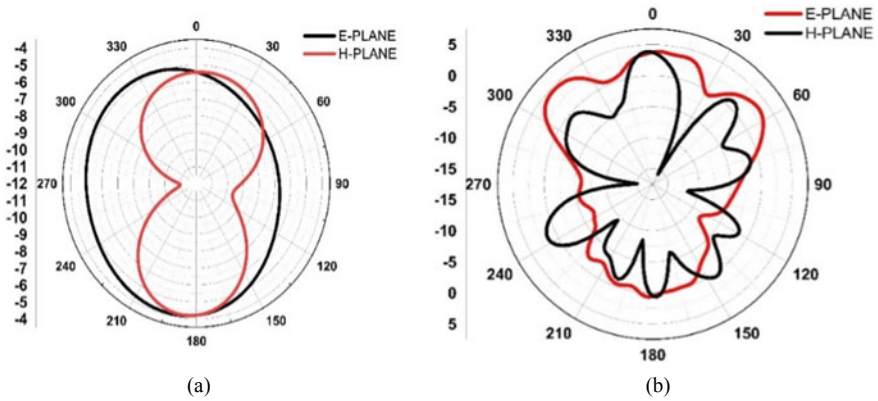


Fig. 8 **a** Radiation pattern at 9.5 GHz **b** radiation pattern at 32.9 GHz

4 Conclusion

In this work, isolation-reduced MIMO antenna is designed and analysed for 5g applications. The proposed antenna shows excellent characteristics up to 40 GHz covers modern applications and advanced communication application. The simulation results, such as ECC, transmission coefficient and DG, show the proposed work will be considered for future ready communication application. The model is analysed using HFSS and exhibits circular polarization in the working band area.

References

1. Yang Q, Wang K, Sun Y (2021) Quad-port miniaturized ultra-wideband MIMO antenna with metal vias. *Prog Electromagn Res Lett* 97:95–104
2. Mchbal A, Touhami NA, Elftouh H, Dkiouak A (2021) Isolation improvement using asymmetric radiators and ground plane diversity mechanism in a six-element UWB MIMO antenna design. *Prog Electromagn Res B* 91:19–38
3. Zhou JY, Wang Y, Xu JM, Du C (2021) A CPW-fed UWB-MIMO antenna with high isolation and dual band-notched characteristic. *Prog Electromagn Res M* 102:27–37
4. Zhang J, Wang L, Zhang W (2020) A novel dual band-notched CPW-fed UWB MIMO antenna with mutual coupling reduction characteristics. *Prog Electromagn Res Lett* 90:21–28
5. Tang ZJ, Liang L, Zhong B, Cheng L, Tan C, Hu S (2020) Uniplanar UWB-MIMO antenna with high isolation based on a radiator-ground shared structure. *Prog Electromagn Res Lett* 93:35–42
6. Irene G, Rajesh A (2020) Dual polarized UWB MIMO antenna with elliptical polarization for access point with very high isolation using EBG and MSR. *Prog Electromagn Res C* 99:87–98
7. Sultan KS, Abdullah HH (2019) Planar UWB MIMO-diversity antenna with dual notch characteristics. *Prog Electromagn Res C* 93:119–129
8. Chaudhary P, Kumar A, Yadav A (2020) Pattern diversity MIMO 4G and 5G wideband circularly polarized antenna with integrated LTE band for mobile handset. *Prog Electromagn Res M* 89:111–120
9. Bait-Suwailam MM, Almoneef T, Saeed SM (2021) Wideband MIMO antenna with compact decoupling structure for 5G wireless communication applications. *Prog Electromagn Res Lett* 100:117–125
10. Jaglan N, Dalal P, Gupta SD, Abdalla MA (2020) Band notched UWB MIMO/diversity antenna design with inductance boosted compact EBG structures. *Prog Electromagn Res C* 105:185–202

Metasurface-Enabled Fork-Shaped Antenna for 2.45 GHz ISM Band Wearable Applications



G. Srilatha, G. S. N. Raju, and P. A. Sunny Dayal

Abstract In this work, a fork-shaped radiator is simulated at 2.45 GHz for wearable applications. Rogers's 3010 is used as dielectric substrate. A fork-shaped design is used as radiating element. The antenna is symmetrical in shape and the two side strip-line dimensions are tuned to achieve the required operating frequency. Four unit cells of metasurface-inspired structure are placed backside of the antenna. The metasurface structure consists of rectangular strip line at center and two E-shaped structures placed at both sides of the rectangular strip in mirroring position. A partial ground with slot is used and placed as metasurface unit cells. The metasurface cells placed under the radiator helped in reduction of the specific absorption rate (SAR). Proposed design performance is presented with the comparative analysis using simulated results of S11, SAR curves, radiation patterns, and VSWR.

Keywords Wearable antenna · Specific absorption rate (SAR) · 2.45 GHz ISM band

1 Introduction

Wearable antennas are being developed for various ranges of frequencies for serving different daily life applications recently. They garnered much attention due to the smart wearable environments being developed for serving entertainment, sports, and health monitoring warbles. One of the main considerations in developing the wearable antennas is human safety. The radiation from the antenna should be maintained low to avoid the effect on human health when used for longer time. It is being analyzed by the average energy absorbed by human body over a selected tissue volume. The specific absorption rate (SAR) value indicates the amount of radiating energy being

G. Srilatha (✉) · G. S. N. Raju · P. A. Sunny Dayal
Department of ECE, CUTM, Vizianagaram 535003, India
e-mail: srilathacrts@gmail.com

G. Srilatha
Department of ECE, SIR CRRCOE, Eluru 534007, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
V. V. S. S. Chakravarthy et al. (eds.), *Advances in Signal Processing, Embedded Systems and IoT*, Lecture Notes in Electrical Engineering 992,
https://doi.org/10.1007/978-981-19-8865-3_23

247

absorbed by human body. As per the Federal Communications Commission, the minimum allowed SAR is 1.6 W/kg, which is over volume of 1 g tissue. For Council of the European Union, the recommended value is 2 W/kg, which is under volume of 10 g tissue. The SAR reduction is achieved using different techniques by several researchers over the years. One of the most commonly used techniques is using the metasurface-inspired structures as secondary layers as part of the ground planes.

Different kinds of metamaterial unit cells are used in different research articles. Some of the metamaterial-inspired structures that were used are electromagnetic bandgap structures (EBGs), split-ring resonators (SRRs), mushroom-like structure, artificial magnetic conductors (AMCs), etc., in [1] for the reduction of specific absorption rate, and split-ring resonator is used as superstrate. EBG based structures are used to reduce SAR for LTE antenna [2]. SAR reduction using metamaterial structures is investigated in [3]. Sometimes, the metamaterial structures are incorporated into the second layers. The second layer substrate used for metamaterial incorporation can be the same substrate used for the radiator layer, or in some cases, cloth materials are used in [4] a single metamaterial-inspired layer incorporated in the same substrate which is used in the first layer to reduce the SAR. In [5], jeans' cloth material holding metamaterial is used to reduce the SAR for CPW-fed antenna. Sometimes, these metamaterial unit cells are placed in the same layer instead of using the secondary layers. The metamaterial-like structures are placed along with antenna in the same plane as ground below the radiating antenna positioned appropriately to reduce the SAR effectively [6, 7].

In the current work, a fork-shaped antenna is designed, and the novel metasurface unit cells are placed under the antenna to reduce SAR to make the antenna usable for wearable applications. The metamaterial unit cell is made up of a single rectangular strip at center and E-shaped structures on both sides of the center rectangular strip positioned in a mirror to each other. The antenna is designed and analyzed to work at 2.45 GHz ISM band applications.

2 Proposed Fork-Shaped Antenna Design and Specifications

The designed wearable antenna dimensions are presented in Fig. 1. Rogers RO3010 with dielectric constant 10.2 and 0.0035 is value of loss tangent which is considered as dielectric substrate. The geometry of the designed antenna which is resonated at 2.45 GHz is optimized by the parametric study, and the dimensions with optimal values are presented in Fig. 1. The overall dimensions of the antenna restricted to only 34 mm * 34 mm * 1.27 mm. The metasurface unit cell size is 11 mm * 11 mm (Table 1).

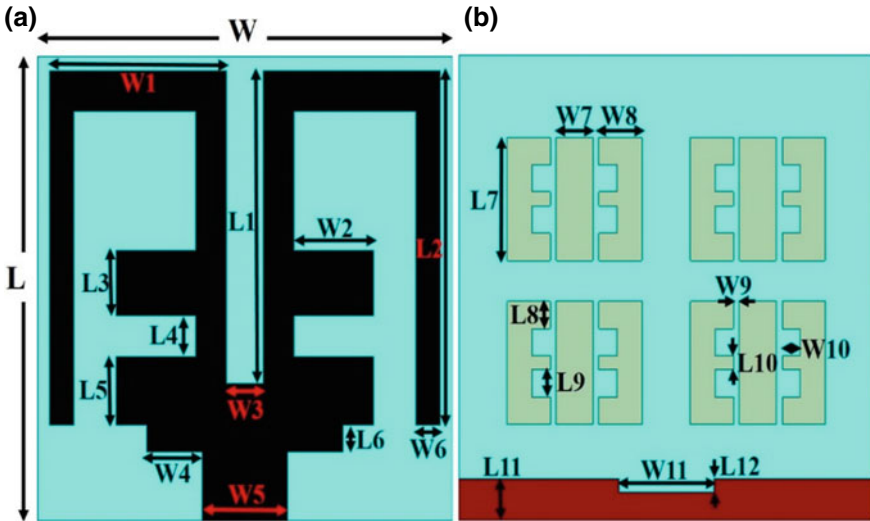


Fig. 1 Proposed 2.45 GHz ISM band wearable antenna **a** top view and **b** bottom view

Table 1 Optimal dimensions (in mm)

$L = 34$	$L1 = 23$	$L2 = 26$	$L3 = 4.8$	$L4 = 3$
$L5 = 5$	$L6 = 2$	$L7 = 9$	$L8 = 2$	$L9 = 2$
$L10 = 1$	$L11 = 3$	$L12 = 1$	$W = 34$	$W1 = 14.5$
$W2 = 6.5$	$W3 = 3$	$W4 = 4.5$	$W5 = 7$	$W6 = 2$
$W7 = 3$	$W8 = 3.5$	$W9 = 0.5$	$W10 = 1.5$	$W11 = 7.8$

3 Off-Body Simulation Analysis

The designed antenna is simulated with HFSS 19.2v. Initially, the antenna is simulated under free-space (off-body) condition, and the simulated analysis is presented in the following Figs. 2 and 3. The reflection coefficient curve presented in Fig. 2a shows that the proposed antenna operating frequency is obtained at 2.45 GHz with return loss -21.48 dB. The operational band covers from 2.42 to 2.47 GHz with band width 50 MHz. The VSWR curve shown in Fig. 2b indicates that, in the operational band, the VSWR is maintained in between 1 and 2. The field distribution curves shown in Fig. 3 represents the E, H and J field distributions at 2.45 GHz the concentration of arrow marks and the color representation explains the maximum or minimum field distributions in the radiating element.

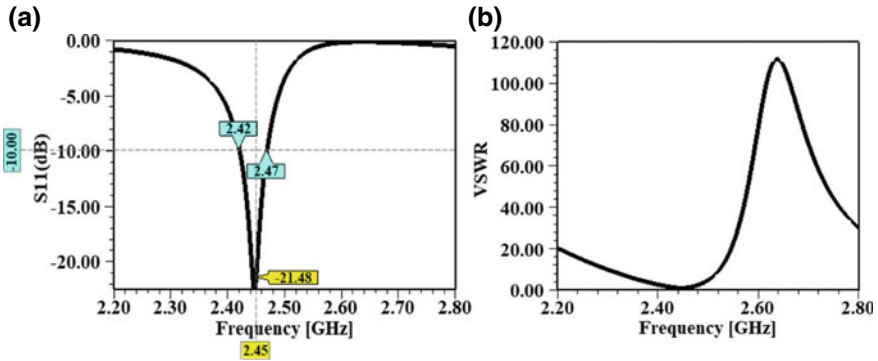


Fig. 2 ISM band wearable antenna proposed at 2.45 GHz **a** S11 versus frequency curve and **b** VSWR curve

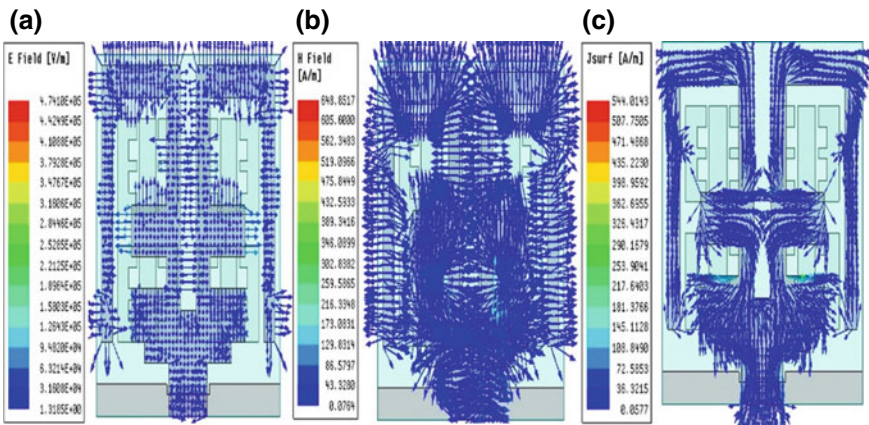


Fig. 3 Field distributions curves for proposed 2.45 GHz ISM band wearable antenna **a** E-field, **b** H-field, and **c** J-field

4 On-Body Simulation and Comparison Discussion

The proposed radiator performance is influenced by human body presence and that effect was evaluated by analyzing the antenna under layered human phantom condition. Here, the proposed antenna and the four-layered phantom model are separated with a minimum distance of 10 mm. To avoid radiation effect on human body the antenna is fed by low power input of 50 mW. To analyze the radiation effect on human body, the designed antenna is placed on top of four-layered tissue model positioned at center and the analysis is carried out in flat condition. The layered human phantom model consists of skin fat, bone, and muscle. Here, in Table 2, the dielectric values of corresponding tissues at 2.45 GHz were listed. The dimensions of these layers 100 mm * 100 mm in XY directions. The thickness of these layers in Z direction is

different for each layer based on average human body, and the thickness is maintained as follows: 2 mm for skin, 5 mm for fat, 10 mm each for muscle and bone.

Proposed antenna along with four-layered human tissue phantom model in top view, diagonal view, and front view is shown in the following Fig. 4. As per the setup in Fig. 4, the analysis carried out for on-body condition and the comparative analysis for off- and on-body conditions are presented in Fig. 5 with the help of S_{11} versus frequency and VSWR curves and in Fig. 6 with the help of 2D radiation patterns and 3D radiation plot.

From Fig. 5 data, one can see that for on-body and off-body conditions, there is little shift in operating frequency, but the intended operating frequency of 2.45 GHz

Table 2 Dielectric values of human body tissues at 2.45 GHz

S. No	Tissue	Relative permittivity (ϵ_r)	Loss tangent	Conductivity (S/m)
1	Bone	11.381	0.2542	0.39431
2	Skin	42.853	0.27255	1.5919
3	Fat	5.2801	0.14524	0.10452
4	Muscle	52.729	0.24194	1.7388

Fig. 4 Antenna on human four-layered phantom **a** top view, **b** diagonal view, **c** front view

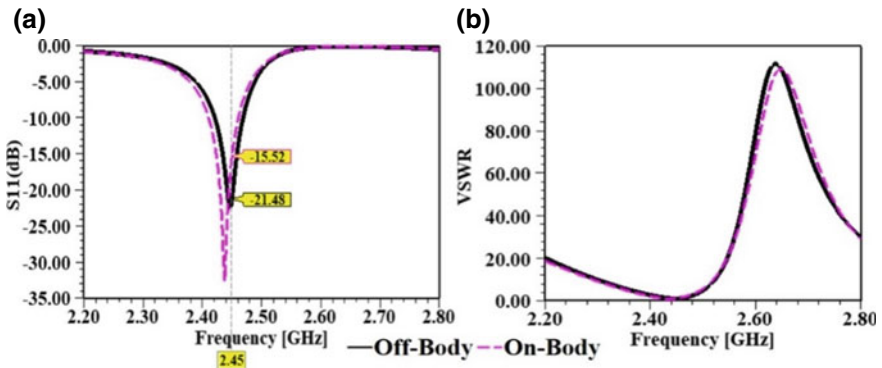
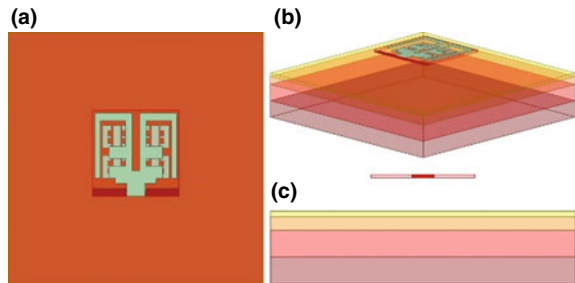


Fig. 5 Simulated comparison for off- and on-body conditions **a** S_{11} versus frequency, **b** VSWR

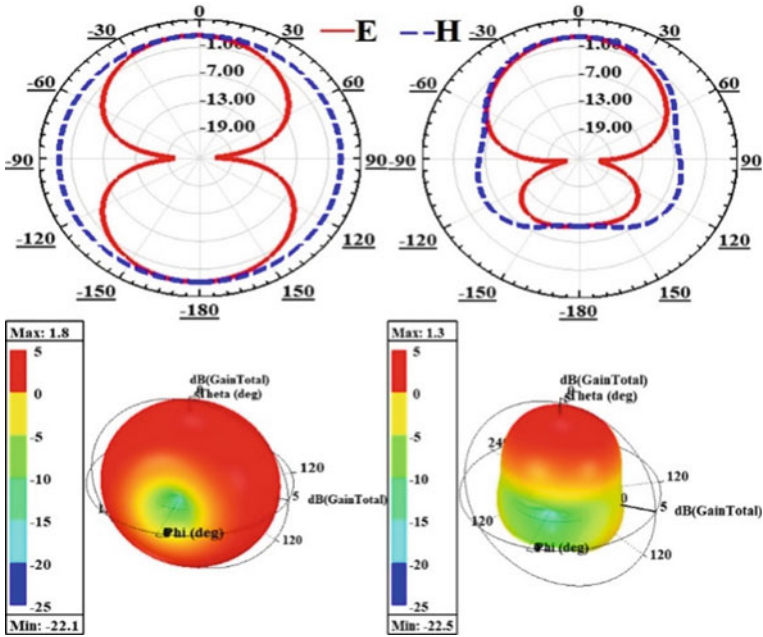


Fig. 6 Simulated comparison of radiation patterns: left side 2D and 3D plots are for off-body condition and right side 2D and 3D plots are for on-body condition

is maintained for both off- and on-body conditions. The shift in operating frequency toward lower resonance for on-body condition is with the presence of variations in the dielectric conditions of the human tissues. The VSWR curve for both off- and on-body conditions is maintained same leaving fewer effects on impedance matching.

From Fig. 6, it can be seen that the presence of the human body at backside of the antenna changed the antenna radiation pattern. The energy at backward of the antenna is affected due to some of the energy which could be absorbed by human body and some energy gets reflected and reconfigured due to human phantom presence.

For the four-layered phantom model shown in Fig. 4, the SAR analysis is carried out and presented in Fig. 7. For the input power 50 mW and with 10 mm minimal distance from phantom, the SAR obtained is 1.0923 W/kg. As per the Federal Communications Commission for 1 g volume tissue, the specific absorption rate value should be below 1.6 W/kg. For the designed antenna, the SAR value is much below than the standard value. It indicates that the designed antenna is suitable for wearable applications. Some other antenna parameters' comparison for on-body and off-body conditions is mentioned in Table 3.

Fig. 7 SAR analysis of proposed antenna

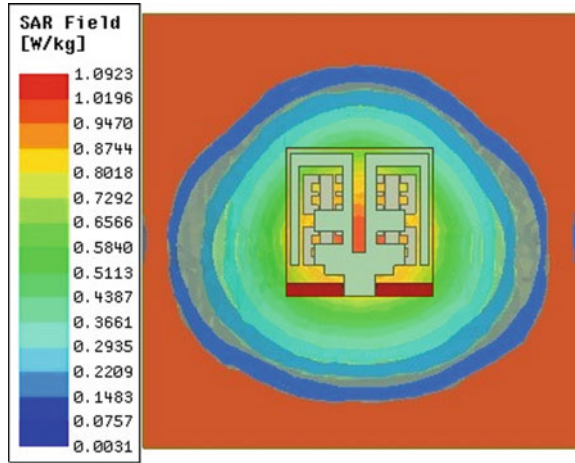


Table 3 Antenna parameters' comparison for off- and on-body conditions

Quantity	Off-body	On-body
Max U mW/sr	120.520419	5.178961
Peak directivity	1.730912	4.404387
Peak gain	1.525388	1.339194
Radiation efficiency	0.881262	0.304059
Front-to-back ratio	1.087684	14.762755

5 Conclusion

A fork-shaped wearable antenna at 2.45 GHz WBAN applications is designed. The simulated analysis presented shown that the antenna performance is good in both off- and on-body conditions. It can be used for normal 2.45 GHz ISM band applications, and Fig. 5 shows that it can also be used for on-body applications as it maintained the same performance with little effect of the human phantom presence. The positioning of the metasurface unit cells under the antenna at backside of the dielectric substrate helped in isolating the on-body loading effects and thus helped in minimizing the shift in operating frequency to minimum. SAR analysis plot shown in Fig. 7 indicates that the antenna maintained 1.0923 W/kg SAR value which is under the specified standard value 1.6 W/kg, which makes it good candidate for wearable applications.

References

1. Imaculate Rosaline S, Raghavan S (2015) Design and analysis of a SRR superstrate for SAR reduction. *J Electromagn Waves Appl* 29(17):2330–2338. <https://doi.org/10.1080/09205071.2015.1091384>
2. Munde M, Nandgaonkar A, Deosarkar S (2019) Low specific absorption rate antenna using electromagnetic band gap structure for long term evolution band 3 application. *Prog Electromagn Res M* 80:23–34. <https://doi.org/10.2528/pierm18102103>
3. Wang M, Yang Z, Wu J, Bao J, Liu J, Cai L, Li E (2018) Investigation of SAR reduction using flexible antenna with metamaterial structure in wireless body area network. *IEEE Trans Antennas Propag* 66(6):3076–3086. <https://doi.org/10.1109/tap.2018.2820733>
4. Janapala DK, Nesasudha M, Neebha TM, Kumar R (2019) Specific absorption rate reduction using metasurface unit cell for flexible polydimethylsiloxane antenna for 2.4 GHz wearable applications. *Int J RF MicrowComput Aided Eng* 29:e21835. <https://doi.org/10.1002/mmce.21835>
5. Abdu A, Zheng H-X, Jabire HA, Wang M (2018) CPW-fed flexible monopole antenna with H and two concentric C slots on textile substrate, backed by EBG for WBAN. *Int J RF Microwave Comput Aided Eng* 28(7):e21505. <https://doi.org/10.1002/mmce.21505>
6. Janapala DK, Nesasudha M, Neebha TM, Kumar R (2019) Flexible PDMS antenna backed with metasurface for 2.4 GHz wearable applications. In: 2019 IEEE 1st international conference on energy, systems and information processing (ICESIP). <https://doi.org/10.1109/icesip46348.2019.8938235>
7. Latha GS, Raju GSN, Sunny Dayal PA (2020) Design and analysis metamaterial inspired wearable antenna for 2.45 GHz ISM band. In: 2020 32nd international conference on microelectronics (ICM), pp 1–4. <https://doi.org/10.1109/ICM50269.2020.9331772>

Segmentation of Cell Periphery from Blood Smear Images Using Dark Contrast Algorithm and K-Medoid Clustering



Siddharth Verma, Vikrant Bhateja, Sourabh Singh, Sparshi Gupta, Ayush Dogra, and Nguyen Gia Nhu

Abstract Computer-Aided Analysis of Blood Smear Images helps to identify several cell features which cannot be analyzed with the existing manual techniques. For this purpose, segmentation of required cell component is very important. The motive of this work is to segment the Cell Periphery, which holds the cytoplasm, from the Blood Smear Images. Primarily, these images are enhanced to increase the observability of various cell components. The enhancement is done using Dark Contrast Algorithm (DCA). This enhanced image is further segmented using K-Medoid Clustering, a technique based on spatial clustering. The number of clusters obtained as output are predefined. This technique clusters the data with the help of Similarity Index, to distribute it according to their similarities or dissimilarities by updating the medoids. The image hence obtained gives us the segmented Cell Periphery. The assessment of this proposed approach is done using parameters like Second Derivative like Measure of Enhancement (SDME) and Measure of Enhancement (EME) for enhancement and Dice Coefficient (DC) for segmentation.

Keywords Blood smear images · DC · DCA · EME · K-Medoid · SDME · Similarity index · Spatial clustering

S. Verma · V. Bhateja (✉) · S. Singh · S. Gupta

Department of Electronics and Communication Engineering, Shri Ramswaroop Memorial College of Engineering and Management (SRMCEM), Faizabad Road, Lucknow, UP 226028, India
e-mail: bhateja.vikrant@gmail.com

Department of Electronics Engineering, Faculty of Engineering and Technology, Veer Bahadur Singh Purvanchal University, Shahganj Road, Jaunpur, Uttar Pradesh, India

A. Dogra

Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India

Ronin Institute, Montclair, NJ 07043, USA

N. G. Nhu

Duy Tan University, 254 Nguyen Van Linh Str, Da Nang City, Vietnam
e-mail: nguyengianhu@duytan.edu.vn

1 Introduction

Cell Periphery is important in identifying several contrasting features of leukocytes (commonly known as White Blood Cells). They help us in identifying the type of cell whether blast or not and in also identifying various types of WBCs. The conventional methods to identify type of cell are through manual observations which is both tedious and is subjected to inaccuracy [1]. Now, to rectify these problems, Computer-Aided Diagnosis is used for analysis of Blood Smear Images. These images are obtained by mounting digital camera on microscope, either optical or electron, used for observing blood slides. These Blood Smear Images are processed to extract cells' characteristics to categorize them into their types [2]. Harun et al. [3] proposed a technique of nucleus segmentation from Blood Smear Images with a median filter and Dark Contrast Algorithm (DCA). Segmentation is done through thresholding and clustering. Classification of WBCs done by Dasariraju et al. [4] used Multi-Otsu Thresholding for segmentation of nucleus and cell boundary. The preprocessing step in this technique was the conversion from RGB to lab scale. The accuracy of this whole proposed work was moderate to high. A technique of segmentation through clustering was proposed by Su et al. [5] for bone marrow smear images. The segmentation here was done in two stages: firstly, K-Means clustering algorithm was used to divide the image into two clusters further and Hidden Markov Random Field (HMRF) was used for final segmentation. This method of segmentation was able to label six types of WBCs. A similar clustering approach was proposed by Acharya et al. [6] for segmentation of Blood Smear Images with the help of K-Medoid Clustering. This method is observed to be much effective than K-Means clustering. Another method of segmentation by morphological filters [7] to extract nucleus from Blood Smear Images is quite effective, but was not able to segment cell boundary. Eckardt et al. [8] proposed a deep learning approach to detect NPM1 mutation with the help of bone marrow smears. It was done using Faster Region-based Convolutional Neural Net (FRCNN) to improvise cell borders and an image annotator tool. These approaches conclude that enhancement preceded by segmentation gives best results in order to identify cell components. Experimentally, the accuracy of K-Medoid when compared to K-Means for segmentation is quite high. K-Means is technique which works on iterative clustering, while K-Medoid works on spatial clustering. The approach opted for this paper uses DCA for contrast enhancement as it works on darker regions of the image and K-Medoid Clustering for segmentation of the Cell Periphery. The organization of the paper in the following parts is as follows: Sect. 2 contains the proposed approach of enhancement and segmentation to extract Cell Periphery, Sect. 3 contains experimental results and analysis and Sect. 4 consists of conclusion.

2 Proposed Approach of Enhancement and Segmentation to Extract Cell Periphery

The proposed approach to segment Cell Periphery from Blood Smear Images follows enhancement preceded by segmentation. To perform enhancement, firstly, the RGB Blood Smear Image is converted to grayscale image so that there is no error to define Cell Periphery. The enhancement done using Dark Contrast Algorithm (DCA) highlights the Cell Periphery further making it easy to segment. The segmentation done by K-Medoid Clustering converts this enhanced image into clusters of similar data.

2.1 Enhancement Using Dark Contrast Algorithm (DCA)

The enhancement approach used in this experimental work is through DCA which is a contrast stretching technique and stretches the contrast of darker regions. The working of DCA is dependent on its parameter which is Threshold Value (TH) and Stretching Factor (NTH). For DCA, the value of TH is less than that of NTH . The control function on which DCA works is given in Eq. (1) [3].

$$\left(O(x, y) = \begin{cases} \frac{p(x,y) - \min_p}{TH - \min_p} \cdot NTH & \text{if } p(x, y) < TH \\ \frac{p(x,y) - TH}{\max_p - TH} \cdot (255 - NTH) + NTH & \text{if } p(x, y) \geq TH \end{cases} \right) \quad (1)$$

Here, $p(x, y)$ and $O(x, y)$ represent the input and output pixels, respectively. The maximum and minimum intensities of input image are represented by \max_p and \min_p , respectively. Equation (1) shows both the cases of DCA, that is, stretching and compression. The assessment of the enhanced image is done using Image Quality Assessment (IQA) parameters like Second Derivative like Measure of Enhancement (SDME) and Measure of Enhancement (EME) [9, 10].

2.2 Segmentation of Cell Periphery Using K-Medoid Clustering

Further, to segment the Cell Periphery from the enhanced image, K-Medoid Clustering is used. This is a type of spatial clustering where the clusters are divided according to the similarity in the data. The difference between K-Means [5] and K-Medoid, both being clustering algorithms, is K-Medoid being more sensitive to dissimilarities. K-Medoid Clustering takes medoids into consideration to cluster the data [6]. Figure 1 shows the graphical representation of data clustered by K-Medoid Clustering into two clusters.

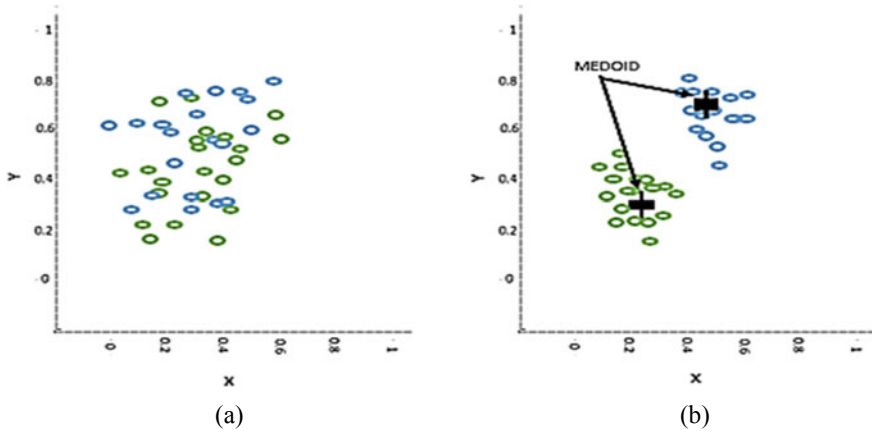


Fig. 1 Graph showing **a** input data and **b** two clusters of data

In K-Medoid Clustering, the clusters are defined as inputs and medoid is first chosen as any random value. Medoid refers to a point in the cluster whose dissimilarities are minimum. These medoids are further updated according to the dissimilarity or similarity of data. The similarity measure on which this technique works is given by Eq. (2) [6].

$$V = \sum_i^k \sum_{x_j \in S_i} (x_j - c_i)^2 \quad (2)$$

Here, k is the number of clusters, C_i is the medoid, x_j is the object and S_i is the cluster. With the help of this similarity, the medoid is updated and is further compared with the medoid which is predefined. No change in medoid indicates similar data, while change in medoid indicates the data of other category [11]. In the proposed approach, the number of clusters are set to be two since the contrast-enhanced image is a gray scale. This clustering helps us to distinguish Cell Periphery from nucleus since both have different levels of intensity. Segmentation of Cell Periphery using K-Medoid gives us information regarding the disorientations of the same which helps us to identify blast, malignant cells. The segmented image of Cell Membrane hence is evaluated using Dice Coefficient (DC) [12]. The complete algorithm for the proposed approach of enhancement and segmentation to extract the Cell Periphery of Blood Smear Images is given in Algorithm I.

Algorithm 1: Procedural Steps for the Proposed Approach of Enhancement and Segmentation.

Begin**Module 1:** Enhancement using DCA.

- Step 1:** *Input image as p .*
Step 2: *Convert p to grayscale.*
Step 3: *Initialize DCA Parameters.*
 $TH \leftarrow 150$
 $NTH \leftarrow 175$
- Step 4:** *Read parameters of image p .*
 $N_r \leftarrow$ Number of rows
 $N_c \leftarrow$ Number of columns
 $max_p \leftarrow$ Maximum Intensity
 $min_p \leftarrow$ Minimum Intensity
- Step 5:** *For $i = 1$ to N_r*
Step 6: *For $j = 1$ to N_c*
Step 7: *If $p(i, j) < TH$*
 $p(i, j) = [(p(i, j) - min_p) / (TH - min_p)] * NTH$
 Else
 $p(i, j) = [(p(i, j) - TH) / (max_p - TH)] * (255 - NTH) + NTH$
 End If
 End For
 End For
- Step 8:** *Display Contrast Enhanced Image p .*

Module 2: Segmentation using K – Medoid Clustering.

- Step 1:** *Initialize Number of Clusters, $k \leftarrow 2$ and object S_i as p .*
Step 2: *Read medoid C_i .*
Step 3: *Calculate similarity index (V) using Eq. (2).*
Step 4: *Update C_i as C_j .*
Step 5: *If $C_j = C_i$*
 Display Segmented Image.
 Else
 Repeat from Step 2.
 End If

End

3 Experimental Results

3.1 Dataset

The database of Blood Smear aspirates for this experiment is obtained from the online repository of Cancer Imaging Archive [4], for single-cell images, and ASH [6], for

multi-cell images. These images are acquired by a digital camera with dimensions 400×400 and 720×960 , respectively. Two images one from each dataset were taken to analyze the results of the proposed approach to segment Cell Periphery. As initial steps of preprocessing, these images were converted from their original color model to gray scale for better identification in segmentation.

3.2 Simulation Results and Analysis

Two images chosen from the dataset were used to observe the proposed approach of enhancement using DCA and segmentation of Cell Periphery using K-Medoid Clustering Algorithm [3, 6]. The effectiveness of the DCA-enhanced image was compared with the original image on the basis of parameters like SDME and EME [9, 10], while the assessment of segmented image was done using DC by comparing it with the ground truth. The results from both the images are shown in Fig. 2.

The visibility of images (c) and (g) is better than that of (b) and (f) as shown in above Fig. 2. The cell components and periphery are much visible after enhancement. Further, the segmented image (d) and (h), when observed visually, shows us the successfully segmented Cell Periphery. These enhanced and segmented images are also evaluated qualitatively. The IQA metrics of enhanced images are given in Table 1.

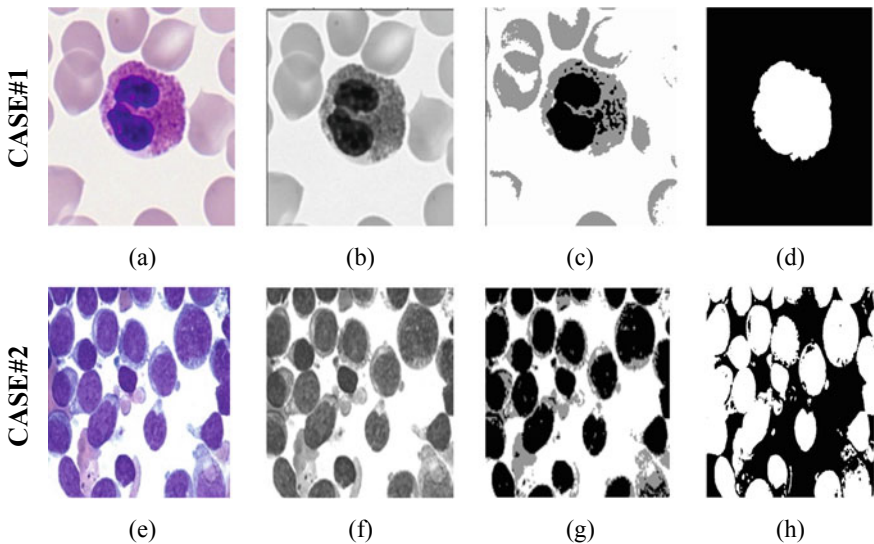


Fig. 2 Experimental results of images showing: **a** and **e** input Blood Smear Image, **b** and **f** grayscale-converted image, **c** and **g** DCA-enhanced image and **d** and **h** segmented image of cell periphery

Table 1 IQA metrics for input and enhanced images

Metric	Single-cell image (CASE#1)		Multi-cell image (CASE#2)	
	input image	Enhanced image	Input image	Enhanced image
SDME	23.45	25.67	24.58	28.77
EME	3.332	4.53	5.56	6.10

Table 2 Performance evaluation of segmentation technique

	Single-cell image (CASE#1)	Multi-cell image (CASE#2)
Metric	Cell periphery	Cell periphery
DC	0.893	0.92

On observing these values, it shows a clear increase in both SDME and EME by enhancement using DCA, thus increasing its metrics qualitatively. This indicates that the enhanced image has better contrast and brightness. Similarly, the assessment of segmented image of Cell Periphery is done using Dice Coefficient (DC) [12] as shown in Table 2.

The value of DC nears to 1 in the both types of images, which clearly indicates that the Cell Periphery is segmented successfully when compared with the ground truth results. So, visually and qualitatively, the performance of segmentation is evaluated and is successful. Therefore, this approach of enhancing the Blood Smear Images for segmenting Cell Periphery is successful in giving the desired results.

4 Conclusion

This paper presents an approach to segment the cell membrane/Cell Periphery from Blood Smear Images. Primarily, DCA was used for enhancement of these images to increase visibility of Cell Membrane. This was successful since the Cell Periphery can be seen in the enhanced image. This image was evaluated using SDME and EME, which shows that the enhancement was successful. Now, for segmenting our required ROI, K-Medoid Clustering was used. Here, the number of clusters were initialized to 2. The clusters hence obtained from segmentation contain the Cell Periphery. The results of segmentation were appraised using DC, which clearly show successful segmentation of Cell Periphery. This segmented image can be further used to calculate features like Nucleus-to-Cytoplasm Ratio.

References

1. Jagadev P, Virani HG (2018) Detection of Leukemia and its types using image processing and machine learning. In: Proceeding of international conference on trends in electronics and informatics (ICTEI 2017). IEEE, Tirunelveli, India, pp 522–526
2. Rezatofighi SH, Zadeh HS (2011) Automatic recognition of five types of white blood cells in peripheral blood. *Comput Med Imaging Graph* 35(4):333–343
3. Harun NH, Bakar JA, Hambali HA, Khair NM, Mashor MY, Hassan R (2018) Fusion noise—removal technique with modified algorithm for robust segmentation of acute leukemia cell images. *Int J Adv Intell Inf* 4(3):202–211
4. Dasariraju S, Huo M, McCalla S (2020) Detection and classification of immature leukocytes for diagnosis of acute myeloid leukemia using random forest algorithm. *Bioengineering* 7(4):120–131
5. Su J, Liu S, Song J (2017) A segmentation method based on HMRF for the aided diagnosis of acute myeloid leukemia. *Comput Methods Prog Biomed* 152(7):115–123
6. Acharya V, Ravi V, Pham TD, Chakraborty C (2021) Peripheral blood smear analysis using automated computer-aided diagnosis system to identify acute myeloid Leukemia. *IEEE Trans Eng Manage* 1–14
7. Hegde RB, Prasad K, Hebbar H, Singh BMK (2018) Development of a robust algorithm for detection of nuclei and classification of white blood cells in peripheral blood smear images. *J Med Syst* 42:110
8. Eckardt JN, Middeke JM et al (2022) Deep learning detects acute myeloid leukemia and predicts NPM1 mutation status from bone marrow smears. *Leukemia* 36:111–118
9. Trivedi M, Jaiswal A, Bhateja V (2013) A No-reference image quality index for contrast and sharpness measurement. In: 3rd IEEE international advance computing conference (IACC). IEEE, India, pp 1234–1239
10. Prajapati P, Narmawala Z, Darji NP, Moorthi SM, Ramakrishnan R (2015) Evaluation of perceptual contrast and sharpness measures for meteorological satellite images. In: Soni AK, Lobiyal DK (eds) 3rd International conference on recent trends in computing (ICRTC), procedia computer science, vol 57. Springer, India, pp 17–24
11. Halder A, Dasgupta A, Ghosh S (2022) Image segmentation using rough—Fuzzy K—medoid algorithm. In: Proceeding of international conference on communications, devices and intelligent systems. IEEE, India, pp 105–108
12. Kumar SN, Lenin Fred A, Ajay Kumar H, Sebastin Varghese P (2018) Performance metric evaluation of segmentation algorithms for gold standard medical images. In: Sa P, Bakshi S, Hatzilygeroudis I, Sahoo M (eds) Recent findings in intelligent computing techniques. *Advances in intelligent systems and computing*, vol 709. Springer. Singapore (2018)

Groundwater Quality Assessment of Raipur City Using Machine Learning Models



Anushree Shrivastava, Mridu Sahu, and D. C. Jhariya

Abstract Groundwater has been a vital source of water consumption across India. Raipur, the capital city of Chhattisgarh, has been utilizing this resource for water consumption and utilization for Irrigation. While water is a necessity for survival, it is also important that water which we uptake is fit for consumption. Groundwater often gets contaminated by the fertilizers and pesticides by affecting the concentration of the major ions and other parameters present in the water. Groundwater Quality Index is a measure to determine the quality of water, which is calculated based on some physicochemical parameters and ions that water contains. The Water Quality Index (WQI) is a really useful measure for assessing the overall water quality. It simplifies the interpretation of information by lowering a huge number of data points to a single value. The WQI is used to assess whether or not groundwater is suitable for drinking. In this paper, the Water Quality Index was calculated based on pH, TA, TH, Chloride, Nitrate, Fluoride, and Calcium. Further, the quality of groundwater was assessed using various Machine Learning Models, namely, Logistic Regression, Decision Tree Classifier, Gaussian NB, Random Forest Classifier, Linear SVC, and XGB Classifier. The best classification was shown by Random Forest Classifier with an outstanding of 100 percent accuracy.

Keywords Logistic regression · Decision tree classifier · Gaussian NB · Random forest classifier · Linear SVC · XGB classifier

A. Shrivastava · M. Sahu (✉) · D. C. Jhariya
Department of Information Technology, National Institute of Technology, Raipur, India
e-mail: mrisahu.it@nitrr.ac.in

A. Shrivastava
e-mail: ashrivastava.mtech2021.it@nitrr.ac.in

D. C. Jhariya
e-mail: dcjhariya.geo@nitrr.ac.in

1 Introduction

Groundwater quality assessment is essential for determining the quality of water in a region as water is not only a basic necessity in every household but also an important resource for survival. Conservation of water has been a need of the hour, as freshwater is getting contaminated due to various factors like the use of fertilizers, pesticides, and industrial wastes. Thus, the study of freshwater resources like groundwater is becoming important.

The study area for this paper is the capital city of Chhattisgarh, i.e., Raipur. In Raipur, the Kharun River is the only supply of raw water currently available. Groundwater is another source of water, having a capacity of 22 million liters per day, in addition to water from the Kharun River.

The quality assessment of water from the groundwater sources is a tedious task as it involves manual work and analysis. The role of Machine Learning can be crucial in automating such tasks, especially in the scenario where lockdown is implemented due to the pandemic. The prediction using machine learning can prevent the ceasing of analysis of such crucial resources.

In this paper, we have used the concept of the Water Quality Index (WQI) for determining the water quality of the 44 different regions of Raipur. The Water Quality Index was calculated based on pH, TA, TH, Chloride, Nitrate, Fluoride, and Calcium. Further, the quality of groundwater was assessed using various Machine Learning Models, namely, Logistic Regression, Decision Tree Classifier, Gaussian NB, Random Forest Classifier, Linear SVC, and XGB Classifier.

2 Methodology

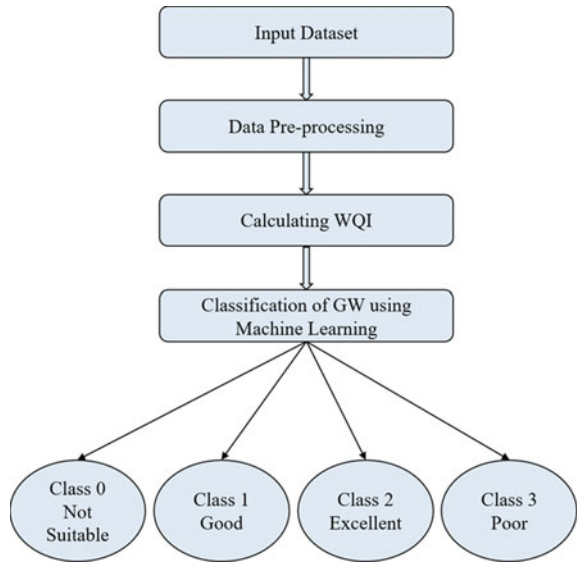
Figure 1 shows the methodology used in this work. The stages include input dataset, data preprocessing, calculating WQI, and classification of groundwater using Machine Learning Models.

2.1 Dataset Description

The dataset has been collected from the Geology Department of NIT, Raipur. It consisted of groundwater details of samples collected from 44 locations within Raipur and 24 features, namely, pH, EC, TDS, TH, TA, HCO_3^- , Cl^- , NO_3^- , SO_4^{2-} , F^- , Ca^{+2} , Mg^{+2} , Na^+ , K^+ , Si^{+4} , SSP, SAR, SAR,

KR, RSC, MR, and CR from which seven features have been used, namely, pH, TA, TH, Cl^- , NO_3^- , F^- , and Ca^{+2} for the calculation of WQI, which further form the basis of classification.

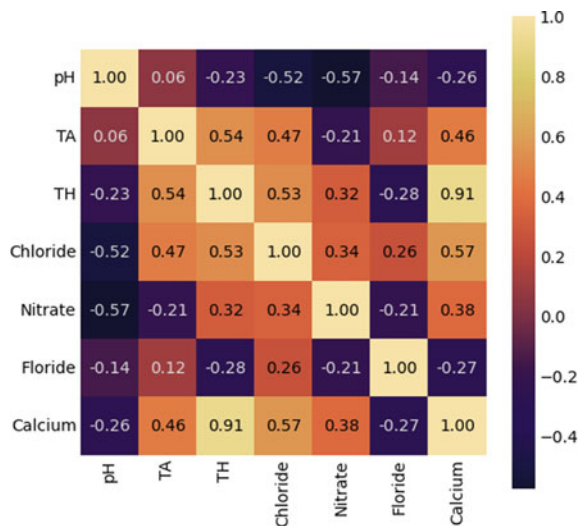
Fig. 1 Methodology used



2.2 Correlation of Features

Figure 2 shows the correlation of the seven parameters used for the calculation of WQI and Water Quality Classification. Correlation shows how the parameters are related to each other and whether the presence or absence of one parameter leads to the presence or absence of another parameter. Correlation is a mechanism used to identify the strength of relationships between features.

Fig. 2 Correlation matrix of the features



Correlation is of three types:

- Positive correlation: If there are two features which are having positive correlation, it means that if the value of one feature is increasing in a scenario, the value of the other feature will also tend to rise in a similar scenario;
- Negative correlation: If there are two features which are having negative correlation, it means that if the value of one feature is increasing in a scenario, the value of the other feature will tend to reduce in a similar scenario;
- No correlation: When two features are said to be in no correlation, it means that the value of the two features is independent, i.e., increasing or decreasing and the value of a particular feature will have no impact on the value of other features.

2.3 Calculation of Water Quality Index

Step 1. Calculation of weightage factor

The weight is assigned for each parameter as per the importance of the parameter in the water consumption. The relative weight is given by:

$$W_i = \Sigma w_i, \tag{1}$$

where W_i denotes relative weight, w_i denotes parameter weight, and n represents the number of parameters. The weight assigned to all the seven parameters and the relative weight is shown in Table 1.

Step 2. Calculating sub-index

For obtaining the Water Quality Index, firstly sub-index is calculated which is given by:

$$SI = W_i * (c/s) * 100, \tag{2}$$

Table 1 Weight, relative, and BIS standard value for each parameter

Parameter	BIS standard	Weightage	Relative weight
pH	6.5–8.5	2	0.133
TA	200–600	1	0.067
TH	200–600	1	0.067
Chloride	250–1000	3	0.133
Nitrate	45	5	0.267
Fluoride	1.0–1.5	4	0.200
Calcium	75–200	1	0.067
Magnesium	30	1	0.067
		Σw_i	ΣW_i

where W_i is the relative weight, c is the value of the parameter in the water sample in mg/l, and s is the standard value of the parameter mentioned in Table 1.

Step 3. Calculating Water Quality Index

$$W_i = \sum SI \tag{3}$$

The Water Quality Index is used to determine the category of water quality, namely, Excellent, Good, Poor, or Not Suitable for consumption. Table 2 shows the different classes of water quality into which the dataset is bifurcated according to the Water Quality Index.

Figure 3 shows the pie chart representation of the occurrence of different categories of water quality. Figure 4 shows the violin plot of the classes of water quality and how they are related to WQI. A violin plot uses density curves to represent numeric data distributions for one or more groups.

Figure 5 shows the distribution and box plot representation of the parameters based on the water quality category. The distribution curves and box plot depict how each feature is contributing to the water quality. In the distribution curve and box plot, green represents excellent water quality, blue represents good water quality, yellow represents poor water quality, and red represents water quality which is not suitable for consumption.

Table 2 Water quality classes

Category	Water quality index	Water quality
0	> 75	Not suitable
1	35–55	Good
2	< 35	Excellent
3	55–75	Poor

Fig. 3 Pie chart representation of the occurrences of different categories of WQ

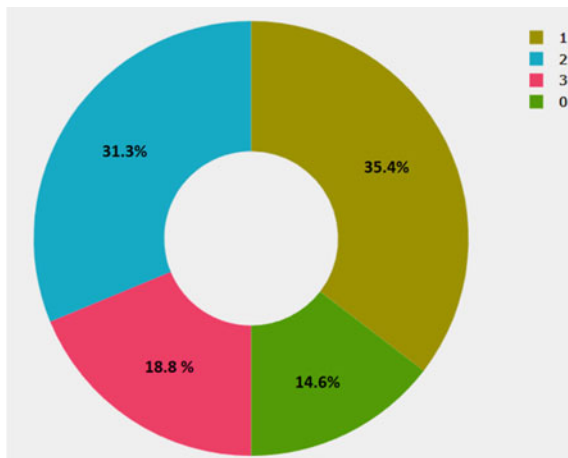
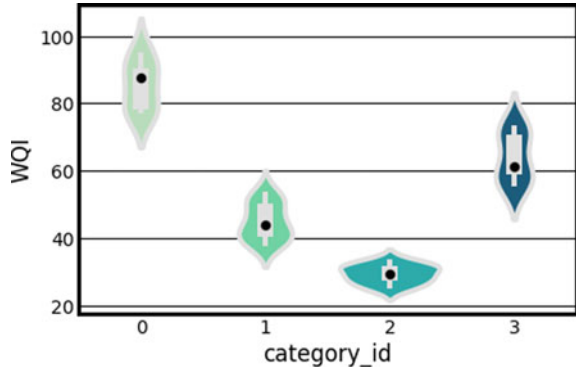


Fig. 4 Violin plot of water quality classes based on WQI



2.4 Machine Learning Models

2.4.1 Logistic Regression

This is a supervised learning-based classification approach. It works with discrete data and produces non-continuous output. However, instead of digital numbers, it outputs probabilistic values between zero and one. The minimum value is the most crucial part of Logistic Regression; it lays the foundation for classification and is used to determine if the outcome is nearer to one or zero. The recall and accuracy settings determine the threshold value. If both recall and accuracy are one, the threshold value is considered to be perfect.

However, this optimum condition does not always occur; thus, there are two possible approaches: one with great precision but low recall and the other with low precision but high recall. The threshold is established [1] based on the system's requirements. The Sigmoid function (see Fig. 6) is an S-shaped curve that is used to calculate the result depending on the threshold value in Logistic Regression. This is accomplished by putting the real values and threshold value onto the curve, then comparing the real values to the threshold value to see if they are closer to 1 or 0. The Logistic Regression model is ideal for binary classes, as can be seen from the preceding description, but it may also be utilized for numerous classes adopting the idea of one versus all [2].

2.4.2 Decision Tree Classifier

A Supervised Machine Learning Approach called a decision tree produces decisions based on a combination of rules. It constructs decision trees using historical data. Construction of the maximum tree, selection of the appropriate tree size, and classification of fresh data using the established tree are the three elements of this technique.

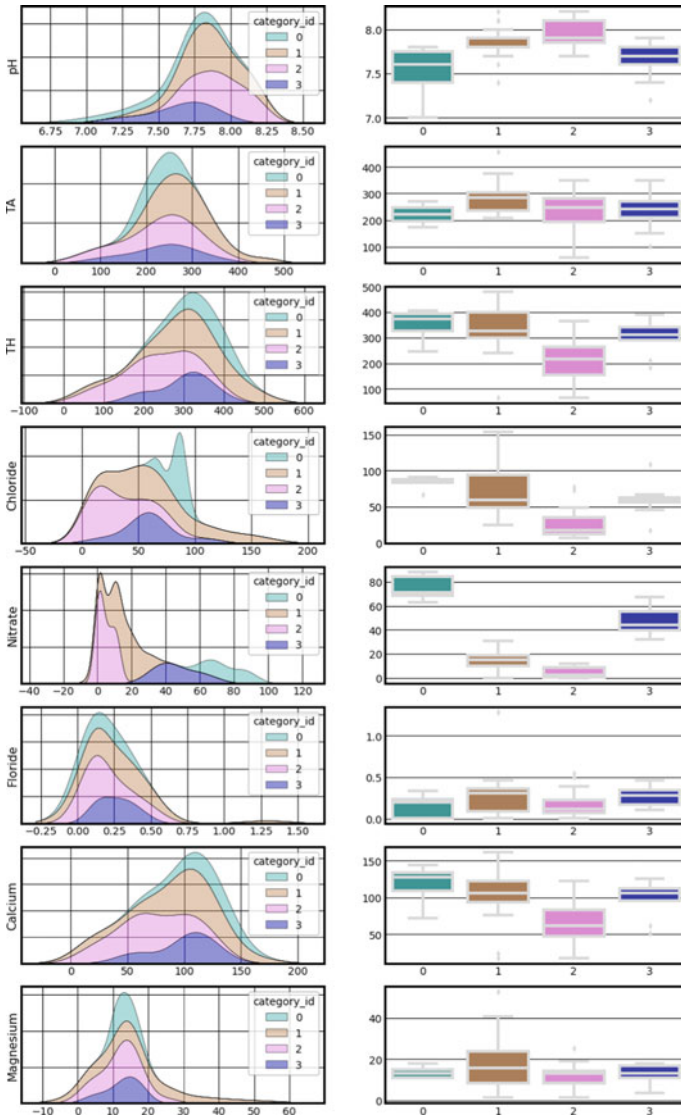


Fig. 5 Density distribution and box plot based on water quality category

Maximum tree creation takes a large amount of time. The observations are separated continually until only one class remains at the lowest possible level. Figure 7 depicts the maximum tree's creation. The figure to the left of the tree's root ought to be smaller than the figure to the right. The values that are referred here are parameters in the learning sample's parameter matrix.

Fig. 6 Logistic regression

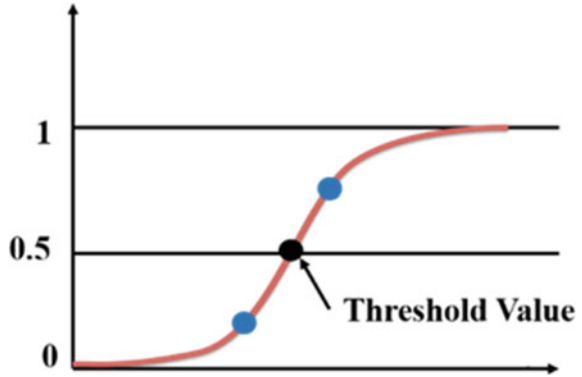
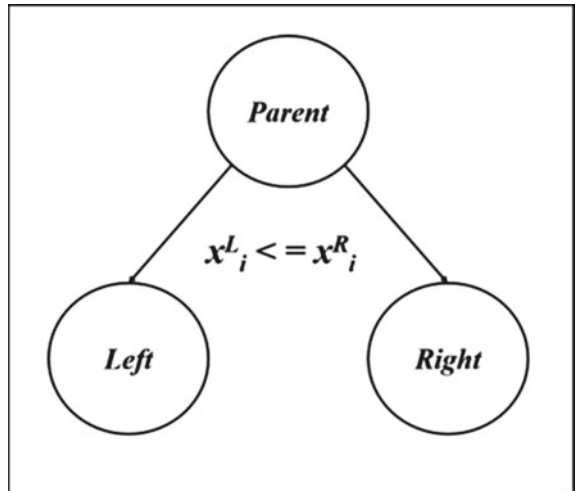


Fig. 7 Construction of maximum tree



The proper size tree must be chosen since the maximum tree might have many layers and be complex; thus, it must be optimized before categorization. This is intended to shrink the tree and remove any extraneous nodes. Finally, classification is completed, with classes allocated to each observation [3].

2.4.3 Gaussian Naïve Bayes

A statistical and supervised learning-based probabilistic classifier is yet another name for GNB. It is based on the probability hypothesis of Bayes. For classification, it employs a probabilistic technique. When used on massive datasets, it produces accurate findings in a shorter amount of time [4].

It has four types of probabilities: likelihood probability, which represents the probability of predictor, such that the class is given; predictor prior probability, which

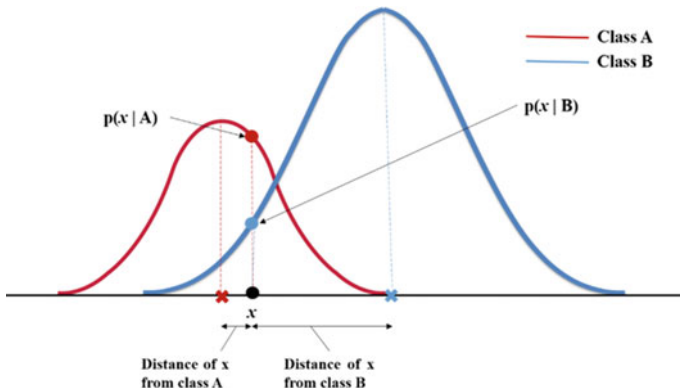


Fig. 8 GNB classifier

represents the probability of the previous predictor; class prior probability, which tells the probability of class; and posterior probability, which represents the probability of class such that the predictor is given. The existence or lack of one characteristic does not impact the existence or lack of another characteristic, according to the Naive Bayes approach. Each characteristic contributes to the categorization process in its own way [5].

We employed GNB, that is a type of NB classifier that is used for continuous data categorization. It is based on probability ideas of Gaussian normal distribution. Classes are believed to be represented by normal distributions with distinct dimensions. It is simple, computationally quick and only takes a modest amount of training data. However, since it presupposes characteristic independence, it may produce inaccurate estimates, thus the label naive. The workings of GNB are shown in Fig. 8. The categorization is based on the shortest distance between each data points (x).

2.4.4 Random Forest Classifier

This approach, as the name implies, uses many trees to arrive at a categorization choice. It is based on the notion of tyranny of the majority. The trees each make their own estimate, and the class with the highest count is chosen. The overall result is successful because the outputs of the multiple decision trees are combined; thus, even if some decision trees produce incorrect predictions, the other trees' outcomes cover up the final estimate [6]. Because it considers different tree options on a majority premise, Random Forest produces superior results. However, when there are a lot of trees, categorization can be stagnant and unproductive. This classifier's training is quick, but predictions are stagnant [7]. Figure 9 depicts the Random Forest Classifier's generic operation.

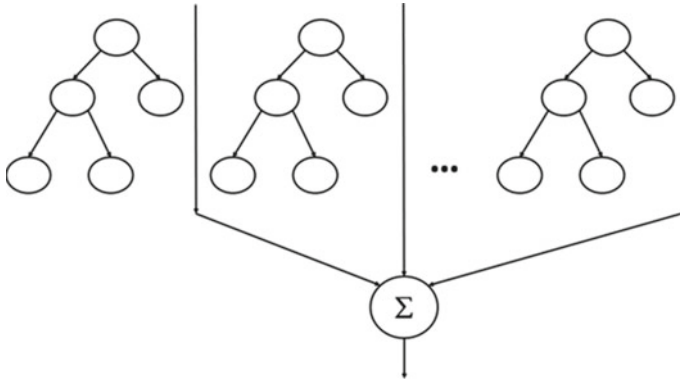
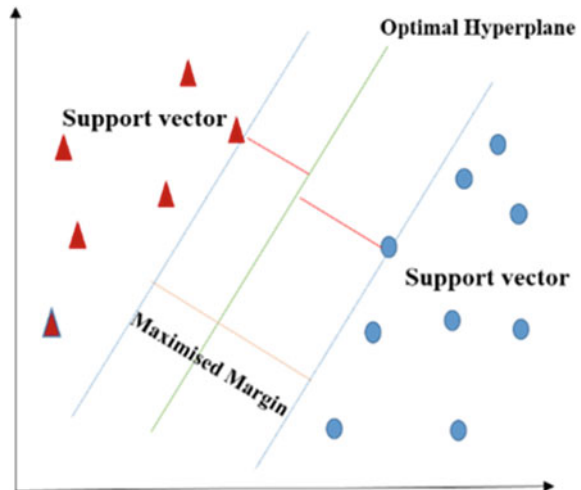


Fig. 9 Random forest classifier

2.4.5 Linear SVC

It is a kernel-based classifier that was designed for linear separation and can divide dataset into two groups. SVM now is utilized to solve a variety of real-world situations [8]. For classification, SVM employs the idea of hyperplane. To partition the data into classes, it creates a line or hyperplane. To achieve maximum gap between the two classes, SVM tries to push the decision boundary much further than possible. It boosts the productivity by using a sub-class of training points called as support vectors. Moreover, because it requires a lot of training time, it performs best with modest datasets. The basic premise of SVM is depicted in Fig. 10, where an optimum hyperplane is used to classify objects into groups based on support vector.

Fig. 10 Linear SVC



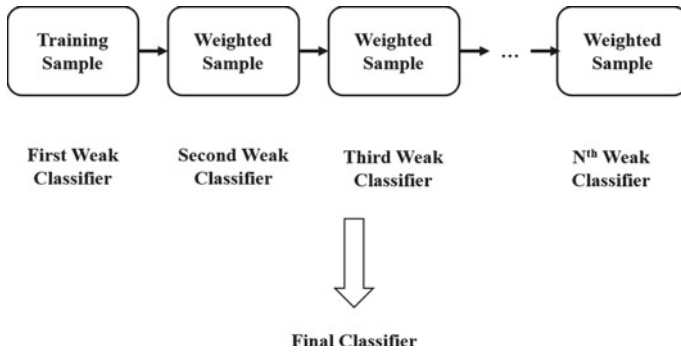


Fig. 11 XGB

2.4.6 XGB

Gradient-boosted decision trees are implemented in XGB. Decision trees are constructed sequentially in this approach. In XGBoost, weights are very significant. All of the independent variables are given weights, which are subsequently put into the decision tree, which predicts outcomes. The weight of variables that the tree predicted incorrectly is raised, and the variables are then put into the second decision tree. Individual classifiers/predictors are then combined to form a more powerful and precise model. Figure 11 shows the working of XGB [9, 10].

2.5 Results

Table 3 shows the comparison of different classifiers used for predicting the groundwater quality using the stated Machine Learning Approach. It can be seen that Random Forest and Decision Tree Classifiers perform the best and give 100 percent accuracy. Also, XGB has performed well in the process of classification, which emphasizes on the fact that when we use labeled data, i.e., supervised learning approach, then classification is better in using XGB, RF, and DT classifiers.

Table 3 Comparative results of different classifiers

Model	Train accuracy	Test accuracy
LR	0.70833	0.666667
DT	0.950000	0.920000
GNB	0.791667	0.833333
RF	0.850000	1.000000
Linear SVC	0.71667	0.583333
XGB	0.941667	0.916667

Fig. 12 Confusion matrix of RF

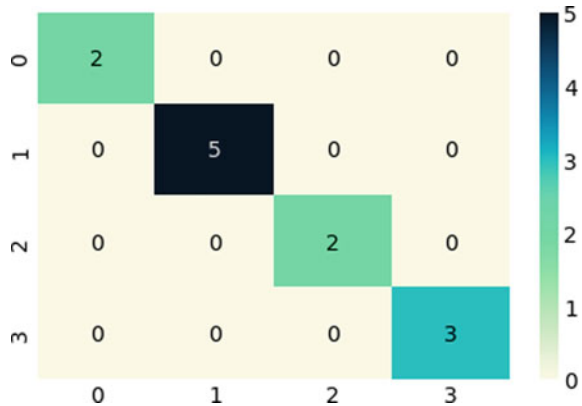


Figure 12 shows the confusion matrix of random forest classification.

2.6 Conclusion and Future Work

The groundwater quality prediction and classification were a multiclass problem. It can be observed that Extreme Gradient Boost, Random Forest, and Decision Tree Classifiers have been proved to perform best in this problem. The dataset for such problems is usually raw data, with some missing values and uneven scales, so data preprocessing needs to be done in order to get the best classification results. Further, Deep Learning techniques can be used in future for this domain, especially when the dataset is large. Time series forecasting.

References

1. "Introduction to Logistic Regression—by Ayush Pant—Towards Data Science." <https://towardsdatascience.com/introduction-to-logisticregression-66248243c148>. Accessed 11 Jan 2021
2. Thomas WE, Manz DO (2017) Chapter 4—exploratory study. In: Thomas WE, Manz DO (eds) Research methods for cyber security, Syngress, 2017. ISBN 9780128053492 pp 95–130. <https://doi.org/10.1016/B978-0-12-805349-2.00004-2>. <https://www.sciencedirect.com/science/article/pii/B9780128053492000042>
3. Timofeev R (2004) Classification and regression trees (CART) theory and applications. Humboldt University, Berlin, p 54
4. Islam MJ, Wu QJ, Ahmadi M, Sid-Ahmed MA (2007) Investigating the performance of naive-bayes classifiers and k-nearest neighbor classifiers. In: 2007 international conference on convergence information technology (ICCIT 2007), IEEE, pp 1541–1546
5. Jiang L, Wang D, Cai Z, Yan X (2007) Survey of improving naive bayes for classification. In: international conference on advanced data mining and applications. Springer, Berlin, Heidelberg, pp 134–145

6. "Understanding Random Forest. How the Algorithm Works and Why it Is... — by Tony Yiu — Towards Data Science." <https://towardsdatascience.com/understanding-random-forest58381e0602d2>. Accessed 11 Jan 2021
7. Chaudhary SK, Kamal R (2016) An improved random forest classifier for multi-class classification. *Inf Process Agric* 3(4):215–222. <https://doi.org/10.1016/j.inpa.2016.08.002>
8. Akhtar AK, Khan SA, Shaikat A (2013) Automated plant disease analysis (APDA): Performance comparison of machine learning techniques. In: *Proceeding—11th international conference frontiers of information technology. FIT 2013*, pp 60–65. <https://doi.org/10.1109/FIT.2013.19>
9. Aggarwal P (2019) ML: XGBoost (extreme gradient boosting). GeeksforGeeks. Retrieved 18 May 2022, from <https://www.geeksforgeeks.org/ml-xgboost-extreme-gradient-boosting/>
10. Friedman JH (2002) Stochastic gradient boosting. *Comput Stat Data Anal* 38(4):367–378

Performance Analysis of Breast Cancer Data Using Mann–Whitney U Test and Machine Learning



Priyanka Khanna and Mridu Sahu

Abstract Breast cancer is the most prevalent cause of mortality among women worldwide. Early detection and treatment can reduce the rate of mortality. Late prognosis and treatment of breast cancer (BC) patients lead to irreparable diseases and even death. As a result, in recent years, early BC diagnosis methods based on pathological breast images have been in high demand. In recent years, various models have been put up by the researcher for the early diagnosis of breast cancer. In this article, Wisconsin breast cancer (diagnostic) dataset (WDBC) is employed to categorize tumors into benign or malignant. Statistical-based Mann–Whitney U Test is applied for feature selection, followed by machine learning models for the classification of tumors. We compare two methods: a machine learning method with feature selection and one without. Finally, the results demonstrate that on selecting pertinent features, enhances the overall performance when tested on the WDBC dataset. The classification accuracy, sensitivity, and specificity obtained were 97.2%, 98.8%, and 94.5% using Random Forest with feature selection.

Keywords Breast cancer · Mann–Whitney U test · Machine learning · Feature selection · Accuracy

1 Introduction and Related Work

Cancer refers to the uncontrolled growth of certain cells in the human body [1]. These cells can spread into the surrounding tissue forming a lump known as tumor or malignancy [2]. After lung cancer, the second most common malignancies and reason of mortality for women worldwide are breast cancer [3]. Breast cancer (BC) is a frequently observed cancer in females of childbearing age. Breast cancer is the

P. Khanna (✉) · M. Sahu
National Institute of Technology, Raipur, India
e-mail: pkhanna.phd2019.it@nitrr.ac.in

M. Sahu
e-mail: mrisahu.it@nitrr.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
V. V. S. S. Chakravarthy et al. (eds.), *Advances in Signal Processing, Embedded Systems and IoT*, Lecture Notes in Electrical Engineering 992,
https://doi.org/10.1007/978-981-19-8865-3_26

277

prevalent diagnosed cancer and is increasing every year very rapidly [4, 5]. According to the changes in the environment, the nature of the breast cancer is also changing day by day [6]. As a result, raising awareness of the benefits of screening and early detection is desirable. Ultrasound (US), mammography, contrast-enhanced (CE), breast tomosynthesis (3D mammography), magnetic resonance imaging (MRI), computed tomography (CT), and positron emission tomography (PET) are the currently used clinical practices for the early diagnosis of BC. These methods are used to examine significant parameters such as the size, shape, location, type of cancer, stage of cancer, or how quickly it is growing. These methods are sometimes combined for a more accurate prognosis.

The most crucial and significant task is classification. Multiple classifiers and feature selection strategies are used in numerous research on datasets related to breast cancer [7]. Different machine learning (ML) classifiers have been developed for classification and employed on medical datasets. Machine learning is a subset of artificial intelligence within the realm of computing. ML is not only confined to computer science, but also extended to many other branches.

1.1 Based on Breast Cancer

Sengar [8] compared machine learning algorithms like Logistic Regression, Decision Tree on taken dataset. Decision Tree reported maximum classification accuracy of 95%. The main limitation of this work is that only two classifiers are evaluated. Anji Reddy Vaka [9] used deep neural networks on collected dataset. They collected data from Mahatma Gandhi Cancer Hospital and Research Institute, Visakhapatnam, India. As the dataset is limited, data augmentation is done to enlarge the dataset. Gaussian filtering is used for removal of noise as preprocessing step, and neglected values are removed using entropy followed by different ML algorithms for classification. Deep neural network reported highest accuracy of 97.01%.

Moh'd Rasoul Al-hadidi [10] used radiography images, and all the images are of equal size, thus making processing easier. Weiner filter is used to remove the image blurriness followed by Logistic Regression and back-propagation neural network. Back-propagation network attained maximum accuracy of 93%. Bazazeh [11] used WBCD dataset to train the model. Different machine learning algorithms like support vector machine, Random Forest, and Bayesian networks are used for evaluation.

Sadhukhan [12] converted images into fine-needle aspiration images which are further converted into grayscale images by removing hue from the images. For segmentation, thresholding is used and radii, smoothness, compactness, texture are calculated. Adel [13] cropped images to separate B-mode images amid elastography images. Different features like signal-to-noise ratio, width-to-height ratio, area, difference, perimeter difference, solidity, contrast-to-noise ratio, and compactness were extracted. Further, dimensionality reduction is done and input is fed to support

vector machine achieving an accuracy of 94.12%. Kaklamanis [14] applied correlation matrix for feature selection. Further, CART, KNN, Naive Bayes, and SVM are used for classification reporting accuracy of 93%, 96%, 89%, and 96%, respectively.

1.2 Based on Feature Selection

Perez [15] used two datasets of breast mammography images. Features were selected using Mann–Whitney U Test and selected feature subset is fed as an input to feedforward back-propagation network. MacFarland [16] emphasized on Mann–Whitney U Test, and it is generally conducted on non-parametric and independent values. It was first started by testing on goats, and two groups of goats in a total of 30 were taken, in which one group received mineral supplement included in the diet, whereas the other group is supplied with normal meal. At the end of the treatment, mineral supplement supplied goats shown to be healthier than the other group. Some facts like details about mineral supplement, how it is added to the meal, cost, and treatment regulation are not disclosed.

1.3 Based on Machine Learning

Bhavsar [17] evaluated different machine learning classifiers, namely support vector machine, Decision Tree, Supervised Learning, and Nearest Neighbor Neural Network. Performance metrics' accuracy, specificity, sensitivity were evaluated. Morgan [18] evaluated the performance using Gaussian process and Gaussian kernel ridge regression. For selecting pertinent features, Leave-Group-Out cross-validation root mean squared error is used. Fatima [19] provides comparative analysis of different machine learning algorithms for prognosis of different diseases. It emphasizes the use of machine learning algorithms for the analysis of disease and its decision-making.

2 Material and Method

2.1 Material

In this article, Wisconsin breast cancer (diagnostic) dataset (WDBC) [20] collected from UCI repository is used to differentiate benign from malignant sample. WDBC has 32 attributes and 569 instances, 357 of which are benign and 212 are malignant. Fine-needle aspirate (FNA) digitized picture was used to calculate features. These features exhibit ten characteristics of each cell nucleus. Excluding ID and diagnosis,

Table 1 Description of the Wisconsin breast cancer dataset (WDBC)

S. No.	Attribute	Description
1	ID	Id number
2	Diagnosis	Diagnosis (b = benign, m = malignant)
3	Radius	The average distance separation between the center and edge points
4	Texture	Standard deviation of values in gray scale
5	Perimeter	Tumor mean size
6	Area	Tumor mean area
7	Smoothness	Mean of local length variation
8	Compactness	Mean of $\text{perimeter}^2/\text{area} - 1$
9	Concavity	Severity of a contour's mean concave portions
10	Concave points	Mean of the concave points on the contour
11	Fractal	Mean of coastline approximation - 1

for each attribute, mean, standard error, and “worst” or largest (mean of the three largest values) are computed. There are no missing data in the dataset. Table 1 shows the description of the WDBC dataset features.

2.2 Method

In this article, we describe a feature selection method for WDBC dataset diagnostic that uses the Mann–Whitney U Test followed by different machine learning classifiers for classification. Firstly, WDBC dataset is taken and unwanted columns are removed as preprocessing step. Secondly, to improve the classification accuracy, feature selection using Mann–Whitney U Test is performed to choose relevant features. To categorize tumor as benign or malignant, selected features are finally fed via machine learning classifiers. Figure 1 demonstrates a proposed method for classifying breast tumor.

The assessment is conducted on the above datasets with and without feature selection method. And, the results are compared and analyzed. Evaluation metrics' sensitivity, specificity, and accuracy are calculated to access different machine learning classifiers. Experimental simulations were conducted using Jupyter Notebook.

2.2.1 Preprocessing

The first and most significant step is preprocessing, which enhances image quality while retaining key elements. Incorrect conclusions can be drawn from radiological images due to artifacts, noise, and other factors. The dataset consists of some unwanted columns which need to be removed for better result. No missing values are

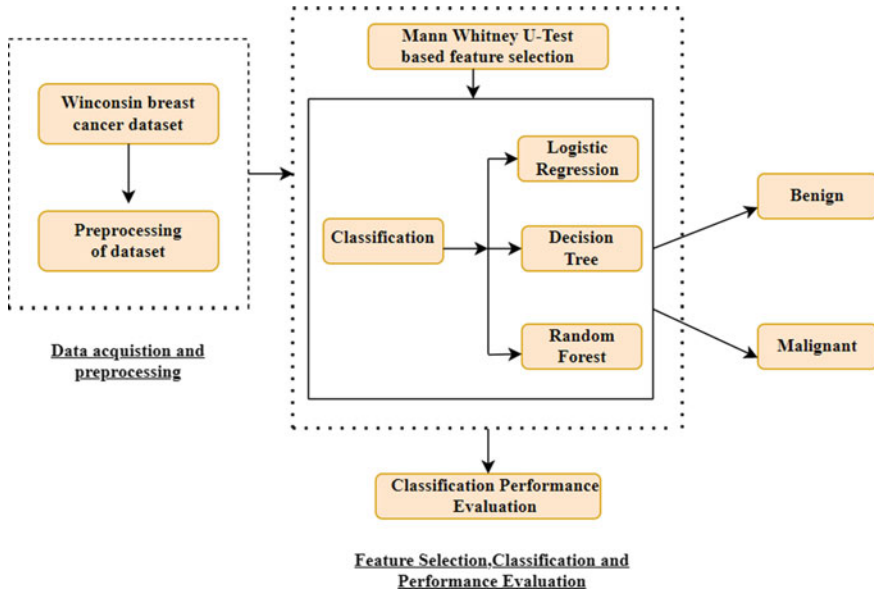


Fig. 1 Proposed methodology for classifying breast tumor

found in this dataset values. The categorical data diagnosis is changed to numerical data for compactness with the Mann–Whitney U Test [21].

2.2.2 Feature Selection

Following preprocessing, we carried out feature selection to select relevant features because they have a direct impact on classifier performance. The size of the feature space and computation time are reduced by removing redundant features. Gain ratio, recursive feature removal, Random Forest, Chi-square test, and searching algorithms are a few techniques frequently used for feature selection. In order to have effective prediction and computationally less costly models, the number of input classifier is limited. Mann–Whitney U Test is used as a feature selection technique in this paper. Mann–Whitney U Test is a statistical method used for non-uniformly distributed data.

In this test, calculation of U is done whose distribution under the null hypothesis is known. The normality of data was verified by U test, results obtained have a significant value less than 0.001 ($p < 0.001$), and 95% confidence interval (CI) marked those features was not normally distributed (non-parametric) [21]. A feature vector with 32 features ($F [1], F [2], \dots, F [32]$) is provided as input, and further, 26 features are chosen using the Mann–Whitney U Test.

2.2.3 Classification

Following feature selection, the classifier uses the pertinent features to categorize breast tumor as benign or malignant. Any automated system's classifier plots feature space as input to produce class labels [22]. The Naive Bayesian (NB), Decision Tree (DT), K-Nearest Neighbor (KNN), support vector machine (SVM), Random Forest, and Logistic Regression are examples of commonly used machine learning classifiers. Random Forest, Logistic Regression, and Decision Tree are evaluated in this study.

Logistic Regression transforms the linear regression model to allow us to probabilistically model the binary variables in consequence. A supervised procedure called Logistic Regression is used to predict the likelihood of a target variable. There are only two useful classes because the goal's or established variable's personality is binary. The established variable is binary in nature, with records encoded as 1 or 0. $P(Y = 1)$ is predicted by the Logistic Regression version as a function of X [8]. Decision Tree is a popular and unsupervised approach used for classification and prediction [23]. It is represented as a recursive partition of the instance, where leaves represent the class labels and branches refer to outcome in the form of features. It is a top-down approach which divides each result of the data into subsets. This predictive paradigm acts as a mapping between the item's qualities and values. Random Forest (RF) algorithm is based on multiple Decision Trees which is merged to produce an accurate and stable prediction [24]. RF is an ensemble of classifiers grown from a certain amount of randomness. RF stands for randomized ensembles of Decision Trees and is defined as a generic principle. Every observation is input into every Decision Tree. The final result is the most common outcome for each observation.

3 Experimental Results and Discussion

This section discusses the findings of the proposed method and compares them with the other related work. Experimental simulations were conducted using Jupyter Notebook. On the WDBC dataset, simulations were used to categorize the breast tumor as benign or malignant. The proposed method employed Mann–Whitney U Test for feature selection, using Statistical Package for Social Sciences (SPSS) software with 95% confidence interval, and the significance level was chosen to be less than 0.001. The values shown in Table 2 are the asymptotic significance values obtained on conducting Mann–Whitney U Test (non-parametric test). If asymptotic significance is greater than 0.001, then the features will be eliminated. The benign and malignant values, which are in categorical form, are converted into ordinal form. Out of 30 features, 26 features are selected based on U test and four features are eliminated. Selected features are further passed through classifier, and result is evaluated with and without feature selection.

Further features selected are fed as an input to Random Forest, Logistic Regression, and Decision Tree classifier. The dataset is split into sections: testing and

Table 2 Statistical analysis using Mann–Whitney U Test

Feature	Asymptotic significance
Perimeter worst	<0.001
Texture worst	<0.001
Radius worst	<0.001
Fractal dimension se	<0.001
Symmetry se	0.028
Concave points se	<0.001
Concavity se	<0.001
compactness se	<0.001
Smoothness se	0.214
Area se	<0.001
Perimeter se	<0.001
Texture se	0.644
Radius se	<0.001
Fractal dimension mean	0.537
Symmetry mean	<0.001
Concave points mean	<0.001
Concavity mean	<0.001
Compactness mean	<0.001
Smoothness mean	<0.001
Area mean	<0.001
Perimeter mean	<0.001
Texture mean	<0.001
Radius mean	<0.001
Area worst	<0.001
Smoothness worst	<0.001
Compactness worst	<0.001
Concavity worst	<0.001
Concave points worst	<0.001
Symmetry worst	<0.001
Fractal dimension worst	<0.001

training, under K-fold cross-validation protocol. Value of $k = 10$ is taken to compute the performance of the system. Table 3 shows evaluation measures used to evaluate the classifiers’ sensitivity, specificity, and classification accuracy. True positive (t_p) represents the quantity of patients who have been correctly classified, while the number of patients who have been correctly classified as negative class is represented by true negative (tn). False positive (f_p) represents the number of incorrectly predicted patients, whereas false negative (f_n) indicates the number of incorrectly predicted patients.

Table 3 Performance measure

Performance measures	Definition
Classification accuracy	$\frac{t_{tp} + t_{tn}}{t_{tp} + t_{fn} + t_{tn} + t_{fp}} \times 100$
Sensitivity	$\frac{t_{tp}}{t_{tp} + t_{fn}} \times 100$
Specificity	$\frac{t_{tn}}{t_{tn} + t_{fp}} \times 100$

Table 4 shows the experimental result in terms of accuracy, specificity, and sensitivity obtained by applying ML algorithms with feature selection using Mann–Whitney U Test. And, Table 5 shows the result obtained without feature selection. As shown in Tables 4 and 5, performance measure under ten-fold cross-validation is evaluated. It is observed that employing Random Forest as a classifier increases accuracy on selecting features. Accuracy of 99.5%, sensitivity of 98.8%, and specificity of 94.5% are obtained.

A comparative analysis of the proposed technique with prior relevant work on the WDBC dataset is shown in Table 6. Accuracy rate of 97.2% for the proposed method was obtained. Asri et al. [25] used C4.5, SVM, NB, and KNN for classification. Saravana Kumar et al. [26] proposed multi-layer perceptron based on deep learning. Performance comparison of proposed work with aforementioned related work is mentioned in Table 6.

Table 4 Performance metric obtained with feature selection

Classification technique	Accuracy (%)	Sensitivity (%)	Specificity (%)
Logistic regression	96.5	97.7	94.4
Decision tree	94.4	97.6	89.4
Random forest	97.2	98.8	94.5

Table 5 Performance metric obtained without feature selection

Classification technique	Accuracy (%)	Sensitivity (%)	Specificity (%)
Logistic regression	95.1	96.6	92.5
Decision tree	93.7	97.6	87.9
Random forest	96.5	97.7	94.4

Table 6 Comparison with related work

Dataset	Method	Accuracy (%)
Wisconsin dataset [25]	C4.5	95.13
	SVM	97.13
	NB	95.99
	KNN	95.27
Wisconsin dataset [26]	MLP	97
Wisconsin dataset	Proposed method	97.2

4 Conclusion and Future Scope

This study offered a thorough methodology for ultrasound-based breast cancer diagnosis. The study's primary contributions are as follows: Firstly, the WBCD dataset were taken and some unwanted columns were removed for better result. Secondly, to pick relevant features, an effective statistical approach Mann–Whitney U Test was used. Thirdly, features are trained using different machine learning classifiers to differentiate class labels. For future scope of this work, we plan to use a substantial dataset to test our proposed study and also to use data augmentation for increasing data size and optimization techniques for feature selection. In conclusion, the potential for the proposed technique to classify breast tumors is apparent, though better optimization techniques and big datasets are still needed.

References

1. James G, Witten D, Hastie T, Tibshirani R (2013) An introduction to statistical learning, vol 112. Springer, New York, p 18
2. Huang MW, Chen CW, Lin WC, Ke SW, Tsai CF (2017) SVM and SVM ensembles in breast cancer prediction. *PLoS ONE* 12(1):e0161501
3. Moon WK, Huang CS, Shen WC, Takada E, Chang RF, Joe J, Nakajima M, Kobayashi M (2009) Analysis of elastographic and B-mode features at sonoelastography for breast tumor classification. *Ultrasound Med Biol* 35(11):1794–1802
4. Moon WK, Huang YS, Lee YW, Chang SC, Lo CM, Yang MC, Bae MS, Lee SH, Chang JM, Huang CS, Lin YT (2017) Computer-aided tumor diagnosis using shear wave breast elastography. *Ultrasonics* 78:125–133
5. Okagbue HI, Adamu PI, Oguntunde PE, Obasi EC, Odetunmbi OA (2021) Machine learning prediction of breast cancer survival using age, sex, length of stay, mode of diagnosis and location of cancer. *Health Technol* 1–7
6. Amrane M, Oukid S, Gagaoua I, Ensari T (2018) Breast cancer classification using machine learning. In: 2018 electric electronics, computer science, biomedical engineering's meeting (EBBT), IEEE, pp 1–4
7. Mušić L, Gabeljić N (2019) Predicting the severity of a mammographic tumor using an artificial neural network. In: International conference on medical and biological engineering. Springer, Cham, pp 775–778

8. Sengar PP, Gaikwad MJ, Nagdive AS (2020) Comparative study of machine learning algorithms for breast cancer prediction. In: 2020 Third international conference on smart systems and inventive technology (ICSSIT). IEEE, pp 796–801
9. Vaka AR, Soni B, Reddy S (2020) Breast cancer detection by leveraging machine learning. *ICT Express* 6(4):320–324
10. Alarabeyyat A, Alhanahnah M (2016) Breast cancer detection using k-nearest neighbor machine learning algorithm. In: 2016 9th international conference on developments in e-systems engineering (DeSE). IEEE, pp 35–39
11. Bazazeh D, Shubair R (2016) Comparative study of machine learning algorithms for breast cancer detection and diagnosis. In: 2016 5th international conference on electronic devices, systems and applications (ICEDSA). IEEE, pp 1–4
12. Sadhukhan S, Upadhyay N, Chakraborty P (2020) Breast cancer diagnosis using image processing and machine learning. In: *Emerging technology in modelling and graphics*. Springer, Singapore, pp 113–127
13. Adel M, Kotb A, Farag O, Darweesh MS, Mostafa H (2019) Breast cancer diagnosis using image processing and machine learning for elastography images. In: 2019 8th international conference on modern circuits and systems technologies (MOCASST). IEEE, pp 1–4
14. Kaklamanis MM, Filippakis ME (2019) A comparative survey of machine learning classification algorithms for breast cancer detection. In: *Proceedings of the 23rd panhellenic conference on informatics*, pp. 97–103
15. Pérez NP, López MAG, Silva A, Ramos I (2015) Improving the Mann-Whitney statistical test for feature selection: an approach in breast cancer diagnosis on mammography. *Artif Intell Med* 63(1):19–31
16. MacFarland TW, Yates JM (2016) Mann–whitney u test. In *Introduction to nonparametric statistics for the biological sciences using R*. Springer, Cham, pp 103–132
17. Bhavsar H, Ganatra A (2012) A comparative study of training algorithms for supervised machine learning. *Int J Soft Comput Eng (IJSCE)* 2(4):2231–2307
18. Lu HJ, Zou N, Jacobs R, Afflerbach B, Lu XG, Morgan D (2019) Error assessment and optimal cross-validation approaches in machine learning applied to impurity diffusion. *Comput Mater Sci* 169:109075
19. Fatima M, Pasha M (2017) Survey of machine learning algorithms for disease diagnostic. *J Intell Learn Syst Appl* 9(01):1
20. Salama GI, Abdelhalim M, Zeid MAE (2012) Breast cancer diagnosis on three different datasets using multi-classifiers. *Breast Cancer (WDBC)* 32(569):2
21. Kirk R (2007) *Statistics: an introduction*. Nelson Education
22. Singh BK (2019) Determining relevant biomarkers for prediction of breast cancer using anthropometric and clinical features: a comparative investigation in machine learning paradigm. *Biocybernetics Biomed. Eng.* 39(2):393–409
23. De Mántaras RL (1991) A distance-based attribute selection measure for decision tree induction. *Mach Learn* 6:81–92
24. Breiman L (2001) Random forests. *Mach Learn J Paper* 45:5–32
25. Asri H, Mousannif H, Al H, Noel T (2016) Using machine learning algorithms for breast cancer risk prediction and diagnosis. *Proc Comput Sci* 83:1064–1069
26. NM SK, Tamilselvi S, Hariprasath K, Kaviyavarshini N, Kavinya A (2022) An efficient multi-layer perceptron neural network-based breast cancer prediction. In: *Principles and methods of explainable artificial intelligence in healthcare*. IGI Global, pp 211–231

Skin Lesion Analysis and Classification Techniques



Dilip Kumar Sharma, Anand Singh Jalal, and Bilal Sikander

Abstract Skin lesion diagnosis using computer-aided systems is a growing field of research. In recent years, researchers have been developing computer-aided diagnosis systems. An analysis of the efficiency of computer-aided diagnosis is examined, synthesized, and evaluated in this work. This study surveyed articles published in ScienceDirect, IEEE Xplore, and SpringerLink during the last five years. Traditional machine learning techniques and deep learning techniques are also covered in the publications. The research is compared in terms of their contributions, methodologies, and outcomes. Given the scarcity of data sets, it is difficult to evaluate skin lesion segmentation and classification methods without resorting to picture selection or introducing racial bias.

Keywords Deep learning · Skin lesion images · CNN-Convolutional neural networks · Skin cancer

1 Introduction

Melanomas are becoming more common, with an annual incidence of 53% rise. In part, this is due to an increase in exposure to ultraviolet (UV) light. Even though melanoma is one of the deadliest kinds of skin cancer, it may be treatable if caught and treated early.

Cancer may arise if cells in the body multiply uncontrollably. The medical word for when cancer spreads to other places of the body is metastasizing [1]. So, skin

D. K. Sharma · A. S. Jalal · B. Sikander (✉)

GLA University, Mathura, India

e-mail: sikander.gla_mt18@gla.ac.in

D. K. Sharma

e-mail: dilip.sharma@gla.ac.in

A. S. Jalal

e-mail: asjalal@gla.ac.in

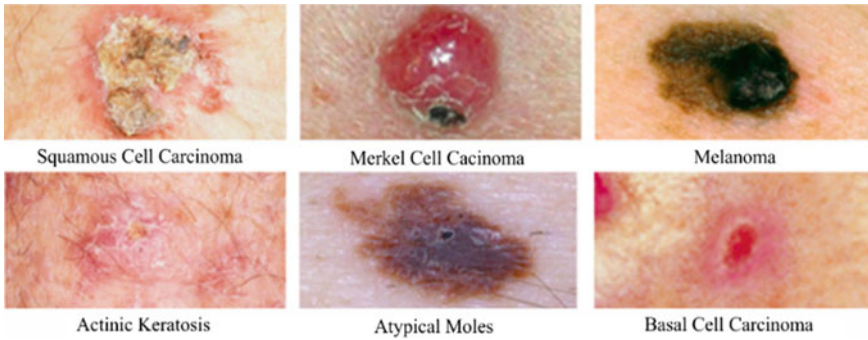


Fig. 1 The skin cancer foundation categorizes many types of skin cancer [3]

cancer is the uncontrolled expansion of abnormal skin cells. DNA damage from sun exposure that isn't repaired may lead to mutations or genetic defects in skin cells, which can then progress into tumors.

There are many different types of benign and malignant melanomas, making it difficult to determine the nature of skin disease. Squamous cell carcinoma (SCC), basal cell carcinoma (BSC), and melanoma are the most often seen forms of skin cancer in clinical settings [2]. The Skin Cancer Foundation (SCF) [3] identifies Merkel cell carcinoma, actinic keratosis (AKIEC), and atypical moles as three rare kinds of aberrant cells. There are six different kinds of skin lesions, which are shown in Fig. 1. Atypical moles are the second-most lethal kind of cell after melanoma cells. According to SCF [3], abnormal tissues may be distinguished in the following ways:

- *Melanoma*: The most deadly kind of skin cancer is this. However, it may come in a variety of colors, such as black or brown. Sunlight's ultraviolet (UV) rays have been linked to the development of cancerous tumors. It's treatable if caught early, but if it spreads to other regions of the body, it might be deadly.
- *Merkel cell carcinoma*: This is a very uncommon and dangerous kind of skin cancer. Despite this, it occurs in the community at a rate 40 times less often than melanoma.
- *Atypical moles*: Dysplastic nevi, often known as irregular moles, are harmless but irregular moles. You are more prone to get melanoma in any moles or other sun-exposed areas if you have one. Malignant melanoma is more common in these people.
- *Squamous Cell Carcinoma (SCC)*: Most incidences of skin cancer are of the basal cell kind. Skin lesions that appear as a result of an infection, such as warts and scaly, red areas, are common.
- *Actinic Keratosis (AKIEC) or solar keratosis*: Keratosis is a frequent skin condition. The skin develops a scaly, crusty growth. It's termed pre-cancerous because it may progress into skin cancer if left untreated.

- *Basal Cell Carcinoma (BCC)*: Basal cell carcinoma is the most typical kind of skin cancer. This kind of skin cancer seldom metastasizes. Some of the most common symptoms are open sores, shiny pimples, red patches, pink growths, or scars.

UV radiation damage to skin tissue is the most common cause of this kind of skin cancer [4]. Dermatologists often use ocular examinations to diagnose melanoma [5]. In certain ways, the accuracy of the clinical diagnosis is deceiving [6]. As a non-invasive approach, dermoscopy enables the exhibition of features that cannot be seen by the naked eye. Possible methods for dramatically enhancing the visibility of morphological characteristics include sun scanning [7], epiluminescence microscopy (ELM), cross-polarization epiluminescence microscopy (XLM), and side transillumination microscopy (STIM). As a result, the dermatologist has new information to work with in making a diagnosis. When compared to an unspecified eye, dermoscopy enhances diagnostic performance by 10–30% [8]. However, [9] found that novice dermatologists had a worse diagnosis accuracy using dermoscopy compared to professional dermatologists since this technique takes a lot of knowledge to detect lesions [10].

Malignant lesions of the skin may be diagnosed by a dermatologist doing a visual examination of the affected area. Diagnoses made using a Computer-Aided System (CAD) are almost identical to those made by a dermatologist with hands-on expertise [11], thus getting it right is crucial. Using just their clinical expertise, dermatologists can accurately detect melanoma in patients with a 65–80% accuracy rate [12]. When there is reason to be concerned, the last step in the visual examination is to obtain a dermatoscopic picture using an extremely high-resolution camera. Skin reflections are reduced and the deeper layers of the skin may be seen by using filtered illumination. As a result of this support, skin lesions were diagnosed 49% more correctly. Absolute accuracy of 75–84% for melanoma diagnosis was achieved in combination with visual examination and dermatoscopy pictures.

Classifying skin lesions has been a long-term goal for the machine learning community. The use of automated lesion classification in clinical tests has several advantages for both patients and doctors, including improved turnaround times and lower costs for potentially life-saving diagnoses. Pre-processing (enhancement), segmentation, feature extraction, and classification [13] made up the standard machine learning pipeline. The following section explains these phases:

- *Image enhancement*: The goal of this process is to clean up dermoscopy images by getting rid of things like hair and blood vessels.
- *Segmentation*: Segmenting the area of interest is crucial in computer-aided design (CAD). The vast number of distinct skin lesions complicates the segmentation procedure for photos of skin cancer. As a result, it rapidly became one of the most difficult and time-consuming CAD assignments.
- *Feature extraction*: Once an ROI has been established, the next step is feature extraction, in which a set of characteristics that can be used to divide a data source into two or more classes are discovered.

- *Classification and detection*: According to its ability to categorize the data set, the suggested system is rated. As a result, getting the best results requires careful consideration of the classifier. The number of classes and the number of extracted features are also factors to consider. Accuracy, specificity, sensitivity, precision, and Receiver Operating characteristics are all classification performance indicators (ROC).

Many CAD systems have been recognized via the application of various boundary detection, extraction, selection, and classification methods. Several publications [14] have recommended looking into and analyzing an image processing method for skin cancer detection, and both artificial intelligence and computer-aided design (CAD) systems' performance has been compared to that of professional dermatologists. Additional work is required to identify and reduce ambiguity within automated decision support systems to increase diagnostic accuracy. The automated skin lesion diagnosis model has never been the subject of an in-depth and current evaluation. A study of this kind would help advance the rapid advancement of dermoscopic research classification algorithms and procedures in recent years.

2 Methods

2.1 Systematic Review

We looked for English-language systematic reviews and primary research papers in ScienceDirect, IEEE Xplore, and SpringerLink. Only studies that were published in scholarly publications and documented properly were included in this study.

Studies were selected based on the following criteria: (i) dividing skin lesions into binaries or multiple classes for analysis, (ii) traditional ML techniques, contemporary DL models, and digital image formats, and (iii) articles in peer-reviewed journals and works originally published in English.

Studies that didn't adhere to the inclusion criteria were weeded out using the following standards: (i) Non-English language publications reviews. (ii) Scholarly pieces that analyze previous research. (iii) The third option is papers presented at conferences. (iv) Textual works. (v) Book chapters are an example of an item.

The preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) article selection procedure is shown in Fig. 2. Initial research revealed a total of 111,701 pieces of literature that met the search parameters. A total of 5757 records were found using alternative approaches. After removing duplicate entries, there were 106,398 records in the database. 801 full-text articles were identified following the application of the inclusion and exclusion criteria. After considering both conventional machine learning and deep learning techniques, these papers were selected as the best available. Only the highest-scoring models were considered for this report.

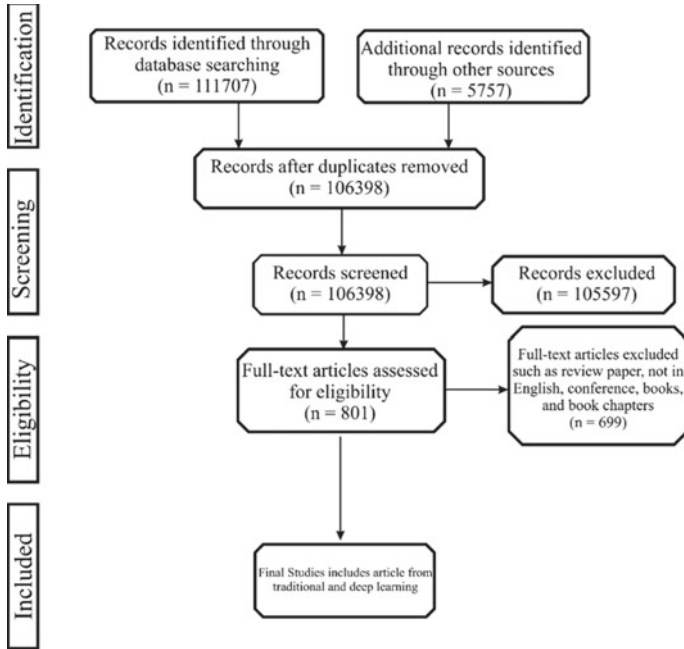


Fig. 2 Using the PRISMA flow approach to choose studies

2.2 Data Sets

It is possible to increase the visibility of skin lesions by using dermoscopy, which is an invasive skin imaging procedure. It is possible to enhance the appearance of deeper skin by reducing surface reflections [15]. The environment, which includes hair, veins, and color calibration charts and rule markings, makes it difficult to identify melanomas due to their texture, size, and color variations, as well as their strong similarity to lesions that are not melanomas.

In this part, we’ll go through some of the most popular data sets utilized in this area of research. We used free and paid data sets like MedNode, DermaIS, DermQuest, ISIC 2018, 2019, Ph2, and even Dermofit.

In MED-NODE, 170 pictures of melanomas and nevus are grouped into 70 and 100 segments, respectively. The data for this research came from the Digital Image Archive of the Dermatology Department at the University of Medicine and Dentistry of the Netherlands. The technology for detecting skin cancer from microscopic images has been designed and tested [16].

The PH2 database was the outcome of a joint effort between the universities of Porto and Lisbon as well as the Pedro Hispano Hospital in Matosinhos. It’s 768 by 560 pixels in size and has 200 pictures in the RGB color space. There are three types of picture sets included in this data set: normal (40 images), melanoma (80 images), and an aberrant nevus (40 images).

One hundred and fifteen photographs were utilized for the training data set and fifteen dozen were used for the test set in the 2018 International Skin Imaging Collaboration (ISIC) data set, also known as the HAM10000 (“Humans versus machines with 10,000 training photos”) data set.

We also have a data set from ISIC 2019, called BCN 20,000 [18], which includes 8 classes of known photos. Included are MEL, NV, AKIEC, BKL, DF, VASC, and SCC. There are 867 pictures in AKIEC, 2624 in BKL, 239 in DF, 12,875 in NV, 4522 in MEL, 628 in SCC, 253 in VASC, and so on.

There are around 1300 skin photos in the Dermofit Image Library collection, each with a matching class name and lesion classification. The ten groups are AKIES, hemangioma, intraepithelial cancer, nevus, SCC, pyogenic granuloma, seborrhoeic keratosis, DF, and MEL. The images in these groups may be of a nevus, MEL, seborrhoeic keratosis, DF, AKIES, hemangioma, SCC, intraepithelial cancer, or Pyogenic Granuloma.

Another skin image collection, the Interactive Atlas of Dermoscopy (EDRA) [19], has 20 labels. Labels such as Clark’s Nevus, DF, melanosis, and VASC were used to describe the many subtypes of the disease. Melanoma was also known by the names of the various subtypes of the disease.

2.3 *Methods and Techniques*

According to Qasim et al. [21], skin lesions should be classified according to their severity. The Gaussian filter’s improved images were utilized to extract ROI using KNN. An SVM was used to classify the ROI after it had been segmented. A blue-white structure was used by Madooei et al. [22] to distinguish melanoma from nevi lesions. According to Saez et al. [23], melanoma and nevi may be differentiated based on the color of the lesions. The colors of the lesions and the surrounding areas were used to determine their classification. A technique for segmenting skin lesions was put out by the researchers Navarro et al. [24]. The lesions were divided into melanoma and nevi by the researchers. Skin lesions might be modeled using a CAD system, as suggested by Riaz et al. [25]. Lesion segmentation was used to obtain ROI from their system. This was done by using Kullback–Leibler divergence.

In addition to traditional machine learning techniques, deep convolutional neural network models were also tested. Table 1 was used only at the end of the systematic review to compare existing approaches.

3 Conclusion

When comparing the various skin lesion categorization systems, it becomes clear that the issue formulations of the various works differ somewhat. Data acquisition (collection), tuning, feature selection, deep learning, and final model construction are

Table 1 A comparison between state-of-the-art methods with novel methods

References	Dataset Involved		Colored Images	Enhancement	Segmentation	Contribution	Contribution Achieved	Methods and Tools	No. of Classes	Accuracy (%)	Sensitivity (%)
	Involved	Free									
[4]	2	Y	Y	N	N	Extraction of features using frequency domain analysis and classify these features	Y	Cross spectrum-based feature extraction, spatial feature extraction, SVM-RFE with CBR	4	98.72	98.89
[5]	1	N	Y	N	N	Strengthen the skin lesion classification bag of dense features	Y	SVM, color features, orientation histogram, and gradient location	2	78	N/A

(continued)

Table 1 (continued)

References	Dataset		Colored Images	Enhancement	Segmentation	Contribution	Contribution Achieved	Methods and Tools	No. of Classes	Accuracy (%)	Sensitivity (%)
	Involved	Free									
[6]	2	Y	N	Y	Y	Medical aid CAD system	N	Models of deformation based on chroma data, velocity functions, SVM, Asymmetry and Compactness Index, Discrete Wavelet Transform, and the Chan-Vese and Wilcoxon Rank Sum Statistics	2	88	95

(continued)

Table 1 (continued)

References	Dataset		Colored Images	Enhancement	Segmentation	Contribution	Contribution Achieved	Methods and Tools	No. of Classes	Accuracy (%)	Sensitivity (%)
	Involved	Free									
[7]	2	Y	N	Y	Y	Fuzzy pixel thresholding and classification are used to segment skin lesions	Y	Histogram thresholding, fuzzy sets, and classification	2	88.4	86.9
[8]	1	Y	Y	Y	Y	Use of feature similarity metrics for lesion categorization in bag-of-features model codebook learning	Y	Scale-invariant feature transform, color histogram, k-means, and codebook learning (SIFT)	2	82	80
[9]	1	Y	Y	Y	Y	Kernel sparse representation for skin lesion segmentation and classification	Y	Code sparsity, kernel dictionary, and kernel singular value decomposition	3	91.34	93.17

(continued)

Table 1 (continued)

References	Dataset Involved		Colored Images	Enhancement	Segmentation	Contribution	Contribution Achieved	Methods and Tools	No. of Classes	Accuracy (%)	Sensitivity (%)
	Involved	Free									
[10]	Y	N	Y	Y	Classifying skin lesions better utilizing equivocal features	Y	SVM, and Feedforward Neural Network (FNN); Gradient-based Histogram Thresholding; Local Binary Patterns Clustering; Euclidean Distance; Discrete Fourier Transform Spectrum; Power Spectral Density;	2	91	68	
[21]	Y	Y	Y	Y	Segmenting skin lesions using an encoder-decoder FCN	Y	GPU, FCN	3	96.92	96.88	

(continued)

Table 1 (continued)

References	Dataset		Colored Images	Enhancement	Segmentation	Contribution	Contribution Achieved	Methods and Tools	No. of Classes	Accuracy (%)	Sensitivity (%)
	Involved	Free									
[22]	3	Y	Y	N	Y	Using extracted ROI and CNNs, classify skin melanomas	Y	Multiclass SVM, SoftMax, Histogram-based windowing method, hierarchical clustering, fine-tune, transfer learning, Augmentation, GPU, AlexNet, ResNet101, GoogleNet,	3	98.14	97.27
[23]	3	Y	N	Y	Y	Fusion of deep characteristics derived from a skin lesion to aid in categorization	Y	PCA, fusion, fine-tune, transfer learning, Augmentation, GPU, Otsu method, Alex, and VGG-16	2	99.9	99.5

(continued)

Table 1 (continued)

References	Dataset Involved		Colored Images	Enhancement	Segmentation	Contribution	Contribution Achieved	Methods and Tools	No. of Classes	Accuracy (%)	Sensitivity (%)
	Involved	Free									
[24]	3	Y	Y	Y	Y	Using a collection of convolutional neural networks (CNNs) across many scales, an ensemble	Y	The terms EfficientNetB0, EfficientNetB1, SeResNeXt-50, fusion, fine-tune, transfer learning, Augmentation, GPU, and so on and so on	7	96.3	N/A
[25]	2	Y	Y	Y	Y	Multi-level multi-class classification of skin lesions with classical machine learning and transfer learning	Y	The use of K-means, Otsu's thresholding, GLCM, ANN, kfold validation, AlexNet, fine-tune, transfer learning, Augmentation, and a Graphics Processing Unit (GPU)	4	93.02	87.87

the five pillars of effective melanoma diagnosis. Clinical data, such as race or age, may be included in classifiers to enhance classification accuracy. Dermatologists might benefit from this additional information. Future research should take these things into account.

References

1. American Cancer Society. Statistics 2020. Available online: <https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2020.html?fbclid=IwAR2gMmnaky1m3LdETjBwoTiRkaxDiaKvWss9UISVx6YqWmR-rrehUjBMpvS>. Accessed on 10 May 2022
2. Kassem MA, Hosny KM, Damaševičius R, Eltoukhy MM (2021) Machine learning and deep learning methods for skin lesion classification and diagnosis: a systematic review. *Diagnostics* 11(8):1390
3. Skin Cancer Foundation. Skin Cancer Information. 2020. Available online: <http://www.skincancer.org/skin-cancer-information>. Accessed on 20 May 2020
4. Cihan M, Ceylan M (2021) Fusion of CT and MR liver images using multiresolution analysis methods. *Avrupa Bilim ve Teknoloji Dergisi* 30:56–61
5. Di Matteo E, Pampena R, Pizzichetta MA, Cinotti E, Chester J, Kaleci S, Manfredini M, Guida S, Dika E, Moscarella E, Lallas A, Farnetani F (2022) Unusual dermoscopic patterns of basal cell carcinoma mimicking melanoma. *Exp Dermatol*
6. Naik PP (2021) Cutaneous malignant melanoma: a review of early diagnosis and management. *World J Oncol* 12(1):7
7. Ring C, Cox N, Lee JB (2021) Dermatoscopy. *Clin Dermatol* 39(4):635–642
8. Zaballos P, Salafranca MÁ, Medina C, Bañuls J, Puig S, Del Pozo LJ, Malveyh J, Karaarslan IK, Thomas L, Landi C, Argenziano G, Flores G (2022) The usefulness of dermoscopy for the recognition of malignant collision tumors. *Dermatology* 238(1):132–139
9. Deda LC, Goldberg RH, Jamerson TA, Lee I, Tejasvi T (2022) Dermoscopy practice guidelines for use in telemedicine. *NPJ Digital Med* 5(1):1–7
10. Young AT, Vora NB, Cortez J, Tam A, Yeniay Y, Afifi L, Yan D, Nosrati A, Wong A, Johal A, Wei ML (2021) The role of technology in melanoma screening and diagnosis. *Pigment Cell Melanoma Res* 34(2):288–300
11. Tognetti L, Fiorani D, Russo F, Lazzeri L, Trovato E, Flori ML, Moscarella E, Cinotti E, Rubegni P (2021) Tele dermatology in 2020: past, present and future perspectives. *Italian J Dermatology Venereology*, 156(2):198–212
12. Meng X, Chen J, Zhang Z, Li K, Li J, Yu Z, Zhang Y (2021) Non-invasive optical methods for melanoma diagnosis. *Photodiagn Photodyn Ther* 34:102266
13. Nie Y, Ferro M, Sommella P, Carratù M, Cacciapuoti S, Di Leo G, Lundgren J, Fabbrocini G (2021) Ensembling CNNs for dermoscopic analysis of suspicious skin lesions. In: 2021 IEEE international symposium on medical measurements and applications (MeMeA). IEEE, pp 1–6
14. Sar- L, Rotemberg VM, Matsoukas K, Halpern AC, Marchetti MA, Hay JL (2021) Interactive skin self-examination digital platforms for the prevention of skin cancer: a narrative literature review. *J Am Acad Dermatol* 84(5):1459–1468
15. Ferlay J, Colombet M, Soerjomataram I, Parkin DM, Piñeros M, Znaor A, Bray F (2021) Cancer statistics for the year 2020: an overview. *Int J Cancer* 149(4):778–789
16. Giotis I, Molders N, Land S, Biehl M, Jonkman MF, Petkov N (2019) MED-NODE: a computer-assisted melanoma diagnosis system using non-dermoscopic images. *Expert Syst Appl* 42(19):6578–6585
17. DermQuest. Available online: <http://www.dermquest.com>. Accessed on 25 Jan 2019

18. Combalia M, Codella NC, Rotemberg V, Helba B, Vilaplana V, Reiter O, Carrera C, Barreiro A, Halpern AC, Malvey J (2019) Bcn20000: Dermoscopic lesions in the wild. arXiv preprint [arXiv:1908.02288](https://arxiv.org/abs/1908.02288)
19. Lio PA, Nghiem P (2019) Interactive atlas of dermoscopy: giuseppe argenziano, MD, H. Peter Soyer, MD, Vincenzo De Giorgio, MD, Domenico Piccolo, MD, Paolo Carli, MD, Mario Delfino, MD, Angela Ferrari, MD, Rainer Hofmann-Wellenhof, MD, Daniela Massi, MD, Giampiero Mazzocchetti, MD, Massimiliano Scalvenzi, MD, and Ingrid H. Wolf, MD, Milan, Italy, 2000, Edra Medical Publishing and New Media. 208 pages. \$290.00. ISBN 88-86457-30-8. CD-ROM requirements (minimum): Pentium 133 MHz, 32-Mb RAM, 24X CD-ROM drive, 800× 600 resolution *J Am Acad Dermatol* 50(5):807–808
20. Rotemberg V, Kurtansky N, Betz-Stablein B, Caffery L, Chousakos E, Codella N, Combalia M, Dusza S, Guitera P, Gutman D, Halpern A (2021) A patient-centric dataset of images and metadata for identifying melanomas using clinical context. *Sci Data* 8(1):1–8
21. Khan MQ, Hussain A, Rehman SU, Khan U, Maqsood M, Mehmood K, Khan MA (2019) Classification of melanoma and nevus in digital images for diagnosis of skin cancer. *IEEE Access* 7:90132–90144
22. Madooei A, Drew MS, Hajimirsadeghi H (2019) Learning to detect blue–white structures in dermoscopy images with weak supervision. *IEEE J Biomed Health Inform* 23(2):779–786
23. Sáez A, Acha B, Serrano A, Serrano C (2019) Statistical detection of colors in dermoscopic images with a texton-based estimation of probabilities. *IEEE J Biomed Health Inform* 23(2):560–569
24. Navarro F, Escudero- M, Bescós J (2019) Accurate segmentation and registration of skin lesion images to evaluate lesion change. *IEEE J Biomed Health Inform* 23(2):501–508
25. Riaz F, Naeem S, Nawaz R, Coimbra M (2019) Active contours based segmentation and lesion periphery analysis for characterization of skin lesions in dermoscopy images. *IEEE J Biomed Health Inform* 23(2):489–500

Dual Band Microstrip Patch Antenna with Annulated Circular Ring



Chirag Arora

Abstract In this article, authors have designed a slotted microstrip patch antenna comprising of a circular ring embedded in the center. Initially, the designed conventional patch antenna operates at 2.45 GHz. However, when two slots are etched on both ends of the patch, an additional resonant frequency band is obtained at 5.8 GHz, which reduces to 5 GHz when an annular ring is inserted between these two slots. The gain and bandwidth of the designed antenna are 4 dBi and 400 MHz at 5 GHz and 8 dBi and 270 MHz at 2.45 GHz. Thus, the proposed antenna can accommodate both IEEE 802.11a Wi-Fi bands operating at 2.45 and 5 GHz. This dual band antenna is designed on 1.48 mm thick FR-4 substrate with dielectric constant of 4.3 and loss tangent of 0.01. The designed structure is excited by a SMA coaxial connector of 50- Ω . High Frequency Structure Simulator software (HFSS) is used for designing and simulating this antenna. This commercially available software is based on finite element method.

Keywords Microstrip patch antenna · Wi-Fi · Dual band · Finite element method

1 Introduction

Due to simple structure and robust nature, there is a huge demand of patch antennas for wireless communication systems. These antennas possess light weight and low profile; thus, they are prominently used in designing of RFID tags and Bluetooth/Wi-Fi enabled devices [1–3]. Thus, it can be concluded that these antennas have brought a revolution in field of wireless communication systems and paved a bright path for antenna engineers to explore their research in this field. Moreover, there are several systems available in the market which are multifunctional and hence, operates on multiple frequencies, such as 575.42/1227.60/1176.45 MHz, which is used for Global Positioning Systems, 900/1800 MHz which is required for Global System

C. Arora (✉)

KIET Group of Institutions, Delhi-NCR, Ghaziabad, UP, India

e-mail: c_arora2002@yahoo.co.in

for Mobile Communication, wireless local area network systems which operates on 2.4/5.2/5.8 GHz, 2.5/3.5/5.5 GHz for Interoperability for Microwave Access, long term evolution system operating at 700/2300/2600 MHz, etc. If separate antennas are used for different operating frequencies, then it may result in coupling problems, along with enhanced cost and size of that device, causing a problem for their existence in the market in this competitive world. Thus, a technique is to be devised by which a single antenna can operate on multiple frequencies. Use of multiple band antennas is one such technique, as it provides a promising solution to achieve multiple resonant frequencies from a single microstrip patch antenna [4, 5].

Several techniques have been adopted to design multiband patch antennas, such as using slotted patch [6, 7], utilizing reconfigurable antennas [8, 9], adopting different feeding techniques [10], etching slots on ground plane [11], loading of metamaterials with traditional antennas [12, 13], and designing multilayered structures [14].

In [6], Faizal designed a slotted multiband square patch antenna with several slots and possessing defective ground. The behavior of designed antenna is highly admirable, but it required double feed, making the antenna fabrication tedious. Abutarbush et al. [9], designed a reconfigurable multiband antenna by using two PIN diodes. In [13], Arora et al. used multiple split ring resonators to obtain multiband operation. However, the integration of pin diodes with patch antenna makes the structure bulky. Loading of conventional antennas with metamaterials results in drastic improvement in antenna performance, but incorporation of these artificially engineered materials with traditional antennas requires precise alignment, which is difficult to acquire. Multilayered structures are also widely used to achieve multiband operation as in [14], where Modi et al. used a multilayer structure for designing the antenna which operates on L-band and S-band frequencies. But antennas which are designed using this technique are quite bulky. However, as per the literature survey done so far, use of slots and slits on the radiating patch appears to be simplest technique for obtaining multiband operation, as it is easy to fabricate, does not require any extra hardware, no integration is required and moreover high gain and considerable bandwidth is obtained. In [7], Venkatachari et al. designed a slot coupled multiband antenna for C-band and X-band purposes with appreciable gain in both the frequency bands.

In this article, authors have designed a microstrip patch antenna with two rectangular slots at two ends of the patch with an annular ring between the two slots. When two slots are milled on the patch, the antenna operates at 2.45 and 5.8 GHz. However, when an annular ring is etched between these two slots, the second resonant frequency shifts to 5 GHz, thus, lowering the profile of the antenna. Also, appreciable gain and bandwidth are obtained at both the resonant frequencies. This technique appears to be novel as it does not require any extra hardware and even the planarity of the structure is maintained.

2 Antenna Design

Figure 1a describes the geometric sketch of a traditional microstrip patch antenna while Fig. 1b presents the same antenna when two slots are milled on each end of the patch, and an annular ring is etched between these two slots. Table 1 shows the measurements of various sides of the proposed slotted antenna. 1.48 mm thick FR-4 substrate possessing dielectric constant of 4.3 and loss tangent of 0.01 is used. SMA coaxial connector of 50Ω is utilized to excite this antenna. The two slots etched on each end are responsible for obtaining dual band operation. The annular ring inserted between the two slots shifts one of the resonant frequencies to the lower value, thus reducing the size of the antenna. Analysis of parameter variation has been used to obtain the positions of two slots so as to achieve wide impedance bandwidth and high gain in desired frequency bands.

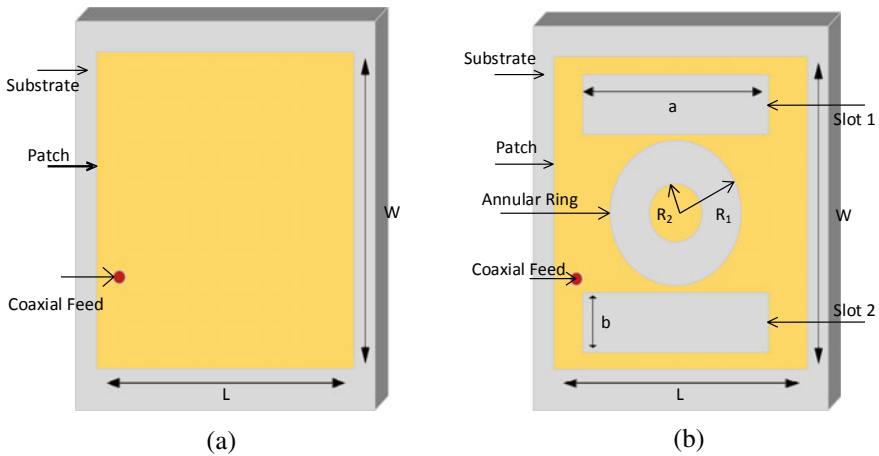


Fig. 1 Structure of **a** traditional patch antenna and **b** slotted patch antenna with annulated ring

Table 1 Proposed slotted patch antenna dimensions

Parameters	Dimensions (mm)
a	16
b	5
W	37
L	28
R ₁	6
R ₂	3

3 Results and Discussion

In this segment, simulated results for the traditional microstrip patch antenna and proposed patch antenna possessing two slots at each end and then after inserting the annular ring between the two slots are discussed. As observed in Fig. 2a, the conventional patch antenna resonates at 2.45 GHz providing the impedance bandwidth of 270 MHz. However, when two slots are etched on each end of this patch antenna, one more resonant frequency is produced at 5.8 GHz, as observed in Fig. 2b. The impedance bandwidth at this resonant frequency is 200 MHz. Now when an annular ring is etched between these two slots, one of the resonant frequencies gets shifted to a lower value of 5 GHz and provides the impedance bandwidth of 400 MHz. This is presented in return loss characteristics shown in Fig. 2c. This reduction in the resonant frequency occurs because insertion of an annular ring between the two slots increases the current path of excited surface [15].

Figure 3a, b depicts the radiation curves of this proposed antenna in elevation plane which is working at frequency of 2.45 GHz and 5 GHz, respectively. As observed from these radiation pattern characteristics, the gain of this antenna is 8 dBi at 2.45 GHz and 4 dBi at 5 GHz. Thus, it is seen that this patch antenna produces an appreciable impedance bandwidth and gain at both the resonant frequencies.

4 Conclusions

In this work, a two band microstrip patch antenna, which operates at 2.45 and 5 GHz has been discussed. Dual band behavior has been obtained with help of two slots etched at two ends of a conventional microstrip patch and then desired frequency band has been tuned by inserting an annular ring between these two slots. Thus, the designed patch antenna can be utilized for both IEEE 802.11a 2.45 and 5 GHz Wi-Fi bands. The uniqueness of this structure is the achievement of lower frequency band for same antenna dimensions. The proposed antenna is yet to be fabricated and once done; the measured results will be published in upcoming articles.

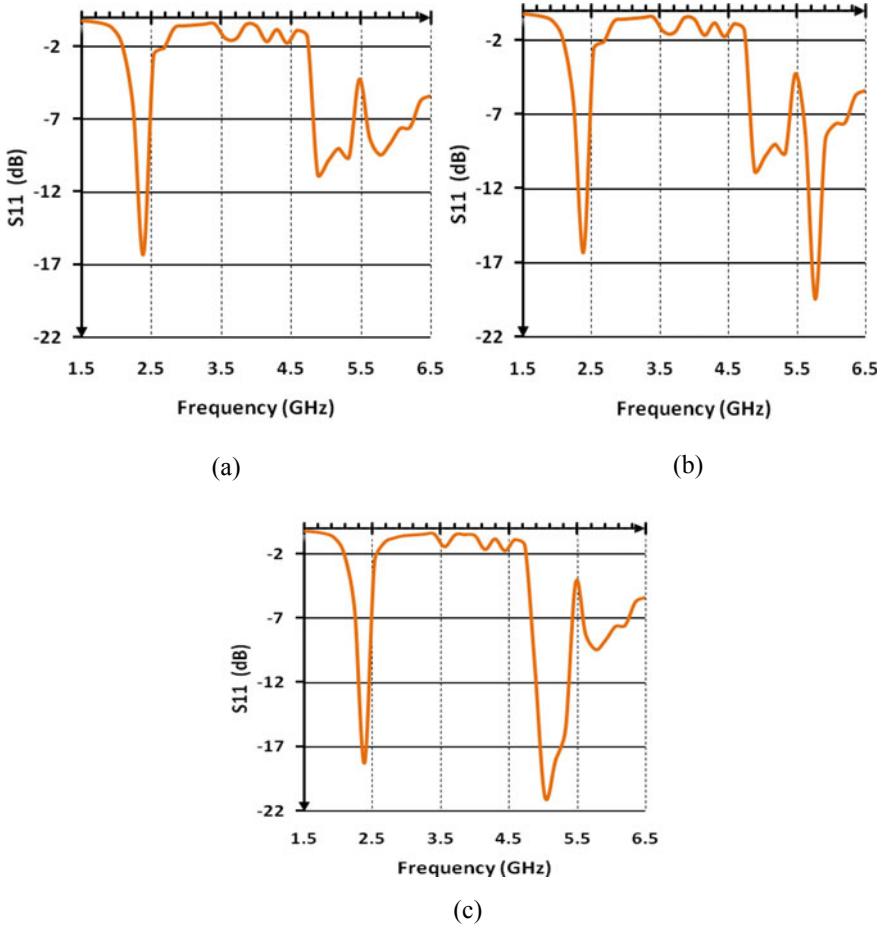


Fig. 2 S_{11} trace of **a** traditional patch antenna, **b** when two slots are etched at each end of the patch, and **c** when an annular ring is etched between the two slots

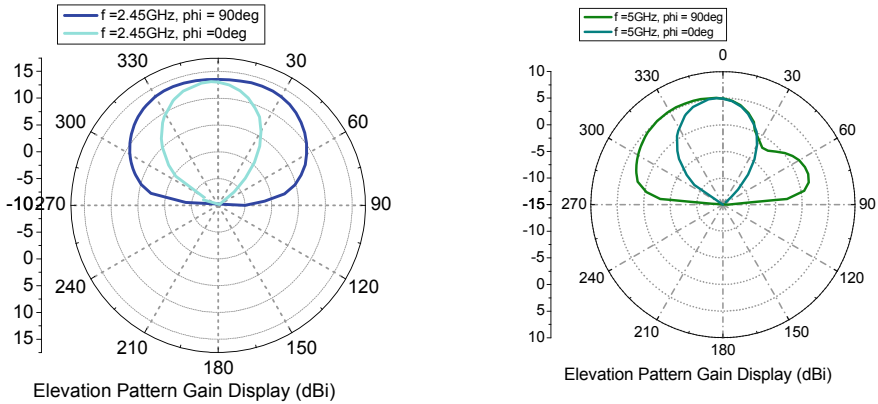


Fig. 3 Elevation plane radiation pattern characteristics of the proposed dual band microstrip patch antenna at **a** 2.45 GHz and **b** 5 GHz

References

1. Burtow M, Paszkowski A, Kurgan P, Nojman M (2020) Microstrip patch antenna with wide beamwidth for AESA TRM diagnostics. In: 23rd international microwave and radar conference (MIKON), pp 80–83
2. Anguera J, Andujar A, Jayasinghe J (2020) High-directivity microstrip patch antennas based on TM odd-0 modes. *IEEE Antennas Wirel Propag Lett* 19(1):39–43
3. Arora C (2021) Metamaterial-Inspired Circularly Polarized Microstrip Patch Antenna. *Computer Communication, Networking and IoT*. Springer, pp 183–190 (2021)
4. Teja PVNR, Raveendra M, Ganesh TV, Mahesh JV, Saravana Kumar U (2021) A multiband triangular slot array based microstrip patch antenna for C & X-band applications. In: Second international conference on electronics and sustainable communication systems (ICESC), pp 582–585 (2021)
5. Abdalrazik A, Gomaa A, Kishk AA (2021) A Hexa band quad-circular-polarization slotted patch antenna for 5G, GPS, WLAN, LTE, and radio navigation applications. *IEEE Antennas Wirel Propag Lett* 20(8):1438–1442
6. Faisal SH, Abbasi H, Shahid S, Saleem S (2021) Multi-band antenna with L-U-E slots for WiMAX and WLAN applications. In: International conference on electrical, communication, and computer engineering (ICECCE), pp 1–4 (2021)
7. Venkatachari D, Kameswararao PS, Shivani S, Ganesh PM, Sumanth S (2021) Design of slot coupled multiband antenna for C-band and X-band applications. In: IEEE 2nd international conference on applied electromagnetics, signal processing, & communication (AESPC), pp 1–5 (2021)
8. Sun M, Zhang Z, Zhang F, Chen A (2019) L/S multiband frequency-reconfigurable antenna for satellite applications. *IEEE Antennas Wirel Propag Lett* 18(12):2617–2621
9. Abutarboush HF, Shamim A (2018) A reconfigurable inkjet-printed antenna on paper substrate for wireless applications. *IEEE Antennas Wirel Propag Lett* 17(9):1648–1651
10. Kumar A, Sankhla V, Deegwal JK, Kumar A (2018) An offset CPW-fed triple-band circularly polarized printed antenna for multiband wireless applications. *AEU Int J Electron Commun* 86:133–141
11. Ali I, Chang RY (2015) Design of dual-band microstrip patch antenna with defected ground plane for modern wireless applications. In: IEEE 82nd vehicular technology conference, pp 1–5 (2015)

12. Arora C, Pattnaik SS, Baral RN (2018) Dual band microstrip patch antenna array loaded with split ring resonators and via holes. *AEU Int J Electron Commun* 93:253–260
13. Arora C, Pattnaik SS, Baral RN (2019) Multiple split ring resonator inspired dual band microstrip patch antenna array. *Microwave Rev* 25:7–14
14. Modi A, Sharma V, Rawat A (2021) Design and analysis of multilayer patch antenna for IRNSS, GPS, Wi-Fi, satellite, and mobile networks communications. In: 12th international conference on computing communication and networking technologies (ICCCNT), pp 1–6 (2021)
15. Bao XL, Ammann MJ (2006) Compact annular-ring embedded circular patch antenna with cross-slot ground plane for circular polarization. *Electron Lett* 42:192–193

Genetically Optimized Quad-Band mm-Wave Microstrip Patch Antenna



Arebu Dejen, Jeevani Jayasinghe, Murad Ridwan, and Jaume Anguera

Abstract The demand of multi-functional antennas with achievable performance from a single device is increasing dramatically. At the same, quad-band rectangular patch antennas are essential for numerous mm-wave applications. This study aims to use a binary-coded GA to enhance the multi-functionality of a microstrip antenna for mm-wave wireless communication. The patch surface was optimized by gridding it into 6×6 tiny rectangular cells and assigning conducting and non-conducting features to them. The proposed method has iteratively modeled the antenna using a Ansys HFSS and MATLAB. The reference antenna model and optimized antenna were presented and compared. The optimized antenna has resonated at 28.3, 38.1, 46.6, and 60.0 GHz center frequency. The antenna realized a peak broadside directivity of 7.8 dB, 8.8 dB, 7.3 dB, and 7.1 dB with total operating bandwidth of 11.5 GHz, respectively. The simulation results of the optimized antenna were compared with references and the proposed antenna outperform in all four distinct frequencies.

Keywords Quad-band · Microstrip antenna · Mm-wave patch antenna · Genetic algorithm

A. Dejen (✉) · M. Ridwan
School of Electrical and Computer Engineering, Addis Ababa Institute of Technology, Addis Ababa, Ethiopia
e-mail: arbdjn@gmail.com

J. Jayasinghe
Department of Electronics, Wayamba University of Sri Lanka, Kuliyaipitiya, Sri Lanka

J. Anguera
Telecommunication Engineering, Universitat Ramon LLull, Barcelona, Spain
Ignion, Barcelona, Spain

1 Introduction

Wireless communication is the fastest expanding and most dynamic area of telecommunications industry. Mobile communication, in particular, has undergone significant changes before reaching the present 5G standard. Future wireless communication will be adversely affected by the introduction of the new 5G technology. It offers a high data rate with minimum of a 1Gbit/sec, allowing several users to connect at the same time [1]. This necessitates a wide bandwidth, which is available at the mm-wave frequency. Figure 1 depicts the estimated mm-wave spectrum available for wireless communication. This band, however, is sensitive to a number of difficulties, including significant propagation loss, obstruction, and atmospheric loss. The performance of the antenna itself must be improved in terms of bandwidth, multi-band, directivity, and gain in order to overcome the propagation difficulties associated with mm-wave technology, make the most use of the available spectrum, and overcome related problems.

Especially, the investigation of multi-band antennas with relative effective quality in terms of bandwidth, gain, and efficiency should meet needs of wireless link. This form of antenna is required to provide a low profile, planar construction, and multi-band properties. Thus, microstrip antenna technologies are a popular alternative due to their low volume, inexpensiveness, simple geometric design, and lightweight [3]. The traditional patch antenna, on the other hand, has a restricted bandwidth, poor gain, and reduced directivity [4, 5]. As a result, creating multi-band microstrip patch antennas is always a challenging problem.

In the literature, many ways for advancing a patch antenna for quad-band operation were published. Quad-band antenna were designed in various mechanisms among them based on novel virus-mimicking fractal structure were presented [6]. Quad-band circularly polarized planar antenna is described [7]. It is suggested to use a patch antenna with a notched ground plane and regularly spaced $\frac{1}{4}$ wavelengths radiating strips for quad-band application [8]. An S-shaped PIFA is advised for implanted quad-band bio-telemetry applications in [9]. In [10], the authors developed analogous circuit model to create a special bowtie antenna layout for quad-band application. In

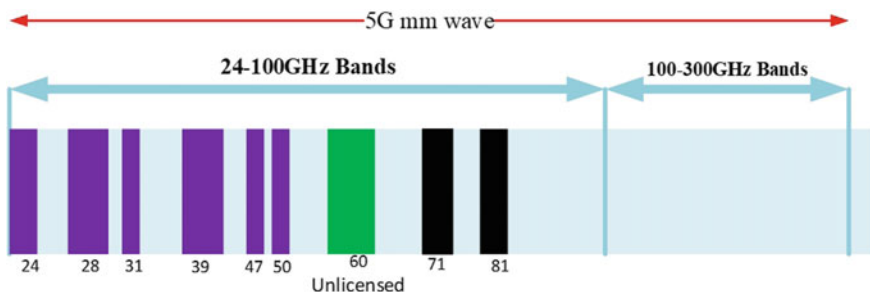


Fig. 1 Identified mm-wave spectrum for wireless application [2]

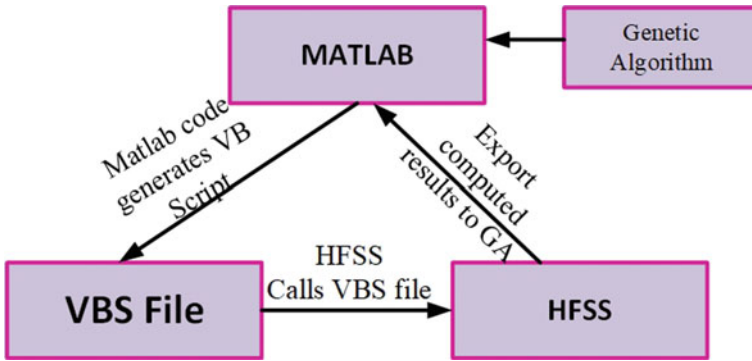


Fig. 2 Interfacing of MATLAB and HFSS in optimization procedure

[12], a quad-band optimization of a patch antenna in microwave frequency using a genetic algorithm is described. The author in [13] describes a multi-band antenna at mm-wave frequency that employs a genetic algorithm optimization.

This study provides a quad-band patch antenna improvement for mm-wave frequencies with increased gain, impedance bandwidth, and directivity utilizing a binary-coded genetic algorithm (GA). The technique is employed on the patch surface by manipulating the gridded rectangular cells to generate a new shape. MATLAB and the Ansys HFSS tools were used to study and evaluate the designated antenna. Ansys HFSS is used to model, compute, and export the resulting parameters to MATLAB’s main function. The antenna performance parameter is sent to the genetic algorithm optimizer code, which evaluates the fitness function, ranks individuals based on the fitness value, and picks the parents for the next generation. As shown in Fig. 2, the optimizer code is written in MATLAB, and HFSS receives the details of the antenna from the VBS file generated by MATLAB.

2 Antenna Modeling

A rectangular microstrip patch antenna designed on a $15 \times 15 \text{ mm}^2$ Roger-RT/duriod 5880(tm) substrate with a thickness of 0.5 mm, dielectric strength 2.2, and loss tangent of 0.0009. The initial dimension of a reference model, rectangular microstrip antenna was calculated using standard formula presented in [14]. However, at the given center frequency and bands, the computed dimensions do not have a suitable antenna performance. As a result, parametric study was performed to establish the radiating patch’s real width and length. Accordingly, the patch length was varied in the four stages from 7.6 to 8.2 mm, whereas, the patch width was adjusted in 0.2 mm increments from 9.2 to 9.8 mm. When the antenna is desired to resonate at 38 GHz center frequency, the optimal length is investigated as 7.8 mm and the width also 9.6 mm. A 50Ω microstrip feed line having length 3.4 mm and width 0.9 mm is

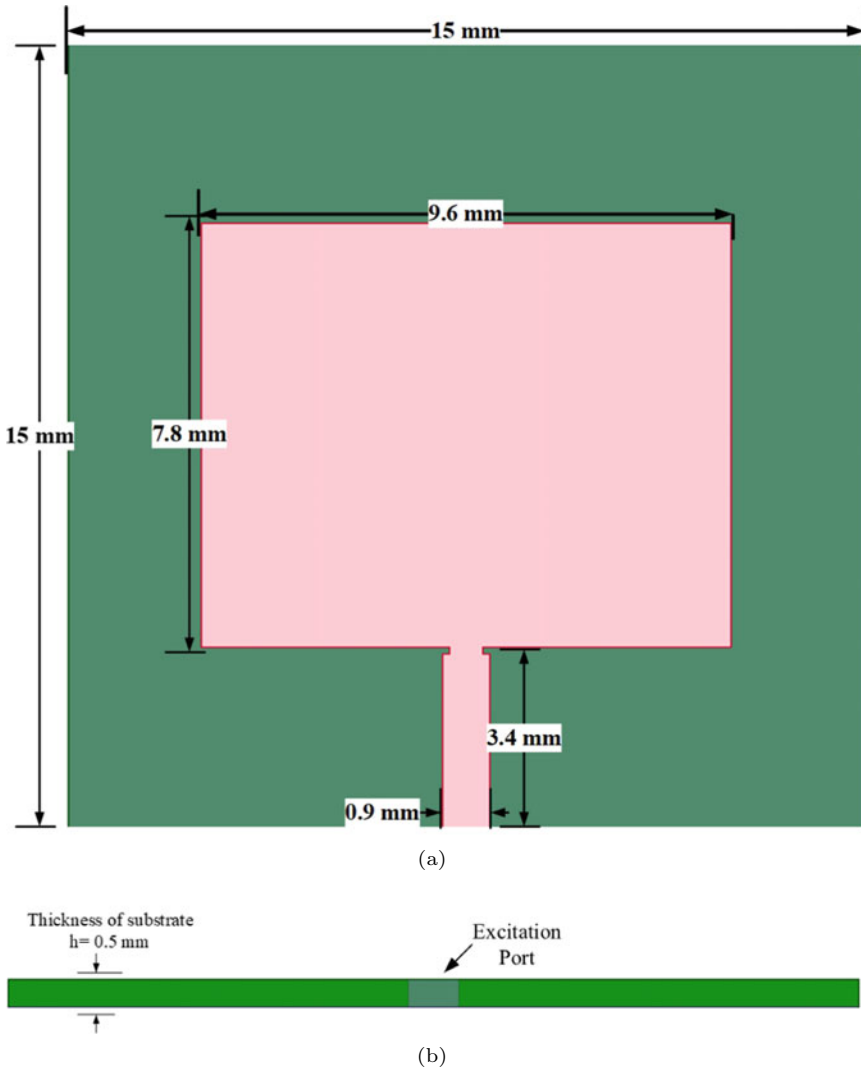
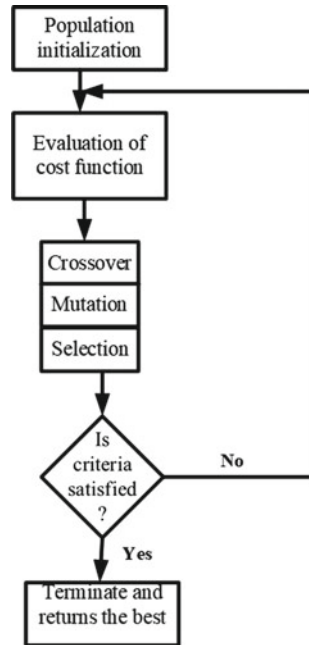


Fig. 3 Proposed reference patch antenna model at 38 GHz **a** top view **b** side view

connected to the patch. The reference antenna model and its dimensions are shown in Fig. 3, which is ready for patch geometry optimization with a binary-coded genetic algorithm.

Fig. 4 Flowchart of genetic algorithm



3 Optimization Procedure

A genetic algorithm (GA) replicates the method of natural selection, where the most fit individuals are picked to generate children for the following generation. In GA, each individual is constituted by gene string [15]. The initial population is generated randomly, while subsequent generations are produced by more fit and chosen individuals who have a higher likelihood of generating better offspring subjected to the fitness value. Crossover is the operator in charge of generating reproduction from two fitted individuals [16]. Mutation, on the other hand, is employed to increase population diversity. This iterative technique generates consecutive generations until a stopping requirement is met [17]. Figure 4 depicts a flowchart of genetic algorithm.

The patch surface is segmented into 6×6 tiny random and uniform rectangular cells to optimize the patch shape, as illustrated in Fig. 5. A binary-coding scheme defines the conducting and non-conducting features of patch cells. If the cell is conducting, a binary “1” is coded to the associated cell, whereas a non-conducting cell is represented by a binary “0” [18]. The goal of this optimization is to create a quad-band single microstrip antenna operating with sufficient performance. In light of this, the cost function is developed in the way described below:

$$\text{Fitness Function} = \frac{-1}{N} \sum_{i=1}^N S_{11}(f_i) \tag{1}$$

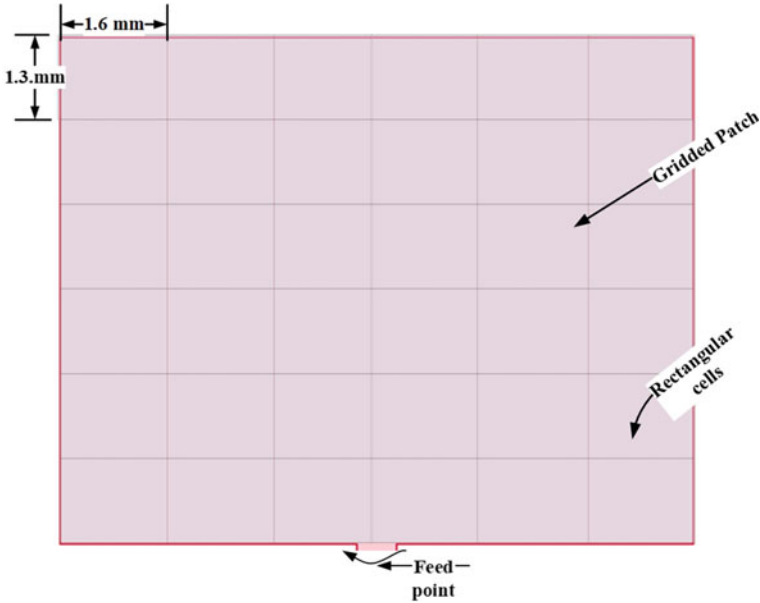


Fig. 5 Arrangement of 36 tiny rectangular cells on the surface of the patch

Where $S_{11}(f_i)$ is designed as.

$$S_{11}(f_i) = \begin{cases} S_{11}(f_i) & \text{if } S_{11}(f_i) \geq -10 \text{ dB} \\ -10 \text{ dB} & \text{if } S_{11}(f_i) \leq -10 \text{ dB} \end{cases}$$

where

- f_i the sample frequency at each 100 MHz interval
- N the total number of sample frequencies in the band
- $S_{11}(f_i)$ the antenna's reflection coefficient.

The optimization setup of GA has been arranged and summarized in Table 1. The solution space has a capacity of $2^{36} = 6.8 \times 10^{10}$ individuals in it. If each individual computing time is 1 s, the overall computation time to find the best-fitting individual will be 2.1×10^3 years. However, owing to a genetic algorithm, the best-fit individuals are picked in around 40.3 h only by utilizing a computer having core I7 with 2.7 GHz processing speed and a RAM of 8 GB.

Table 1 GA optimization setup

No.	Parameter	Values
1	Size of population	30
2	Chromosomes size	36 gens
3	Number of generation	100
4	Type of crossover	Single-point
5	Crossover probability	0.8
6	Type of mutation	flip bit
7	Mutations rate	0.01
8	Type selection	Tournament

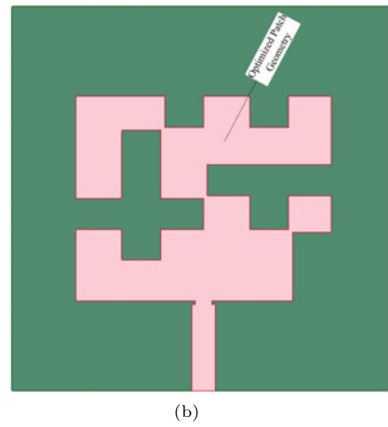
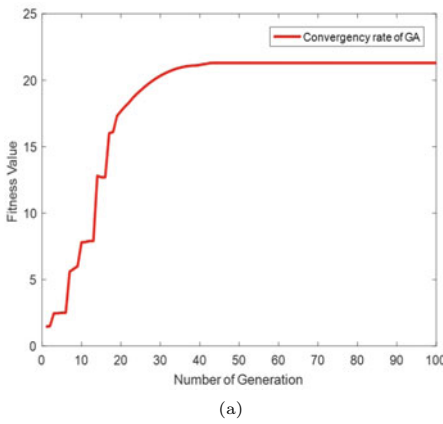


Fig. 6 **a** Average fitness value versus number of generation **b** Optimized patch geometry top view

4 Result and Discussions

The simulation converges after 40 iterations, as shown in Fig. 6a, and the iteration continues for the next 60 generations to demonstrate the consistency of the convergence. Finally, Fig. 6b depicts the most fitted individual antenna based on the algorithm.

Both the reference model and the genetically engineered antenna were simulated using ANSYS HFSS. The reference antenna resonated at 38 GHz single frequency band and had a covered impedance bandwidth of 1.2 GHz. The peak S_{11} value of the reference model at a 39 GHz is -38.4 dB, whereas the antenna’s maximum directivity is 6.1 dB. On the other hand, the proposed genetically optimized antenna works at four distinct frequencies: 28.3, 38.1, 46.6, and 60.0 GHz. The outcomes demonstrated that the antenna’s bandwidth has also improved. The directivity and gain are sufficiently better than the reference antenna and good for mm-wave wireless applications. As shown in Fig. 7, the antenna operates in quad-band, with a

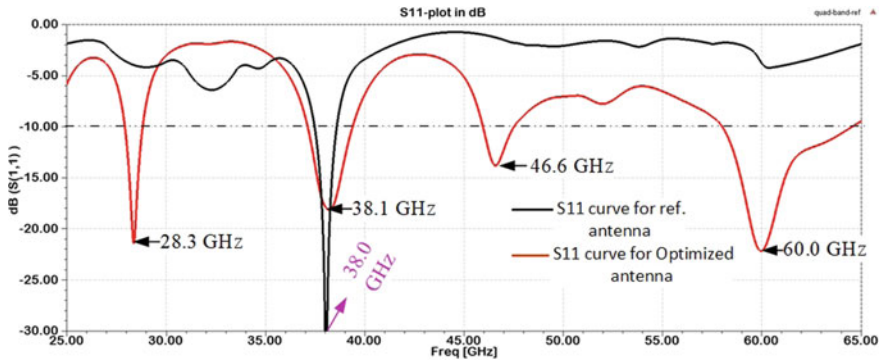


Fig. 7 Simulated results of S_{11} for optimized and reference model

peak value of return loss $S_{11} = -21.4$ dB, $S_{11} = -18.1$ dB, $S_{11} = -13.8$ dB, and $S_{11} = -22.1$ dB, respectively. When $S_{11} \leq -10$ dB is taken into account, bandwidth improvement of antenna in the operating bands is visible. At 28.3 GHz center frequency, the antenna has a fractional bandwidth of 3.2% or 0.9 GHz, 6.0% or 2.3 GHz, 3.4% or 1.6 GHz and 11.1% or 6.7 GHz at 28.3, 38.1, 46.6, and 60.0 GHz, respectively. VSWR is less than 2 in all operational bands which are: 27.9–28.8 GHz, 37.1–39.4 GHz, 45.9–47.5 GHz, and 57.8–64.5 GHz.

At four operational frequencies, the 3D gain plot of optimized antenna is presented in Fig. 8. The antenna’s maximum gain (8.6 dB) is achieved at 38.1 GHz, and at 60.0 GHz, the optimized antenna minimum gain is observed as 6.9 dB. The graphs indicate that the radiation patterns in all working band are toward broadside with some distortion in the upper frequency bands, whereas the rest radiation patterns are well shaped. Figure 9 depicts the antenna’s directivity pattern in both $\phi = 0^\circ$ and $\phi = 90^\circ$, which is almost projected in a broadside orientation. Furthermore, directivity enhancement in all working bands was observed as compared to the reference model. At 38.1 GHz, the maximum peak directivity of the optimized antenna is visualized as 8.8 dB at $\theta = 0^\circ$, while at 28.3 GHz 7.8 dB, 46.6 GHz 7.3 dB, and 60.0 GHz 7.1 dB broadside directivity was illustrated.

The distribution current on the surface of the genetically optimized antenna at four resonating frequency is presented in Fig. 10. Higher amount of current distribution is visualized on the edge of cells and the surface of optimized patch close to the feed.

5 Conclusions

This article synthesized a quad-band mm-wave microstrip antenna using binary-coded GA optimization. The proposed antenna was resonated at four distinct bands: 27.9–28.8 GHz, 37.1–39.4 GHz, 45.9–47.5 GHz, and 57.8–64.5 GHz, whereas the reference model was resonate only 38.0 GHz. A genetically modified microstrip

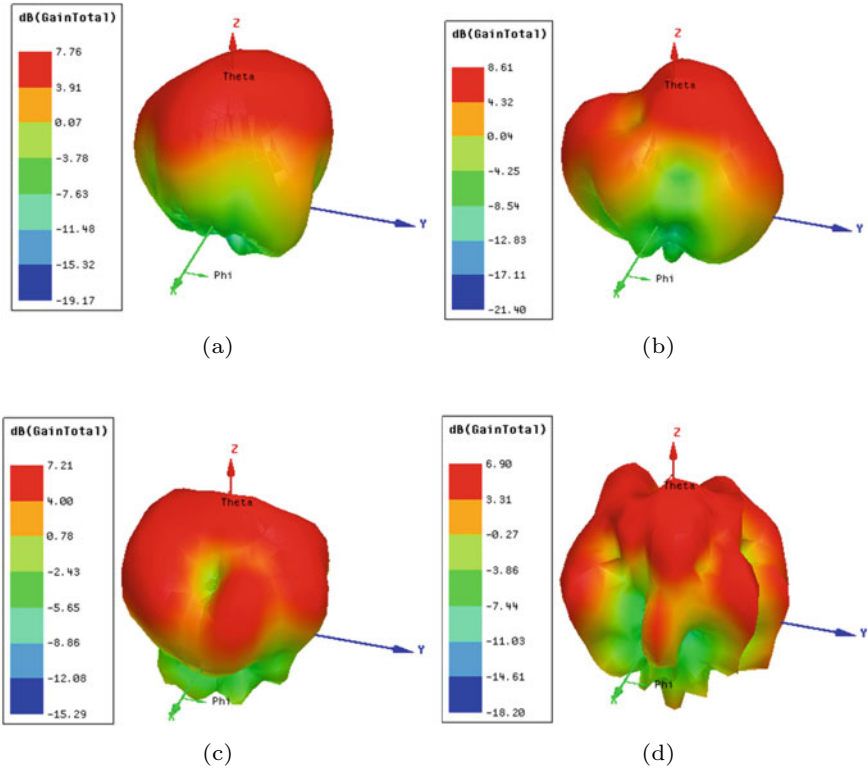


Fig. 8 The 3D plot of proposed antenna gain **a** 28.3 GHz **b** 38.1 GHz, **c** 46.6 GHz, **d** 60.0 GHz

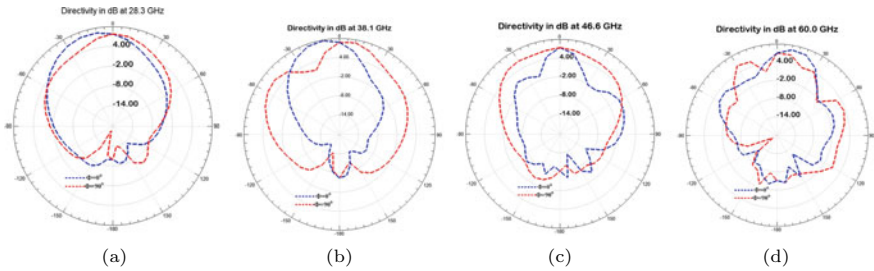


Fig. 9 The 2D radiation pattern plot in dB of proposed antenna **a** 28.3 GHz **b** 38.1 GHz, **c** 46.6 GHz, **d** 60.0 GHz

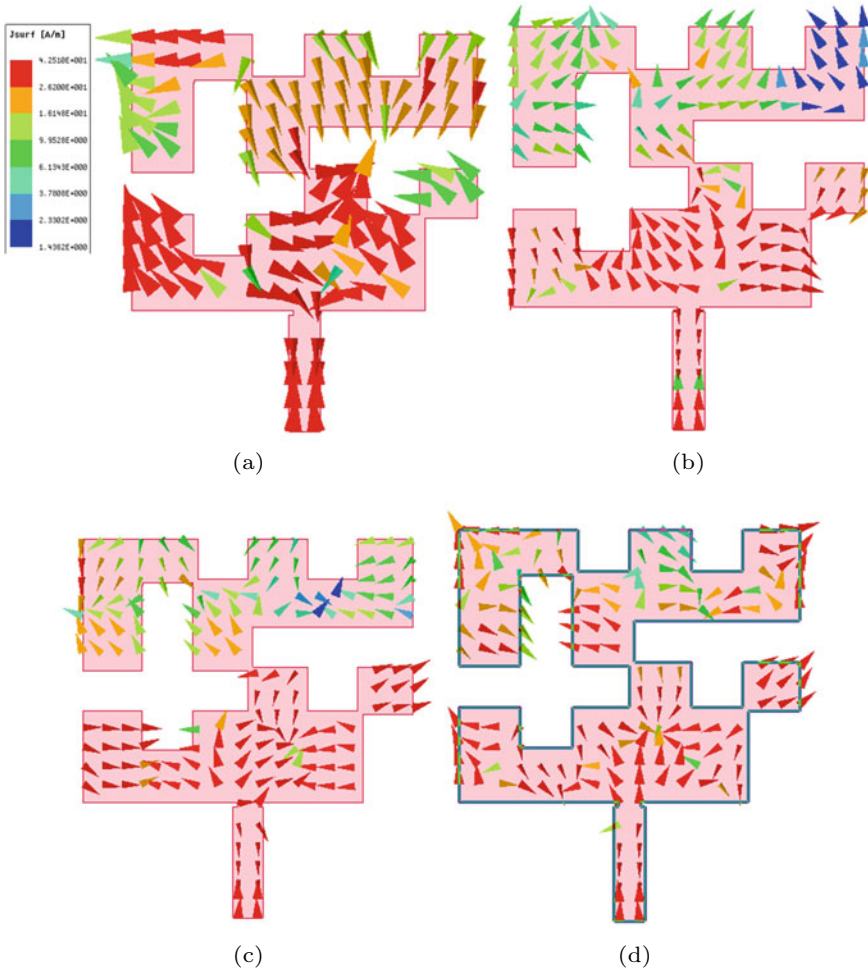


Fig. 10 Current distribution on optimized patch surface **a** 28.3 GHz **b** 38.1 GHz, **c** 46.6 GHz, **d** 60.0 GHz

antenna delivered outstanding performance in the targeted operating bands. With 7.1 dB of directivity, the optimized antenna has a maximum bandwidth of 6.7 GHz at 60.0 GHz. At 38.1 GHz with 2.3 GHz of bandwidth, the antenna also obtained a maximum directivity of 8.8 dB. To sum up, the optimized antenna outperforms the reference model in terms of bandwidth, gain, and directivity in all operating bands. These makes the antenna a viable alternative for mm-wave wireless applications.

References

1. Filipa SS, Fernandes CR, Rodrigues A, Svendsen S, Jagielski O (2021) Two-step method for millimeter-wave antenna performance assessment in 5G smartphones. *Int J Antennas Propag* Hindawi 2021:18 , Article ID 6641501. <https://doi.org/10.1155/2021/6641501>
2. Dejen A, Jayasinghe J, Ridwan M, Anguera J (2021) Optimization of dual-band microstrip mm-wave antenna with improved directivity for mobile application using genetic algorithm. In: 9th EAI International conference on advancements of science and technology Bahirdar, Ethiopia
3. Rappaport TS, Xing Y, MacCartney Jr GR, Molisch AF, Mellios E, Zhang J (2017) Overview of millimeter wave communications for fifth-generation (5G) wireless networks-with a focus on propagation models. *IEEE Transactions on Antennas and Propagation* 2017; vol. Special Issue on 5G
4. Hu C-N, Chang D-C, Yu C-H, Hsiao T-W, Lin D-P (2016) Millimeter-wave microstrip antenna array design and an adaptive algorithm for future 5G wireless communication systems. *Hindawi, Int J Antennas and Propag* 10(1)
5. Jaume A, Alejandro F, Carles P, Aurora A, Jaap G (2022) Antenna boosters versus flexible printed circuit antennas for IoT devices. *MDPI Sig* 3(2):326–340
6. Zhang L, Feng Q (2022) Design of a novel fractal quad-band-notched UWB antenna with bionic structure. *Progress Electromagnet Res M* 111:41–52
7. Dhara R, Yadav S, Sharma MM (2021) A circularly polarized quad-band annular ring antenna with asymmetric ground plane using theory of characteristic modes. *Prog Electromagnet Res M* 100:51–68
8. Wong HS, Kibria S, Islam MT, Mandeep JS, Misran N (2014) Quad band handset antenna for LTE MIMO and WLAN application. *Int J Antennas Propag* 2014:7. Article ID 341574. <https://doi.org/10.1155/2014/341574>
9. Wu C-K, Chien T-F, Yang C-L, Luo C-H (2012) Design of novel S-Shaped quad-band antenna for MedRadio/WMTS/ISM implantable biotelemetry applications. *Int J Antennas Propag* 2012:12, Article ID 564092. <https://doi.org/10.1155/2012/564092>
10. Moulay M, Abri M, Badaoui HA (2015) Quad-band bowtie antenna design for wireless communication system using an accurate equivalent circuit model. *Int J Microwave Sci Technol* 2015:7, Article ID 637607. <http://dx.doi.org/10.1155/2015/637607>
11. Almashhdany MB, Al-Ani OA, Sabaawi AMA (2020) Design of multi-band slotted mm-wave antenna for 5G mobile applications. In: 3rd International conference on sustainable engineering techniques (ICSET 2020), Canada (2020)
12. Jayasinghe JM JW, Uduwawala DN, Anguera J (2012) Design of dual band patch antennas for cellular communications by genetic algorithm optimization. *Int J Eng Technol* 1(1):26–43
13. Dejen A, Ridwan M, Jayasinghe J, Anguera J (2022) Multi-band mm-wave wearable Antenna synthesized with a genetic algorithm. *Int J Antennas Propag* 2022:17, Article ID 1958247. <https://doi.org/10.1155/2022/1958247>
14. Balanis CA (2013) *Modern antenna handbook*. A John Wiley and Sons inc, Canada
15. Haupt RL (1995) *An introduction to genetic algorithms for electromagnetics*. *IEEE Antennas Propag Magaz* 37(2)
16. Sarareh OA, Al Saraira AA, Alsafasfeh QH, Arfoa A (2016) Bio-inspired algorithms applied on microstrip patch antennas: a review. *Int J Commun Antenna Propag (I.Re.C.A.P.)*, 6(6)
17. Michael Johnson J, Rahmat-Samii Y (1997) Genetic algorithms in engineering electromagnetics. *IEEE Antennas Propag Magaz* 39(4)
18. Jmjw J, Anguera J, Uduwawala DN (2012) A simple design of multi band microstrip patch antennas robust to fabrication tolerances for GSM, UMTS, LTE, and Bluetooth applications by using genetic algorithm optimization. *Prog In Electromagnet Res M* 27(1):255–269
19. El-Hassan MA, Farahat AE, Hussein KFA (2022) Millimetric-wave quad-band MIMO antennas for future generations of mobile communications. *Prog Electromagnet Res B* 95:41–60

20. Kumar R, Kartikeyan M (2019) Design and simulation of multi band compact microstrip patch antenna. IEEE Indian Conference on Antennas and Propagation (InCAP), pp 1–4. <https://doi.org/10.1109/InCAP47789.2019.9134545>

Blackhole Attack Prediction in Wireless Sensor Networks Using Support Vector Machine



Niharika Panda and M. Supriya

Abstract One of the most pressing concerns in the Internet of Things (IoT) is security. The majority of IoT deployments rely on the creation of a wireless sensor network (WSN), which builds Low Power and Lossy Networks (LLNs) between a large number of constrained devices. WSN is an IPv6 low-power personal area network (6LoWPAN) communication protocol that uses the routing protocol for Low Power and Lossy Networks (RPL) for routing. Providing security to RPL network devices is difficult due to the resource constraints of the devices attached to RPL/IPv6. This work introduces attack detection using SVM (ADSVM) protocol to detect attacks before they have a significant impact on the IoT. Different WSNs are created considering the mobility factor of nodes. To forecast the attack, a new dataset centered on IoT traits is created and analyzed. Using eightfold cross-validation for attack prediction, the dataset has an accuracy of 84.37%.

Keywords WSN · RPL · Blackhole attack · ML · SVM

1 Introduction

The Internet of Things has ignited a tremendous technological revolution and received a lot of research attention in recent years [1]. IoT requires low-power, resource-constrained devices with mobility and global communication. 6LoWPAN is an example of a small-scale IoT network that interacts over IPv6 [2]. IoT nodes have limited resources, so routing helps in finding a solution to transport packets between nodes along the most efficient path [3]. The routing protocol for Low Power and Lossy Networks, which is based on the 6LoWPAN protocol, is a standard routing protocol for IoT networks [4]. The topology of RPL makes it easier to build a network with a

N. Panda (✉) · M. Supriya
Department of Computer Science and Engineering, Amrita School of Engineering, Amrita
Vishwa Vidyapeetham, Bengaluru, India
e-mail: p_niharika@blr.amrita.edu

M. Supriya
e-mail: m_supriya@blr.amrita.edu

large number of IoT nodes in static and mobile environments. RPL, on the other hand, is a security risk since it lacks the security protections necessary to prevent unauthorized access to data traveling across an IoT network [5]. Variety of applications have been analyzed w.r.t security including smart home [6], smart vehicle [10], smart farming [13], robotics [17], wearables, health care, environmental surveillance [18] and so on. The goal of this paper is to predict attacks in the generated novel dataset of blackhole attacks. The size of the network, the attacker's position, and the number of neighbors the attacker have been explored in this work [22]. In each topological design, one node is chosen as the attacker node, and the network is compared to one without a malicious node.

On the other hand, machine learning (ML) could be utilized for intrusion detection, and identification of attacks [14]. ML is applied to detect blackhole attacks based on network matrices in this study. In addition, due to the scarcity of acceptable IoT datasets, a dataset based on IoT properties and attack parameters is built to conduct tests. This paper is organized as follows.

2 Literature Survey

The SRPL-RPL protocol has been devised and applied in this study, which can overcome the inadequacies by lowering the impact of these attacks. The suggested SRPL-RP is more secure and efficient in terms of network performance and accuracy and holds an accuracy of 95% in a variety of network topologies [1]. Intrusion Detection System (IDS) is one of the security infrastructures established by researchers to defend IoT networks from compliance [2]. Ambarkar et al. [3] demonstrate hybrid attacks using a framework built on cooja and conduct RPL attacks on a few nodes picked at random. The work proposed in [7] discusses a new secure RPL routing (MLRP) protocol that uses a support vector classifier (SVC) for categorizing the attacks learned, an improvement of which is discussed in [14]. Jamalipour et al. in [8] present systematic research on RPL, its threats, and applicable IDS approaches. A survey on attacks concludes that blackhole, sinkhole, greyhole, and DAO are the issues that must be considered in order to obtain superior security [9]. The Internet of Things bridges the gap between the real and virtual worlds, erasing time and space constraints. It has caught the interest of cybercriminals, as one might expect [12].

Based on mobility-based trust metrics, Muzammal et al. [15] proposed SMTrust, a conceptual architecture for routing protocol security in the Internet of Things. Advanced machine learning-based IDS might be recommended as a viable option for predicting unfamiliar and unique attack instances while evaluating the huge data traffic in IoT [16]. The system obtained a test accuracy of 99.4% for decision tree, random forest, and artificial neural networks (ANN). The impact of a lowered rank attack on total network performance has been explored in this study on various network scenarios [21].

This research presents a security paradigm to detect threats. Prediction of attack is done using Support Vector Machine (SVM) algorithm.

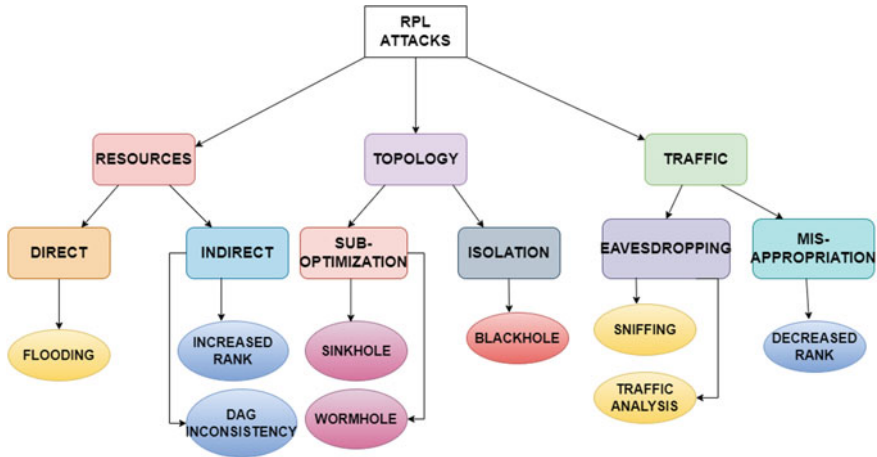


Fig. 1 Attacks category

3 Working Environment

The Internet of Things has been subject to several security vulnerabilities due to the widespread deployment of resource-restricted Low Power and Lossy Networks (LLN) [7]. There are several attacks, which may be designed to compromise network resources, performance [4], topology, and traffic in WSN. The main goal of the attacker node is to disrupt the network’s topology-building mechanism. It essentially disconnects one or more nodes from the network and prevents data communication between them.

3.1 Types of Attack

Attack in WSN is otherwise known as RPL attack. It is divided in to three categories, namely resources, topology, and traffic. Resource attack is of direct and indirect types. Topology attack is subgrouped to sub-optimization and isolation types. Lastly, the traffic attack is divided in to eavesdropping and misappropriation attacks. Classification of various attacks is shown in Fig. 1.

1. Flooding—The attacker consumes battery power by repeatedly delivering a large number of discovery packets or DIS in a network to parent nodes [19].
2. Increased Rank—The malicious node advertises a fake rank node’s location or a fake route through the root node in RPL control messages to deceive close nodes and lure them to forward their packets through it [1].

3. DAG Inconsistency—The attacks try to boost the DIO message’s version number value and then deliver it to its neighbors. As a result, the DODAG had to be rebuilt from the ground up, resulting in network congestion, substantial data packet loss [22].
4. Sinkhole—This attack attempts to gather as much traffic as possible by using bogus informational adverts to divert people to it. As a result, network performance will suffer.
5. Wormhole—When two malicious nodes utilize legitimate ways to establish a malicious tunnel between the sender and destination nodes and persuade nearby nodes to use the malicious tunnel, it’s called a wormhole attack [20].
6. Blackhole—The blackhole attack persuades its surrounding nodes to deliver packets to the destination node through an immorally shorter path, and so removes all or chosen packets passed by neighboring nodes [19].
7. Sniffing—This attack entails listening in on packets sent over the network by nodes, compromising the transmission’s confidentiality.
8. Traffic Analysis—The goal of this attack is to gather enough information to carry out another type of attack [19].
9. Decreased Rank—Nodes that are closer together appear to be more desired than ones that are farther apart. False DIO packets could give malicious nodes a lower rank value, allowing them to draw more traffic and carry out attacks [1].

3.2 *Blackhole Attack*

A malicious node draws other nodes in a blackhole attack by falsely offering better routes to their destination.. Later, it drops all packets that the attacker node was supposed to forward. If the attacker node’s position is correctly chosen, it can isolate numerous nodes from the network [19]. This attack can also be considered as a denial of service attack. Figure 2 describes the blackhole attack with a network topology. After the network topology is established, node 4 is considered as the attacker node. Following the concept of blackhole attack, the node number 4 is detaching all of its child nodes from the topology. Due to which nodes 6, 7, 9, 10 are disconnected from the topology and will not be able to participate in interoperability. In this paper, the effect of a blackhole attack is demonstrated with and without mobility in different networks.

4 Model Description and Simulation

4.1 *Topology Considered*

There are ten sender nodes and one sink node in the network presented in Fig. 3. The network consists of 20 sender nodes and one sink node with two different topologies as can be seen in Figs. 4 and 5. Each network is simulated in both the absence of an

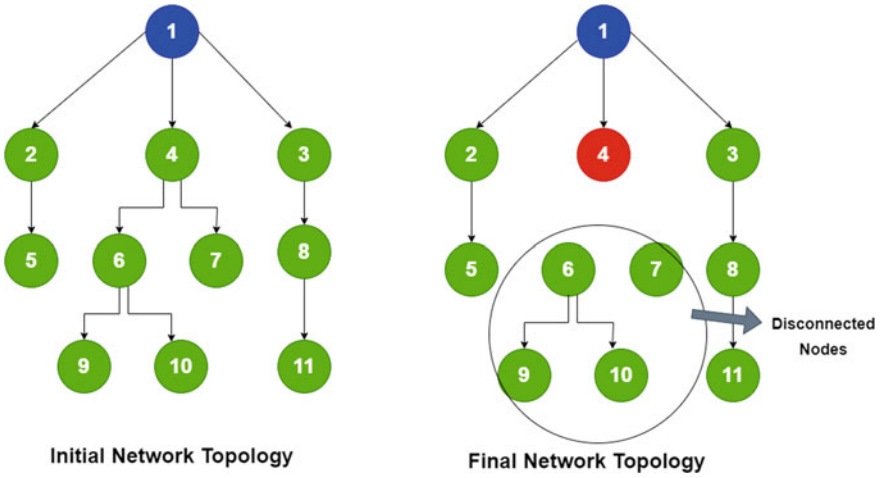
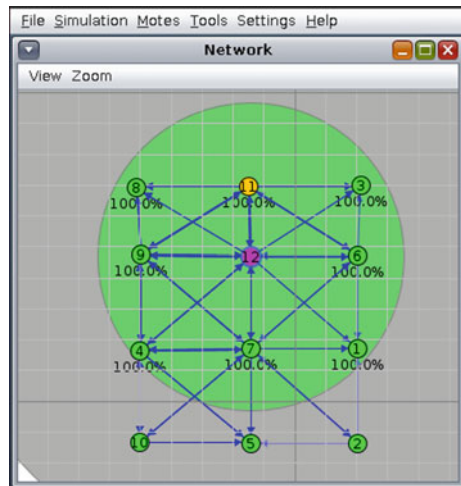


Fig. 2 Blackhole attack

Fig. 3 12 nodes topology



attacker node and the presence of a single attacker node. The above-mentioned scenario is evaluated in a mobile environment as well, where all sender nodes, including the attacker, are considered to be in motion as shown in Fig. 6.

4.2 Dataset

During dataset exploration, a range of methodologies were investigated, including publicly available datasets, privately kept datasets, and an IoT simulated dataset. The

Fig. 4 22 nodes topology type 1

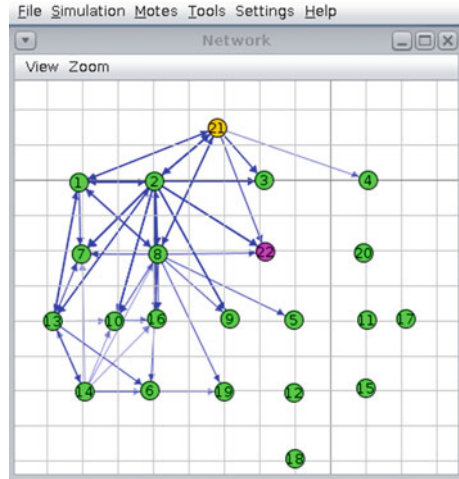
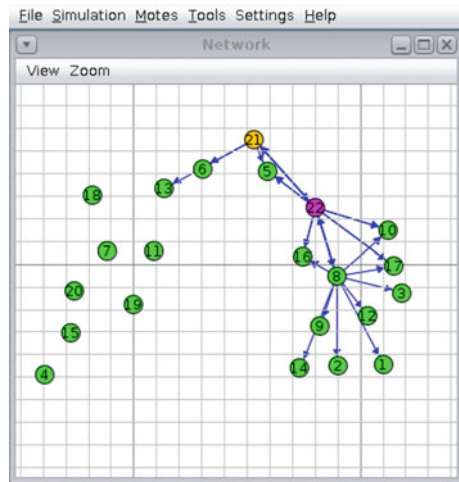


Fig. 5 22 nodes topology type 2



DARPA 1998/1999/2000, KDDCUP 99, and NSL-KDD datasets were chosen at first because they are commonly used in ML research [11]. These datasets, despite being publicly accessible and well-labeled, lack examples of attacks that have occurred since the datasets were created. The existing dataset has a number of issues, including a large number of redundant records, properties that aren't correctly characterized, massive data sizes, and fake attacks are to name a few [11]. When it comes to attack in a WSN or RPL with objective functions OF0 and MRHOF, there is no such dataset available. As a result, this study considers creating a novel dataset from a IoT simulated environment using Contiki-NG. After analyzing the performance parameters, there are 11 attributes selected with 264 sets of tuple in the dataset. In

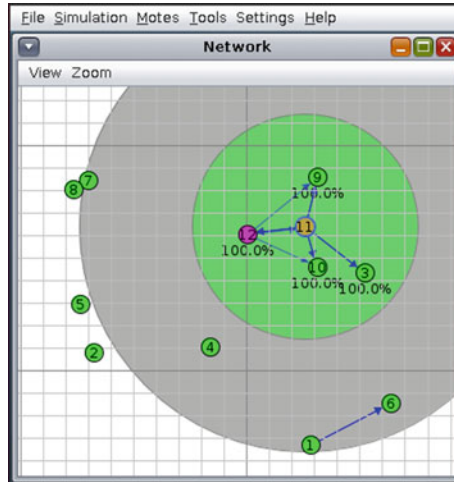


Fig. 6 12 nodes topology with mobility

	IDs	DIOs	DISSs	DAOs	Precv	Pforw	Pdrop	Psent	Network	OF	Attack
0	1.0	1177.0	82.0	1123.0	394.0	149.0	0.0	313.0	Static	MRHOF	1.0
1	2.0	187.0	10.0	0.0	201.0	0.0	0.0	266.0	Static	MRHOF	1.0
2	3.0	236.0	2.0	2.0	409.0	177.0	0.0	242.0	Static	MRHOF	1.0
3	4.0	253.0	41.0	1.0	600.0	305.0	0.0	474.0	Static	MRHOF	1.0
4	5.0	285.0	21.0	0.0	311.0	0.0	0.0	282.0	Static	MRHOF	1.0

Fig. 7 Dataset sample

this work attack prediction is done w.r.t different control messages, packets send and received with different objective functions in static and mobile environment.

A sample of the novel dataset used for the work is presented in Fig. 7.

4.3 Attack Detection Using SVM

It is critical to protect the IoT system by anticipating attack behavior and adapting to changing network conditions. A machine learning technique is the best answer for IoT because of the heterogeneity and plenty of data creation. Supervised learning, unsupervised learning, deep learning, and reinforcement learning are the four key aspects of machine learning algorithms for IoT security [5]. A labeled dataset is used in the supervised learning-based technique to assess the characteristics in order to learn the behavior of the network. Each input is mapped to a class label output

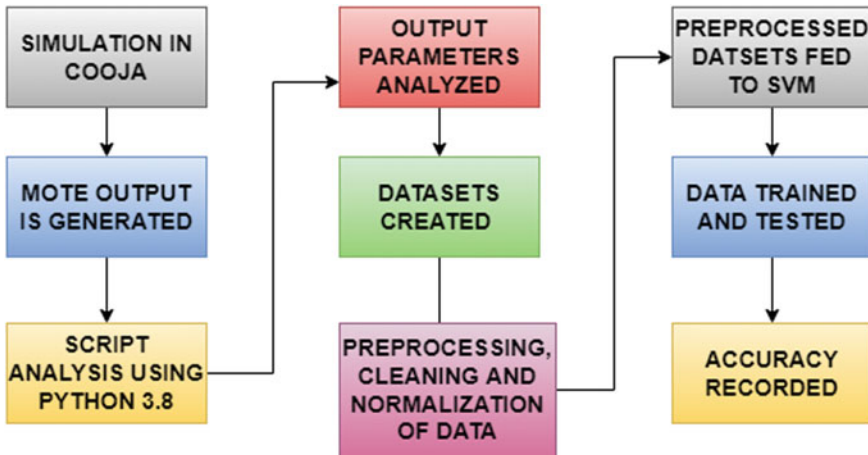


Fig. 8 Step of work

in essence. The basic aim of a supervised classifier with respect to this work is to predict intrusion, given a set of network flows and a training class label in the form of attacks. The learning algorithm is expected to learn the scenarios considered and is expected to predict the attack in the future. Attack Detection using SVM (ADSVM) is proposed as a protocol to detect the attack in the network. The phases pertaining to the attack prediction are shown in Fig. 8. The environmental setup is done initially in Contiki-NG with the help of cooja simulator. It is emulated using 12 and 22 nodes in two separate networks, each with one attacker node. In various configurations, nodes are classified as both static and movable. The simulation lasts for 30 min, documenting various operations carried out by the client, server, and attacker nodes. Three different datasets are generated from result parameters, one for a network of 12 nodes, another 22 nodes, and the third for the combined dataset. Preprocessing has been applied to each dataset. Cleaning of attributes that are not a factor in attack is carried out. The attack attribute is normalized, with values of 0 and 1 representing no-attack and attack in the network, respectively. Later, the machine learning algorithm is fed the final dataset. Furthermore, based on the learning process with respect to the attack attribute, the SVM in ADSVM determines and classifies RPL attacks. There are 11 attributes and 264 tuples in all. 90% of the data is trained, with the remaining being tested afterward. Using eightfold cross-validation procedures, the final result for the combined dataset is collected.

5 Results and Discussion

This section analyzes the results of the simulation study. The simulation begins with a simple attack environment consisting of ten clients, a sink node, and an attacker node.

Later, the simulation continues on to a 20-client, one-sink, and one-attacker node. In this effort, mobile nodes are integrated into each type of network. To categorize the dataset and determine the accuracy of attack detection, the proposed ADSVM protocol is used. To train, test, and find accuracy, ML codes are developed using Jupyter notebook and Python. Using the generated confusion matrix, the evaluation metrics are obtained as presented in Eqs. 1 through 4. Table 1 illustrates the values w.r.t each network.

A confusion matrix is a table that summarizes the results of a classification problem. It helps in evaluating the performance of the classifier with respect to the chosen test dataset. It determines the positives and negatives for each of the true and false cases. A confusion matrix can also be generated for cross-validation in which the model accuracy and other performance parameters are represented with the average of all the k -folds. The performance parameters estimated from the confusion matrix are accuracy, precision, recall, and f_1 score [7].

$$\text{Accuracy} = \frac{TP + TN}{P + N} \quad (1)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{P} \quad (3)$$

$$\text{F}_1 \text{ Score} = \frac{2 * TP}{2 * TP + FP + FN} \quad (4)$$

where,

$$P = TP + FN$$

$$N = TN + FP$$

TP denotes True positive, Prediction is positive for positive values in dataset

TN denotes Prediction is negative for negative values in dataset

FP denotes Prediction is positive for originally negative values

FN denotes Prediction is negative for originally positive values.

The precision and recall values are 75% and 93%, respectively, for the initial network dataset with ten client nodes. Similarly, 20 client nodes have the highest precision and recall, as shown in, with 79% and 92% precision and recall, respectively. The accuracy of the combined dataset, which has 264 cases, is 70%. Cross-validation, a re-sampling process used to evaluate machine learning models on a restricted data sample, is used to improve the performance of the primary dataset. As a result, k -fold cross-validation is a common name for the procedure. When a precise value for k is chosen, it can be used in the model reference instead of k . In this work, k is assumed to be 8. When employing cross-validation, the accuracy of the combined dataset is raised to 84.37% for blackhole attack prediction.

Table 1 Evaluation of metrics using SVM

Network		SVM classification			
Nodes	Class	Precision	Recall	F1-Score	Accuracy (in %)
12	W/o attack	93	93	93	89
	With attack	75	75	75	
22	W/o attack	75	50	60	78
	With attack	79	92	55	
Combined	W/o attack	50	75	60	70
	With attack	87	68	76	

6 Conclusion

The current study's experimental analysis provides a realistic understanding of how IPv6-based IoT networks respond to blackhole attacks. The goal of the article is to run simulations for various network sizes with the set transmission, interference, and reception ranges. To demonstrate the performance degradation of the matrices in the network, an attacker node is added to both the static and mobile scenarios. Finally, using ADSVM, a fresh dataset is constructed from the simulated environment to detect attacks. The accuracy of blackhole attack detection is determined using the SVM classification approach. The accuracy of the total dataset of all the communication networks illustrated in the simulation is 84.37% when using eightfold cross-validation.

In addition, a new mechanism for detecting attacks and increasing the performance of matrices inside the same network could be proposed. More datasets should be added to ML-based IoT research. Various other machine learning algorithms could also be implemented to predict the attacks in wireless sensor networks.

References

1. Almusaylim A, Jhanjhi Z, Alhumam A (2020) Detection and mitigation of RPL rank and version number attacks in the internet of things: SRPL-RP. *Sensors* 20(21):5997
2. Alreshoodi M (2020) An experimental study of iot networks under internal routing attack. Available at SSRN 3690813
3. Ambarkar SS, Shekokar N (2021) Impact analysis of rpl attacks on 6lo wpan based internet of things network. In: 2021 IEEE International conference on electronics, computing and communication technologies (CONECCT), IEEE, pp 1–5
4. Duttagupta S, Kumar M, Ranjan R, Nambiar M (2016) Performance prediction of iot application: an experimental analysis. In: Proceedings of the 6th international conference on the internet of things, pp 43–51 (2016)
5. Foley J, Moradpoor N, Ochenyi H (2020) Employing a machine learning approach to detect combined internet of things attacks against two objective functions using a novel dataset. *Secur Commun Netw*

6. Gupta K, Krishna GG, Anjali T (2020) An iot based system for domestic air quality monitoring and cooking gas leak detection for a safer home. In: 2020 International conference on communication and signal processing (ICCSP). IEEE, pp 0705–0709
7. Hasan M, Islam MM, Zarif MII, Hashem M (2019) Attack and anomaly detection in iot sensors in iot sites using machine learning approaches. *Internet of Things* 7(100):059
8. Jamalipour A, Murali S (2021) A taxonomy of machine learning based intrusion detection systems for the internet of things: a survey. *IEEE Internet Things J.* <https://doi.org/10.1109/JIOT.2021.3126811>
9. Johnraja JI, Leelipushpam GJ, Jebadurai JJ, Santhosam IB et al (2021) Security and privacy issues in the internet of things—a survey. In: 2021 Fifth international conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC). IEEE, pp 56–63
10. Kiran VV, Santhanalakshmi S (2019) Raspberry pi based remote controlled car using smart-phone accelerometer. In: 2019 International conference on communication and electronics systems (ICES). IEEE, pp 1536–1542
11. Lalduhsaka R, Khan A, Roy AK (2021) Issues and challenges in building a model for intrusion detection system. In: 2021 5th International conference on information systems and computer networks (ISCON). IEEE, pp 1–5
12. Lin H, Bergmann NW (2016) Iot privacy and security challenges for smart home environments. *Information* 7(3):44
13. Manideep M, Thukaram R, Supriya M (2019) Smart agriculture farming with image capturing module. In: 2019 Global conference for advancement in technology (GCAT). IEEE, pp 1–5
14. Momand MD, Mohsin MK et al (2021) Machine learning-based multiple attack detection in rpl over iot. In: 2021 International conference on computer communication and informatics (ICCCI). IEEE, pp 1–8
15. Muzammal SM, Murugesan RK, Jhanjhi NZ, Jung LT (2020) Smtrust: proposing trust-based secure routing protocol for rpl attacks for iot applications. In: 2020 International conference on computational intelligence (ICCI). IEEE, pp 305–310
16. Pasikhani AM, Clark JA, Gope P, Alshahrani A (2021) Intrusion detection systems in rpl-based 6lowpan: a systematic literature review. *IEEE Sens J*
17. Praveen A, Radhika R, Rammohan M, Sidharth D, Ambat S, Anjali T (2020) Iot based smart bin: a swachh-bharat initiative. In: 2020 International conference on electronics and sustainable communication systems (ICESC). IEEE, pp 783–786
18. Raviteja K, Supriya M (2020) Greenhouse monitoring system based on internet of things. In: International conference on communication, computing and electronics systems. Springer, pp 581–591
19. Roy AK, Khan AK (2019) Performance degradation in wireless mesh networks via external and internal attacks. In: 2019 2nd International conference on innovations in electronics, signal processing and communication (IESC). IEEE, pp 258–262
20. Roy AK, Khan AK (2020) Privacy preservation with rtt-based detection for wireless mesh networks. *IET Inf Secur* 14(4):391–400
21. Sahay R, Geethakumari G, Modugu K (2018) Attack graph-based vulnerability assessment of rank property in rpl-6lowpan in iot. In: 2018 IEEE 4th world forum on internet of things (WF-IoT). IEEE, pp 308–313
22. Tonapa YT, Wahidah I, Karna NBA (2020) Performance testing of routing protocol for low power and lossy networks (rpl) against attack using cooja simulator. *eProc Eng* 7(2)

Identifying Exoplanet Candidates with Machine Learning



**Chaitanya Virmani, Ria Singhla, Priyanka Gupta,
and Hardeo Kumar Thakur**

Abstract An exoplanet or an extrasolar planet is a planet that orbits around the axis of a star that is not a part of our solar system. Apart from providing or disagreeing with the scientific theories that have been followed on Earth, the exploration of exoplanets has much more to tell. Scientists find it amusing to discover something that no one has ever seen before and can answer many of the unresolved questions. This can help us answer whether there is an Earth-like planet out there and whether life can exist on them. In this study, the agenda is to identify exoplanet candidates among the various celestial or other space orbiting objects that had been captured during the K2 mission. We have applied machine learning and deep learning algorithms along with feature selection and feature extraction techniques and found that recursive feature elimination followed by the fully connected model to be the most efficient one.

Keywords Exoplanets · Extrasolar planets · Kepler mission · Planet

1 Introduction

One of Astronomy's most dynamic research includes exoplanets. These planets are hidden by the bright glare of the stars that they orbit and thus are discernible. So, astronomers find different ways to detect and study these exoplanets. Some of the ways by which astronomers detect them are—microlensing, transit event observation, radial velocity, and direct imaging.

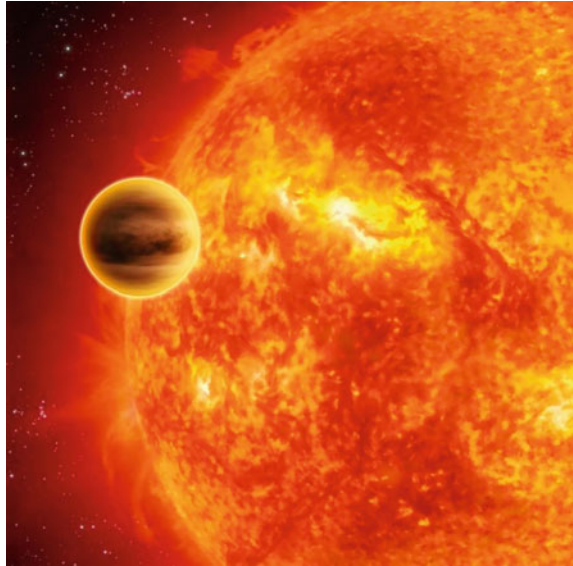
NASA's popular Kepler mission aimed at surveying the region of the Milky Way galaxy which discovered thousands of Earth-sized and smaller planets. One such

C. Virmani (✉) · R. Singhla · P. Gupta · H. K. Thakur
Manav Rachna University, Faridabad, Haryana, India
e-mail: chaitanyavirmani@gmail.com

R. Singhla
e-mail: riasinghla1@gmail.com

P. Gupta
e-mail: priyankagupta@mru.edu.in

Fig. 1 51 Pegasi b [1]



exoplanet is shown in Fig. 1. These planets are discovered around the habitable zone. The objective of this mission includes:

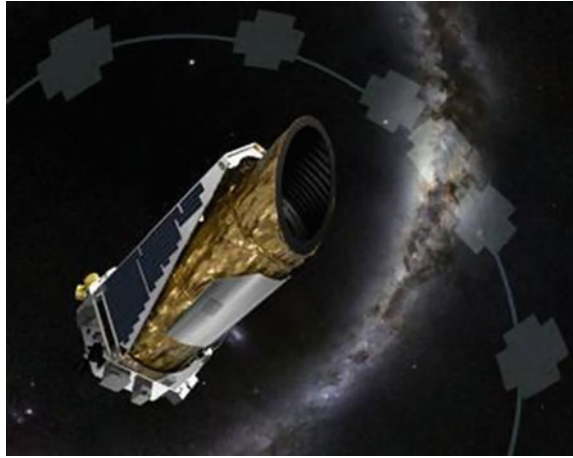
- Determining the percentage of planets in habitable zones with a wide variety of stars.
- Determining orbit size and shape.
- Determining the properties of surrounding stars.
- Determining sizes, masses, planet reflexivity, etc.

Kepler's second mission known as the K2 was a community-driven mission. It started after the Kepler mission's end and by May 2014, became fully operational. The Kepler Space telescope shown in Fig. 2, continued the space observations and monitored 150,000 stars that were studied during the Kepler mission. It used photometry to detect planets transiting their parent star.

The K2 mission came to an end after NASA announced the retirement of the telescope on October 30, 2018. By the end of the K2 mission, more than 50,000 stars and more than 2000 exoplanets were discovered.

The discovery of exoplanets can answer various scientific questions. One of these is "How our planets have formed?". Scientists find it amusing to discover something that no one has ever seen before and can answer many of the unresolved questions. This can help us answer whether there is an Earth-like planet and whether life can exist on it. Studying about these planets can enhance the journey of exploration. One of the most profound questions, "Are we alone?" can be answered. Many scientists research exoplanets to know how things have been formed and can answer various questions that have been raised while studying the process in our own solar system.

Fig. 2 Kepler Space telescope [2]



Much work has already been done in automating the identification process for exoplanet candidates, so scientists do not have to sift through each data point manually, reducing their time and effort. With the same curiosity, we have chosen to make the process of determining whether the given object is an exoplanet candidate or not more efficient. In this research, the work that has been done is briefly described in various sections. Section 2 explains the state of the art, and Sect. 3 explicates the dataset that has been used. Preprocessing, model building, and model architecture are discussed in Sect. 4. Section 5 talks about the results that are achieved during the model building along with hyperparameter tuning. Section 6 provides a conclusion to our result and scope of future work.

2 State of the Art

Exoplanet identification is a topic of wide interest, and a lot of useful research has already been done using various classical machine learning and deep learning models. Kepler collected photometry data between 2009 and 2018. The goal of the Transient Exoplanet Survey Satellite (TESS) was to find transiting extrasolar planets. Various datasets have been used that contain images, light signals, and numerical features. [3–6] used light curves to detect exoplanets in space. CNN, SVM, and DNN models are used on these light curves.

Pearson et al. [7] have used an artificial neural network to learn the photometric features of a transiting exoplanet. Using a phase folding methodology that produces a constraint when fitting for the orbital period, they developed a technique for locating periodic transits. For evaluating the model, they divided the data into overlapping light curves and calculated the cumulative probability phase series (PPS).

Yip et al. [8] used GANs to create a synthetic dataset and used the ConvNet model and captured an accuracy of 97% on coronagraphic images captured manually in 2020. Dattilo et al. [9] used the AstroNet model and captured an accuracy of 97.84% in 2019 on K2 data.

Yu et al. [3] worked on TESS data. They refined the data by identifying threshold crossing events (TCEs). TESS data do not come with assigned labels, and therefore, the data is manually labeled by visually inspecting the light curves. A CNN-based model was trained which captured an accuracy of 97.8%.

Mathur et al. [10] generated synthetic data using the synthetic minority over-sampling technique (SMOTE) to balance the dataset used. Along with identifying exoplanets, the probability of habitability on these exoplanets is studied and they are also classified as mesoplanets, psychroplanets, and non-habitable planets. CNN model has been used to identify these exoplanets based on TCE and light curves. The model is evaluated in three different scenarios—without preprocessing, with Gaussian filter, and with Savitsky Golay filter. It captured an accuracy of 89.12%, 71.83%, and 88.99%, respectively, on training sets.

Many of the previous studies identifying exoplanets were done using light curves. We are using numerical features instead, along with feature reduction and feature extraction techniques to improve the performance.

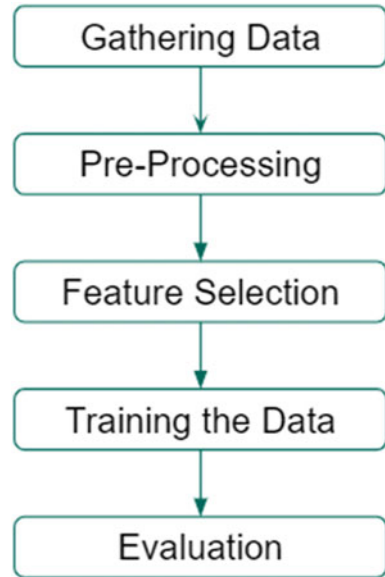
3 Description of the Dataset

The data is provided by NASA Exoplanet Science Institute in the NASA Exoplanet Archive [11] which stores K2 exoplanet candidates and non-candidates. The dataset is stored in a table format that contains all K2 planets and candidates.

This contains 138 features. The features that are given in the dataset were recorded during the mission, some of which include orbit period, transit epoch, impact parameter, equilibrium temperature, stellar surface gravity, stellar radius, and much more along with a feature description documentation. Using this document, all the irrelevant features are dropped and a model is built that will automate the process of finding an exoplanet candidate more efficiently.

4 Proposed Method

The preprocessing stage before passing the data to any ML model is really important. In this section, the work done before passing our data through the training phase is discussed in detail. Figure 3 shows the model pipeline.

Fig. 3 Model pipeline

4.1 Data Preprocessing

The tabular data that is used comprises 139 features and 8054 instances. After going through the details of each feature provided by NASA in the description document, it is clear that there were many redundant features that were not required for example 'kepid', 'kepoi_name', 'kepler_name', and 'koi_disposition'. After dropping all of those redundant features, the feature size was reduced to 38.

The next important step is detecting outliers. Outliers can affect the performance of a model in drastic ways. IsolationForest library is used to detect the outliers. It is present in the sklearn module that detects outliers by finding the anomaly in the data. These outliers are dropped after identification.

When the data is balanced, it is easier for the model to learn the differences clearly. Our dataset is balanced after removing the outliers and the null values that are present in the dataset, with a ratio of 3739:3985.

Some of the features are highly correlated such as upper temperature bound and lower temperature bound. One of the correlated features is taken, others are dropped out of all correlated features, and the data is split in 75:15:10 into train, test, and validation sets, respectively.

At last, scaling of data is done using StandardScaler to scale the values. Scaling is an important task in preprocessing which reduces the scale of the data and hence, the values are normalized.

4.2 *Feature Selection*

Feature selection refers to the removal of noise or redundant data and using only relevant features. Most of the feature selection task is already done in the preprocessing segment. Further, in order to predict the outcome, dimensionality reduction techniques like recursive feature elimination (RFE) and principal component analysis (PCA) are used with the models.

4.3 *Hyperparameter Tuning*

Hyperparameters are the arguments that are chosen before the learning process begins. It controls the learning process and is done to make the predictions better.

In the models, the batch size is set to an optimal value such that values of weights are more accurate and overfitting is avoided. Early stopping is used as a callback function. This function automatically stops the model from further training as soon as the model starts to overfit. Validation loss is monitored in this function. Therefore, choosing the number of epochs was not a manual task. Best weights are restored when the training is stopped. Adam is used as an optimizer and binary cross-entropy as the loss function, and accuracy is monitored with each epoch. In PCA, the explained variance is set to 95%. This avoided the manual process of selecting the best features.

4.4 *Fully Connected Model*

The fully connected model contains only dense layers, i.e., every neuron of the previous layer is connected with every neuron of the current layer. The activation function used in hidden layers is rectified linear unit (ReLU), and the output layer has Sigmoid as an activation function.

The number of neurons is selected by testing for valid values in each layer. The model was not showing satisfactory performance when the number of neurons was less. With the increase in the number of neurons, weight got increased and it became more clear for the model to learn the differences and thus, captured better performance. Increasing the number of neurons to a greater value was overfitting the model.

The model is built for different numbers of hidden layers ranging from 1 to 5. Adam and binary cross-entropy is used as optimizer and loss function, respectively. Early stopping is used as a callback function that monitors validation loss and the patience level is set to 5. It automatically stops training when the model starts overfitting.

4.5 Random Forest Model

Random forest is one of the most popular ensemble models that is used in supervised learning for classification as well as regression. In this, multiple decision trees are used to calculate the result and the final result is obtained by considering the majority in the case of classification.

In our model, `n_estimators = 100` and `max_depth = 10` are passed as parameters. These values were found to be optimum, balancing bias, and variance.

4.6 PCA Followed by Fully Connected Model

Fully connected model is used again, but this time features are extracted using principal component analysis (PCA).

PCA is a dimensionality reduction technique. It creates new features based on the original features. The number of features to train the model can be manually selected. These features are generated before they are passed to the dense model to predict the output.

23 features are generated using PCA. Instead of selecting the number of features manually, a variance value of 95% is chosen and it has generated 23 features that are linear combinations of input features. Due to the selection of fewer features, the model was not able to learn properly in accordance with small amounts of data. On the other hand as well, due to the selection of a greater number of features, the data became sparse and was again not able to perform efficiently.

The model is tested for a different number of hidden layers ranging from 1 to 5. Three dropout layers of threshold values, i.e., 0.25, 0.2, and 0.4 are used, and the output is calculated on the grounds of the sigmoid function. Adam is used as an optimizer and binary cross-entropy as the loss function. Results are captured for varying batch size shown in Table 2, and early stopping is used as the callback function.

4.7 PCA Followed by Random Forest Model

In this model, PCA followed by the random forest model is used. `n_components` is again set to 95%, and 23 features are obtained for prediction. For random forest parameters, `n_estimators` and `max_depth` are set to 100 and 10, respectively.

4.8 RFE Followed by Fully Connected Model

Recursive feature elimination (RFE) is a feature selection technique that ranks the dependent features based on their dependency on the target variable recursively. It eliminates the weakest features based on their correlation and feature importance till the number of features to be selected is attained by choosing a subset starting with all the features.

Six features are captured using RFECV, a cross-validation method, on which the response variable depends. The fully connected model is tested for a variation of hidden layers in the range of 1–5. Four dropout layers are used with 0.25, 0.25, 0.25, and 0.4 as threshold values, respectively. The optimizer used is Adam, and the loss function is binary cross-entropy. Results are obtained for varying batch sizes that are shown in Table 2 and early stopping is used as a callback function with the same parameters as in the fully connected model.

4.9 RFE Followed by Random Forest Model

In this model, the six best features using RFECV are selected. It uses cross-validation method to select the best features automatically and then, those are passed to the random forest model. 100 decision trees are used as estimators with each having a maximum depth of 10.

5 Result Analysis

Various evaluation metrics are used to evaluate the performance of the models. In this section, hyperparameter tuning results are also discussed.

5.1 Varying Numbers of Hidden Layers

The number of hidden layers is varied in every deep learning model used. The performance of the model was calculated for a minimum of one hidden layer and a maximum of five hidden layers. The results can be seen in Table 1. It is inferred that the accuracy increases with the increase in the number of hidden layers. There are no hidden layers in random forest.

Table 1 Varying the number of hidden layers

Number of hidden layers	Fully connected model	PCA followed by fully connected model	RFE followed by fully connected model
1	98.36	98.41	98.61
2	98.18	98.36	98.61
3	98.27	98.44	98.70
4	98.61	98.87	98.87
5	98.40	98.53	99.05

Table 2 Varying batch size

Batch size	Fully connected model	PCA followed by fully connected model	RFE followed by fully connected model
32	98.61	98.36	98.87
64	98.44	98.87	99.05
128	98.44	98.44	98.79
256	98.44	98.36	98.87

5.2 Varying Batch Size

Both higher and lower values of batch size result in lower accuracy, and therefore, an optimum value of batch size was chosen and is set to capture the best results. The results recorded for batch sizes 32, 64, 128, and 256 are recorded in Table 2. The accuracy lies between 98.18 and 99.05%. Random forest is independent of batch size.

5.3 Varying Number of Neurons

Different numbers of neurons are tested for each layer in a model and the best accuracy that was achieved with the optimum number of neurons in each layer is chosen.

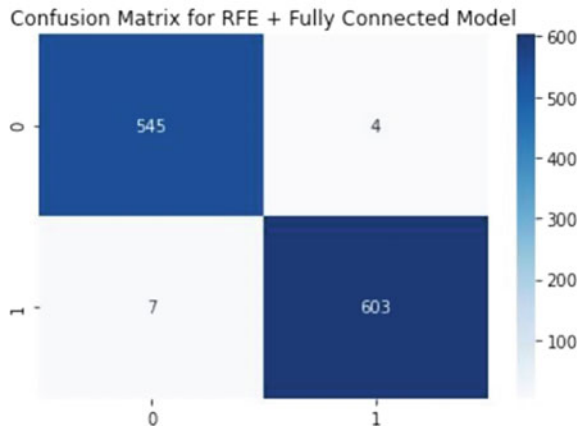
5.4 Comparison of Models

Each model is evaluated using evaluation metrics like accuracy, precision, recall, F1-score, and area under curve (AUC). The comparison of models is given in Table 3 (Fig. 4).

Table 3 Comparison of models

Models	Accuracy	Recall	Precision	F1_score	AUC
Fully connected model	98.61	98.52	98.39	98.44	98.35
Random forest	98.79	98.52	99.17	98.84	98.80
PCA followed by fully connected model	98.87	98.52	98.68	98.60	98.53
PCA followed by random forest	98.69	98.59	98.31	97.09	97.04
RFE followed by fully connected model	99.05	98.85	99.34	99.09	99.06
RFE followed by random forest	98.79	98.68	99.01	98.85	98.79

Fig. 4 Confusion matrix for RFE followed by fully connected model



RFE followed by fully connected model has shown a slightly better performance giving an accuracy of 99.05% and an AUC of 99.06%. We have captured better results using numerical data for identifying exoplanets rather than using light curves. Disposition score, orbital period, planetary radius, and stellar effective temperature are some of the features that are highly correlated with the identification process of these exoplanets.

6 Conclusion

In this study, we have built various models to predict exoplanet candidates. RFE followed by fully connected model has shown the best results with an accuracy of 99.05% and an AUC of 99.06%. Results of other evaluation metrics can be seen in Table 3.

Other models like PCA followed by Fully connected model and RFE followed by random forest were also performing well, capturing accuracy of 98.87% and 98.79%,

respectively. These models can be worked upon in the future by other researchers to improve the accuracy by manually removing outliers.

This study has shown that machine learning and deep learning can significantly capture exoplanet candidates more efficiently by selecting the best features and thus, can help in contributing to an efficient automated process of finding these exoplanet candidates.

References

1. Pegasi b. Available at <https://www.nasa.gov/feature/jpl/20-intriguing-exoplanets>
2. Credits—NASA/Ames/JPL-Caltech/T Pyle. Available at <https://www.nasa.gov/kepler/mission-timeline>
3. Yu L, Vanderburg A, Huang C, Shallue CJ, Crossfield IJM, Scott Gaudi B, Daylan T, Dattilo A, Armstrong DJ, Ricker GR, Vanderspek RK, Latham DW, Seager S, Dittmann J, Doty JP, Glidden A, Quinn SN (2019) Identifying exoplanets with deep learning. III. Automated triage and vetting of TESS candidates. *Astron J* 158(1):1–15
4. Shallue CJ, Vanderburg A (2018) Identifying exoplanets with deep learning: a five-planet resonant chain around Kepler-80 and an eighth planet around Kepler-90. *Astron J* 155(2):1–21
5. Beerer IM (2020) Deep learning approach to finding extra-solar planets from TESS light curves. Stanford University report, CS230
6. Jeevarathinam N (2020) Detection of exoplanets system in Kepler light curves using deep learning. Masters thesis, Dublin, National College of Ireland
7. Pearson KA, Palafox L, Griffith CA (2018) Searching for exoplanets using artificial intelligence. *Monthly Notices R Astron Soc* 474(1):478–491. <https://doi.org/10.1093/mnras/stx2761>
8. Yip KH, Nikolaou N, Coronica P, Tsiaras A, Edwards B, Changeat Q, Morvan M, Biller B, Hinkley S, Salmond J, Archer M, Sumption P, Choquet E, Soummer R, Pueyo L, Waldmann IP (2019) Pushing the limits of exoplanet discovery via direct imaging with deep learning. In: Machine learning and knowledge discovery in databases: European conference, ECML PKDD 2019, Würzburg, Germany, 16–20 Sept 2019, Proceedings, Part III Sep 2019, pp 322–338. https://doi.org/10.1007/978-3-030-46133-1_20
9. Dattilo A, Vanderburg A, Shallue CJ, Mayo AW, Berlind P, Bieryla A, Calkins ML, Esquerdo GA, Everett ME, Howell SB, Latham DW, Scott NJ, Yu L (2019) Identifying exoplanets with deep learning. II. Two new super-earths uncovered by a neural network in K2 data. *Astron J* 157(5)
10. Mathur S, Sizon S, Goel N (2021) Identifying exoplanets using deep learning and predicting their likelihood of habitability. In: Patnaik S, Yang XS, Sethi I (eds) *Advances in machine learning and computing intelligence. Algorithms for intelligent systems*. Springer, Singapore. https://doi.org/10.1007/978-981-15-5243-4_34
11. Nasa Exoplanet Archive. Dataset available at <https://exoplanetarchive.ipac.caltech.edu/cgi-bin/TblView/nph-tblView?app=ExoTbIs&config=koi>

Feature Selection in Corporate Bankruptcy Prediction Using ML Techniques: A Systematic Literature Review



Mohd Raagib Shakeel, Taufeeque Ahmad Siddiqui, and Shahzad Alam

Abstract Systematic literature reviews have long been acknowledged as an important tool for evaluating knowledge theoretically, bridging gaps, and establishing the groundwork for future endeavours. However, in the realm of corporate bankruptcy prediction, systematic literature reviews are few and far between, and none exist for key attributes such as feature selection. This study takes a holistic approach to systematically and qualitatively analyse the feature selection techniques employed by researchers, based on the protocol of the preferred reporting items for systematic reviews and meta-analyses (PRISMA). Furthermore, our goal is to bring together all of the key characteristics, such as feature selection techniques, different machine learning techniques, evaluation criteria, evaluation criteria outcomes, and the research end result, in one place for any interested researcher who wants to do an analysis or view a summary. For the period 2015–2021, the Scopus database and the reference lists of the selected research papers were used to extract a total of 36 articles. The result indicates the split of feature selection approaches under the heading “filter, wrapper, and embedded,” which were employed independently and in combination by various authors. The filter approach is clearly the most preferred amongst researchers due to its simple structure and superior results. Furthermore, we discovered that to acquire the most important variables, multiple feature selection methods were used within the feature selection techniques categories.

M. R. Shakeel (✉) · T. A. Siddiqui
Department of Management Studies, Faculty of Management Studies, Jamia Millia Islamia,
New Delhi 110025, India
e-mail: raagib14@gmail.com

T. A. Siddiqui
e-mail: taufeeque@gmail.com

S. Alam
Department of Computer Engineering, Faculty of Engineering and Technology, Jamia Millia
Islamia, New Delhi 110025, India
e-mail: shahzad5alam@gmail.com

Keywords Feature selection · Variable selection · Corporate Bankruptcy prediction · Financial distress prediction · PRISMA · Systematic review · Machine learning

1 Introduction

The structured techniques involved in the amalgamation of results, which give impartial searches with superior standard of coherency, are the key characteristics that set systematic reviews apart from other types of reviews [1]. Under the names of “company failure prediction”, “financial distress prediction”, “business failure prediction”, financial literature has long been interested in and investigated corporate bankruptcy prediction. Because business failure does not occur overnight, bankruptcy prediction modelling has evolved through time [2, 3].

Feature selection is one of the crucial factors that has a substantial influence on the success of a prediction model. In research studies, the term “feature selection” has been used interchangeably with terms like “variable selection,” “attribute selection,” and “variable subset selection”. The meaningless data in the dataset makes it difficult for machine learning algorithms to distinguish between important and irrelevant data, which has an effect on the final prediction model’s performance. In such a situation, feature selection comes to the rescue, distinguishing between useless and meaningful data, speeding up the process, improving data quality, and allowing the model to forecast more accurately. Filter, wrapper, and embedded approaches are three common feature selection techniques. Filter-based technique employs a ranking criterion and subset selection. The wrapper technique, on the other hand, uses a predetermined predictor to improve the quality of the selected features. In embedded methods, the qualities of filter and wrapper methods are integrated [4].

Although there have been several literature studies and systematic reviews in the field of corporate failure or bankruptcy prediction [5–8], systematic or general reviews of significant factors such as feature selection and its impact, which aid in not only obtaining maximum accuracy but also bringing in data quality, are rare. As far as we are aware, [9] is currently the only research that has conducted a comprehensive evaluation and analysed the performance of eight leading tools for forecasting bankruptcy using 13 essential criteria. However, their research has certain limitations. Their research is limited to journal articles published between 2010 and 2015. Rather than research that employed feature selection techniques to improve prediction outcomes, their major focus was on picking the best AI approach performance out of eight based on 13 essential criteria.

By adhering to PRISMA [1] protocol requirements, our research was designed to provide a holistic approach to systematically analyse feature selection techniques, different machine learning techniques, evaluation criteria, evaluation criteria outcomes, and the research’s final conclusion. Our goal is to bring all of the key traits together for any interested researcher who wants to do an analysis or view a summary in one place. PRISMA [1] is a methodology for doing a systematic review

that contains a 27-item detailed list and a 4-stage step-by-step diagram to improve lucidity, uniformity, and precision across reviews.

2 Methods

The study's inclusion criteria were carefully chosen to assist in a proper assessment and promising results. We used the phrase "bankruptcy prediction" as a lead. All papers that used insolvency prediction, financial distress prediction, financial default prediction, or financial failure prediction as synonyms for our string were included. Only final research articles published between 2015 and 2021 were considered. Only peer-reviewed journal publications were chosen to improve the study's credibility and quality. We excluded research articles that didn't use machine learning approaches to predict bankruptcy. We did not include bank insolvency and credit score predictions in our research. In total, 36 research articles were used in our study. We used the Scopus database and the reference list of the selected research papers to retrieve articles. We decided against using Google Scholar as a database because of some of its shortcomings. The lack of desired filters and limitless outcomes make it difficult to manage. We chose to select research papers from the reference list of selected published articles because it provided us with more relevant and desirable articles, thus improving the quality of our research. Items pertaining to effect measurement, synthesis methods, consistency measurements, assessing risk of bias owing to missing results, confidence evaluation, and meta-analysis results (items 12–15 and 19–22) were excluded because meta-analysis was not performed in this study.

3 Results

The procedure for selecting studies has been explained in Fig. 1. A total of 36 research papers were examined and analysed qualitatively in our study. Figure 2 depicts the number of research articles published over the years that used feature selection techniques. It was discovered that just three papers using feature selection techniques were published in 2015, and this number continued to rise in subsequent years. Tables 1-a, 1-b, 1-c, 1-d shows the names of the authors, the year of publication, the feature selection and machine learning techniques used, the researchers' evaluation metrics, the outcome of the evaluation criteria, and the study's final result.

4 Synthesis

The total number of authors who used various feature selection techniques is depicted in Fig. 3. The pie chart reveals that 36% of the authors used the filter method, 34%

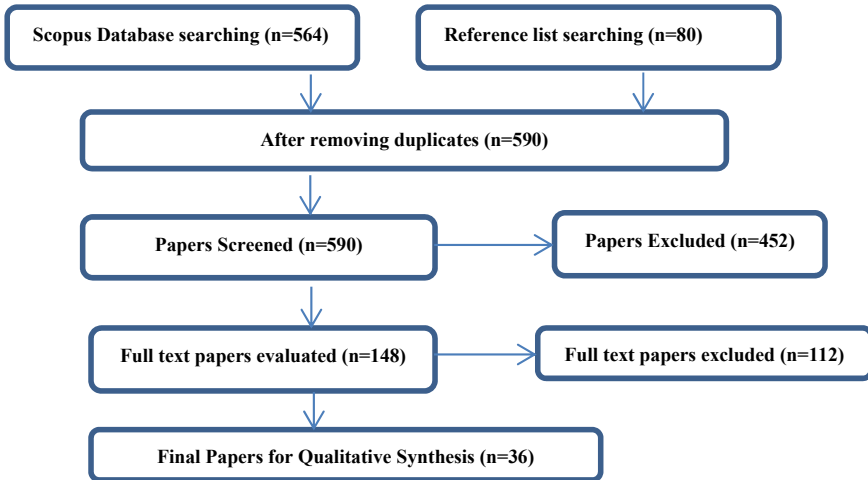


Fig. 1 Flowchart of the selection process

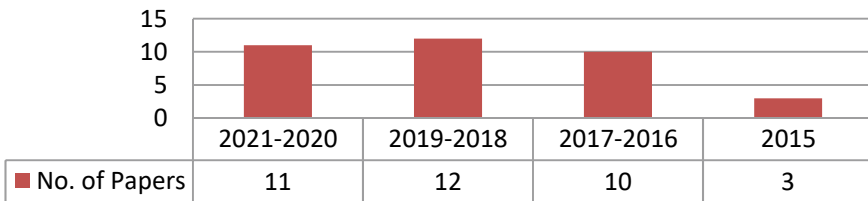


Fig. 2 Number of research papers using feature selection techniques published over the years

used the wrapper method, 14% used both the filter and wrapper techniques, and the remaining 16% applied the embedded method of feature selection technique out of a total of 36 authors. Figure 4 shows the total number of feature selection methods employed by authors that fall under the headings of head filter, wrapper, or embedded.

5 Discussion and Conclusion

The protocol and procedure used in systematic reviews ensure that the study is meticulous, resulting in actionable findings and discussion. However, systematic literature reviews with a protocol are rare in this field. To ensure consistency and precision amongst review studies, we used the PRISMA protocol items [1] in our study. Researchers’ usage of feature selection techniques has increased with time, from 3% in 2015 to 30.5% in 2020–2021. The filter approach was used by 36% of the researchers, followed by the wrapper method, which was used by 34% of the researchers. Furthermore, for feature selection, all of the authors applied a total of

Table 1-a Summary of Reviewed articles with year, author, feature selection method, ML techniques, evaluation criteria and their outcome and study's final result

Author/year	Feature selection method	ML techniques used	Evaluation criteria	Evaluation criteria result	Final result
Jabeur et al. [10]	Embedded: Catboost	Discriminant analysis, LR, SVM, Neural Network, RF, GBM, Deep neural network, XGBOOST, CATBOOST	Accuracy and AUC	Avg Accuracy-82.90% Avg AUC value -87.6%	Catboost outperformed all other models in all years horizon(1-3) in terms of accuracy and AUROC
Abid et al. [11]	Wrapper: HVS-AUC(Heuristic for variable selection procedure, Forward-AUC procedure	Forward-AUC, Linear Discriminant Analysis, HVS-AUC	Accuracy, Type I and Type II error, AUC	AUC- HVS-AUC-88%, Forward-AUC-85% Accuracy—HVS-AUC-84.5% Forward-AUC-85.5%, Type I error—HVS-AUC-14% Forward-AUC-14%, Type II error -HVS-AUC-17% Forward-AUC-15%	HVS-AUC and Forward-AUC is better than LDA
Tsai et al. [12]	Filter: T-test, PCA Wrapper: GA	Bagging and boosting based LR, SVM, ANN, DT	AUC and type II errors	AUC-AP + Bagging DT-93.4% Type II error -PCA-AP + LR- 28%	AP + bagging DT attained the highest AUC and PCA-AP + LR attained the lowest type II error
Noviantoro and Huang [13]	Wrapper: Forward selection	DT, RF, K-NN classifier, Native Bayes, LR, Neural net, Deep learning, SVM	Accuracy, F-score, AUC	Random forest Accuracy-97.16%,F-score-98.55%,AUC-90.5%	Highest Accuracy- RF Highest F-score -RF AUROC- RF
Tabbakh et al. [14]	Filter: Select Percentile,PCA Wrapper: Sequential feature selection	Gradient boosting, DT, Balanced bagging, RF, SVM, and ADABOOST	F1-Score, Accuracy, Recall, RMSE	Adaboost classifier + SFS Accuracy-98.8%,F1-score-98.8% Recall-99.3%,RMSE- 10.8%	Adaboost classifier with SFS feature selection provides better result as compared to other combination (all ML + other 2 Feature selection method)

(continued)

Table 1-a (continued)

Author/year	Feature selection method	ML techniques used	Evaluation criteria	Evaluation criteria result	Final result
Zeng et al. [15]	Filter: Sparse PCA	OF-SVM, PCA-SVM, LDA-KNN, KPA-SVM, KDA-KNN, SPCA-SVM, GSPCA-SVM	Accuracy, Precision, Recall, and F1 value	DATASET 1-GSPCA-SVM-Accuracy-81.94%,Precision-81.08%,Recall-83.33%,F-1 value-82.19% DATASET 2—GSPCA-SVM Accuracy-81.94%,F-1 value- 76% PCA-SVM Precision—94.59% KPCA-SVM Recall-83.05%	Dataset 1- GSPCA-SVM is superior to other 6 algorithm, Dataset 2-Accuracy and F1 value-GSPCA-SVM is superior Precision value—PCA-SVM Recall value -KPCA-SVM
Yan et al. [16]	Embedded: Lasso penalty	For financial ratios only Neural Network, DT, Lasso-SVM, Lasso logistic in one year time frame VS LSVMDLmodel, LLDL Financial ratios plus macroeconomic indicators Neural Network, Decision Tree, Lasso SVM, Lasso-logistic models vs LSVMDL and LLDL models	AUC, G-mean, Kolmogorov–Smirnov (KS) statistics	(1) Financial ratios Period t-3- NN- AUC-93.56% and KS -88%.DT G-mean-88.74%.Period t-4—NN AUC-92.24%.DT G-mean- 90.87%. DT and LSVMDL KS-86% Period t-5—LLDL AUC-91.52%,DT G-mean-90.87%,DT and LSVMDL KS- 86% 2) Financial ratios + macroeconomic variables Period t-3 LLDL AUC-95.08%,LSVMDL G-mean-93.98%, LSVMDL KS-92%.Period t-4 LLDL AUC-95.08%,LSVMDL-G-mean-93.98%.DT and LSVMDL KS-92%.Period t-5 LLDL AUC-95.08%,LSVMDL KS-92%	(1) Financial ratios only Period t-3- AUC and KS -NN outperforms, G-mean-DT outperforms, Period t-4- AUC-NN, G-mean-DT, KS- DT AND LSVMDL, Period t-5 AUC-LLDL, G-mean- DT, KS—DT AND LSVMDL 2) Financial ratios + macroeconomic variables Period t-3-AUC- LLDL, G-mean- LSVMDL, KS -LSVMDL, Period t-4-AUC LLDL, G-mean -LSVMDL, KS- DT AND LSVMDL, Period t-5—AUC- LLDL, G-mean- LSVMDL, KS-LSVMDL

(continued)

Table 1-a (continued)

Author/year	Feature selection method	ML techniques used	Evaluation criteria	Evaluation criteria result	Final result
Marso and Merouani [17]	Wrapper: RF-RFE algorithm	MLP linked to ABC algorithm called ABCNN, MDA, and MLP associated to the Brodyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm called BPNN	Accuracy, AUC, Type I error, Type II error	One year before bankruptcy ABCNN Accuracy-92.04%, type I-6.81% and type II error -9.19%, AUC-96.26%	ABCNN outperforms other 2 models when data is one year before bankruptcy
Liang et al. [18]	Filter: T-test, SDA Wrapper: SLR	Stacking ensembles	DET curve and misclassification cost	Stacking ensembles misclassification cost > 7	Without feature selection—Combining FR and CGI indicator does not enhance the prediction performance of stacking ensembles. With Feature selection selected FRs and CGIs can make stacking ensembles outmatch the baseline model of using FR alone and OR ensembles when the cost ratios > 7

Table 1-b Summary of Reviewed articles with year, author, feature selection method, ML techniques, evaluation criteria and their outcome and study's final result

Author/year	Feature selection method	ML techniques used	Evaluation criteria	Evaluation criteria result	Final result
Cao et al. [19]	Embedded: Lasso	LR, DT, SVM, DNN (deep neural network) and Bayesian network (BN)	Accuracy and AUC	DNN(50,30,20) Avg of 12 quarters Group 1 variables Accuracy-80.09%,AUC-84.96% Group 2 variables Accuracy-69.32%, AUC-76.94%	DNN(50,30,20) with 3 hidden layer FOR Group 1 and 2 variables outperforms all other models
Altman et al. [20]	Filter: R-squared Wrapper: SLR	Decision Tree, GB, Logistic Regression, Neural network with multi-layer perceptron (NN),SVM	AUC	LR-All variable Avg AUC for all horizon (year 1-10)-82.30%	LR model with all the variables outmatched the LR models where wrapper or filter variable selection methods were used. For both LR and NN, wrapper method (SLR) gave higher AUCs than the filter method (the R-squared method) for all the horizons. The greater the difference in AUCs between the selection procedures, the longer the horizon. SVM beat the filter versions of LR and NN on average
Tang et al. [21]	Wrapper: Minimal-redundancy-maximal-relevance	EPNN, PNN and MLP; Other classification method KNN, Radial basis Function, RF, Decision Tree, SVM, and Discriminant analysis	Accuracy, sensitivity, specificity, and AUC	EPNN Avg accuracy-99.57%, sensitivity-99.76%, specificity-97.75%, AUC-99.84%	On both the Qualitative Bankruptcy dataset and the Distress dataset, the proposed EPNN outperforms the PNN and MLP in terms of average accuracy and AUC

(continued)

Table 1-b (continued)

Author/year	Feature selection method	ML techniques used	Evaluation criteria	Evaluation criteria result	Final result
Farooq and Qamar [22]	Wrapper: Wrapper subset approach	DTNB hybrid model, LMT and A2DE (alternative and dependence estimator) and other 21 techniques	TPR, FPR, F-measure and AUROC	DTNB TPR-81.7%, FPR-61%, F-measure-818%, ROC area-95%	The DTNB hybrid model is the most accurate in predicting class variables
Devi and Radhika [23]	Filter: Multivariate Gaussian distribution	Naive Bayes classifier	Accuracy	Naive Bayes classifier Accuracy ratio-78.4%	With a ratio of 0.784, the Naive Bayes classifier approach combined with the Hadoop big data tool is successful in predicting bankruptcy
Son et al. [24]	Embedded: Xgboost, LightGBM	LR, RF, XGBoost, LightGBM, ANN	Accuracy, TPR, FPR, and AUROC	1) full dataset (ALL)-XGBoost and LightGBM AUC-83% 2) externally audited(+EX) XGBoost AUC-88% 3)Not externally audited (NOEX) LightGBM AUC-83%	Machine learning techniques such as XGBoost, LightGBM, and Artificial Neural Network outperformed the classic logistic regression model in terms of accuracy
Faris et al. [25]	Filter: CFS, Relief, Information Gain, CAE, Correlation Attribute Evaluator	Basic classifier-J48, Random tree, Rep tree, KNN, NAIVE Bayes, MLP, Ensemble classifier-Adaboost, bagging, random forest, Rotation forest, DECORATE	Accuracy, G-mean, Type I error, Type II error and AUROC	This composite model has achieved an accuracy of 98.3%, Type I error of 0.6%, Type II error of 45%, and G-mean of 0.73	With highly imbalanced datasets, the SMOTE oversampling approach combined with an AdaBoost ensemble method using the REP tree as a fundamental classifier can produce promising results
Hu [26]	Filter: Grey relational analysis	Grey prediction models, MLP, SVM, C4.5 decision tree and LR	Accuracy	Grey prediction model Avg accuracy (year 1-3)—80.25%	Grey prediction models outperformed all other classifiers from year 1 to year 3

(continued)

Table 1-b (continued)

Author/year	Feature selection method	ML techniques used	Evaluation criteria	Evaluation criteria result	Final result
Valencia et al. [27]	Embedded: Generalized additive model selection	RF:SVM with radial kernel, GAM with embedded variable selection, GAM without variable selection, SVM with linear kernel, LR with lasso type penalization, LDA	AUC	RF AUC-85.76%	RF outperformed all other models, however GAM with embedded variable selection was pretty close to RF
Joshi et al. [28]	Wrapper: GA	Random forest, DT	Accuracy	RF-Accuracy-100%	Random forest outperforms
Uthayakumar et al. [29]	Wrapper: ACO, GA, PSO, GWO	SVM, bagging, Decision Tree, LR, RBF, random forest (RF), MLP, AdaBoost and Olex-GA, ACO-DC (data classification) algorithm	Accuracy, FPR, FNR, sensitivity, specificity, Accuracy, F-score, Discriminant power (DP), Mathew correlation coefficient (MCC), FDR, FOR and kappa value	Data set ACO-FS-ACO-DC FPR-0.09%,FNR-2.36%,Sensitivity-97.63%,specificity-99.90%,accuracy-97.69%,F-score-98.80%,DP-4.11,MCC-0.71,FDR-0.02,FOR-47.84%,KAPPA-67.45%	Feature selection result ACO feature selection method is superior to other FS method such as GA, PSO, and Grey Wolf Optimization (GWO). Classification result

Table 1-c Summary of Reviewed articles with year, author, feature selection method, ML techniques, evaluation criteria and their outcome and study's final result

Author/year	Feature selection method	ML techniques used	Evaluation criteria	Evaluation criteria result	Final result
Lin et al. [30]	Filter: Information gain Wrapper: GA	LR, NAIVE BAYES, BACK PROPAGATION NN, C4.5 DT, SVM, KNN	Type I error	Single classifier Australian dataset GA + NB-Type I error-4.97% German dataset GA + SVM-Type I error-8.58% Taiwanese dataset GA + SVM Type I error 0% Ensemble classifier Australian dataset GA + NB-bagging-5.24% German dataset GA + SVM-bagging 8.01%, Taiwanese dataset GA + SVM-bagging/boosting-0%	Single classifier Lowest error rate obtained by wrapper FS (GA) with NB and SVM Ensemble Classifier lowest error rate obtained by GA with NB-bagging, SVM-bagging, SVM-boosting
Le et al. [31]	Filter: Principal component algorithm	RF, DT, MLP, and SVM	AUC	Without oversampling RF AUC-82.4% With oversampling SMOTE + ENN + RF AUC-84.2%	Without oversampling techniques RF is the best classifier With oversampling Technique SMOTE + ENN and Random Forest is the best oversampling technique and classifier model
Veganzones and severin [32]	Filter: Correlation, Fisher-F test, chi square test Wrapper: Backward search procedure	LDA, LR, NN, SVM, RF	sensitivity, specificity, G-mean, AUC	Avg of all imbalanced datasets SVM and RF Avg sensitivity-60.46% and 53.04%, avg specificity-90.98% and 90.9%. SVM avg G-mean-72.2%. SVM avg AUC-83.68%	For unbalanced datasets, SVM and RF appear to be more appropriate algorithms

(continued)

Table 1-c (continued)

Author/year	Feature selection method	ML techniques used	Evaluation criteria	Evaluation criteria result	Final result
Garcia et al. [33]	Filter: ReliefF, Pearson's correlation	FLD, LDC, a support vector machine (SVM) with a linear kernel and the soft-margin constant C = 1.0, and the logistic regression (logit) model	Accuracy, TPR, FPR	RC-DS (Avg of all % prototypes) Accuracy-FLD-98.48%, LDC-98.08%.LOGIT-96.08%.SVM-98.88%, TPR-FLD-96.64%.LDC-97.01%.LOGIT-94.02%.SVM-98.36% FPR-FLD-99.58%.LDC-98.88%.LOGIT 97.63%.SVM -99.30%	In terms of accuracy, true-positive rate, and true-negative rate, all of the linear models performed better on the dissimilarity space (DS) than on the feature space (FS)
Jones [34]	Embedded: Gradient boosting	TreeNet (Gradient boosting model),LR	Average log-likelihood, AUC, Accuracy, Lift	TreeNet (Gradient boosting) (All data) Avg Log Likelihood-04.190%.ROC-99.68%, lift-1.134,accuracy-95.78%.TreeNet (One Year Prior to Bankruptcy) Avg Log Likelihood-6.297%.ROC-99.12%, lift-1.136,accuracy-93.87%.TreeNet (Three Years Prior to Bankruptcy) Avg Log Likelihood-8.59%.ROC-98.28%, lift-1.136,accuracy-91.25%	In a high-dimensional (and multi-dimensional) framework, which is analysed using TreeNet, business bankruptcy is better predicted. A gradient boosting model based on all available predictors outperforms no single dimension
Kim et al. [35]	Filter: Two sample t-test	(Data depth) DD-SVM, LOGIT, MDA, ANN	Accuracy, Type I error and Type II error	DD-SVM(Year 1-10) Avg Accuracy-82.75%.Avg Type I-13.79%. Avg Type II-8.965%	DD-SVM outperforms all other models from 1 to 6 years

(continued)

Table 1-c (continued)

Author/year	Feature selection method	ML techniques used	Evaluation criteria	Evaluation criteria result	Final result
Fallahpour et al. [36]	Wrapper: Sequential floating forward	Feature selection + classifier SFFS + SVM, SFS + SVM, Artificial bee colony (ABC) + SVM, GA + SVM, PCA + SVM IG + SVM, RELIF + SVM, SVM	Specificity, Sensitivity and Average accuracy	SFFS + SVM (all 3 years) Avg Specificity-93.69%, avg Sensitivity-96.29%, avg accuracy-94.99%	In all cases, the accuracy attained by SFFS + SVM was larger than that of the other approaches for the three years t, t-1, and t-2
Zelenkov et al. [37]	Wrapper: Genetic Algorithm	TSCM, 5 bagging methods (with DT, kNN, SVM, NB and LR base classifiers), 3 AdaBoost methods (with DT, SVM, and NB base classifiers), methods of stochastic design of ensembles of the decision trees (Random Forest and Extra Trees), gradient boosting and voting classifier	Accuracy, precision, and recall	TSCM Accuracy-93.4%, precision-91%, recall-95.3%	The suggested TSCM approach has accuracy comparable to the top methods. Precision and recall ratios suggest that TSCM is precise

Table 1-d Summary of reviewed articles with year, author, feature selection method, ML techniques, evaluation criteria and their outcome and study's final result

Author/year	Feature selection method	ML techniques used	Evaluation criteria	Evaluation criteria result	Final result
Kim et al. [38]	Wrapper: Genetic Algorithm	GA-ANN, GA-ANN_CBEUS, ANN_None, ANN_RUS, GA-ANN_RUS, GA-ANN_EUS	Sensitivity, specificity, G-mean, AUC, H-measure	GA-ANN_CBEUS Sensitivity-87.407%, specificity-75.556%, g-mean-84.259%, Auroc-88.778%, H-measure-56.158%	Other models were outscored by GA-ANN CBEUS
Zhou et al. [39]	Filter: PCA	KELM, ELM, SVM, RF, PSOFKNN	Accuracy, Type I error, Type II error and AUC	KELM Avg Accuracy-82.50%, Avg Type I-15.15%, Avg Type II-19.48%, AUC-82.68%	The KELM model outperforms the five other advanced bankruptcy prediction models
Liang et al. [40]	Filter: SDA, T-test Wrapper: GA, RFE, SLR	SVM, KNN, NAIVE BAYES, CART, MLP	Accuracy, type I error and type II error	SDA + SVM (FC variables (FR + CGI) Avg accuracy-81.5%, Type I error-16.3%, Type II error-20.8%	Taiwan dataset When compared to a model based only on FRs, combining FRs and CGIs can enhance model performance. The optimal combination is SDA + SVM. Chinese Market The results produced by combining FRs and CGIs are no better than those obtained by utilising FRs alone
Hosaka and Takata [41]	Embedded: RealAdaBoost	RealAdaBoost	Leave-one-out cross-validation	Identification rate-89.3%	With RealAdaBoost identification rate of 0.893 was achieved with four financial ratios
Karan and Kumar [42]	Filter: Pearson correlation	DT, SVM, LR	Accuracy, AUC, Precision	LR Accuracy -94.73%, AUC-0.9697, precision-93.97%	Logistic Regression is chosen as it has outperformed other machine learning algorithms
Nagaraj and Sridhar [43]	Filter: Pearson correlation, IG	LR, RF, NAIVE BAYES, NEURAL NETWORK, (radial basis function) RBF-based SVM	Accuracy, TPR, FPR, Precision	RBF-Based SVM Accuracy-99.6%, TPR-0.996, FPR-0.004, Precision-99%	RBF-based SVM outperforms other ML techniques

(continued)

Table 1-d (continued)

Author/year	Feature selection method	ML techniques used	Evaluation criteria	Evaluation criteria result	Final result
Dellepiane et al. [44]	Filter: Correlation based subset Evaluation	SVM, LOGIT, LDA	Accuracy	SVM + Correlation based subset Accuracy-around 90%for both one year and 2 years ahead prediction	SVM + Correlation based subset evaluation method improves the accuracy as compared to other methods
Tian et al. [45]	Embedded: Lasso	LASSO selected reduced form model and CHS(2008) model	In sample(1980-2009): AUC, AUC Out of sample(2003-2009): Accuracy and AUC, Brier score	In Sample Lasso selected reduced model AUC-14683, AUC-71.1%, Out of Sample Lasso selected reduced model Accuracy ratio-68.2%,AUC-84.1%,Brier score-40.8%	The LASSO-selected models have superior in-sample fits in terms of AIC than the CHS (2008) model for all prediction horizons. The performance of the LASSO selected models in terms of in-sample AUC is equivalent to, if not better than, that of the CHS (2008) model. Except for the one-month-ahead prediction, the LASSO-selected models continuously outperform the CHS model in terms of accuracy ratio and AUC for out-of-sample forecasting

DA discriminant analysis, LR logistic regression, ANFIS adaptive neuro-fuzzy inference system, SVM support vector machine, GASVM-libsvm GASVM with libsvm model, NN neural network, RF random forest, GP-OLS genetic programming with orthogonal least square, GBM gradient boosting machine, AUC area under curve, GA-SVM-lsrbf GASVM with rbf kernel func., LDA linear discriminant analysis, TNR true negative rate, FGS Broyden-Fletcher-Goldfarb-shanno, ANN artificial neural network, DT decision tree, FNR false negative rate, MLP multilayer perceptron, ISCBR instance selection using GA for CBR, MDA multiple discriminant analysis, GA-ANN_CBEUS cluster based evolutionary undersampling, FR financial ratios, DP discriminant power, SFF sequential floating forward, PLS partial least square, CGI corporate governance indicators, PNN pruning neural network, GEP gene expression programming, RBF radial basis function, TPR true positive rate, FS feature space, SDA stepwise discriminant analysis, FPR false positive rate, KNN K-nearest neighbour, AW A wrapper, FSCGACA fitness-scaling chaotic GACA, CART classification and regression trees, SOM self-organizing map, SD standard deviation, LSVMDL lasso-svm distributed lag, FS feature selection, FISCBR feature and instance selection using GA for CBR, CHAID chi-square automatic interaction detection, LVQ learning vector quantization, GP-SA genetic programming with simulated annealing, CBR case based reasoning, SFFS sequential floating forward selection, MCC mathew correlation coefficient, NPV negative predictive value, FWCBR feature weighting using FA for CBR, FDR false discovery rate, PCA principal component analysis, ANN_RUS ANN without sampling, GA genetic algorithm, RF-RFE random forest-recursive feature elimination, GRNN generalised regression NN

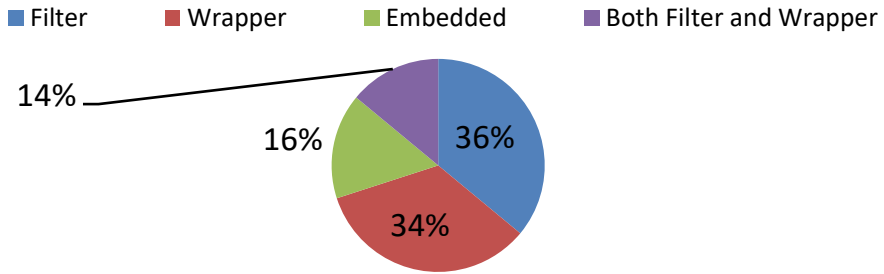
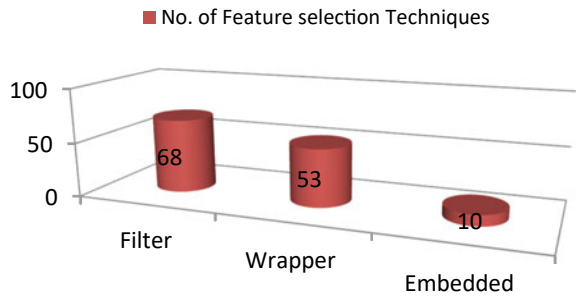


Fig. 3 Authors using different types of feature selection techniques

Fig. 4 Total number of feature selection techniques used by all the authors



68 filter methods. It is evident that researchers have preferred the filter approach because of its simple structure and superior outcomes. Considering the limitations of our study, we only used the Scopus database to extract data, with the rest coming from the reference lists of selected papers. As a result, in the future, more database searches will be possible. Because we eliminated bank insolvency from our research, there is potential for banks to be added. We did not pursue quantitative data synthesis, which can be pursued in future research. We recommend future researchers perform systematic reviews using the PRISMA protocol since it ensures lucidity, uniformity, integrity, and precision.

References

1. Liberati A, Altman D, Tetzlaff J, Mulrow C, Gøtzsche P, Ioannidis J, Moher D (2009) The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *J Clin Epidemiol* 62:e1–34. <https://doi.org/10.1016/j.jclinepi.2009.06.006>
2. Alam S et al (2021) Dual secure robust watermarking scheme based on hybrid optimization algorithm for image security. *Pers Ubiquit Comput* 1–13. <https://doi.org/10.1007/s00779-021-01597-2>
3. Alam S et al (2021) A lossless digital watermarking scheme based on a new 3D-hyper chaotic map. *Int J Electron Bus* 16(1):88–102. <https://doi.org/10.1504/IJEB.2021.112769>

4. Kumar V, Minz S (2014) Feature selection: a literature review. *Smart Comput Rev* 4(3):211–229. <https://doi.org/10.6029/smarterc.2014.03.007>
5. Appiah KO, Chizema A, Arthur J (2015) Predicting corporate failure: a systematic literature review of methodological issues. *Int J Law Manage* 57(5):461–485. <https://doi.org/10.1108/IJLMA-04-2014-0032>
6. Kezelj T, Guenbichler R (2021) A systematic literature review on corporate insolvency prevention using artificial intelligence algorithms. *J Strategic Innov Sustain* 16(4):12–21. <https://doi.org/10.33423/jsis.v16i4.4618>
7. Shi Y, Li X (2019) An overview of bankruptcy prediction models for corporate firms: a systematic Literature review. *Intangible Capital* 15(2):114–127. <https://doi.org/10.3926/ic.1354>
8. Alam et al (2020) A novel hybrid watermarking scheme with image authentication based on frequency domain, 2- level SVD using chaotic map. *EAI Endorsed Trans Energy Web*. <https://doi.org/10.4108/eai.13-7-2018.165512>
9. Alaka HA, Oyedele L, Owolabi HA, Kumar V, Ajayi SO, Akinade O, Bilal M (2018) Systematic review of bankruptcy prediction models: towards a framework for tool selection. *Expert Syst Appl* 94:164–184. <https://doi.org/10.1016/j.eswa.2017.10.040>
10. Jabeur SB, Gharib C, Meftah-Wali S, Arfi WB (2021) CatBoost model and artificial intelligence techniques for corporate failure prediction. *Technol Forecasting Soc Change* 166(C):120658. <https://doi.org/10.1016/j.techfore.2021.120658>
11. Abid I, Ayadi R, Guesmi K, Mkaouar F (2021) A new approach to deal with variable selection in neural networks: an application to bankruptcy prediction. *Ann Oper Res*. <https://doi.org/10.1007/s10479-021-04236-4>
12. Tsai C-F, Sue K-L, Hu Y-H, Chiu A (2021) Combining feature selection, instance selection, and ensemble classification techniques for improved financial distressed prediction. *J Bus Res* 130(C):200–209. <https://doi.org/10.1016/j.knosys.2014.10.010>
13. Noviantoro T, Huang J-P (2021) Comparing machine learning algorithms to investigate company financial distress. *Rev Bus Account Finance* 1(5):454–479. <https://fortunepublishing.org/index.php/rbaf/article/view/70>
14. Tabbakh A, Rout JK, Sahoo KS, Jhanjhi NZ (2021) Bankruptcy prediction using robust machine learning model. *Turkish J Comput Math Educ* 12(10):3060–3073. <https://doi.org/10.17762/turcomat.v12i10.4957>
15. Zeng S, Li Y, Yang W, Li Y (2020) A financial distress prediction model based on sparse algorithm and support vector machine. In: *Mathematical problems in engineering*, pp 1–11. <https://doi.org/10.1155/2020/5625271>
16. Yan D, Chi G, Lai KK (2020) Financial distress prediction and feature selection in multiple periods by Lassoing unconstrained distributed lag non-linear models. *Mathematics* 8(8):1275. <https://doi.org/10.3390/math8081275>
17. Marso S, Merouani ME (2020) Bankruptcy prediction using hybrid neural networks with artificial bee colony. *Eng Lett* 28(4)
18. Liang D, Tsai CF, Lu HYR, Chang LS (2020) Combining corporate governance indicators with stacking ensembles for financial distress prediction. *J Bus Res* 120(5):137–146. <https://doi.org/10.1016/j.jbusres.2020.07.052>
19. Cao Y et al (2020) A two-stage Bayesian network model for corporate bankruptcy prediction. *Int J Financ Econ* 27(1):455–472. <https://doi.org/10.1002/ijfe.2162>
20. Altman EI, Iwanicz-Drozdowska M, Laitinen EK, Suvas A (2020) A race for long horizon bankruptcy prediction. *Appl Econ* 52(37):4092–4111. <https://doi.org/10.1080/00036846.2020.1730762>
21. Tang Y, Ji J, Zhu Y, Gao S, Zheng T, Todo Y (2019) A differential evolution-oriented pruning neural network model for bankruptcy prediction. In: *Complexity*, pp 1–21. <https://doi.org/10.1155/2019/8682124>
22. Farooq U, Qamar MAJ (2019) Predicting multistage financial distress: reflections on sampling, feature, and model selection criteria. *J Forecast* 38(7):632–648. <https://doi.org/10.1002/for.2588>

23. Devi S, Radhika Y (2019) An analytical approach for bankruptcy prediction using big data and machine learning technique. *J Theor Appl Inf Technol* 97(6):1823–1832
24. Son H, Hyun C, Phan D, Hwang HJ (2019) Data analytic approach for bankruptcy prediction. *Expert Syst Appl* 138. <https://doi.org/10.1016/j.eswa.2019.07.033>
25. Faris H, Manaseer W, Saadeh M, Mora A (2019) Improving financial bankruptcy prediction in a highly imbalanced class distribution using oversampling and ensemble learning: a case from Spanish market. *Progr Artif Intell* 9:31–53. <https://doi.org/10.1007/s13748-019-00197-9>
26. Hu Y-C (2019) A multivariate grey prediction model with grey relational analysis for bankruptcy prediction problems. *Soft Comput* 24:4259–4268. <https://doi.org/10.1007/s00500-019-04191-0>
27. Valencia C, Cabrales S, Garcia L, Ramirez J (2019) Generalised additive model with embedded variable selection for bankruptcy prediction: prediction versus interpretation. *Cogent Econ Finance* 7(1). <https://doi.org/10.1080/23322039.2019.1597956>
28. Joshi S, Ramesh R, Tahsildar S (2018) A bankruptcy prediction model using random forest. In: 2018 second international conference on intelligent computing and control systems (ICICCS), pp 1–6. <https://doi.org/10.1109/ICCONS.2018.8663128>
29. Uthayakumar J, Metawa N, Shankar K, Lakshmanaprabu SK (2018) Financial crisis prediction model using ant colony optimization. *Int J Inf Manage* 50:538–556. <https://doi.org/10.1016/j.ijinfomgt.2018.12.001>
30. Lin WC, Lu YH, Tsai CF (2018) Feature selection in single and ensemble learning based bankruptcy prediction models. *Expert Syst* 36(1). <https://doi.org/10.1111/exsy.12335>
31. Le T, Lee MY, Park J, Baik SW (2018) Oversampling techniques for bankruptcy prediction: novel features from a transaction dataset. *Symmetry* 10(4):79. <https://doi.org/10.3390/sym1004007>
32. Veganzones D, Severin E (2018) An investigation of bankruptcy prediction in imbalanced datasets. *Decis Support Syst* 112:111–124. <https://doi.org/10.1016/j.dss.2018.06.011>
33. Garcia V, Marques AI, Sanchez JS, Dominguez HJO (2017) Dissimilarity-based linear models for corporate bankruptcy prediction. *Comput Econ* 53(3):1019–1031. <https://doi.org/10.1007/s10614-017-9783-4>
34. Jones S (2017) Corporate bankruptcy prediction: a high dimensional analysis. *Rev Acc Stud* 22(3):1366–1422. <https://doi.org/10.1007/s11142-017-9407-1>
35. Kim S, Mun BM, Bae SJ (2017) Data depth based support vector machine for predicting corporate bankruptcy. *Appl Intell* 48(3):791–804. <https://doi.org/10.1007/s10489-017-1011-3>
36. Fallahpour S, Lakvan EN, Zadeh MH (2017) Using an ensemble classifier based on sequential floating forward selection for financial distress prediction problem. *J Retail Consumer Serv* 34:159–167. <https://doi.org/10.1016/j.jretconser.2016.10.002>
37. Zelenkov Y, Fedorova E, Chekrizov D (2017) Two step classification method based on genetic algorithm for bankruptcy prediction. *Expert Syst Appl* 88(C):393–401. <https://doi.org/10.1016/j.eswa.2017.07.025>
38. Kim H-J, Jo N-O, Shin K-S (2016) Optimization of cluster-based evolutionary undersampling for the artificial neural networks in corporate bankruptcy prediction. *Expert Syst Appl* 59:226–234. <https://doi.org/10.1016/j.eswa.2016.04.027>
39. Zhao D, Huang C, Wei Y, Yu F (2016) An effective computational model for bankruptcy prediction using kernel extreme learning machine approach. *Comput Econ* 49(2):325–341. <https://doi.org/10.1007/s10614-016-9562-7>
40. Liang D, Lu C-C, Tsai C-F, Shih G-A (2016) Financial ratios and corporate governance indicators in bankruptcy prediction: a comprehensive study. *Eur J Oper Res* 22(2):561–572. <https://doi.org/10.1016/j.ejor.2016.01.012>
41. Hosaka T, Takata Y (2016) Corporate bankruptcy forecast using RealAdaBoost. *Information* 19(6):2285–2298
42. Karan A, Kumar P (2016) Predicting bankruptcy using machine learning algorithms. *Int J Cybern Inf* 5:91–105. <https://doi.org/10.5121/ijci.2016.5110>
43. Nagaraj K, Sridhar A (2015) A predictive system for detection of bankruptcy using machine learning techniques. *Int J Data Mining Knowl Manage Process* 5(1):29–40. <https://doi.org/10.5121/ijdkp.2015.5103>

44. Dellepiane U, Marcantonio MD, Laghi E, Renzi S (2015) Bankruptcy prediction using support vector machines and feature selection during the recent financial crisis. *Int J Econ Finance* 7(8). <https://doi.org/10.5539/ijef.v7n8p182>
45. Tian S, Yu Y, Guo H (2015) Variable selection and corporate bankruptcy forecast. *J Banking Finance* 52:89–100. <https://doi.org/10.1016/j.jbankfin.2014.12.003>

Activity Recognition from Videos Using Semantic Motion Patterns



M. Sivarathinabala and R. Jothi Chitra

Abstract In today's world, human activity recognition plays a major role in video surveillance applications. The novelty of this work involves in identifying the activity of the person in the recursive manner using high-level semantics. Recursive activity is defined as the activity that is performed by the persons continuously or recursively. The motion of each and every pixel in the video can give more semantics while identifying the activity. In this work, recursive motion pattern has been identified based on the optical flow algorithm. Initially, the person in the video has been detected and the visual words are extracted. The global and local motion fields are extracted from the video, and this motion pattern gives more semantic meaning in identifying the activity. The semantic gap has been covered using high-level semantic descriptor, and the descriptor output is directly fed into the conditional random field (CRF). Introducing latent variables in the CRF, the descriptor is learned using discriminative model, and the interaction is recognized. The semantic similarity from the interaction has been identified, and the recursive activity is recognized. The new recursive activities such as greeting, fighting, loitering, and buying have been defined from the dataset videos. The importance of the recursive activity is to identify the behavior of the person in the particular scenario. The videos are considered from BIT Interaction dataset, UT Interaction dataset, and collective activity dataset. This work shows the promising results in identifying the recursive activity.

Keywords Recursive activity · Global and local flow motion pattern · Field · Semantic descriptor · Conditional random field

M. Sivarathinabala (✉) · R. Jothi Chitra
Department of Electronics and Communication Engineering, Velammal Institute of Technology,
Chennai, India
e-mail: msrb@velammalitech.edu.in

R. Jothi Chitra
e-mail: rjc@velammalitech.edu.in

1 Introduction

Human activity recognition has been classified in [1, 2], ‘single layered and hierarchical approaches’. Human activity can be recognized using model-based approaches such as dynamic Bayesian networks and HMM. Handshaking, hugging, and pushing are some of the activities have been recognized. From the previous works, it is known that the limitations in identifying the interactions. The activities done by the multiple people in the complex environments are very difficult and challenging. In this interaction recognition, the persons in the videos are separately segmented and their features are extracted. In most of the research works, the features considered are HOG [3] and HOF [4–9]. The classifier used in the previous research works are hidden Markov model and conditional random field. Using these approaches, activity modeling with multiple agents is very difficult.

From our daily experience, the behavior can be recognized from collective activity, concurrent activity, group activity, crowd activity, and recursive activity. In this work, we have proposed recursive activity from the interaction between two persons. Our theoretical contributions on modeling four key aspects of activity. (1) We have been proposed recursive activity recognition from the video images using fully connected latent conditional random field model. In this model, we consider all the interactions that are happened among the people in a video clip. (2) We have defined new recursive activities such as greeting and fighting from the interactions. (3) We describe the person-to-person interactions using the semantic features: location, velocity, and time sequence of the motion attributes. The interaction of the people is different according to the degree of similarity in the features. (4) We perform the inference using belief propagation, and the pair-wise potentials are modeled using Gaussian kernels. (5) We evaluate our model on two interaction datasets, and the experimental results show that our model outperforms the other graphical models.

This research article has been organized as follows: Sect. 2 gives the related works Sect. 3 gives the overview of the recursive activity recognition model. Under the Sect. 3, the action descriptor, interaction recognition, and recursive activity recognition are discussed. Sections 4 and 5 explain the experimental results and discussion, and Sect. 6 concludes the paper.

2 Related Works

Interaction is defined from the successive action that takes place between two persons. Interactions can be recognized in the previous works [4–6]. Ali et al. [3] designed human activity recognition systems and analyzed the detected movement accordingly. The motion pattern of each user is stored over certain period and trained to identify the activity. Zebhi et al. [10] presented a new method to identify the time variations of the motion. The descriptor is named as time-sliced averaged gradient

boundary magnitude (TAGBM). Using these templates, spatial and temporal information can be extracted. For classification, transfer learning technique has been used.

Kong et al. [11] proposed an interactive phrases approach to recognize human interactions using interactive phrases. The interactive phrases explain the motion relationship between the interacting people, and this phrase is considered as latent variables. The authors proposed a discriminative learning model to learn the interactive phrases. The interdependencies between the phrases are considered in this model to overcome the motion ambiguity and partial occlusions. Li et al. [1] proposed the dynamic characteristics of human motion using fuzzy theory. Simple actions are recognized using spatial coordinate of joint points, and the complex actions are recognized using regularization algorithm. Nguyen et al. [2] proposed a framework for HAR application and emerging trends in HAR.

3 Proposed Methodology

The recursive activity has been recognized, and the model has been shown in the Fig. 1. The input video is processed using pre-processing method. The global motion field is calculated for the video fully combining all the video frames. The bag of visual words has been identified, and the optical flow vector has been measured for each visual word. Each visual word represents a motion attribute that contributes toward activity. The global and local motion field vector combines to give a semantic meaning. Then the semantic motion pattern has been extracted. Latent conditional random field has been used to learn the motion patterns. Here, the motion attributes are represented as latent variables. Using L-CRF, the type of interaction has been recognized. Then motion pattern similarity has been calculated from the flow vectors, whenever the same motion pattern is repeated the recursive activity has been recognized. The recursive activity framework has been shown in Fig. 1.

Our task is to learn the motion pattern M to a single interaction label Y . Motion pattern X represented as, $X = \{x_1, x_2, \dots, x_T\}$ in which the observation x_t represents the movement of a person from time $t - 1$ to t , where $t = \{1, 2, \dots, T\}$. x_t is represented by the local feature $\varphi(x_t) \in \mathbb{R}^D$. The Y is the interaction label, and it is given by a set of constants. Let (X_1, X_2) are the observations of two persons and Y interaction as, $y \in \{1, 2, \dots, Y\}$. We have noted the motion pattern observation as X , for our convenience. Based on the latent CRF, latent variables $H = \{h_1, h_2, h_3, h_4, h_5\}$ are introduced to model the intermediate motion attributes contained in the activity. The set H consists of all the five motion attributes. Let us consider the undirected graphical model (UGM), $G = (V, E)$.

In latent CRF, the vertices are the latent variables and (i, j) corresponds to the edges. In edges, the links between h_i and h_j are defined. The dependencies between the random variables are represented by edges. The temporal dynamics have been found out with the linear structure. The maximal cliques are the pair of neighboring states (h_{t-1}, h_t) . Long range dependency has been observed using the connections

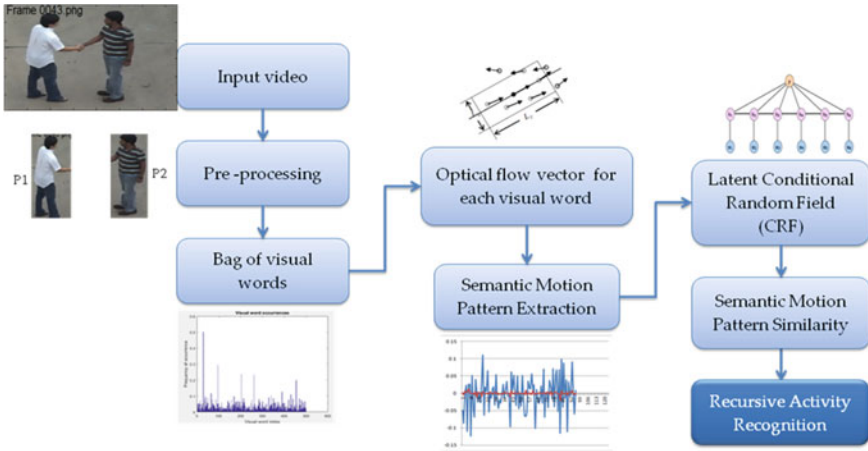


Fig. 1 Recursive activity model

between the latent state and the observations. We model the motion in CRF such as, $p(y, H|X; \theta) = \frac{1}{Z(X; \theta)} \exp\left(\sum_{t=1}^T F(y, h_{t-1}, h_t, X; \theta)\right)$ marginalizing the latent variables, $H = \{h_1, h_2, \dots, h_T\}$ yields the following LCRF form

$$P(y|x; \theta) = \sum_H P(y, H|x; \theta) = \frac{1}{Z(X; \theta)} \exp\left(\sum_{t=1}^T F(y, h_{t-1}, h_t, X; \theta)\right); \quad (1)$$

where $Z(X)$ is the the normalization factor represented as,

$$Z(X; \theta) = \sum_{y', H} \exp\left(\sum_{t=1}^T F(y', h_{t-1}, h_t, X; \theta)\right) \quad (2)$$

The feature function F is defined as

$$F(y, h_{t-1}, h_t, X; \theta) = \sum_{a \in A} \theta_a f_a(y_i, y_j, h_{(i)t-1}, h_{(i)t}, h_{(j)t-1}, h_{(j)t}, X) + \sum_{b \in B} \theta_b f_b(y_i, h_{i(t)}, X) \quad (3)$$

where $f_a(y_i, y_j, h_{(i)t-1}, h_{(i)t}, h_{(j)t-1}, h_{(j)t}, X)$ is pairwise feature function, A is the set of edge features, and f_a is predefined state function which depends on the pair of latent variables. The pairwise feature functions are calculated for all interacting persons, and unary feature function is calculated for individual person. B is the set of node features and f_b is predefined state function which depends on the single latent variable known as unary feature function. From the training data, the parameters $\theta = \{\theta_a, \theta_b\}$ are estimated.

3.1 Learning and Inference

Our training data consists of (X_1, X_2) observations of the interacting persons and the interaction classes (y_1, y_2, \dots, y_N) and motion pattern, $M = \{(X_1, Y_1), (X_2, Y_2), \dots, (X_N, Y_N)\}$. The conditional log-likelihood of the training data is optimized, and the parameters are obtained $L(\theta) = \sum_{i=1}^N L_i \theta = \sum_{i=1}^N \log P(y_i | X_i; \theta)$. Likelihood maximization leads to an optimization task that can be solved using gradient ascent method [2]. For testing, given a new observed motion pattern X , we want to classify the motion pattern into one of the interaction categories.

$y^* = \arg \max_{y \in Y} P(y | X, \theta^*)$ where the values of θ^* are learned from the training data. Inference tasks can be solved using belief propagation [2].

3.2 Recursive Activity Recognition

The interaction is recognized using low-level action descriptor and high level semantic descriptor. The recursive activity is mainly depends on the interaction that happened regularly for a particular time period. The semantic motion pattern has been identified for each and every interaction pattern. Whenever the interaction repeated for a particular time the motion pattern has been same for that particular time. The time $t - 2$, $t - 1$, and t periods interactions are analyzed using the motion patterns as y_i, y_j, y_k respectively. During the time period $t + 1$, the pattern y_i is repeated, then the activity is recognized as recursive activity. For example, fighting is a recursive activity in which the interactions such as kicking, pushing, and boxing are done repeatedly for the period of time. Consider a video x has the motion pattern of interactions (y_i, y_j, y_k) . We first define the similarity between the three motion patterns. The similarity between two motion patterns y_i and y_j is measured as $\frac{(w_i + w_j) \exp -d(y_i, y_j)}{\sum_{i,j} (w_i + w_j)}$.

Where $w_i = \frac{Arcleng(y_i)}{\sum_{k=1}^n Arcleng(y_k)}$, $d(y_i, y_j)$ distance computed by performing dynamic time warping of the motion vectors y_i and y_j , w_i represents the weight associated to each pattern y_i . Here, arc length represents the maximum motion in the pattern, and the activities are recognized with the help of calculated weights.

4 Experimental Setup

Using latent CRF classification technique, the recursive activity recognition system has been proposed, and the system performance has been compared based on different parameters and with different classification techniques. The experiments on recursive activity recognition are based on the observations of interactions between two persons. The real-time activity recognition system that recognizes the interaction

and identifies the activity that is happened from the surveillance videos. The system extracts the various features such as global and local flow field in order to identify the activity more semantically. The features are fused together to form a final vector used to recognize the interaction. Here in this experiment, interactions along with the motion pattern have been considered. The input videos are considered from different datasets such as UT interaction dataset [12], BIT interaction dataset [13], and collective activity dataset [14].

In this work, we have considered six interactions such as handshaking, hi-five, and hugging for greeting activities, and push, kick, and box interactions for fighting activities. The collective activity dataset [14, 15] consists of 44 short video clips, including five activity classes: crossing, waiting, queuing, walking, and talking. From this dataset, loitering and buying are the two new recursive activities defined. Loitering activity is defined as continuous walking activity for the period of time. If the interaction of waiting and queuing continuously happens during the particular time interval, then the recursive activity buying is defined.

5 Results and Discussion

Preprocessing is the first and foremost step to identify the person. In the pre-processing stage, the persons are segmented separately, and their features are extracted separately. The global and local flow fields are calculated in order to identify the motion pattern of the particular interaction. The global flow field is calculated collectively from all the frames. The position and location of the motion features are collectively shown in the global flow field. In the local flow field, the attributes are selected manually using bag of visual words. The motion in the corresponding attributes is identified in order to identify the semantics. The motion relationships can give more meaning in identifying the interactions. Figure 2 shows the global motion pattern for the handshaking and hi-five interactions.

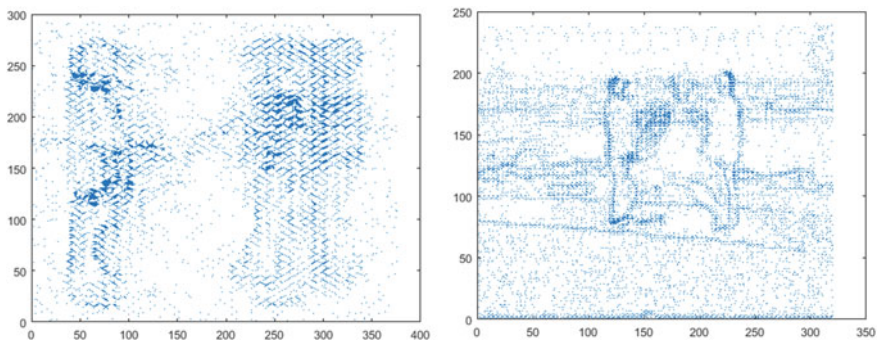


Fig. 2 Global motion pattern from optical flow algorithm

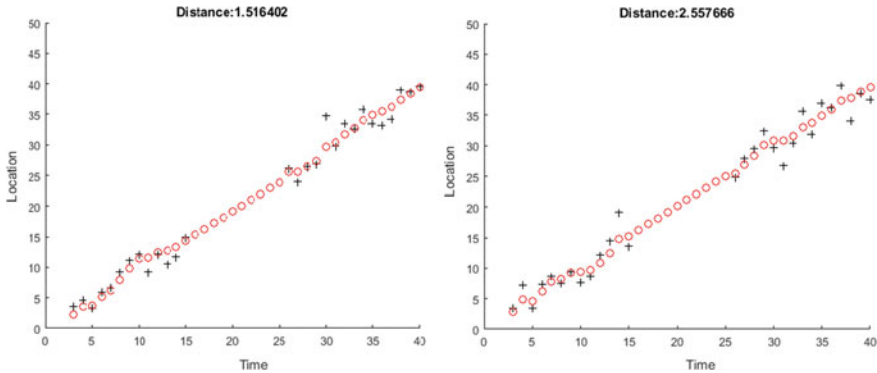


Fig. 3 Motion attributes location and velocity

Motion pattern extraction includes the extraction of flow features using both global and local approaches. In the local flow field computation, each visual word location and velocity are computed and shown in figure. The latent states are initialized based on data-driven approach [3]. Figure 3 shows the motion attributes such as location and velocity.

In this work, recursive motion pattern is the major contribution. It is well known from the previous techniques that, motion pattern is derived from the different descriptors. The simple motion pattern [3, 6] can recognize the simple actions such as walking, jogging, running, and waving. Some interactions for a particular period of time can represent a behavior of the person in the particular scenario. Thus, recursive motion pattern is more important in identifying the recursive activity. Figure 5 shows the semantic motion pattern of greeting activity. The semantic motion pattern similarity has been measured for every time period t . Whenever the similar pattern is repeated after $t + 1$ time period, the activity is recognized. The maximum arc length of the motion pattern is measured and the weights are calculated. Human behavior has been recognized based on his activities or interactions between other persons or interactions between other agents such as bag. In this work, behavior of the person has been recognized based on the interactions between other persons (Fig. 4).

The fighting activity is defined from the interactions such as boxing, kicking, and pushing. Figure 5 shows the semantic motion pattern of fighting activity.

6 Conclusion

In this paper, a novel recursive activity recognition system was implemented and presented for identifying activities by employing semantics descriptor. Here, the recursive motion patterns are learned using latent CRF that are fully connected. The proposed model depends on location, velocity, and time. Recursive activity is modeled using fully connected latent CRF model is proposed. The pairwise potentials

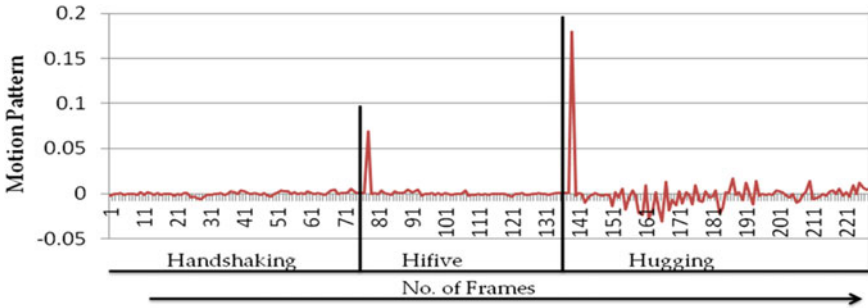


Fig. 4 Semantic motion pattern of greeting activity

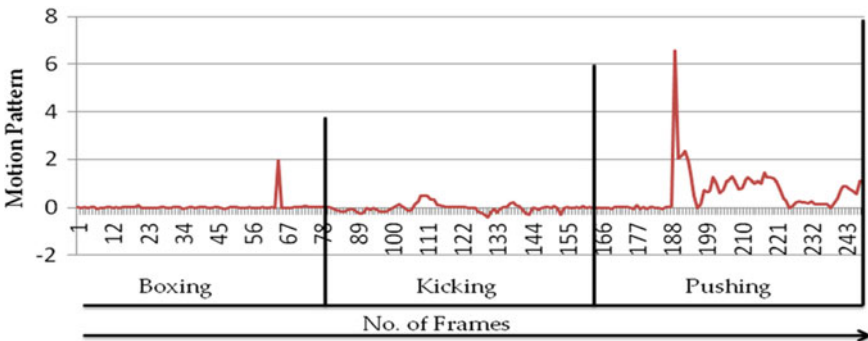


Fig. 5 Semantic motion pattern of fighting activity

and efficient approximation method is used for this model. The experimental results showed that proposed model can able to deal with the recursive activities in the video. The evaluation has been done on the two publicly available datasets showed that latent CRF and fully connected model outperforms state-of-the-art models. Future direction includes the investigation of other useful contexts for recursive activity recognition, extension to online processing for online applications and research on efficient learning techniques.

References

1. Li X, Sun G, Li Y (2021) Human motion representation and motion pattern recognition based on complex fuzzy theory. Complexity, ArticleID 9923748, 12 pages. <https://doi.org/10.1155/2021/9923748>
2. Nguyen B, Coelho Y, Bastos T, Krishnan S (2021) Trends in human activity recognition with focus on machine learning and power requirements. Mach Learn Appl 5:100072, ISSN 2666-8270. <https://doi.org/10.1016/j.mlwa.2021.100072>
3. Ali A, Samara W, Alhaddad D, Ware A, Saraereh OA (2022) Human activity and motion pattern

- recognition within indoor environment using convolutional neural networks clustering and Naive Bayes classification algorithms. *Sensors* 22:1016. <https://doi.org/10.3390/s22031016>
4. Ryoo MS, Agarwal JK (2011) Human activity analysis: a survey. *ACM Comput Surv* 43(3):16:1–16:43
 5. Vahdat A, Gao B, Ranjbar M, Mori G (2011) A discriminative key pose sequence model for recognizing human interactions. In: *IEEE international conference on computer vision*
 6. Chen S, Liu J, Wang H (2013) A hierarchical human activity recognition framework based on automated reasoning. In: *IEEE international conference on systems, man, and cybernetics*, pp 3495–3499
 7. Bruhn A, Weickert J (2005) Lucas/Kanade meets Horn/Schunck: combining local and global optic flow methods. *Int J Comput Vis* 61(3):211–231
 8. Jain M, Jégou H, Bouthemy P (2013) Better exploiting motion for better action recognition. In: *IEEE international conference on computer vision and pattern recognition*, pp 2555–2562
 9. Huang K, Wang S, Tan T, Maybank S (2009) Human behaviour analysis based on new motion descriptor. In: *IEEE transactions on circuits and systems for video technology*
 10. Zebhi S, Almodarresi SMT, Abootalebi V (2021) Human activity recognition based on transfer learning with spatio-temporal representations. *Int Arab J Inf Technol* 18(6). Electrical Engineering Department, Yazd University, Iran
 11. Kong Y, Jia Y, Fu Y (2012) Learning human interaction by interactive phrases. In: *European conference on computer vision*, vol 7572, pp 300–313
 12. Ryoo MS, Aggarwal JK (2010) UT interaction dataset, ICPR contest on semantic description of human activities (SDHA).
 13. Sadek S, Al-Hamadi A, Michaelis B, Sayed U (2012) A fast statistical approach for human activity recognition. *Int J Comput Inf Syst Ind Manage Appl* 4:334–340. ISSN 2150-7988
 14. Khamis S, Morariu VI, Davis LS (2012) A flow model for joint action recognition and identity maintenance. In: *IEEE conference on computer vision and pattern recognition (CVPR)*
 15. Lan T, Wang Y, Mori G (2010) Retrieving actions in group contexts. In: *International workshop on sign gesture activity (SGA)*

Machine and Deep Learning Technique for Depression Detection Using EEG Data



Shikha Tiwari, Kiran Pandey, Vivek Sharma, and Bhupendra Verma

Abstract Electroencephalogram (EEG) signal-based emotion recognition has attracted wide interests in recent years and has been broadly adopted in medical, affective computing, and other relevant fields. Depression has become a leading mental disorder worldwide. Evidence has shown that subjects with depression exhibit different spatial responses in neurophysiologic signals from the healthy controls when they are exposed to positive and negative. Depression is a common reason for an increase in suicide cases worldwide. EEG plays an important role in e-healthcare systems, especially in the mental healthcare area, where constant and unobtrusive monitoring is desirable. EEG signals can reflect activities of the human brain and represent different emotional states. Mental stress has become a social issue and could become a cause of functional disability during routine work. This paper proposed an adaptive approach based on machine and deep learning for detecting depression using EEG. The algorithm first extracts features from EEG signals and classifies emotions using machine and deep learning techniques, in which different parts of a trial are used to train the proposed model and assess its impact on emotion recognition results.

Keywords EEG · Emotion · Stress · Deep learning · Machine learning · E-healthcare

1 Introduction

This WHO estimates that the burden of mental health problems in India is 2443 disability-adjusted life years (DALYs) per 10,000 population; the age-adjusted suicide rate per 100,000 population is 21.1. The economic loss due to mental health conditions, between 2012 and 2030, is estimated at USD 1.03 trillion.

S. Tiwari (✉) · K. Pandey · V. Sharma · B. Verma
Department of CSE, Technocrats Institute of Technology, Bhopal, India
e-mail: rovintiwari@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
V. V. S. S. Chakravarthy et al. (eds.), *Advances in Signal Processing, Embedded Systems and IoT*, Lecture Notes in Electrical Engineering 992,
https://doi.org/10.1007/978-981-19-8865-3_34

375

Stress is commonly recognized as a state in which an individual is expected to perform too much under sheer pressure and in which he/she can only marginally contend with the demands. These demands can be psychological or social. It is known that psychosocial stress exists in daily life, which has resulted in poor quality of life by affecting people’s emotional behavior, job performance, and mental and physical health [1]. Psychosocial stress is a leading cause of several physiological disorders. For example, it increases the likelihood of depression, stroke, heart attack, and cardiac arrest [4].

Electroencephalography (EEG) is an efficient modality which helps to acquire brain signals corresponds to various states from the scalp surface area. These signals are generally categorized as delta, theta, alpha, beta, and gamma based on signal frequencies ranges from 0.1 Hz to more than 100 Hz. It is a test that detects electrical activity in the brain using small, metal disks (electrodes) attached to the scalp. Routinely, EEG is used in clinical circumstances to determine changes in brain activity that might be useful in diagnosing brain disorders, especially epilepsy or another seizure disorder.

The types of EEG waves [2, 3] are identified according to their frequency range-delta: below 3.5 Hz (0.1–3.5 Hz), theta: 4–7.5 Hz, alpha: 8–13 Hz, beta: 14–40 Hz, and gamma: above 40 Hz. The EEG may show unusual electrical discharge when some abnormality occurs in the brain (Fig. 1).

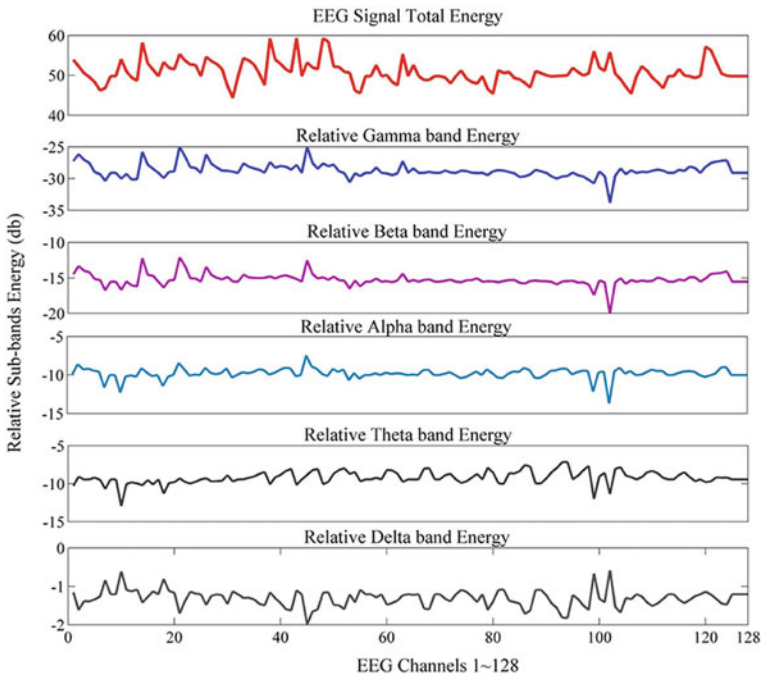


Fig. 1 EEG signal [1]

PREVALENCE PER 100,000			
DEPRESSIVE DISORDERS		CONDUCT DISORDERS	
TamilNadu	4,796	Jharkhand	983
Andhra Pradesh	4,563	Bihar	974
Telangana	4,356	Meghalaya	961
Odisha	4,159	Uttar Pradesh	927
Kerala	3,897	Nagaland	924
ANXIETY DISORDERS		IDIOPATHIC DEVELOPMENTAL INTELLECTUAL DISABILITY	
Kerala	4,035	Bihar	6,339
Manipur	3,760	Uttar Pradesh	5,503
West Bengal	3,480	Madhya Pradesh	5,216
Himachal Pradesh	3,471	Assam	5,121
Andhra Pradesh	3,462	Jharkhand	4,940

Fig. 2 Mental health data (Indian health report)

Stress is your body's reaction to a challenge or demand. In short bursts, stress can be positive, such as when it helps you avoid danger or meet a deadline. EEG nonlinear dynamics features and frontal asymmetry of theta, alpha, and beta bands have been selected as biological indicators for chronic stress, showing relative greater right anterior EEG data activity in stressful individuals (Fig. 2).

Emotions often facilitate interactions among human beings, but the big variation of human emotional states make a negative effect on the reliable emotion recognition [6].

Multimodal emotion recognition is an emerging interdisciplinary field of research in the area of affective computing and sentiment analysis. It aims at exploiting the information carried by signals of different nature to make emotion recognition systems more accurate. This is achieved by employing a powerful multimodal fusion method [8].

The main motivation of this work is to make a optimize model of prediction of the depression using EEG signal.

This paper is organization into five sections; Section 1 describes the overview or introduction of the concept. Section 2 presents the literature review or previous work done in this field. Section 3 shows the proposed methodology and the steps involving it. Section 4 presents the simulation work and results analysis. Section 5 presents the conclusion and future scope of the research.

2 Literature Review

Seal et al. [1] suggest that CNN trained on recordwise split data gets overtrained on EEG data with a small number of subjects. The performance of DeprNet is remarkable compared with the other eight baseline models.

Sun et al. [2] present combined multi-types features (All: L + NL + PLI + NM) outperformed single-type features for classifying depression. Analyzing the optimal features set we found that compared to other types features, PLI occupied the largest proportion of which functional connections in intra-hemisphere were much more than that of in the inter-hemisphere.

Zheng et al. [3] investigate stable patterns of electroencephalogram (EEG) over time for emotion recognition using a machine learning approach. Up to now, various findings of activated patterns associated with different emotions have been reported.

Fang et al. [4] present the operation was validated using ADVANTEST V93000 PS1600, and the training process and real-time classification processing time took 0.12495 ms and 0.02634 ms for each EEG image, respectively.

Bota et al. [5] show the affective computing is a multidisciplinary field of research spanning the areas of computer science, psychology, and cognitive science. Potential applications include automated driver assistance, healthcare, human–computer interaction, entertainment, marketing, teaching, and many others.

Wang et al. [6] the proposed method for emotion recognition is verified on the common facial expression datasets, the Extended Cohn-Kanade (CK+) dataset and the Japanese female facial expression (JAFFE). The results are satisfactory, which shows cloud model is potentially useful in pattern recognition and machines learning.

Khalil et al. [7] present the emotion recognition from speech signals is an important but challenging component of human–computer interaction (HCI). Deep learning techniques have been recently proposed as an alternative to traditional techniques in SER (Fig. 3).

Nemati et al. [8] presents a hybrid multimodal data fusion method is proposed in which the audio and visual modalities are fused using a latent space linear map and then, their projected features into the cross-modal space are fused with the textual modality using a Dempster-Shafer (DS) theory-based evidential fusion method.

3 Methodology

- Firstly, download the EEG dataset from kaggle website, which is a large dataset provider and machine learning repository provider company for research (Fig. 4).

The EEG dataset has 2133 rows and CTA columns. The features of the dataset shows like # mean_0_a mean_1_a mean_2_a mean_3_a mean_4_a mean_d_0_a mean_d_1_a mean_d_2_a mean_d_3_a etc.

The sample dataset of selected dataset is taken form `fft_0_b':'fft_749_b.`

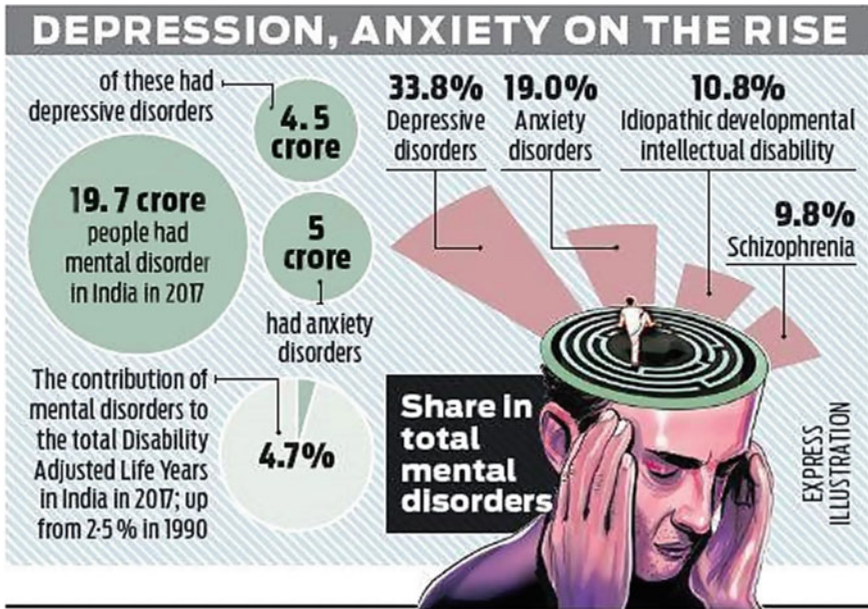


Fig. 3 Depression statics

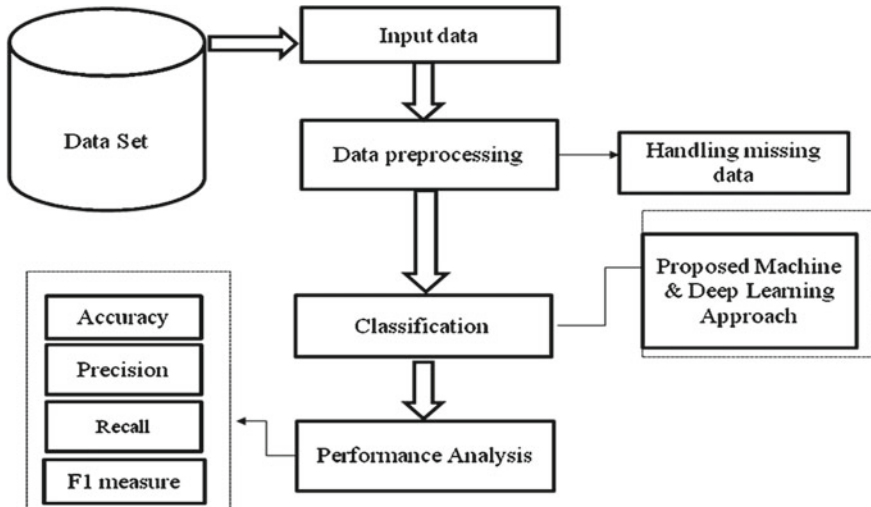


Fig. 4 Flowchart

- Now apply the preprocessing of the data, here handling the missing data, removal null values.
- Now extract the data features and evaluate in dependent and independent variable.
- Now apply the classification method based on the machine learning (KNN) and deep learning (LSTM) approach.

KNN: KNN algorithm assumes the similarity between the new case/data and available cases and put the new case into the category that is most similar to the available categories. It classifies a new data point based on the similarity. This means when new data appears then it can be easily classified into a well-suited category by using KNN algorithm.

RNN-LSTM: Long short-term memory (LSTM) is an artificial recurrent neural network (RNN) architecture used in the field of deep learning.

It can process not only single data points (such as EEG signal o EEG images), but also entire sequences of data.

The long short-term memory (LSTM) cell can process data sequentially and keep its hidden state through time.

A recurrent neural network (RNN) is a class of artificial neural networks where connections between nodes form a directed or undirected graph along a temporal sequence. This allows it to exhibit temporal dynamic behavior (Fig. 5).

Recurrent Layer

The independently recurrent neural network addresses the gradient vanishing and exploding problems in the traditional fully connected RNN. Each neuron in one layer only receives its own past state as context information (instead of full connectivity to all other neurons in this layer), and thus, neurons are independent of each other’s history (Fig. 6).

Mathematically, the recurrent operation can be expressed as:

$$g(x, y) = w * f(x, y)$$

where

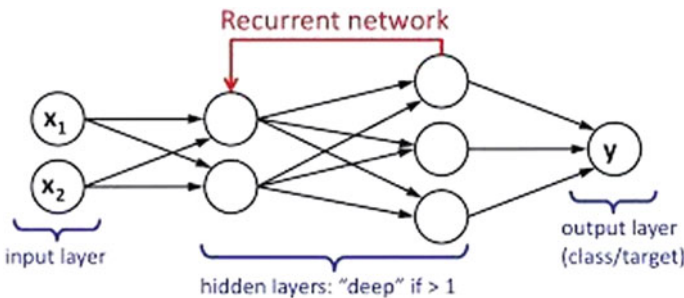


Fig. 5 Basic RNN architecture

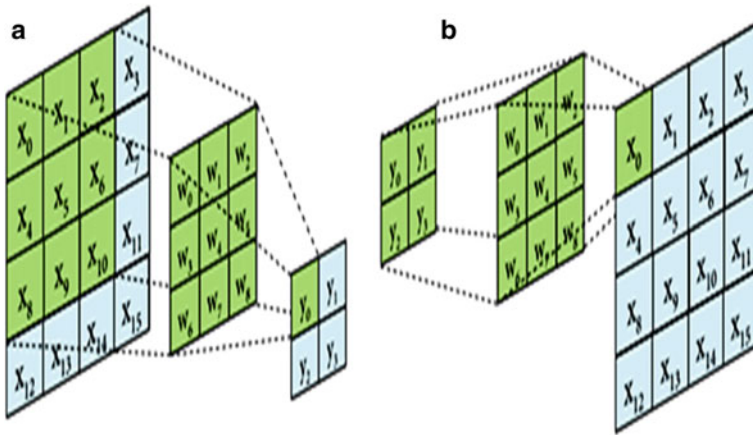


Fig. 6 **a** Recurrent operation, **b** equivalent transposed recurrent operation

$g(x, y)$ = Image after filtering,
 w = Filter kernel,
 $f(x, y)$ = Input image

Initially, the computer vision experts were designing the filters which were applied to the images for the analysis of various features of the image. Nowadays, during the training process, the weights, i.e., the values in the filter, are updated automatically, this is the innovation of using the recurrent layer in the neural network. During the training process, the network learns which features it needs to be extracted from the image.

- Now generate a confusion matrix and show all predicted classes like true positive, false positive, true negative, and false negative.
- Now calculate the performance parameters by using the standard formulas in terms of the precision, recall, F_measure, accuracy, and error rate.
- Precision is a measure of the accuracy, provided that a class label has been predicted. It is defined by:

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$$

- Recall is the true positive rate:

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$$

- F1 Score is needed to a balance between Precision and Recall

$$F1_Score = \frac{2 \times (Precision \times Recall)}{Precision + Recall}$$

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

$$Classification\ Error = 100 - Accuracy$$

4 Simulation Results

The simulation is performed using Python Spyder software (Table 1).

Figure 7 is showing the dataset of this research. The dataset is taken from the kaggle website.

Figure 8 is showing the training dataset which is used to train the model.

Figure 9 is showing the test dataset which is used to test the proposed model.

Table 1 Specification of the system

Sr. No	Parameter	Value
1	Windows	7
2	RAM	8 GB
3	CPU	Intel
4	Processor	Core i3
5	Graphics Memory	2 GB
6	Software	Python
7	IDE	Sypder
8	Version	3.7

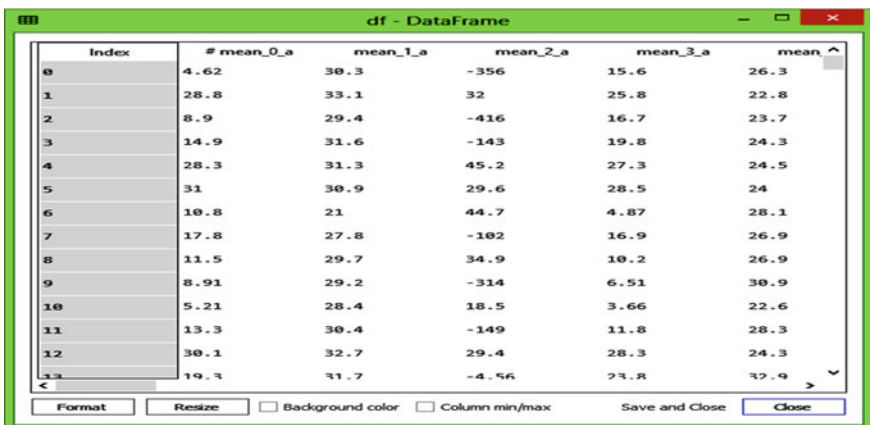


Fig. 7 Dataset

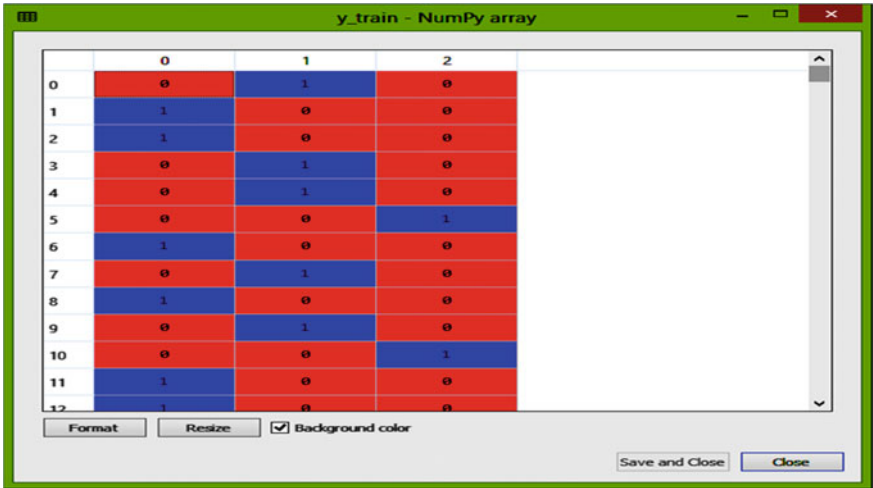


Fig. 8 Train dataset

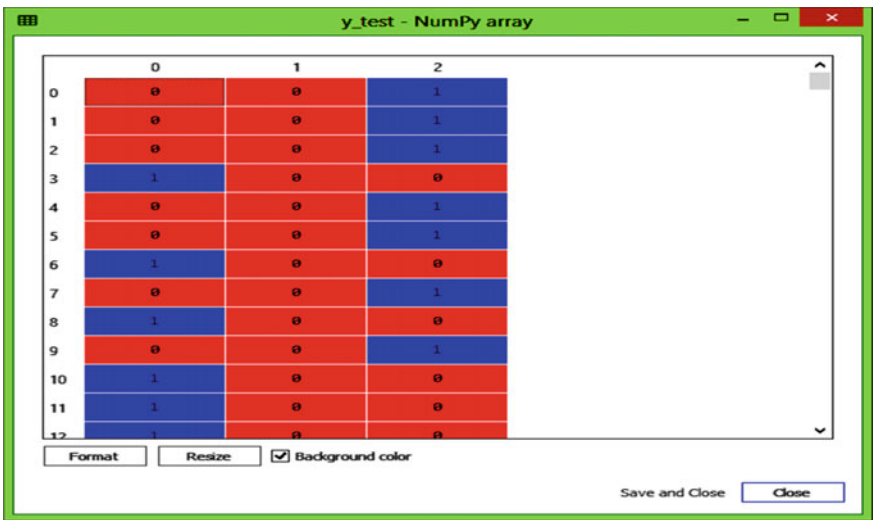


Fig. 9 Test dataset

Figure 10 is showing the confusion matrix of the proposed model (Tables 2, 3 and 4).

Fig. 10 Confusion matrix

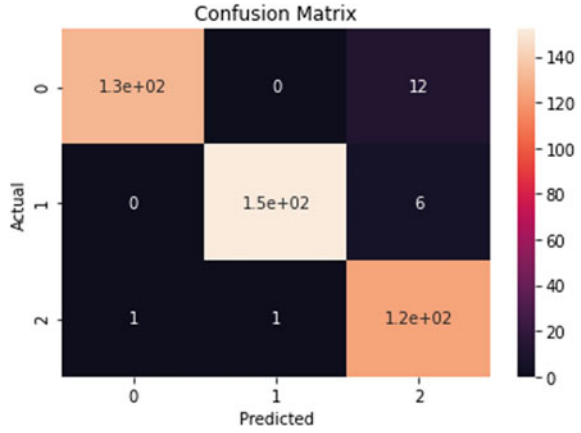


Table 2 Simulation results of KNN

Sr. No	Parameters	Proposed approach (%)
1	Accuracy	94.14
2	Classification error	5.86
3	Precision	97
4	Recall	94
5	F-measure	95

Table 3 Simulation results of LSTM

Sr. No	Parameters	Proposed approach (%)
1	Accuracy	96.48
2	Classification error	3.52
3	Precision	99
4	Recall	94
5	F-measure	97

Table 4 Result comparison

Sr. No	Parameters	Accuracy (%)	Classification error (%)
1	Work [1]	91	9
2	Work [3]	91.07	8.93
3	Work [9]	88.56	11.44
4	Proposed work	96.48	3.52

5 Conclusion

This research proposed the machine learning and deep learning technique to identify the prediction from given dataset. The Python Spyder ISE 14.7 software is used to simulate the work. Machine learning KNN classifier approach is achieved 94% accuracy, while deep learning LSTM classifier approach is achieved 96% accuracy. Therefore, the simulation results show that the proposed approach gives significant better results than existing work. In future the other set can be taken and apply more classification and the regression methods and predict various other parameters.

References

1. Seal A, Bajpai R, Agnihotri J, Yazidi A, Herrera-Viedma E, Krejcar O (2021) DeprNet: a deep convolution neural network framework for detecting depression using EEG. *IEEE Trans Instrumentation Meas* 70:1–13, Art no. 2505413. <https://doi.org/10.1109/TIM.2021.3053999>
2. Sun S, Chen H, Shao X, Liu L, Li X, Hu B (2020) EEG based depression recognition by combining functional brain network and traditional biomarkers. *IEEE Int Conf Bioinformatics Biomed (BIBM) 2020*:2074–2081. <https://doi.org/10.1109/BIBM49941.2020.9313270>
3. Zheng W, Zhu J, Lu B (2019) Identifying stable patterns over time for emotion recognition from EEG. *IEEE Trans Affective Comput* 10(3):417–429. <https://doi.org/10.1109/TAFFC.2017.2712143>
4. Fang W, Wang K, Fahier N, Ho Y, Huang Y (2019) Development and validation of an EEG-based real-time emotion recognition system using edge AI computing platform with convolutional neural network system-on-chip design. *IEEE J Emerging Selected Topics Circuits Syst* 9(4):645–657. <https://doi.org/10.1109/JETCAS.2019.2951232>
5. Bota PJ, Wang C, Fred ALN, Plácido Da Silva H (2019) A review, current challenges, and future possibilities on emotion recognition using machine learning and physiological signals. *IEEE Access* 7:140990–141020. <https://doi.org/10.1109/ACCESS.2019.2944001>
6. Wang S, Chi H, Yuan Z, Geng J (2019) Emotion recognition using cloud model. *Chin J Electronics* 28(3):470–474. <https://doi.org/10.1049/cje.2018.09.020>
7. Khalil RA, Jones E, Babar MI, Jan T, Zafar MH, Alhussain T (2019) Speech emotion recognition using deep learning techniques: a review. *IEEE Access* 7:117327–117345. <https://doi.org/10.1109/ACCESS.2019.2936124>
8. Nemati S, Rohani R, Basiri ME, Abdar M, Yen NY, Makarenkov V (2019) A hybrid latent space data fusion method for multimodal emotion recognition. *IEEE Access* 7:172948–172964. <https://doi.org/10.1109/ACCESS.2019.2955637>
9. Zhang H, Jolfaei A, Alazab M (2019) A face emotion recognition method using convolutional neural network and image edge computing. *IEEE Access* 7:159081–159089. <https://doi.org/10.1109/ACCESS.2019.2949741>

Next-Generation Firewall with Intelligent IPS



Parth Barot, Sharada Valiveti, and Vipul Chudasama

Abstract These days, Internet is used for almost everything. There are several mediums (channels) available for people to search for information that is relevant either in terms of business or entertainment perspective. These mediums can be used by a hacker to steal information and perform an attack on any network infrastructure using various techniques. Hence, securing the network is essential for cybersecurity. Accordingly, various types of attacks like denial of service (DoS) attacks, backdoors, buffer overflow, guess the password, Smurf DoS, etc., are quite prevalent and are commonly used by an attacker. This paper presents the complete process to detect and re-mediate/mitigate these attacks using automatic detection techniques. The paper includes ways to improve network security using the feature of intrusion prevention system in firewall by detecting and analyzing the packet flows. The main objective of the paper is to detect the attack, analyze it, and if it is risky then drop the attacking packet before it harms the network. Also, the paper presents the error and accuracy rate of detecting attacks using the different machine learning algorithms which can be used in the IPS.

Keywords Firewall · Intrusion detection system · Intrusion prevention system · Attacks · KDD CUP dataset

1 Introduction

Provisioning security aims to build a framework comprising of a set of methods and that help defend against various cyberattacks [14]. A cyberattack is a harmful act that aims to harm data, steal data, and disrupt digital life. Malware, data breaches,

P. Barot · S. Valiveti (✉) · V. Chudasama
Institute of Technology, Nirma University, Ahmedabad, Gujarat 382481, India
e-mail: sharada.valiveti@nirmauni.ac.in

P. Barot
e-mail: 20mcei01@nirmauni.ac.in

V. Chudasama
e-mail: vipul.chudasama@nirmauni.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
V. V. S. S. S. Chakravarthy et al. (eds.), *Advances in Signal Processing, Embedded Systems and IoT*, Lecture Notes in Electrical Engineering 992,
https://doi.org/10.1007/978-981-19-8865-3_35

denial of service (DoS) attacks, various attack vectors, etc., [13] are a few examples of cyberattacks. Attackers may initially be aimless sometimes. They continuously monitor the system (network). As and when they find any vulnerability, they try to exploit it. Cyber-offenders gain huge benefits from this. Therefore, cyber-threat is now an industry for money-making.

Network-based intrusion detection system (NIDS) inspects the packets flowing through the network and detects any suspicious activity in the network [5]. Every packet entering the network is checked to make sure that the packets are not containing any malicious content. The detection techniques used by NIDS are divided into two categories: anomaly based [7] and signature based [9]. When the system deviates from its typical behavior, anomaly-based NIDS issues an alert. When the analyzed data matches the known attack pattern, a signature-based NIDS generates an alarm (signature). Most of the NIDS are signature based. Anomaly-based NIDS offer one major advantage over signature-based NIDS - they can detect threats for which no signature exists yet, such as zero-day and “targeted assaults” according to [3]. As a result, signature-based NIDS are unable to detect unknown assaults.

Enclosing a malicious link in email, phishing Web sites, social engineering, etc., is some common methods employed by attackers. Various types of malware like adware/spyware, ransomware, trojans, worms, viruses, etc., also impact performance of the system followed by the network [16]. Firewalls are considered to be the first line of defense in a network [1]. Hence, it is mandatory to provide an intelligent detection and prevention system with the firewalls so as to evade these attacks.

This paper takes us to a lucid introduction of the state of the art in Sect. 2, followed by a systematic study of various metrics introduced to evaluate measures of network security. The proposed change management in firewall and methodology are discussed in Sect. 3. Discussion about the implementation results is elucidated in Sect. 4. Finally, the conclusions are briefed in Sect. 5.

2 Literature Review

A study of recent work is summarized in this section. Chou et al. [3] presented a two-phase IDS technique. They created a correlation-based feature selection system to select important features like port number, protocol from a large dataset. The second part is to build an intrusion detection system to address the ambiguity caused by insufficient and ambiguous information.

Peddachigari et al. [10] proposed a hybrid intrusion detection system that combines decision tree and support vector machine (SVM) in a hierarchical manner. Also, an ensemble technique is combined with base classifiers to improve accuracy and reduce processing costs. It was observed that a hybrid system delivers a higher accuracy and detection rate than a single SVM or decision tree.

Guo et al. [6] introduced a hybrid intrusion detection system that combines a two-class support vector mechanism with a decision tree, resulting in a higher detection rate than a single SVM. Packets are segregated into four parts—TCP, UDP, ICMP,

and application layer. Each packet consists of a two-class support vector mechanism and a decision tree. Through this, the training time is significantly reduced. However, the accuracy of detecting novel attacks is extremely low.

Daneshgadeh et al. [4] suggested an information gain-based model with a useful feature selection facility. Because KDD is the most extensively used dataset for system evaluation, Mahbod Tavallaee et al. suggested a complete analysis of the KDD Cup 99 dataset using classification. Redundant records are found in the KDD data. The method has a high classification rate and explores the 41 significant features in relation to the dataset labels.

Katkar et al. [8] created a hybrid IDS (HIDS) that combines the C5.0 decision tree classifier with the one class support vector machine (OC-SVM). The C5.0 decision tree classifier was used to create the SIDS (State-based IDS). This framework detects both evident intrusions and zero-day assaults with high detection accuracy and low false-alarm rates.

Senthilnayaki et al. [12] proposed an enhanced network intrusion detection system. K -means clustering and Naive Bayes classification algorithms are used in this detection strategy. Clustering with K -means in order to create clusters for normal and anomalous traffic, an algorithm is used.

Chandola et al. [2] presented an anomaly detection model that is used to detect anomalies in network's nodes. It contain various forms of data, which must be classified into critical information that includes intrusions, defects, and system failures.

In many ways, the work proposed in this study is more effective than all other studies in the literature. To begin with, the trials are carried out using a standard benchmark dataset, the KDD cup dataset. Secondly, using the important feature like protocol, flag of the dataset tries to detect attack with different machine learning algorithm to make better classification decision. Finally, it lowers the rate of false positives and errors.

Researchers used various metrics to evaluate the network security measures employed by them. The following subsection explores these metrics.

2.1 Network Security Measures

There are several network security measures used by organizations, which safeguard the enterprise network either by restricting access or by protecting the network boundary. These include the firewalls and the IDS/IPS.

- **Firewalls:**

The function of a firewall is to monitor every transacted packet in the network and decide whether it should be sent to the destined receiver or not. Software-based firewalls are mostly host based which protects one host or system from the security threats whereas the hardware firewalls or the physical firewalls are devices which are mostly kept between the network to be protected and its gateway. Details about types of firewalls are as mentioned in Table 1.

Table 1 Types of firewalls

Firewalls	Strength	Weakness
Packet filtering firewalls	There are two types of packet filtering firewalls, namely, stateless and stateful firewalls. In stateless firewalls, all the packets are examined independently. On the other hand, stateful firewalls store the information about the session and are more secure	Packet filtering firewalls are effective but have limited capabilities; they are not intelligent enough to safeguard a network from today’s attacks [4]
Next-generation firewalls	NGFW solutions were typically based on signatures, meaning they were trying to understand patterns of how malicious traffic behaved. In this approach, a group of researchers would observe a new attack, analyze it, and once it was understood, they would develop a pattern or signature of the attack	This was very similar to how IPS systems operated with the difference that they were optimized to catch DoS and DDoS attacks and not necessarily exploits, worms, bots, or other malware traversing the network
Proxy firewalls	Most notably, most prominently, the intermediary firewalls screen traffic for layer 7 conventions, for example, HTTP and FTP, and use both stateful and deep packet inspection to detect malicious traffic	The customer must send a solicitation to the firewall, where it is then assessed against a lot of security rules and afterward allowed or blocked

● **Intrusion Detection/Prevention System:**

Intrusion Detection/Prevention Systems are also necessary to detect intrusions in a network. IDS inspects each packet by traversing the content it is carrying to detect any malicious activity. Table 2 elucidates various types of IDS/IPS systems with their strengths and weaknesses.

Scalability and risk assessment are mandated in implementation of any network infrastructure. The proposed approach, being presented in the next section presents the implementation scenario of change management in a firewall. This change management suggests integration of an intelligent IDS/IPS into the firewall for real-time removal of malicious traffic in the network.

Table 2 Types of IDS/IPS

IDS	Advantages	Disadvantages
HIDS	HIDS can actively analyze encrypted data and conversation HIDS inform us whether an attack was successful or not. It is simple to set up because it doesn't require any more hardware, and it doesn't interfere with the current architecture	If the attack causes the OS to fail, HIDS will fail. Network scanning and DOS attacks are not detectable by HIDS. HIDS are known for requiring a lot of work
NIDS	Because NIDS is not affected by the operating environment, they will not affect the host's performance	Encrypted traffic cannot be analyzed. Inside the host machine, NIDS has extremely less visibility

3 Change Management in Firewall

Scalability is an intrinsic need of a business. Hence, there is a need for change in firewall configurations on a regular basis. Sometimes the changes are needed on an hourly basis which may require continuous change management, monitoring and maintenance so as to ensure that the security and network performance remains intact. With high frequency of changes, there is always a chance of mis-configuration in firewall, which can lead to the vulnerable system or breaking the existing connections and hence affecting the business. Manual change management processes consumes a lot of time and also impeded business agility [6].

Risk assessment deals with whether it is appropriate for the application to access a particular server or database. An application which is in a less secure domain should not be allowed to access a database server which is in a highly secure domain. After the risk assessment is complete, the engineer can work to provision the connections and implement the request. This process of manual change implementation is very time-consuming and keeps the network administrators engaged for even small requests. The complex requests can take a lot more time and energy.

Hence, the integration of IDS/IPS with firewall to facilitate scalability and risk assessment will benefit the futuristic firewall trends manifold. The next subsection proposes the methodology.

3.1 Proposed Methodology

There are 3 options for IPS we can put in firewall

- IPS before Firewall: Load on IPS will be much higher
- IPS After firewall: Incoming traffic will be dropped and hence more applicable
- IPS inline firewall: IPS will be the part of the firewall.

Figure 1 shows the approach of implementation.

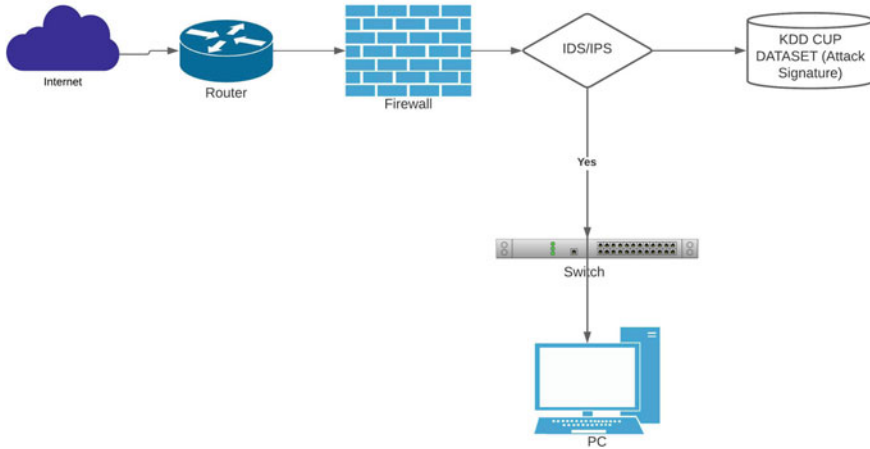


Fig. 1 Proposed methodology

The experiment is based on the KDD CUP 99 dataset [15]. It has 42 properties that are divided into three categories: numeric, nominal, and binary. There are two types of classes: normal and anomaly. Some attacks in the anomaly class are classified as DoS, Smurf, and buffer overflow [11]. This dataset contains sufficient records for training and testing. The model is trained for various network attacks on the KDD Cup 99. The approach will be able to forecast any unknown attacks, which will aid in network intrusion detection. The classifier can assist in tracing the decision tree based on netropy. The protocol feature in any packet is very useful because it contains TCP, UDP, or ICMP using which protocol-based attacks like DDoS can be detected. Service field is used to determine the request which the user requests for. If the access is risky like remote desktop or file sharing, the related attack can be detected. From flag feature, information about whether the connection is established with attacker or not is determined. So, these 3 feature are very important and based on which we are trying to predict attack types.

3.2 Determining Normal and Abnormal Traffic

Normal traffic can be analyzed by the data feature. It is like a normal behavior of packets. The protocol and service they want are also not risky, so it can be defined as normal traffic. Abnormal traffic can be analyzed by matching it with any abnormal predefined signatures which tries to connect with some risky ports. The next section focuses on the implementation of proposed approach and analyzes the working of the model.

4 Implementation Results

The proposed approach aims to train the learning algorithm with live traffic. The related Algorithm is shown in Algorithm 1. Firewall and IPS collects the network metrics and classifies using the incoming traffic using machine learning algorithm. The algorithm classifies the traffic either as an attack or as a normal traffic. If it is an attack, firewall will take action. For the implementation purpose, data is split into 80:20 ratio as training and test data, respectively. Decision tree, random forest, MLP classifier and Gaussian Naive Bayes classifiers are used. There are total 42 classes from which 3 classes have been used as discussed in the previous section.

Algorithm 1 Determine if packet is normal or abnormal

```

Require: Packets
Ensure: Determine if packets need to be dropped
Use Anomaly or Signature Based algorithm
if Are the analyzed packets abnormal? then
  It is an Abnormal Packet
  if Are the anomalies corresponding to one of the set of attacks? then
    Anomaly is matching with attack parameters
    Drop the Packet
  else
    Anomaly is not matching with attack parameters
  end if
else
  The analyzed packet is normal
end if

```

When any user tries to access any repository or try to connect with different server, it goes via packet in the firewall. Firewall checks the packets against the set of rules and determines whether it is normal or abnormal using signature/anomaly-based algorithm included with IDS/IPS. If the packet is normal and from firewall, rule set is already allowed; it will allow the flow of the connection. If the packet is abnormal, then it checks using the signature parameter to identify if the packet is vulnerable or not. If the packet is matching with the parameters which we set in IDS/IPS, it drops the packet. If the packet parameters do not match with the signature parameters, and also from firewall if the rule is already there for allowing the connection, the said flow connection goes through.

Accuracy is measured based on true positives while error is measured based on false positive as shown in Table 3. The performance of the decision tree classifier algorithm also depends on the value of its “entropy.” Random forest has the highest accuracy among all algorithms where we have set the value of n-jobs as 1 and random state is set to 3. MLP classifier has an accuracy of 99.3% for which activation function is tanh and hidden layer size is set as 13. The last algorithm which has the lower accuracy rate among all is the Gaussian Naive Bayes. We can say that random forest has the highest accuracy with 99.7% among all the four algorithms as mentioned in Table 3.

Table 3 Comparison of different classifiers

Classifier technique	Accuracy	Error	Recall	f_1 score
Decision tree	0.999504	0.000495	1.00	1.00
Random forest	0.999797	0.000202	1.00	1.00
MLP classifier	0.993006	0.006993	0.99	0.97
Gaussian Naïve Bayes	0.945983	0.054016	0.95	0.93

5 Conclusions

In this paper, the next-generation firewall with IPS feature is discussed. It uses features from the KDD Cup dataset, and the IPS is trained with different machine learning algorithms. The experimental results uncovered that the proposed model has high accuracy, a high detection rate, and a low false-alarm rate. The proposed model compares the highest accuracy and error rate based on different machine learning algorithms. The proposed model surpasses other models in use. Based on experimental results, it can hence be suggested to have the IPS placed just after the firewall.

References

1. Bellovin SM, Cheswick WR (1994) Network firewalls. *IEEE Commun Mag* 32(9):50–57
2. Chandola V, Banerjee A, Kumar V (2009) Anomaly detection: a survey. *ACM Comput Surv (CSUR)* 41(3):1–58
3. Chou T-S, Yen KK, Luo J (2008) Network intrusion detection design using feature selection of soft computing paradigms. *Int J Comput Intell* 4(3):196–208
4. Daneshgadah S, Baykal N, Ertekin Ş (2017) DDoS attack modeling and detection using SMO. In: 2017 16th IEEE international conference on machine learning and applications (ICMLA). IEEE, pp 432–436
5. Di Pietro R, Mancini LV (2008) *Intrusion detection systems*, vol 38. Springer Science & Business Media
6. Guo C, Ping Y, Liu N, Luo S-S (2016) A two-level hybrid approach for intrusion detection. *Neurocomputing* 214:391–400
7. Jyothsna VVRPV, Prasad R, Munivara Prasad K (2011) A review of anomaly based intrusion detection systems. *Int J Comput Appl* 28(7):26–35
8. Katkar V, Bhirud SG (2012) Novel DoS/DDoS attack detection and signature generation. *Int J Comput Appl* 47(10):18–24
9. Kumar V, Sangwan OP (2012) Signature based intrusion detection system using snort. *Int J Comput Appl Inf Technol* 1(3):35–41
10. Peddabachigari S, Abraham A, Grosan C, Thomas J (2007) Modeling intrusion detection system using hybrid intelligent systems. *J Network Comput Appl* 30(1):114–132
11. Sanmorino A, Yazid S (2013) DDoS attack detection method and mitigation using pattern of the flow. In: 2013 International conference of information and communication technology (ICoICT). IEEE, pp 12–16

12. Senthilnayagi B, Venkatalakshmi K, Kannan A (2013) An intelligent intrusion detection system using genetic based feature selection and modified j48 decision tree classifier. In: 2013 Fifth international conference on advanced computing (ICoAC). IEEE, pp 1–7
13. Simmons C, Ellis C, Shiva S, Dasgupta D, Wu Q (2014) AVOIDIT: a cyber attack taxonomy. In: 9th Annual symposium on information assurance (ASIA'14), pp 2–12
14. Stallings W (2006) Cryptography and network security, 4 edn. Pearson Education, India
15. Tavallaee M, Bagheri E, Lu W, Ghorbani AA (2009) A detailed analysis of the KDD CUP 99 data set. In: 2009 IEEE symposium on computational intelligence for security and defense applications. IEEE, pp 1–6
16. Zolkipli MF, Jantan A (2010) Malware behavior analysis: learning and understanding current malware threats. In: 2010 Second international conference on network applications, protocols and services. IEEE, pp 218–221

Security Challenges During Handoff Authentication Operation for Wireless Mesh Network



Vanlalhrauia and Ajoy Kumar Khan

Abstract In wireless mesh network, a client moves in any direction outside the coverage area of its home access point and results in a handoff operation so that the next foreign access point can handle the communication as the client enter its coverage area. During handoff operation, it is important that the handoff authentication process should be performed securely between the clients and the access points. All the information exchanged during the handoff authentication process must be carried out securely without compromising the security so that a seamless communication is experienced by the client. However, due to the absence of trusted centralized authority in WMN, achieving a secure handoff authentication process is a challenging task. Enormous protocols have been proposed in the past to overcome the security challenges during the handover operation but with certain limitations. In this paper, we analyzed some of the existing protocols related to handoff authentication process and discuss the limitations present in the existing protocols.

Keywords Handoff · Authentication · Ticket · Wireless mesh network

1 Introduction

The mesh network is a type of network where the devices are connected by multiple connecting devices. Typically, there are two types of mesh networks; static or dynamic. In a static mesh network, the connecting devices are stationary, whereas in dynamic, the devices are mobile. The former type can be wired, wireless or both where the latter used portable devices to access the network such as smart phones, laptops and tablets [1]. A wireless mesh network (WMN) is scalable and flexible network architecture providing a wide coverage area. It is a promising candidate for the replacement of traditional wired networks due to its cost effectiveness [2] for deployment and maintenance. WMN provides wide range of applications such

Vanlalhrauia (✉) · A. K. Khan

Department of Computer Engineering, Mizoram University, Aizawl, Mizoram, India
e-mail: hruai_a56@yahoo.com

as a last-mile broadband network for home, enterprises, an extension of Wi-Fi and WiMAX [3] video and audio conferencing, TV and gaming [4]. They transfer data from one device to another by the concept called multi-hop. This is achieved by forwarding the data packet received to the neighbor node which is further forwarded toward the destination until it arrives at the destination node. Forwarding of the packet can be carried out through unicast or multicast routing of data or by flooding [1].

The wireless feature provides support for mobility of client within the network. A client is attached to a single access point at a time. When it moves to the area of the new access point, the handoff process occurs. Handoff within a network may cause a change in the packet routing path. Improper handoff handling can cause long delays, packet lost and interrupted service, which will degrade the performance of applications such as Voice over Internet Protocol (VoIP) and Transmission Control Protocol (TCP) throughput [2]. Traditional Extensible Authentication Protocol—Transport Layer Security (EAP-TLS) was proved to be inefficient for WMN for handoff authentication as it initiates full authentication with nine messages exchanged in multi-hop fashion [5]. It is one of the challenging tasks to develop a protocol for efficient handoff operation. The absence of centralized trusted authority and open medium of data exchange raises security challenges for WMN. It is prone to several kinds of security attacks like eavesdropping [1], unauthenticated access, malicious attack [4], etc. Illegitimate nodes external to the network may gain access to the network resources by breaking through an authentication process. So, strong authentication and access control should be placed [6] to protect an external attack.

WMN gains its popularity due to its robustness, self-healing and self-organization capability. It comprises of three types of nodes such as (i) gateway router, which forward the data packet to the Internet [7], (ii) mesh router, which forward the packets in a multi-hop fashion within the network and act as an access point and (iii) mesh client, which are non-static and access the Internet through mesh routers. Mesh routers are the backbone of WMN and are free from resource constraint. A mesh router covers a certain area and provides services for mesh clients in its coverage area. Mesh client is limited to resources constraint such as computational power, battery and bandwidth [8] (Fig. 1).

During handoff, there is a possibility of temporary interruption and illegal use of the network when a mesh client leaves its coverage area and enters to a new coverage area under the new mesh access point (MAP). Extensive works were done in the past to offer secure and efficient handoff authentication for WMN. Among them, ticket-based authentication method has taken a large step. A single authentication process during handover via ticket-based allows neighboring MAPs to share a pre-stored symmetric key. A group key authentication process during handover allows the authentication server (AS) to divide the network into multiple groups. A broadcast authentication process during handover allows AS to authenticate the clients while maintaining every MAP. The illustration for the situation where handoff operation is needed as MC is moving away from home mesh router (HMR) and toward foreign mesh router (FMR) is shown in Fig. 2, where MC is mesh client, HMR is home mesh router and FMR is foreign mesh router.

Fig. 1 Architecture of wireless mesh network

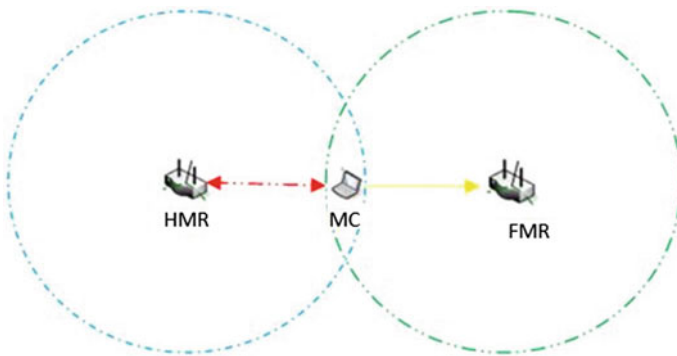
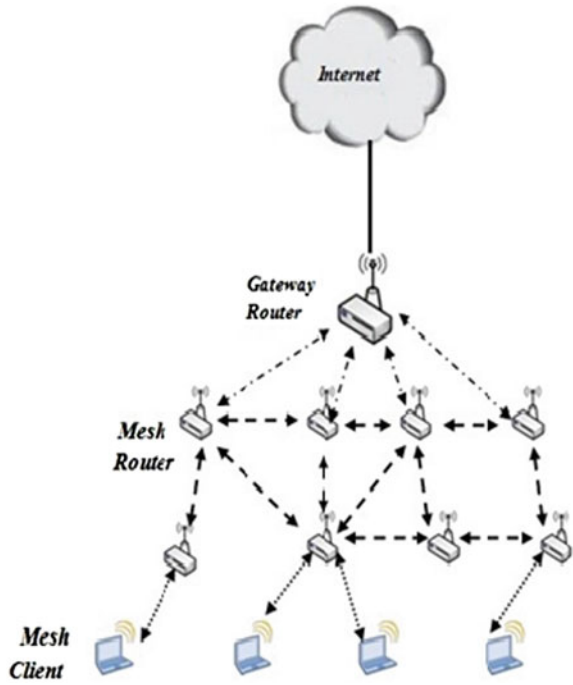


Fig. 2 Handoff scenario

2 Related Works

Several protocols have been proposed for authentication in WMN, some of which are discussed in this section with their limitations.

In 2013 [9], ticket-based efficient authentication for fast handoff in WMN was proposed. In this paper, after successfully login, HMR sends message containing it's ID, Client ID, MAC key, PMK (between HMR and client) to all its neighbor in

encrypted form for future handover for the client. Each neighboring MR decrypts the message with their own private key. When a client moves into the area of foreign mesh router, it sends a message containing transfer ticket $\theta_c = \{\mu, \text{VK}_{\text{MAC}}(\mu)\}$, where $\mu = \{I_R, I_c, I_A, T_{\text{exp}}, \text{MAC}_{\text{Algo}}\}$, nonce N_c and $\text{VK}_{\text{MAC}}(N_c)$. A foreign mesh router (FMR) verifies the correctness of MAC key and validity of the ticket, if the verification is success FMR generates a nonce N_R and sends $\text{VK}_{\text{MAC}}(N_C, N_R)$ to a client. When a client receives this message, it computes a MAC value $\text{VK}_{\text{MAC}}(N_C, N_R)$. If $\text{VK}_{\text{MAC}}(N_C, N_R) = \text{VK}_{\text{MAC}}(N_C, N_R)$, a FMR is authenticated by the client. A client sends $\text{VK}_{\text{MAC}}(N_R)$ as an authentication challenge to FMR, if the verification is succeeded FMR authenticate the client. If all the above steps are carried out successfully both the client and the FMR authenticate each other in handoff. The above protocol is secure against several types of attacks like replay attack, forgery attack, DoS attack, etc. But the drawback of the protocol is distribution of the transfer ticket. The ticket is generated for each authenticated client whether the client needed it or not and the client kept the ticket during the communication session and the neighboring MAP also keep the ticket for possible handoff. This can cause MAP keep several unnecessary ticket which cannot be discarded until timer expired.

In 2016 [10], ticket-based handoff authentication protocol was proposed. In this protocol, it is assumed that mesh routers established symmetric key with each neighbors. After successful login, the HMR generates temporary MAC key (K'_{MAC1}) for each neighbor and pseudo ID(I_{C1}) for client, then it sends transfer ticket(θ_C 's), K'_{MAC1} and I_{C1} encrypted with symmetric key to the corresponding neighbor. Each neighboring mesh router decrypts the message to prepare for handoff authentication in the future. When the roaming client enter the area of FMR, it calculate K_{MAC1} and I_{C1} by itself, then it submits I_{C1} , $H(\theta_C || I_{C1})$, N_{C2P} and the MAC value $\text{VK}'_{\text{MAC1}}(I_{C1}, H(\theta_C || I_{C1}), N_{C2P})$ to the foreign MAP(θ_C and N_{C2P} are transfer ticket and random nonce, respectively). A FMR checks the correctness of $H(\theta_C || I_{C1})$ and K'_{MAC1} , then it generates formal MAC key from temporary MAC key and new Nonce. Two more messages are being exchanged one for MAC key generation for a client and verification for FMR. This protocol suffers from the same problem stated as in the previous paper. Beside that the random Nonce is generated from the elliptic curve points that need elliptic curve point multiplication that cause extra burden and curve parameter to consider.

In 2016 [11], fast handoff technique for WMN was proposed. In the proposed protocol, several number of symmetric keys are maintain for authentication (i.e., server-mesh client, server-mesh routers, mesh routers-mesh clients), a group master key is also maintain for the communicating mesh routers. The ticket is generated by server and distributed to the mesh routers based on the group master key. When the mesh client enters the area of FMR, it requests a ticket from the server. The received ticket from the server is submitted to FMR, then FMR matches both the tickets which are from the client and HMR. If the matching is success client is authenticated. In this protocol, the ticket is send from server at the time of handoff request, this request message and the ticket may travel multiple hops depending on the distance between the client and server which can cause significant delay.

In 2019 [12], handoff with reduction delay was proposed. In this protocol, a secure channel is established via master key between mesh client and authentication server, a group master key is generated for mesh routers which are one hop away from each other. An authentication server generates transfer ticket for the client and keep it ready to forward the ticket upon request. When a mesh client enters FMR, it sends a handoff request to HMR, the HMR forwarded the request. Then, authentication server sends the handoff ticket and subsequently sends to FMR. A client also requests for ticket to authentication server. A roaming mesh client submits the ticket to FMR for authentication, if the ticket from the client and from the HMR are matched, the client is authenticated otherwise rejected. This protocol suffers from handoff delay due to multi-hop transfer of ticket and key management overhead.

In 2020 [13], ID-based handoff with Diffie-Hellman was proposed. In the proposed protocol, a ticket is transferred in encrypted form, this causes a computational overhead at HMR and FMR as the ticket is to be encrypted and decrypted at the sender side and receiver side, respectively.

In 2020 [14], authentication with privacy preservation was proposed. In the proposed protocol, HMR sends message authentication code (MAC) key to FMR in encrypted form for verification of the ticket. This causes computational overhead and each time the client moves out of the currently serving mesh router, a new MAC key has to be generated.

In 2019 [15], probabilistic-based authentication was proposed. In the proposed protocol, an AS responsible for generating and distributing tickets to other entities, various types of trust is maintained between the network entities such as

AS-HMR: Trust based on group master key

MR-MR: Trust based on group master key

MR-MC: Trust based on ticket

AS-MC: Trust based on pair wise master key

In the local verification step, mesh client and mesh router authenticate each other by exchanging messages containing their IDs and $\text{Sig}_{\text{server}}$. After successful authentication of client, AS will generate a handoff ticket for every client and distribute it randomly to each mesh router. Only one ticket is given to each mesh router. If there are 1000 mesh clients and 100 mesh routers in the network, only 10% is distributed. A handoff process is triggered by signal strength.

$$\text{Signal strength} = \text{Distance}/\text{SNR}$$

After finding good signal strength by mesh client, it requests HMR for handoff ticket to the HMR, then HMR forwards the request to the AS. The AS sends the ticket to the HMR and the targeted FMR. The HMR sends the ticket to the mesh client, then the mesh client submits the ticket to the FMR. If the ticket received from mesh client and AS are matching, then FMR authenticates the mesh client and store the ticket for future authentication of the same client. However, the success rate of one-hop authentication is too low as there are only 10% tickets distributed. In 90%, the ticket is requested from the AS, which will cause significant delay.

3 Security Requirements

The true capacity of WMN cannot be accomplished without considering the involved security issues. The security requirements are similar to other types of networks which are listed below:

- (1) **Confidentiality:** When the information is exchanged between the sender and receiver, then the information should be revealed only to the intended receiver.
- (2) **Integrity:** All the contents of the information exchanged between the sender and receiver should be preserved from modification.
- (3) **Authentication:** Both sender and receiver should authenticate each other before initiating the information exchange.
- (4) **Access Control:** Only the authorized entities are permitted to access the network.
- (5) **Non-repudiation:** Neither sender nor receiver can deny later that they have not initiated the information exchange.
- (6) **Availability:** Only the authorized actions are allowed to initiate between the sender and receiver.

4 Attacks on Wireless Mesh Network

Several kind of attacks can be launch in WMN, it can be from external to the network where an attacker does not belong to the network or from the compromised node within the network. Different types of attack are listed below:

- A. Impersonation: Over the insecure network, the intruder tries to own the identity of the legitimate node and acts on behalf of it.
- B. Flooding attack: The intruder forces the legitimate node to drain all its resources unnecessarily by sending enormous data toward the target node.
- C. Blackhole attack: Malicious node over the insecure network convinces its neighboring node to forward the packets through malicious node as a shortest path and drops all the packets.
- D. Wormhole attack: It involves two malicious nodes which form a fake shortest tunnel to convince the legitimate nodes that the packets can be forwarded through the shortest path via the tunnel and results in packet drop and delay.
- E. Sleep deprivation: An attacker request services from a victim over and over again, so it cannot go into an idle that will drain battery.
- F. DoS and flooding: DoS attacks are launch by flooding a victim node so that it cannot perform an useful work as the victim node had to drop legitimate data packet.
- G. Greedy attack: This attack forces the legitimate node to utilize all its shares unnecessarily for a repeated transmission while violating the rules of standard routing protocol. As a result, the node wastes all its power and bandwidth.

- H. Sybil attack: In this attack, the intruder owns the identity of multiple legitimate nodes and can perform action in more than one place at the same time.
- I. Eavesdropping: An attacker listen data in transit to discover information about the location of the nodes and their relation and about the network topology.

5 Analysis of Existing Protocol

Some of the proposed protocols have been analyzed with certain parameters which is shown in Table 1.

Table 1 Comparison of related works

Protocol	Authentication process	Privacy	Transfer ticket generation	Authentication delay	Parameter used
EAFH [9]	One hop	YES	HMR	Low	Computational and communication cost
FHT [11]	Multi hop	YES	AS	High	Computational cost and handoff latency
NEHA [10]	One hop	YES	MR	Low	Computational and communication cost
SHRAD [12]	Multi hop	YES	AS	High	Authentication delay, false authentication and correct authentication
HATD [13]	One hop	YES	MR	Low	Handoff delay and computational cost
APPP [14]	One hop	YES	MR	Low	Computational cost, authentication delay, number of message exchange
PHV [15]	Multi hop	YES	AS	High	Authentication delay, false authentication and correct authentication

6 Challenges

Mesh client is usually small device which is having limited computational power and limited power supply as it is battery operated. Besides this the mobility and the limited bandwidth of the device imposed the following challenges to be addressed without compromising security issues for handoff in WMN.

- (a) **Computational Cost:** The amount of cryptographic computation that involves in handoff should be minimal, it is desirable to employ hash or MAC algorithm rather than symmetric or asymmetric cryptography due to the lower computational overhead especially in client side.
- (b) **Message Exchange:** The number of message exchange for handoff should be minimal, increase in message exchange impose handoff delay, not only this the hop count also need to be considered if a hop count is comparatively high the moving mesh client may move out of the coverage area of home mesh router that will cause repetition of login authentication causing interruption in an ongoing communication session.
- (c) **Key Management:** Cryptographic operation usually required a key(s), some may be public other may be private or secretly shared. It may be pair wise or group key, these keys must be generated and distributed in an efficient manner to consume minimal bandwidth and storage space.
- (d) **Storage:** Keeping redundant data at the mesh entities consumes device resources, so it should be avoided, unnecessary data should not be kept at the mesh entities for long period of time.

7 Conclusion

In this manuscript, we have highlights all the issues and problems associated with the handover operation. In wireless mesh networks due to the absence of centralized authority, it had been found that achieving security in this network is a challenging task. Therefore, we have addressed all the issues that had been remained in the existing protocols. In the future, we aim to propose new methods which can resolve the issues during handover operation.

References

1. Nanda A, Nanda P, He X, Puthal D, Jamdagni A (2018) A novel hybrid authentication model for geo location oriented routing in dynamic wireless mesh networks. In: International conference on system sciences
2. Wei HY, Kim S, Ganguly S, Izmailov R (2006) Seamless handoff support in wireless mesh networks. In: 2006 1st workshop on operator-assisted (wireless mesh) community networks, pp 1–8

3. Alamri N, Chow CE, Aljaedi A, Elgzil A (2018) UFAP: Ultra-fast handoff authentication protocol for wireless mesh networks. In: 2018 wireless days (WD), pp 1–8
4. Sharma PK, Mahajan R (2017) A security architecture for attacks detection and authentication in wireless mesh networks. *Clust Comput* 20(3):2323–2332
5. Xu Q, Wan C, Hu A (2008) The performance analysis of fast EAP re-authentication protocol. In: 2008 international symposium on computer science and computational technology, vol 1, pp 99–103
6. Roy AK, Khan AK (2019) Efficient authentication and key management scheme for wireless mesh networks. *Int J Internet Technol Secured Trans* 9(1–2):184–200
7. Karthika KC (2016) Wireless mesh network: a survey. In: 2016 international conference on wireless communications, signal processing and networking (WiSPNET), pp 1966–1970
8. Sgora A, Vergados DD, Chatzimisios P (2016) A survey on security and privacy issues in wireless mesh networks. *Security Commun Netw* 9(13):1877–1889
9. Li C, Nguyen UT, Nguyen HL, Huda N (2013) Efficient authentication for fast handover in wireless mesh networks. *Comput Secur* 37:124–142
10. Lai YM, Cheng PJ, Lee CC, Ku CY (2016) A new ticket-based authentication mechanism for fast handover in mesh network. *PLoS ONE* 11(5):e0155064
11. Rathee G, Saini H (2016) A fast handoff technique in wireless mesh network (FHT for WMN). *Proc Comput Sci* 79:722–728
12. Rathee G, Saini H (2019) Secure handoff technique with reduced authentication delay in wireless mesh network. *Int J Adv Intell Paradigms* 13(1–2):130–154
13. Roy AK, Khan AK (2021) Handoff authentication through Diffie-Hellman approach for wireless mesh networks. In: Proceedings of the international conference on computing and communication systems, pp 615–622
14. Roy AK, Khan AK (2020) Authentication protocol with privacy preservation for handover in wireless mesh networks. In: First international conference on sustainable technologies for computational intelligence, pp 383–395
15. Rathee G, Jaglan N, Saini H, Gupta SD, Kanaujia BK (2019) Probabilistic verification scenarios with reduced authentication delay for handoff clients in mesh networks. *Wireless Pers Commun* 104(4):1553–1571

Implementation of Photon Hash Function on FPGA



W. Heera, K. Bhagyashree, Rohan Patil, Nalini Iyer, and Shraddha Hiremath

Abstract Every day, a large amount of data is transmitted through the Internet. It is essential that a message be sent to the recipient directly from the sender through the network without the interference of a third party. Hash algorithms can be used to verify the sender and receiver's identities. Hash algorithms are used to ensure the sender and receiver's validity and secrecy. The photon algorithm is used to construct a lightweight hash algorithm. When compared to other hash algorithms, the method has an extremely tiny footprint. The report includes the details of the implementation of proposed hash algorithms on the ARTY A7 FPGA board using Vivado 2018.1. The proposed photon algorithm is described in detail, as well as the results produced. When compared to other photon variations, the 80/20/16 variant has the lowest footprint, making it ideal for resource-constrained devices. Our proposed design of photon hash algorithm uses the design of photon 80/20/16 and includes a custom design of standard 8-bit AES S-box. The simulation of our algorithm provides the implementation for very small area-constrained devices with high rate of hash security.

Keywords Lightweight cryptography · Hash function · Sponge construction · Hash value · Photon

W. Heera (✉) · K. Bhagyashree · R. Patil · N. Iyer · S. Hiremath
School of Electronics and Communication Engineering, KLE Technological University, Hubli
580031, India
e-mail: heerawali@kletech.ac.in

K. Bhagyashree
e-mail: bhagyashree@kletech.ac.in

R. Patil
e-mail: rp.vpatil09@gmail.com

N. Iyer
e-mail: nalinic@kletech.ac.in

S. Hiremath
e-mail: shraddha_h@kletech.ac.in

1 Introduction

The Internet of Things (IoT) provides a framework for the safe and friction-less flow of data and instructions across constrained devices. Many IoT security techniques are unsuitable for constrained devices. Because constrained devices do not have a consistent source of power, they are unable to perform big and complex algorithms. Many researchers are working on security measures for constrained devices based on lightweight cryptography right now. It is challenging for security developers where private information could be leaked and modified with these resource-constraint devices. It is very essential to include security measures in communication aspects of IoT devices. The two essential elements of lightweight cryptography are random number generation and hash design. The cryptographic methods and high-level authentication security to avoid third-person access require high-processing hash functions do not suit such small devices. As a result, there were several lightweight authentication designs were introduced [1–4] and implemented on hardware and software, where some of the designs were compact and efficient on hardware and some on software [5–10]. Photon hash function by Guo et al. [4] is one of the lightweight cryptographic algorithms. Hash functions are mathematical operations that take arbitrary length of input and provide fixed length output as hash value [11]. The photon hash function is compact in hardware and efficient in its software implementation. The photon hash function is sponge-based lightweight cryptographic hash. The working design of photon is combination of sponge construction in absorption phase and AES like permutation in permutation phase which can be design with tiny area utilization for hardware [12].

In this paper, a new approach to increase the confusion and diffusion of photon is proposed. The new design addresses the uniquely modified S-box. The proposed design uses the standard 8-bit AES S-box in addition to the existing design. The methodology of this implementation increases the chance of occurrence of collision and re-image of the algorithm. This technique successfully improves the overall security of the photon hash function in terms of an increase in confusion and diffusion giving more resistance to collision and re-imaging.

2 Related Work

Our work focuses on the FPGA implementation of the photon hash function, a lightweight cryptographic method. Survey of lightweight cryptographic functions [14] and study on quark and photon [13] are two related research projects. Three forms of architectural structures, namely sponge construction, Merkle–Damgard construction, and Davies–Meyer construction, are briefly discussed in [14]. In our work, we focus on sponge construction. The data is absorbed in the sponge architecture; the mathematical functions are run, and the hash value is returned as an output when the sponge is squeezed. Photon, spongent, Keccak, MAC and HMAC, among others,

Table 1 Abbreviations

Notation	Description
n	The bit size of the hash output
t	The bit size of the internal state
c	The bit size of the capacity part of internal state
r	The bit size of the bitrate part of input and internal state
d	The number of cell columns and rows in the internal permutation
r'	The bit size of the bitrate part of output
s	The bit size of one cell in the internal permutation

Table 2 Variants of Photon

	r	r'	c	s	d	t
Photon-80/20/16	20	16	80	4	5	100
Photon-128/16/16	16	16	128	4	6	144
Photon-160/36/36	36	36	160	4	7	196
Photon-224/32/32	32	32	224	4	8	256
Photon-256/32/32	32	32	256	8	6	288

are some of the lightweight hash functions described in the paper [13], compares the quark and photon lightweight hashing algorithms and claims that neither of them has been broken yet. Software optimization is lacking in quark, but the hardware implementation is more efficient and is chosen for IoT application. When compared to other hash algorithms, photon and quark have the most efficient performance. Photon is known to be most secure and smallest among lightweight hash algorithm.

3 Photon Hash Function

3.1 Photon Hash Function

The photon hash function was designed with the main objective of being lightweight and compact in size. The domain extension algorithm is based on sponge construction and has five variants with hash sizes of 80, 128, 160, 224, and 256 bits, as shown in Table 1.

In our work, we use photon 80/20/16 variant. The algorithm is broken down into three stages: initialization, absorption, and squeezing. The sponge construction is as shown in Fig. 1. The arbitrary input is padded until the message length is in multiple of ‘ r ’ bits and then separated into blocks of r bits during the initialization phase (Table 2)

$$\text{pad}(m) = m || 1 || 0^k \tag{1}$$

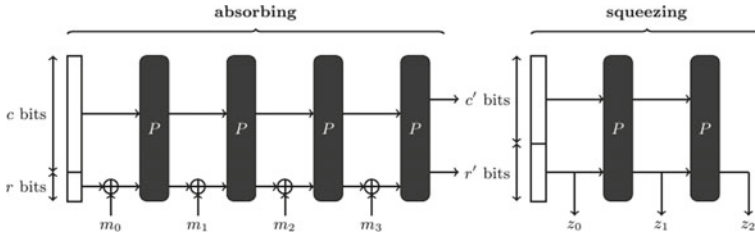


Fig. 1 Sponge construction

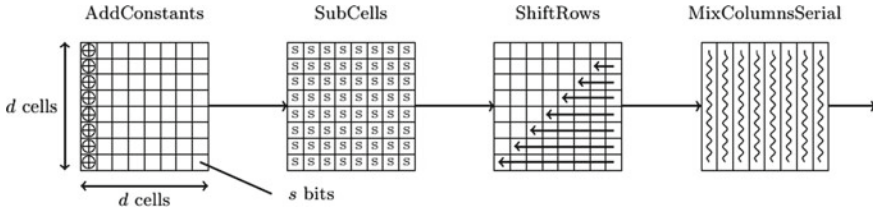


Fig. 2 Internal permutation of sponge construction

The padded message is absorbed in the absorption phase by XORing the first ‘ r ’ bits of the initialization vector (IV) of total size ‘ t ’ bits with the r -bits of the message and keeping the remaining bits of IV to complete the absorption.

$$IV = 0^{(t-24)} || n/4 || r || r' \tag{2}$$

The internal state is passed to permutation block. The permutation is based on rounds. The permutation of photon is very much like AES and comprises of four modules: AddConstants, SubCells, ShiftRows, MixColumns as shown in Fig. 2 Add-Constant: The internal state of the absorption or previous round’s MixColumns is XOR-ed with the round constant (RC) generated by the linear feedback shift register (LFSR) and a pre-defined internal constant (IC). The RC is set to a fixed constant and updated using LFSR every round. This procedure will only affect the first column of internal state, leaving the remaining columns unaffected. Equation gives the updated state, where $h[i, j]$ is the current state, i is the row number, j is the column number, and N is the the round number. The equation for AddConstant can be given as,

$$\begin{aligned} h'[i, 0] &= h[i, 0] \oplus RCN(i) \oplus IC(i) \text{ for } j = 0 \\ h'[i, j] &= h[i, j] \text{ for } 0 < j < d; \end{aligned} \tag{3}$$

SubCells: Photon’s second function is to do cell substitution. Because $s = 4$ bits in photon 80/20/16, the present S-box is utilized. Each state cell is replaced with a corresponding nonlinear S-box cell. In a later part, we will discuss how our work contributed to the introduction of a new S-box design. SubCell’s overall function is described as

$$h'[i, j] = \text{Sbox}(h[i, j]) \quad (4)$$

ShiftRows: This function is nearly equivalent to AES's shift rows function. All the rows of the state are rotated to left by i cells, where i is row index and starts from 0. The first row is not updated in this case, but the remaining rows are. ShiftRows has the following equational notation:

$$h'[i, j] = h[i, (i + j) \bmod d] \quad (5)$$

MixColumns: It is a maximum distance separable (MDS) matrix-based finite field multiplication. Internal state columns are multiplied individually with a pre-defined matrix based on the dimension size d . The architecture of MC for PHOTON is similar to that of AES, but it focuses on consuming the least amount of space possible. For photon 80/20/16, Galois field with the irreducible polynomial $x^4 + x + 1$ for $\text{GF}(2^4)$ [12] is used for matrix multiplication for the MC operation.

The absorption phase ends with the completion of the permutation block. The squeeze phase would come next. The internal state is compressed to yield r' bits output before being transmitted via the permutation block. Iteratively, this operation is repeated until the total overall output is n bits. These n bits value is known as hash value. The squeeze phase can be equated as

$$z[i, j] = h[i, j - 1] \quad (6)$$

The overall architecture design for photon 80/20/16 is shown in Fig. 3.

3.2 Proposed S-Box Design for Photon

S-boxes in cryptography offer confusion by entirely substituting the block of data with the nonlinear value of the S-box, whereas diffusion in cryptography is achieved by shuffling the bits of data inside the block.

The original S-box architecture for photon 80/20/16 is based on a 4-bit present S-box substitution technique, in which a 20-bit ' r' '-size block is passed through a 4-bit S-box, which takes 4-bit input and outputs 4-bit, five times for a total of 20 bits. This process is carried out in parallel. The new S-box design transformation works with two different types of S-boxes. The first type is an 8-bit AES S-box that was added to the design, while the second type is the original 4-bit S-box. This proposed photon design increases the amount of confusion and diffusion.

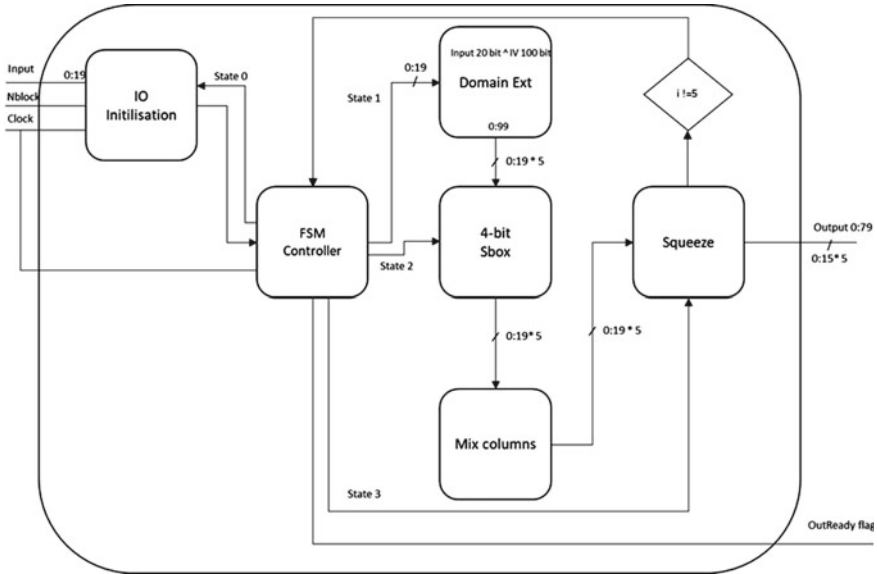


Fig. 3 Data path of Photon 80/20/16

While the other processes in the proposed photon architecture stay the same, the functioning of the S-box is as follows:

1. First, the 8-bit AES S-box receives the modified internal state from AddConstant.
2. The 20-bit block is split into three parts: two 8-bit blocks (19:12 and 11:4), and one 4-bit block (3:0). The 8-bit S-box is used to control the two 8-bit blocks, while the 4-bit block is left unoperated.
3. The output of 8-bit S-box plays a major role in improving confusion and diffusion of the design.
4. The first 8 bits (19:12) of the output of an 8-bit S-box are the substituted values of the first sub-block, the following 4 bits (11:8) are the unoperated 4-bit sub-block, and the final 8 bits (7:0) are the second substituted value of the second 8-bit sub-block. Substitution and permutation are both used in this way to generate confusion and diffusion. The simple illustration of 8-bit S-box is shown in Fig. 4.
5. The output of this block of 8-bit S-box is fed into the current 4-bit S-box, which continues to operate in regular manner. The block illustration of existing 4-bit S-box is shown in Fig. 5.

The proposed photon architecture is used to produce more confusion and diffusion, which also in turn helps to reduced the occurrence of collisions and pre-image of the algorithm.

The overall architecture design for photon 80/20/16 is shown in Fig. 6.

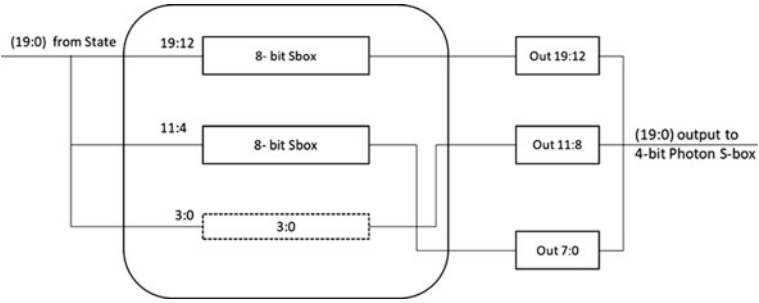


Fig. 4 Design proposed using 8-bit S-box

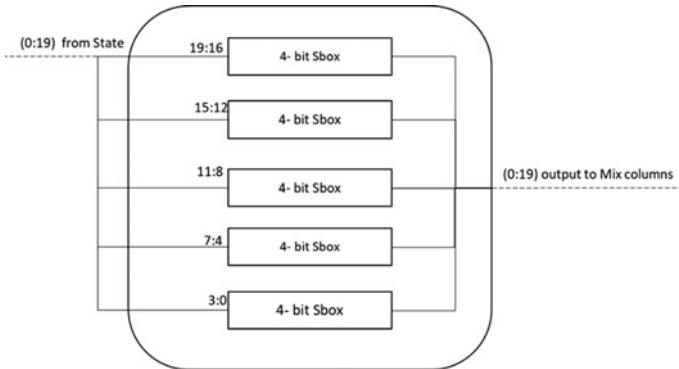


Fig. 5 4-bit S-box of Photon 80/20/16

4 Implementation

The photon 80/20/16 hash function and proposed photon hash function were implemented on FPGA devices. The device was configured using Xilinx Vivado and simulated in Xilinx Vivado, focusing the target board as DIGILENT ARTY A7-35T. The photon 80/20/16 follows serialized architecture focusing on low-area optimization and low-processing devices. The proposed photon hash function also follows serialized architecture but takes more resources as there is an addition of 8-bit AES S-box. The output for the algorithm is received after 137 clock cycles. The algorithm is programmed in such a way that it follows finite state machine. The LFSR is used to update round constant round wise, control counter in FSM. At State 0, the algorithm is in idle state, State 1, State 2, and State 3 performs operations, and at State 3, we get the output hash value. The representation of FSM is shown in Fig. 7.

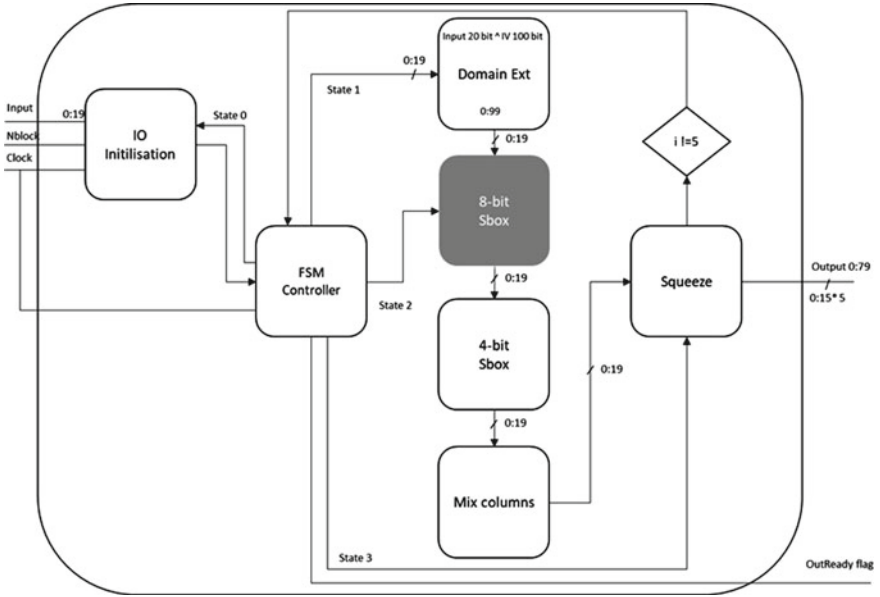


Fig. 6 Architecture of proposed photon hash function

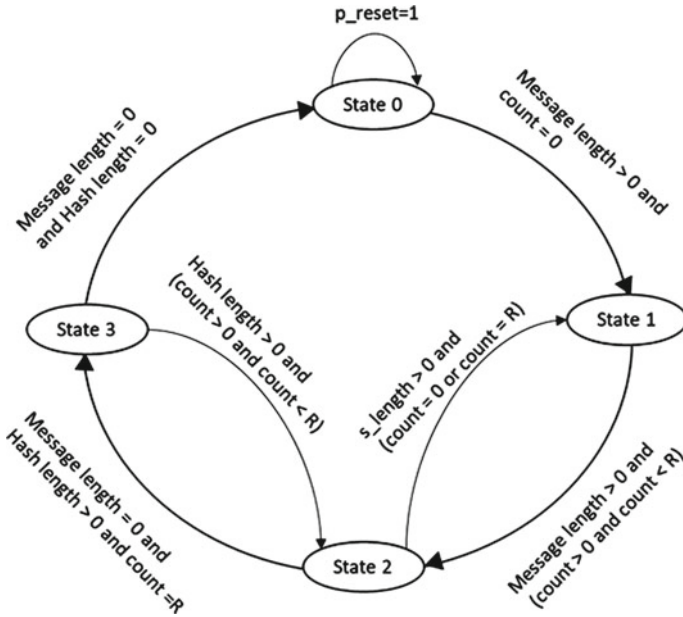


Fig. 7 FSM for photon hash function

Table 3 Test vector

Design	IV	m	$P(m)$
Photon 80/20/16	0 0 0 0 0	0 0 0 0 0	3 3 D 5 F
	0 0 0 0 0		6 2 9 B 9
	0 0 0 0 0		5 C 4 8 1
	0 0 0 0 0		6 5 C E 7
	0 0 0 0 0		B 7 7 0 C

Table 4 Resource utilization of Photon 80/20/16 and proposed Photon design

Design	Hardware	LUT	Slices	Flip flop	Max OpFreq (MHz)	Frequency (MHz)	Clock cycles
Photon 80/20/16	Arty A7 35T	156	39	132	305	72.99	137
Proposed Photon design	Arty A7 35T	233	71	132	291	72.99	137

5 Results and Discussion

In this section, the simulation results will be discussed. The implementation of photon 80/20/16 and proposed photon architecture are implemented on Xilinx FPGA board (ARTY A7-35T) is used to test the functionality and compare the results. The architecture contributes in providing low-area utilization, and additional 8-bit S-box contributes in securing the algorithm by increasing confusion and diffusion of the system which also helps to reduced collisions and re-imaging of the algorithm. Figure below represents the algorithmic flow of photon hash function. The results of photon 80/20/16 and proposed photon architecture are simulated using Xilinx Vivado 2018.1. The simulated output for photon is validated using test vector [15] shown in Table 3.

Table 4 depicts the resource utilization for the photon 80/20/16 and proposed design of photon hash function, mentioning the number of LUTs and slices.

The proposed design’s performance in terms of resource utilization for photon 80/20/16 and proposed photon design is compared with ‘FPGA-Based Lightweight Hardware.’

Architecture of the PHOTON Hash Function for IoT Edge Devices’ [12] work as shown in Table 5.

The output of photon 80/20/16 and proposed photon architecture is simulated for input ‘00000,’ and the results of the following results are shown in Figs. 8 and 9.

Table 5 Performance comparison of FPGA implementations of lightweight cryptographic

Design	Hardware	LUT	Slices	Flip flop	Max OpFreq (MHz)	Clock cycles
Photon 80/20/16	Arty A7 35T	156	39	132	365	137
Proposed Photon design	Arty A7 35T	233	71	132	291	137
Related work	Arty A7 100T	363	145	188	373.43	60
Related work	Spartan-3 S50	510	265	200	262.07	60

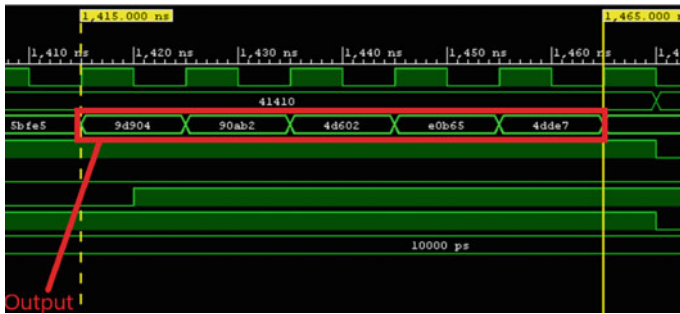


Fig. 8 Simulation result for Photon 80/20/16

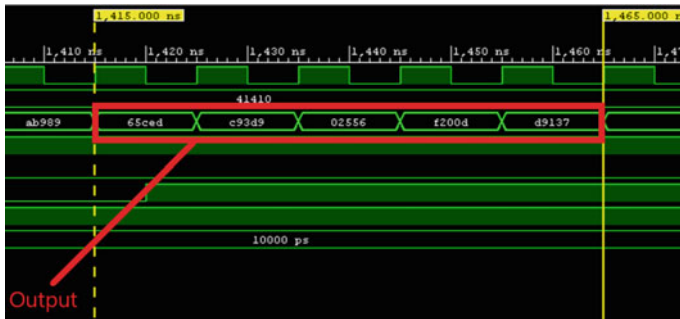


Fig. 9 Simulation result for proposed photon design

6 Conclusion and Future Scope

6.1 Conclusion

On the Arty A7 35T target platform, the proposed design is simulated, synthesized, and implemented. On the Arty A7 35T, the chosen design for the lightweight hash function photon for the version 80/20/16 and the proposed design for the lightweight

hashing algorithm photon for the revised design are implemented. Both the photon and proposed photon designs were built within board limitations. Before the design was integrated with 8-bit AES S-box, the output from photon 80/20/16 was evaluated using test vectors to ensure that it had a lower footprint, efficient throughput, and resource efficiency. When compared to Mohammed Al-Shatari's work [2], the area in terms of slices and flip flops is much smaller. The suggested architecture includes an extra standard 8-bit AES S-box, which handles data substitution and block permutation. With this change, the algorithm's confusion and diffusion increases, while the possibility of collision and pre-image occurrence decreases dramatically, making it less vulnerable to cyber attacks. Hence, with less resource utilization and higher range of collision and pre-image of output, the proposed algorithm can be used for securing the resource-constraint IoT edge devices.

6.2 Future Scope

The proposed photon architecture features a bit rate of 20 for input and 16 for output. The internal state is 100 bits, and the output hash size is 80 bits. There is an extra 8-bit conventional AES S-box with this, which adds to the algorithm's confusion and diffusion. The method has become more robust to collisions and pre-images since the introduction of the 8-bit S-box. Photon hash is one of the most secure lightweight hash functions known to date. The proposed design may be utilized in a wide range of applications that require greater security, such as low-power high-speed RF ID tags in military applications, signal authentication, and high-speed communication via the controller area network (CAN Bus) in industrial and automotive applications.

References

1. Bogdanov A, Knežević M, Leander G, Toz D, Varc K, Verbauwhede I (2011) SPONGENT: a lightweight hash function. In: Proceedings of international workshop on cryptographic hardware and embedded systems. Springer, Berlin, pp 312–325
2. Bertoni G, Daemen J, Peeters M, Van Assche G (2009) Keccak sponge function family main document. Submission NIST 3(30):320–327
3. Aumasson J-P, Henzen L, Meier W, Naya-Plasencia M (2010) Quark: a lightweight hash. In: Proceedings of international workshop on cryptographic hardware and embedded systems, pp 1–15
4. Guo J, Peyrin T, Poschmann A (2011) The PHOTON family of lightweight hash functions. In: Proceedings of annual cryptology conference, pp 222–239
5. Anandakumar NN, Peyrin T, Poschmann A (2014) A very compact FPGA implementation of LED and PHOTON. In: Proceedings of international conference on cryptology, pp 304–321
6. Jungk B, Lima LR, Hiller M (2014) A systematic study of lightweight hash functions on FPGAs. In: Proceedings of international conference reconfigurable computing FPGAs, pp 1–6
7. Kahri F, Mestiri H, Bouallegue B, Machhout M (2016) High speed FPGA implementation of cryptographic KECCAK hash function cryptoprocessor. *J Circ Syst Comput* 25(4), Art. no. 1650026

8. Aerabi E, Bohlouli M, Livany MHA, Fazeli M, Papadimitriou A, Hely D (2020) Design space exploration for ultra-low-energy and secure IoT MCUs. *ACM Trans Embedded Comput Syst* 19(3):1–34 Jul
9. Martino R, Cilaro A (2020) SHA-2 acceleration meeting the needs of emerging applications: a comparative survey. *IEEE Access* 8:28415–28436
10. Alzahrani A, Gebali F (2018) Multi-core dataflow design and implementation of secure hash Algorithm-3. *IEEE Access* 6:6092–6102
11. Abbas YA, Jidin R, Jamil N, Z'aba MR, Mohamed MA (2018) Photon: a new mix columns architecture on FPGA. *Int J Eng Technol*
12. Al-Shatari M, Hussin FA, Witjaksono AAAG, Tran X-T (2020) FPGA-based lightweight hardware architecture of the PHOTON hash function for IoT edge devices. *IEEE Access* 2020.3038219
13. Gupta DN, Kumar R (2021) Sponge based lightweight cryptographic hash functions for applications. In: *International conference on intelligent technologies*
14. Hammad BT, Jamil N, Rusli ME, Z'aba MR (2017) A survey of lightweight cryptographic hash. *Int J Sci Eng Res*
15. Guo J, Peyrin T, Poschmann A (2011) The PHOTON family of lightweight hash functions. In: *Proceedings of annual cryptology conference*, pp 222–239

Predictive Generation Trajectory Clustering Model for Google Online Learning: OGMM



Vishnu Kumar Mishra, Megha Mishra, Sameera Khan, Sheela Verma,
and Ashish Kumar Tamrakar

Abstract The learner-oriented teaching methodology through which the familiarity of cracking an undetermined problem which is restricted for stimulated content is performed. Learners gain knowledge of a subject in a deep manner. The PBL method does not concentrate for explaining the problems by means of a given answer, other than it allows additional desirable abilities and attributes to be shaped. This engages the achievement of details, improved synchronization and link the assembly and communication. Pedagogic systems have been exaggerated universally by the COVID-19 pandemic, and this pandemic is responsible to the complete shutting down of schools, universities and colleges. Online learning comprises courses delivered 100% interactive by postsecondary institutions, except massively open online courses (MOOCs). Compared to conventional courses that work as a brick-and-mortar devoting for school house, online learning or virtual classes offered over the Internet. The approach helps learners to improve the abilities and turns it to opportunity by doing practice. It improves condemnatory assessment, retrieval of anthologized and facilitates continuous knowledge surrounded by a team. This study also explores the downside of PBL by using the online platform with its benefits.

Keywords Massively open online courses (MOOCs) · GTM · OGMM · Wireless communication and problem-based learning

V. K. Mishra (✉) · M. Mishra · S. Verma
SSTC-FET, Bhilai, Chhattisgarh, India
e-mail: vshn123mshr@gmail.com

S. Khan
Verdhaman College of Engineering, Hyderabad, India

A. K. Tamrakar
BIT, Raipur, India

1 Introduction (OGMM)

Wireless communication now provides advanced features of artificial intelligence, and sensing has created enormous patterns for the creation of Google-meet monitoring [1–4]. Multiple biosensors have been used to continuously track and record personal meeting records, video records and data, interact with remote Google workplace service centers' via wireless networks and also help meet metrics, locate and debug meetings, display network statistics (jitter, packet loss and congestion). The online Google-meeting monitoring (OGMM) provides various applications for evaluation of record of individuals and in group [5, 6]. There is a lot of privacy data and data obtained by OGMM applications in the personal meeting assistant.

2 Literature Study

The common apps for social networks, even credit card transaction records and ubiquitous Internet of Things (IoT) applications, used in real life problems [7–11]. To distinguish users by means of specific course, they depict sufficient surrounding information for cyberpunk. They are also threatened with Zoom and Doodle Sensitive personal data in this intractable situation (e.g., home address and final information). The OGMM service providers never collect and assemble the identity information of the customer.

The likelihood to differentiate the trajectory is known as the accuracy of probability that use for re-identification. This research work analyzes the course differentiation jeopardy that is proportional toward the individuality of OGMM, especially; the focus is the percentage of particular course. An anonymous customer through a specific course for specified data-records is also able to differentiate by connection attacks, as shown in Ref. [12]. As verified by a 15-month mobile device trajectory review in Refs. [12], 95% of users can uniquely identify only four positions of spatiotemporal in a trajectory. Although academia and industry have frequently cited the above mentioned conclusion, and it is vague whether that summary is indeed appropriate for additional relevance circumstances, for example the GOMM scenario. The uniqueness of trajectories would also influence the quantity of solitary and the spatiotemporal concentration of solitary in cellular phone interaction circumstances, overlooks the difficulty and variability of cell phone learning surroundings [13–17]. For example, no research has shown that the assumption is also valid in cities where small cells are densely deployed [18, 19]. Inside this article, we recommend the Google Trajectory Mock-up (GTM), an isolated jeopardy assessment replica of online user trajectory. It is applied for analytical extraction trajectory distinctiveness, commencing cumulative statistics widely published via OGMM service contributor. And normally envelop the number of models of trajectory risk assessment. In particular, the GTM can be used to define trajectory privacy threats for OGMM situations, users plus isolated advocates. In particular, the GTM can be used to define trajectory

privacy threats for OGMM situations, users and privacy advocates. Following is the overview of remaining article.

Inside Sect. 3, estimated, related and connected effort is included. The implementation of GTM covered in Sect. 4. And the statistical trajectory uniqueness characteristics of hitting the leaning area are formulated in Sect. 5, (e.g., assumptions and variances) experimental results are given. The paper eventually finishes with Sect. 6.

3 Edges of OOMG Formulation

Google Meet formulates it impossible for technically clout conference IDs via using cipher extended ten fonts, by means of using twenty five fonts within the box (this is when a malicious person tries to guess a conference ID and create an illegal effort in the direction of connecting it). The individuality of fact records is typically distinct as the amount of exclusive trials that will calculate the probability of distinguishable ones [12, 20]. In a re-identified or anonymous dataset, isolated specific information can also be distinguishable through connection assault, where title, contact details and additional understandable information are excluded. For instance, the unidentified therapeutically information of Massachusetts was distinguishable in Sweeney's study [21] by linking them to the list of voter registrations. In addition, Google Meet clients are able to register their description with our Sophisticated Defense Curriculum (SDC), which is exclusively intended and designed for the elevated likelihood descriptions and provides our best phishing and account hijacking safeguards.

There have also been more and more models and technologies proposed to address privacy issues [22, 23]. In Ref. [20], Tu et al. created a molest arrangement that can restore client path through an exactness and the rate of 73%–91%. Without any prior information, as of two collective portable records containing merely the amount of singles enclosed by everyone. For another example, in Ref. [10], Montjoye et al. emphasized to facilitate four comprehensive locations that is able to exclusively classify 90% of bank card owner transaction history. Course Replica for User GTM.

Collective record may be created by combining unique course data supported by time periods to distribute merely special, fascinating arithmetical records (e.g., the quantity of individuals protected through every minute unit) [20]. A re-establishing jeopardy is defined by the course individuality, which is directly linked through the collective record. Consequently, we can determine the re-establishing jeopardy by means of combined records.

The GTM, a solitude jeopardy assessment replica, is designed to forecast the individuality of course datasets in which three parameters can define aggregated data (i.e., quantity of singles, amount of time period and spatiotemporal concentration of respective). The effect of this already mentioned factors on course individuality is subsequently analyzed. Spontaneously, new people could share a unique trajectory since the volume of the unique record set amplified. The GTM, a model of solitude jeopardy assessment, is designed in the direction of forecasting the individuality of

Table 1 Key notations GTM

R	Prediction set
E	Set of all trajectory set database D
\forall	Belonging all to cell value
\in	Belong to value set
gi	Spatial temporal tuple value individual
si	Set value
D	Possible trajectories for set
jDj	Possible trajectory of number sets
jTj	Amount of the entire probable trajectories
T	Set of all possible trajectories

course dataset in which combined record can be illustrated by different factors. Table 1 represents the key notation of this working model.

3.1 Working Procedure

We denote as D the set of all promising course records described through the grouping of GTM parameters. It is impractical to list all of them since the large quantity of records in D. However, constrained by the same combination of parameters, the majority of the records in D comprise extremely analogous individuality. Consequently, in order to measure the uniqueness of the initial trajectories, the GTM utilizes the arithmetical uniqueness of course inimitability in D (i.e., assumption along with variation). Figure 1 demonstrates the overall working of this model. As time increases, the rising amount of period concentration of spatiotemporal entity also increases. Grouping of the three factors shall be designated because of the amalgamation of the factors of GTM. The grouping of GTM parameters, obviously, is capable of, to be derived as of the primary record set and represents distant additional single course record set. Owing to the large amount of records in D, it is impractical to catalog every one of them. Though, most of the records in D contain extremely analogous individuality, limited by the same combination of parameters. The GTM therefore utilizes the statistical features of trajectory uniqueness in D in order to calculate the uniqueness of the initial trajectories (i.e., assumption and variation).

Definition 1 (Course): The course T of an individual is an structured set of comprehensive tuples: $S \ 1/4 \ \langle\langle g1 \ 1; s1 \rangle; \langle g2; s2 \rangle; \dots; \langle sjgj; sjgj \rangle\rangle$, where $sj1 < j \ll jTj$ denotes the small cell accessed, $Gt < Gj$ is the universe in time slot tj of all accessible miniature units (i.e., tiny units allocating geo-assigned specifications) and $|T|$ represents the course span (i.e., the amount of time periods in T).

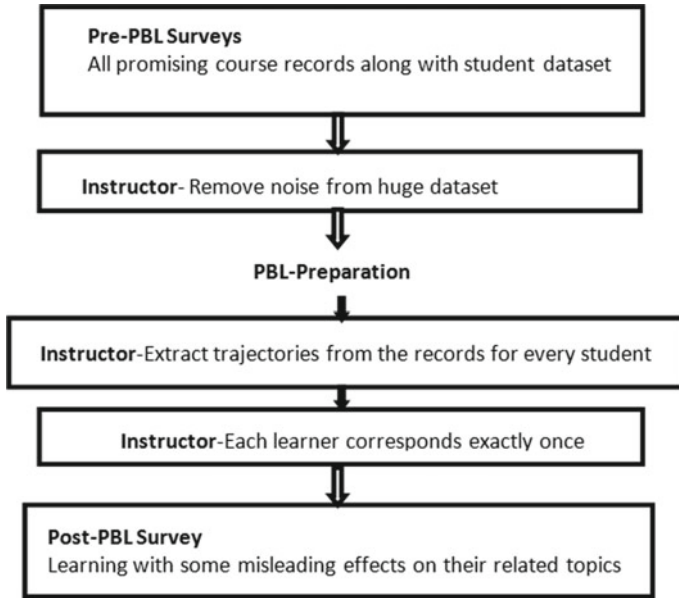


Fig. 1 Workflow diagram of proposed model

4 Trajectory Dataset and Privacy Evaluation

The dataset taken for 540 students with a total of 100,000 records, our dataset contains one week of mobility results. All records contain one user identifier and one spatiotemporal point. True identities have been replaced by digital identities. The location of a person was reported every fifteen minutes. The spatial accuracy of the data is equal to that of a fixed number of distinct locations instead of the exact locations of the users.

4.1 Implementation Process

In the entire dataset, there are about 1700 different student positions. Distances between two positions can be determined by their Euclidean distance (2-norm) that $\|p_i - p_j\|$. The smallest distance to any two points ($\min\|p_i - p_j\|$) is approximately 0.11 km, while the largest distance to the mobile location ($\max\|p_i - p_j\|$) is approximately 200 km. In two stages, we preprocess the dataset on two-step dataset preprocessing. Second, we filter the raw data to boost the consistency of the data. In order to arrive at a more meaningful analysis, the original data, like any huge, real world data set, contains some noise that should be removed. We found that in the original dataset were several duplicate records as well as some “singleton” students with

only one position during the entire week. We filter out the redundant records from the dataset and delete records for all students with only one record with one position in order to improve the effectiveness of uniqueness computation. In the filtered dataset, 0.63% of learners are included. Second, we extract trajectories from the records for every student, i.e., we extract the ordered set of spatiotemporal points. Each learner corresponds exactly once to the tri-trajectory.

To measure the achievement of the students, the teachers have to adapt modern online evaluation methods. With updated online article query, online practical tests, peer and self-evaluation through online mode, etc., they have to integrate online written examinations. Online problem-based learning has found to be marginally more beneficial to feminine contributor although having misleading effects on their male contributor related to topic-based learning.

5 Experimental Result

We calculate the average amount of spots required to exclusively classify the trajectory of a person by calculating the uniqueness of the dataset both before and after anonymous records. To estimate the uniqueness of a user's trajectory, we use a sampling-based approach (Tables 2, 3, 4 and 5).

To calculate the dataset's on Fig. 2. The x -axis represents the number of users who have the same spatiotemporal points selected for the target user. The y -axis is the normalized number of individuals that have identical spatiotemporal points with a certain number of individuals. Note the large difference between the frequencies

Table 2 Spatiotemporal tuples trajectory algorithm

Algorithm: New Representative CLCSTRpar presents Generation
Input: (1) Consider $G \in T$ as CLCSTRpar present = $\{g1, g2 \dots Cnindexstart\}$; (2) MLins (3) A smoothing parameter α
Output: The demonstrative PLRi as CLCSTRpar present = $\{s1, s2 \dots pnindexstart\}$; Algorithm:
01: Find max value of direction vector field \vec{v} ; 02: Replace the hatchets X axis is equivalent to \vec{v} ; 03: set value CLCSTRpar present = currentindex; 04: iterative ($S \in R$) do;
05: until;
06: if ($G \in r$) repeat
07: suppose amt be the quantity of the row division that holds the X' - rate for the position p ;
08: if ($amt \geq G$ MLins) then
09: calculate variation in X' —values among p and its nearest position;
10: if ($diff_rence \geq \alpha$) then
11: calculate the standard synchronized avg_p; 12: Unwrap the spin and find out avg_p;
13: Add avg_p to the closing stage of iTRj;
14: otherwise, repeat predictive generation trajectory clustering;

Table 3 New representative CLCSTRpar presents generation algorithm

Algorithm generation (spatiotemporal tuples trajectory)
Input: Find trajectories generation data $T = \{G_1, \dots, G_n\}$ Output: (1) set prediction cell data $R = \{S_{i1}, \dots, S_{j1}\}$
(2) $A S = \{\text{set of trajectories moving set of datas } 1 \dots s_j\}$
Algorithm:
/* CLSTR clustering prediction */ 1: $\forall (G \in T)$ recursive until termination;
2: Predictive generation trajectory clustering; find L_j and $N_\epsilon(L)$;
3: Accumulate L_s into a set E ;
/* Grouping Phase */
4: Execute line segment clustering for E ; find value of R clusters;
5: search belong ($S \in R$) do;
6: Repeat predictive generation trajectory clustering;

Table 4 Individuals report for spatiotemporal density with time (N¼ 250, d¼ 4, every location supplied through a solitary unit)

Locations	t1 (minute)	t2 (minute)	t3 (minute)	t4 (minute)
East	45	60	45	60
West	45	60	45	60
North	45	60	45	60
South	45	60	45	60
Northeast	45	60	45	60
South-west	45	60	45	60

Table 5 Widespread spatiotemporal density of individuals at the Chhattisgarh location (N¼ 250, d¼) Location are 1: Bhilai, 2: Korba, 3: Bilaspur, 4: Durg, 5: Chharoda, and 6: Aambikapur

Location No	East	West	North	South
1	61	75	49	29
2	40	20	51	50
3	65	45	54	32
4	20	49	37	23
5	39	25	25	45
6	49	37	41	50

of uniquely identified individuals, which decreases from 0.6 in the initial data to 0.4 in the anonymized data. The same results are shown on the logarithmic scale and with a wider x-axis spectrum as in Fig. 3. We take 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1 2 3 4 5 6 as uniqueness initial random number of points Anonymous 6 h Anonymous

8 h Anonymous as 2 h that shown on Fig. 4: Contrast of uniqueness before and after anonymisation.

The x-axis indicates the number of random points selected. The y-axis indicates the importance of measured uniqueness. Lower values are stronger, i.e., is 0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0 10 20 30 40 50 the frequency number of persons initially anonymous Fig. 3: Distribution of the number of persons concurrently residing in the same place. There are more people who have at least two other people who have the same spatiotemporal points we got in result.

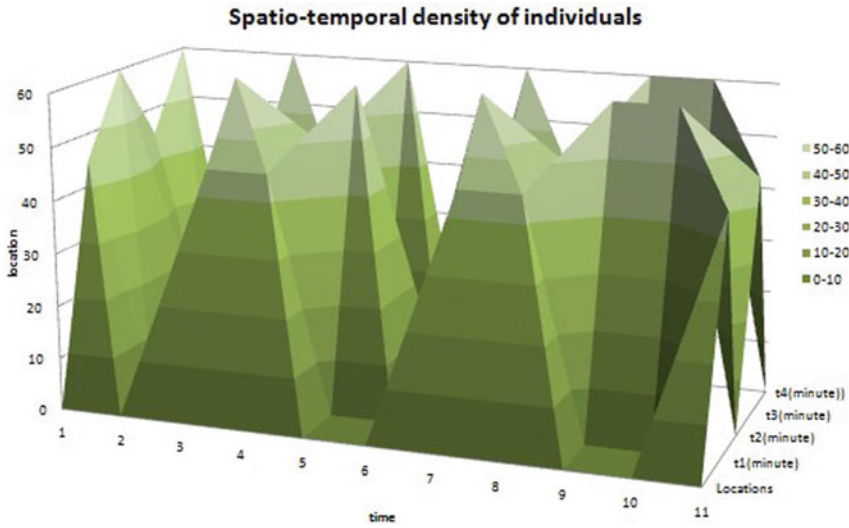


Fig. 2 Comprehensive solidity of individual

Fig. 3 Course individuality versus amount of time period (prospect of individuality for a distorted allocation earlier represented through red and later represented through blue)

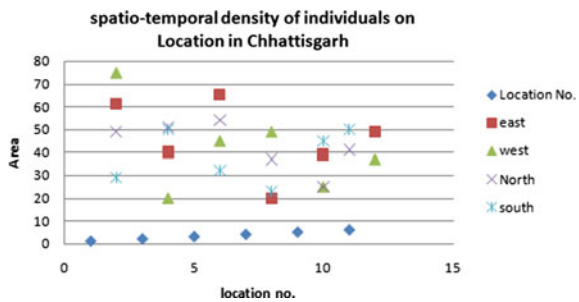
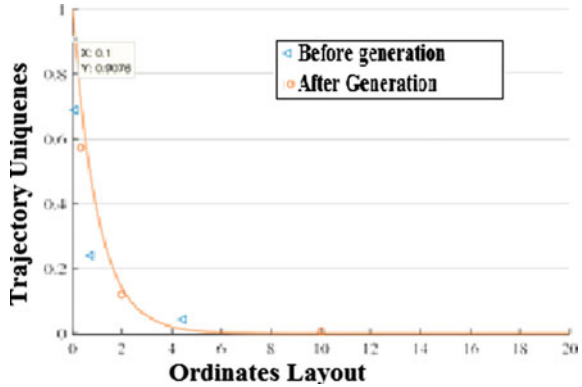


Fig. 4 Course individuality versus predictability key (potential of individuality for irregular allocation earlier is represented through blue triangle and later represented through orange circle in the simplification, while the FPC is depicted by orange line)



5.1 Analysis

We are anonymizing the dataset with three different time window sizes: 6, 8 and 2 h. The uniqueness of the anonymized dataset decreases dramatically, particularly when the time is 6 h. In the case of two random points, the uniqueness falls by more than 0.2, from 0.6 to around 0.4.

6 Conclusion

The proposed GTM suggested that the confirmation of that recognized jeopardy of unidentified course record can be estimated from summative record. Certainly, if the aggregated information is released by the OGMM service providers, the OGMM service providers disclose the aggregated information. The OGMM model help consumers and isolation advocates will employ our replica to determine the risk of re-classification.

This article focuses on Google-meeting online scenarios and highlight analyzing the course of privacy risks. Present suggestion focuses the analysis that shows the numerical distinctiveness of course individuality and can be created using amalgamation of factors, i.e., quantity of persons, amount of time period, and comprehensive concentration of persons that described by an aggregated data. We scan for all other students based on the selected points and count the number of students having the same points. That is, we are looking for other students who were concurrently at the same places.

References

1. Sun Y, Peng M, Mao S (2019) Deep reinforcement learning-based mode selection and resource management for green fog radio access networks. *IEEE Internet Things J* 6(2):1960–1971. <https://doi.org/10.1109/JIOT.2018.2871020>
2. Yang B, Wu D, Wang R (2019) Cue: an intelligent edge computing framework. *IEEE Netw* 33(3):18–25. <https://doi.org/10.1109/MNET.2019.1800316>
3. Angriyan E, Sulisworo D (2022) A digital worksheet based on a problem based learning to improve creative thinking skills in mathematics learning. In: Conference: 5th international European conference on interdisciplinary scientific research, Valencia
4. Chang Z, Hu Y, Chen Y, Zeng B (2019) Cluster-oriented device-to-device multimedia communications: joint power, bandwidth, and link selection optimization. *IEEE Trans Veh Technol* 67(2):1570–1581
5. Li Z, Jiang Y, Gao Y, Sang L, Yang D (2019) On buffer-constrained throughput of a wireless-powered communication system. *IEEE J Sel Area Commun* 37(2):283–297. <https://doi.org/10.1109/JSAC.2018.2872374>
6. Wu D, Shi H, Wang H, Wang R, Fang H (2019) A feature-based learning system for internet of things applications. *IEEE Internet Things J* 6(2):1928–1937
7. Wu D, Liu Q, Wang H, Yang Q, Wang R (2019) Cache less for more: exploiting cooperative video caching and delivery in D2D communications. *IEEE Trans Multimedia* 21(7):1788–1798. <https://doi.org/10.1109/TMM.2018.2885931>
8. de Montjoye Y-A, Radaelli L, Singh VK, Pentland AS (2015) Unique in the shopping mall: on the reidentifiability of credit card metadata. *Science* 347(6221):536–539. <https://doi.org/10.1126/science.1256297>
9. Zhang Y, Chen M, Mao S, Hu L, Leung VCM (2014) Cap: community activity prediction based on big data analysis. *IEEE Netw* 28(4):52–57. <https://doi.org/10.1109/MNET.2014.6863132>
10. Srikan P, Pimdee P, Leekitchwatana P, Narabin A (2021) A problem-based learning (PBL) and teaching model using a cloud-based constructivist learning environment to enhance thai undergraduate creative thinking and digital media skills. *Int J Interactive Mobile Technol (IJIM)* 15(22):68. <https://doi.org/10.3991/ijim.v15i22.24963>, November
11. Luo C, Ji J, Wang Q, Chen X, Li P (2019) Channel state information prediction for 5G wireless communications: a deep learning approach. *IEEE Trans Netw Sci Eng* 1–1 <https://doi.org/10.1109/TNSE.2018.2848960>
12. Wu D, Zhang Z, Wu S, Yang J, Wang R (2019) Biologically inspired resource allocation for network slices in 5G-enabled internet of things. *IEEE Internet Things J* 1–1. <https://doi.org/10.1109/JIOT.2018.2888543>
13. Zhang Z, Wang C, Gan C, Sun S, Wang M (2019) Automatic modulation classification using convolutional neural network with features fusion of SPWVD and BJD. *IEEE Trans Signal Info Process Over Netw* 1–1. <https://doi.org/10.1109/TSIPN.2019.2900201>
14. Cui H, Luo C, Wu J, Chen CW, Wu F (2013) Compressive coded modulation for seamless rate adaptation. *IEEE Trans Wireless Commun* 12(10):4892–4904
15. Cao Y, Zhao N, Yu FR, Jin M, Chen Y, Tang J, Leung VCM (2018) Optimization or alignment: secure primary transmission assisted by secondary networks. *IEEE J Sel Area Commun* 36(4):905–917. <https://doi.org/10.1109/JSAC.2018.2824360>
16. Tan B, Lu J, Wu J, Zhang D, Zhang Z (2018) Toward a network slice design for ultra high definition video broadcasting in 5G. *IEEE Wireless Commun* 25(4):88–94. <https://doi.org/10.1109/MWC.2018.1800021>
17. Agyapong PK, Iwamura M, Staehle D, Kiess W, Benjebbour A (2014) Design considerations for a 5G network architecture. *IEEE Commun Mag* 52(11):65–75. <https://doi.org/10.1109/MCOM.2014.6957145>
18. Tu Z, Xu F, Li Y, Zhang P, Jin D (2018) A new privacy breach: user trajectory recovery from aggregated mobility data. *IEEE/ACM Trans Netw* 26(3):1446–1459. <https://doi.org/10.1109/TNET.2018.2829173>

19. Sweeney L (2022) k-anonymity: a model for protecting privacy. *Int J Uncertain Fuzziness Knowl-Based Syst* 10:557–570. <https://doi.org/10.1142/S0218488502001648>
20. Machanavajjhala A, Kifer D, Gehrke J, Venkatasubramanian M (2007) L-diversity: privacy beyond k-anonymity. *ACM Trans Knowl Discov Data* 1. <https://doi.org/10.1145/1217299.1217302>
21. Li N, Li T, Venkatasubramanian S (2007) t-closeness: privacy beyond k- anonymity and l-diversity. In: 2007 IEEE 23rd international conference on data engineering, pp 106–115. <https://doi.org/10.1109/ICDE.2007.367856>
22. Xiong J, Ren J, Chen L, Yao Z, Lin M, Wu D, Niu B (2019) Enhancing privacy and availability for data clustering in intelligent electrical service of IoT. *IEEE Internet Things J* 6(2):1530–1540. <https://doi.org/10.1109/JIOT.2018.2842773>
23. Mustofa RF, Hidayah Y (2020) The effect of problem-based learning on lateral thinking skills. *Int J Instr* 13(1):463–474

A Review on Fake News Identification in Online Social Networks



A. B. Athira, S. D. Madhu Kumar, and Anu Mary Chacko

Abstract In the modern world, people’s major source of information has changed to online news, as social media has expanded in popularity across all generations. It has also resulted in the dissemination of fake news. The proliferation of fake news is a growing threat to academics and businesses alike. The exponential growth of fake news has heightened the demand for automated fake news identification in recent years. Several approaches to fake news have yielded promising outcomes. On the other hand, these detection systems cannot explain why they made a prediction. The capacity to discover bias and discrimination in detection algorithms is a crucial benefit of explainability. This survey highlights recent advancements in the detection of explainable fake news. We discuss the shortcomings of existing fake news detection models based on explainable AI (XAI). A multimodal explanationable fake news detection model is also described in our paper.

Keywords Misinformation · Explainable fake news identification · Fake news detection · Fake news identification · Explainable AI

1 Introduction

In today’s world, Internet social media platforms have surpassed traditional news sources as the primary source of information. Most of the news material on social media is intentionally misleading. We use the phrase “fake news” to represent them. The dissemination of fake news influenced the 2016 presidential election in the USA.

A. B. Athira (✉) · S. D. Madhu Kumar · A. M. Chacko
Computer Science and Engineering, NIT Calicut, Calicut, India
e-mail: athira_p190140cs@nitc.ac.in

S. D. Madhu Kumar
e-mail: madhu@nitc.ac.in

A. M. Chacko
e-mail: anu.chacko@nitc.ac.in

Following the election, the phrase has become commonplace. For many reasons, fake news identification is a technological challenge. The main problem is that there is no universally accepted definition of fake news, which makes identifying fake news difficult. Content is quickly created and disseminated through social media platforms, resulting in a large amount of data to evaluate. The vastness of online material, which covers many topics, adds to the task’s difficulty.

Practical ways of identification of fake news include the use of fact-checking websites. FactCheck.org, TruthOrFiction.com, Politifact.com, Snopes.com, and Hoax-Slayer.com are examples of fact-checking websites. The majority of fact-checking websites utilize human resources-based methodologies. On the other hand, human fact-checking is time-limited since a vast volume of data is generated daily. It also has scalability concerns. Thus, the challenges imposed by manual fact-checking approaches paved the way toward the automatic detection of fake news.

Figure 1 explains the existing research on fake news identification approaches, which is grouped into two categories (based on the features used for data analysis). The majority of previous research works treat the identification of fake news as a binary classification problem. Our literature analysis concluded that the challenge of fake news identification could be classified as either single or multimodal, depending on the number of characteristics involved. In fake news identification models using single-modality, a single type of feature is analyzed. Various feature combinations are examined in the case of fake news identification models, which uses multimodality.

All current research is seeking AI-based solutions, regardless of the modality of the features used for training, because AI systems have shown indisputable performance when compared to traditional machine learning approaches. However, because of its black-box nature, a normal user would struggle to understand the reason for

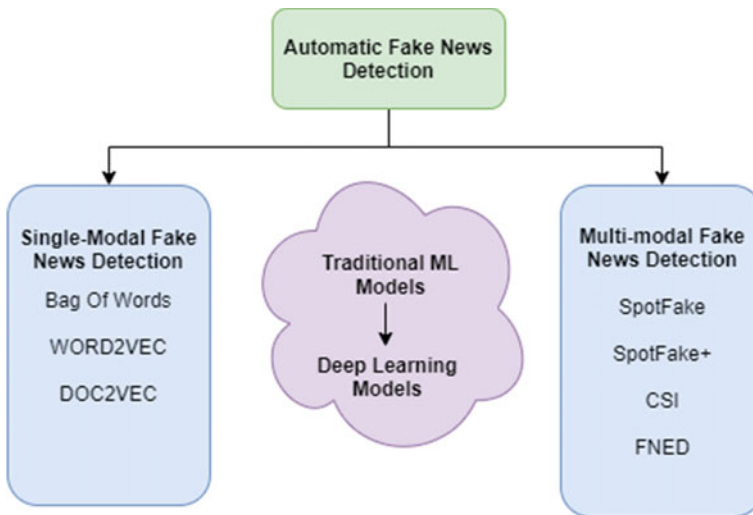


Fig. 1 Automatic fake news identification: an overview

prediction. The AI system's effectiveness is limited by its lack of explainability when compared to the normal machine learning model. The detection model must have transparency into the decision-making process. Thus, explainable based fake news identification models are needed to reduce bias and increase performance.

The remainder of the paper is organized as follows: Sect. 2 highlights studies on applying explainable AI (XAI) in detecting fake news. We also discuss related works in this section. Section 3 discusses our model to detect fake news by analyzing image features and linguistic content of news propagated on social media platforms. Section 4 concludes the paper.

2 Fake News Detection Using XAI

People use social networking sites to stay updated on news and information, and their part in spreading misinformation cannot be ignored. Automatic tools are available for buying likes, followers, and comments on social media posts. This leads to the creation and propagation of fake news through social media platforms. The motivation for creating and spreading fake news can range from monetary gain to criminal or political gain. It is hard to spot such news on social media platforms.

Several researchers have recently created AI-based methods to prevent the spread of fake news. Several models use a large dataset and complex machine learning models to identify fake news. Deep learning algorithms may pick up biases in the training data. The model's effectiveness is impacted if users rely on classification accuracy while neglecting the reasons behind specific predictions. A system whose predictions are supported by an explanation is easier to believe. Recently, explainable AI has been used by several academics to identify fake news. The most advanced techniques for identifying fake news that use explainable AI are listed in Table 1.

When we reviewed the table, we discovered that most existing models in explainable fake news identification focus on textual content of the social media news items for providing explanations. Only [2–4] provided explanations based on news content and social engagements. The authors of [7] took a different approach, they identified the patterns in the networks to build an explainable fake news identification system. The authors of [11] recently proposed a methodology for identifying explainable fake news that takes advantage of user information, linguistic, and temporal aspects in the propagation pattern. Recently, many researchers have combined various elements of social media news items to identify fake news. We could not find any study that uses visual cues in explainable fake news identification.

We are currently developing an explainable fake news detection algorithm based on visual features. The model's details are discussed in the following section.

Table 1 Overview of recent explainable AI approaches for detecting fake news

Research articles	Features analyzed	Model used	Explainability model used	Accuracy
[1]	News attributes, semantic meanings of news statements, and linguistic features of news statements	MIMIC, PERT, and ATTN frameworks	Mimic learning (or explanation by simplification), perturbation-based explanation, self-attention-based explanation, and visual explanation	67.1%, 67.3%, and 53.2% accuracy for MIMIC, ATTN, and PERT, respectively
[2, 3]	News content and user comments	RNN and bidirectional GRU	Co-attention captures intrinsic explainability and visual explanation	Outperforms related works by at least 5.33%
[4]	News content and user comments (filtered)	Multi-nominal knowledge base and RNN	Co-attention mechanism	Accuracy increased above 90% for all machine learning algorithm
[5]	Interpretable textual features	CoCoGen, BRNN, and GRU	Ablation explanations	ISOT:99% LIAR dataset:27.4%
[6]	Textual features of news content	BiDir-LSTM-CNN and BERT	Visual explanation and explanation based on attention	Accuracy achieved 85% of accuracy
[7]	Network-based patterns	Supervised learning algorithms with fivefold cross-validation	Empirical studies on social psychological theories	93% accuracy for PolitiFact and 86% for BuzzFeed dataset
[8]	Textual contents from news	TM: Tstelin Machine	Local and global interpretability	Outperforms baseline works by at least 5% accuracy
[9]	Semantic contents of news articles	Knowledge graph framework (KGF)	Structured explainability (relational level)	71.4, 86.0, and 73.3% accuracy for datasets Celebrity, PolitiFact, and GossipCop, respectively
[10]	Tweets from source and retweet users	Graph attention network	Attention-based explanation	Improvement of 2.5% than previous method
[11]	User features, textual features, and temporal features in propagation pattern	Propagation network	Attention-based explanation	Outperformed 89.2% for dataset GossipCop

(continued)

Table 1 (continued)

Research articles	Features analyzed	Model used	Explainability model used	Accuracy
[16]	User features and textual features	GCAN	Attention-based explanation	Outperformed state-of-the-art methods by 16% in accuracy

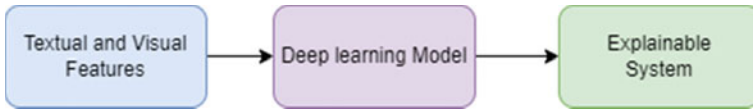


Fig. 2 Workflow diagram

3 Analysis of Textual and Image Content

The textual feature analysis of news content is the primary focus of the present XAI-based models for fake news identification. These strategies will only be effective if the news articles are written in a consistent manner. When the writing style changes, these methods become irrelevant. The answer to this issue is aided by the inclusion of additional elements such as social interaction, user profile data, and image features.

Social media users are more likely to find fake news postings appealing with multimedia content than ones with only text. The accuracy of fake news detection techniques can be increased with the use of these visual components. Our study focuses on the analysis of social media news stories’ visual and linguistic content. The relationship between the text and visual features is employed to provide a reason for the prediction. There has been good progress in the research, and preliminary findings have been encouraging.

Experimentation was carried out using the FakeNewsNet dataset. Figure 2 outlines a high-level view of our proposed model. The linguistic and image features were retrieved from the social media news posts. These are the inputs to *XLNet* [20] and *VGG19* [19], respectively (deep learning models). Classification of news as fake or real is done by the deep learning models. An additional module is used to provide an explanation for the predicted output. The explainable system will provide the reason for the prediction.

4 Conclusion

In this paper, we present the importance of artificial intelligence in finding and preventing the proliferation of fake news on social media. This paper presents an overview of our work in this area, along with the significance of explainable AI

in detecting fake news. We have listed a detailed analysis of recently developed methods for identifying fake news using explainable AI. The preliminary results of our model based on explainable AI are encouraging, and we are moving forward with developing a multimodal fake news identification system with adequate explanations.

References

1. Yang F, Pentylala SK, Mohseni S, Du M, Yuan H, Linder R, Ragan ED, Ji S, Hu XB (2019) Xfake: explainable fake news detector with visualizations. In: The world wide web conference, WWW '19. Association for Computing Machinery, New York, NY, USA, pp 3600–3604
2. Shu K, Cui L, Wang S, Lee D, Liu H (2019) Defend: explainable fake news detection. In: Proceedings of the 25th ACM SIGKDD inter-references 38 national conference on knowledge discovery and data mining, KDD '19. Association for Computing Machinery, New York, NY, USA, pp 395–405
3. Cui L, Shu K, Wang S, Lee D, Liu H (2019) Defend: a system for explainable fake news detection. In: Proceedings of the 28th ACM international conference on information and knowledge management, CIKM '19. Association for Computing Machinery, New York, NY, USA, pp 2961–2964
4. Sharma DK, Sharma S (2021) Comment filtering based explainable fake news detection. In: Singh PK, Wierzhón ST, Tanwar S, Ganzha M, Rodrigues JJPC (eds) Proceedings of second international conference on computing, communications, and cyber-security, Singapore. Springer Singapore, pp 447–458
5. Qiao Y, Wiechmann D, Kerz E (2020) A language-based approach to fake news detection through interpretable features and BRNN. In: Proceedings of the 3rd international workshop on rumours and deception in social media (RDSM), Barcelona, Spain (Online). Association for Computational Linguistics, pp 14–31
6. Kurasinski L, Mihailescu R-C (2020) Towards machine learning explainability in text classification for fake news detection. In: 2020 19th IEEE international conference on machine learning and applications (ICMLA), pp 775–781
7. Zhou X, Zafarani R (2019) Network-based fake news detection: a pattern-driven approach. SIGKDD Explor Newsl 21:48–60
8. Bhattarai B, Granmo O-C, Jiao L (2021) Explainable tsetlin machine framework for fake news detection with credibility score assessment
9. Wu K, Yuan X, Ning Y (2021) Incorporating relational knowledge in explainable fake news detection. In: Advances in knowledge discovery and data mining. Springer International Publishing, Cham, pp 403–415
10. Ni S, Li J, Kao H-Y (2021) Mvan: multi-view attention networks for fake news detection on social media. IEEE Access 9:106907–106917
11. Silva A, Han Y, Luo L, Karunasekera S, Leckie C (2021) Propagation2vec: embedding partial propagation networks for explainable fake news early detection. Inf Process Manage 58(5):102618
12. Yang Z, Yang D, Dyer C, He X, Smola A, Hovy E (2016) Hierarchical attention networks for document classification. In: Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: human language technologies, June 2016. Association for Computational Linguistics, San Diego, California, pp 1480–1489
13. Guo H, Cao J, Zhang Y, Guo J, Li J (2018) Rumor detection with hierarchical social attention network. In: CIKM '18. Association for Computing Machinery, New York, NY, USA
14. Ahmed H, Traoré I, Saad S (2017) Detection of online fake news using n-gram analysis and machine learning techniques. In: ISDDC, 2017
15. Wang WY (2017) “liar, liar pants on fire”: a new benchmark dataset for fake news detection

16. Lu Y-J, Li C-T (2020) Gcan: graph-aware co-attention networks for explainable fake news detection on social media
17. Wikipedia. Fake news (2021). Accessed on 20 October 2021. [Online]. Available <https://en.wikipedia.org/wiki/Fakenews>
18. Zhou X, Zafarani R (2020) A survey of fake news: fundamental theories, detection methods, and opportunities. *ACM Comput Surv* 53
19. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv preprint [arXiv:1409.1556](https://arxiv.org/abs/1409.1556)
20. Yang Z, Dai Z, Yang Y, Carbonell J, Salakhutdinov R, Le QV (2019) Xlnet: generalized autoregressive pretraining for language understanding (2019). <https://doi.org/10.48550/ARXIV.1906.08237>, <https://arxiv.org/abs/1906.08237>

Comparative Assessment of State-of-the-art Image Fusion Techniques for Fusion of Multi-modality Medical Images



Vineeta Singh and Vandana Dixit Kaushik

Abstract Clinical imaging is a vital part in clinical field for analysis of illnesses, for example, growth recognition, malignant growth identification and a lot more where inside organ infections are required to have been analyzed. Clinical instruments are exorbitant and accompanied impediment. For instance, CT images give better insights concerning hard designs in a human body, while MRI portrays better insights concerning delicate tissues. Be that as it may, there might be sure cases where both of the subtleties are expected by a clinical specialist to productively analyze. Here, machine vision, human insight, PC vision, clinical imaging, minute imaging and a lot more are such encouraging fields where clinical images combination plays out a vital job. In this examination paper, creators have assessed various images amalgamation strategies on clinical picture dataset for exhibition of best performing techniques through specific execution measurements. Trial results outlined that pulse coupled neural network has performed better when contrasted with non-sub sampled contourlet transform, sparse representation as well as NSCT+SR methodologies including execution evaluators: average gradient, standard deviation and spatial frequency with a best performing values of 10.4, 82.2 and 8.1 separately.

Keywords NSCT · SR · PCNN · Performance analysis · Fusion strategies · Medical imaging · Multi-modality medical image fusion · Comparative analysis of image fusion algorithms · Image fusion algorithms

1 Introduction and Literature Survey

There is a huge demand of visual effect in medical imaging field involving clinical-assisted diagnosis of the diseases inside a human body. Clinical advancement for diagnosing diseases inside a human body involves medical imaging and day by day

V. Singh (✉) · V. D. Kaushik

Department of Computer Science and Engineering, Harcourt Butler Technical University, East Campus, Nawabganj, Kanpur, Uttar Pradesh 208002, India

e-mail: vineeta.singh.cs@gmail.com

advancements, researches, inventions attracting more opportunities to utilize medical imaging in diagnosis purpose and in clinical investigation [1, 2].

Imaging of a modality demonstrates specific information and it varies from modality to modality for the lesion. For instance, CT image of a modality reflects clearer illustration of a bone image, while images of soft tissues are not clearly visible, i.e., blurred. MRI image reflects information in perspective of multi-angle, multi-plane details for a soft tissue while image of a skeleton is not clear here. PET image illustrates information related to metabolic activity in context with human cells, while anatomy structure is blurred here. Hence, the images of different modality undergo a fusion process to enhance the accurate as well as identification of location of lesion, as a result of which, much effective medical image is obtained to diagnose a disease for modern medicine [3]. NSCT transform has merit of multi-direction, multi-scale analysis in comparison with wavelet as well as contourlet transform. NSCT involves merit of anisotropy as well as translation invariance [4]. But the demerit of NSCT is that low-frequency sub-bands generated here are not sparse, by which to directly fuse those coefficients is not helpful in maintaining the source image characteristics, while SR method is capable of extracting deeper structural features for low-frequency sub-bands as well as further approximates them in a linear form [5]. As compared to different artificial neural networks (ANNs), PCNN involves promising merits [6]. PCNN technique involves global coupling as well as pulse synchronization which results in combining high-frequency sub-bands as per visual characteristics of human being further yielding richer details and information [7].

In the research paper [5], an image fusion technique was devised relied on sparse representation as well as non-subsampled contourlet transform. Although efficiency of image fusion was enhanced but the demerit was to provide four directional sparse representations for low-frequency sub-groups coefficients as a result of which full representation of details as well as characteristics of source image was not possible. In [8], a fusion scheme was devised to fuse multi-modal images, and scheme was based on NSCT+SR. Here, it was possible to approximate low-frequency sub-bands effectively. But for high-frequency sub-groups, rules relied on broader local energy/variance resulted to reduce effectiveness of image detail smoothness issue generated via SR. Different image fusion techniques to fuse multi-modal medical images have been studied as well as analyzed by the researchers and further a future direction has also demonstrated. Multi-modal medical image fusion techniques with a hybrid approach as well as enhancements of such hybrid approaches also have been analyzed via researchers [9, 10]. Medical image fusion scheme with an optimum weight as well as particle swarm optimization technique was devised [11]. Researchers devised a fusion technique for fusing medical images relied on non-subsampled contourlet transform as well as convolution neural networks [12]. Image fusion strategies have a wider area of applications along with medical field, such as remote sensing, concealed weapon detection, undersea object detection, image quality enhancement [13, 14], clinical diagnosis, classification, pattern recognition, defense area and many more.

The other sections of the research paper illustrate as mentioned here: Sect. 2 demonstrates image fusion techniques involved in the study, while Sect. 3 demonstrates allied key indicators, while Sect. 4 describes simulated work and comparative discussion involving visual analysis, graphical analysis and comparative description and Sect. 5 includes conclusion and illustrates future scope.

2 Current Fusion Strategies

2.1 Non-subsampled Contourlet Transform

In non-subsampled transform technique, it involves two stages: At first stage, non-downsampling pyramid, i.e., NSP decomposition takes place, while second stage involves non-downsampling direction filter bank, i.e., NSTDVFB. Here, NSP decomposition involves decomposition of input images further to coefficients of low-frequency sub-groups, moreover coefficients of high-frequency sub-groups via help of NSTFB, i.e., non-subsampling tower filter bank for facilitating the characteristics of non-subsampled contourlet transform, i.e., multi-scale as well as multi-directional analysis of images [15].

2.2 Sparse Representation

Sparse representation is abbreviated as SR, it means approximation of a natural signal may be done by a linear composition involving few atoms in an over-complete dictionary $\mathbf{R} \in \mathbf{P}_{n \times k}$, moreover for signal s sparse coefficient is yielded through expression $\min \mathbf{Z} \|\mathbf{Z}\|_0$, s.t. $\|X - \mathbf{RZ}\|_2$; here $\epsilon > 2$ as well as \mathbf{R} represents a pre-defined dictionary as well as \mathbf{Z} shows an sparse coefficient vector, moreover $\|\mathbf{Z}\|_0$ signifies the count for nonzero entries found in \mathbf{Z} while finally ϵ denotes the error for bounded representation [16].

2.3 PCNN

PCNN simplest version involves feedback neural network model devised via signal processing technique in reference to cat visual cortex [17]. Simple version involves partial simplification of the parameters which results to well guarantee generality for the technique. Whereas a strong disparity occurs in the visual system reaction for various varying featured image locations. In this technique, this variation is basically involved with the parameters settings where the considerable change in

parameter settings will influence the fused outcomes as well. Here, common discrete mathematical iterative model has been utilized [18].

2.4 NSCT+SR

Image fusion involves fusion of multiple images captured from a single scene to produce a fused final image involving essential information. This objective may be achieved via NSCT and SR collectively. Whereas in SR technique to yield accurate illustration about image details is cumbersome process because of limited numbered dictionary atoms which results to difficult handling of large calculations. Moreover in NSCT technique (multi-scale transform-based technique), low-frequency sub-bands are difficult to demonstrate sparsely by which extraction of specific features out of images is not possible. In [19], a fusion technique involving NSCT+SR was devised. NSCT technique utilized to carry out multi-scale decomposition for input images for expressing image details and further a dictionary learning technique in NSCT domain was utilized, by which representation of low-frequency information of image in a sparsely manner was possible and further extraction of salient features for images was done [19].

3 Performance Evaluators

3.1 AvGr

The performance metric average gradient is utilized for measuring gradient information for fused image and it yields texture detail for an image. Higher value denotes better performance result [20].

3.2 StDv

Standard deviation (StDv) metric is utilized to measure contrast for fused output image. Greater value represents high contrast fused final image [21]. Equation form of illustration is depicted as:

$$\text{STDV} = \sqrt{\frac{1}{NM} \sum_{i=1}^N \sum_{j=1}^M (I_{\text{fuse}}(i, j - m'))^2} \quad (1)$$

3.3 *SpFr*

Sharpness of a fused image reflected via spatial frequency (SpFr). Greater value of SpFr denotes higher image resolution has been achieved [22]. SpFr may be illustrated via following equation:

$$SF = \sqrt{RF^2 + CF^2} \quad (2)$$

where RF stands for row frequency and CF refers to column frequency of the images consecutively.

4 Experimental Related Details

The medical image group has been extracted out of <http://www.med.harvard.edu/AANLIB/home.html> [23] for experimental related work. The methods NSCT, PCNN, SR and NSCT+SR have been evaluated on the color medical image dataset for the three performance metrics AvGr [20], SpFr and StDv [21]. Here, four comparative algorithms have been considered for evaluation of the performance for fusing multi-modal medical images. First one is NSCT [15]. Second method is relied on fragmented SR [16]. Third one is PCNN [18]. Fourth one is NSCT+SR [19]. Parameter settings involved in NSCT method: class of decomposition = 4 and scale decomposition filter = “pyrexc” while direction filter = “vk”.

4.1 *Simulation Details*

Comparative analysis has been demonstrated through Table 1 for the methods NSCT, PCNN, SR and NSCT+SR for the three metrics AvGr, StDv and SpFr. It is evident from Table 1 that PCNN is best performing method with highest values for all three metrics for both the images. Bold values denote the best performing results. Experimental results has shown that pulse coupled neural network (PCNN) has performed better as compared to NSCT, SR and NSCT+SR methods utilizing performance evaluators average gradient (AvGr), standard deviation (StDv) and spatial frequency (SpFr) with a best performing value of 10.4, 82.2 and 8.1, respectively. Graphical analysis has been depicted through Fig. 1a, b, respectively, while Fig. 2 represents visual analysis where fusion of MRI image and PET image has been demonstrated for PCNN, NSCT, SR and NSCT+SR fusion method. Visual analysis (see Fig. 2) involves an important place in medical image fusion research as well as facilitating medical practitioners to yield decisions involving diagnosis.

Table 1 Analysis using AvGr, StDv and SpFr

Images	Metrics	PCNN	SR	NSCT	NSCT+SR
1	AvGr	6.9	4.9	4.9	4.8
	StDv	71	50.1	70.3	50.8
	SpFr	7	5.7	6.1	5.9
2	AvGr	10.4	8.0	6.8	7.2
	StDv	82.2	54.3	70.7	56.6
	SpFr	8.1	7.0	6.7	6.7

4.1.1 Comparative Assessment

See Table 1.

4.1.2 Graphical Analysis

See Fig. 1. Here we have included graphical analysis for demonstrating comparisons of these four fusion strategies namely PCNN, SR, NSCT and NSCT+SR. As we can see through Fig. 1a and b PCNN is showing better performance for all three metrics.

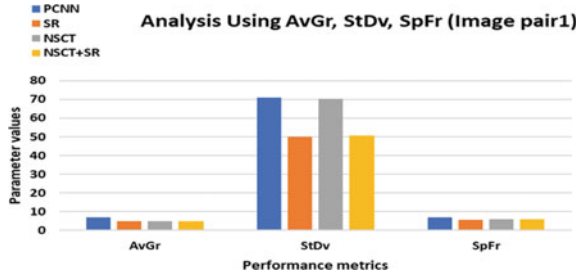
4.1.3 Visual Analysis

Please see Fig. 2.

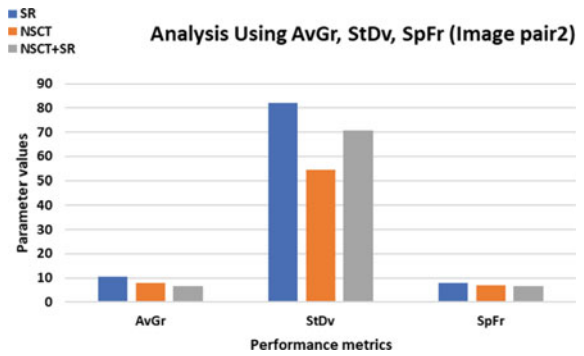
5 Conclusion and Future Directions

Through this research, we have evaluated four image fusion techniques PCNN, SR, NSCT and NSCT+SR utilizing three performance key indicators AvGr, StDv and SpFr on medical image dataset. Experiment related outcomes demonstrated that pulse coupled neural network (PCNN) has performed better as compared to NSCT, SR as well as NSCT+SR methods utilizing metrics average gradient (AvGr), standard deviation (StDv) and spatial frequency (SpFr) with a best performing value of 10.39, 82.15 and 8.06, respectively. In the future, more hybrid approaches may be adapted to evaluate the performance outcomes.

Fig. 1 a Analysis using AvGr, StDv, SpFr (Image pair 1). **b** Analysis using AvGr, StDv, SpFr (Image pair 2)



(a) Analysis using AvGr, StDv, SpFr (Image pair 1)



(b) Analysis using AvGr, StDv, SpFr (Image pair 2)

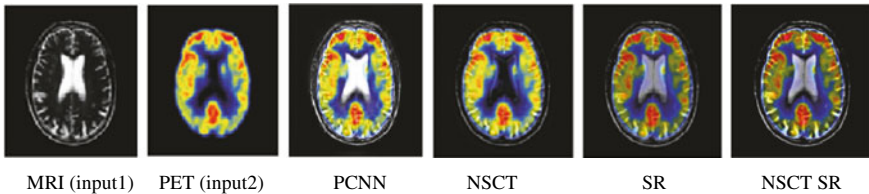


Fig. 2 Visual analysis

References

1. Fei Y, Wei G, Zongxi S (2017) Medical image fusion based on feature extraction and sparse representation. *Int J Biomed Imaging*
2. Singh V, Kaushik VD (2022) DTCWTASODCNN: DTCWT based weighted fusion model for multimodal medical image quality improvement with ASO technique and DCNN. *J Sci Ind Res* 81(08):850–858
3. Lei K et al (2022) Artifact-and content-specific quality assessment for MRI with image rulers. *Med Image Anal* 102344
4. Zago GT et al (2018) Retinal image quality assessment using deep learning. *Comput Biol Med* 103:64–70

5. Zhao C, Guo Y (2016) Fast image fusion algorithm based on sparse representation and non-subsampled contourlet transform. *J Electron Inf Technol* 38(7):1773–1780
6. Wang H et al (2022) Application of deep learning image reconstruction in low-dose chest CT scan. *Br J Radiol* 20210380
7. Reitboeck HJ et al (1990) A model for feature linking via correlated neural activity. In: *Synergetics of cognition*. Springer, Berlin, Heidelberg, pp 112–125
8. Ibrahim RW et al (2022) A medical image enhancement based on generalized class of fractional partial differential equations. *Quant Imaging Med Surg* 12(1):172
9. Das KP, Chandra J (2022) Multimodal classification on PET/CT image fusion for lung cancer: a comprehensive survey. *ECS Trans* 107(1):3649
10. Alagic Z et al (2022) Deep learning versus iterative image reconstruction algorithm for head CT in trauma. *Emerg Radiol* 1–14
11. Bhutto JA et al (2022) CT and MRI medical image fusion using noise-removal and contrast enhancement scheme with convolutional neural network. *Entropy* 24(3):393
12. Liu J, Lian J, Sprott JC, Liu Q, Ma Y (2022) The butterfly effect in primary visual cortex. *IEEE Trans Comput*
13. Singh V, Kaushik VD (2021) WeAbDeepCNN: weighted average model and ASSCA based two level fusion scheme for multi-focus Images. *J Sci Ind Res* 80(10):905–914
14. Singh V, Kaushik VD (2021) HoEnTOA: Holoentropy and Taylor assisted optimization based a novel image quality enhancement algorithm for multi-focus image fusion. *J Sci Ind Res* 80(10):875–886
15. Zhang Y-J (2022) A selection of image processing techniques: from fundamentals to research front. CRC Press
16. Ch M et al (2022) Medical image fusion using non subsampled contourlet transform and iterative joint filter. *Multimedia Tools Appl* 81(3):4495–4509
17. Li Y et al (2018) An image fusion method based on sparse representation and sum modified-Laplacian in NSCT domain. *Entropy* 20(7):522
18. An FP, Ma XM, Bai L (2022) Image fusion algorithm based on unsupervised deep learning-optimized sparse representation. *Biomed Sig Process Control* 71:103140
19. Shehanaz S et al (2021) Optimum weighted multimodal medical image fusion using particle swarm optimization. *Optik* 231:166413
20. Wang Z et al (2021) Medical image fusion based on convolutional neural networks and non-subsampled contourlet transform. *Exp Syst Appl* 171:114574
21. Sengupta D, Gupta P, Biswas A (2022) A survey on mutual information based medical image registration algorithms. *Neurocomputing* 14(486):174–188
22. Li S, Kwok JT, Wang Y (2001) Combination of images with diverse focuses using the spatial frequency. *Inf Fusion* 2(3):169–176
23. Medical image dataset taken from <http://www.med.harvard.edu/AANLIB/home.html>. Accessed on January 2022

Audio and Text-Based Emotion Recognition System Using Deep Learning



Palash Thakur, Ronit Shahu, and Vikas Gupta

Abstract Emotion is inherent in people, and hence, emotion comprehension is a critical component of human-like Artificial Intelligence (AI). Because of its ability to mine opinions from a plethora of publicly available conversational data on platforms such as Facebook, YouTube, Reddit, Twitter, and human computer interaction (HCI), assessing emotions is becoming increasingly popular as a new research frontier in natural language processing (NLP). This paper demonstrates the implementation of audio- and text-based emotion recognition algorithm. We use MELD data set to train the models. We implement both audio-based and text-based models individually as well as the combined model and show their respective results. Our results show that the Audio + Text-based model outperforms audio- or text-based models in terms of weighted average F1-score. The Audio + Text-based model gives the weighted average F1-score of 70%.

Keywords Emotion recognition · Human computer interaction · Natural language processing · Classification

1 Introduction

In this modern age of technological advancement, people want everything to be smart, whether it is their phone, Television (TV), Air Conditioner (AC), or any other electronic device. While making a lifeless machine smart, one of the key factors is emotion recognition. Machines have to be smart enough to understand when the user is happy or sad. These critical issues can be addressed by HCI. Such smart machines should not ignore any of the emotions as that would lead to a failure of the system. Emotion recognition using audio and text processing via deep learning techniques is a challenging field in the domain of AI. A speech-based emotion recognition system works in two parts. In the first part, we extract features from a given audio input, and in the second part, we classify the emotions based on the extracted features. We

P. Thakur (✉) · R. Shahu · V. Gupta
Shri Ramdeobaba College of Engineering and Management, Nagpur, Maharashtra, India
e-mail: thakurpp@rknec.edu

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
V. V. S. S. Chakravarthy et al. (eds.), *Advances in Signal Processing, Embedded Systems and IoT*, Lecture Notes in Electrical Engineering 992,
https://doi.org/10.1007/978-981-19-8865-3_41

447

have proposed two deep learning-based models for the classification of the emotions. The first proposed model takes an audio file (.WAV) as input. We extract the features from the audio file using OpenSMILE [1]. The features extracted from the audio-based model is then passed to fully connected layer for classification. The second proposed model takes a text file (.TXT) as input. We use word2vec algorithm for feature extraction. These extracted features are then passed through fully connected layer for classification.

2 Literature Survey

Many authors have done very intriguing work in the domain of emotion recognition using deep learning. Most of the authors have discussed methodologies which mainly include text, audio as well as facial features while classifying human emotions.

In [2], Heysem has proposed a model which is divided into five parts namely face alignment, pre-processing/CNN-training, feature extraction, modelling, and prediction. The facial alignment is done using principal component analysis (PCA). It basically dealt with false positives and rotated faces. After the PCA, the images with high mean reconstruction error per image is discarded since the image is probably poorly detected or misaligned. For the CNN training part, the authors used a pre-trained model via transfer learning and fine tuning it with the FER 2013 data set. For the pre-trained model, the authors used VGG-Face model that is used for face recognition. For feature extraction from the already fine-tuned CNN model, the authors re-scaled the images in 224×224 pixels and then normalized them by subtracting the average image of the VGGFace network. Heysem concluded that the multi-modal system diminishes the results for natural/semi-natural data.

In [3], the authors initialized each token with pre-trained 300 dimensional GloVe vectors and then fed them to 1D-CNN to extract 100 dimensional textual features. For the feature extraction of audio input, OpenSMILE toolkit is used. It extracted 6373 dimensional features. As the audio representation is high dimensional, the authors employed L2-based feature selection with sparse estimators like support vector machines (SVMs), to get a dense representation of the overall audio segment. Later, the audio and textual features are concatenated to obtain a bi-modal feature. Strong benchmark was to be given to MELD. In order to do that different models such as textcnn, bclstm and Dialogue RNN was used. Among these, the DialogueRNN was suited the best as it can handle multi-party conversations. The results obtained from such experiments are that the DialogueRNN with multi-modal variant gave the best performance, i.e. 67.59% F-Score. It surpassed the bclSTM which has a F-Score of 66.68%. The authors also concluded that the textual modality (57.03% F-Score) is better than the audio modality(41.79% F-Score) by 17%. For the positive sentiments, the audio model performed poorly. The multi-modal fusion did increase the emotion recognition by 3%. However, the multi-modal classifier performed worse than the textual classifier while classifying sadness. The performance of some particular

classes like sadness, disgust and fear are very poor. The main cause of this is the imbalance of the data points in these classes in the MELD data set.

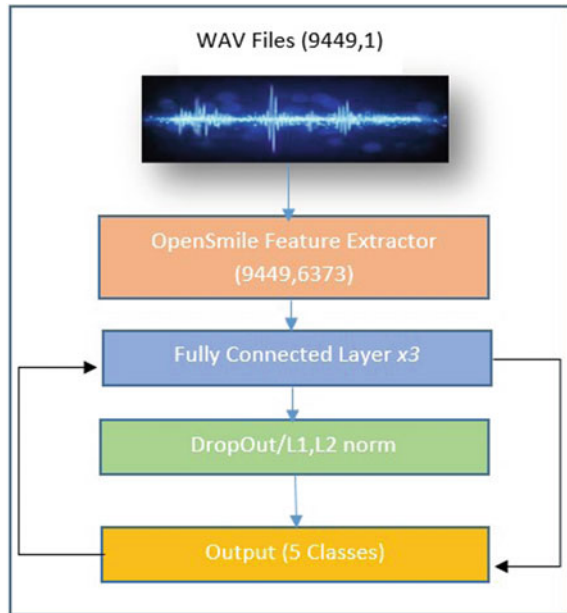
In [4], the authors divided the complete system into four parts, namely video+audio pre-processing, feature extraction, feature fusion and classifier. For the video pre-processing part, the authors used the Dlib toolbox and extended the face bounding box with a ratio of 30%. The cropped faces were then resized to 224×224 pixels. If no face was detected in an image, the entire frame was passed in the network. For audio feature extraction, the speech spectrogram was passed through a hamming window with a 40 ms window size and 10 ms window shift. The 200 dimensional low frequency parts of the spectrogram were used for audio modality. The authors used three different CNN models to extract the facial features namely VGGFace, ResNet18, IR50. For the feature extraction of audio files, they extract the feature maps of the audio from the last pooling layer of AlexNet. The size of a three-dimensional feature map is $H \times W \times C$, where H , W are the height and width of the feature map and C is the number of the channel of the feature map. The feature maps are then split into n vectors ($n = H \times W$). The authors used four different emotion data sets, namely AffectNet, RAF-DB, FER+, AFEW. The AffectNet data set had emotion labels. The RAF-DB data set had both 7-class basic and 12-class compound emotion labels in them. Only the basic emotion label were used. The FER+ and AFEW did not have such labels. When all the three trained models, i.e. the ResNet18, the IR50 and the VGGFace were compared, the authors found that the IR50 was superior to the rest. The well-trained IR50 was then used on all the four above-mentioned data set, and it was found that the IR50 when pre-trained on AffectNet gave the highest result of 53.78%. The authors used speech spectrogram for Audio CNN which gave 38% on AFEW validation set. Log Mel-Spectrogram was also used which gave a slight better accuracy. The authors concluded that the three main intra-modal fusion techniques, namely self-attention, relation-attention and transformer-attention was used which gave better accuracy. For fusion of audio and video feature, feature concatenation and factorized bi-linear pooling was used. The result obtained was 62.48% and ranked 2nd in EmotiW 2019 challenge.

3 Data Set

3.1 Selection of Data Set

Many available data sets in multi-modal sentiment analysis and emotion recognition are non-conversational. But IEMOCAP [5], SEMAINE [6] and MELD [3] are one of the most popular conversational data sets where each utterance in a dialogue corresponds to an emotion (a label). The MELD was created by extending EmotionLines data set and enhancing it further more. It consists of audio and visual modality along with text. MELD has more than 1400 dialogues and 13,000 utterances from a famous TV series named "F.R.I.E.N.D.S". Multiple speakers participated in some dialogues.

Fig. 1 Audio model architecture



Each utterances in any dialogue is labelled as any of the seven emotions, viz. anger, disgust, sadness, fear, happy, surprise and neutral.

MELD has seven emotions classes, namely “anger”, “surprise”, “neutral”, “disgust”, “fear”, “joy” and “sadness”. The class data distribution of each of the above-mentioned class is shown in Fig. 1. The data set includes video clips of the same show “F.R.I.E.N.D.S”. These clips range from 2 to 13 s having a resolution of 1280×720 pixels. The total number of video files is 15,907. The data set was clearly imbalanced as the majority of data points belonged to the “neutral” class. The classes “anger”, “fear”, “disgust” and “sadness” are considered negative. The class “joy” is considered positive and the “neutral” class is labelled neutral. Surprise is a complex emotion since it can be considered as both positive and negative sentiment. The entire task of sentiment annotation reaches a Fleiss’ kappa score of 0.91. On average, three emotions were present in each of the dialogues of data set. The average utterances duration is 3.59 s.

3.2 Preparation of MELD Data Set for Training

MELD data set is a conversational data set. It has visual, audio and textual information provided for each speaker. Since this paper mainly focuses on audio and textual features of speaker, the visual data is omitted, and only audio is extracted from video. Every video file has a dialogue associated with it. These dialogues are used

Table 1 Distribution of data set for training, validation and testing

Training	Validation	Testing
7559	1890	2000

Table 2 Distribution of class samples for testing

Emotions	Count
Anger	226
Joy	390
Neutral	1005
Sadness	115
Surprise	264

for preparation of text data set. The data set is highly imbalanced. “Disgust” and “fear” are two emotions which contribute to less than 3% of total data. Training the model with entire data set showed that these two classes are resulting in zero recall values. These two emotions are removed, and training and testing is done on only five classes, namely “anger”, “surprise”, “neutral”, “sadness” and “joy”. 9449 samples are used for training and validation. Testing is done on 2000 samples. The distribution of data set for training, validation and testing is explained in Table 1. 20% data is used for validation. Table 2 refers the distribution of class samples for testing.

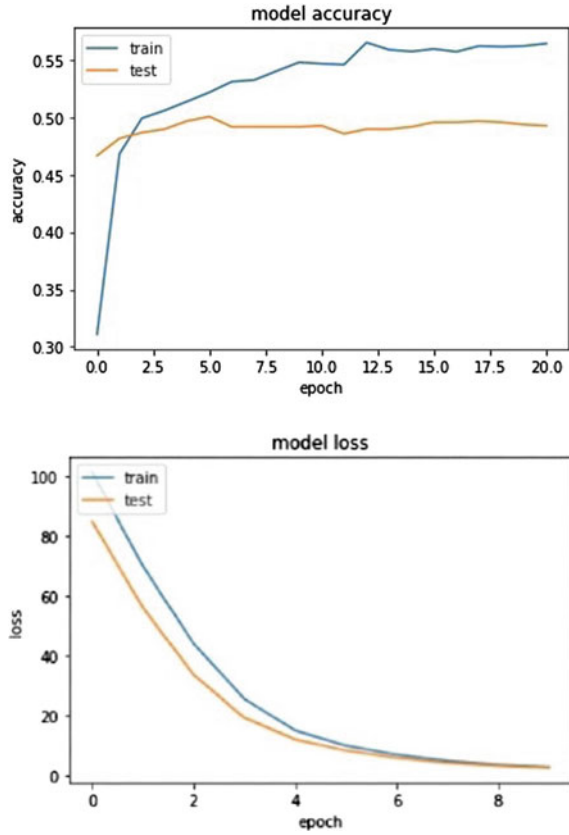
4 Audio Model

The extracted audio files are in Waveform Audio File (WAV) format. By doing so, we created a entirely new audio data set with corresponding emotion to it. There are total 9449 audio files in the data set. The sole task of this model is to detect emotion entirely from audio of a person. MELD provides us with conversational data set where more than two people are talking with each other but one at a time. While training audio model, the base assumption is to predict emotion of a single person while he is speaking to other person. OpenSMILE is used as feature extractor for the purpose of transfer learning.

4.1 Training of Audio Model

The audio files are passed through OpenSMILE for feature extraction. OpenSMILE expects input to be a WAV file. Dimension of feature vector after feature extraction is

Fig. 2 Accuracy and loss



(9449, 6373). Since the feature vector is high in dimension, we normalized data for faster training and to improve accuracy. The block diagram representation of audio model is shown in Fig. 1.

The input to the model is 6373 dimensional vector space and output is probability distribution among five classes, namely sadness, fear, neutral, surprise and joy. Due to high dimensionality of feature vector, model was overfitting resulting in high variance between training and validation data. To prevent model from overfitting, dropout and L_1 and L_2 norm like regularization techniques are used. The activation function used is ReLU activation, and Adam optimizer is used for optimization of model with learning rate of 0.0001 and reducing learning-rate by a factor of 0.2 once learning stagnates, and there is no improvement even after few epoch. The audio model is trained for ten epochs. The training accuracy of audio model is 52.22% and validation accuracy is 50.40%. The graph of accuracy and loss is shown in Fig. 2.

$$new_learning - rate = 0.2 * old_learning - rate \tag{1}$$

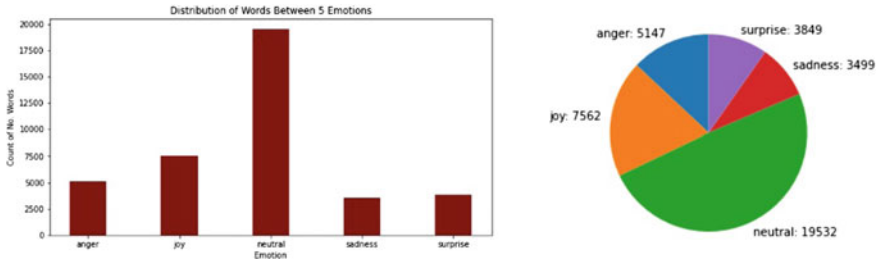


Fig. 3 Bar graph and pie chart representation of distribution of words among five classes

Table 3 Tabular representation of raw text data

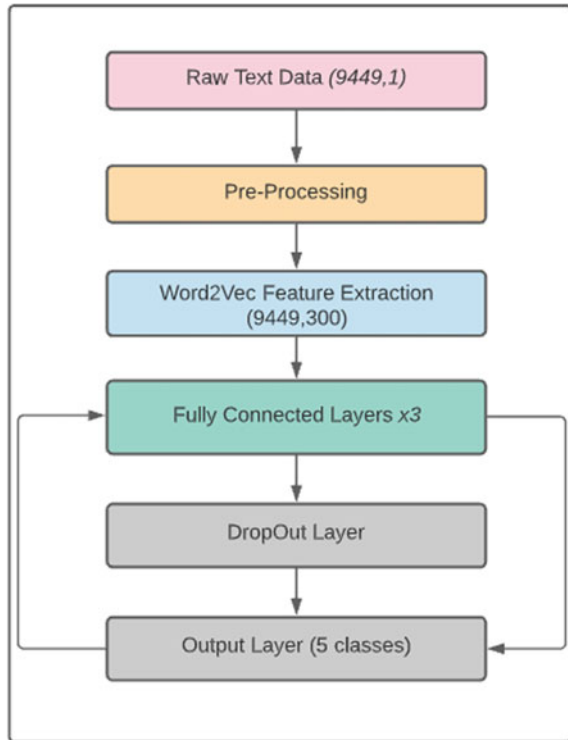
Utterances	Emotions
Also I was the point person on my company’s transition from the KL-5 to GR-6 system	Neutral
My duties? All right	Surprise
What?! What is with everybody? It’s Thanksgiving, not... Truth-Day!	Anger
Do I ever	Joy
Oh, totally. Oh, God, oh, she seemed so happy too	Sadness

There are 260 different speakers in the MELD utterance data. The aim of this model is to solely predict emotion simply from text irrespective of speaker’s name or expression. While making this as base assumption for building data pipeline, we considered only utterance and emotion corresponding to that person. There are total 39,589 words in all utterances. The distribution of words across every emotion is given in bar chart as well as pie chart in Fig. 3. It is clear from this graphical representation that the neutral class emotion is dominant and is occupying almost 50% of the total data. Sample utterances from training data across every emotion is given in Table 3.

4.2 Training of Text Model

We used the word2vec model for transfer learning for training text model and feature extraction. The word2vec algorithm uses a neural network model to learn word associations in a large collection of text. Word2Vec is a state-of-the-art model and requires pre-processing of text before being fed into the word2vec model. This pre-processing includes removing stop words and punctuation, tokenizing and then averaging.

Fig. 4 Text model architecture



The extracted features are then fed into fully connected layers and tested with several newly introduced activation functions. The output of this model is the probability distribution among the five classes obtained using the “SoftMax” activation function. Due to high dimensionality of feature vector, model was prone to overfitting, which happened quickly. Overfitting causes reduced generalization ability of a model and results in high variance of test data. To avoid overfitting and the problem of high variance, some of the regularization techniques such as dropout as well as L_1 and L_2 norm are introduced in the model to converge model’s training accuracy and validation accuracy.

The block diagram of text model is shown in Fig.4. The dimension of feature vector before pre-processing was (9449, 1) while the dimension after pre-processing became (9449, 300). The text model follows the same data distribution as explained in Table 1. The text model is trained for 20 epochs. The training accuracy of audio model is 57%, validation accuracy being 49.65%. The graph of accuracy and loss is shown in Fig.5.

Fig. 5 Accuracy and loss

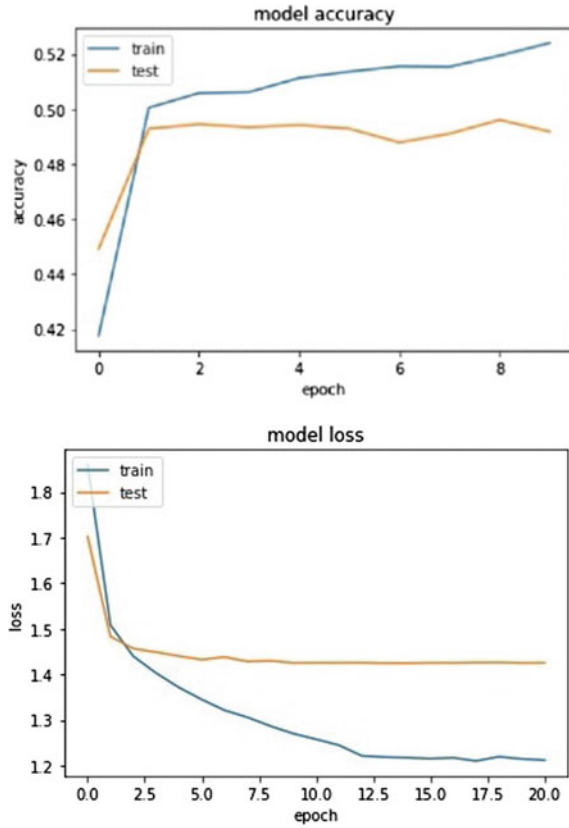


Table 4 Result summary of audio, text and audio + text model

Model	Emotions						Correct predictions (%)
	Anger (%)	Joy (%)	Neutral (%)	Sadness (%)	Surprise (%)	Weighted f1 (%)	
Audio	72	45	84	8	62	68	1328
Text	19	38	71	23	38	52	1040
Audio + Text	63	64	80	18	67	70	1070

5 Audio + Text Model

In the multi-modal approach, we aim to combine both the models and create a robust and more accurate emotion detection system which can identify emotions into five classes with more confidence. The combination of models can be done on the basis of the their output. Using ensemble-based techniques is other option

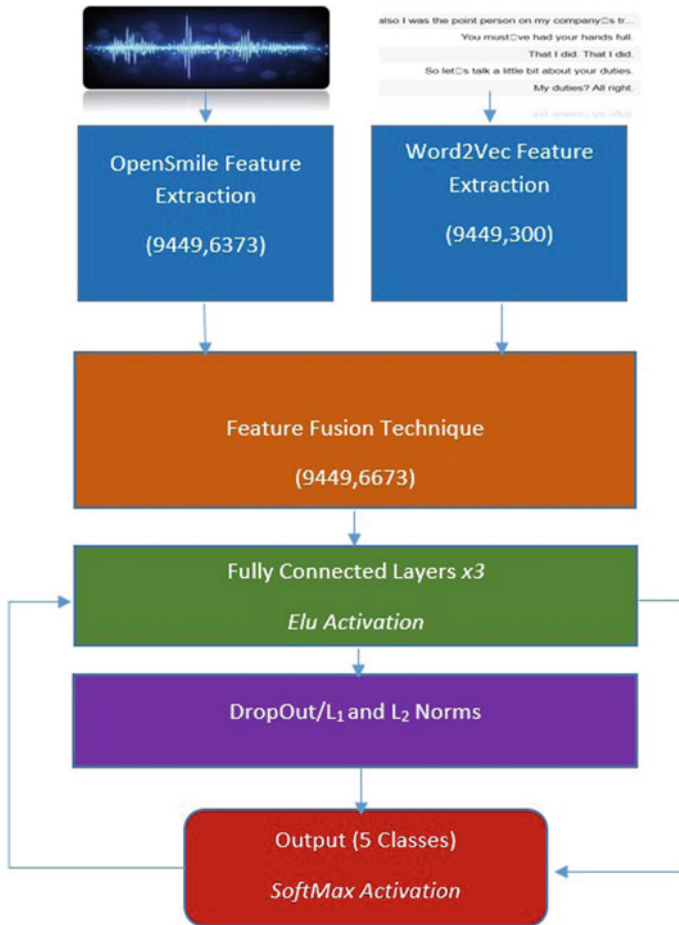
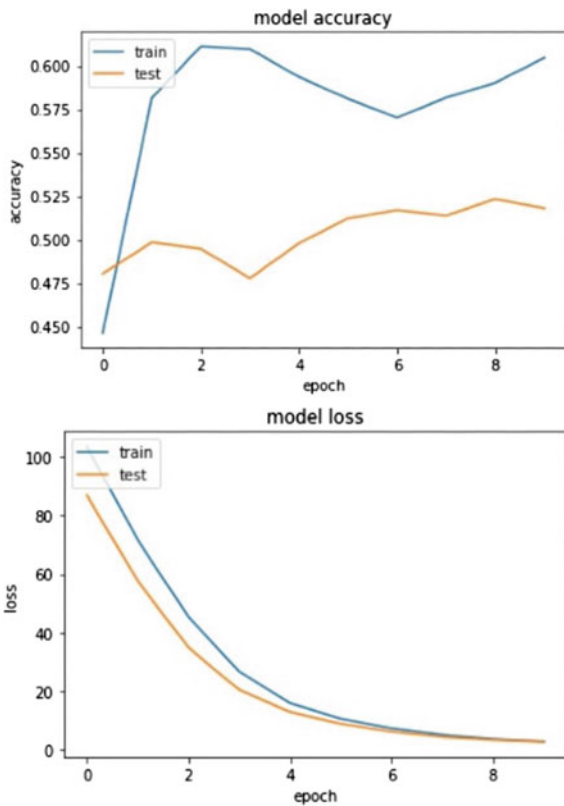


Fig. 6 Audio + text model architecture

to average out the output probability of both the models, so that the output with highest probability is obtained as the result. But the problem with this approach is that the model will not be able to learn the relationship between the audio and the text features, as well as the corresponding emotions. To avoid this problem, this paper uses feature fusion strategies. Feature fusion strategies are some mathematical equations used to merge different linear features from different models. There are some strategies predefined for feature fusion. Based on the performance of different strategies, feature concatenation is used to build the model. Some of the feature fusion strategies are

1. Feature Concatenation Concatenating two features by their columns.

Fig. 7 Accuracy and loss



$$X := X_{Audio}(9449, 6373) + X_{Text}(9449, 300) \tag{2}$$

$$Dimension(X) = (9449, 6673)$$

2. Maximum value extraction

$$X[i][j] = MAX(X_{Audio}[i][j], X_{Text}[i][j]) \tag{3}$$

$$Dimension(X) = (9449, 6373)$$

3. Product of two features

$$X[i][j] = X_{Audio}[i][j] * X_{Text}[i][j] \tag{4}$$

$$Dimension(X) = (9449, 6373)$$

After testing above strategies while building this model, feature concatenation is outperforming all the other strategies, and it is used to as a final feature fusion method

5.1 Training of Audio + Text Model

The block diagram representation of the audio and text model is given in Fig. 6. Features are extracted at individual stages from both audio and text files. Extracted features are then stacked by column, extending entire feature vector length to 6673 containing both audio and text features. The features that are extracted belong to different categories, and hence, they have different data distribution. Since the features are combined, a new normal distribution must be defined. Hence, entire feature space containing 6673 features are normalized again making normal distribution over 6673 features. After normalizing data, the next step is training the model. Fully connected layers are used to simply train the model. The activation function used in fully connected layers is “ELU” activation. The α value used here is 1.0. The optimizer used is “Adam” with a predefined learning rate of 0.0001, reducing the learning-rate by a factor of 0.2 once learning stagnates, and there is no improvement even after a few epoch. The dimension of feature vector is very high, resulting in model overfitting and high variance between training and validation data. To avoid this problem, dropout as well as L_1 and L_2 norm have be used for regularization. The final layer of model outputs probability distribution among five classes using “SoftMax” activation function. The audio + text model is trained for ten epochs. The training accuracy of audio model is 60%, and validation accuracy is 53.60%. The graph of accuracy and loss is shown in Fig. 7.

6 Result and Discussion

Every model is tested on 2000 samples having different number of test samples across each class. The accuracy metric used in this paper to compare model’s performance is weighted F1-score. The weighted average of each model is calculated by multiplying F1-score of each class with count of samples across that class then adding them and finally dividing them by total number of samples which is 2000. The audio model has highest prediction accuracy with correct prediction of 1328 out of 2000. But the model is not confident while making prediction. The audio+text model has predicted 1070 samples out of 2000 samples correctly and has highest F1-score of 70% as given in Table 4. The model is most confident when making a decision in this case. The detailed result summary of all the models is given in Table 4. Audio + Text model is able to pick samples across every class with moderate confidence. One of the reason for low accuracy is imbalanced class data set. If data set can be balanced, then accuracy will improve. Good feature extractor’s output is higher dimensional vector

space resulting in model's overfitting over a few epochs. The controlled training is done in this paper along with regularizing the model, hence avoiding overfitting even with high dimensional input. The text model results in lowest F1-score of 52%.

7 Conclusion

In this paper, we illustrate implementation of a multi-modal emotion recognition system using MELD data set. The multi-modal approach is implemented using audio and text features of the MELD data set. Audio model is giving good accuracy as compared to the text model. The reason of text model performing poor is due to the fact that the data set is highly imbalanced. Overall multi-modal accuracy can be improved with some advanced and enhanced feature extractors for both audio as well as text features. If the text model's accuracy is improved, then the overall Audio + Text model's accuracy can be improved.

References

1. Navyasri M, Ramakrishnamurty M et al (2017) Robust features for emotion recognition from speech by using Gaussian mixture model classification. In: International conference and published proceeding in SIST series, August 2017, vol 2. Springer, pp 437–444
2. Kaya H, Gürpınar F, Ali Salah A (2017) Video-based emotion recognition in the wild using deep transfer learning and score fusion. *Image Vis Comput* 65:66–75
3. Poria S et al (2018) Meld: a multimodal multi-party dataset for emotion recognition in conversations. arXiv preprint [arXiv:1810.02508](https://arxiv.org/abs/1810.02508)
4. Zhou H et al (2019) Exploring emotion features and fusion strategies for audio-video emotion recognition. In: 2019 international conference on multimodal interaction
5. Busso C et al (2008) IEMOCAP: interactive emotional dyadic motion capture database. *Lang Resour Eval* 42(4):335–359
6. McKeown G et al (2011) The semaine database: annotated multimodal records of emotionally colored conversations between a person and a limited agent. *IEEE Trans Affect Comput* 3(1):5–17
7. Eyben F, Wöllmer M, Schuller B (2010) Opensmile: the Munich versatile and fast open-source audio feature extractor. In: Proceedings of the 18th ACM international conference on multimedia
8. Praneel ASV, Srinivasa Rao T, Ramakrishna Murty M (2019) A survey on Accelerating the classifier training Using various boosting schemes within cascades of boosted ensembles. In: International conference with Springer SIST series, vol 169, pp 809–825
9. Mikolov T et al (2013) Efficient estimation of word representations in vector space. arXiv preprint [arXiv:1301.3781](https://arxiv.org/abs/1301.3781)

Land Use Land Cover Segmentation of LISS-III Multispectral Space-Born Image Using Deep Learning



Nirav Desai and Parag Shukla

Abstract Remote sensing information provides important and sensed data. The study represents semantic segmentation using a fully convolutional network (FCN) for semantic segmentation. Semantic segmentation is a pixel-level classification of images where each pixel is assigned to a respective class. In this present study, four classes—Water Bodies, Vegetation, Uncultivated Land, and Residential areas—were identified. There are various types of machine learning (ML) models as well as deep learning (DL) models to handle segmentation tasks. In this study, deep neural network was used. A fully convolutional network (FCN) with skip connections is trained to take an input image of size $256 * 256 * 3$ and outputs a matrix of shape $256 * 256 * 4$, i.e., a one-hot encoded version of the mask. The experiment showed that the FCN classifier has a very good capability for land use land cover class detection. The model identifies four classes with 81% of OAA.

Keywords Remote sensing · Multispectral image · Land use land cover classification · Deep learning · Fully convolutional network (FCN)

1 Introduction

In recent eras, the consumption of land resources has become a serious problem. Remote sensing (RS) is the art of finding and understanding data from a long distance, using sensors without communication with the object being observed [1]. Land use land cover classification aims to organize space-born images into an exact class, which were reliant on the distribution of recognized land use land cover classes. Land use land cover classification aims to organize space-born images into an exact class, which was reliant on the distribution of recognized land use land cover classes and involve attention in the last few years due to a variety of applications like urban development, monitoring of natural tragedies, and land use land cover analysis [2–5]. The

N. Desai (✉) · P. Shukla

Department of Computer Applications, Atmiya University, Rajkot, India

e-mail: niravdesai.research@gmail.com

remote sensed image consists of Residential areas, agricultural land, Uncultivated Land, Water Bodies, and other open areas.

Hinton [6] et al. projected deep learning and demonstrated that training difficulties of a deep neural network can be solved using one-by-one layer initialization and effectively applicable to the field of video and image processing, a field of data analysis [7]. In the DL mechanism, the machines learn from the information themselves by growing more layers of a network [8].

Fully convolutional networks are applied for the classification of the land use land cover of the South Gujarat region, India. The success of the classifier was tested on physical data. In the present study, a total of four classes have been classified—Water Bodies, Vegetation, Uncultivated Land, and Residential areas. Semantic segmentation is defined as a pixel-level classification of images where a class is allotted to an individual pixel of the image. There are various types of machine learning models and deep learning models to handle segmentation tasks. In this study, deep neural network was applied to handle this task. A fully convolutional network (FCN) with skip connections is trained to take an input image of size $256 \times 256 \times 3$ and outputs a matrix of shape $256 * 256 * 4$, i.e., a one-hot encoded version of the mask. The model achieved an OAA of 81%.

2 Literature Review

Deep learning (DL) belongs to machine learning techniques, where various layers of data processing phases in ordered architectures are overburdened by unsupervised learning and pattern classification [9]. Chen et al. [10] presented hyperspectral image classification with a hybrid framework that contains deep learning and logistic regression. Zuo et al. applied deep belief networks [11], and the result exposed the success of the model. Liu et al. and Piramanayagam et al. [12, 13] used CNN for land use land cover segmentation, where designated training data in each repetition with DL achieved good results. The deficiency of labeled data was managed by applying data augmentation techniques [14]. Yang et al. [15] is merging CNN and multi-scale feature fusion on the controlled data. Guo et al. [16] studied Saliency Dual Attention Residual for achieving good performance. Xu et al. [17] projected classification technique which includes the neural network with random forest (RF) for land use land cover classification. Alichiri et al. EfficientNet-B3 CNN [18] tested it on six common land cover–land use datasets and showed the efficiency in space-borne image scene classification. This resulted in an important enhancement in overfitting and better accuracy. Pires de Lima and Marfurt [19] proposed strategies that improved convolutional neural networks for aerial image segmentation.

Semantic segmentation is the job of appointing every pixel in an image to the defined various classes. It identifies which substances are displayed and where they are obtainable [20]. DL and CNN have transformed the image classification part over a long period and now became a central method for image classification [16]. Lateef and Ruichek [21] gives an organized and comprehensive evaluation of various groups

of applying DL methods. Hao et al. [22] placed diverse stress on the command during training. Many techniques are abridged in [23], and new advances are given detailed in [24, 25].

3 Materials and Methods

3.1 Data Collection

Multispectral space-born image (LISS-III) was used for the study. It has more than 100 nm resolution and less the ten bands. The LISS-III image contains a total of four bands; the spatial resolution is 30 m. Quadrats of 30 m * 30 m size were laid down across the study area. Data were collected from the website <https://bhuvan-app3.nrsc.gov.in>. The data were processed using the software ENVI4.7. A wide field study was completed to collect environmental landscapes and circulation patterns of different land uses and land covers. And, the longitude and latitude of the location for the respective class are recorded (GCPs). To record GCPs, the GPS device Garmin-eTrex 30 is used for GCPs' collection. The GCPs reserved for the respective category were reliant on the allocation of recognized land use classes within the study area. A LISS-III multispectral remote sensing image consists of four different bands in separate.tiff files and the number of bands is Band-2, 3, 4, and 5 (blue, green, red, and near-infrared).

3.2 Preprocessing

For the land use land cover classes' inspection in the study, the LISS-III space-born multispectral images were merged into a false-color composite (FCC) image. False-color composites (FCCs) are created by stacking these multiband. TIFF images are on top of each other and take a stacked grouping of Band-4, Band-3, and Band-2 to generate FCC. After creating FCCs for each image, ground truth masks were created and used to train the deep learning model. These masks are created using the maximum likelihood algorithm on a small region of interest for each class (Fig. 1).

Both the FCCs and their corresponding masks are resized to 1024 * 1024 pixels and then divided into patches of size 256 * 256 pixels with a striding of 128. This will create 49 patches of size 256 * 256 for each image.

Preprocessing steps:

1. False-color composites (FCCs) are created by stacking these multiband TIFF images on top of each other and stacked Band-4, Band-3, and Band-2 to generate FCC.

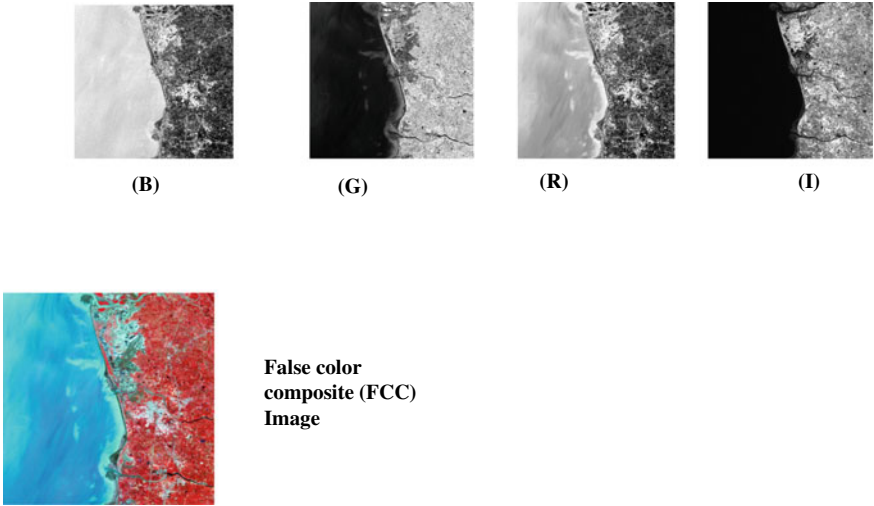


Fig. 1 FCC creation

2. After creating FCCs for each image, ground truth masks were generated which were used to train the deep learning model. These masks are created using the maximum likelihood algorithm on a small region of interest for each class.
3. Both the FCCs and their corresponding masks are resized to $1024 * 1024$ pixels and then divided into patches of size $256 * 256$ pixels with a striding of 128. This will create 49 patches of size $256 * 256$ for each image. Thus, the size of our dataset would be $30 * 49 = 1470$ images.
4. These images and masks are separated into two subgroups, one for training and another for validation. About 1255 images are used to train the model, while 215 images are reserved for validation and evaluation (Fig. 2).

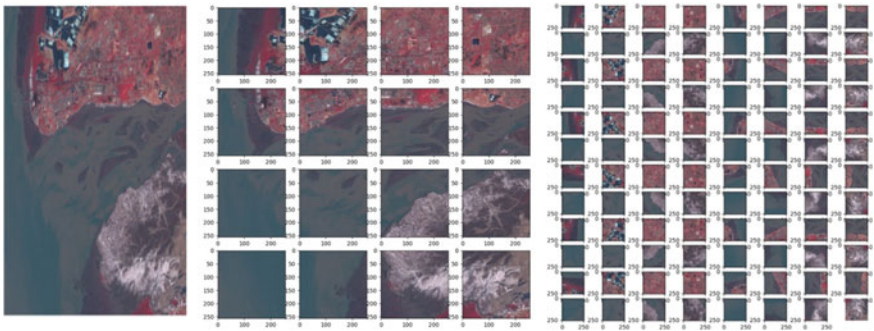


Fig. 2 FCC (1024×1024) (256×256)

3.3 Methods

A maximum likelihood classifier (MLC) classification is used with space-borne image data, in which a pixel with the maximum likelihood is classified into the corresponding class. In MLC, a pixel is selected for a class based on its chance of fitting. Mean vector and covariance metrics are the main constituent of MLC that can be recovered from training data [6].

The following is discriminant functions calculated for each pixel:

$$g_i(x) = \ln p(\omega_i) - 1/2 \ln |\Sigma_i| - 1/2(x - m_i)^t \Sigma_i^{-1}(x - m_i) \tag{1}$$

where i is class, and x is n -dimensional data in which n represents the total number of bands. $p(\omega_i)$ represents the chance that class ω_i occurs in the image, $|\Sigma_i|$ is the determinant of the covariance matrix, and Σ_i^{-1} is an inverse matrix the mean vector represents by m_i (Fig. 3).

After creating FCCs for each image, ground truth masks were created that will be used to train a model. These masks are created using the maximum likelihood algorithm on a small region of interest for each class (Fig. 4).

Fully convolutional networks [27] are effectively applied in various fields, such as segmentation of an image [28, 29], medicinal image analysis [26, 30], character recognition [31]. FCN contains relative within an entirely linked layer in the prior credit for each activation while seen during neural networks [32]. CNN and FCN

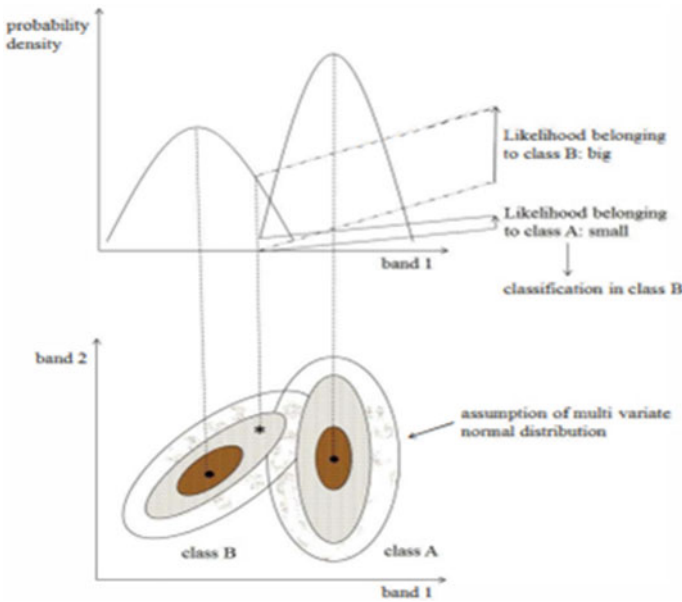


Fig. 3 Basic concept of ML [26]

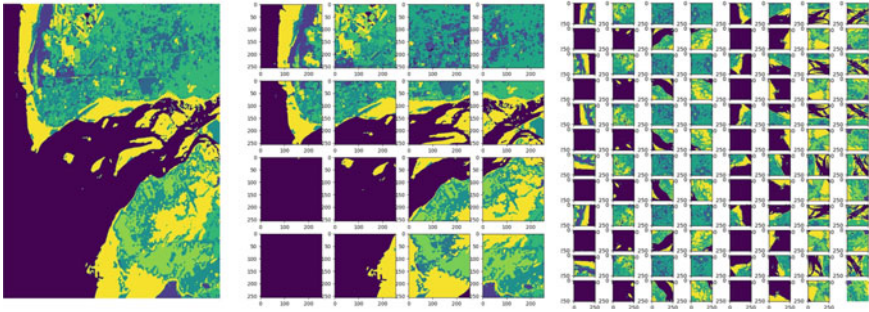


Fig. 4 Ground truth mask

were applied for the mangrove classification [33–35]. FCN is broadly applied in pixel-based classification [36]. And used an encoder for feature extraction and a decoder to reestablish the. FCN uses an encoder for feature extraction and a decoder to re-establish the input resolution by deconvolutional or upsampling layers [37].

3.4 Training Configuration

While data ingestion, i.e., before passing the images and masks to the model for training, we normalize the input images by clipping them to [0.0, 255.0], while the masks are one-hot encoded according to the total number of classes, i.e., 4. In addition, random augmentations are also applied to the batch of images and masks before passing them to the model for training. This expands our dataset and makes the model robust enough to encounter different orientations than just the training data. Data augmentation is a factor that prevents overfitting when it is done correctly. A custom image data generator is created to fulfill the requirements of this data ingestion pipeline.

The following is the table of hyperparameters and other configurations used for training the model.

3.5 Model Structure

Semantic segmentation is defined as a pixel-level classification of images where a class is allotted to each pixel of the image. In the present study, there are four classes—Water Bodies, Vegetation, Uncultivated Land, and Residential areas. A deep neural network was used to handle this task. A fully convolutional network (FCN) with skip connections is trained to take an image input of size $256 * 256 * 3$ and outputs a matrix of shape $256 * 256 * 4$, i.e., a one-hot encoded version of the mask. The FCN is a U-Net architecture that contains an encoder part and a decoder part. The encoder

part contains five blocks, and each block is two (convolution + batch normalization + ReLU) layers stacked on top of one another and followed by a max-pooling except for the last block. The output of this encoder part is then inputted to the decoder containing four blocks. Each block in the decoder starts with an upsampling of the input followed by a $1 * 1$ convolution operation. A skip connection is also used that concatenates the output of the corresponding encoder block to the output of the upsampling and convolution operation. The concatenated tensor is then again passed to two convolution layers similar to that of the corresponding encoder block. The output of the decoder part is finally fed to a $1 * 1$ convolution with the amount of filters equivalent to the amount of classes which is four.

The arrows denote the various processes, the black containers denote the feature map, and the gray containers denote the cropped feature maps from the contracting path.

$$E = \sum w(x) \log(P_{k(x)}(x)) \quad (2)$$

where p_k is the pixel-wise softmax function applied over the final feature map.

$$P_k(x) = \frac{e^{a_k(x)}}{\sum_{k=1}^k e^{a_k(x)}}, \quad (3)$$

And, $a_k(x)$ denotes the activation in channel k .

3.6 Algorithm Steps

The steps of the FCN algorithm are as follows:

1. The preprocessing steps include remotely sensed image alteration, registration, and the masking of the image. The images contain Band-2 to Band-4.
2. For training and testing, a total of four types of classes were selected, such as Water Bodies, Vegetation, Uncultivated Land, and Residential areas. This study used the ENVI image processing software (ROI Tool) to pick the four types of classes.
3. Total of four types of classes were used for building and training of the model. The model structure is shown in Fig. 5. The model structure and parameters were kept for successive image segmentation or classification after training.
4. The FCN model which was trained in step-3 was applied for classification.
5. Find whether all the classifications were finished or not. If all is done, then the classification outcome will be displayed and terminated the algorithm (Fig. 6).

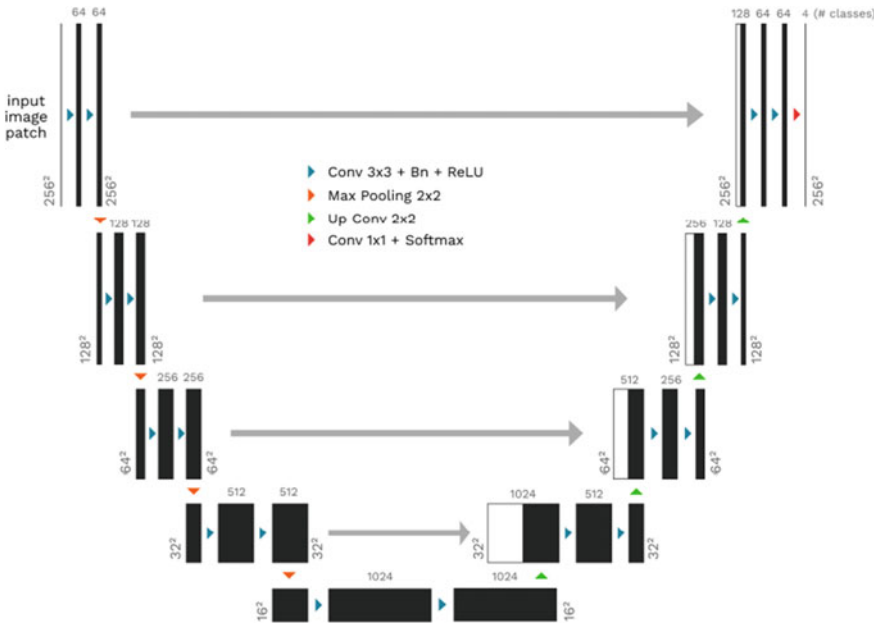


Fig. 5 U-Net architecture

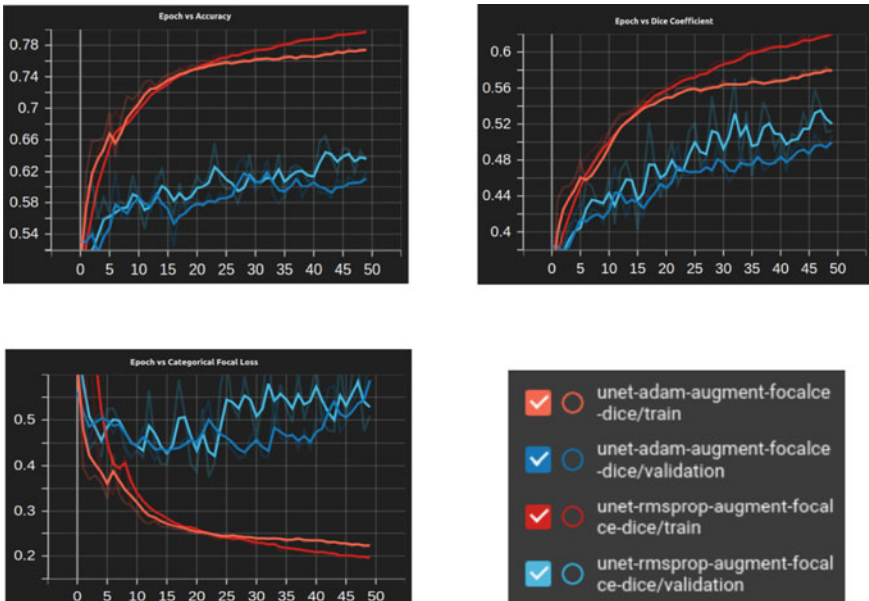


Fig. 6 Training logs

4 Result and Discussion

The environment for the experiment is as follows: Windows 11 operating system and ENVI 4.7 are used to process the remote sensing information, such as selection of training sample, generation of masking. The algorithm for classification developed in fully convolutional network (FCN) was coded with Python + OpenCV.

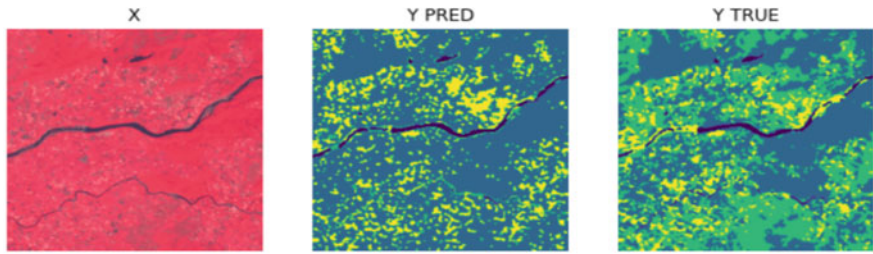
Experiment figured out the best fine-tuning parameters for U-Net with RGB bands of the dataset of LISS-III multispectral remote sensing images. The parameters which give vast performance are used to develop the final model. The model also used data augmentation. Table 2 shows the experiment results and accuracy with different epochs. And, from the outcomes, it was detected that U-Net gives better results in classifying land use land cover classes.

Figure 1 shows an FCC image combined by LISS-III multispectral space-born image in Band-4, Band-3, and Band-2 using tools of ENVI 4.7, in which the class Water Bodies are in blue, class Vegetation is in red, class Uncultivated Land in light red color, and class Residential area is in white. Figure 7a shows the land cover–land use classification results in the South Gujarat region, India, by FCN algorithm, respectively. In Fig. 7a, Water Bodies are black, Vegetation is in light green, Uncultivated Land is in light blue, and Residential area is in yellow. Figure 7b shows the FCC image and Fig. 7c shows the ground truth mask generated via the maximum likelihood classifier. And, Fig. 7d shows the model prediction for the four specified classes.

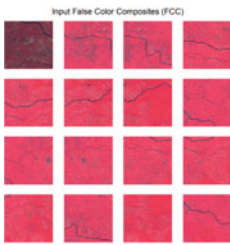
The model classified the image into four classes, i.e., Water Bodies, Vegetation, Uncultivated Land, and Residential areas with an OAA of 81% accuracy. The combination of maximum likelihood for ground truth masking and FCN for classification gives better results. Table 1 shows the training configuration for the model, and Fig. 5 shows the training logs. Figure 7a shows the FCC image, predicted, and truth image, respectively. Figure 7d shows the classified image which is very near to Fig. 7c regarding ground truth masking. This research work was done on a total of 30 FCC images. FCC images and their matching masks are resized to 1024 * 1024 pixels and then divided into patches of size 256 * 256 pixels with a striding of 128. This will create 49 patches of size 256 * 256 for each image. Thus, the size of the dataset would be $30 * 49 = 1470$ images. A total of 1255 images were used for purpose of training the model, while 215 images are reserved for validation and evaluation.

5 Conclusion

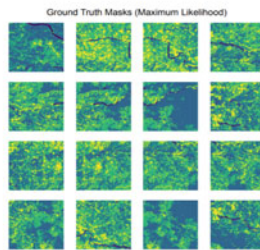
The advantages of using intelligent systems like deep learning for LULC classification are becoming more evident. It will provide a cost-effective and time management solution than the visual interpretation or other machine learning techniques currently obtainable today. In this paper, a land use land cover classification model is presented,



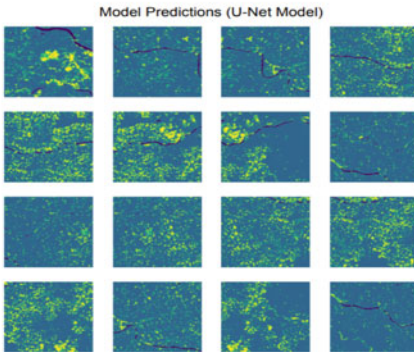
(a.)



(b.)



(c.)



(d.)

Fig. 7 Classification of land cover–land use in FCN algorithm

built on an FCN classifier which is trained and tested on land use land cover LISS-III multispectral space-born image dataset. Experiments show that model is able to detect land use land cover classes. A model can detect different classes with very good accuracy. Outcomes confirmed that the FCN classifier holds massive potential for accurate detection of land use land cover classes. The model identifies four classes with very good accuracy.

Table 1 Hyperparameters and other configurations used for training the model

Hyperparameters and configurations	Values
Train batch size	16
Validation batch size	16
Input image shape	(256, 256, 3)
# of classes	4
Epochs	50
Loss	Categorical focal loss*
Optimizer	Adam
Metrics	Dice coefficient*
Class weights	[1.69941, 0.53043, 1.23977, 1.38949]

Table 2 Experiment result

Model	Optimizer	Epoch trained	Total time	Accuracy
U-Net	Adam	25	1 h 30 min	79
U-Net	Adam	50	2 h 45 min	81

References

- Goslee SC (2011) Analyzing remote sensing data in R: the landsat package. *J Stat Softw*, 43. <http://www.jstatsoft.org>
- Cheng G, Zhou P, Han J (2016) Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images. *IEEE Trans Geosci Remote Sens* 54(12):7405–7415
- Gmez-Chova L, Tuia D, Moser G, Camps-Valls G (2015) Multimodal classification of remote sensing images: a review and future directions. *IEEE Proc* 103(9):1560–1584
- Zhang L, Zhang L, Du B (2016) Deep learning for remote sensing data: a technical tutorial on the state of the art. *IEEE Geosci Remote Sens Mag* 4(2):22–40
- Jaiswal RK, Saxena R, Mukherjee S (1999) Application of remote sensing technology for land use/land cover change analysis. *J Indian Soc Remote Sens* 27(2):123–128
- Hinton GE, Osindero S, Teh YW (2006) A fast learning algorithm for deep belief nets. *Neural Comput* 18(7):1527–1554
- Wen-bo F, Tao S, Ji L, Bao-wei Y, Fu-Xin F (2018) Review of principle and application of deep learning. *Comput Sci* 45(S1):11–15+40
- Richards JA, Richards JA (1999) *Remote sensing digital image analysis*. vol 3, pp 10–38. Springer, Berlin
- Mishra C, Gupta DL (2017) Deep machine learning and neural networks: an overview. *IAES Int J Artif Intell* 6(2):66
- Chen Y, Lin Z, Zhao X, Wang G, Gu Y (2014) Deep learning-based classification of hyper-spectral data. *IEEE J Sel Top Appl Earth Observations Remote Sens* 7:2094–2107. <https://doi.org/10.1109/JSTARS.2014.2329330>
- Yin X, Chen W, Wu X, Yue H (2017) Fine-tuning and visualization of convolutional neural networks. pp 1310–1315. <https://doi.org/10.1109/ICIEA.2017.8283041>

12. Liu P, Zhang H, Eom K (2016) Active deep learning for classification of hyperspectral images. *IEEE J Sel Top Appl Earth Observations Remote Sens*, pp 1–13. <https://doi.org/10.1109/JSTARS.2016.2598859>
13. Piramanayagam S, Schwartzkopf W, Koehler FW, Saber E (2016) Classification of remote sensed images using random forests and deep learning framework. In: *Image and signal processing for remote sensing XXII*, vol 10004, pp 205–212. SPIE
14. Yu X, Wu X, Luo C, Ren P (2017) Deep learning in remote sensing scene classification: a data augmentation enhanced convolutional neural network framework. *GISci Remote Sens* 54:1–18. <https://doi.org/10.1080/15481603.2017.1323377>
15. Yang Z, Mu X-D, Zhao F-A (2018) Scene classification of remote sensing image based on deep network and multi-scale features fusion. *Optik*, 171. <https://doi.org/10.1016/j.jjleo.2018.06.024>
16. Guo D, Xia Y, Luo X (2020) Scene classification of remote sensing images based on saliency dual attention residual network. *IEEE Access* 8:6344–6357. <https://doi.org/10.1109/ACCESS.2019.2963769>
17. Xu X, Chen Y, Zhang J, Chen Y, Anandhan P, Manickam A (2020) A novel approach for scene classification from remote sensing images using deep learning methods. *Eur J Remote Sen*, 54:1–13. <https://doi.org/10.1080/22797254.2020.1790995>
18. Alhichri H, Alsuwayed A, Bazi Y, Ammour N, Alajlan N (2021) Classification of remote sensing images using efficientNet-B3 CNN model with attention. *IEEE Access*. pp 1–1. <https://doi.org/10.1109/ACCESS.2021.3051085>
19. Pires de Lima R, Marfurt K (2019) Convolutional neural network for remote-sensing scene classification: transfer learning analysis. *Remote Sens*, 12:86. <https://doi.org/10.3390/rs12010086>
20. Mohanty SP, Czakon J, Kaczmarek KA, Pyskir A, Tarasiewicz P, Kunwar S, Rohrbach J, Luo D, Prasad M, Fleeer S, Göpfert JP (2020). Deep learning for understanding satellite imagery: an experimental survey. *Front Artif Intell*, 3:534696
21. Lateef F, Ruichek Y (2019) Survey on semantic segmentation using deep learning techniques. *Neurocomputing* 338:321–348. <https://doi.org/10.1016/j.neucom.2019.02.003>
22. Hao S, Zhou Y, Guo Y (2020) A brief survey on semantic segmentation with deep learning. *Neurocomputing* 406:302–321. <https://doi.org/10.1016/j.neucom.2019.11.118>
23. Thoma M (2016) A survey of semantic segmentation. *CoRR*. abs/1602.06541
24. Atif N, Bhuyan M, Ahamed S (2019) A review on semantic segmentation from a modern perspective. In: *2019 international conference on electrical, electronics and computer engineering (UPCON)*, 8–10 Nov 2019, Aligarh, India. IEEE, Piscataway, NJ, pp 1–6
25. Minaee S, Boykov Y, Porikli F, Plaza A, Kehtarnavaz N, Terzopoulos D (2020) Image segmentation using deep learning: a survey. *arXiv preprint arXiv: 2001.05566*
26. Fan J, Cao X, Yap P-T (2018) Birnet: brain image registration using dual-supervised fully convolutional networks. *Med Image Anal* 54:02
27. Shelhamer E, Long J, Darrell T (2016) Fully convolutional networks for semantic segmentation. *IEEE Trans Pattern Anal Mach Intell*, 39:1–1
28. Bi L, Kim J, Ahn E, Kumar A, Feng DDF, Fulham M (2019) Step-wise integration of deep class-specific learning for dermoscopic image segmentation. *Pattern Recogn* 85:78–89
29. Huang Y, Zhou F, Gilles J (2019) Empirical curvelet based fully convolutional network for supervised texture image segmentation. *Neurocomputing* 349:04
30. Li C, Wang X, Liu W, Latecki LJ, Wang B, Huang J (2019) Weakly supervised mitosis detection in breast histopathology images using concentric loss. *Med Image Anal* 53, p 02
31. Such F, Pillai S, Brockler F, Singh V, Hutkowski P, Ptucha R (2018) Intelligent character recognition using fully convolutional neural networks. *Pattern Recogn* 88:12
32. Kumar S, Kumar N, Rishabh IK, Keshari V (2021) Automated brain tumour detection using deep learning via convolution neural networks (CNN). *Int J Cur Res Rev* 13(02):148
33. Guo Y, Liao J, Shen G (2021) Mapping large-scale mangroves along the maritime silk road from 1990 to 2015 using a novel deep learning model and landsat data. *Remote Sens* 13:245. <https://doi.org/10.3390/rs13020245>

34. Guo M, Yu Z, Xu Y, Huang Y, Li C (2021) ME-Net: a deep convolutional neural network for extracting mangrove using sentinel-2A data. *Remote Sens* 13:1292. <https://doi.org/10.3390/rs13071292>
35. Wan L, Zhang H, Lin G, Lin H (2019) A small-patched convolutional neural network for mangrove mapping at species level using high-resolution remote-sensing image. *Ann GIS* 25:45–55. <https://doi.org/10.1080/19475683.2018.1564791>
36. Hoese T, Bachofer F, Kuenzer C (2020) Object detection and image segmentation with deep learning on earth observation data: a review—part II: applications. *Remote Sens* 12:3053. <https://doi.org/10.3390/rs12183053>
37. Hosseiny B, Mahdianpari M, Brisco B, Mohammadimanesh F, Salehi B (2021) WetNet: a spatial-temporal ensemble deep learning model for wetland classification using sentinel-1 and sentinel-2. *IEEE Trans Geosci Remote Sens*, pp 1–14
38. Castelluccio M, Poggi G, Sansone C, Verdoliva L (2015) Land use classification in remote sensing images by convolutional neural networks. Aug 2015. [arXiv:1508.00092](https://arxiv.org/abs/1508.00092)
39. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. IEEE, San Juan, Puerto Rico, pp 770–778
40. Li J, Lin D, Wang Y, Xu G, Zhang Y, Ding C, Zhou Y (2020) Deep discriminative representation learning with attention map for scene classification. *Remote Sens* 12:1366. <https://doi.org/10.3390/rs12091366>
41. Sonune N (2020) Land cover classification with EuroSAT dataset. <https://www.kaggle.com/nilesh789/land-cover-classification-with-eurosat-dataset>
42. *Remote Sensing Notes* edited by Japan Association of Remote Sensing © JARS 1999
43. Xu J, Song L, Zhong D, Zhao Z, Zhao K (2013) Remote sensing image classification based on a modified self-organizing neural network with a priori knowledge. *Sens Transducers* 153:29–36

Particle Swarm Optimization for Web Document Retrieval Based on Term-Document Matrix



Vishal Kumar, Manoj Kumar, Simran Sirohi, Ravi Kumar,
and Santosh Kumar

Abstract Retrieving relevant information from the enormous volume of data is the top most priority of information retrieval systems. As technology arises, various search engines help us to track the relevant data and information on the user's needs, but the amount of information available is huge and complex. Sometimes, user finds it difficult to find the data according to their need. When people provide the same queries, the same set of relevant documents will be returned by classical search engines. It gets hard for the users to retrieve the relevant documents. To solve this problem, a developed document search technique has been proposed, which is based on Particle Swarm Optimization. It helps to optimize web document information retrieval. PSO algorithm is easy to use and can have better results as it is a faster method compared to other methods. PSO algorithm is being applied on Term-Document Matrix which consists of weights calculated by TF-IDF weighting formula. Term-Document Matrix is the search space on which the PSO is applied by considering the query as a particle and the position of the particles is initialized randomly. This method is tested experimentally on a big TREC 2019 dataset. By applying PSO, the results are compared with the results obtained through traditional approaches.

V. Kumar (✉) · M. Kumar · S. Sirohi · R. Kumar
ABES Engineering College, Ghaziabad, UP, India
e-mail: vishal.18bcs1128@abes.ac.in

M. Kumar
e-mail: manoj.18bcs1105@abes.ac.in

S. Sirohi
e-mail: simran.18bcs1010@abes.ac.in

R. Kumar
e-mail: ravi.kumar@abes.ac.in

S. Kumar
Galgotias University, Greater Noida, India
e-mail: santosh.kumar@abes.ac.in

Keywords Term-Document Matrix · Web information retrieval · Fitness function · Particle Swarm Optimization · Cosine similarity · Vector space model

1 Introduction

The acquisition, organization, storage, search, and selection of data are all aspects of information retrieval (IR), a branch of computer science. A user can obtain information from an information retrieval system (IRS) by providing his information requirements in the form of a query, which is a list of keywords. But, with the development of personal computers, electronic media, advancement of the internet, it is becoming tough to extract appropriate and latest information. IR has several applications in a variety of domains [1], and numerous research has been conducted on the subject of information retrieval, particularly document retrieval [2, 3]. Different information retrieval algorithms are being used by search engines to provide relevant information according to queries given by the user, but sometimes these algorithms are not able to provide relevant documents quickly. Hence, optimization of the information retrieval process for searching is required. Many algorithms like genetic algorithm [4, 5], bees swarm optimization [6], ant colony optimization [7], relevance feedback [8], etc., are being used for the optimization purpose of the information retrieval process. Other approaches like SVM-based relevance feedback [9] and dimensionality reduction are also used to improve the efficiency of IR process. Dimensionality reduction entails projecting a high-dimensional document into a lower-dimensional subspace in order to employ typical retrieval techniques in the reduced lower-dimensional document space. Two essential techniques for dimensionality reduction and feature extraction in document retrieval systems are Latent Semantic Indexing (LSI) [10] and Linear Discriminant Analysis (LDA) [11]. For enhancing information retrieval, better algorithms can be applied either on document set or query set, talking about documents; then, clusters of the relevant document can be used by web document clustering using FP growth [12]; similarly, queries can also be optimized using some algorithms like memetic algorithm [13] which is quite similar to GA except the local improvement heuristic used which improves the result. Particle Swarm Optimization (PSO) [14] is a metaheuristic algorithm widely used in the optimization of NP-Hard problems such as the Traveling Salesman Problem, No-Wait Flow Shop Scheduling Problems, and so on. Positive points of PSO are its robustness for parameter control and its efficiency of computation. Hence, PSO algorithm is used to build a system that will compute the similarity of documents and the query entered by the user to enhance the accuracy and information extraction time. The suggested approach will be utilized to compare the performance of the PSO algorithm to that of the standard information retrieval method. This method can be used for efficient information retrieval on small as well as the large dataset. Our *main* aim is to examine the efficiency of PSO algorithm by comparing it with traditional algorithm using the similarity score of documents with queries. The dataset is based on the TREC 2019 Deep Learning

Track document ranking dataset. This dataset has nearly 3.5 lakhs queries and a collection of 32 lakh pages.

2 Related Previous Work

Wang et al. [15] use PSO to optimize user queries and show that PSO can enhance the correctness of information retrieval as compared to other algorithms like genetic and relevant feedback. To accelerate the capabilities of PSO, a neighborhood search strategy has been adopted to get efficient query vector near to the original query vector in less time. TREC collection containing approximately 742,600 documents was used to test performance, and the 1000 top-ranked relevant documents for 50 queries were retrieved. In four of the five iterations' PSO, the number of documents retrieved by PSO is greater than genetic algorithm. Drias [16] developed two different versions of PSO algorithms PSO1-IR and PSO2-IR for document retrieval. PSO1-IR is tested on the whole collection of documents, and PSO2-IR is tested on the inverted file (reduced input size). Author has shown that the designed algorithms are much superior and robust as compared to traditional approaches. The experimental results have shown a great decrease in response time as compared to classical approaches. Experiments were conducted on two collections: (1) CACM is a tiny collection of article abstracts from ACM journals published between 1958 and 1979. There are 3204 documents and 6468 terms in all. The typical file size is 2 K bytes. (2) RCV1 is a large collection of 804,414 documents and 47,236 terms containing archives published by Reuters. The average document size is 2 K bytes. To obtain more comparable documents, Ramya et al. [17] used PSO combined with a unique similarity measure known as Similarity Measure for Documents' Retrieval (SMDR). In terms of accuracy, precision, sensitivity, F-measure, and specificity, the findings were compared to the existing system. By assigning a rank to a document, SMDR indicates how similar it is to the issued query. When compared to an existing system, PSO with SMDR as its fitness function significantly reduces the proposed system's reaction time. Dataset used is CACM. To enhance the web information retrieval process, Ramya and Shreedhara [18] suggest a PSO approach hybridized with stimulated annealing. PSO hybridization will alleviate PSO's shortcomings, such as premature convergence, which will slow down the search process around the global optimum. The experiment's results were analyzed using a variety of metrics, including accuracy, Mean Reciprocal Ranking (MRR), Mean Average Precision (MAP), Discounted Cumulative Gain (DCG), IDCG, F-measure, and specificity. Computed similarity values were first given to stimulated algorithm to get the optimized similarity list and then that similarity list is fed to PSO. The achieved precision and recall rates suggest that the hybridized PSO has a significant influence on lowering query response time when compared to the conventional one, hence enhancing system efficiency. The RCV1 dataset is utilized since it contains a large number of XML documents issued by Reuters. Alloui et al. [19] reformulate the user query and increase the number of relevant results using a relevance feedback technique.

3 Proposed Model

The suggested technique assigns weights to words in all documents using a vector space model and the TF-IDF weighting algorithm. A Term-Document Matrix (TDM) is being created. PSO algorithm is being applied on TDM. The query is also represented in vector form using the vector space model, and it is treated as a particle for PSO, with the position and velocity initialized at random. The personal best solution for each particle is computed, followed by the global best solution, and the location and velocity of the particles are updated until a certain number of repetitions are accomplished. The architecture for web documents retrieval is given in Fig. 1.

3.1 Preprocessing

Text preprocessing is a method which is used for cleaning and preparing text data so that it can be used to achieve specific tasks. All data in the dataset are processed for indexing purposes which helps in the process of information retrieval. The steps for information retrieval are given in Algorithm 1.

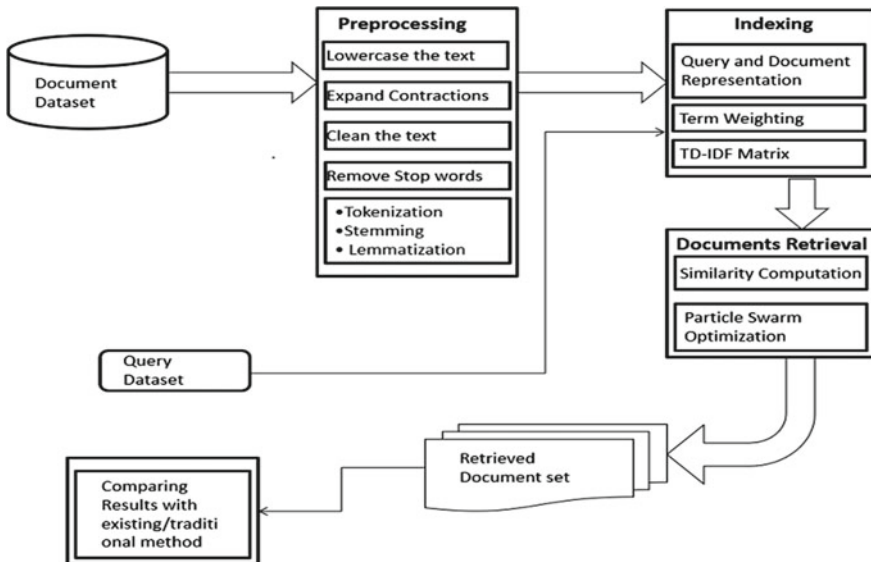


Fig. 1 Architecture for web documents' retrieval

Algorithm 1. Steps of Preprocessing

<p><i>Documents:</i></p> <ol style="list-style-type: none"> 1. Convert the text to lowercase 2. Converting Contractions to complete words 3. Cleaning the data (Remove numbers, punctuation marks, URLs, new line characters and extra spaces) 4. Remove Stopwords 	<ol style="list-style-type: none"> 5. Tokenization 6. Stemming 7. Lemmatization <p><i>Queries:</i></p> <ol style="list-style-type: none"> 1. Lowercase the query 2. Converting Contractions to complete words 3. Cleaning the query
--	---

3.2 Indexing

In information retrieval systems, indexing is a critical step. It is the primary purpose of the IR process. It is essentially a procedure in which both the document and the query are represented as a collection of words (index terms). Indexing is performed by the indexer module. Traditional style of indexing or manual indexing is a time taken process and requires huge hours to index a repository. Nowadays, automatic indexing which is much faster and less prone to errors is a common practice on large datasets.

3.3 Indexing with Vector Space Model

The vector space model [20] is employed in the majority of information retrieval techniques. Both documents and queries are represented as vectors of characteristics in this approach. It is not like the Boolean model in which there is no concept of ranking. Vector space model provides a structure which includes Algorithm 2.

Algorithm 2. Vector Space Model

<p>Steps used in Vector Space Model</p> <ol style="list-style-type: none"> i. Weight assigned to every term ii. Rank assigned to every retrieved document iii. Possibility of relevance feedback <p>A document D_j and query Q in vector space is represented like this–</p> $D_j = (t_{1j}, t_{2j}, \dots, t_{mj})$ $Q = (t_{1q}, t_{2q}, \dots, t_{hq})$	<p>t represents term weight in documents so there are m terms in a document and h terms are in query</p> <p>Where t_{ij} and t_{iq} represent the weight of the jth term in the document and query respectively</p>
---	---

In a vector space model, term weight is stated as a function of changes in the term frequency (TF) or term frequency–inverse document frequency (TF-IDF). TD-IDF is used to assess the significance of a certain term in a collection of documents.

The TD-IDF is computed by multiplying the normalized frequency of a term i (TF_{*i*}) in document D_j by the inverse document frequency of the term i (IDF_{*i*}). The

idea is that words that capture the essence of a document appear often in the 29 documents (i.e., their TF is high), but if such a term is a good term that distinguishes the document from others, it should exist in just a few documents in the entire population (that is, its IDF ought to be high also). IDF is computed by the expression given in Algorithm 3.

Algorithm 3. Computation of IDF	
<p>1. Compute frequency of terms</p> $TF_{ij} = f_{ij} / \sum_{i=1}^{i= V } f_{ij}$ <p>2. Compute inverse document frequency of the term i</p> $IDF_i = \log(N/n_i)$	<p>Where, TF_{ij} represents the normalized term frequency of term i in document D_j f_{ij} is the number of times term i appears in document D_j The inverse document frequency weight for term i is denoted by IDF_i N is number of documents in the collection n_i is the number of documents that include the term i</p>

3.4 Similarity Computation in Vector Space Model

The similarity assessment function is used to compute the similarity between the document and the query. There is no intrinsic similarity assessment function in the model; instead, alternative similarity functions can be employed. Nonetheless, cosine similarity is the widely used similarity function for assessment. When the angle between the query and document vectors decreases, the cosine of the approaching angle equals one, indicating that the similarity between the query and document vectors expands. Algorithm 4 is used for similarity computation.

Algorithm 4. Similarity Computation
1. Find document vector, d_j
2. Find query vector, q
3. Calculate Cosine $(d_j, q) = (d_j \times q) / d_j \cdot q $

3.5 Particle Swarm Optimization (PSO)

It is a swarm movement and intelligence-based stochastic optimization algorithm [14]. Kennedy and Eberhart suggested it in 1995. The main idea behind the PSO is that in a flock of birds seeking for food at random, they do not know the location of food is, but they do know how far away it is. The ideal strategy is to approach the birds who are closest to the food particle. (i) A flock of birds is represented by a set

of particles with varying positions and velocities that seek the best solution in the search space. (ii) Each particle in the group attempts to locate its local best-known location, which is known as the personal best solution, or PbS. (iii) The other best value discovered by the PSO is known as GbS or global best solution. That is the best possible value discovered thus far by any particle in the region of that particle. (iv) In each iteration, particles try to improve the solution by updating the position and velocity.

3.6 Proposed PSO-Based Algorithm

PSO-based approach for document retrieval with improved similarity score is given in Algorithm 5.

Algorithm 5. PSO based Algorithm	
<p>1. Make a Term-Document Matrix with TD-IDF weighting for all the documents</p> <p>2. Initialize all the particles by assigning the following-</p> <ul style="list-style-type: none"> • current position as indexes of documents • velocity • personal best position • personal best fitness value <p>$pbest_cost = f(pos)$</p> <ul style="list-style-type: none"> • global best position <p>$gbest_pos = indexOf(max(pbest_cost))$ in $pbest_pos$</p> <ul style="list-style-type: none"> • and global best fitness value <p>$gbest_cost = indexOf(max(pbest_cost))$ in $pbest_cost$</p> <p>3. compute the new velocity and position for each particle k</p> <p>$r1 = random\ value\ from\ [0, 1]$</p> <p>$r2 = random\ value\ from\ [0, 1]$</p> <p>$velo[k] = w*velo[k] +$</p> <p>$c1*r1*(pbest_pos[k]-pos[k]) +$</p> <p>$c2*r2*(gbest_pos-pos[k])$</p> <p>$pos[k] = pos[k] + velo[k]$</p> <p>4. compute the fitness value for new/updated position of each particle k</p> <p>5. update the best solution according to fitness value for each particle k</p> <p>if $f(pos[k]) > pbest_cost[k]$ then</p> <p>$pbest_pos[k] = pos[k]$ and $pbest_cost[k] = f(pos[i])$</p>	<p>if $pbest_cost[k] > gbest_cost$ then</p> <p>$gbest_pos = pbest_pos[k]$ and $gbest_cost = pbest_cost[k]$</p> <p>6. Repeat the steps 3, 4 and 5 until maximum iterations are reached</p> <p>Where,</p> <p>k is the loop counter</p> <p>pos is the initial positions of particles</p> <p>$velo$ is velocity</p> <p>$pbest_pos$ is the personal best positions of particles</p> <p>$gbest_pos$ is the global best position of particles</p> <p>$pbest_cost$ is the personal best fitness values of particles</p> <p>$gbest_cost$ is the global best fitness value of particles</p> <p>f is the fitness function</p> <p>max finds the maximum value</p> <p>$indexOf$ is to find the index</p> <p>$r1$ and $r2$ are random numbers in the range $[0,1]$</p> <p>$c1$ and $c2$ are acceleration coefficients</p> <p>w is the inertia weight</p>

4 Results and Discussion

The proposed algorithm is implemented in Python. TREC collection of documents is used to find the results. The suggested approach enhances the document’s resemblance to the query. Table 1 displays the similarity score for five different queries for both the proposed and existing algorithms. Table 1 shows that there is an increase in similarity values, indicating that the suggested approach outperforms the existing technique. Figures 2 and 3 compare the performance of proposed and existing algorithms.

Table 1 Similarity score

Algorithm	Query	Similarity score
Existing	1	0.395
	2	0.287
	3	0.18
	4	0.483
	5	0.461
Proposed	1	0.394
	2	0.313
	3	0.19
	4	0.534
	5	0.469

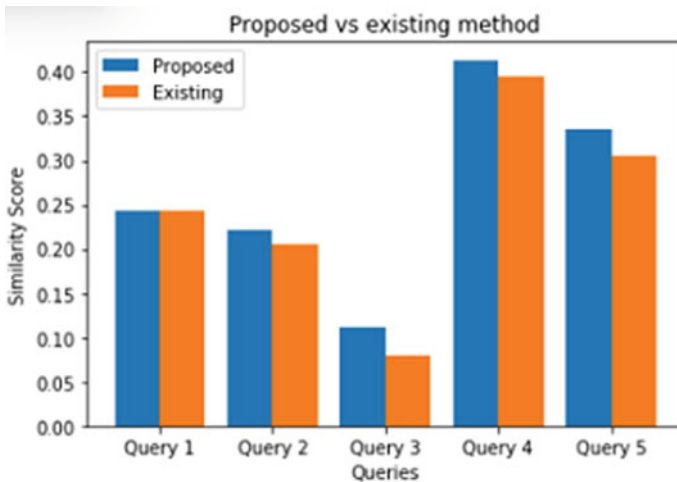


Fig. 2 Performance comparison

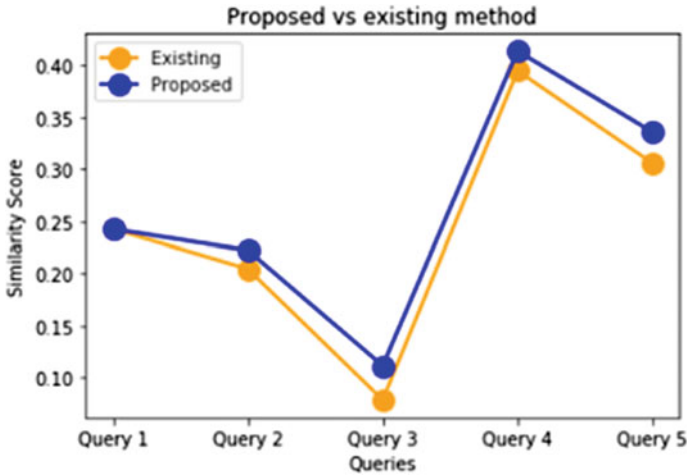


Fig. 3 Query versus similarity score

5 Conclusion and Future Work

The proposed work presents a unique PSO optimization method for document retrieval. The method is capable of increasing similarity score. When compared to existing methods, our technique has been shown to outperform them in terms of stability and robustness. The results show that our method is well adapted to dealing with large collections. It is designed in such a way that it may be used as an application for the retrieval of documents. It is preferable to use PSO in the future to get highly similar documents. Future study will include investigating ways to enhance PSO to address IR optimization difficulties such as tailored source selection, distributed information retrieval, and so on.

References

1. van Rijsbergen CJ (1979) Information retrieval. Butterworths, London
2. Singhal A (2001) Modern information retrieval: a brief overview. IEEE Comput Soc Tech Committee Data Eng 24(4):35–43
3. Yates RB, Neto BR (1999) Modern information retrieval. Addison Wesley
4. Jiao LC, Wang L (2000) A novel genetic algorithm based on immunity. IEEE Trans Syst Man Cybern-Part A 30(5):552–561
5. Hornig J-T, Yeh C-C (2000) Applying genetic algorithms to query optimization in document retrieval. Inf Process Manage 36(5):737–759. [https://doi.org/10.1016/s0306-4573\(00\)00008-x](https://doi.org/10.1016/s0306-4573(00)00008-x)
6. Drias H, Mosteghanemi H (2010) Bees swarm optimization based approach for web information retrieval. In: 2010 IEEE/WIC/ACM international conference on web intelligence and intelligent agent technology
7. Dorigo M, Gambardella LM (1997) Ant colony system: a cooperative learning approach to the traveling salesman problem. IEEE Trans Evol Comput 1(1):53–66

8. Salton G (ed) (1971) Relevance feedback in information retrieval. PrenticeHall, Englewood Cliffs, N.J., pp 313–323
9. Drucker H, Shahrar B, Gibbon DC (2001) Relevance feedback using support vector machines. In: Proceedings of the eighteenth international conference on machine learning, pp 122–129
10. Berry MW, Drmac Z, Jessup ER (1999) Matrices, vector spaces, and information retrieval. *SIAM Rev* 41(2):335–362
11. Duda RO, Hart PE, Stork DG (2000) *Pattern Classification*, 2nd edn. Wiley-Interscience, Hoboken
12. Pamba RV, Sherly E (2016) A novice approach for web document clustering using FP growth based fuzzy particle swarm optimization. In: 2016 3rd International conference on soft computing & machine intelligence (ISCMI). <https://doi.org/10.1109/iscmi.2016.36>
13. Wang Z, Sun X, Zhang D (2008) Web document query optimization based on Memetic algorithm. In: 2008 IEEE Pacific-Asia workshop on computational intelligence and industrial application. <https://doi.org/10.1109/paciia.2008.100>
14. Eberhart R, Kennedy J (1995) A new optimizer using particle swarm theory. In: MHS'95. Proceedings of the sixth international symposium on micro machine and human science. <https://doi.org/10.1109/mhs.1995.494215>
15. Wang Z, Li X, Zhang D, Wu F (2006) A PSO-based web document query optimization algorithm. In: The semantic web—ASWC 2006, pp 609–615. https://doi.org/10.1007/11836025_59
16. Drias H (2011) Web information retrieval using particle swarm optimization based approaches. In: 2011 IEEE/WIC/ACM international conferences on web intelligence and intelligent agent technology. <https://doi.org/10.1109/wi-iat.2011.225>
17. Ramya C, Shreedhara KS (2019) A PSO strategy of finding relevant web documents using a new similarity measure. *Int J Darshan Inst Eng Res Emerg Technol* 7(2):29
18. Ramya C, Shreedhara KS (2019) A bio-inspired modified PSO strategy for effective web information retrieval using RCV1 datasets. 8(10):779–785. <https://doi.org/10.35940/ijitee.j8903.0881019>
19. Alloui T, Boussebough I, Chaoui A (2015) A particle swarm optimization algorithm for web information retrieval. *IJIT* 11(3):15–29
20. Lee DL, Chuang H, Seamons K (1997) Document ranking and the vector-space model. *IEEE Softw* 14(2):67–75

Light-Weight Deep Learning Models for Visual Malware Classification



E. Akshay Kumar and Jothi Rangasamy

Abstract Malware attacks are on the rise every day in the Internet-based digital world. Regular Internet users are at risk due to the evolution of new infections. In recent years, the use of machine learning algorithms to identify malware has gained popularity because numerous studies have demonstrated its efficacy. This work provides two deep learning models to categorize the malware turned into images. Our method uses fewer resources and takes less time to accomplish the same performance as state-of-the-art results. The primary advantage of malware images is that no additional feature engineering is required. Our models for categorizing image-based malware are less complex and can be used in computational systems with limited computational capabilities, such as Android devices.

Keywords Malware classification · Cyber security · Deep learning

1 Introduction

Malicious software designed to take advantage of the system and its resources is called malware [1]. The sensitive information present in information systems must be protected for user's privacy and security. Exposing sensitive information may lead to even some crucial impacts on the security of a country [2]. Ever-rising developments in creation of new malware are always a concern to Internet users. This makes malware analysis a significant and challenging problem. Recent development in data science paves the way toward automated malware analysis.

Due to the COVID pandemic, we have witnessed many businesses migrating to online and some sectors have transformed their functioning through the Internet. This attracted cyber criminals to become more active, thereby increasing the number of cyber-attacks and cybercrimes. Since the frequency of new variants of malware

E. Akshay Kumar (✉) · J. Rangasamy

Department of Mathematical and Computational Sciences, NIT-K, Surathkal, Karnataka, India
e-mail: akshaynair.e@gmail.com

J. Rangasamy

e-mail: jothiram@nitk.edu.in

being launched has increased, keeping pace with the attacks itself has become a major challenge. Malware analysis mainly deals with dissecting a binary file to understand its working and frame remission methods to detect the same and prevent it from entering the system. Also, it deals with finding the actions caused by the malware. Malware analysis is mainly used for detecting any unwanted malicious software. Malware detection deals with checking the presence of the malware, and malware classification deals with classifying the malware into respective groups.

2 Related Works

Usage of static features of malware is common in malware classification systems. Roseline and Geetha [3] proposed an oblique random forest-based system that uses static features collected from malware for malware detection. A hierarchical convolutional neural network (CNN) for malware classification [4] is proposed for considering the hierarchical structure of the program also in malware classification. The hierarchical structure is not considered when we do classification based on frequent n-grams features. A study on the effect of feature selection on malware analysis using machine learning [5] mainly discusses the effect of feature selection on malware analysis. Using the cuckoo sandbox, they collected static features from virustotal and dynamic features to trace the trend. From their study, it is clear that feature selection has a considerable effect on the classifier's performance. They have used information gain to select the most prominent features, and it gave maximum performance when an MLP model was used for finding the pattern.

A four-layer model for malware classification has been proposed using convolutional gated neural network [6]. The first layer contains a convolutional neural network, divided into three sub-layers: a conv layer, an activation layer, and a pooling layer. The activation layer uses tanh, relu, or sigmoid activation functions. The GRU layer contains user-defined GRU units. Each GRU in the layer took the values of all CNNs from the previous level. Each GRU gives a single output value for an input, resulting in the number of output values being the same as none of the CNNs in the previous layer. The third layer is a network of DNNs that receives output from GRU and then passes through the sigmoid layer.

An analysis of malware prediction based on infection rate using machine learning technique [7] utilized Microsoft Malware Prediction Dataset. They constructed several models, including LIGHTGBM, RF, and neural networks. In that, LIGHTGBM fitted into a sparse matrix gave the best results, obtained the highest AUC score, and took less time to run with excellent efficiency. Jhu-Sin Luo and Dan Chia-Tien Lo [8] proposed a malware classification system using image processing techniques. Local binary pattern (LBP) features are extracted from malware images reorganized into 3×3 grids. Finally, machine learning is used to classify the same with LBP features.

Vinayakumar et al. [9] mainly evaluated the classical machine and deep learning approaches for malware classification and detection. Their major contribution is the development of a novel image processing-based technique with the ideal amount of

parameters for achieving a successful zero-day malware mitigation model. Cui et al. [10] proposed a novel convolutional neural network-based architecture for malware classification. They converted the malware code into images that can be fed into a CNN network to identify whether the image is an image of malware or not by taking care of data imbalance.

3 Proposed Work

3.1 Dataset

Maling dataset [11] was created by converting malware binary into grayscale images. There are 9339 malware samples in the Maling collection from 25 different malware families. Figure 1 shows some sample images from the dataset. By converting malware binaries into a matrix, the dataset was created. The matrix is made up of unsigned 8-bit integers. The created matrix can be seen as a grayscale image with values ranging from 0 to 255, where 0 represents white and 255 represents black. We randomly divided the dataset into 70 percent training and 30 percent testing sets.

3.2 Baseline Models

We compare the proposed deep learning models with the baseline models composed of Resnet-50 [12], DenseNet-121 [13], MobileNet [14], Inception-V3 [15], and Xception [16]. The comparison's primary goal is to demonstrate that the recommended architectures perform admirably regarding the model's complexity and other performance criteria.

3.3 Proposed Deep Learning Models

We propose two deep learning models for categorizing malware images. The suggested solution's major goal is to apply malware detection and classification techniques for resource-constrained computing devices.

3.3.1 CNN Model

The first model we introduced is a simple CNN model. The model mainly aims to reduce the malware detector's complexity without lowering its performance. We transform the malware binary into images thereby converting the malware classifi-

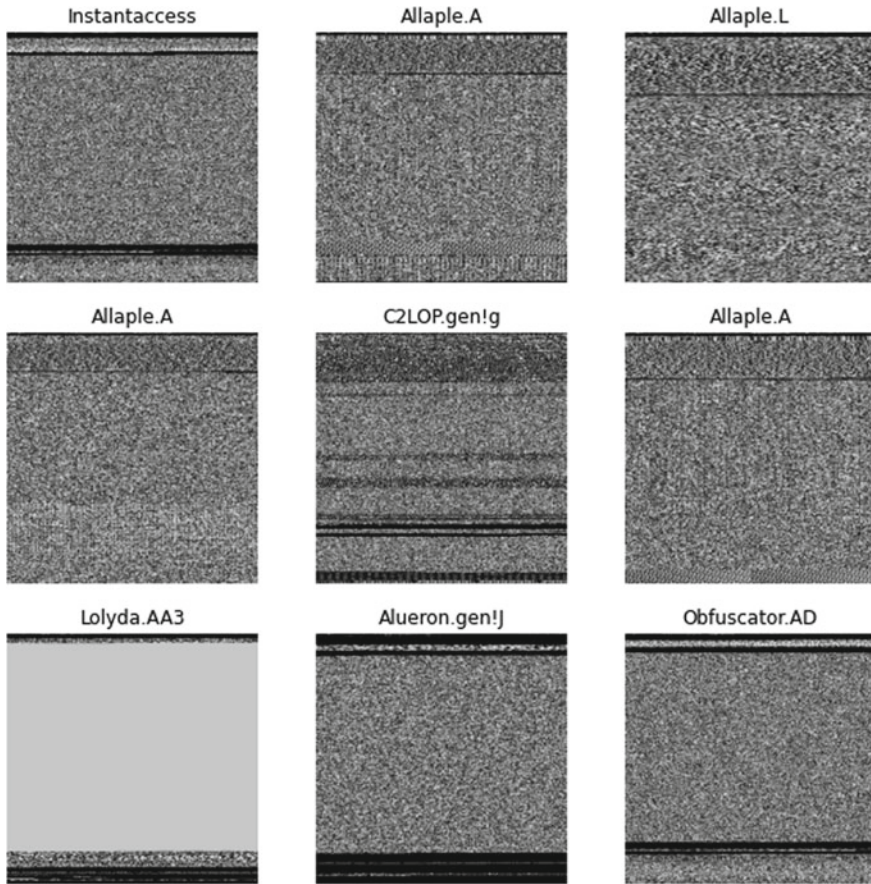


Fig. 1 Some sample malware images from Maling dataset

cation into an image classification problem. The key benefit of this transformation is that the dataset does not need to undergo any feature engineering work. A brief explanation of CNN layers [17] is provided below.

Convolutional Layer: Between the input image and a particular filter of a specific size $M \times M$, the mathematical operation of convolution is done. The dot product between the various locations of the input image and the filter after the filter size is calculated by sliding the filter over the input images ($M \times M$) to get the feature map.

Pooling Layer: The pooling layer's primary goal is to lower the size of the feature map produced by the convolution process to reduce computational expenses. There are various sorts of pooling procedures depending on the technique utilized. The average of the elements in a predefined sized image section is determined by average pooling.

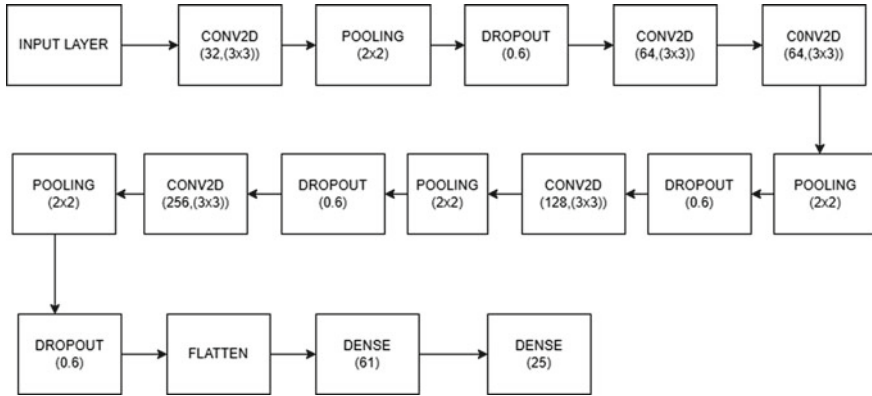


Fig. 2 Proposed CNN model

Fully Connected Layer: Its primary function is to connect neurons between two layers. It is composed of the weights and biases as well as the neurons. The final few layers of a CNN model are often built with fully connected layers and frequently positioned before the output layer.

Dropout: Dropout is an effective method used to tackle the problem of overfitting. Here, a few neurons from the model are dropped during the training process, which results in reduced network size.

Activation Function: It is to add nonlinearity to the deep learning network. We mainly used the ReLU at the convolution layers and the Softmax at the fully connected layers.

ReLU: The largest value between zero and the input is the ReLU output. When the input value is negative, the output equals zero, and when the input value is positive, the output equals the input value.

$$f(x) = \max(0, x) \quad (1)$$

Softmax: The Softmax regression is a kind of logistic regression that converts an input value into a vector of values that adheres to a probability distribution and whose sum equals 1.

$$\sigma(z_i) = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}} \quad \text{for } i = 1, 2, \dots, K \quad (2)$$

All of the z_i values in the equation are input vector elements and can take any real value. The normalization factor in the denominator of the equation ensures that all of the function's output values sum to 1, resulting in a proper probability distribution (Fig. 2).

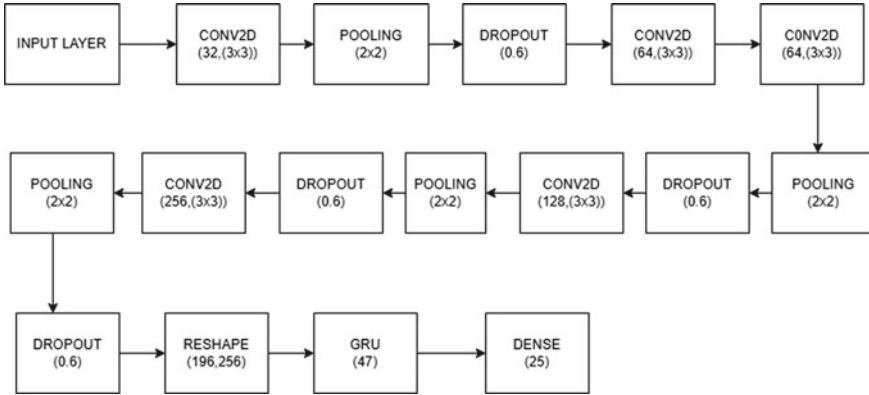


Fig. 3 Proposed CNN-GRU model

3.3.2 CNN-GRU Model

The architecture of the gated recurrent unit (GRU) is simpler than that of the long short-term memory (LSTM) cell [18]. In a GRU [19] model, there are two major gates: a reset gate and an update gate, respectively. The reset gate and update gate at the t -th time step, respectively, are represented by r_t and z_t in the following equations, which are computed as follows:

$$r_t = \sigma(W_r x_t + U_r h_{t-1}) \tag{3}$$

$$z_t = \sigma(W_z x_t + U_z h_{t-1}) \tag{4}$$

where W_r , U_r , W_z , and U_z are the weight matrices of the reset gate and update gate, respectively. The hidden state h_t can be computed as follows:

$$h_t = z_t h_{t-1} + (1 - z_t) \tilde{h}_t \tag{5}$$

$$\tilde{h}_t = \tanh(W_r x_t + U_r (r_t \odot h_{t-1})) \tag{6}$$

In the above equation, \odot represents the Hadamard product [20].

The second model is a CNN-GRU hybrid model. The GRU unit is utilized for classification, whereas the outer CNN layers are mostly employed to extract features from malware images. Because of the GRU unit structure, the model takes more training and testing time. Nevertheless, the space consumed and the total number of parameters are much less than the proposed CNN model and baseline models (Fig. 3).

4 Experimental Results

We used the Python framework and the Keras v2.7.0 [21] deep learning package to implement the suggested architectures. Several binary images of size 224×224 were subjected to experimental analysis. We employed a 32 GB NVIDIA Tesla V100 GPU for the experiments shown in the outcome.

True Positive Rate (TPR): TPR is a measure of the measure of positive cases in the data that are correctly classified as such (i.e., positive).

$$\text{TPR} = \frac{\text{TP}}{(\text{TP} + \text{FN})} \quad (7)$$

False Positive Rate (FPR): FPR is the ratio of negative cases wrongly classified as positive classes in the data.

$$\text{FPR} = \frac{\text{FP}}{(\text{FP} + \text{TN})} \quad (8)$$

Accuracy: It is defined as the fraction of predictions our model done correctly. In simple words, it is a ratio of correct predictions to that of total predictions.

$$\text{Accuracy} = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{FP} + \text{TN} + \text{FN})} \quad (9)$$

Precision: It is the fraction of correctly classified positive values to the total predicted positive values.

$$\text{Precision} = \frac{\text{TP}}{(\text{TP} + \text{FP})} \quad (10)$$

Recall: It reflects the measure of our model correctly identifying True Positive Values.

$$\text{Recall} = \frac{\text{TP}}{(\text{TP} + \text{FN})} \quad (11)$$

F1-score: It is the harmonic mean of Precision and Recall. The use here is to measure the performance value shown by the classifiers.

$$\text{F1-Score} = 2 * \text{Precision} * \frac{\text{Recall}}{(\text{Precision} + \text{Recall})} \quad (12)$$

Table 1 Performance comparison with baseline models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Resnet-50	98.25	98.33	98.25	98.21
Xception	98.21	98.28	98.21	98.24
MobileNet	97.53	97.70	97.54	97.51
DenseNet-121	98.57	98.67	98.57	98.56
Inception-v3	89.18	87.80	89.18	87.67
Proposed model 1 (CNN)	98.60	98.66	98.61	98.61
Proposed model 2 (CNN-GRU)	98.60	98.59	98.60	98.58

Table 2 Complexity comparison with baseline models

Model	Training time (s)	Testing time (s)	Total number of parameters	Size of model (MB)
Resnet-50	1643.68	5.62	26,096,537	314
Xception	3843.68	6.94	23,370,305	281
MobileNet	1678	3.69	4,483,289	51.4
Densenet-121	2011	6.74	8,291,929	100
Inception-v3	1493.90	6.03	23,082,809	278
Proposed model 1 (CNN)	479.59	3.14	3,487,691	41.9
Proposed model 2 (CNN-GRU)	5934.5	7.01	469,549	5.72

In Table 1, the effectiveness of the suggested deep learning models is contrasted with that of cutting-edge models primarily employed for image classification tasks. We trained all the models for 70 epochs for measuring the training time, testing time, and the number of parameters with a learning rate of 0.0001. We used the Adam optimizer with categorical cross-entropy as the loss function. From Table 2, it can be inferred that our first CNN model outperforms all the baseline models in terms of all complexity parameters. The second proposed model has fewer parameters, and the trained model's size is significantly less than all other models, including the baseline models. Because of the GRU unit, the CNN-GRU model requires a long period of training and testing.

The accuracy as well as loss graphs for the suggested CNN model are shown in Fig. 4, while the plots for the proposed CNN-GRU model are shown in Fig. 5. We have not performed any augmentation methods to deal with data imbalance. Even without that, our system achieves excellent performance with a low complexity. We achieved an accuracy of 98.6% for both the models without doing much hyperparameter optimization. The size of the saved model in h5 format is taken for comparing the

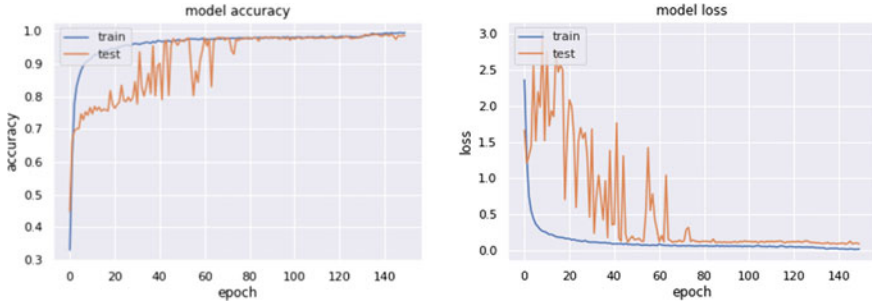


Fig. 4 Accuracy and loss of proposed CNN model

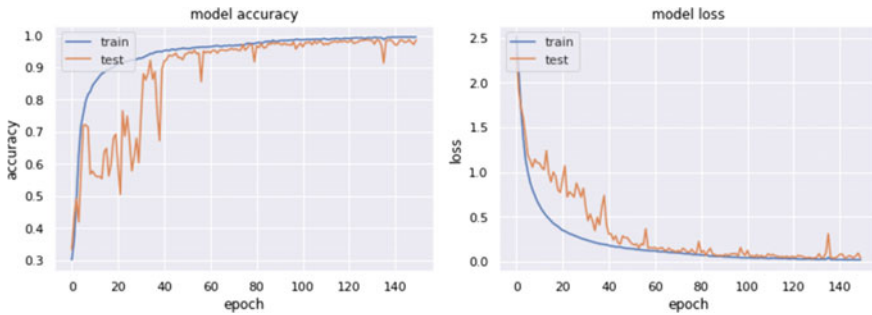


Fig. 5 Accuracy and loss of proposed CNN-GRU model

space complexity. The CNN-GRU model only takes 5.72 MB in memory for storing the model which signifies its applicability in resource-constrained environments.

5 Conclusion and Future Work

This project primarily proposes lightweight deep learning models for categorizing malware that has been turned into images. For resource-limited devices, the computational complexity of the network is as important as the accuracy. We proposed CNN and CNN-GRU-based networks having less complexity when compared to the state-of-the-art models for malware classification while matching their performance with other parameters. Another advantage of the proposed model is that there is no need for any preprocessing; the obfuscated malware also shows the same pattern when converted to images; hence, the same model suffices. The future work will be toward manually using the same models for many samples converted to images.

References

1. Namanya AP, Cullen A, Awan IU, Disso JP (2018) The world of malware: an overview. In: 2018 IEEE 6th international conference on future internet of things and cloud (FiCloud), pp 420–427
2. Cristea LM (2020) Current security threats in the national and international context. *J Account Manage Inform Syst* 19(2):351–378
3. Roseline SA, Geetha S (2018) Intelligent malware detection using oblique random forest paradigm. In: 2018 International conference on advances in computing, communications and informatics (ICACCI), pp 330–336
4. Gibert D, Mateu C, Planes J (2019) A hierarchical convolutional neural network for malware classification. In: 2019 International joint conference on neural networks (IJCNN), pp 1–8
5. Babaagba KO, Adesanya SO (2019) A study on the effect of feature selection on malware analysis using machine learning. In: Proceedings of the 2019 8th international conference on educational and information technology, pp 51–55
6. Kim CH, Kabanga EK, Kang SJ (2018) Classifying malware using convolutional gated neural network. In: 2018 20th International conference on advanced communication technology (ICACT), pp 40–44
7. bin Asad A, Mansur R, Zawad S, Evan N, Hossain MI (2020) Analysis of malware prediction based on infection rate using machine learning techniques. In: 2020 IEEE region 10 symposium (TENSymp), pp 706–709
8. Luo JS, Lo DCT (2017) Binary malware image classification using machine learning with local binary pattern. In: 2017 IEEE International conference on big data (big data), pp 4664–4667
9. Vinayakumar R, Alazab M, Soman KP, Poornachandran P, Venkatraman S (2019) Robust intelligent malware detection using deep learning. *IEEE Access* 7:46717–46738
10. Cui Z, Xue F, Cai X, Cao Y, Wang GG, Chen J (2018) Detection of malicious code variants based on deep learning. *IEEE Trans Ind Inform* 14(7):3187–3196
11. Nataraj L, Karthikeyan S, Jacob G, Manjunath BS (2011) Malware images: visualization and automatic classification. In: Proceedings of the 8th international symposium on visualization for cyber security, pp 1–7
12. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 770–778
13. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ (2017) Densely connected convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 4700–4708
14. Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, Andreetto M, Adam H (2017) Mobilenets: efficient convolutional neural networks for mobile vision applications. arXiv preprint [arXiv:1704.04861](https://arxiv.org/abs/1704.04861)
15. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z (2016) Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 2818–2826
16. Chollet F (2017) Xception: deep learning with depthwise separable convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1251–1258
17. Gurucharan M (2020) Introduction to convolutional neural networks (CNN). Retrieved from <https://www.upgrad.com/blog/basic-cnn-architecture/>
18. Yu Y, Si X, Hu C, Zhang J (2019) A review of recurrent neural networks: LSTM cells and network architectures. *Neural Comput* 31(7):1235–1270
19. Cho K, Van Merriënboer B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H, Bengio Y (2014) Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv preprint [arXiv:1406.1078](https://arxiv.org/abs/1406.1078)

20. Wikipedia contributors.: Hadamard product (matrices). In: Wikipedia, The Free Encyclopedia. Retrieved 08:31, May 5, 2022, from [https://en.wikipedia.org/w/index.php?title=Hadamard_product_\(matrices\)&oldid=1079702186](https://en.wikipedia.org/w/index.php?title=Hadamard_product_(matrices)&oldid=1079702186)
21. Chollet F et al (2015) Keras. GitHub. Retrieved from <https://github.com/fchollet/keras>

IndoorGML Modeling for WiFi-Based Indoor Positioning and Navigation



Manjarini Mallik and Chandreyee Chowdhury

Abstract With the advancement in indoor localization and navigation, indoor spaces are represented using distinct forms of spatial information. Communication and interoperability among different systems comprising different technologies demand a common standardized representation. From this demand, OGC published IndoorGML as a standard spatial data model. However, sensor-based indoor localization research hardly utilizes such spatial data model for better interoperability. The main aim of this work is to bridge this gap and propose a framework to (i) extract IndoorGML representation of the benchmark dataset of indoor localization; (ii) coarse grained localization based on sensory data with IndoorGML representation. We performed a case study on a WiFi fingerprint-based benchmark dataset for indoor localization based on our university campus. The result of the case study validates the robustness of the proposed framework.

1 Introduction

Indoor Geographic Information System (GIS)-based applications enhance the indoor location-based services [1]. The growing applications of sensor-based indoor localization systems demand a similar standardization too. Indoor areas are nowadays becoming more complex both infrastructure wise and ambience wise due to rapid urbanization. Thus, searching destination and tracking users are becoming complicated and time consuming. Various technologies have been used for this purpose, for example, WiFi, Bluetooth, inertial sensors, RFID and many more [2]. Researchers have evaluated their proposed localization works using benchmark datasets [3], others have collected new datasets on their own [4]. The available datasets are mostly found in the form of excel, csv or plain text format. One needs to understand the predefined terms and descriptions manually and arrange the dataset as needed. This process may be difficult for complex building structures, specifically if spatial information is very important for the service.

M. Mallik (✉) · C. Chowdhury
Department of Computer Science and Engineering, Jadavpur University, Kolkata, India
e-mail: manjarini.mallik@gmail.com

Hence, representation of an indoor space in terms of spatial information is important.

Few well-known indoor spatial data models including CityGML [5] of Open Geospatial Consortium (OGC) and Industrial Foundation Classes (IFC) [6] of buildingSMART mainly focus on the data exchange property, but the detailed spatial properties are not completely reflected. OGC published a new spatial data model, IndoorGML [7] in 2014 which contains standardized geometrical symbolic and topological information that are sufficient to understand the indoor spatial structure. Some recent works on IndoorGML have been found where researchers have integrated IndoorGML with some other spatial data models as in [8–10].

The work on machine learning and statistical model-based indoor positioning and the indoor geospatial modeling have taken independent research paths that need to be merged. We could not find any indoor localization datasets describing the space in standard formats for information exchange. Thus, though for localization, they report signal fingerprints, navigation strategies could not be analyzed or deduced through spatial analysis of them.

In this paper, we have attempted to merge the aforementioned research directions of indoor positioning with the spatial analysis in order to enable a complete analysis of the indoor geospatial space that can create many potential urban social applications. We have taken a benchmark dataset of indoor positioning for case study. The dataset is WiFi fingerprint datasets collected from our university campus for indoor localization purpose. The JUIndoorLoc [11] dataset consisting of RSSIs collected using smartphones from several floors of a building has been built and used for indoor localization. The fingerprints are collected for the several WiFi Access Points (APs) spread across the floor plan. Representing the dataset using IndoorGML would allow the dataset to exchange information with other applications without manual effort, leading toward the implementation of improved indoor navigation applications combining different technologies. Another important aspect of this paper is that it introduces the liaison between indoor localization based on sensory data and IndoorGML.

The contributions of this work are summarized as follows.

- A framework is proposed to map fingerprint datasets for sensor-based indoor localization to IndoorGML, and the prerequisites are identified for that.
- The framework is implemented on the JUIndoorLoc benchmark dataset, and the applicability has been shown.

Rest of the paper is structured as follows. The proposed framework is detailed in Sect. 2. The IndoorGML representation of JUIndoorLoc dataset is given in Sect. 3 with specific terms and explanations of GML followed by a conclusion in Sect. 4.

2 Proposed Framework for IndoorGML Representation from any Floorplan

To generate an IndoorGML representation from any given floor map, maintaining specific terms and structures is important. So, the necessary rules are identified, and we have proposed the framework through which IndoorGML representation can be generated for any floorplan. The workflow is represented in Fig. 1.

Since IndoorGML mainly gives structured indoor navigation pathways, so it is relevant only for that part of the floorplan which is kept to be navigable. Polygonal cells from the navigable floorplan should be divided into rectangular (or any other simple geometric shapes) cells having unique identifiers for the IndoorGML representation. Here, two components should be identified: (i) the doors between cells that enables navigation and (ii) the walls that surrounds a cell and prevents navigation. Thus, once this list of cells and their terminal points have been figured out as indicated in the figure, now the workflow can be divided into three branches:

1. generating a dual space graph to find out navigable routes between any two terminal points
2. generating an inter-layer connection graph that helps to design the emergency evacuation strategies
3. generating sensor coverage graph that helps in locating a new user in that floorplan at runtime.

The dual space graph, inter-layer connection graph and sensor coverage graph are explained in detail with the case study in the subsequent section.

3 Case Study : IndoorGML Representation of JUIndoorLoc Dataset

The JUIndoorLoc¹ [11] dataset has been chosen because we have measured the dimensions of the experimental region covered for collecting the data. IndoorGML implementation is represented in the following subsection.

3.1 Brief Overview of the Floorplan

The data were collected from a floor which is 42 meter broad along the x -axis and 21 meter long along the y -axis. The complete region was gridded into $1\text{ m} \times 1\text{ m}$ squares to measure the RSSI values in each location point (each $1\text{ m} \times 1\text{ m}$ square denotes one location point). There are total 12 accessible rooms, positioned on both sides of

¹ https://drive.google.com/open?id=1_z1qhoRIcpineP9AHkfVGCfB2Fd_e-fD.

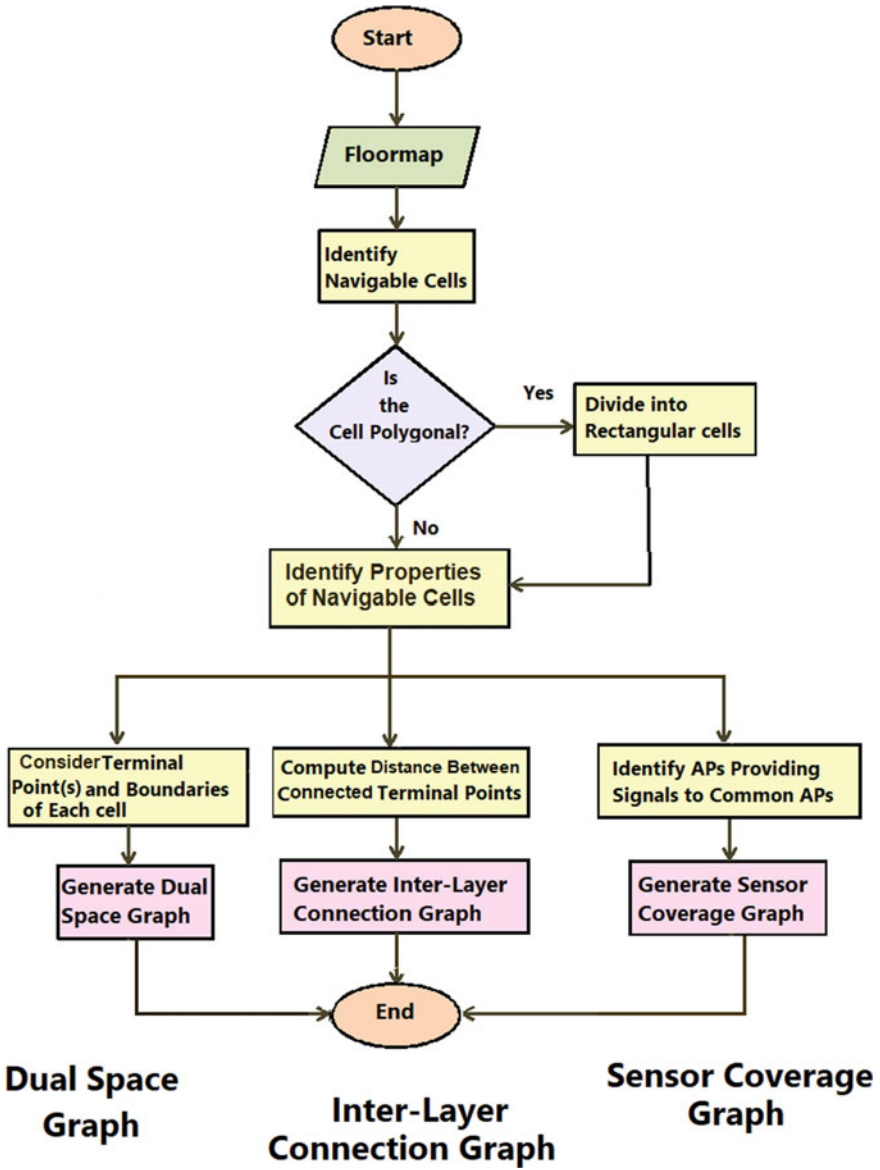


Fig. 1 Workflow for IndoorGML representation from a given floorplan

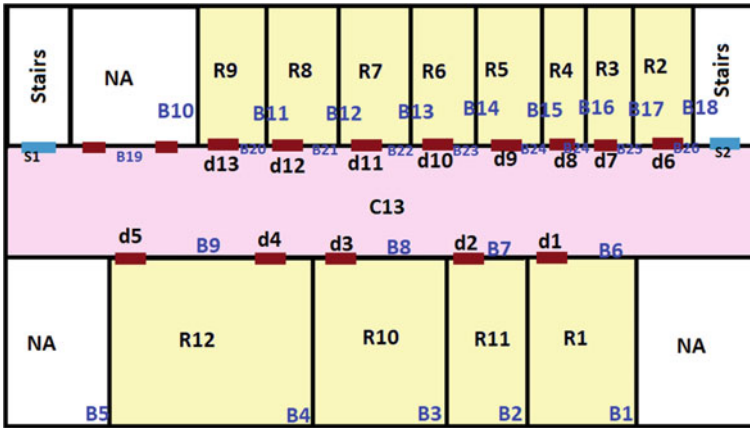


Fig. 2 JUIndoorLoc: topographic space

the corridor. Geospatial information for each room has been stored. The floorplan is shown in Fig. 2. The bold lines indicate walls (treated as boundaries) of the rooms. The doors are also indicated in the figure. Users can enter and exit from each room with corresponding door(s), specified by specific names. Two stairs connect to the outer layer of the region through which a user can exit.

3.2 Topographic Space and Dual Graph Representation

The topographic space of JUIndoorLoc is represented in Fig. 2 and its corresponding dual space graph is represented in Fig. 3. The rooms and corridor are cells. If one can navigate from one cell to another cell, these two cells are connected by a *connectivity link* in the dual space graph. The floor is mapped in such a way that one can navigate to each and every room through the corridor C13. For this reason, in the dual space graph, each node is connected by a bold line that is connectivity link, to the node C13. Again, these rooms are separated from the corridor by some boundaries, indicated by dotted line that is adjacency link between corridor and each room. Each room is adjacent to two other rooms, and separated by two boundaries, this property is also represented in the graph by adjacency link. For example, the room R10 is connected to the corridor by the door d3, and separated from corridor by the boundary B8, indicated by connectivity and adjacency links respectively. Again, R10 is separated and cannot be reached because of the boundaries B4 and B3 from the hall R12 and the room R11 respectively, represented by adjacency links between (R10, R12) and (R10, R11). s1 and s2 denote the two stairs through which one can exit to the outer region.

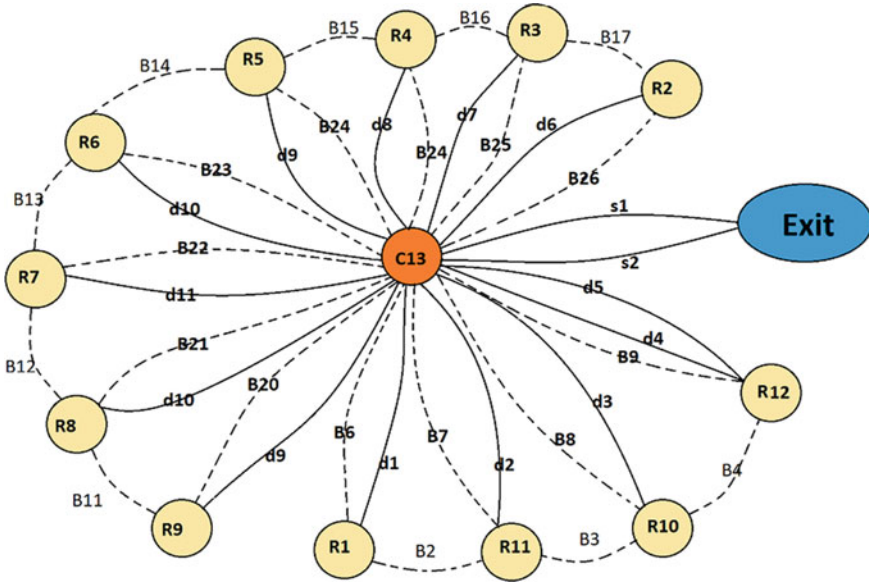


Fig. 3 JUIndoorLoc: dual space graph

3.3 Horizontal Distance Computation

Horizontal distance is the distance between any two points on the same floor, connected by a straight line. To compute the distance between two such points, we consider the doors to which these points are connected through straight lines and compute the distance between those two (or more) doors. The summation of all distances gives the actual distance of the path between the two points. In the JUIndoorLoc dataset, some doors are present in same row or column of the grid layer, and the distance between these pair of doors can be measured easily. To compute the distance between a pair of doors which are not in the same row or column, we have used Pythagoras theorem [12], as shown in Fig. 4.

Since all the rooms are positioned side by side on both sides of the corridor, one can reach from any door to any other door. For this reason, the door-to-door graph of JUIndoorLoc is a complete graph connecting 13 doors by 78 edges, as shown in Fig. 5 representing the inter-layer connection graph with two parts. One is the door-to-door graph (shown in lower part of the graph) representing connectivity among doors, and other is the dual graph for topographic space layer (shown in upper part of the graph) representing connectivity among cells, i.e., rooms and corridor. These two layers are connected through dotted links, indicating which cell is directly connected to which door. The bold lines between each room/hall to corridor indicate that all these cells can be reached from the corridor. The distance between each door-to-door pair is the weight of the link connecting those two doors. To avoid clumsiness, the distances are not shown in Fig. 5.

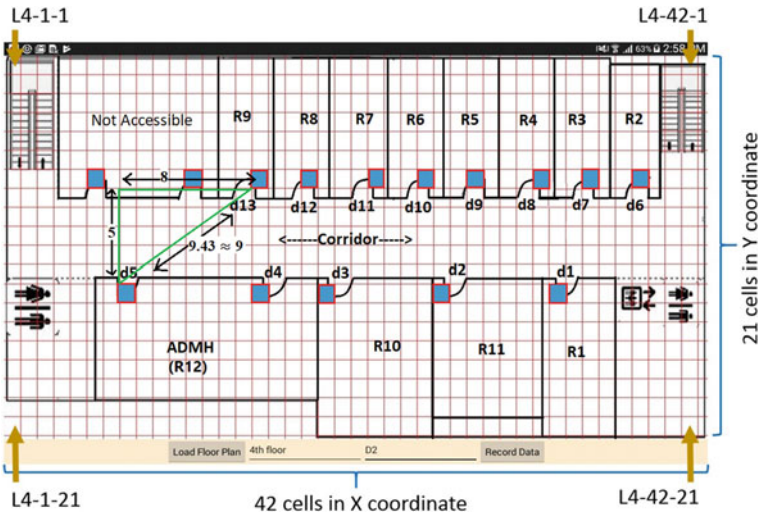


Fig. 4 JUIndoorLoc: computing distance between two doors

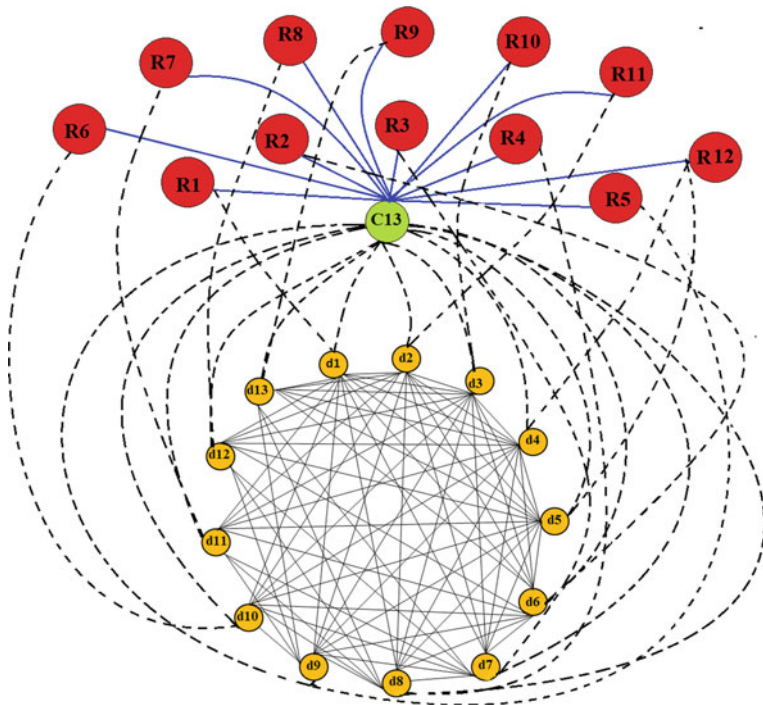


Fig. 5 JUIndoorLoc: inter-layer connections

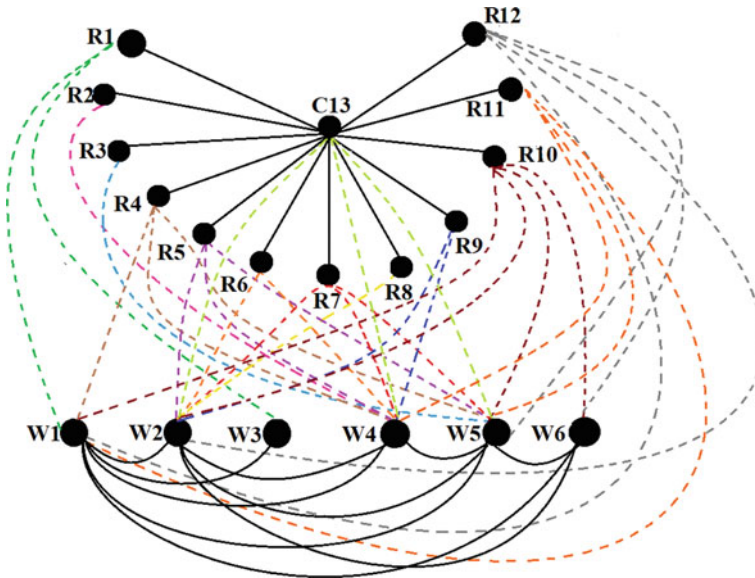


Fig. 6 JUIndoorLoc: sensor coverage graph

3.4 Sensor Coverage Graph

In order to establish a liaison between the spatial representation of IndoorGML and the WiFi fingerprint data representation for indoor positioning, a sensor coverage graph is constructed. The graph indicates the connectivity among the access points (APs) and cells. The complete dataset of JUIndoorLoc contains 172 APs, each covering more than one cells. We have selected six important APs for which RSSI values are received for 90% or more sample points for each of the cells covered by the AP. Two APs are connected by bold edge if they provide signals to at least one common region (Fig. 6). The APs providing signal to cells are connected using dotted link. For example, if someone is receiving signals from the APs W1 and W3, it may be assumed that he/she is positioned near to the room R1. In the sensor coverage graph, the cells which can be navigated from one to another are also connected using bold edges. Since in JUIndoorLoc each cell is navigable from corridor C13, each of them is connected to C13 by individual bold edge. The original id of these six APs is stored in Table 1, along with the information of their signal covering area.

3.5 Representation of JUIndoorLoc using IndoorGML

The details of JUIndoorLoc dataset are stored in XML file using GML format. For each cell, the information including its boundary coordinate, navigation property through connectivity and adjacency links, its geometry type is stored. In Fig. 7, the

Table 1 Access points and their covering area

AP	ID	Covering area
W1	AP005	R1, R4, R10, R11, R12
W2	AP023	R5, R6, R7, R8, R9, R10, R12, C13
W3	AP009	R1
W4	AP029	R2, R4, R5, R6, R7, R9, R11, C13
W5	AP019	R3, R4, R5, R6, R7, R10, R11, R12, C13
W6	AP025	R10, R12

Table 2 GML tags and their meanings

GML tag	Source	Meaning
<code>gml:name</code>	<code>gmlBase.xsd</code> ^a	A label or identifier for any object, usually a descriptive name
<code>gml:boundedBy</code>	<code>feature.xsd</code> ^b	Minimum bounding area in square or rectangular in shape that encloses the whole feature
<code>gml:Box</code>	<code>feature.xsd</code>	Defines a specific square/rectangular space with coordinates
<code>gml:Edge</code>	<code>topology.xsd</code> ^c	The 1-dimensional primitive used to define topology among cells
<code>gml:description</code>	<code>gmlBase.xsd</code> ^a	A simple textual description of any object

^a<http://www.datypic.com/sc/niem21/s-gmlBase.xsd.html>

^b<http://www.datypic.com/sc/niem21/s-feature.xsd.html>

^c<http://www.datypic.com/sc/niem21/s-topology.xsd.html>

details of room R1 are displayed. The complete region is bounded by the coordinates (0, 42) and (0, 21). The room R1 is bounded by the coordinate (34, 13) and (38, 21). The room is connected to cell C13 through door d1, which is represented by connectivity link in the dual space graph in Fig. 3. R1 is non-navigable to/from the cells R11 and C13 because of the boundary walls B2 and B6, respectively. These are represented as adjacency link in Fig. 3. There is another file named “R1 data.xml” linked to this xml file that store the RSSI value received from different APs in each 1 m × 1 m square area of room R1. One can reach to this file for localization with fine granularity. Details of all other cells are stored in a similar way. The file has been validated with online validator.²

A GML feature is defined as a representation of identifiable real-world objects in specific domains of communication. Various schema terms for GML defined by OGC are stored³ in different files. The terms used in the above case studies are explained in brief in the following Table 2 along with their sources.

² <https://codebeautify.org/xmlvalidator/cb9d16d8>.

³ <http://www.datypic.com/sc/niem21/ss.html>.

```

<gml:name>JUIndoorloc</gml:name>
  <gml:boundedBy>
    <gml:Box srsName="http://www.opengis.net/gml/srs/epsg.xml#4326">
      <gml:coord><gml:X>0.0</gml:X><gml:Y>0.0</gml:Y></gml:coord>
      <gml:coord><gml:X>42.0</gml:X><gml:Y>21.0</gml:Y></gml:coord>
    </gml:Box>
  </gml:boundedBy>
  <Room>
    <gml:name>R1</gml:name>
    <gml:boundedBy>
      <gml:Box srsName="http://www.opengis.net/gml/srs/epsg.xml#4326">
        <gml:coord><gml:X>34.0</gml:X><gml:Y>13.0</gml:Y></gml:coord>
        <gml:coord><gml:X>38.0</gml:X><gml:Y>21.0</gml:Y></gml:coord>
      </gml:Box>
    </gml:boundedBy>
    <gml:GeometryType>Polygon</gml:GeometryType>
    <gml:Edge>
      <gml:name>d1</gml:name>
      <gml:description>Connectivity link to C13</gml:description>
    </gml:Edge>
    <gml:Edge>
      <gml:name>B2</gml:name>
      <gml:description>Adjacency link to R11</gml:description>
    </gml:Edge>
    <gml:Edge>
      <gml:name>B6</gml:name>
      <gml:description>Adjacency link to C13</gml:description>
    </gml:Edge>
    <xi:include href="R1data.xml"/>
  </Room>
</FloorModel>

```

Fig. 7 Information of room R1 of JUIndoorLoc stored using GML format

4 Conclusion

To make a connection between different indoor navigation and communication systems, we need some common standardized representation format. Geospatial information is widely accepted for defining fine-tuned location information, but unfortunately, it is mainly used for outdoor localization. We have made an effort to represent a fingerprint-based WiFi dataset for indoor areas, using IndoorGML format describing the indoor areas with geospatial information. Before explaining the datasets with IndoorGML format, we have briefly described how datasets can be represented in general with IndoorGML format.

This is the initial step toward introducing IndoorGML in implementing sensory data-based navigation approach. In the future, we aim to propose a graph-based shortest path navigation approach and identify the user's location using these datasets stored in IndoorGML format. User's location can be determined from sensory data using sensor coverage graph and based on that the system can recommend suitable navigation path for users.

Acknowledgements This research work is partially supported by the project entitled—“Developing Framework for Indoor Location Based Services with Seamless Indoor Outdoor Navigation by expanding Spatial Data Infrastructure,” funded by the Ministry of Science and Technology, Department of Science and Technology, NGP Division, Government of India.

References

1. Qingxiang C, Jing C, Wumeng H (2020) Method for generation of indoor GIS models based on BIM models to support adjacent analysis of indoor spaces. *ISPRS Int J Geo Inform.* <https://doi.org/10.3390/ijgi9090508>
2. Priya R, Chandreyee C (2021) A survey of machine learning techniques for indoor localization and navigation systems. *J Intell Robot Syst.* <https://doi.org/10.1007/s10846-021-01327-z>
3. Priya R, Chandreyee C, Mausam K, Dip G, Sanghamitra B (2020) Novel weighted ensemble classifier for smartphone based indoor localization. *Exp Syst Appl.* <https://doi.org/10.1016/j.eswa.2020.113758>
4. Salamah AH (2016) Mohamed Tamazin, Maha Sharkas, Mohamed Khedr? an enhanced WiFi indoor localization system based on machine learning. *Int Conf Indoor Position Indoor Navig (IPIN).* <https://doi.org/10.1109/IPIN.2016.7743586>
5. OGC, OGC CityGML Encoding Standard, Document No. 12-019, 2012. <http://www.opengeospatial.org/standards/citygml>
6. BuildingSMART, IFC Standard. <http://www.buildingsmart-tech.org/specifications/ifc-overview>
7. OGC, OGC IndoorGML, Document No. 14-005r4, 2014. <http://www.opengeospatial.org/standards/indoorgml>
8. Claridades AR (2019) Inhye Park: Integrating IndoorGML and Indoor POI Data for Navigation Applications in Indoor Space. *J Kor Soc Surv Geodesy Photogram Cartography* 359–366. <https://doi.org/10.7848/ksgpc.2019.37.5.359>
9. Kim JS, Yoo SJ, Li KJ (2014) Integrating IndoorGML and CityGML for indoor space 184–196. https://doi.org/10.1007/978-3-642-55334-9_12
10. Alattas A, Zlatanova S, Oosterom P, Chatzinikolaou E, Lemmen C, Li KJ (2017) Supporting indoor navigation using access rights to spaces based on combined use of IndoorGML and LADM models. *ISPRS Int J Geo Inform.* <https://doi.org/10.3390/ijgi6120384>
11. Roy P, Chowdhury C, Ghosh D, Bandyopadhyay S (2019) JUIndoorLoc: a ubiquitous framework for smartphone-based indoor localization subject to context and device heterogeneity. *Wireless Pers Commun* 739–762. <https://doi.org/10.1007/s11277-019-06188-2>
12. Weisstein EW, Pythagorean theorem, Wolfram Research, Inc. <https://www.mathworld.wolfram.com>

Assessing Damage of Natural Disasters from Satellite Imagery Using a Deep Learning Model



Shubham Tikle, P. Jidesh, and A. Smitha

Abstract Natural disasters are events that arise anywhere on the planet. It causes enormous devastation and places entire cities in need of significant support. The capability to swiftly and precisely deploy rescue services in the affected regions is critical for reducing the impact and saving lives. A two-step model is developed in an attempt to resolve this problem by using satellite images as input. The model draws attention to the structures such as buildings, which are severely damaged. The current deep learning-based computer vision models use pre- and post-disaster satellite images to semantically infer the level of damage to individual buildings after natural disasters. This model alleviates an important roadblock in disaster managerial decisions by simplifying the evaluations of damages caused to the building. We used DeepLabv3+ for semantic segmentation and a custom CNN model for image classification to analyze disaster-related images. This paper describes how satellite data and proficient image analysis may effectively be used to conduct disaster and crisis management to assist jobs that require fast mappings. This model's performance and accuracy are sub-optimal and are being studied to further improvisations. However, it surpasses the current cutting-edge model.

Keywords Damage classification · DeepLabV3+ · Custom CNN · Natural disasters · Crisis-management

S. Tikle (✉) · P. Jidesh

Department of Mathematical and Computational Sciences, National Institute of Technology
Karnataka, Karnataka, India

e-mail: shubhamrameshtikle.203cd003@nitk.edu.in

P. Jidesh

e-mail: jidesh@nitk.edu.in

A. Smitha

Department of Information and Communication Technology, Manipal Institute of Technology,
Manipal, India

e-mail: smitha.a@manipal.edu

1 Introduction

The satellite images provide a real-time image of the Earth system. That is one of the reasons why it has become a valuable source for various geophysical studies by the research community. Satellite images provide a wide range of uses and benefits. One such application that requires accuracy is the detection and classification of natural disaster damage using satellite images. Detecting where damage has occurred and sent rescue teams to the location of event of a natural disaster can save many lives and also assists in determining where the actions should be concentrated. As a result, the damage evaluation based on the satellite image is the most important part of the process.

This is vital for organizations to identify the whereabouts of injured persons in the initial hours of post-disaster in order to speed up relief efforts and the deployment of response activities [1]. Building damage is frequently used as a proxy for population concentration [2]. Because of its vast coverage area and data availability, remote sensing is an efficient technique for locating damaged buildings. Humanitarian aid, on the other hand, mostly relies on conventional computerization of structural damages, which also persists as a most successful way. Manual digitization is time-consuming, requires skilled image observers, is unsuitable for large areas, and is vulnerable to inconsistencies and mistakes caused by fatigue or poor quality regulation. The time it takes to produce damage assessment reports would be considerably reduced if this procedure can be automated.

To help humans cope with the massive workload of handling huge quantities of the satellite images, computer vision is utilized to simplify the identification of objects and situations crucial to crisis management. Despite this capability, the existing computerized systems appear to be incapable of recognizing all intensities of building damages with a good degree of certainty [3].

A lot of research has been done in recent years, with the swift advancement in the area of machine learning for computer vision, especially using deep neural networks (DNNs) [4]. DNNs also obtains good-level of performance on computer vision tasks, including image analysis and pattern segmentation. As a result, these techniques are useful for instantly retrieving useful information from satellite data.

2 Related Works

Authors have explored various methods for classifying natural disasters using satellite images from various datasets. Some of these are discussed further below.

Seismological factors cause millions of buildings to fall when earthquakes strike [5]. Numerous machine learning techniques were used for forest fire forecasts since the 1990s. They employ a machine learning approach in a recent study. This study in Italy used the random forest methodology to map wildfire sensitivity [6]. It discovered

that the presented method could recognize locations that can be impacted by forest fires. The limitation of this work is that it deals with only one type of disaster.

Flood susceptibility mapping [7] used two new combination ensemble models, dagging and remote sensing, along with three machine learning models to model flood risk mapping in the Teesta River basin, northern Bangladesh. Artificial neural networks, random forests, and support vector machines are three examples. The fact that their dataset only covers water disaster categories is a significant limitation because it does not cover other types of disasters.

In [8], the researchers describe the techniques that the team used to solve disaster identification problems. They created two flood catastrophe datasets: One was created via social media images; the other is based on satellite photographs. The photos were trained using the GoogleNet architecture. The fact that their set of data only covers water disaster categories is a significant limitation.

Photos shared on social media portals throughout natural calamities are investigated in the work [9] to determine the extent of conditions caused by the disasters. The presented method can successfully adapt deep-CNN features to assess the degree of damage from social media images captured after a disaster strikes. The authors gathered images from the Internet of various disasters, including Typhoon Ruby, Hurricane Matthew, and the Nepal Earthquake. The researchers also used Web search to gather images of collapsed buildings, bridges, and roads, among the other things. The dataset's limitation is that it only includes images of broken infrastructure.

Convolutional neural networks have been used extensively in this damage detection problem (CNN). One method combined ordinal regression and CNN [10].

Support vector machines and Naive Bayes are two types of machine learning algorithms used to classify natural calamities on different parameters, but it is bound to only the initial stages of natural disaster [11]. Text mining and regular log mining techniques are used to detect earthquakes on seismological data, but the drawback is that earthquake detection is dependent on public feedback [12].

Recent research has looked into using CNNs to detect disasters from satellite photos [13–16]. In [13–15, 17], the researchers employ CNNs to identify damaged infrastructure by categorizing them into two groups: damaged and non-damaged.

The paper only identifies the presence or absence of buildings following a natural disaster or examines damage for a single disaster type [18, 19].

3 Research Gaps

From the detailed literature review, it is evident that the majority of the works are based on advanced deep learning with hybrid architectures. The previous literature has primarily focused on picture classification as issue damage or non-damage type, i.e., binary classification, and most of the literature focuses on only one type of disaster. The ability to detect the level of damage present in the image is limited to no-damage, minor-damage, major-damage, or destroyed. In an emergency, however, the satellite images can provide more specific information about things, people, and situations.

3.1 Motivation

Damage assessment from satellite photography is difficult, especially for different sorts of disasters process. After a disaster, logistics, resource planning, and damage estimation are tough tasks, and putting first responders in post-disaster conditions is risky and expensive. The work's primary motivation is to automate the process of detecting damage from satellite images and also to further classify the damage into various levels. The use of passive technologies for damage assessment, such as satellite imagery analysis, saves time and money, reduces relief time, lowers risk, and expedites an otherwise dangerous process. Every year, natural disasters happen, and they are inevitable events. So, this is a field that requires relevant and continuous research.

4 Approach and Methodology

The overall process pipeline of the classification of damages is shown in Fig. 1 which is self explanatory.

4.1 Dataset Collection

The set of data comes from xView2.org and is labeled with xBD (Building Damage). In order to fully support deep learning for building damage assessment, datasets of appropriate scope, scale, size, and standard must be available. For this reason, we use an xBD labeled dataset. xBD addresses the limitations enumerated in others by collecting data across six disaster types, 15 countries, and thousands of square kilometers of imagery, and Fig. 2 illustrates a sample of segmented images. Furthermore, they introduce a joint damage scale that provides guidance and an assessment scale to label building damage in satellite imagery, as shown in Fig. 3 [20].

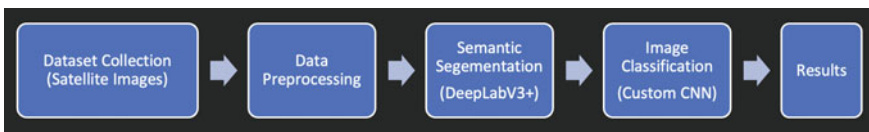


Fig. 1 Process pipeline

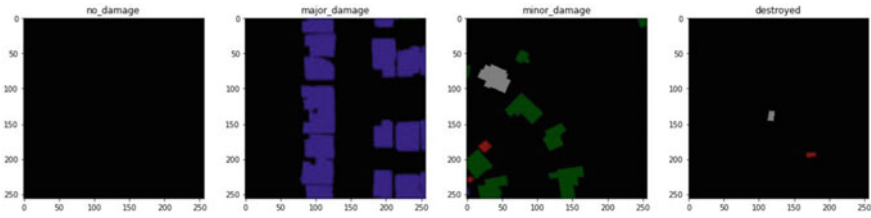


Fig. 2 Sample of segmented images

Disaster Level	Structure Description
0 (No Damage)	Undisturbed. No sign of water, structural or shingle damage, or burn marks.
1 (Minor Damage)	Building partially burnt, water surrounding structure, volcanic flow nearby, roof elements missing, or visible cracks.
2 (Major Damage)	Partial wall or roof collapse, encroaching volcanic flow, or surrounded by water/mud.
3 (Destroyed)	Scorched, completely collapsed, partially/ completely covered with water/mud, or otherwise no longer present.

Fig. 3 Joint damage scale

4.2 Data Preprocessing

It is an integral step in deep learning, especially dealing with images, to preprocess the data before proceeding further. In an emergency, however, the photos can provide more specific information about the things, people, and situations that suppresses undesired distortions or enhances some apparent features relevant for feature processing and analysis tasks. The images and their corresponding labels are given in JSON format. Before proceeding to the next step, all JSON files were converted to PNG files. Extracted all pixel properties, created a polygon for the damaged area, and assigned different colors to each damaged area, such as white cyan, dark green, causal blue, and maroon red for no-damage, minor-damage, major-damage, and destroyed, respectively. The image resolution in the dataset is 1024×1024 . However, it is resized into 256×256 for proper training. Before processing the model, image segmentation and data augmentation are done.

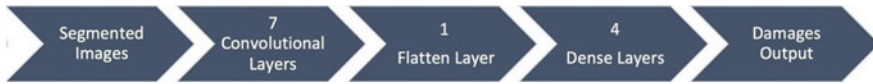


Fig. 4 Customized CNN model

4.3 Model Selection and Creation

The aim is to develop a replicable model. That can be used to aid in rescue efforts in emerging regions. The overall procedure for this project consists of three major steps. The initial step is data preprocessing, in which image processing is being performed to break down images to a standard size and normalize it, and then feed them into our model. The second step is the semantic segmentation of buildings. The first step's algorithm is fed by a satellite image, which is then processed using a semantic segmentation like Deeplabv3+, which masks the damaged areas with different colors like white cyan, dark green, causal blue, and maroon red for no-damage, minor-damaged, major-damaged, and destroyed, respectively. Finally, the semantic segmentation results are fed into a custom CNN model Fig. 4 consisting of seven convolutional layers, one flatten layer, and four dense layers with softmax as the activation function. The Adam optimizer is employed, with a learning rate of 0.0001, which estimates the damages caused by each structure and produces a damage classifier—no-damage, minor-damage, major-damage, or destroyed. From these details, the appropriate authorities can be informed, and immediate assistance can be provided.

4.4 Experiments and Outcomes

Hyperparameter tuning performed on the custom CNN model as shown in Table 1.

4.5 Results and Observations

To develop a model for our damage classification, the buildings had to be labeled in four different ways, and the final layer had four output nodes. As mentioned earlier, for segmentation of images, the model used is DeepLabV3+. ResNet50 layer is used as a backbone in the DeepLabV3+. In this model, loss Fig. 5a and Intersection of Union (IOU) Fig. 5b are the important parts. The output segmented images of the DeepLabV3+ model are the input to the custom CNN model, where the results for 12 epochs are 70.84% accurate on training data and 72.06% accurate on testing data. The CNN model F1-score and loss as shown in Fig. 6a, b, respectively. The model was trained using 5500 images. The ratio of training to testing is 80:20. In the previous

Table 1 Hyperparameters tuning on custom CNN model

Sr. No	Convolutional layers	Learning rates	Batch Size	Kernel size	Accuracy (%)
1	4(Filters = 128)	0.001	8	2*2	66.07
2	4(Filters = 128)	0.002	8	2*2	65.11
3	5(Filters = 128)	0.001	16	2*2	67.00
4	5(Filters = 128)	0.002	16	2*2	67.50
5	6(Filters = 128)	0.001	16	2*2	69.00
6	6(Filters = 128)	0.002	32	2*2	68.86
7	6(Filters = 128)	0.003	32	2*2	69.01
8	7(Filters = 128)	0.0003	32	2*2	69.74
9	7(Filters = 128)	0.0001	32	2*2	69.84
10	7(Filters = 128)	0.0001	32	2*2	70.00
11	7(Filters = 128)	0.0001	64	2*2	71.05
12	7(Filters = 128)	0.0001	128	2*2	72.00
13	7(Filters = 256)	0.0001	256	2*2	72.06
14	7(Filters = 256)	0.0002	128	2*2	70.07
15	7(Filters = 256)	0.0002	256	2*2	70.50
16	7(Filters = 256)	0.0003	128	2*2	70.87
17	7(Filters = 256)	0.0003	256	2*2	71.10

Bold value represents the Best/Maximum output accuracy or F1-score after training the model

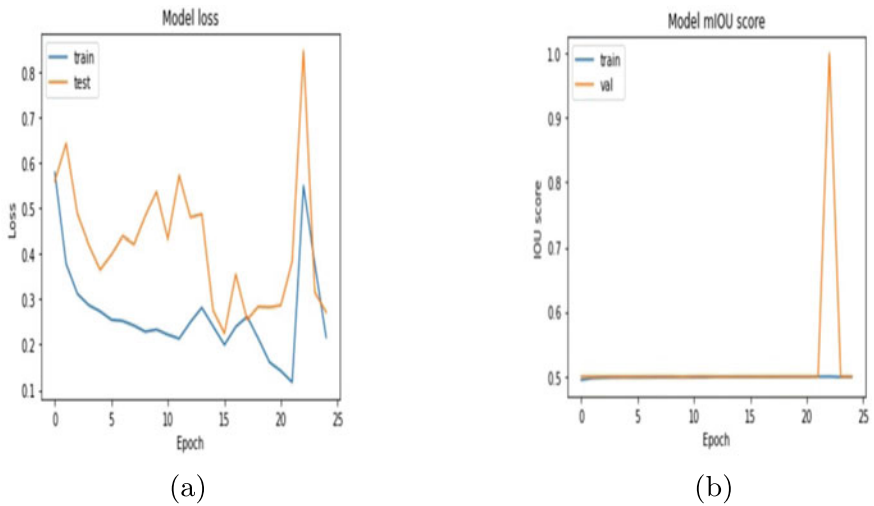


Fig. 5 Performance metrics of Deeplabv3+ model, **a** DeeplabV3+ model loss, **b** DeeplabV3+ model IOU

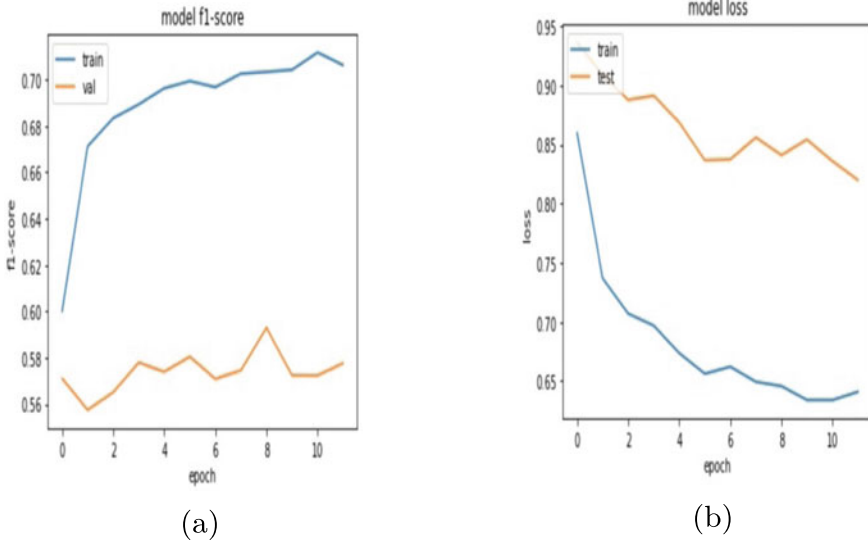


Fig. 6 Performance metrics of Custom CNN model, **a** Custom CNN F1-score, **b** Custom CNN model loss

Table 2 Accuracy on training and testing dataset

Sr. No	Models	Training (%)	Testing (%)
1	Mask R-CNN & ResNet50	72.1	68.2
2	DeepLabV3+ & Custom CNN	70.84	72.06

Bold value represents the Best/Maximum output accuracy or F1-score after training the model

work on this particular dataset in 2020, with a value of a learning rate of 0.0003, on the training dataset, the model achieved 72.1% accuracy and 68.2% accuracy on the testing dataset [21]. The results are tabulated in Table 2. In their work, they combined four classes into two. Such as putting ‘no-damage’ and ‘minor-damage’ into one class and ‘major-damage’ and ‘destroyed’ into another. One of the reasons they are getting higher training accuracy than the proposed model could be because it is a binary classification.

5 Conclusion

Using deep learning algorithms, the proposed methodology attempts to build an efficient and accurate damage classification system. Deep learning models are typically tuned using various hyper-parameter tuning techniques. Other issues such as over-fitting and class imbalance are also taken into account. Classifying the damage into different levels is difficult, and the damages are usually very different in each image.

It does not adhere to any specific patterns or structures is the main challenge during training the custom CNN Model. In summary, the proposed damage detection and classification system detects damages and can further classify the type of damages. This opens the door for future research in this area.

References

1. Beamon BM, Balcik B (2008) Performance measurement in humanitarian relief chains. *Int J Public Sector Manage*
2. Dell'Acqua F, Gamba P (2012) Remote sensing and earthquake damage assessment: experiences, limits, and perspectives. *Proc IEEE* 100(10):2876–2890
3. Albrecht CM, Elmegreen B, Gunawan O, Hamann HF, Klein LJ, Lu S, Schmude J (2020) Next-generation geospatial-temporal information technologies for disaster management. *IBM J Res Dev* 64(1/2)
4. Ghazouani F, Farah IR, Solaiman B (2019) A multi-level semantic scene interpretation strategy for change interpretation in remote sensing imagery. *IEEE Trans Geoscience Remote Sens* 57(11):8775–8795
5. Gupta R, Hosfelt R, Sajeev S, Patel N, Goodman B, Doshi J, Gaston M et al (2019) xbd: a dataset for assessing building damage from satellite imagery. *arXiv preprint arXiv:1911.09296*
6. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521(7553):436–444
7. Mignan A, Broccardo M (2020) Neural network applications in earthquake prediction (1994–2019): meta analytic and statistical insights on their limitations. *Seismological Res Lett* 91(4):2330–2342
8. Tonini M, D'Andrea M, Biondi G, Degli Esposti S, Trucchia A, Fiorucci P (2020) A machine learning-based approach for wildfire susceptibility mapping. The case study of the Liguria region in Italy. *Geosciences* 10(3):105
9. Rahman M, Chen N, Islam MM, Mahmud GI, Pourghasemi HR, Alam M, Dewan A (2021) Development of flood hazard map and emergency relief operation system using hydrodynamic modeling and machine learning algorithm. *J Cleaner Prod* 311:127594
10. Zaffaroni M, Lopez-Fuentes L, Farasin A, Garza P, Skinnemoen H (2019) AI-based flood event understanding and quantification using online media and satellite data
11. Kradolfer U (2013) SalanderMaps: a rapid overview about felt earthquakes through data mining of web-accesses. In: *EGU general assembly conference abstracts*, p EGU2013-6400
12. Amit SNKB, Aoki Y (2017) Disaster detection from aerial imagery with convolutional neural network. In: *2017 international electronics symposium on knowledge creation and intelligent computing (IES-KCIC)*, IEEE, pp 239–245
13. Cao QD, Choe Y (2020) Building damage annotation on post-hurricane satellite imagery based on convolutional neural networks. *Nat Hazards* 103(3):3357–3376
14. Duarte D, Nex F, Kerle N, Vosselman G (2018) Satellite image classification of building damages using airborne and satellite image samples in a deep learning approach. *ISPRS Ann Photogrammetry Remote Sens Spatial Inf Sci* 4(2)
15. Fujita A, Sakurada K, Imaizumi T, Ito R, Hikosaka S, Nakamura R (2017) Damage detection from aerial images via convolutional neural networks. In: *2017 fifteenth IAPR international conference on machine vision applications (MVA)*, IEEE, pp 5–8
16. Ignatiev V, Trekin A, Lobachev V, Potapov G, Burnaev E (2019) Targeted change detection in remote sensing images. In: *Eleventh International Conference on Machine Vision (ICMV 2018)*, vol 11041, SPIE, pp 677–682
17. Frolking S, Qiu J, Boles S, Xiao X, Liu J, Zhuang Y, Qin X (2002) Combining remote sensing and ground census data to develop new maps of the distribution of rice agriculture in China. *Global Biogeochem Cycles* 16(4)

18. Xu JZ, Lu W, Li Z, Khaitan P, Zaytseva V (2019) Building damage detection in satellite imagery using convolutional neural networks. arXiv preprint [arXiv:1910.06444](https://arxiv.org/abs/1910.06444)
19. Saito K, Spence RJ, Going C, Markus M (2004) Using high-resolution satellite images for post-earthquake building damage assessment: a study following the 26 January 2001 Gujarat earthquake. *Earthquake Spectra* 20(1):145–169
20. Gupta R, Goodman B, Patel N, Hosfelt R, Sajeev S, Heim E, Gaston M et al (2019) Creating xBD: a dataset for assessing building damage from satellite imagery. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops, pp 10–17
21. Crow C, Juwono M, Royesh A, Tesfaye TT. Satellite image analysis for building detection and natural disaster damage classification

Fundamental Graphical User Interface Design of an Educational Android Application



Vandana, Kapil Kumar Nagwanshi, Anil Kumar, and Manish Paliwal

Abstract A graphical user interface (GUI) design of an application presents a user-friendly, elegant, concise, and attractive mechanism for interacting with visual components. Android is a powerful Linux-based mobile operating system, and these days android devices are becoming the most sought-after and widely used devices than other types of mobile devices in the mobile application market. Nowadays, android applications (herein as ‘android apps’) have also become popular in the educational sector because of their compatible, efficient, and effective interface to access the content. In android apps, graphical designs are built from GUI components which is an object with which mobile user interacts via various forms of input. This paper presents the basic structured GUI design of an educational android application. These days, the GUI design of any android is an increasingly important factor in deciding whether your app will be accepted by mobile users to achieve success in the market and it also builds a connection between users and applications. In addition, users have more expectations and demands for good quality GUI design. GUI design contains three phases: informational architecture, interaction, and visual design. The aim of this paper helps to explain how to create a basic GUI design of an educational android application by using an integrated development environment (Android Studio) and programming languages such as Extensible Markup Language (XML) with different graphical components such as visual layouts and widgets. To assess user experience, this study has carried the survey using a questionnaire set. The main motive (purpose)

Vandana
ASET, Amity University Rajasthan, Jaipur, India

K. K. Nagwanshi (✉)
CSE, SoS E&T, Guru Ghasidas Vishwavidyalaya, Bilaspur, CG, India
e-mail: dr.kapil@ieee.org

A. Kumar
Faculty of Engineering and Technology, Wachemo University, Hossana, Ethiopia

M. Paliwal (✉)
Pandit Deendayal Energy Univeristy, Raysan, Gandhinagar, Gujarat, India
e-mail: manish.paliwal@cotpdpu.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
V. V. S. S. Chakravarthy et al. (eds.), *Advances in Signal Processing, Embedded Systems and IoT*, Lecture Notes in Electrical Engineering 992,
https://doi.org/10.1007/978-981-19-8865-3_47

of the survey was to assess (evaluate) the effectiveness and the user experience of the interface design.

Keywords Graphical user interface · GUI design · Android · Android application · Educational application

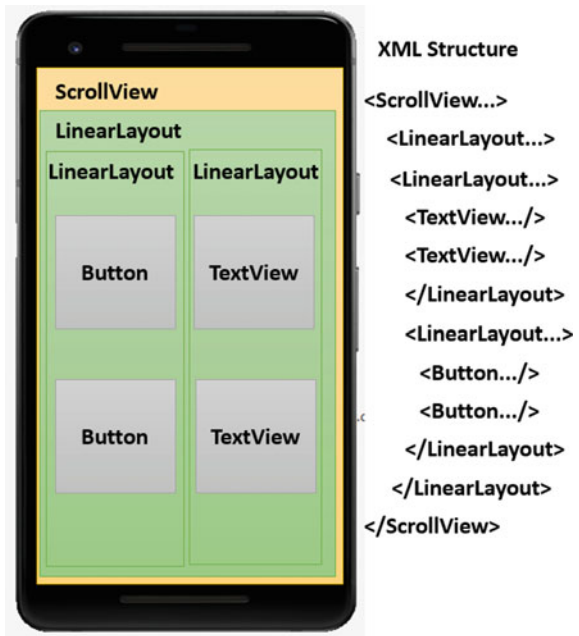
1 Introduction

Android devices are increasingly becoming popular in the education sector for learning and teaching students because of their robust, high quality, interactive and comprehensive user interface (UI) design including icons, layouts, components, styles, colors, patterns, etc. [18]. Many colleges and universities are using android devices to support students in constructive, communicative, supportive, and collaborative activities. In the android market, several high-quality educational applications are not successful because of their complex UI. Hence, ease of use is one of the key components of an acceptable and successful educational application [11]. Thus, when designing an educational android application, the developer needs to focus on efficiency, learnability, effectiveness, understandability, usefulness, ease of use, attitude, and intention to use that will ensure a high level of acceptance. The innovation in educational android apps has increased interest among learners and educators because it facilitates visual learning and teaching method [17].

2 Background and Literature Survey

Android UI design tools (e.g., Android Developers, Mockplus, iDoc, etc.) provide the latest standards to create effective UI design with the innovation of technology [16]. For an educational android app, the experience of end-users (e.g., learner and educator) is extremely important [6]. Therefore, the main purpose of this study is to describe how to create a basic GUI design of an educational android application. In addition, a prototype application has been developed composed of seven activities including Study Tour Reports, Digital Library, Events Info, Legends Corner, Online Practice Test, FAQs Portal, and Mentors Corner. These activities are developed with XML (for UI designing) and Java Programming Language by using the Android Software Development Kit (SDK) [13]. The paper is structured as follows: First, we discuss the basic UI structure of an android app. Then, we introduce UI design and explain the basic prototype design of an educational android application. Later, we discuss the result analysis, and finally, we conclude this paper with future work.

Fig. 1 Basic UI structure of an android application with XML format



2.1 Basic UI Structure of An Android Application

As compared to the desktop application, the GUI design of an android application is easier because it fulfills user’s expectations and demands by providing them high-quality user interface through android development tools [7]. For example, drag and drop widgets (visual UI elements such as text boxes and buttons) with a visual design interface. In android application development, a layout file is the face of the application and it appears at the startup of the screen. As can be seen in Fig. 1 which shows the basic UI structure of an android application which includes four common UI elements like ScrollView, LinearLayout, Button, and TextView [15].

In the XML layout files, ScrollView provides the layout for the scrolling view and enables users to combine UI elements (such as LinearLayout, TextView, and Button) within a scrolling view. A LinearLayout is a view group (base class) that arranges all children views (elements) in a single direction either vertically or horizontally by specifying the orientation attribute. TextView is a UI control that displays text to the user based on requirements. It is basic building block for user interface components and buttons are GUI components which is used to perform an action whenever the user taps/clicks on it [2].

3 UI Design Methodology

3.1 UI Design Principles

- **Affordance:** It acts as signals and a feature or property of an object that can be used to allow people to know how to use it.
- **Visibility and Natural Mapping:** In human-computer interaction (HCI), it plays an important role and it is one of the most important aspects in design because when something is out of sight, it is difficult to know about and use. Natural Mapping refers to the relationship between control and effect.
- **Low Physical Effort:** It emphasizes design where individuals can be used comfortably and efficiently and with a minimum of fatigue or exertion [12].
- **Learnability:** It is the ease with which a software application can be understood by users and it is also a vital element that defines usability in UI design [19].
- **User Satisfaction:** It consists of all aspects of the end-users relationship with the organization, its products, and its services.
- **Feedback:** Feedback guides users through the step-by-step process and tells us whether or not we are approaching the goal
- **Error Tolerance:** It is the design that reduces the difficulty of incidental actions.
- **Reducing Cognitive Burden:** By using certain methods, cognitive burden can be reduced like remove an irrelevant task, ignore useless elements, reduce choices, and use iconography with caution.
- **Flexibility:** It allows that software designs should keep up the ability to customize certain tasks based on the various needs of specific users [8].
- **Simplicity:** It is the discipline of reducing, rectifying, or revert a design.

3.2 Activities Stack in Android

Android activity is a single screen of the android app where on one activity, users can take various actions. An android app may consist of one or more activities where the user navigates inside the app, it opens and closes applied activities [1]. When redirecting from one page (activity) to another, the page that the end-user quits is not destroyed but is appended to a stack and its condition is preserved. Hence, the user can achieve this activity using the “Back” button of the mobile device and the activity will open in entirely the same state as at the time of its stack, previous activity task being destroyed. This activity life cycle can be seen in Fig. 2. The developed prototype app has seven activities: Study Tour Reports, Digital Library, Events Info, Legends Corner, Online Practice test, FAQs Portal, and Mentors Corner. Above-mentioned activities are designed for the android platform using XML. In XML file, main element is `LinearLayout` which is present in all activities and additionally added further UI elements from the android library such as `ScrollView`, `TextView`, `ImageView`,

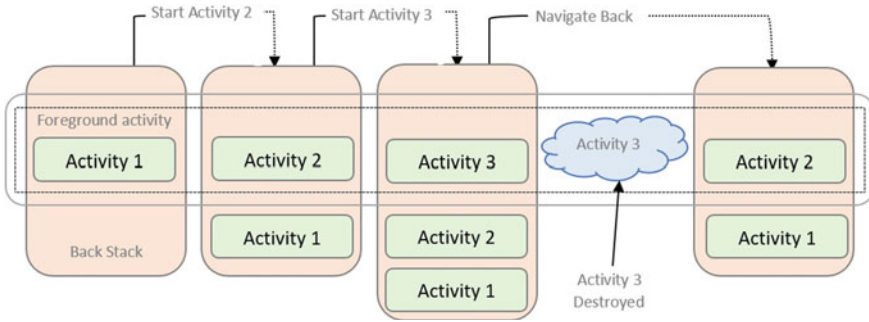


Fig. 2 Activities stack in android

ProgressBar, and ExpandablePlaceHolderView. The ExpandablePlaceHolderView design by using RecyclerView, ExpandableListView, and develop on top of PlaceHolderView with the parent-child structure. ExpandablePlaceHolderView has two types of item views: Parent item view (defines the header view and contains the child item views in collapsed/expanded state) and Child item view (associated with their corresponding parent item views).

4 Basic Prototype Design of an Educational Android Application

This activity will show study tour reports of students which will be a unique travel experience that integrates learning with traveling. This interface can be seen as a moment for students to learn and increase their technical knowledge and acquire different experiences by visiting national or international universities. The study tour is defined with two terms: theory and practice. The basic graphical interface design of the study tour activity (see Fig. 3) includes 3 sections: national tour, international tour, and other tours. In this activity, users can view photos of their tour experience with the ImageView widget and read information from the TextView widget. The main aim of this interface is to increase the skill development and knowledge building of students [5]. It is also useful for student’s self-reformation.

4.1 Digital Library

This activity is about a digital library which is an electronic store where books, research papers, journals, and other information stored in digital formats and accessible by electronic devices such as smartphones, computers, and tablets. It is also called a type of information retrieval system (IRS) because its content may be saved

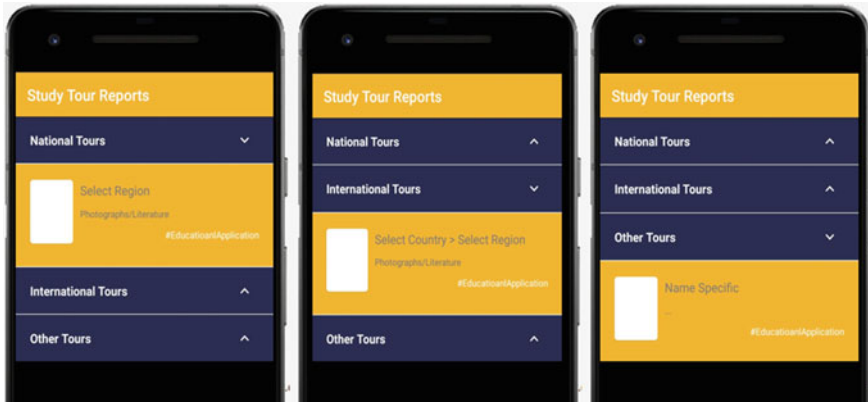


Fig. 3 Basic prototype graphical interface design of the study tour activity



Fig. 4 GUI design: **a** prototype of the digital library activity, **b** the events info activity, and **c** the legends corner activity

on the local computer, or accessed online via computer networks [20]. As shown in Fig. 4a, basic user interface created of digital library in which user can access multiple data such as course books, reference books, encyclopedia (online Dictionary with videos, pictures, and facts), standard books, and other information from a single platform (android application) [9]. Pressing one of the arrow icons from above-mentioned activity, the GUI will display associated information in another view. In this activity, the user can also scroll down the page to view the entire information through the ScrollView widget.

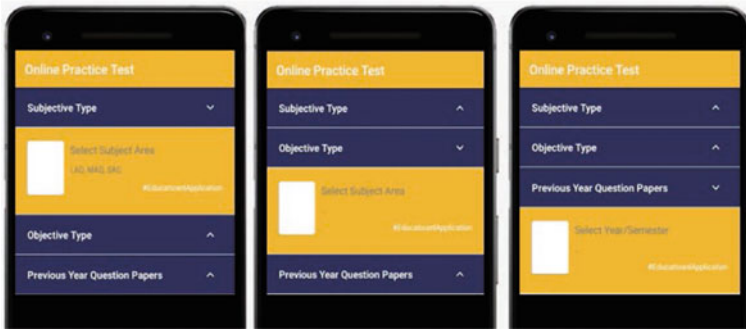


Fig. 5 Basic graphical interface design of the online practice test activity

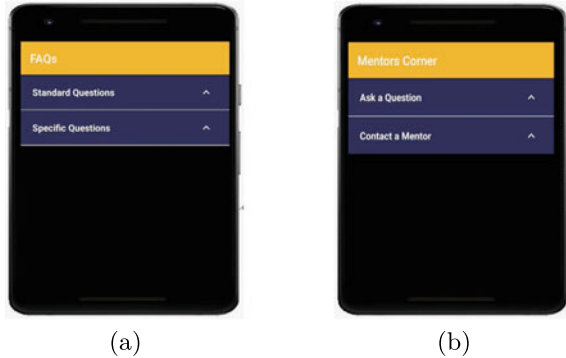
4.2 Events Info

This activity will provide students an affordable and accessible way to engage them with the latest events and programs [10]. As shown in Fig. 4b, events information available like conference alert (displays details of upcoming national/international academic conferences on a distinct subject), seminar, webinar, competition, annual Function, convocation, alumni meet, jury, and study visit.

4.3 Legends Corner and Online Practice Test

In legends corner activity, living and non-living legends life history, philosophy, and their famous work available so that user can view data of very popular living legends like Raj Rewal, B.V. Doshi, and Hafeez Contractor and non-living legends like Charles Correa, Zaha Hadid, and Le Corbusier. Students can read about legends milestone achievements like their construction and design. `LinearLayout`, `ImageView`, `TextView`, and `RecyclerView` widget used in this activity to show their information interactively. As mentioned in Fig. 4c, name-wise legends are listed in `ExpandablePlaceHolderView`, and their information is displayed in the child item view. Online practice test activity facilitates architectural aspirants in learning, practicing, and improving their knowledge skills. They can prepare themselves for the national and provincial level test directly through smartphones in hand [3]. This helps aspirants with their entrance exam preparation and they can check their level of preparation through this activity [4]. As shown in Fig. 5, the online practice test interface contains `LinearLayout`, `ImageView`, `TextView`, `ExpandablePlaceHolderView`, `RecyclerView` widget which displays question type sets in the main menu. In the submenu, further information is available. This interface is very simple and interactive for assessment of end-users in online practice tests such as National Aptitude Test in Architecture (NATA).

Fig. 6 GUI design for: **a** FAQs portal activity, and **b** the mentors corner activity



4.4 *FAQs Portal and Mentors Corner*

In this layout, two options are given, namely: Standard Questions and Specific Questions. In the first option, standard questions like “What are the benefits of this app?” and answers available. In the second option “what is NATA”, “why it is important”, and “how to prepare for NATA” and their answers available. As shown in Fig. 6a, FAQs consist of one layout in which two categories are mentioned with ExpandablePlaceHolderView and LinearLayout widget. Mentors are good problem-solvers, good listeners, and observers and they make an effort to respect the interest and goals of a student [14]. As mentioned in Fig. 6b, the basic graphical interface design of the mentor’s corner activity includes two options “Ask a Question” and “Contact a Mentor” by using the interactive android widget. From the “Ask a Question” menu, students and aspirants can raise their questions, e.g., “how do you motivate innovative thoughts”, and from “Contact a Mentor”, they can contact them from the given list of experienced mentors.

5 Result Analysis

The prototype proposed here was developed having in mind the mobile learning with a user-friendly design interface using the android platform. To assess user experience, this study has carried the survey using a questionnaire set. The main motive (purpose) of the survey was to assess (evaluate) the effectiveness and the user experience of the interface design. Based on the survey result, from the total of ninety-three respondents (58% Male and 42% Female). Figure 7a and b shows percentage distribution according to gender and age in which almost half of the respondents (49.0%) aged between 19 and 23, 42.0% respondents aged 16–18, and the rest (9.0%) of them aged 24–30. Figure 7c pie chart questionnaire shows that from the total of sixty-six respondents, a substantial percentage of respondents (77.4%) find out that they are satisfied with the application’s user interface design, and (22.6%) respondents are

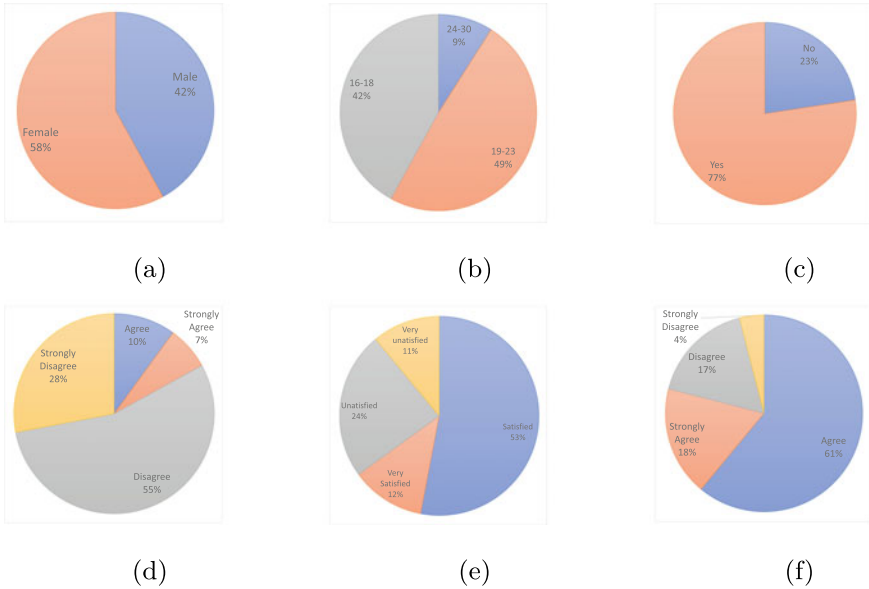


Fig. 7 Questionnaire: **a** gender, **b** age, **c** user satisfaction for the design of the application, **d** efforts and practice required, **e** efficiency of the application, and **f** clear understandable interaction

not satisfied with the interface. Figure 7d pie chart questionnaire shows that from the total of seventy-two respondents, 55.0% of respondents disagreed with the statement that it would take much time to learn and use the application with proficiency and 28.0% of respondents strongly disagreed but 10.0% of respondents agreed with the above-mentioned statement and 7.0% strongly agreed. Figure 7e pie chart questionnaire shows that from the total of ninety-three respondents, 53.0% of respondents satisfied with the prototype of the educational application to help them in the preparation of the national level test, and 12.0% of respondents very satisfied but 24.0% of respondents unsatisfied with the above-mentioned statement and 11.0% very unsatisfied. Figure 7f pie chart questionnaire shows that from the total of seventy-eight respondents, 61.0% of respondents agreed with the statement that the interaction with the prototype of the educational application is clear and understandable for them and 18.0% of respondents strongly agreed but 17.0% of respondents disagreed with the above-mentioned statement and 4.0% strongly disagreed.

6 Conclusion and Future Work

We proposed a prototype design of an educational application to support online learning through the android platform. This design is composed of software components, i.e., Study Tour Reports, Digital Library, Events Info, Legends Corner, Online Prac-

tice Test, FAQs Portal, and Mentors Corner. This application has proved its usefulness, as the analysis presented above section. As discussed in Sects. 2 and 3 basic UI structure and design of an android application, we found a significant improvement in the development of the educational android application (Result Analysis Section). The percentage obtained from the survey “Result Section” by users confirms that an interactive, user-friendly, clear, and understandable design interface is the need of all users (students and aspirants). The aim of this research was double-fold; first to explain how to create a basic UI design of an educational android application by using the android platform. In this regard, it was obvious from the analysis of the questionnaire that effective UI design was required to design a successful and acceptable android application. Based on the user’s perspective, these include efficiency, responsiveness, effectiveness, and accessibility required for good UI, where the aforementioned features represent the basic requirement of users, respectively. The second motive of this paper was to propose a high-quality educational application to support students in supportive, constructive, collaborative, and communicative activities. Furthermore, and based on the user’s perspective, these include flexibility and scalability that they can fulfill users’ needs as required. In the future, we want to implement some other features, intended to make this application functional for users so that they can use proposed features as discussed in the paper. Also, we want to add career and upload portfolio options to this application to make this application more useful and effective. We believe that this will be possible, as we described above, it will increase the skill development and knowledge building of students and also useful for student’s self-reformation.

References

1. Adinugroho TY, Gautama JB et al (2015) Review of multi-platform mobile application development using webview: Learning management system on mobile platform. *Proc Comp Sci* 59:291–297
2. Adipat B, Zhang D (2005) Interface design for mobile applications. *AMCIS 2005 Proc* 494
3. Astra IM, Nasbey H, Nugraha A (2015) Development of an android application in the form of a simulation lab as learning media for senior high school students. *Eurasia J Mathe Sci Technol Educ* 11(5):1081–1088
4. Cavus N (2016) Development of an intelligent mobile application for teaching English pronunciation. *Proc Comp Sci* 102:365–369
5. Cieza E, Lujan D (2018) Educational mobile application of augmented reality based on markers to improve the learning of vowel usage and numbers for children of a kindergarten in trujillo. *Proc Comp Sci* 130:352–358
6. Furió D, Juan MC, Seguí I, Vivó R (2015) Mobile learning versus traditional classroom lessons: a comparative study. *J Comp Assis Learn* 31(3):189–201
7. Georgiev T, Georgieva E (2009) User interface design for mobile learning applications. *e. Learning* 9:145–150
8. Ines G, Makram S, Mabrouka C, Mourad A (2017) Evaluation of mobile interfaces as an optimization problem. *Proc Comp Sci* 112:235–248
9. Johnsson BA, Weibull G (2016) End-user composition of graphical user interfaces for palcom systems. *Proc Comp Sci* 94:224–231

10. Kleinwort R, Semm T, Falger PM, Zaeh MF (2018) Integration of an android application into the learning factory for optimized machining. *Proc Manuf* 23:9–14
11. Kocakoyun S, Bicen H (2017) Development and evaluation of educational android application. *Cypriot J Educ Sci* 12(2):58–68
12. Łobaziewicz M (2015) The design of b2b system user interface for mobile systems. *Proc Comp Sci* 65:1124–1133
13. Ma L, Gu L, Wang J (2014) Research and development of mobile application for android platform. *Int J Multimed Ubiquit Eng* 9(4):187–198
14. Malhotra R, Kumar D, Gupta D (2020) An android application for campus information system. *Proc Comp Sci* 172:863–868
15. Moran K, Bernal-Cárdenas C, Curcio M, Bonett R, Poshyvanyk D (2018) Machine learning-based prototyping of graphical user interfaces for mobile apps. *IEEE Trans Softw Eng* 46(2):196–221
16. Ozdamli F, Cavus N (2011) Basic elements and characteristics of mobile learning. *Proc Soc Behav Sci* 28:937–942
17. Papadakis S, Kalogiannakis M (2017) Mobile educational applications for children: what educators and parents need to know. *Int J Mobile Learn Organ* 11(3):256–277
18. Papadakis S, Kalogiannakis M, Zaranis N (2018) Educational apps from the android google play for Greek preschoolers: a systematic review. *Comp Educ* 116:139–160
19. Sánchez-Morales LN, Alor-Hernández G, Rosales-Morales VY, Cortes-Camarillo CA, Sánchez-Cervantes JL (2020) Generating educational mobile applications using uidps identified by artificial intelligence techniques. *Comp Stand Interf* 70:103407
20. Yahaya NS, Salam SNA (2014) Mobile learning application for children: Belajar bersama dino. *Proc Soc Behav Sci* 155:398–404

Personality Detection Using Signature Analysis



Garapati Jaya Surya Koushik, Kasukurthi Tirudeepak, and Dwijen Rudrapal

Abstract Every human is having unique personality. Personality plays vital role for a human to be placed in dream position always. Different research work proposed different automated personality prediction system based on different measures like face, behaviour, attitude, GAIT and handwritten signature. Predicting personality using handwritten signature is a difficult problem. Signature has different features to assess the pattern and predict. In this paper, we propose a personality prediction system based on handwritten signature pattern using neural network. We have analysed various features, classify and extracted their values from signatures in the corpus. We devise a computational strategy for minimising mistakes in each class feature value selection for the entire system. We have also conducted a statistical significance test by computing alpha value and improved the accuracy of our proposed system.

Keywords Computational analysis · Graphology · Human personality · Statistical analysis of human personality · Theoretical analysis of human personality

1 Introduction

Graphology [3] is basically a study of a particular person handwriting qualities and patterns [Miguel] in order to find or assume his or her psychological state of mood at the time of writing. Handwriting is a projective test in which the unconscious emerges and communicates itself to the conscious. It works because our hands are controlled by our subconscious mind when writing [1]. A writer will not consciously draw each letter by hand when writing, just as a person will when typing. Humans have always been fascinated by the diversity and uniqueness of each individual. A handwriting expert can roughly explain a person's personality and character traits by analysing the handwriting [9]. Here, we use a person's signature instead of handwriting for analysis. The signature is a special reflection of the clot of everything on a person's body, and his natural mark expresses all nature. Signature analysis is

G. J. S. Koushik · K. Tirudeepak · D. Rudrapal (✉)
Computer Science and Engineering, NIT, Agartala, India
e-mail: dwijen.rudrapal@gmail.com

a part of the handwriting analysis, To understand the personality of a person, the relationship between individual signature and handwriting analysis [4] plays crucial role. A person's signature is mainly one of most unique things they process. We generally do not find any two persons having the exact same signature. From a signature personality characteristics of any individual, like his mind being balanced, kindness to himself and others, anger issues and his way of behaviour with others and himself, all these things can be found in a person's signature.

The goal of this project is to implement our proposed system which recognises the personality of a person using deep learning models. Signature is the smallest text you can write, and it tells a lot about an individual. In this implementation, we tried to figure out a number of signature features which are very unique. As we get the features and good accuracy with neural network models, then we merged the primary features. We divided the data into several classes and used statistical significance to improve the model's accuracy. By considering features like angle, size, we then merged into four different classes, and then, we compute the alpha value. Then, we compute the statistical significance accuracy.

2 Related Work

In the last decade, several research works related to prediction of personality [2, 11] have been carried out by various researchers. Prior researches achieved promising result with comparable accuracies.

Most of the recent promising researches are based on machine and deep learning approaches. The work by [8] proposed a machine learning-based system using various extracted features like space, baseline, slant and pressure, etc. The work experimented on 3 datasets: CEDAR, GPDS Synthetic Signature, BHSig260. The work carried out by [3, 6] analysed signature images for personality identification using features like margin, dot structure, breaks, start, end streak, underline. Authors proposed a prediction model based on artificial neural networks (ANNs). The work by [5] predicted personality of human handwritten signature based on histogram oriented gradient (HOG) and backpropagation neural networks (BPNNs). The work showed that the use of ANN on large input dimensions causes a long processing time.

The approaches proposed based on deep learning offer better performance and accuracy using extensive and interconnected feature extraction. However, these approaches need state-of-the hardware and huge amount of training data. Convolutional neural networks (CNNs) are a popular deep learning network for personality prediction from handwritten signature. Prior research work proposed by many researchers [7, 10, 12] is based on CNN and achieved state-of-the-art accuracy.

3 Proposed Approach

In this section, we have explained our dataset preparation, features selection and proposed model.

3.1 Dataset Preparation

We have collected handwritten signature corpus from Kaggle¹. The corpus contains signatures of 30 people. Each person has 5 signatures which they made themselves. Each signature is saved in image form for processing task. In addition of the collected corpus, we added more signatures from undergraduate students. Finally, our handwritten signature dataset includes 204 signatures.

3.2 Feature Selection

The basic features of the signatures are discussed in this section for further analysis and classification system development. The main features of the signatures that are used in our current work are angle, size, underline, circle, last big letter, last small letter, presence of dot. The primary features are the ones which are definitely present in a signature. Those are angle and size. There are three sub-classes in angle feature. They are ascending, descending and straight.

Similarly, size is also having two sub-classes. These are identical and non-identical.

3.3 Feature Classification

Personalities based on signatures may be divided into four classes. Identification of each personality characteristics depends on the presence of one or more features of signature. Table 1 shows the mapping of personality classes, features and human characteristics.

3.4 Classifier Selection

We have proposed CNN and ANN for all feature classifications and personality prediction. CNN models and an ANN models have been trained with various images

¹ <https://www.kaggle.com/datasets/divyanshrai/handwritten-signatures>.

Table 1 Classification of signature features

Personality classes	Human characteristics	Features
Complex person	Unpredictable	Non-identical
Shy and stable	Shy and modest, emotionally stable	Identical, straight
Modest and confident	Optimistic or confident	Identical, ascending
Lethargy and modest	Lack of energy	Identical, descending

Table 2 Confusion matrix for the maximum dataset ($N = 204$)

Complex	Shy and stable	Modest and confident	Lethargy and modest
37	20	1	2
24	44	2	5
16	18	22	0
4	5	0	4

of their respective properties. We have used the “ReLU” activation function for convolution layers. For binary classification, “Sigmoid” activation function is used for output layers, and for categorical classification, “Softmax” activation function is used. The optimiser used in the compile function was “adam”. These models give us whether a particular property’s presence is there in a signature or not. Based on the features value presence, proposed prediction system predicts personality of signature owner’s.

4 Experiment Result and Analysis

Table 2 represents the confusion Matrix for the proposed approach. The neural network technique displays the diagonal terms 37, 44, 22, 4 in the table.

Rateria [8] shows the amount of samples and the machine learning technique’s adjusted projected data. In the table below, we generated sample sizes ranging from 204 to 80. A maximum accuracy rate of 53% was obtained. The points represent the data generated by a neural networks technique. The line through the points is the equation of reciprocals of number of sample sizes N . The intercept of Fig. 1 is almost 0.5302 at the limit of $N \rightarrow \infty$. As N reaches infinity, then accuracy reaches 53%. The slope of this curve is around -4.514 . Although the slope may change due to a variety of variables, the intercept will remain constant (Table 3).

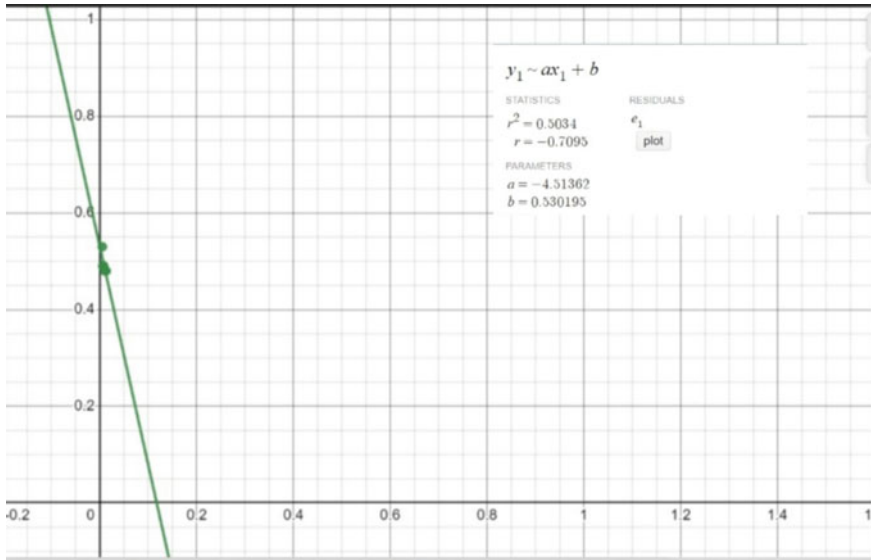


Fig. 1 Accuracy rate of proposed method

Table 3 Accuracy rate for various sample size

Signature count	1/N	Predicted correctly	Accuracy
204	0.004901	107	53
180	0.005556	89	49
150	0.006667	74	49
130	0.007692	64	49
100	0.010000	48	48
80	0.012500	39	48

5 Proposed Statistical Method

We employed an error calibration mode to detect and rectify errors. By defining and analysing faults in each class, a statistical metric called alpha was introduced. Although manual errors are possible, they may be caused by a lack of accuracy. By using the experiment data, we developed a system classification and decreasing the mistakes in this experiment further. We used the neural network to create the classification system and then check its performance. We assumed that the signatures in the data that moved into a different classes were part for that class. As there are four different classes, we computed the value of α is approximate to 0.173.

The accuracy rate graph is plotted as a function of the reciprocals of total number of samples to confirm this value of alpha. The accuracy rate enhances when in a limit of N tends to infinity because the standard deviation will be reduced to a tiny

number. The signatures were moved in the correct class. With the premise that N reaches infinity in the limit, because the standard deviation is reduced to a minimal amount, the accuracy rate will improve and error margin.

$$(\Delta_x)^2 = (z_{score})^2 * \sigma_{est}(1 - \sigma_{est})/N \tag{1}$$

(Δ_x) Margin of Error and σ_{est} is the value of standard deviation. The accuracy rate reaches 0.53 when we have a huge dataset. It is unaffected by the group of persons that choose the category for those 204 signatures. We used these 204 signatures as our starting point.

The values α_{ki} can be written in the following way,

$$C_i^j(N_i) = \sum_{k=1}^{k=4} \alpha_{kj} A_{ij}^k(N_i) \tag{2}$$

The value of α is calculated based on the value of the accuracy per cent of the curve when $N \rightarrow \infty$ this is shown in Fig. 1. This value is shown at 0.53. The final value of α is also calculated as the limit of the number of sample sizes goes to infinity. α is dependent on the number of sample sizes (N_i).

The original truth set was complex = 60, shy and stable = 75, modest and confident = 56, lethargy and modest = 13. We have seen in the limit of $N \rightarrow \infty$, the accuracy rate moves to 0.53. In complex person category, the machine predicted 37 signs out of 60 signatures. The machine also incorrectly predicted 24 out of 75 shy and stable signs as complex. Similarly, 16 out of 56 signatures from modest and confident and 4 signs out of 13 from lethargy and modest to be in complex. to machine. In our method, there is a probability “*alpha*” in which 24 signs *24alpha* will be complex signs where, “*alpha*”. Similarly, from modest and confident , there is a probability *16alpha* will be in complex, and for lethargy, *4alpha* can go into complex. So, the probability of total number of signs will be in complex category is

$$\begin{aligned} C_1^1 &= A_{11}^1 + \alpha A_{21}^2 + \alpha A_{31}^3 + \alpha A_{41}^4 \\ &= 37 + \alpha 24 + \alpha 16 + \alpha 4 + \alpha 1 \\ &= 37 + 44\alpha \end{aligned}$$

$37 + [24 + 16 + 4]\alpha$. When $N_i = 204$ for counting the number of complex person. Where, $N_1 = 204$. $j = 1$ (complex category) similarly for for shy and stable where $j = 2$ $44 + [18 + 5 + 20] \alpha = 44 + 43 \alpha$ *alpha* and for modest and confident ($j = 3$), $22 + 3\alpha$, same for lethargy and modest category $j = 4$, number of signatures will be $4 + 7\alpha$. The total number of corrected signatures for all categories will be $107 + 97\alpha$. From Fig. 1 above, we know the value will be 0.53. Thus, $[(107 + 97\alpha)/204] = 0.53$. From this, we can get value of alpha as $\alpha = 0.012$. for $N_1 = 204$. We done this for all values of N_i and then determined the average for alpha as and we got average probability is 0.068 (Figs. 2 and 3; Table 4).

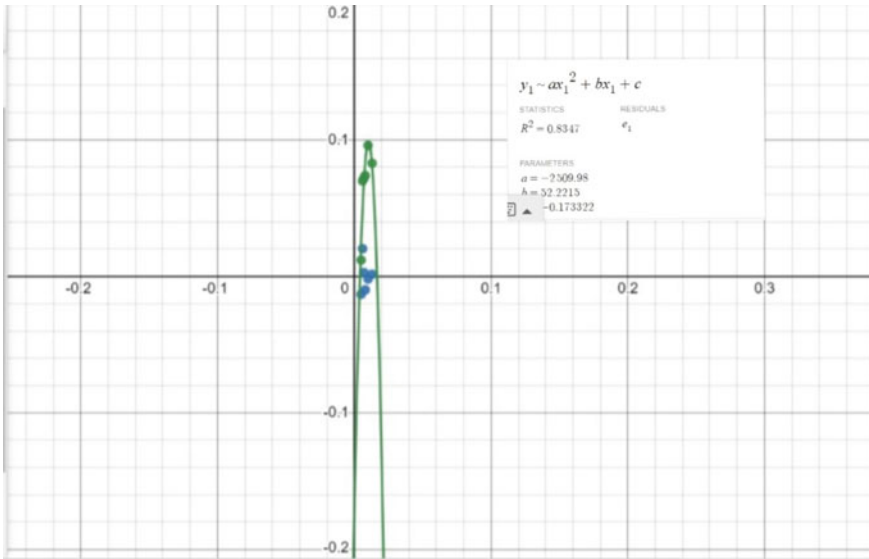


Fig. 2 Value of alpha as a function of (1/N)

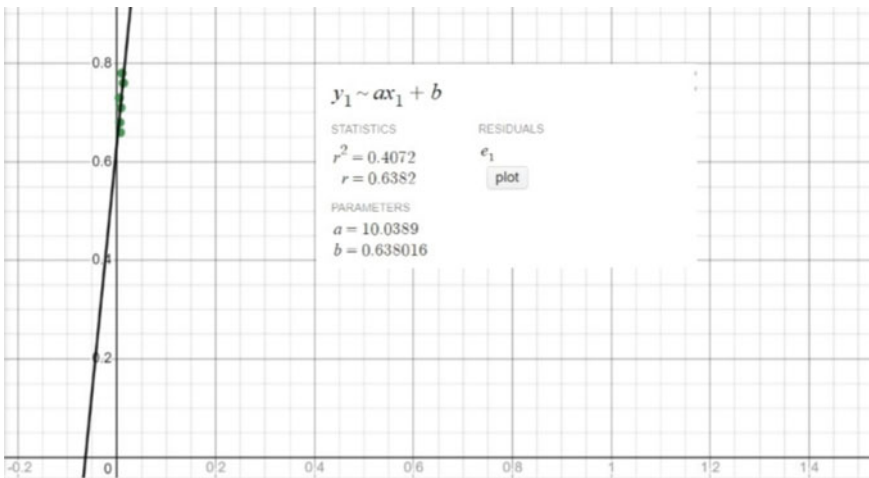


Fig. 3 The accuracy rate of complex class using statistical method

It has been observed that when alpha is introduced, the accuracy rate improves when compared to the neural networks technique. Table 5 displays the correctly predicted signatures by our new statistical approach for class complex; column 1 means the column 2 is correctly evaluated signatures by statistical significance. The sum of accuracy rates remains at 0.53 for all classes and for all sample sizes, while the sample size increased from 80 to 204. This shows that the statistical significance was

Table 4 Redistribution of signatures in other category with an assumption of equal weight

N	$1/N(x)$	Complex	Shy	Modest	Lethargy	Prediction	Accuracy	$(\alpha k_j = \alpha)k$
204	0.005	37.53	44.5	22.04	4.08	108.12	0.53	0.012
180	0.006	31.94	41.52	18.49	3.42	95.4	0.53	0.007
150	0.007	27.49	36.16	14.43	1.43	79.5	0.53	0.072
130	0.008	23.27	30.70	13.44	1.37	68.9	0.53	0.074
100	0.010	18.88	22.34	10.38	1.38	53	0.53	0.096
80	0.013	14.07	17.75	9.25	1.33	42.4	0.53	0.083

Table 5 By including alpha, the value of complex was corrected

N	Extracted data	Predicted by Stat. method	Predicted by NN method
204	60	44	37
180	53	36	29
150	47	31	25
130	38	27	21
100	27	21	16
80	21	16	12

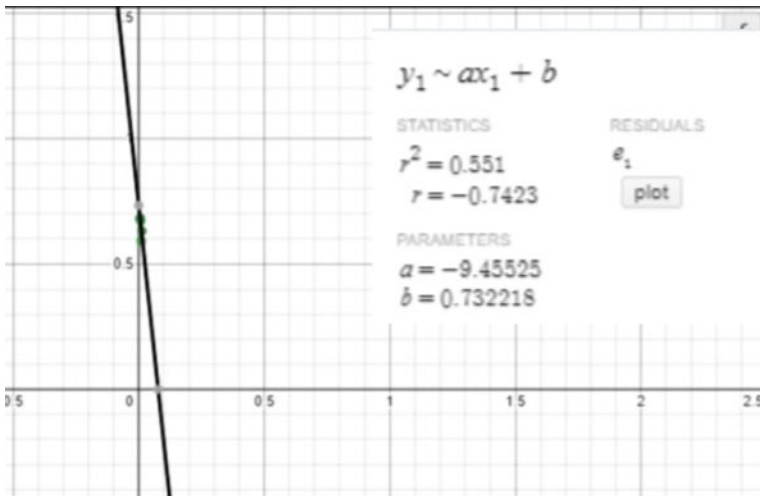


Fig. 4 Individual shy and stable accuracy rates calculated using statistical method

used to move signatures from one category to another as the sample size was changed. So, if there are more shy and stable signs, this could explain why some categories, such as lethargy and modesty, modesty and confidence, have low accuracy rates.

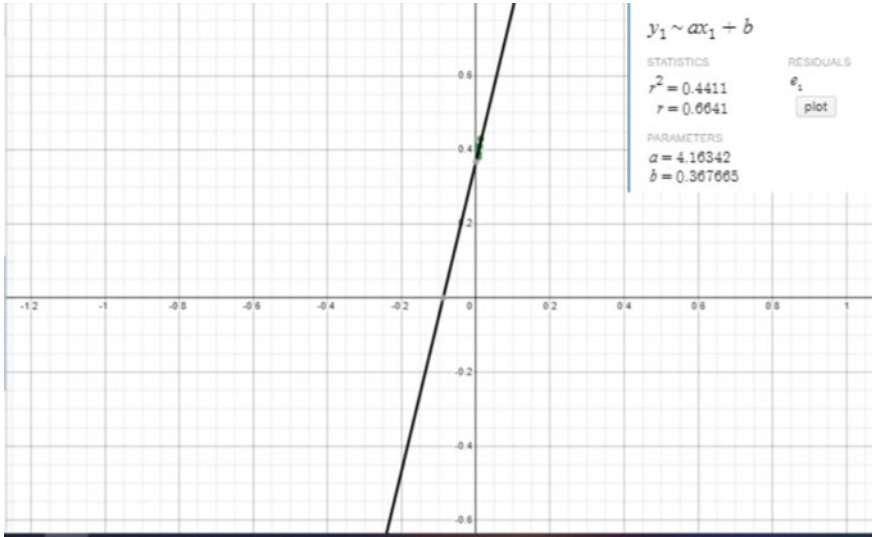


Fig. 5 Statistical method to determine individual modest and confident accuracy rates

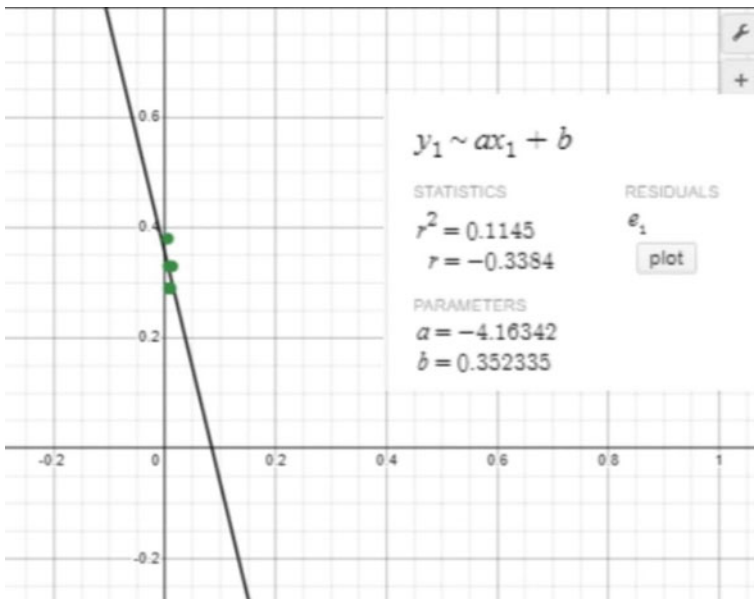


Fig. 6 Individual lethargy and modest by statistical method

Figure 1 shows the accuracy rate of only the complex person class. Following data analysis, we discovered that the complex category's intercept is 0.638, indicating that the chance of receiving signatures for the complex category for this class of signatures is high. Figure 4 depicts the data for the shy and stable groups. Even the slope is negative, the intercept is 0.74 limit of N tends to boundless value, which is an accuracy rate for the Shy and stable category. The accuracy rate for the modest and confident category is 0.37 in Fig. 5, and the accuracy rate for lethargy and modest is close to 0.35 in Fig. 6.

6 Conclusion

In this paper, we did a signature classification. We gathered a lot of signatures from various resources for the experiment. When the sample size was 204, we used a machine learning technique to construct a classification system and a neural network method to reach a 49% accuracy rate. To find individual importance of features, we done statistical method. Computed how to improve accuracy. We introduced a new parameter *alpha* in our method, which is found to be based on the overall number of signatures that went into distinct classes than an actual class. The amount of signatures from one class in a given class is nearly ± 0.173 . In a limit N tends to boundless value, then it leads to a higher accuracy rate. We chose the findings of our own perspective for our initial work on signatures. Generally, for individual features, we got a good accuracy more than 70%.

References

1. Chandra S, Maheskar S (2016) Offline signature verification based on geometric feature extraction using artificial neural network. In: 2016 3rd international conference on recent advances in information technology (RAIT), IEEE, pp 410–414
2. Chaudhari K, Thakkar A (2019) Survey on handwriting-based personality trait identification. *Expert Syst Appl* 124:282–308. <https://www.sciencedirect.com/science/article/pii/S0957417419300284>
3. Djamal EC, Darmawati R, Ramdhan SN (2013) Application image processing to predict personality based on structure of handwriting and signature. In: 2013 International Conference on Computer, Control, Informatics and Its Applications (IC3INA), pp 163–168
4. Hafemann LG, Sabourin R, Oliveira LS (2017) Learning features for offline handwritten signature verification using deep convolutional neural networks. *Pattern Recogn* 70:163–176. <https://www.sciencedirect.com/science/article/pii/S0031320317302017>
5. Inan Y, Sekeroglu B (2018) Signature recognition using backpropagation neural network. In: International conference on theory and applications of fuzzy systems and soft computing. Springer, Heidelberg, pp 256–261
6. Lokhande VR, Gawali BW (2017) Analysis of signature for the prediction of personality traits. In: 2017 1st international conference on intelligent systems and information management (ICISIM), IEEE, pp 44–49

7. Mohapatra RK, Shaswat K, Kedia S (2019) Offline handwritten signature verification using cnn inspired by inception v1 architecture. In: 2019 fifth International Conference on Image Information Processing (ICIIP), IEEE, pp 263–267
8. Rateria A, Agarwal S (2018) Off-line signature verification through machine learning. In: 2018 5th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON), pp 1–7. <https://doi.org/10.1109/UPCON.2018.8597090>
9. Ruiz V, Linares I, Sanchez A, Velez JF (2020) Off-line handwritten signature verification using compositional synthetic generation of signatures and siamese neural networks. *Neurocomputing* 374:30–41. <https://www.sciencedirect.com/science/article/pii/S0925231219313098>
10. Sronothara AB, Hanmandlu M (2018) Off-line signature verification using CNN. *Int J Future Revolution Comput Sci Commun Eng* 28(12):1407–1414
11. Varshney A, Puri S (2017) A survey on human personality identification on the basis of handwriting using ann. In: 2017 International Conference on Inventive Systems and Control (ICISC), pp 1–6
12. Wen Cheng C, Xiaopeng G, Hong S, Limin Z (2017) Offline Chinese signature verification based on alexnet. In: International conference on advanced hybrid information processing. Springer, Heidelberg, pp 33–37

Trajectory Tracking and Link Vibration Reduction of Flexible Manipulator in the Presence of Matched Uncertainty and External Disturbances Using Lyapunov-Based Controller



Sanjay Thakur, Ranjit Kumar Barai, and Anagha Bhattacharya

Abstract In this study, a Lyapunov-based controller (LBC) has been developed for the Two-Link Flexible Manipulator (TLFM) to reduce link vibration and track joint positions while dealing with model uncertainty and disturbances. An exponential trajectory has been considered the desired trajectory for both the joints. Since the controller has been designed using Lyapunov stability condition; therefore, it will guarantee the system stability. Model uncertainty and external disturbances have been taken into account. It has been considered that the equivalent viscous damping coefficient (EVDC) is uncertain. The system's state-dependent external disturbances have been considered. Proof of the matching condition of the considered model uncertainty has been provided. The lumped parameter method has been used to derive the system dynamics. The effectiveness of the proposed controller has been obtained by varying the uncertain parameter as mentioned in the simulation section.

Keywords Two-link flexible manipulator · Lyapunov-based controller · Exponential trajectory · Matching condition · Equivalent viscous damping coefficient

1 Introduction

Two-Link Flexible Manipulators (TLFMs) have many advantages like being light in weight, low power consumption can be operated at high operation speed, and longer links are possible. These features are not available in the case of rigid link manipulators. Therefore, flexible manipulators (FMs) have been highly used in industries, military operations, and medical research fields. In [1], the author introduced a sliding mode controller (SMC) for joint trajectory tracking and vibration suppression

S. Thakur (✉) · R. K. Barai

Department of Electrical Engineering, Jadavpur University, Kolkata, WB 700032, India
e-mail: sanjaythakur.ee.rs@jadavpuruniversity.in

A. Bhattacharya

National Institute of Technology, Aizawl, Mizoram, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
V. V. S. S. Chakravarthy et al. (eds.), *Advances in Signal Processing, Embedded Systems and IoT*, Lecture Notes in Electrical Engineering 992,
https://doi.org/10.1007/978-981-19-8865-3_49

543

in presence of uncertainty and external disturbances. Cao et al. [2] used the barrier Lyapunov function for designing a boundary controller for joint position tracking and link vibration suppression. In [3], the author decomposed the system into flexible and rigid dynamic subsystems and designed a fuzzy SMC composite controller for joint trajectory tracking and vibration suppression. Author in [4] designed a backstepping controller with a neural network for vibration reduction of TLFM in presence of model uncertainty. Ge et al. introduced LBC for the tip position tracking of the flexible manipulator. Optimization techniques like artificial bee colony optimization have been used by researchers [5, 6]. In [6], for position and vibration control, a very simple energy-based controller has been developed. The unknown parameters used in the controller have been tuned using the artificial bee colony optimization technique. For tuning the gains used in the design of LBC, the author used a genetic algorithm [7]. The author in [5] proposed an LBC only for vibration reduction in the presence of matched uncertainty. Siti et al. developed an adaptive PID controller based on the sliding mode approach for hub angular position control under the actuator faulty condition [8]. In [9], the author designed a non-singular fast terminal SMC for the trajectory tracking and vibration reduction of the TLFM. In [10], for the motion control and vibration suppression, PID controller has been introduced. Yiwei et al. [11] introduced the backstepping method for trajectory tracking and control for a flexible manipulator. A finite-time tracking has been obtained using the command filtering technique with the proposed controller. Wei et al. [12] used dynamic surface control for link position tracking of FM in presence of disturbances, uncertainty, and input saturation. In [13], Adaptive Model Predictive Controller (AMPC) along with integral sliding mode controller (ISMC) has been introduced for the vibration reduction and position tracking of the TLFM in the presence of matched uncertainty and external disturbances. In most of the above-mentioned articles, to fulfill the objective, either more than one controller has been used or unknown gains have been introduced with the controller. To obtain the proper value of the unknown gains, some tuning methods have been used. Using extra methods or controllers, put an extra burden on the actuators, which is a drawback. Therefore, in this work, single controller has been designed without any unknown gain terms.

Due to its flexible nature, vibration is a very common phenomenon. Along with vibration, FMs can have external disturbances and model uncertainty. It becomes a very challenging task for a controller to perform satisfactorily under such challenges and also ensure asymptotic stability and error convergence. To overcome such issues in this work, LBC has been introduced. Some notable contributions of this work are:

1. LBC has been designed using by considering the model uncertainties and error dynamics. The external disturbances have been added later with the system model. Instead of that, the controller can perform satisfactorily.
2. There are no unknown gain terms in the suggested controller. As a result, no further tuning process is required.
3. Matching condition of the model uncertainty has been presented.

The uncertain value has been modified to demonstrate the efficiency of the suggested controller. The obtained results have been compared with a published work [1] and found better.

At first, the mathematical model of the TLFM has obtained. EVDC has been identified as model uncertainty. The proof of the model uncertainty's matching condition has been provided. Then, the controller has been designed for the joint position tracking of each link, using Lyapunov stability theory and considering the error dynamics in presence of model uncertainty. Error dynamics are nothing but the difference between the actual position of the joints and the desired position of the joints. Closed-loop stability proof has been presented. External disturbances have not been considered while designing the controller. After that, the uncertain parameter's value has been changed, and different simulation results have been obtained. The entire process is presented in Fig. 4. The rest of the paper has been arranged as: Dynamic modeling has been shown in Sect. 2. Controller design and stability proof have been shown in Sect. 3. Conclusions have been presented in Sect. 4.

2 Mathematical Modeling and Problem Formulation

In this work, mathematical modeling has been obtained using lumped parameter method [1, 5].

$$\begin{aligned}\ddot{\Theta}_1 &= -\frac{b_{eq}}{j_{eq1}}\dot{\Theta}_1 + \frac{k_{stiff}}{j_{eq1}}\delta_1 + \frac{1}{j_{eq1}}\tau_1 + \Omega_1 \\ \ddot{\Theta}_2 &= -\frac{b_{eq}}{j_{eq2}}\dot{\Theta}_2 + \frac{k_{stiff}}{j_{eq2}}\delta_2 + \frac{1}{j_{eq2}}\tau_2 + \Omega_2,\end{aligned}\quad (1)$$

$$\begin{aligned}\ddot{\delta}_1 &= \frac{b_{eq}}{j_{eq1}}\dot{\Theta}_1 - k_{stiff}\left(\frac{1}{j_{eq1}} + \frac{1}{j_{link1}}\right)\delta_1 - \frac{1}{j_{eq1}}\tau_1 + \Omega_3 \\ \ddot{\delta}_2 &= \frac{b_{eq}}{j_{eq2}}\dot{\Theta}_2 - k_{stiff}\left(\frac{1}{j_{eq2}} + \frac{1}{j_{link2}}\right)\delta_2 - \frac{1}{j_{eq2}}\tau_2 + \Omega_4,\end{aligned}\quad (2)$$

where Θ is the joint angle, b_{eq} is the equivalent viscous damping coefficient, k_{stiff} is the link stiffness, δ is the link deflection, τ is the joint torque, j_{eq} is the equivalent moment of inertia, j_{link} is the link moment of inertia, and Ω is the external disturbance. The external disturbances externally added to the system have been given as

$$\begin{aligned}\Omega_1 &= 0.1 \times \sin(\pi\Theta_1) \sin(\pi\dot{\Theta}_1) \\ \Omega_2 &= 0.1 \times \sin(\pi\delta_1) \sin(\pi\dot{\delta}_1) \\ \Omega_3 &= 0.1 \times \sin(\pi\Theta_2) \sin(\pi\dot{\Theta}_2) \\ \Omega_4 &= 0.1 \times \sin(\pi\delta_2) \sin(\pi\dot{\delta}_2).\end{aligned}\quad (3)$$

2.1 Matching Condition

In this work, b_{eq} has been considered uncertain. Its value depends on the temperature of the surrounding, speed of operation, etc. [14]. Temperature and the speed of operation always keep on changing which makes the b_{eq} uncertain. Let, \tilde{b}_{eq} represents the nominal value of the uncertainty and b_{eq} is the actual value of the uncertainty. Without taking into account the terms of deflection, (1) can be rewritten as

$$\dot{X} = A(b_{eq})X + BU, \quad (4)$$

where $X = \begin{bmatrix} \dot{\Theta}_1 \\ \dot{\Theta}_2 \end{bmatrix}$, $A(b_{eq}) = \begin{bmatrix} -\frac{b_{eq}}{j_{eq1}} & 0 \\ 0 & -\frac{b_{eq}}{j_{eq2}} \end{bmatrix}$, $B = \begin{bmatrix} \frac{1}{j_{eq1}} & 0 \\ 0 & \frac{1}{j_{eq2}} \end{bmatrix}$, and $U = \begin{bmatrix} \tau_1 \\ \tau_2 \end{bmatrix}$.

For the nominal value of the uncertain parameter, (4) can be written as

$$\dot{X} = A(\tilde{b}_{eq})X + BU. \quad (5)$$

It can be observed that the matrix A alone contains the uncertain parameter. Therefore, uncertainty present in the matrix A can be given as

$$A(b_{eq}) - A(\tilde{b}_{eq}) = -(b_{eq} - \tilde{b}_{eq}) \begin{bmatrix} \frac{1}{j_{eq1}} & 0 \\ 0 & \frac{1}{j_{eq2}} \end{bmatrix}. \quad (6)$$

Equation (6) can also be written as

$$A(b_{eq}) - A(\tilde{b}_{eq}) = -(b_{eq} - \tilde{b}_{eq})B. \quad (7)$$

From (7), it can be observed that the uncertainty present in the matrix A is in the range of B . Hence, the uncertain parameter b_{eq} is satisfying the matching condition.

2.2 Error Dynamics

Let, Θ_{d1} and Θ_{d2} are the desired joint positions of the link1 and link2, respectively. Therefore, the error dynamics can be given as

$$\begin{aligned} e_1 &= \Theta_1 - \Theta_{d1} \\ e_2 &= \Theta_2 - \Theta_{d2}. \end{aligned} \quad (8)$$

3 Controller design (LBC)

LBC has been developed in this section, for trajectory tracking and vibration suppression of the TLFM in the presence of uncertainty. Using (1), (2), and (8), the expression of the controllers has been given as

$$\begin{aligned}\tau_1 &= j_{eq1}(\ddot{\Theta}_{d1} - e_1 - \Delta B_{eq}\dot{e}_1) + b_{eq}\dot{\Theta}_1 - k_{stiff}\delta_1 \\ \tau_2 &= j_{eq2}(\ddot{\Theta}_{d2} - e_2 - \Delta B_{eq}\dot{e}_2) + b_{eq}\dot{\Theta}_2 - k_{stiff}\delta_2,\end{aligned}\quad (9)$$

where $\Delta B_{eq} = (b_{eq} - \tilde{b}_{eq})$.

3.1 Proof of Stability

At first, the stability proof has been presented for the first joint torque (τ_1). Let, the Lyapunov candidate function can be written as

$$V = \frac{1}{2}e_1^2 + \frac{1}{2}\dot{e}_1^2. \quad (10)$$

Differentiating (11) and putting the respective parameter values, the following expressions have been obtained:

$$\begin{aligned}\dot{V} &= (e_1 + \ddot{e}_1)\dot{e}_1 \\ &= (e_1 + \ddot{\Theta}_1 - \ddot{\Theta}_{1d})\dot{e}_1 \\ &= -\Delta B_{eq}\dot{e}_1^2.\end{aligned}\quad (11)$$

From (11), it can be observed that if the necessary condition $\Delta B_{eq} > 0$ is satisfied, then it can be concluded that $\dot{V} \leq 0$, which is satisfying the Lyapunov condition. Similarly, the stability proof using τ_2 can be shown.

4 Simulation Results

The various simulation results have been obtained in this area by varying the b_{eq} as 0.5, 1.0, and 1.5. The initial value of the system's state variables has been taken as $(\Theta_1, \dot{\Theta}_1, \Theta_2, \dot{\Theta}_2) = (0, 0.1, 0, 0.1)$. An exponential desired trajectory has been considered ($\Theta_d = \frac{\pi}{6}e^{-t}$). The parameter values of the TLFM have been taken from Table 1. Different simulation results are already shown in Figs. 1, 2, and 3. The RK4 approach has been used to run the simulation in MATLAB for 20 s.

Table 1 Parameters of TLFM [1]

Parameter name	Values
k_{stiff}	6.4π Hz
\dot{j}_{eq1}	0.099 kg m^2
\dot{j}_{eq2}	0.092 kg m^2
\dot{j}_{link1}	0.00195 kg m^2
\dot{j}_{link2}	0.00933 kg m^2
b_{eq}	1.99

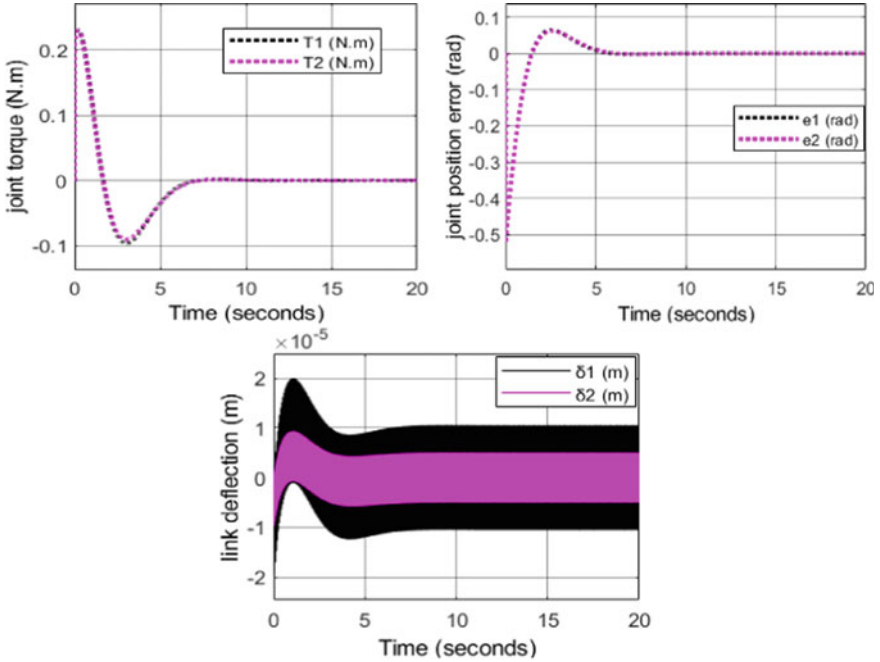


Fig. 1 Results obtained using LBC when $\tilde{b}_{eq} = 0.5$

Figure 1 describes the condition when $\tilde{b}_{eq} = 0.5$. The maximum torque required by the motor placed at the first joint ($\tau_{1\max}$) and second joint ($\tau_{2\max}$) is 0.2303 N.m and 0.2297 N.m , respectively. The error R.M.S of the joint position error has been obtained as $e_{1\text{rms}} = 7.082 \times 10^{-2} \text{ rad}$ and $e_{2\text{rms}} = 7.079 \times 10^{-2} \text{ rad}$. The maximum deflection produced by link1 and link2 has been obtained as $\delta_{1\max} = 2 \times 10^{-5} \text{ m}$ and $\delta_{2\max} = 9.617 \times 10^{-6} \text{ m}$, respectively. The uncertain parameter's value was then raised to 1.0 and 1.5. From the observation table (Table 2), it has been observed that the maximum torque needed for the joint rotors, the R.M.S error of the joint position tracking, and the maximum link deflection all grow as the uncertain parameter value increases.

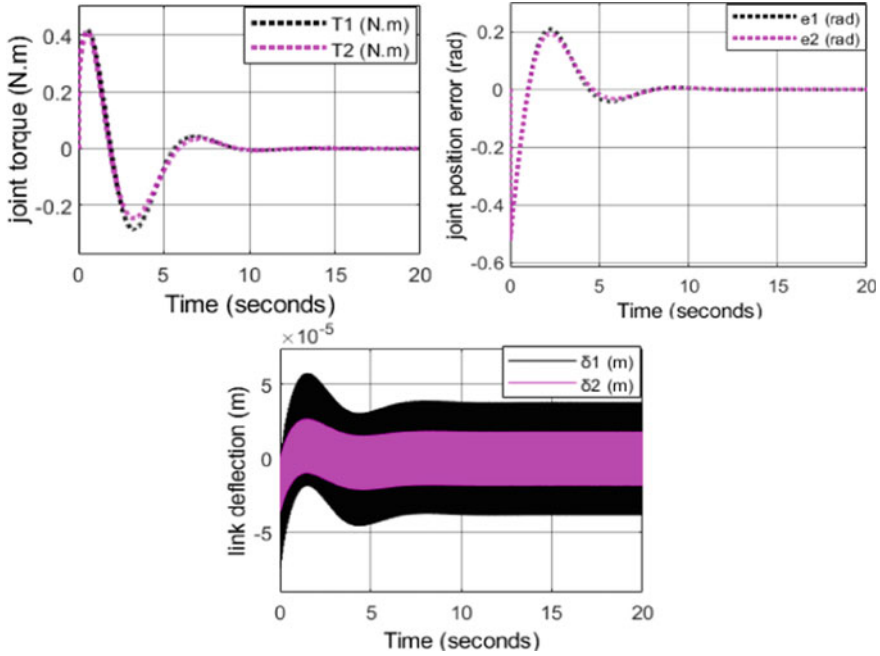


Fig. 2 Results obtained using LBC when $\tilde{b}_{eq} = 1.0$

5 Conclusions

In this work, LBC has been designed for the exponential trajectory tracking and vibration reduction of TLFM in presence of external disturbances and matched uncertainty. EVDC has been considered an uncertain parameter. Proof of the matching condition has also been presented in Sect. 2.1. While tracking the exponential trajectory, the maximum joint control torque requirement is very less approximately in the order of 10^{-1} N.m, which is very less compared to the joint motors' maximum torque requirement in [1] using the sliding mode controller (SMC). A lesser value of the controller indicates that the burden on the controller is very less and too high rating actuators are not required. Therefore, it reduces the cost of the design and power consumption. A clear comparison is presented in Table 3. Here, the LBC has been designed by only considering the joint position dynamics (Eq. (1)) and uncertain parameters, and still, it can reduce the vibration of the links efficiently along with externally added disturbance. The vibration of the links has been decreased approximately in the order of 10^{-5} m. R.M.S errors of the joint position tracking have also been reduced significantly (presented in Table 2). It can be concluded that the designed LBC is alone sufficient in minimizing the vibration of the links and can track the desired joint position. Therefore the designed LBC is better compared to

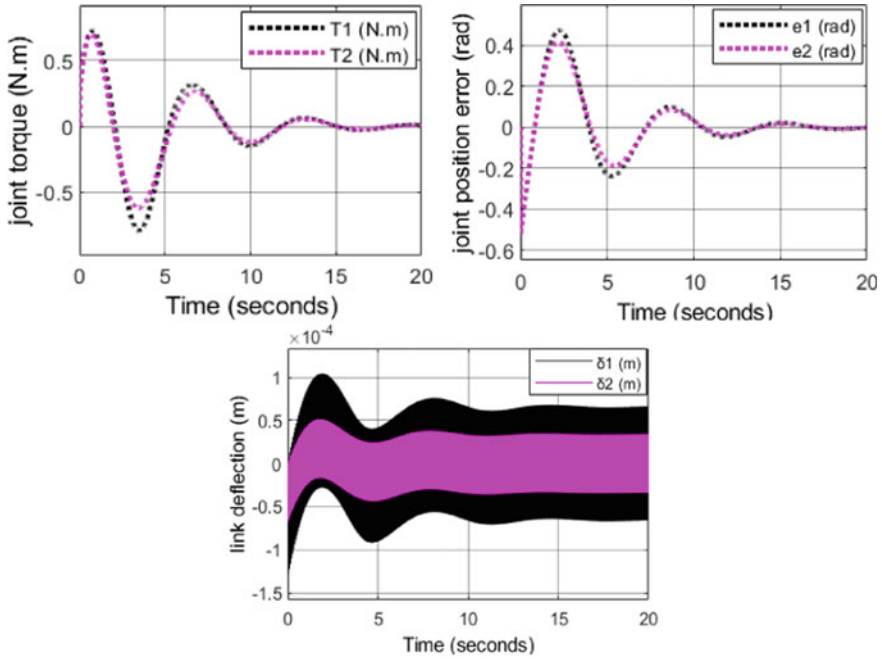


Fig. 3 Results obtained using LBC when $\tilde{b}_{eq} = 1.5$

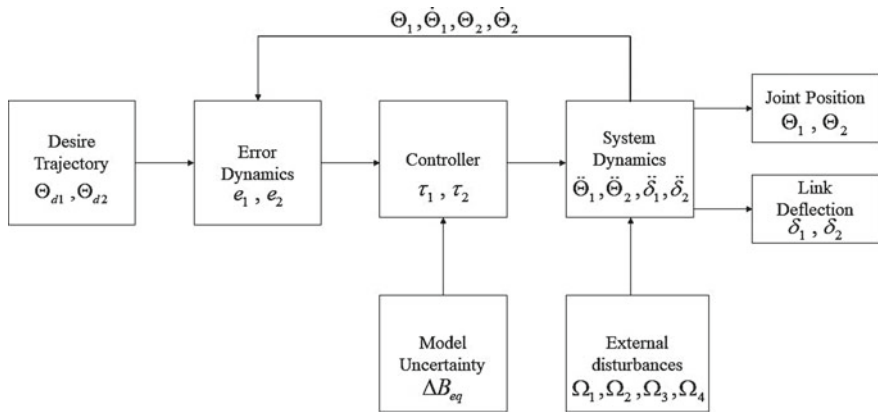


Fig. 4 Work flow diagram of TLFM

the SMC designed in [1]. The design of various types of controllers, as well as their implementation in real-time, is preserved as a future extension.

Table 2 Observation table

Uncertain parameter	Maximum torque required (N.m)		R.M.S joint position tracking error (rad)		Maximum link deflection (m)	
	$\tau_{1 \max}$	$\tau_{2 \max}$	e_{1rms}	e_{2rms}	$\delta_{1 \max}$	$\delta_{2 \max}$
0.5	2.303×10^{-1}	2.297×10^{-1}	7.082×10^{-2}	7.079×10^{-2}	2×10^{-5}	9.617×10^{-6}
1.0	4.184×10^{-1}	4.122×10^{-1}	8.752×10^{-2}	8.487×10^{-2}	7.398×10^{-5}	3.578×10^{-5}
1.5	7.939×10^{-1}	6.231×10^{-1}	1.628×10^{-1}	1.444×10^{-1}	1.050×10^{-4}	4.787×10^{-5}

Table 3 Maximum torque required by the joint actuators

Uncertain parameter	Using designed LBC		Using SMC in [1]	
	$\tau_{1 \max}$ (N.m)	$\tau_{2 \max}$ (N.m)	$\tau_{1 \max}$ (N.m)	$\tau_{2 \max}$ (N.m)
0.5	2.303×10^{-1}	2.297×10^{-1}	2.172	2.572
1.0	4.184×10^{-1}	4.122×10^{-1}	4.012	4.355
1.5	7.939×10^{-1}	6.231×10^{-1}	5.434	5.797

Acknowledgements The author is grateful for the financing support provided by the AICTE-NDF (India).

References

1. Thakur S, Barai RK (2021) Joint trajectory tracking of two-link flexible manipulator in presence of matched uncertainty. In: IEEE international conference on distributed computing, VLSI, electrical circuits and robotics (DISCOVER), pp 151–154
2. Cao F, Liu J (2017) Vibration control for a rigid-flexible manipulator with full state constraints via Barrier Lyapunov Function. *J Sound Vib* 406:237–252
3. Huang H, Tang G, Chen H, Han L, Xie D (2022) Dynamic modeling and vibration suppression for two-link underwater flexible manipulators. *IEEE Access* 10:40181–40196
4. He W, Kang F, Kong L, Feng Y, Cheng G, Sun C (2017) Vibration control of a constrained two-link flexible robotic manipulator with fixed-time convergence. *IEEE Trans Cybern*
5. Thakur S (2021) Vibration reduction of flexible manipulator using lyapunov based controller in presence of matched uncertainty. *Webology* 18(4):420–426
6. Eser S, Çetin ST (2022) Optimum control of a flexible single link manipulator with Artificial Bee Colony Algorithm. In: Proceedings of the institution of mechanical engineers, Part C: *J Mech Eng Sci* 236(7):3731–3742
7. Ge SS, Lee TH, Zhu G (1996) Genetic algorithm tuning of Lyapunov-based controllers: an application to a single-link flexible robot system. *IEEE Trans Ind Electron* 43(5):567–574
8. Abd Latip SF, Rashid Husain A, Mohamed Z, Mohd Basri MA (2019) Adaptive PID actuator fault tolerant control of single-link flexible manipulator. *Trans Inst Measur Control* 41(4):1019–1031

9. Lei RH, Chen L (2021) Finite-time tracking control and vibration suppression based on the concept of virtual control force for flexible two-link space robot. *Defence Technol* 17(3):874–883
10. Mahamood RM, Aweda JO, Ajao KR, Popoola OT, Odusote JK (2015) PID controller design for two link flexible manipulator. *LAUTECH J Eng Technol* 9(2):17–22
11. Zhang Y, Zhang M, Fan C, Li F (2022) A finite-time trajectory-tracking method for state-constrained flexible manipulators based on improved back-stepping control. *Actuators* 11(5):139. MDPI.
12. Yao W, Guo Y, Wu YF, Guo J (2022) Robust adaptive dynamic surface control of multi-link flexible joint manipulator with input saturation. *Int J Control Automat Syst* 20(2):577–588
13. Bakhti M, Hannane A, Bououlid Idrissi B (2022) Robustifying adaptive model predictive control for a one-link flexible manipulator using super-twisting integral sliding mode control. *Int J Dyn Control*, pp 1–9
14. Aloisio A, Alaggio R, Fragiaco M (2021) Equivalent viscous damping of cross-laminated timber structural archetypes. *J Struct Eng* 147(4):04021012–04021013

Multimodal Approach for Code-Mixed Speech Sentiment Classification



S. Keshav, G. Jyothish Lal, and B. Premjith

Abstract Sentiment analysis is a natural language processing (NLP) technique used to classify a statement into three polarities, namely, positive negative and neutral. Thus, speech sentiment analysis invariably becomes a NLP task rather than a direct speech processing task, making the available speech recognition model just a tool to carry out NLP. Nevertheless, we do not have a way to directly use the pre-trained speech models to carry out sentiment analysis on speech utterances. The present study proposes to achieve this by directly using the speech signal to carry out sentiment analysis. We evaluate two such approaches on our custom made dataset consisting of movie and political reviews. The first approach used a fully connected neural network (FCNN) model and the second one used a 3-shot few shot learning (FSL) framework. For the FCNN model, the proposed framework provides a classification accuracy of 61.53%, whereas we get an accuracy of 99.83% for the 3-shot FSL framework. The performance is comparable to the current state-of-the-art (SOTA) system in place.

Keywords Few shot learning · Code-mixed data · Speech sentiment · Sentiment classification · Multimodal approach · Hinglish · FCNN · Wav2vec2 · BERT · ResNet

1 Introduction

At present, almost all successful speech processing models work on the basis of conversion of speech-to-text (STT), and then natural language processing (NLP) is carried out on the text for either classification or generation. They are definitely successful as proven by the high level of accuracy achieved at various levels in which they are implemented, but these supervised models are data intensive while training.

S. Keshav (✉) · G. Jyothish Lal · B. Premjith
Center for Computational Engineering and Networking (CEN), Amrita School of Engineering,
Amrita Vishwa Vidyapeetham, Coimbatore, India
e-mail: keshavsainath10@gmail.com

G. Jyothish Lal
e-mail: g_jyothishlal@cb.amrita.edu

Moreover, they work on monolingual tasks, that is, they are not very adaptable to multilingual tasks if they don't have data available for training the model. The present study is motivated by this issue of having low resource data and tries to bring an unsupervised approach to resolve the issue. In the next section, a brief review of the related literature toward sentiment analysis is given.

1.1 Related Literature Review

Text-based sentiment analysis has been a very popular approach to carry out sentiment classification with high prediction accuracy with availability of countless pre-processing tools [1] and the contributing factor to its success is the wide range of availability of content for NLP models from social media [2]. Many machine learning models [3] exist to classify sentiments, but to make the prediction even more accurate, aspect-based deep learning (DL) models [4] exist that are able to capture the true sentiment without being confused due to multiple polarities present in the statement. Recently, we are exposed to speech-based content on a regular basis in the form of audio message forwards or audio-visual content. In the near future, speech sentiment analysis [5] will become a necessity to filter out the contents which earlier was applicable only to textual data [6] and will not only capture emotions but the polarity of the speech content. Voice recognition software [7] will be able to filter the right sentiment and deliver the content you need to hear. Conversion of speech-to-text (STT) will not be needed for sentiment analysis [8]. There is a gap at present that utilizes NLP to achieve these tasks making the models at present complex and data-hungry [9]. The present work is a novel approach to make the sentiment classification more human-like and efficient.

1.2 Formulation of the Study

To begin with, transformer-based pre-trained models, BERT and wav2vec2's XLSR [10] in this case, are used to obtain NLP and speech embedding respectively. We use these embedding as a dataset to train a simple FCNN model to carry out the task. Embeddings chosen are generated by the *Feature extractor* layer of the transformers as embeddings generated by any of the layers of a transformer do not affect the task at hand by a huge degree [11]. To follow up this process and in order to improve the classification, few shot learning (FSL) is chosen. Learning to learn is what humans do in order to find similarities and differences between objects in real time, this was the motivation behind choosing FSL. The idea of using this human concept into machines is what makes few shot learning a success when handling small samples as real world data also generally consists of very little resource to learn. FSL is successfully tested to give a high level of accuracy for image [12] classification for the omniglot dataset and mini-imagenet datasets. Transfer learning has been abundantly used in the present

study as the goal of the study was to classify low resource code-mix content-based speech signals for sentiment analysis. These models are evaluated on a custom dataset created with the intent of testing the classification capability of the model solely on the basis of speech signals rather than following the conventional approach.

The rest of the paper is organized as follows: Section 2 explains the proposed approach where the architecture of both FCNN and FSL are presented. Section 3 introduces the dataset used and pre-processing steps applied on it. Section 4 discusses the results obtained on proposed approaches, and finally, Sect. 5 concludes the paper.

2 Proposed Approach

Present study is carried out in two phases. In phase I, we train a FCNN using pre-trained wav2vec2 and BERT embedding generated by the speech signal and its transcript, and after the training, testing and validation is carried out using the testing set to see the accuracy of the model. One important aspect to keep in mind is that the embedding generated by BERT and wav2vec2 are of different shapes, so while we try to combine them, there are certain challenges that need to be dealt with so that the FCNN can take the embedding as input and carry out the task.

2.1 Architecture for FCNN

Using Wav2vec2 embedding

The reason for choosing the model in Fig. 1 is to check the accuracy of the model using only the speech embedding, as the focus in this study is speech signal. The embedding obtained from wav2vec2 is split into test and train sets. Initially, we train the FCNN with the wav2vec2 embedding as input and then using the weights of the trained FCNN rebuild another FCNN model to which we pass the test set in order to obtain the classified sentiment as the output.

Using BERT and Wav2vec2 embedding in combination

We chose the model in Fig. 1, so that we can see the effect of training the model as a combination of BERT embedding and wav2vec2 embedding of the dataset. This is achieved in two ways, first, by up-scaling the wav2vec2 embedding to the shape of the NLP embedding or, second, by down-scaling the NLP embedding to the shape of the wav2vec2 embedding. This is followed by the same steps mentioned above.

Using a projected model

Just to make sure that we have not left any space for suspicion regarding the loss of information while down-scaling or up-scaling the dataset, we project both the wav2vec2 embedding shape and the BERT embedding shape to $(x, 1000)$ and then carry out the same task as mentioned above.

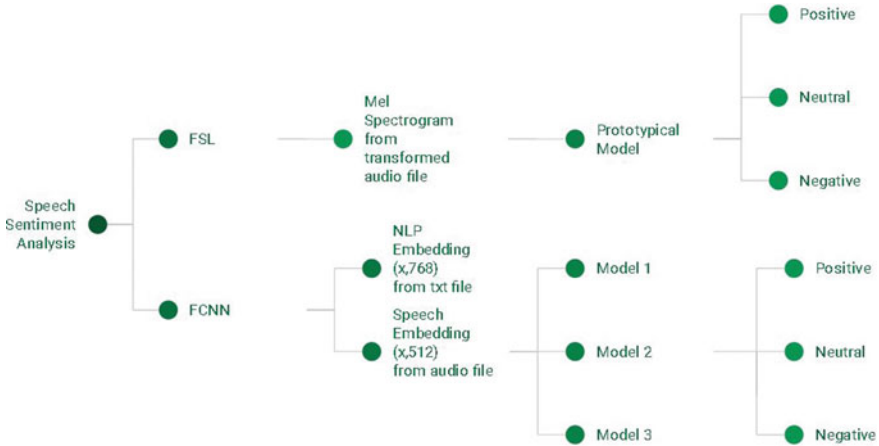


Fig. 1 Flow diagram representing the multi-modal approach

2.2 Architecture for FSL

Using Wav2vec2 embedding

This model is chosen to use the wav2vec2 embedding as the input to train the prototypical model [12] with various pre-trained torchaudio models as its input in order to check its capability to give us the accuracy of the predictions made by learning the scores which is computed using cosine distance calculated between the support set and the query set.

Using Mel

Here, the reason for choosing the model in Fig. 1 is to obtain an image-like [13] input for the prototypical model as it takes in a 3-channel RGB image as an input for training the model and predicting the sentiment. The choice for mel spectrogram instead of MFCC is just because of the fact that MFCC is more accurate for predicting the tonal textures of the input like the pitch information, whereas mel spectrogram succeeds in mimicking the working of a human ear [14].

3 Dataset Description and Pre-processing

A custom dataset is used to carry out this study. The dataset consists of 14 youtube videos which contain content related to movie and political review in Hinglish (Hindi + English). These videos were chosen considering the fact that they contain statements of the three polarities, namely positive, negative and neutral. Audio extracted [13] from these videos are separated into single sentence speech files (the range of

length of each file varies from 2 to 28s). 898 files are obtained in total. Each of these speech files was manually transcribed for extracting NLP embeddings [15]. The pre-trained model used for generating the embeddings is trained on multiple languages which includes the target language, i.e., Hindi and English. The dataset was manually verified for this project.

3.1 Pre-processing

Pre-processing for FCNN

The features extracted from the speech signal using pre-trained wav2vec2 XLSR53 [10] model is of the shape $(x, 512)$ and for that of pre-trained BERT is $(x, 768)$ where x refers to the number of speech files in the dataset. There are three FCNN models in total, the first one accepts only speech embedding of shape $(x, 512)$. The second model taken in a combination of the embedding which is achieved by

Down-scaling Down-scaling the BERT embedding of shape $(x, 768)$ to $(x, 512)$, matching them to wav2vec2 embedding and then performing the add function to get a final embedding shape of $(x, 512)$.

Up-scaling Up-scaling the wav2vec2 embedding of shape $(x, 512)$ to $(x, 768)$, matching them to wav2vec2 embedding and then performing the add function to get a final embedding shape of $(x, 768)$.

The third model in this section basically uses random projection to upscale both the wav2vec2 embedding and BERT embedding to $(x, 1000)$.

Pre-processing for FSL

The prototypical model used for FSL uses a pre-trained model as its input and is designed to accept the data input in the form of 3-channel images. For the FSL model to take the dataset as input, certain transformations need to be performed as depicted in Fig. 1.

- T1** Re-sampling all the files 22,050 Hz.
- T2** Converting all the tracks to a single channel (mono) instead of stereo to maintain uniformity.
- T3** Finding the track with maximum length and converting it to 22,050 samples.
- T4** Right padding the tracks that are shorter in terms of duration than the longest track.
- T5** Applying mel spectrogram transform with the parameters set to $n_mels = 64$, $n_fft = 1024$ and $hop_length = 512$.

A 64-dimension speech descriptor is extracted using the Mel Spectrogram transformation function from the Torchaudio toolkit. The tensor obtained finally is a single channel tensor which now needs to be converted to three channels.

- T6** Copying the single-channel information to the other two channels to make the input a 3-channel input.
- T7** Normalizing the above tensor to maintain consistency in the dynamic range of the input data.

3.2 *Pre-trained Models*

For the present study, we have relied heavily on the use of pre-trained models from huggingface [16] and pytorch [17] library to carry out classification tasks at various stages.

FCNN

Here, we have used numerous pre-trained models from *huggingface*¹ in order to create the dataset. Both BERT and wav2vec2 pre-trained models have helped to generate the desired embedding for the sentiment classification task.

FSL

In this case, we used the Pytorch pre-trained models not to generate a dataset but to use the forward function of the pre-trained model for generating the embeddings for the support set and query set to compute scores taken from Pytorch.² Another important role of these pre-trained models is for calculating the loss function while training and help in improving the accuracy of the model. These pre-trained models include resnet18, resnet34, resnet50, wide_resnet50_2, resnext50_32x4d, vgg16, vgg11_bn, googlenet, densenet161 and squeezenet1_0.

4 Experimental Results and Discussion

In this section, we discuss the experimental setup and its corresponding result in terms of identifying the speech sentiment.

4.1 *FCNN*

One important aspect in the present project is that the FCNN model parameters are constant throughout the classification process, the only change that is applied is in the feature generation process to obtain the dataset from the pre-trained models which is the input for the FCNN.

¹ <https://huggingface.co/models>.

² <https://pytorch.org/vision/stable/models.html>.

Table 1 Accuracy of embedding generated from different wav2vec2 pre-trained models

Embedding from pre-trained model	# of epochs	Accuracy
Harveenhadha/vakyansh-wav2vec2-hindi-him-4200 (M1)	400	59.61
skylord/wav2vec2-large-xlsr-hindi (M2)	400	60.43
theainerd/Wav2Vec2-large-xlsr-hindi (M3)	400	61.53
tanmaylaud/wav2vec2-large-xlsr-hindi-marathi (M4)	400	58.65
shiwangi27/wave2vec2-large-xlsr-hindi (M5)	400	60.57

Table 2 Accuracy of combined embedding generated from different BERT pre-trained models and M3

NLP+speech pre-trained model	# of epochs	Up-scaling acc	Down-scaling acc	Train acc–val acc
monsoon-nlp/muril-adapted-local (N1)	200	0.5865	0.5961	0.7215–0.5962
rohanrajpal/bert-base-multilingual-codemixed-cased-sentiment (N2)	200	0.6346	0.2980	0.7627–0.5096
ai4bharat/indic-bert (N3)	200	0.2596	0.2019	0.8765–0.5000
setu4993/LaBSE (N4)	200	0.3846	0.2788	0.7772–0.6154

Wav2vec2 Pre-trained Model

Here, we have used five different wav2vec2 XLSR53 pre-trained models to generate the embedding for the 898 speech files. We evaluate the performance of the dataset generated using the proposed models. In this case, the model is a simple FCNN which is rebuilt from training a FCNN using the embedding of shape $(x, 512)$.

As given in Table 1, the performance of embeddings generated by M3, i.e., the wav2vec2 XLSR53 model *theainerd/Wav2Vec2-large-xlsr-hindi* outperforms all the other chosen pre-trained models by achieving a 61.53% accuracy.

Combination of BERT and Wav2vec2

In the case of combination, we consider two ways of combining the embedding before it is sent for training and testing. The first way is to downscale the BERT embedding from $(x, 768)$ to $(x, 512)$. This is done so as to get the final embedding shape as $(x, 512)$. Here, we are not concatenating the embedding but using the Add() to achieve an embedding that has the information of both BERT and wav2vec2.

As given in Table 2, the performance of embeddings generated by the combination of M3 along with N1, i.e., the BERT model *monsoon-nlp/muril-adapted-local* outperforms all the other chosen pre-trained models by achieving a 58.65% (up-scaled) and 59.62% (down-scaled) accuracy and also maintaining a relatively high train and validation accuracy of 72.15% and 59.62%, respectively.

Table 3 Accuracy of projected embedding generated from M3 and N1

NLP+speech pre-trained model	# of epochs	Up-scaled PM (768)	Down-scaled PM (512)	Scaled PM (1000)
M3+N1	200	0.5673	0.5384	0.5865

Projected Model

We are using a similar approach as the previous model, the only difference is that instead of down-scaling or up-scaling either of the embeddings, both are up-scaled using random projection to $(x,1000)$ by setting the random state parameter to 42. This remains common throughout the study just to maintain uniformity.

Table 3 just proves the fact that we can directly use wav2vec2 embedding and get a better result compared to the combination of BERT and wav2vec2 embeddings.

4.2 FSL

Since we have three classes in our dataset, i.e., *positive*, *negative* and *neutral*, we choose the 3-way n-shot system and in order to get the best performance from the model, we perform the task with $n = 1, 2, 3$.

Checking with Embeddings as the input

This particular part was unsuccessful but an eye-opener at the same time. We could not get the FSL to work no matter what we tried only to realize the following:

- 1 The input needed to be 3-channel
- 2 The pre-trained model from Pytorch had 2D convolution layers but the embeddings generated from wav2vec2 had 1D convolution layers
- 3 Mis-match in the masking length and sequence length.

All these issues resulted in the failure of this approach of using embedding as an input into FSL model but gave an insight into the working of the Prototypical model

Checking with various Pre-trained Models

To make this model work, now a new input was put in place, which was the mel spectrogram of each speech file. As seen in Table 4, we can easily see that the resnet pre-trained model gives the most accurate classification of the sentiment and surpasses the results of other models by a huge margin.

Checking with Resnet Model variants

Next objective is to check if the other variants of Resnet pre-trained models have a better advantage than the resnet18 pre-trained model. We repeat the same process that we followed in the previous step to get the result given in Table 5.

Table 4 FSL for different Pytorch pre-trained models

Pre-trained model	3-way 1-shot	3-way 2-shot	3-way 3-shot
resnet18	50.74	96.68	98.34
vgg16	33.13	33.33	33.32
vgg11_bn	33.13	34.33	32.82
googlenet	34.24	34.97	42.67
densenet161	33.11	37.30	40.78
squeezenet1_0	33.33	33.33	33.33

Table 5 FSL for different resnet Pytorch pre-trained models

Pre-trained model	3-way 1-shot	3-way 2-shot	3-way 3-shot
resnet18	50.74	96.68	98.34
wide_resnet50_2	35.52	34.14	41.12
resnext50_32x4d	35.61	52.58	89.51
resnet34	32.98	33.36	44.28
resnet50	70.84	95.56	99.83

Table 5 indicates that resnet50 is the best option among the resnet pre-trained models probably because it is effectively able to carry out episodic learning with minimum over-fitting while assigning scores. This result is indicative of the fact that FSL can be used for speech sentiment analysis and that it can produce high accuracy with low resource speech content without using the traditional approach.

5 Conclusion and Future Scope

The present work is focused to obtain a high level of accuracy for classification of speech sentiment of low resource code-mix audio content. We were able to achieve an accuracy of 99.83% for a 3-way 3-shot FSL model using the custom made dataset. Though the study managed to get a very high accuracy for the dataset, the entire process is still a black-box and we don't know exactly what the model learnt. The lack of pre-trained of 1D models and the lack of in-depth understanding of wav2vec2 models in terms of the masking and sequencing features made the use of embedding for FSL impossible at this point but with more effort, ways can be developed to solve this issue. A huge difference in the result between FSL and FCNN is very evident with this experiment, and one of the main reasons is because of the fact that FSL learns to predict the similarity or the differences between the input embeddings, whereas the FCNN does a linear comparison between what it has learnt to what is provided to it which may lead to immediate rejection even incase of a small difference. This present study can be stretched in many directions, not only to improve the prediction

capability in terms of discriminative models but also in terms of generative models, given the fact that these models could possibly be lightweight in terms of their application. Since, this is just the introduction to code-mix, and its implementation is still at a nascent stage; a published dataset can be used with many more data points to improve the quality of the model as such.

References

1. Haddi E, Liu X, Shi Y (2013) The role of text pre-processing in sentiment analysis. *Procedia Comput Sci* 17:26–32
2. Tulasi Sasidhar T, Premjith B, Sreelakshmi K, Soman K (2021) Sentiment analysis on Hindi-English code-mixed social media text. *Innov Comput Sci Eng*, pp 615–622
3. Kumar S, Premjith B, Kumar M, Soman K (2015) AMRITA_CEN-NLP@ SAIL2015: sentiment analysis in Indian Language using regularized least square approach with randomized feature learning. In: International conference on mining intelligence and knowledge exploration, pp 671–683
4. Sai Aparna T, Simran K, Premjith B, Soman K (2021) Aspect-based sentiment analysis in Hindi: comparison of machine/deep learning algorithms. *Inventive Comput Inf Technol*, pp 81–91
5. Maghilnan S, Kumar M (2017) Sentiment analysis on speaker specific speech data. In: 2017 International conference on intelligent computing and control (I2C2), pp 1–5
6. Patra B, Das D, Das A (2018) Sentiment analysis of code-mixed indian languages: an overview of sail_code-mixed shared task@ icon-2017. *ArXiv Preprint [ArXiv:1803.06745](https://arxiv.org/abs/1803.06745)*
7. Lu Z, Cao L, Zhang Y, Chiu C, Fan J (2020) Speech sentiment analysis via pre-trained features from end-to-end ASR models. In: ICASSP 2020-2020 IEEE international conference on acoustics, speech and signal processing (ICASSP)
8. Ezzat S, El Gayar N, Ghanem M (2012) Sentiment analysis of call centre audio conversations using text classification. *Int J Comput Inf Syst Ind Manage Appl* 4:619–627
9. Pravalika A, Oza V, Meghana N, Kamath S (2017) Domain-specific sentiment analysis approaches for code-mixed social network data. In: 2017 8th International conference on computing, communication and networking technologies (ICCCNT), pp 1–6
10. Conneau A, Baevski A, Collobert R, Mohamed A, Auli M (2020) Unsupervised cross-lingual representation learning for speech recognition. *ArXiv Preprint [ArXiv:2006.13979](https://arxiv.org/abs/2006.13979)*
11. Shah J, Singla Y, Chen C, Shah R (2021) What all do audio transformer models hear? probing acoustic representations for language delivery and its structure. *ArXiv Preprint [ArXiv:2101.00387](https://arxiv.org/abs/2101.00387)*
12. Laenen S, Bertinetto L (2021) On episodes, prototypical networks, and few-shot learning. *Adv Neural Inf Process Syst* 34
13. Bui K, Oh H, Yi H (2020) Traffic density classification using sound datasets: an empirical study on traffic flow at asymmetric roads. *IEEE Access* 8:125671–125679
14. Mittal A, Bharadwaj S, Khare S, Chemmengath S, Sankaranarayanan K, Kingsbury B (2021) Representation based meta-learning for few-shot spoken intent recognition. *ArXiv Preprint [ArXiv:2106.15238](https://arxiv.org/abs/2106.15238)*
15. Pratapa A, Choudhury M, Sitaram S (2018) Word embeddings for code-mixed language processing. In: Proceedings of the 2018 conference on empirical methods in natural language processing, pp 3067–3072

16. Wolf T, Debut L, Sanh V, Chaumond J, Delangue C, Moi A, Cistac P, Rault T, Louf R, Funtowicz M, Others (2019) Huggingface's transformers: state-of-the-art natural language processing. ArXiv Preprint [ArXiv:1910.03771](https://arxiv.org/abs/1910.03771)
17. Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, Killeen T, Lin Z, Gimelshein N, Antiga L, Others (2019) Pytorch: an imperative style, high-performance deep learning library. Adv Neural Inf Process Syst, 32

Despeckling of Ultrasound Imagery with Qualitative Filtering Techniques



Satwinder Kaur, Bhawna Goyal, and Ayush Dogra

Abstract Sound waves' usage gives rise to ultrasound imagery as it is an indispensable tool for identification and diagnosis of the inner structure of the body for the detection of eventual sickness or anomalous tissues. Ultrasound imaging is a secure, reliable and economical diagnostic reading test. However, ultrasound imagery has an intrinsic property, i.e., speckle noise which diminishes the contrast and resolution of any imagery as it abolishes adequate details or boundaries which are mandatory to be well-preserved. In recent times, many researchers have contemplated distinct filtering schemes for mitigating noisy speckles to achieve good estimation in biomedical operations which include filtration schemes that transform the multiplicative noise into additive with the usage of algorithmic operations. This paper examines and compiles several schemes that are mostly utilized for the eradication of speckling from ultrasound imageries, and a comparison has been done for all the procedures studied based on experiments to illustrate the advantages of every technique. Since genuine ultrasound imageries are already degraded and actual noiseless imageries do not exist. So, synthetic imageries have been formed for testing the procedures. Nevertheless, the scheme has been offered for despeckling in real imageries.

Keywords Speckle noise · Noise removal/reduction · Ultrasound imageries · Despeckling

1 Introduction

There is a dire need for timely detection and examination of diseases which constantly pushes the improvements of imaging modalities. Medical imagery which includes

S. Kaur (✉) · B. Goyal
Department of Electronics & Communications Engineering, Chandigarh University, Mohali, India
e-mail: erskaur@gmail.com

A. Dogra
Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India

Ronin Institute, Montclair, NJ 07043, USA

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
V. V. S. S. Chakravarthy et al. (eds.), *Advances in Signal Processing, Embedded Systems and IoT*, Lecture Notes in Electrical Engineering 992,
https://doi.org/10.1007/978-981-19-8865-3_51

565

MRI, ultrasound imaging, CT scan, X-rays, etc., is a very valuable tool for the clinical diagnosis and study of different types of diseases. Ultrasound imageries are safe, secure, portable, reliable and widespread medical care used for biomedical examination such as cardiopathy, urology, gynecology, obstetrics and fetology [1]. Its key parameters are its non-invasive character, the nonexistence of ionization radiation and its real-time estimation for human tissues, muscles, blood vessels, joints, the existence of cysts and organs like liver, kidneys, stomach, gall bladder, uterus, ovaries and so on [2]. It is also used for the inspection of the behavior of pregnant ladies in a safe and real-time manner.

Ultrasound imageries are intrinsically of inferior pictorial quality because of fluctuating acoustic along with temporal variations [3, 4]. The pictorial quality of these imageries is additionally dropped by the existence of several types of noises, i.e., impulsive noise, amplifier noise/Gaussian noise and speckle noise. Gaussian noise is an amplifier noise that takes place thermally or electronically in electronic circuits.

The other major type of noise that prevails in ultrasound imageries is speckling noise which is a granular type of noise that appears due to constructive along with destructive intrusion among the temporal coherent ultrasonic waves in the device. Speckling noise has the specific characteristic of multiplicative noise as well as Rayleigh distribution which disrupts the resolution along with the contrast of imagery [5]. Speckle noise in biomedical ultrasound imageries inhibits physicians from the accomplishment of perfect analysis as they obstruct the extraction, examination as well as identify the tissues [6]. To obtain an accurate diagnosis of ultrasonic imaging, a good despeckling technique is a crucial preprocessing constraint.

1.1 Speckle Modeling

The word ‘‘Speckle’’ itself means little dots just as natural color marks on the skin. Therefore, speckle noise describes by a random production of several little dots in the digital image. As the speckle noise is in multiplicative nature [8], some researchers establish the speckle noise in mathematical behavior. The final signal comprises the echo signal which is perceived by the reception instrument for ultrasound image formation, and it is demonstrated as the reflection of the ultrasonic waves that are inserted into the inside of the body [9]. The recovered echo signal is composed of two factors, in which one is the valuable signal (reflected from the body) and the other one is multiplicative [10] along with additive noise [11]. The mathematically modeling behavior of speckle noise demonstrates as

$$u(x, y) = m(x, y)n(x, y) + l(x, y), \quad (1)$$

where $n(x, y)$ and $l(x, y)$ show the mechanisms of multiplicative along with additive noise, respectively. (x, y) shows the spatial coordinates of two dimensions. $M(x, y)$ displays the original input signal, and $u(x, y)$ displays the detected signal. Due to the less influence of additive noise on ECG, Eq. (1) can be intended by (2).

$$U(x, y) = m(x, y)n(x, y). \quad (2)$$

2 Despeckling Methodologies

2.1 Several Filtration Schemes

To mitigate the noises is a cumbersome task that has been always a research topic for researchers while conserving the edges. The list of schemes that are conferred below is proposed to give the reader a view of the diversity of prevailing procedures.

2.1.1 Median Filtering Scheme

The median scheme of filtration is a nonlinear procedure that replaces the intensity mean of the imagery with the median when the spatial noise dispersion is asymmetrical among the window matrix. This scheme suppresses the variance among the pixel intensities within the imagery and it retains the edges even though it eradicates noise in a big amount. However, it is a robust filtering scheme that works on a pixel by pixel for substituting the entries with a median of nearby entries [12]. This filtering scheme is usually utilized for the eradication of speckle because it considers only that particular area in which there are any impulsive artifacts present.

2.1.2 Fourier Filtration Scheme

This name of the scheme automatically dictates that it utilizes the properties of the Fourier and it is the most widely used transform as it provides information about the frequency–amplitude of any arbitrary signal. This filtration scheme usually utilizes the two filters that are Butterworth filter along with Ideal filter that reduces the components of high frequency and the recovered image gets acquire after taking the inverse of Fourier transform [13].

2.1.3 Homomorphic Scheme of Filtration

This scheme of filtration mechanism is for frequency domain by the declining the low-frequency elements with the usage of high-pass filtration scheme. This is the most prevalent scheme utilized for intensifying the corrupted imageries by irregular illumination. As the speckle noise in ultrasound imagery is of multiplicative in nature which is a cumbersome task to remove, so this technique eradicates the speckle noise on the authentic image by the usage of logarithmic transformation and it modifies

Fig. 1 **a** Real/authentic ultrasound image **b** noise-corrupted ultrasound imagery [7]

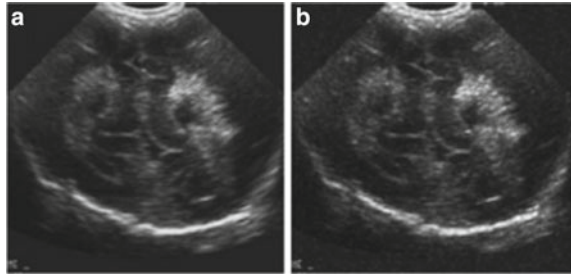


Fig. 2 Sequence of steps followed in homomorphic scheme of filtration [15]

noise into additive [14]. The mathematical calculation is shown in the given equation:

$$\log f(i, j) = \log g(i, j) + \log h(i, j). \quad (3)$$

The steps which are being considered for the conversion of noise from multiplicative to additive are shown in the given Fig. 2.

3 Literature Review

There is plethora of algorithms for the eradication of speckle noise, but two basic schemes are most prominent. The first one is the compounding methodology, and the other one is postprocessing methodology [16]. The compounding methodology amends the data acquisition process to form various imageries for the same region and merge them to make a single imagery [17]. The other one is preprocessing methodology in which different schemes apply on B-mode (2D) imageries after they have been made [18].

The compound methodology introduces a homomorphic filtration scheme for speckle mitigation. Postprocessing filtration schemes implement straight on the authentic imageries. Approaches in this methodology consist of various fixed along with adaptive filtration schemes like adaptive filtration reduction (ASR) [19], adaptive weighted median filter (AWMF) [20], nonlinear diffusion [21], maximum a posteriori (MAP) estimation [22], etc.

Adaptive filtration schemes are well distinguished for mitigation of speckle [23]. It considers a homogeneous region for filtration. Furthermore, arithmetic average/mean filtration schemes operate homogenous areas. Moreover, nonlinear median filtration

methods rectify the edge pixels. However, Huang et al. [24] introduced a interpolation technique known as square distance weighted (SDW) which lessen the smudging in standard interpolation technique of distance weighted, i.e., DW. SDW works according to the inverse distance which is actual weight of that particular pixel and by taking the square of that inverse distance.

4 Experimental Analysis

Noise can be mitigated by the help of two methods, in which first one is to place all the distinct types of algorithms consecutively for achieving the required flawless imagery. On this condition, there is only noisy imagery existed and the genuine noise less imagery does not exist, so the quality cannot be achieved by the usage of traditional metrics. Due to that, a synthetic imagery has been generated by considering a noise less imagery and by the usage of noise models to degrade it. Therefore, rectifying schemes can be utilized for comparison with noise less imagery. Since the contemporary methods for the mitigation of noise are very less, so researchers employ modified and improving those schemes.

Moreover, those schemes can be merged with one another like in homomorphic filtration scheme. For instance, Fig. 3 illustrates the process of filtration with the usage of wavelet transform in homomorphic scheme, and Fig. 4 illustrates the complicated technique which is a merged method of two AWMF along with wavelet schemes. The main reason behind this grouping is to make a pairing that is utilized for the intensification of imageries. This grouping can be of two types. First one is the series pairing in which one output is gone to next, and the second one is parallel grouping in which all the serial combinations are arranged and function in a parallel manner. The final recovered imagery from the entire process is computed by considering the average of all the final imageries from the series pairing.



Fig. 3 Serial combination algorithm process [11]

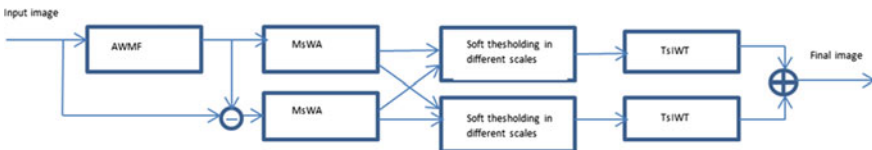


Fig. 4 Parallel combination algorithm process—full forms of adaptive weighted median filtration (AWMF), multi-scale wavelet transformation (MsWA), two-dimensional inverse wavelet transformation (TsIWT) [25]

The first one grouping, i.e., series pairing is solely employed for the utilization of homomorphic filtration in which several functions are performed. Logarithmic procedure is carried out at the origin, and an exponential procedure is being carried out at the ending. Nonetheless, it is also being considered that this filtration scheme can be opened for cascading schemes. By the usage of parallel pairing, the destructive and constructive aspects can be remunerated for every series pairing in order to compensate the flaws of one procedure by the second one. For achieving the best outcomes, a tradeoff will have to be made.

This paper implements an upgraded program Field II [26] for all categories of ultrasound machine’s transducers and its accompanying imageries. This algorithm utilizes the approach of spatial impulse which obtains ultrasound field for pulsed along with constant wave, and it majorly works on linear methodology. The impulse response is known as spatial impulse response because the impulse response is being modified according to the location of the transducer.

The benchmarks utilized for experimental analysis with ultrasound imageries are the mean square error (MSE), PSNR along with SNR [27]. In addition, the beta metric, i.e., β is also used for experimental analysis [25]. The statements for the corresponding terms are stated as:

$$MSE = \frac{\sum_{i=0}^{I-1} \sum_{j=0}^{J-1} [F(i, j) - F'(i, j)]^2}{I \cdot J}, \tag{4}$$

$$SNR = 10 \log_{10} \frac{\frac{\sum_{i=0}^{I-1} \sum_{j=0}^{J-1} F^2(m, n)}{I \cdot J}}{MSE}, \tag{5}$$

$$PSNR = 10 \log_{10} \frac{255^2}{MSE}. \tag{6}$$

For these statements, F states the input authentic imagery; F' demonstrates the image contaminated with the noise and $I \cdot J$ represents the estimation of measurements of imagery.

Beta metric is determined as:

$$\beta = \frac{\sum_{i=0}^{I-1} \sum_{j=0}^{J-1} [\Delta F(i, j) - \Delta' F][\Delta F'(i, j) - \Delta' F']}{\sqrt{\sum_{i=0}^{I-1} \sum_{j=0}^{J-1} [\Delta F(i, j) - \Delta' F]^2 [\Delta F'(i, j) - \Delta' F']^2}}. \tag{7}$$

The Δ operator used in this statement represents a filter that sieves high-pass frequencies for the imagery and Δ' depicts the average value of the image after the Δ operator has been applied.

The MSE benchmark is being utilized for comparison among the imageries like if its value is less, so the error is less and it is almost equal to the original imagery. Moreover, it shows how superior is the performance of technique. Nonetheless, the SNR and PSNR benchmarks demonstrate the connection between original imagery and the approximation error, and large value shows that there is a need of upgrading.

The beta metric conserves the edges in the approximated imagery, and if its value is rising, then it means that its performance is superior otherwise not.

5 Databases and Results

We have simulated our results on three different databases of ultrasound imageries [28] by using different filters and compare them for different parameters. Firstly, we have created a noisy imagery by utilizing the upgraded Field II algorithm, and the benchmarks for the given databases are discussed in the given paragraphs.

5.1 Database: Breast Cyst

- (i) Median filtration: Firstly, original RGB imagery has been considered which is noisy as shown in Fig. 5a, and after that, it is to be converted into grayscale double class in Fig. 5b. In Fig. 5c–f, the imageries have been generated with distinct sizes of windows, i.e., 3×3 , 5×5 , 7×7 and 9×9 . It is clearly visible that while the size of the window surges, the imagery gets flawless or fine along with color also gets consistent. For instance, the imagery acquired from filter 9 window size is barely visible any deviation in the framework of the white region. But, on the contrary side, the destructive consequences are more apparent. For instance, the edges of the first imagery are well distinguished, but if the window size surges, the edges of the imageries get more and more fainted.

The benchmarks for different sizes of windows are shown in Table 1. The table states that the median filtration technique with a 3×3 size mitigates noise while having a superior quality by showing all the values of MSE, SNR and PSNR in table. Figure 6 shows the performance of benchmarks while magnifies the window sizes.

To summarize, if the dimension of the window is less, then noise gets lessen in a small amount but conserves the significant features. In contrast, if the dimension of

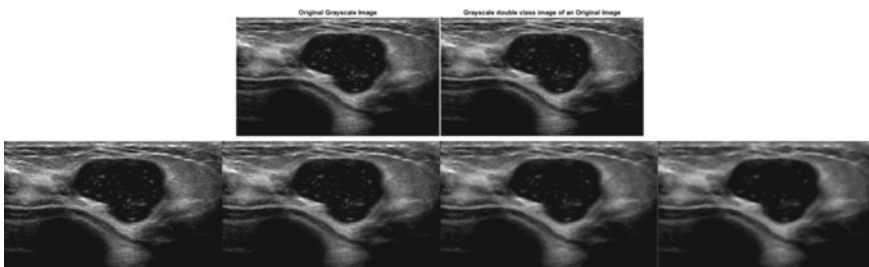
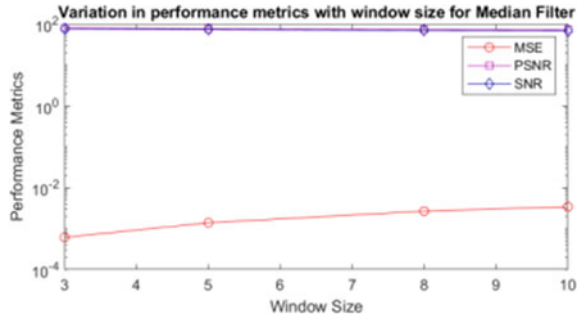


Fig. 5 a Original RGB, b grayscale double class, filtered images having window size c 3×3 , d 5×5 , e 7×7 and f 9×9

Table 1 Objective evaluation for different window sizes for median filtration scheme

Window	MSE	SNR	PSNR
3 × 3	0.000615	77.06138	80.24305
5 × 5	0.001395	73.50398	76.68566
7 × 7	0.00213	71.66478	74.84646
9 × 9	0.002948	70.25331	73.43499

Fig. 6 Benchmarks for median filter for distinct dimensions of window



the window is more, then noise will get more mitigated at the cost of destruction of edges.

(ii) Fourier filtration scheme

Ideal filter

The next filtration is Fourier by using the Ideal filter. The assessment attribute is only one, i.e., cutoff/minimal frequency. The outcomes are revealed in Figs. 7 and 8. The benchmarks are shown in Table 2.

In the first place, the cutoff frequency will be considered as 10%, and after that, it will be increasing from 30 to 50%. By watching the images, it is cleared that if the cutoff frequency gets drop down, then smoothness get surges but sharpness get diminishes. Therefore, the minimal cutoff frequency has to be fixed because the destructive effects will be more significant. For instance, if the cutoff frequency is considered to be 30%, then there has been a significant change, and similarly, if the cutoff frequency is to be considered at 40% and 50%, then there would also be

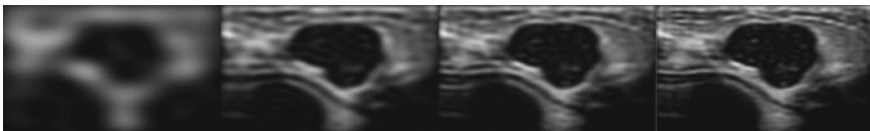


Fig. 7 Filtered images using ideal Fourier filtration having cutoff frequency a 10%, b 30%, c 40% and d 50%

Fig. 8 Benchmarks for ideal Fourier filtration for distinct cutoff frequencies

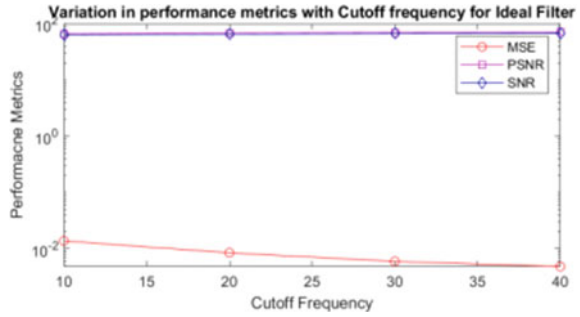


Table 2 Objective evaluation for different cutoff frequencies for ideal Fourier filtration scheme

Cutoff frequency (%)	MSE	SNR	PSNR
10	0.013767	63.56075	66.74242
30	0.005952	67.20262	70.38429
40	0.004837	68.10295	71.28463
50	0.00391	69.0274	72.20907

a noteworthy improvement. The benchmarks, i.e., MSE PSNR together with SNR have also been intensified while increases the minimal frequency.

Butterworth filtration scheme

The prior outcomes from Ideal filtration scheme give sharp-edged cutoff frequencies. These sharp edges can be flattened by this Butterworth filtration and keep down the extreme large frequency imagery details at the cost of lessening noise. Table 3 reveals the benchmarks for the imageries created with this filtration scheme which are shown in Fig. 9.

Table 3 Benchmarks, i.e., MSE, SNR and PSNR for cutoff frequencies for Butterworth filtration scheme

Cutoff frequency (%)	MSE	SNR	PSNR
10	0.012534	63.96818	67.14986
30	0.005115	67.86073	71.04241
40	0.003937	68.99757	72.17925
50	0.003138	69.98283	73.16451

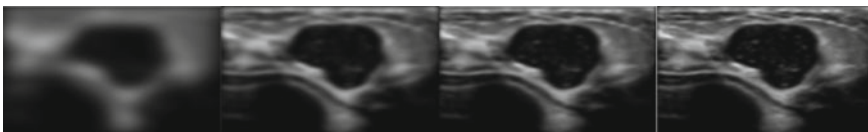


Fig. 9 Filtered images using Butterworth filtration having cutoff frequency a 10%, b 30%, c 40% and d 50%

As we discussed in previous Ideal filtration scheme, the Butterworth filtration scheme will also be considered first minimal frequency of 10% for assessing the outcomes. Figure 9a has superior results of benchmarks than the Ideal filtration scheme with the same minimal frequency. But, the drawback of this filter at this minimal frequency is that imageries got fainted. As the minimal frequency surges like for 30, 40 and 50% which are shown in Figs. 9b and 11c, d, there would be a drop in quality of imagery just like Ideal filter for the same minimal frequencies. The mass is well cleared, and a lesser amount of faintness is exposed in Fig. 9b.

The next minimal frequency is to be considered 40%, and there would be a damage in picture quality. The benchmarks, i.e., PSNR, SNR and MSE also show the destruction in image. When the minimal frequency is taken as 50%, there would be a little change in imagery.

As we increase the minimal frequency, there would be additional noise that diminishes the clarity in edges, but in addition with that, there would be more sharpness also while surges the minimal frequency.

After seeing these result outcomes of Fourier filtration scheme, it is well cleared that Ideal filtration scheme is not acceptable. But this is not the case Butterworth filtration scheme, the imagery gets smoother. The graph for the performance benchmarks is also shown in Fig. 10.

(iii) Homomorphic filtration scheme

The similar tests are performed with homomorphic filtration scheme for the same attributes excluding median filtration scheme. The difference between homomorphic filtration methods and the median filtering method is that Homomorphic filtering methods combine with median filtration method as a nucleus would not alter quality

Fig. 10 Benchmarks for Butterworth filtration scheme for distinct cutoff frequencies

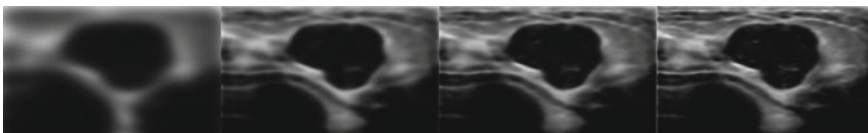
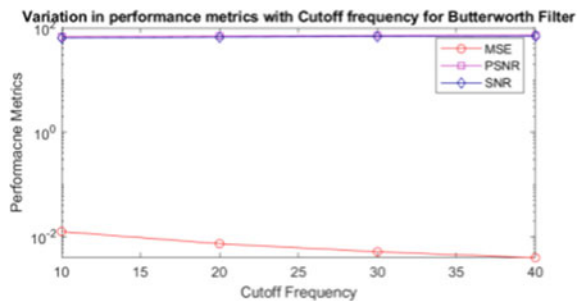


Fig. 11 Filtered images using homomorphic filtration having cutoff frequency a 10%, b 30%, c 40% and d 50%

Table 4 Benchmarks, i.e., MSE, SNR and PSNR for cutoff frequencies for homomorphic filtration scheme

Cutoff frequency (%)	MSE	SNR	PSNR
10	0.014204	63.42511	66.60679
30	0.005207	67.78333	70.965
40	0.003643	69.33504	72.51672
50	0.002668	70.68684	73.86852

of the imagery, but on the contrast, the single median filtration scheme can change the quality as it depends upon the pixels' values present in the imagery. Therefore, usage of logarithm operation to each one would not alter the particular function. The filtered imageries from homomorphic filtration scheme are shown in Figs. 11a–d for different cutoff frequencies. The table for benchmarks is revealed in Table 4.

6 Comparative Investigation

See Table 5.

7 Conclusion

This paper deals with the speckle noise which is the crucial segment in the biomedical imaging. Biomedical imaging is a very vital tool for the diagnosis and examination of any sickness like MRI, ultrasound imageries, X-rays, CT scan, etc. It is really imperative that these imageries have to be flawless so far as possible by considering its usage. This paper emphasizes on the mitigation schemes of the noise which is existent in the biomedical imageries, i.e., stated as speckle noise.

Moreover, this paper has been done comparison among some filtration schemes that are currently utilized for flattening the noise in ultrasonic imageries of biomedical imaging. There has been also a discussion of methods which have been utilized in these experiments. The comparative analysis for the despeckling of the imageries has been accomplished on a noise-less synthetic imagery which is being corrupted by a Field II program. After that, experiments have been performed.

The foremost conclusion is that the finest quality imageries are accomplished with Fourier filtration scheme. The other filtration schemes also give some advancement but in that amount. In addition to that, Fourier filtration scheme is uncomplicated than other filters as it accounts only single parameter. Furthermore, the homomorphic filtration scheme is also advantageous only when it is being utilized with Fourier filtering method. The novelty in this paper is that we have compared different filters on different parameters such as cutoff frequency and different size windows. Our future work is to design an effective and qualitative algorithm that not only removes

Table 5 Comparative analysis among different filters

Serial no.	Median	Fourier filtration		Homomorphic
		Ideal	Butterworth	
1	Comparison has been done on different window sizes	Comparison has been done on different cutoff frequencies	Comparison has been done on different cutoff frequencies	Comparison has been done on different cutoff frequencies
2	If the dimension of the window is less, then the noise gets lesion in a small amount but conserves the significant features	If the cutoff frequency gets drop down, then the smoothness get surges, but sharpness get diminishes. Therefore, the minimal cutoff frequency has to be fixed as the destructive effects will be more significant	Ideal filtration gives sharp-edged frequencies, and these sharp-edged frequencies can be flattened by this filter while simultaneously lessening the noise	The main advantage of the scheme is that it would not alter the quality of the imagery as a nucleus while alone median filtration can change the quality of the imagery as it depends upon the pixels' values present in the imagery
3	If the dimension of the window is more, then the noise will get more mitigated at the cost of the destruction of edges	If the cutoff frequency is set to be at 40 or 50%, then there would be a noteworthy improvement	If the minimal frequency gets increases, then there would be lesser amount of faintness and mass is well cleared, and it is almost little change in original image	
4	Uncomplicated but it depends upon the pixels, so it gives less clarity	Uncomplicated	Uncomplicated while giving the best image quality as it requires only one single parameter Advantage—superior results than the Ideal filtration for the same cutoff frequency Disadvantage—at minimal cutoff frequency, images got faded	Complicated but it is advantageous only when it is being utilized with Fourier filtering method

speckles and enhances edges but also computational cost and time could be less and accuracy should be high.

References

1. Rosa R, Monteiro FC (2016) Performance analysis of speckle ultrasound image filtering. *Comput Methods Biomech Biomed Eng: Imaging Visual* 4(3–4):193–201
2. Ortiz SH, Chiu T, Fox MD (2012) Ultrasound image enhancement: a review. *Biomed Sig Process Control* 7(5):419–428
3. Goyal B, Agrawal S, Sohi BS (2018) Noise issues prevailing in various types of medical images. *Biomed & Pharmacol J* 11(3):1227
4. Wang S, Huang TZ, Zhao XL, Mei JJ, Huang J (2018) Speckle noise removal in ultrasound images by first-and second-order total variation. *Numer Algorithms* 78(2):513–533
5. Alex DM, Christinal AH, Chandy DA, Singh A, Pushkaran M (2020) Speckle noise suppression in 2D ultrasound kidney images using local pattern based topological derivative. *Pattern Recogn Lett* 131:49–55
6. Elyasi I, Pourmina MA (2016) Reduction of speckle noise ultrasound images based on TV regularization and modified bayes shrink techniques. *Optik* 127(24):11732–11744
7. El-Said SA, Azar AT (2012) Speckles suppression techniques for ultrasound images. *J Med Imaging Radiat Sci* 43(4):200–213
8. Loizou CP, Pattichis CS (2011) Despeckle filtering of ultrasound images. In: *Atherosclerosis disease management*. Springer, New York, NY, pp 153–194
9. Zhang J, Wang C, Cheng Y (2015) Comparison of despeckle filters for breast ultrasound images. *Circ Syst Sig Process* 34(1):185–208
10. Wagner N, Eldar YC, Feuer A, Danin G, Friedman Z (2011) Xampling in ultrasound imaging. In: *Medical imaging 2011: ultrasonic imaging, tomography, and therapy*, 25 Mar 2011, vol 7968, p 796818. International Society for Optics and Photonics
11. Zong X, Laine AF, Geiser EA (1998) Speckle reduction and contrast enhancement of echocardiograms via multiscale nonlinear processing. *IEEE Trans Med Imaging* 17(4):532–540
12. Goyal B, Dogra A, Agrawal S, Sohi BS, Sharma A (2020) Image denoising review: from classical to state-of-the-art approaches. *Inf Fusion* 55:220–244
13. Kuttan DB, Kaur S, Goyal B, Dogra A (2021) Image denoising: pre-processing for enhanced subsequent CAD analysis. In: *2021 2nd international conference on smart electronics and communication (ICOSEC)*, 7 Oct 2021. IEEE, pp 1406–1411
14. Kaur S, Goyal B, Dogra A (2022) Denoising hyperspectral imageries with split-Bregman iteration scheme. In: *International conference on intelligent systems design and applications*. Springer, Cham, pp 986–1001
15. Georgieva V (2015) Homomorphic filtering approach for narrowband images enhancement. *J Appl Electromagn (JAE)*, in print. 1 Dec 2015
16. Demmig-Adams B, Adams III WW (2006) Photoprotection in an ecological context: the remarkable complexity of thermal energy dissipation. *New Phytol* 172(1):11–21
17. Behar V, Adam D, Friedman Z (2003) A new method of spatial compounding imaging. *Ultrasonics* 41(5):377–384
18. Stippel G, Philips WR, Lemahieu IL (2002) New denoising technique for ultrasound images using morphological properties of speckle combined with tissue classifying parameters. In: *Medical imaging 2002: ultrasonic imaging and signal processing*, 11 Apr 2002. SPIE, vol 4687, pp 324–333
19. Martinek R, Zidek J (2010) Use of adaptive filtering for noise reduction in communications systems. In: *2010 International conference on applied electronics*, 8 Sep 2010. IEEE, pp 1–6
20. Wang CY, Li LL, Yang FP, Gong H (2010) A new kind of adaptive weighted median filter algorithm. In: *2010 International conference on computer application and system modeling (ICCAASM 2010)*, 22 Oct 2010. IEEE, vol 11, pp V11–667

21. Yu Y, Acton ST (2002) Speckle reducing anisotropic diffusion. *IEEE Trans Image Process* 11(11):1260–1270
22. Wainwright MJ, Jaakkola TS, Willsky AS (2005) MAP estimation via agreement on trees: message-passing and linear programming. *IEEE Trans Inf Theory* 51(11):3697–3717
23. Chen Y, Yin R, Flynn P, Broschat S (2003) Aggressive region growing for speckle reduction in ultrasound images. *Pattern Recogn Lett* 24(4–5):677–691
24. Huang QH, Zheng YP, Lu MH, Chi ZR (2005) Development of a portable 3D ultrasound imaging system for musculoskeletal tissues. *Ultrasonics* 43(3):153–163
25. Hao X, Gao S, Gao X (1999) A novel multiscale nonlinear thresholding method for ultrasonic speckle suppressing. *IEEE Trans Med Imaging* 18(9):787–794
26. Jensen JA (2001) Users' guide for the Field II program. Technical University of Denmark. 2800, p 28
27. Johnson DH (2006) Signal-to-noise ratio. *Scholarpedia* 1(12):2088
28. <https://www.mathworks.com/matlabcentral/fileexchange/67408-image-processing-dataset>

An Analysis of Different Noise Removal Techniques in Medical Images



Jaspreet Kaur, Bhawna Goyal, and Ayush Dogra

Abstract In medical application like diagnosis of diseases, image processing plays crucial role. The process of segmenting the medical images is followed by careful analyzation to detect the ailments ranging from minor to serious. The presence of noise hinders the accurate diagnosis, and denoising becomes important part of image processing in medical images. The denoising approach can be built upon understanding the noise models and using denoising algorithm accordingly. This research aims at studying different medical images, noises and noise removal methods followed by experimental and comparative analysis of the prominent ones.

Keywords Medical images · Noise · Denoising · Filters

1 Introduction

Images are no longer just form of representing information, they are crucial sources of data and these data are key component of many application ranging from robotics, remote sensing to medical diagnosis. The issue of noise in images persists in each application as a result of method of acquisition or process of transmission. Noise suppression has become more crucial with hike in number of applications dependent on image data.

J. Kaur (✉) · B. Goyal

Department of Electronics and Communications Engineering, Chandigarh University, Mohali, India

e-mail: jaspreetkaur11xxi@gmail.com

A. Dogra

Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India

Ronin Institute, Montclair, NJ 07043, USA

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

579

V. V. S. S. Chakravarthy et al. (eds.), *Advances in Signal Processing, Embedded*

Systems and IoT, Lecture Notes in Electrical Engineering 992,

https://doi.org/10.1007/978-981-19-8865-3_52

1.1 Noises in Medical Images

In medical images, there is possibility of occurrence of more than one noise as multi-sensor imaging is deployed. In addition to this, common occurrences of noise emerge as result of insufficient lights or inefficient exposure [1–4].

Some of the most prominent noises that corrupt the medical images are Quantization noise, Gaussian noise, Poisson noise and Speckle noise. Detecting and analyzing the type of noise are key step in denoising as each noise affects the images in its own way [5–10]. Some noise affects the image uniformly while some distorts only part of the pixels [10–14].

2 Literature Review

The denoising techniques depend on the type of noise and the medical image on which it has to be implemented. The number of methods has been proposed by various researchers to implement the efficient denoising algorithm. In this section, survey of different techniques of noise removal has been discussed to get an overview of all the possible techniques that can be executed or optimized to get a methodical approach for removing noise from medical images [14–18].

2.1 Denoising Method Based on Statistical Models

In order to resolve the reconstruction of MR image of brain, minimization of non-convex total is implemented [19]. Minimum–Convex TV (MCTV) is deployed to obtain the MR Image. MR with identical information of structure and patterns of noise are dealt with parallel imaging technique. The structure of images can be identical, but there is difference in the contrast of each image [20–22]. The approach followed in this research is based on learning.

2.2 Denoising Method Based on Noise Derivative

Machine learning can be used to statistically enhance the quality of image and thus remove the effects of the noise. In one such machine learning-based method, CLFAHE is used to process the images before analysis [23–25]. This technique does not remove the noise entirely but just minimizes the distortion caused due to noise.

In another method, Partial Differential Equation (PDE) along with GCV is used to denoise the image [26, 27]. GCV is executed to obtain the threshold that can efficiently remove the noise by analyzing the statistical characteristics of the noisy

image. PMI combined with GCV gives high PSNR values when the intensity and power of the noise are quite high.

2.3 Denoising Method Using Wavelet Technique

Homomorphic wavelet provides the facilitation of extending the threshold that can provide more efficient results [28–31]. This method mainly aims at Speckle noise but is complex process to implement. In order to minimize the effect of noise by softening it, method of soft-thresholding is implemented [32–34]. The model used in this technique is that of optical recovery but while implementing this method, the values of threshold which are quite large might exclude the necessary coefficients. Speckle noise is dealt with another method which deploys adaptive wavelet domain [35–37]. The technique efficiently preserves the details of image, but the noise is not removed vaguely and thus fails to be counted as an effective technique. PSNR value of high range can be obtained with the implementation of Curvelet method [38, 39].

3 Noise Removal Methods

Filtering techniques deployed with algorithms are most prominently used for removal of noise. The neighboring information is used to detect the noise, and based on that noise, best denoising filter is deployed while taking care that quality of the image is not hampered.

In this section, different denoising methods or filters are discussed.

3.1 Median Filter

This filter is prominently used for removal of impulse noise as it traverses through each pixel of image and replaces the pixel with median of the neighboring pixels, and thus, it transforms all the pixels, both noise and noise free. But the drawback is that pixels which are not corrupted are noise also tend to be infected with noise after applying median filters. This filter is capable of denoising without hampering the sharpness of the image [39].

3.2 Wiener Filter

Wiener filter is type of linear filter that deploys statistical method to remove noise from the pixels of and image and it. The spectral properties of the original methods

are studied, and LTI filter is used to obtain the result closer to the original image. It can be used for MRI images. The local neighborhood of each pixel is deployed to narrow down the statistical estimates, and these estimates are used to design adaptive Wiener filter.

3.3 *Gaussian Filter*

Speckle noise is removed with the help of Gaussian filter. The Speckle noise is mostly present in ultrasound images and MRI brain images. This method is based on the technique in which average value of neighboring pixels is used to replace the pixels which are noisy based on Gaussian distribution. It prominently removes blurriness from the images. Therefore, it can be used as linear smoothing filter [40].

4 Experimental Analysis

In this section, the results of applying different filters on different types of noises are represented.

We used images as shown in Fig. 1 for this analysis (Fig. 2).

The mean and variance of Gaussian filter here are 0.1 and 0.1, respectively.

The variance of Speckle noise is 0.5, and noise density of Impulse noise is 0.25.

Now, different filtering techniques discussed above are applied on each dataset to obtain the result and conclude that which filtering technique suits which type of noise best.

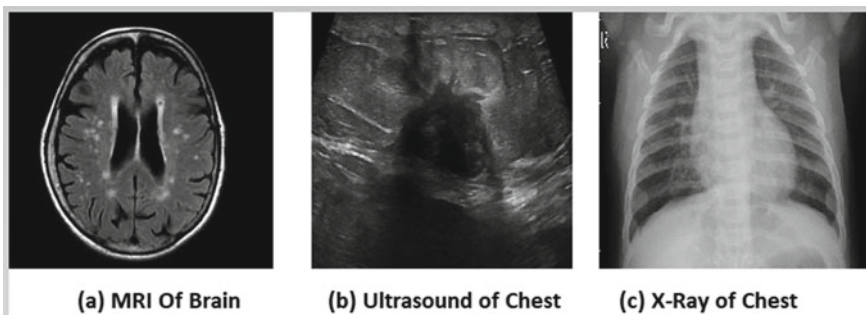


Fig. 1 Dataset for analysis

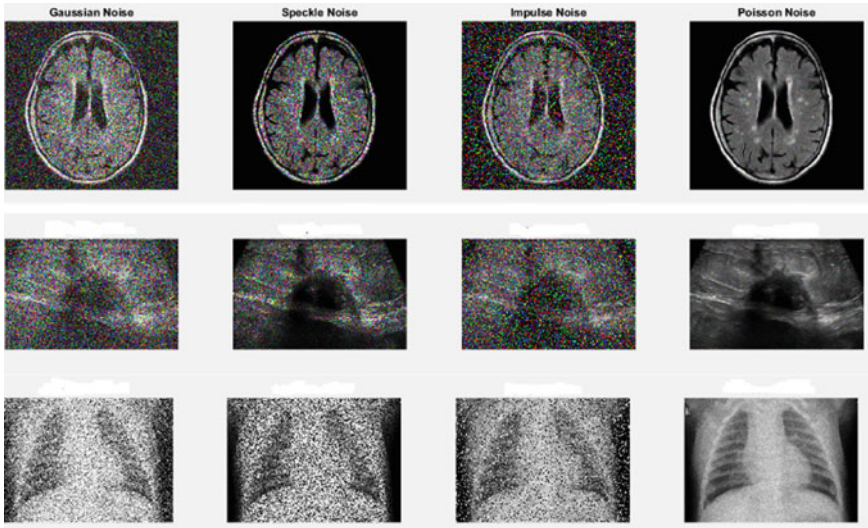


Fig. 2 All kinds of noises added to the dataset. (Left to right) Gaussian noise, Speckle noise, Impulse noise, Poisson noise

4.1 Gaussian Filter

From the above results, we obtained that Gaussian filter has most effect on ultrasound image with Speckle noise. In other images and noises, there is no significant reduction in noise (Fig. 3).

4.2 Wiener Filter

If the results of Gaussian filter and Wiener filter in Fig. 7 are compared, it can be concluded that Wiener filter has produced better results than the Gaussian filter. The best results by Wiener filter are obtained in MRI images and in removal of Impulse noise (Fig. 4).

4.3 Median Filter

Based on Fig. 8, it can be concluded that the median filter has worked best on the Impulse noise for MRI and ultrasound images (Fig. 5).

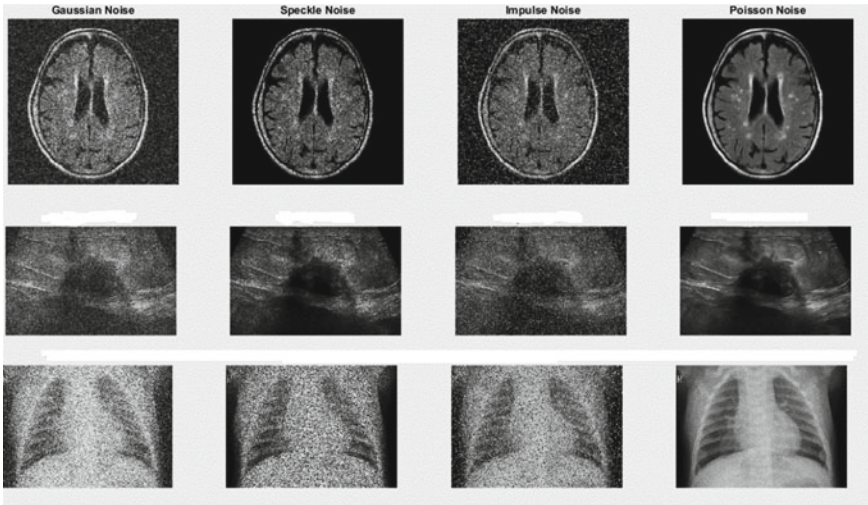


Fig. 3 Result of Gaussian filter

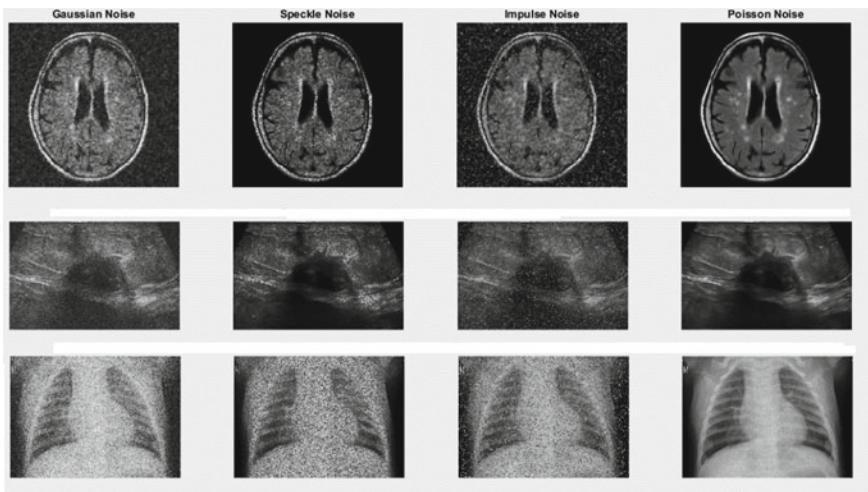


Fig. 4 Result of Wiener filter

5 Comparative Analysis

In this section, the results of different filters on different images are compared based on parameters like Mean Square Error, Signal-to-Noise ratio and Peak Signal-to-Noise Ratio.

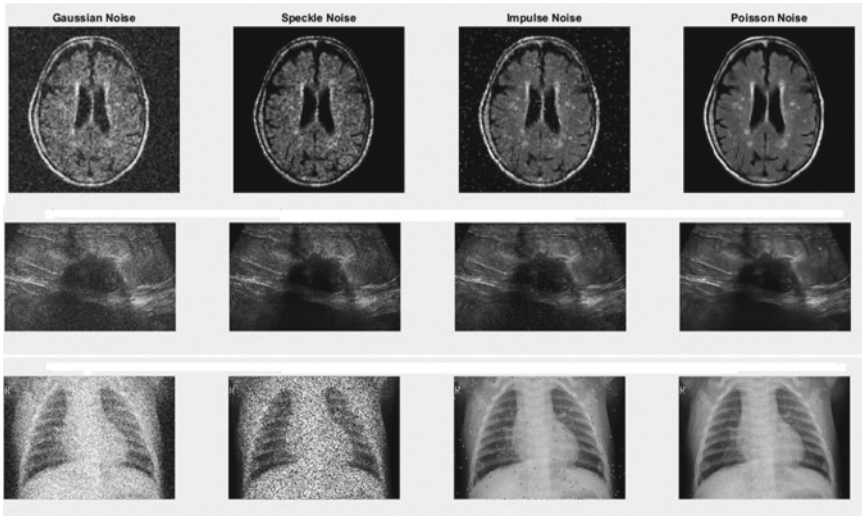


Fig. 5 Result of median filter

Table 1 records MSE, SNR and PSNR for each result of image with Gaussian noise.

Table 1 clearly shows that median filter and Wiener filter are most capable of removing Gaussian noise (Table 2).

Table 1 Measure of parameters of denoising in images with Gaussian noise

Filter	MRI image			Ultrasound image			X-ray image		
	MSE	SNR	PSNR	MSE	SNR	PSNR	MSE	SNR	PSNR
Gaussian filter	0.048	48.946	61.24	0.045	60.60	61.5169	0.036	69.76	62.5142
Median filter	0.027	51.449	63.74	0.026	63.04	63.9603	0.025	71.33	64.0901
Wiener filter	0.035	50.341	62.63	0.029	62.56	63.4782	0.020	72.22	64.9736

Table 2 Measure of parameters of denoising in images with Speckle noise

Filter	MRI image			Ultrasound image			X-ray image		
	MSE	SNR	PSNR	MSE	SNR	PSNR	MSE	SNR	PSNR
Gaussian filter	0.013	54.531	66.829	0.012	66.129	67.039	0.044	68.914	61.668
Median filter	0.01	55.273	67.571	0.009	67.634	68.544	0.036	69.796	62.549
Wiener filter	0.01	54.10	66.40	0.01	66.96	67.87	0.023	71.6887	64.4421

Table 3 Measure of parameters of denoising images with Impulse noise

Filter	MRI image			Ultrasound image			X-ray image		
	MSE	SNR	PSNR	MSE	SNR	PSNR	MSE	SNR	PSNR
Gaussian filter	0.048	48.946	61.244	0.039	61.212	62.123	0.032	70.211	62.964
Median filter	0.004	59.353	71.652	0.001	74.322	75.233	0.001	82.932	75.685
Wiener filter	0.034	50.4963	62.7946	0.025	63.127	64.037	0.019	72.43	65.19

On average, median filter followed by Wiener filter removes the Speckle noise from medical images effectively.

Measure of Parameters of denoising Images with Impulse noise (Table 3).

As evident from Table 4, median filter removes Impulse noise most effectively.

Wiener filter surpasses others when it comes to removing Poisson noise.

The following graph shows MSE values for different filters and noises in different images (Fig. 6).

The above graph shows that in MRI images, the noise is best removed by median filter.

6 Conclusion

The removal of noise from medical images is not only limited to uncovering the crucial information hidden by the noise. Preserving the details of the image is also key aim of denoising when it comes to medical images. This research performed detailed analysis of all kinds of noises and implemented denoising with various filters. Each noise is efficiently removed with one or more types of filters. However, no filter is capable of removing all kinds of noise. This paves way for future researchers to devise algorithms and techniques that can remove all kinds of noises.

Table 4 Measure of parameters of denoising images with Poisson noise

Filter	MRI image			Ultrasound image			X-ray image		
	MSE	SNR	PSNR	MSE	SNR	PSNR	MSE	SNR	PSNR
Gaussian filter	9.5536e-04	66.0308	78.3291	2.4988e-04	83.2430	84.1534	4.9791e-04	88.4059	81.1593
Median filter	9.7561e-04	65.9703	78.2686	2.5036e-04	83.2113	84.1217	4.9683e-04	88.4202	81.1736
Wiener filter	5.2936e-04	68.5928	80.8911	2.5493e-04	83.1648	84.0753	4.0279e-04	89.3040	82.0575

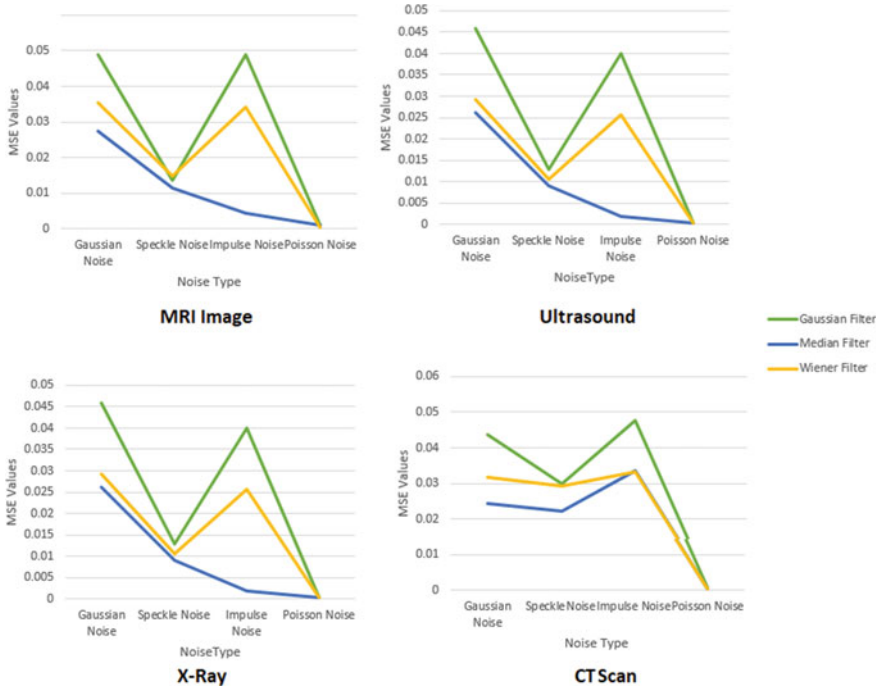


Fig. 6 MSE values’ comparison for different noises and filters in MRI

References

1. Franquet T (2001) Imaging of pneumonia: trends and algorithms. *Eur Respir J* 18(1):196–208
2. Becker M, Magnenat-Thalmann N (2015) Muscle tissue labeling of human lower limb in multi-channel mDixon MR imaging: concepts and applications. *IEEE/ACM Trans Comput Biol Bioinf* 14(2):290–299
3. Patil RC, Bhalchandra AS (2012) Brain tumour extraction from MRI images using MATLAB. *Int J Electron Commun Soft Comput Sci Eng (IJECSCE)* 2(1):1
4. Kumbhakarna V, Patil VR, Kawathekar S (2015) Review on speckle noise reduction techniques for medical ultrasound image processing. *Int J Comput Tech* 2(1):15
5. Magud O, Tuba EVA, Bacanin N (2017) Medical ultrasound image speckle noise reduction by adaptive median filter. *WSEAS Trans Biol Biomed* 14:38–46
6. Dance DR, Christofides S, Maidment ADA, McLean ID, Ng KH (2014) Diagnostic radiology physics: a handbook for teachers and students. Endorsed by: American Association of Physicists in Medicine, Asia-Oceania Federation of Organizations for Medical Physics, European Federation of Organisations for Medical Physics
7. Razifar P, Sandström M, Schnieder H, Långström B, Maripuu E, Bengtsson E, Bergström M (2005) Noise correlation in PET, CT, SPECT and PET/CT data evaluated using autocorrelation function: a phantom study on data, reconstructed using FBP and OSEM. *BMC Med Imaging* 5(1):1–23
8. Scholl I, Aach T, Deserno TM, Kuhlen T (2011) Challenges of medical image processing. *Comput Sci-Res Dev* 26(1):5–13

9. Afshan N, Qureshi S, Hussain SM (2014) Comparative study of tumor detection algorithms. In: 2014 International conference on medical imaging, m-health and emerging communication systems (MedCom). IEEE, Nov 2014, pp 251–256
10. Nanthagopal AP, Sukanesh R (2013) Wavelet statistical texture features-based segmentation and classification of brain computed tomography images. *IET Image Proc* 7(1):25–32
11. Andria G, Attivissimo F, Lanzolla AML (2013) A statistical approach for MR and CT images comparison. *Measurement* 46(1):57–65
12. Hendee WR, Ritenour ER (2003) *Medical imaging physics*. Wiley
13. Kumbhakarna V, Patil VR, Kawathekar S (2015) Review on speckle noise reduction techniques for medical ultrasound image processing. *Int J Comput Tech*
14. Dogra A, Goyal B, Agrawal S, Sohi BS, Anatomical and functional imaging modalities: a brief review. In: International conference, IRSD
15. Sanchez MG, Sánchez MG, Vidal V, Verdú G, Mayo P, Rodenas F (2012) Medical image restoration with different types of noise. In: 2012 Annual international conference of the IEEE engineering in medicine and biology society. IEEE, Aug 2012, pp 4382–4385
16. Olsen SI (1993) Noise variance estimation in images. In: Proceedings of 8th SCIA
17. Goyal B, Agrawal S, Sohi BS (2018) Noise issues prevailing in various types of medical images. *Biomed Pharmacol J* 11(3):1227
18. Gudbjartsson H, Patz S (1995) The Rician distribution of noisy MRI data. *Magn Reson Med* 34(6):910–914
19. Liu Y, Du H, Wang Z, Mei W (2018) Convex MR brain image reconstruction via non-convex total variation minimization. *Int J Imaging Syst Technol* 28(4):246–253
20. Kwon K, Kim D, Park H (2016) Multi-contrast MR image denoising for parallel imaging using multilayer perceptron. *Int J Imaging Syst Technol* 26(1):65–75
21. Chen G, Zhang P, Wu Y, Shen D, Yap PT (2016) Denoising magnetic resonance images using collaborative non-local means. *Neurocomputing* 177:215–227
22. Khan A, Waqas M, Ali MR, Altalhi A, Alshomrani S, Shim SO (2016) Image de-noising using noise ratio estimation, K-means clustering and non-local means-based estimator. *Comput Electr Eng* 54:370–381
23. Fenshia Singh J, Magudeeswaran V (2017) A machine learning approach for brain image enhancement and segmentation. *Int J Imaging Syst Technol* 27(4):311–316
24. Zhang C, Chen Y, Duanmu C, Yang Y (2016) Image denoising by using PDE and GCV in tetrolet transform domain. *Eng Appl Artif Intell* 48:204–229
25. Rafsanjani HK, Sedaaghi MH, Saryazdi S (2016) Efficient diffusion coefficient for image denoising. *Comput Math Appl* 72(4):893–903
26. Lysaker M, Lundervold A, Tai XC (2003) Noise removal using fourth-order partial differential equation with applications to medical magnetic resonance images in space and time. *IEEE Trans Image Process* 12(12):1579–1590
27. Kollem S, Reddy KRL, Rao DS (2019) A review of image denoising and segmentation methods based on medical images. *Int J Mach Learn Comput* 9(3):288–295
28. Gupta S, Chauhan RC, Saxena SC (2005) Homomorphic wavelet thresholding technique for denoising medical ultrasound images. *J Med Eng Technol* 29(5):208–214
29. Donoho DL (1995) De-noising by soft-thresholding. *IEEE Trans Inf Theory* 41(3):613–627
30. Gupta S, Chauhan RC, Saxena SC (2005) Locally adaptive wavelet domain Bayesian processor for denoising medical ultrasound images using speckle modelling based on Rayleigh distribution. *IEEE Proc-Vis Image Sig Process* 152(1):129–135
31. Kamble VM, Parlewar P, Keskar AG, Bhurchandi KM (2016) Performance evaluation of wavelet, ridgelet, curvelet and contourlet transforms based techniques for digital image denoising. *Artif Intell Rev* 45(4):509–533
32. Wang S, Summers RM (2012) Machine learning and radiology. *Med Image Anal* 16(5):933–951
33. Ravishankar H, Prabhu SM, Vaidya V, Singhal N (2016) Hybrid approach for automatic segmentation of fetal abdomen from ultrasound images using deep learning. In: 2016 IEEE 13th International symposium on biomedical imaging (ISBI). IEEE, Apr 2016, pp 779–782

34. Bin-Habtoor ASY, Al-amri SS (2016) Removal speckle noise from medical image using image processing techniques. *Int J Comput Sci Inf Technol* 7(1):375–377
35. Al-Nuaimy W, El Rabaie S, Taha TE, Zahran O, El-Samie A (2016) Mixed curvelet and wavelet transforms for speckle noise reduction in ultrasonic B-mode images. *Inf Sci Comput*
36. Ragesh NK, Anil AR, Rajesh R (2011) Digital image denoising in medical ultrasound images: a survey. In: *Icgst Aiml-11 Conference, Dubai, UAE, vol 12, p 14*
37. Jeyalakshmi TR, Ramar K (2010) A modified method for speckle noise removal in ultrasound medical images. *Int J Comput Electr Eng* 2(1):54
38. Diwakar M, Kumar M (2014) CT image noise reduction based on adaptive Wiener filtering with Wavelet Packet Thresholding. In: *International conference on parallel, distributed and grid computing*. <https://doi.org/10.1109/PDGC.2014.7030722>
39. Ning CY, Liu SF, Qu M (2009) Research on removing noise in medical image based on median filter method. In: *2009 IEEE international symposium on IT in medicine & education*. IEEE, vol 1, pp 384–388
40. Reddy KS, Jaya T (2021) De-noising and enhancement of MRI medical images using Gaussian filter and histogram equalization. *Mater Today: Proc*

Study and Analysis of Classification Techniques for Specific Plant Growths



Riya Sharma, Rashmi Vashisth, and Nidhi Sindhvani

Abstract There are various methods often regard to done for analysis of the subject of a particular group of specimen. There may be more or fewer, but by using classification techniques with a machine learning algorithm, a rate of trend will be shown, and a data representation is provided to examine how a specific group of plants uses to grow on the various levels of surroundings and nature aspects and to demonstrate the idea of plant growth by using classification techniques in machine learning for specific group of plants. In the below research, various studies have been termed that how the plant breed with different aspects of nature and what machine learning algorithms we can have to trace those effects.

Keywords Neural networks · Deep learning · Machine learning · Regression · Farming · Phenotyping artificial intelligence · Genomes · Support vector machine

1 Introduction

Ongoing significant advances in high-throughput field phenotyping have given plant reproducers reasonable and effective instruments for assessing countless genotypes for significant agronomic characteristics at early development stages [1]. In any case, the execution of enormous datasets produced by high-throughput phenotyping devices, for example, hyperspectral reflectance in cultivar advancement programs is as yet testing because of the fundamental requirement for escalated information in computational and measurable examinations [2]. In this review, the heartiness of three normal AI (ML) calculations, multi-facet perceptron (MLP), support vector machine (SVM), and irregular timberland also known as RF were assessed for foreseeing soybean (*Glycine max*) seed yield utilizing hyperspectral reflectance [3].

R. Sharma · R. Vashisth · N. Sindhvani (✉)
Amity Institute of Information Technology, Amity University, Noida, Uttar Pradesh, India
e-mail: nidhiece15@gmail.com

R. Vashisth
e-mail: rvashisth@amity.edu

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
V. V. S. S. Chakravarthy et al. (eds.), *Advances in Signal Processing, Embedded Systems and IoT*, Lecture Notes in Electrical Engineering 992,
https://doi.org/10.1007/978-981-19-8865-3_53

591

Essentially, a plant aggregate, which compares to the biochemical and actual appearance attributes, is impacted by the collaborations between hereditary properties and also natural conditions. Since it varies as per plant species, it is essential to quantify the connection among aggregates and natural conditions for each plant species [4]. To take care of this issue, the advancement of plant phenotyping frameworks for different plant species has been led for quite a long time. This shows that how a plant behaves according to the various types of surrounding and since it can moderate various aspects of the nature phenomenon [5].

The remainder of this paper is organized as follows: In first section, introduction is followed by the example of various researched supervisions followed by Sect. 2 of learning the plant growth using various techniques, and Sects. 3 and 4 talks about usage of proposed algorithm with various technologies while providing examples in Sects. 5 and 6 along with the conclusion.

2 Learning of Plant Growth and Breeding

As the word of “big data” and its evolution time in plant sciences, a test in both essential and applied explorations (for example, rearing applications) is to clarify or foresee aggregates from the hidden genotypes under various natural conditions [6]. Genotypic variety prompts contrasts in the biochemical cosmetics of cells, which thus along with the climate impact organ development, plant development, and at last, characteristics important in horticulture, like yield and resistance to stresses and irritations [7]. Disentangling the impacts of genotypic variety and climate on aggregates yields central experiences into the guideline of significant cycles in plant advancement and physiology and the capacity to anticipate yield and quality attributes from genotypes in explicit conditions, which is fundamental in present-day sub-atomic plant rearing [8]. Breaking down aggregates estimated at these various levels or connecting these aggregates to genotypes progressively calls for handling and coordination of enormous, loud, and heterogeneous informational indexes. The time series is a new type of fundamental to describe the anomaly of screening the plant traits and examining their growth, physiology, ecology, etc. [9] (Fig. 1).

This has been termed as a great initiative to using machine learning algorithms and using Deep Learning architecture to utilizing the resources more profoundly [11]. This has been achieved by using traditional image processing and moreover on the machine learning-based simulations of providing such imaging computations. Computer vision preferably the OpenCV resources is known as the intervention part to classifying the generation of images and the factual representation of the difference.

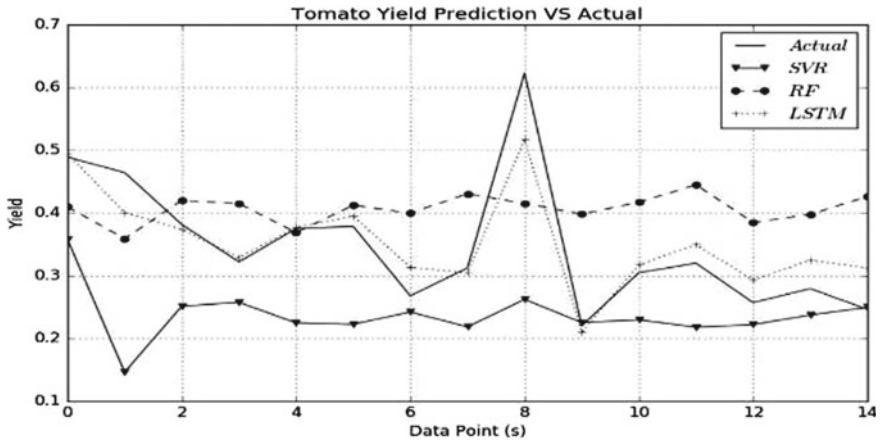


Fig. 1 Tomato yielding using prediction mechanism [10]

3 Methods of Classifying the Plant Ecosystem Using Technologies

Plant psychometry is often very conjugational and hence derives its various properties which is hard to understand easily. Thus, researchers are made a various evolution in this field using various technologies to find a prospectus solution for this architecture [12]. The classification of plants is also termed as the assessment of the various complex plant traits consisting of growth, resistance, ecology, physiology, and what not. By and large, plant attributes have been estimated physically in phenotyping research [13] (Fig. 2).

This limits throughput and confines extensive examination a thought alluded to as the phenotyping bottleneck. Picture-based phenotyping has been proposed as an answer for this bottleneck, as it has shown extraordinary potential in expanding the scale, throughput, productivity [14], and speed of phenomics research. It is currently ordinarily contended that profound learning strategies for picture division, highlight extraction, and information investigation are the key for progress in picture-based high-throughput plant phenotyping [15]. Phenotyping is itself a large acknowledged term for assessing these traits of the plant evolution.

3.1 Traditional Image Processing for Generalizing the Plant System

In this type of simulation method, it has been occurred to see a great deal of using the plant phenotyping which is recently based on the machine learning systems which shows an extraordinary accuracy for checking the plant types [16].

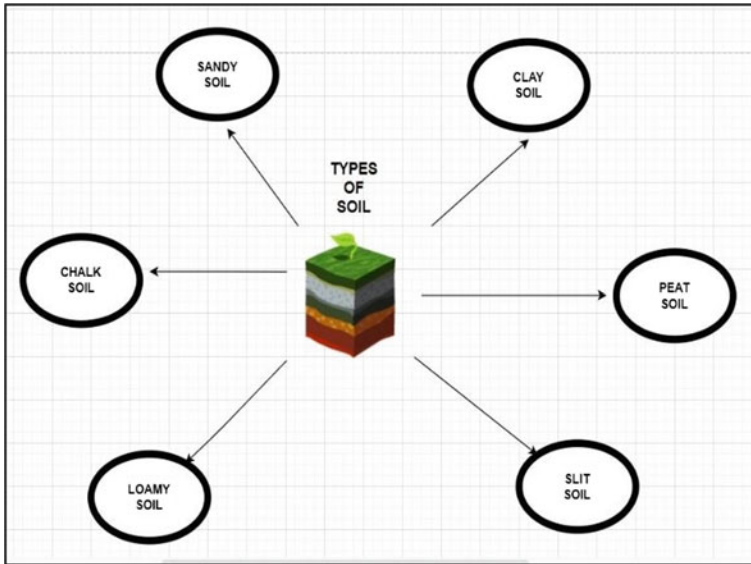


Fig. 2 Types of soil for plant growths

3.2 Computer Vision Using Learning Methods

These methods are being executed particularly at research standards more preferably in the static plants of analysis, demonstrating that they are being made learn to compile the plant growth patterns and can develop the learning-based approaches from them coined as to be “*genetic phenotyping*” [17]. Moreover, the generalization with the fact that using the conventional neural networks are often give more promising results in case of computer vision of Convolution Neural networks [18]. Such methodologies have shown superior execution when analyzed with customary picture-based phenotyping approaches, and researcher is depending more and to a greater degree toward these remarkable outcomes to catch complex elements and designs of plants both above and underneath the ground [19].

Below is the example of showcase of the fresh weight of the plants using the machine learning. The correlation ratio can be shown as a visualization result for a particular harvest rate by taking a factor of PA with value of 180 h [20]. The set of light green or dark green shows the set of experiments of different plants [21].

4 Analysis of Simulation of Plant Growth in 3D/2D

This is an innerving machine learning method for the strategic image analysis which has to be fully explored on the various time series data [22]. An analysis plant growth

of traditional has been performed for the modeling of temporal analysis. Several tools and approaches for the deterministic methods are being used for such type of classifications which use these types of 2D/3D determinations [23]. By Providing these simulations using machine learning algorithms for providing the greenhouse growers, such as biologists of plant scientists to analyze the future of growing nature of the patterns [24]. This shows the understanding of the behavior of environmental efforts on biotic and abiotic factors. Numerous static examination frameworks might be reached out into transient information by essentially running them for each edge premise [25].

4.1 Unconventional Mapping of Predicting Plant Growth

This involves the tasks of mapping and predicting the plant growth which sometimes seems to be challenging. Apart from these Generative Adversarial Networks is the new intuitive proposed by the researcher of Goodfellow which is an emerging format of machine learning methods being developing for the generation of the realistic images to forecast the prediction [26] (Figs. 3 and 4).

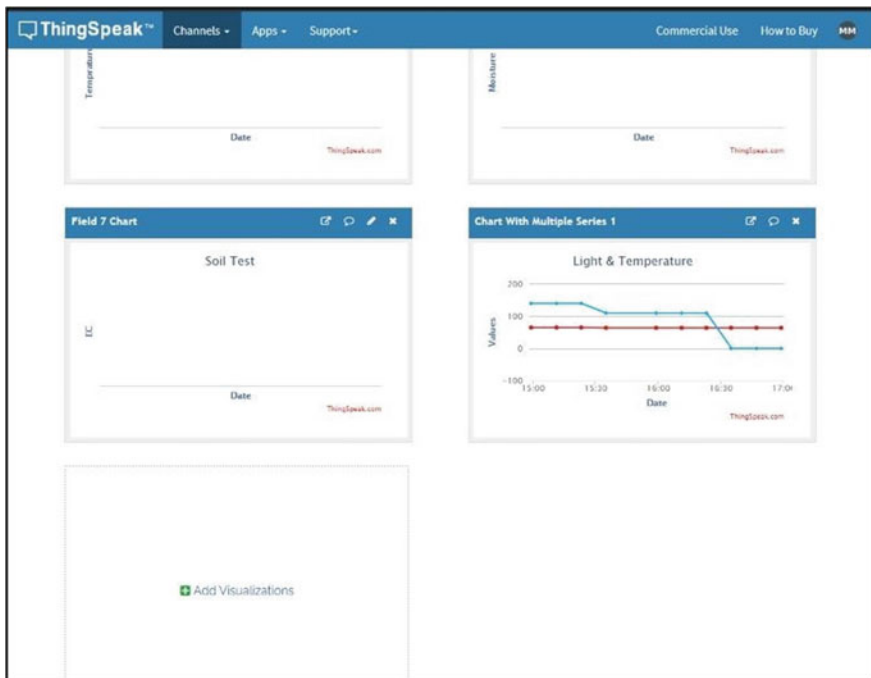


Fig. 3 Time-to-time check for the data measures

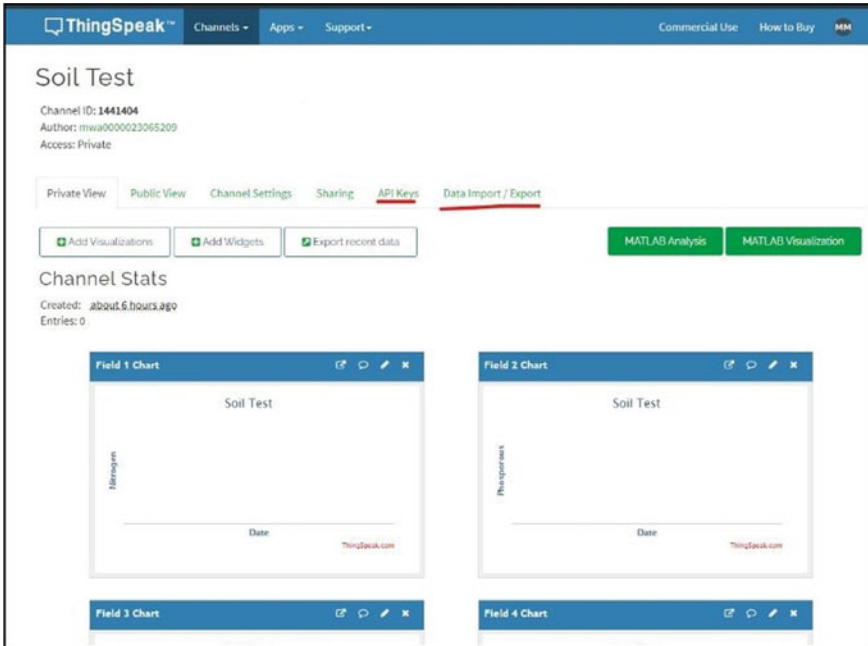


Fig. 4 ThinkSpeak data visualization dashboard

Perhaps, using the various tools comes along with many distinguishable techniques; likewise, below are the data visualization tool for monitoring soil values for the specific plants’ growth with analytical charts (Fig. 5).

5 Data and Machine Learning Types With Contribution

This work is planned as a characteristic augmentation of the work introduced in RootNav 2.0 into transient groupings, fusing a GAN that guides plant development designs and estimates plant development [27, 28]. The proposed framework brings about conceivable plant improvement expectations, hypothetically permitting tests like this to close early, accelerating the trial cycle [29]. Moreover, there are several types and key contributions that have been made throughout the specific contribution with image processing and data amplification for predicting a several species of growth are as follows (Fig. 6).

Fig. 5 Mapping example using GAN algorithm using machine learning [30]

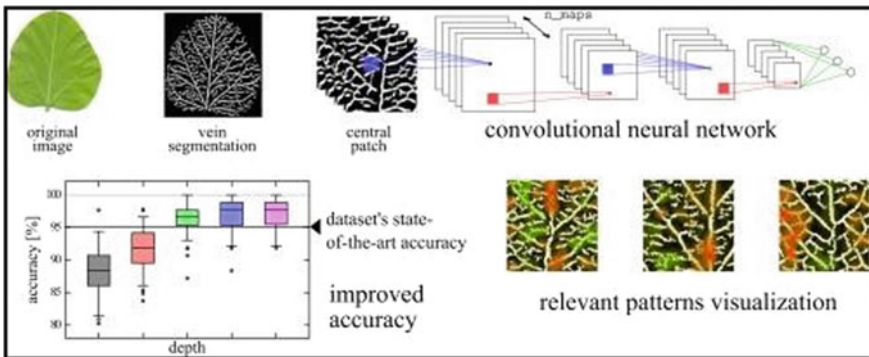
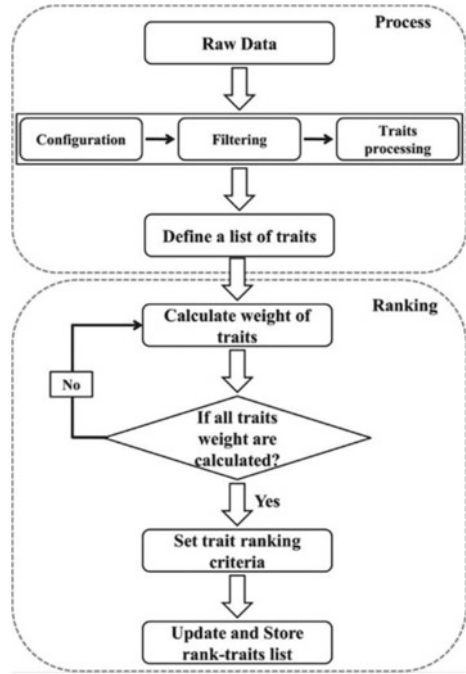


Fig. 6 Leaf prediction using neural network algorithm [34]

5.1 Innovative Datasets and Novel Preprocessing

The greater part of the investigations in plant phenotyping put together remote detecting centers just with respect to a couple of accessible plant datasets. We gained the Arabidopsis dataset and utilized the most recent AI programming to clarify it. The general comment was programmed, and it produces picture and XML-based explanations that can be interoperable between various programming devices [31].

Along these lines, the recorded plant information (particularly, roots) can be utilized for future examination and tests.

5.2 Deep Learning Models with Specification Innovation

The field of remote detecting continually improves through the turn of events and utilization of imaginative AI models. GANs are one of the least explored different avenues regarding AI techniques for plant phenotyping. The proposed explored key target is to exhibit the strength and variety of GANs and use it to improve phenotyping efficiency [32]. The higher goal yield created by the continuously developing GAN design embraced and improved is additionally one of the proposed techniques of key commitments [33].

5.3 Comparative Result Articulations

The discussed framework is intended to consolidate and use both spatial and worldly plant information to conjecture development. It offers an exact and effective prescient division of plant information (root/leaf). Precise forecast of future plant development could significantly diminish the time expected to direct development tests, with plants requiring less developing time and new test cycles starting sooner [35].

5.4 Productive Reproductive and Generalization

It is a vigorous AI-based framework that might be reapplied to any dataset with just minor adjustments [36]. We show this through the use of this framework on two totally different datasets of plant shoots and roots separately. Below, we can see that also there are various substitute methods of the various sequential methods of machine learning algorithms such as regression and classification with which one can understand the behavior of the [36] demographics and all other such type of mechanism of a plant, crop, or behavior. Based on the various methodologies of machine learning or algorithms demonstrated can be effortlessly adjusted to changing conditions and have worked on the vigor of picture-based phenotyping [37]. Various analyses have shown that profound learning-based techniques can productively catch complex elements of plants both above and underneath the ground [38]. This is assisting with accelerating trial cycles, which is fundamental to streamline plant usefulness and, thus, further develop food security and yield protection from ecological burdens. Great above and subterranean picture information is restricted; it is essential to use in vitro assets to advance the innovations prior to applying them in the field [39]. Various arising

works have investigated the utilization of both spatial and fleeting elements of plant development to deliver more helpful and instructive phenotypic data [40].

6 Plants' Ecosystem and How They Develop

Plants are the nurturing part of the ecosystem as it has been totally depending on its environment and the products being enriched with the soil. If not taken care of, there may be numerous items to clean from the air in the area that serves as a shaded lounge [41]. Plants are much more than giving oxygen to the surroundings the other roles can be categorized with it are in carbon cycle, controlling global warming, etc. They are the most important producers in which they use the sunlight to convert the CO₂ into glucose or some other parts of sugar in which the classification technique helps in determining the growth pattern of plants [42]. This is done by examining the features like height, width, color, and shape of plants. The eight ways' plants are classified based on plant similarities, including kind of stem, size of plant, stem growth form, kind of fruit, life cycle, foliage retention, temperature tolerances, and number of cotyledons and leaf venation [43]. It also covers binomial nomenclature and scientific classification. Plants provide oxygen, food, and shelter for other organisms in the ecosystem. They provide us with oxygen, food, medicine, fuel, and many other things. Classification technique is a process that helps to identify a particular plant growth by its characteristics and traits [44]. It can be done by looking at different parts of a plant such as leaves or flowers or even fruit. The classification technique for specific plant growths can be done through their leaves which have certain shapes like ovate, elliptic or serrated, etc. They also have certain colors like green, brownish-green or purple, etc. They also have certain sizes like small to large, etc. [45]. The classification technique is very important because it determines how plants will grow in different environments. It also helps identify which type of plant is best suited for a specific environment. Plants can be classified into four types based on their growth pattern, which are:

- Annuals,
- biennials,
- perennials,
- shrubs.

7 Type of Soil Testing For Checking Plants' NPK Values From Soil Based on Moisture Values

The main idea for checking soil moisture and NPK values to plant growth and nutrients is that to allow farmers of the workers to have the proper knowledge of their plants [46] (Figs. 7 and 8).

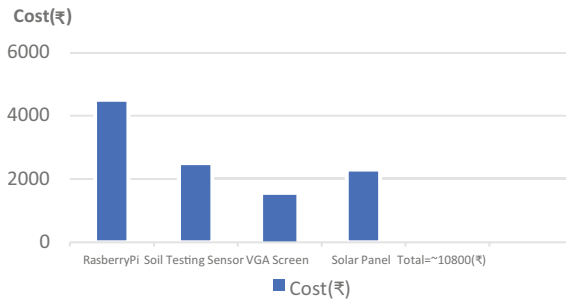


Fig. 7 Price distribution for making one system for checking soil moisture and NPK values

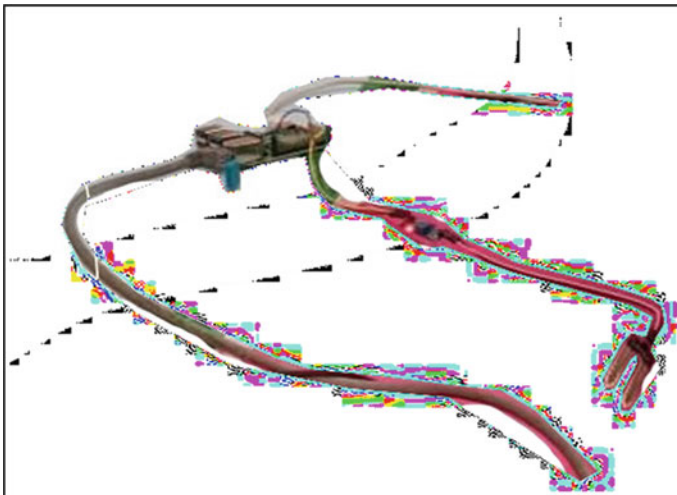


Fig. 8 Soil moisture sensor for checking moisture values

7.1 Hardware

- This part consists the soil checking sensor which will be installed with the harrow or tiller which will go inside with the contact of the soil [47]. Just when the harrow goes into the contact of the soil, the sensor will provide the reading for the different values which consist of pH value, fertility, humidity, fertility, acidity, etc. [48].
- Along with it, a smart screen panel is also installed near the steering wheel to get informed about the soil activity [49]. Beneath with the soil sensor on the wheel panel of tractor, there is a solar panel which is also installed along with the battery system which will provide the electricity separately for the system [50].

7.2 Software

- Here, the main functions will perform its task in which first the smart screen panel is made up of using the Raspberry Pi mini size computer which has the ability to handle the AI and machine learning algorithms and can be used to provide various analytical operations using cloud-based and offline technology.
- The readings are recorded with the use of this smart screen panel which is connected throughout the integration of offline- and online-based connectivity which will inform the farmer about the soil components and ratio.
- Artificial Intelligence Predictions will provide the information about the crops and the machine learning will store and compute all the possible actions which can be done to grow better crops within just matter of time. This will help the farmer not to waste so much time to get the result for testing and can start farming in time (Fig. 9).

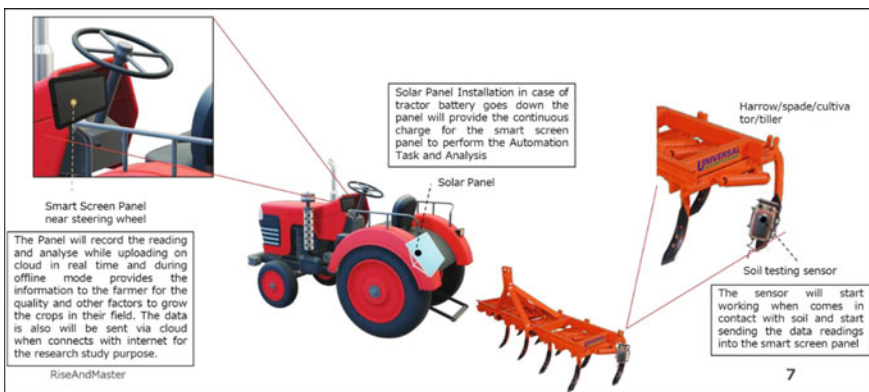


Fig. 9 Proposed framework for soil testing resulting for farmer benefits for plant growth and their nutrients

7.3 Previous Methods for Soil Testing/Traditional Ways



8 Conclusion

This study shows that how can a plant structure can be studied using dependency regardless to their nature but with their articulated growth system. Often, the study also provides a specified problem statement with findings of the individual study resource in accordance to its significance of using various technologies [42]. There can be various technologies stated of fundamentals of the nature of machine learning, sectional regression of the classifications, etc. This study also tried to showcase the study of neural networks on the plants' genome and phenotyping introduction with the involvement of the specified data capturing methods using image processing and preprocessing. Also, the outcomes introduced here have shown that AI and GANs specifically can perform complex determining of plant development. These methodologies can be retrained and adjusted into different areas, possibly helping a wide scope of analysts.

References

1. Srivani P, Yamuna Devi CR, Manjula SH (2022) Prediction and comparative analysis using ensemble classifier model on leafy vegetable growth rates in DWC and NFT smart hydroponic system. In: IOT with Smart Systems, pp 795–804. Springer, Singapore
2. Li W, Zhu D, Wang Q (2022) A single view leaf reconstruction method based on the fusion of ResNet and differentiable render in plant growth digital twin system. *Comput Electron Agric* 193:106712
3. He Y, Yang T, Yan S, Niu S, Zhang Y (2022) Identification and characterization of the BEL1-like genes reveal their potential roles in plant growth and abiotic stress response in tomato. *Int J Biol Macromol*

4. Jung J-Y, Lee S-H, Kim T-H, Oh M-M, Kim J-K (2022) Shape based deep estimation of future plant images. *IEEE Access*
5. Asaf S, Jan R, Khan AL, Bilal S, Asif S, Al-Harrasi A, Kim KM (2022) Unraveling the genome sequence of plant growth promoting *aspergillus niger* (CSR3) Provides insight into the synthesis of secondary metabolites and its comparative genomics. *J Fungi* 8(2):107
6. Shukla R, Dubey G, Malik P, Sindhwani N, Anand R, Dahiya A, Yadav V (2021) Detecting crop health using machine learning techniques in smart agriculture system. *J Sci Ind Res (JSIR)* 80(08):699–706
7. Bhatta M, Sandro P, Smith MR, Delaney O, Voss-Fels KP, Gutierrez L, Hickey LT (2021) Need for speed: manipulating plant growth to accelerate breeding cycles. *Curr Opin Plant Biol* 60:101986
8. Manggau FX, Parennden D (2021) Interpretation of food crop growth progress visualization and prediction of drone based production estimates based on histogram values in government areas-case study of Merauke regency. *IOP Conf Ser: Mater Sci Eng* 1125(1):012102
9. Alhnaity B, Pearson S, Leontidis G, Kollias SD (2019) Using deep learning to predict plant growth and yield in greenhouse environments. *ArXiv*, abs/1907.00624
10. Singh J, Kumar V, Kumar P (2021) Kinetics and prediction modeling of heavy metal phytoremediation from glass industry effluent by water hyacinth (*Eichhornia crassipes*). *Int J Environ Sci Technol*, pp 1–12
11. Sharma R, Saxena R, Rana A (2021) Unsupervised learning in accordance with new aspects of artificial intelligence. In: *Machine learning approach for cloud data analytics in IoT*, pp 429–459
12. Sharma R, Mogha M, Tanwar S, Rana A (2021) Revolution of IoT in healthcare during covid-19. In: *2021 9th International conference on reliability, infocom technologies and optimization (trends and future directions) (ICRITO)*, pp 1–5. *IEEE*
13. Rizkiana A, Nugroho AP, Salma NM, Afif S, Masithoh RE, Sutiarso L, Okayasu T (2021) Plant growth prediction model for lettuce (*Lactuca sativa*.) in plant factories using artificial neural network. *IOP Conf Ser: Earth Environ Sci* 733(1):012027
14. Wang X, Miao Y, Dong R, Zha H, Xia T, Chen Z, Kusnierek K, Mi G, Sun H, Li M (2021) Machine learning-based in-season nitrogen status diagnosis and side-dress nitrogen recommendation for corn. *Eur J Agron* 123:126193
15. Fuentes S, Tongson E, Unnithan RR, Gonzalez Viejo C (2021) Early detection of aphid infestation and insect-plant interaction assessment in wheat using a low-cost electronic nose (E-Nose), near-infrared spectroscopy and machine learning modeling. *Sensors* 21(17):5948
16. Yousefi S, Pourghasemi HR, Avand M, Janizadeh S, Tavangar S, Santosh M (2021) Assessment of land degradation using machine-learning techniques: a case of declining rangelands. *Land Degrad Dev* 32(3):1452–1466
17. Gupta D, Sharma P, Choudhary K, Gupta K, Chawla R, Khanna A, Albuquerque VH (2021) Artificial plant optimization algorithm to detect infected leaves using machine learning. *Exp Syst* 38(6):e12501
18. Zheng C, Abd-Elrahman A, Whitaker V (2021) Remote sensing and machine learning in crop phenotyping and management, with an emphasis on applications in strawberry farming. *Remote Sens* 13(3):531
19. Yu J, Wang J, Leblon B (2021) Evaluation of soil properties, topographic metrics, plant height, and unmanned aerial vehicle multispectral imagery using machine learning methods to estimate canopy nitrogen weight in corn. *Remote Sens* 13(16):3105
20. Nagano S, Moriyuki S, Wakamori K, Mineno H, Fukuda H (2021) Leaf-movement-based growth prediction model using optical flow analysis and machine learning in plant factory. *Front Plant Sci*, 10, ISSN 1664-462X
21. Awika HO, Mishra AK, Gill H, DiPiazza J, Avila CA, Joshi V (2021) Selection of nitrogen responsive root architectural traits in spinach using machine learning and genetic correlations. *Sci Rep* 11(1):1–13
22. Shrivastava VK, Pradhan MK (2021) Rice plant disease classification using color features: a machine learning paradigm. *J Plant Pathol* 103(1):17–26

23. Abioye EA, Hensel O, Esau TJ, Elijah O, Abidin MS, Ayobami AS, Yerima O, Nasirahmadi A (2022) Precision irrigation management using machine learning and digital farming solutions. *AgriEngineering* 4(1):70–103
24. Gao F, Shen Y, Sallach JB, Li H, Zhang W, Li Y, Liu C (2022) Predicting crop root concentration factors of organic contaminants with machine learning models. *J Hazard Mater* 424:127437
25. Shahoveisi F, Riahi Manesh M, del Río Mendoza LE (2022) Modeling risk of *Sclerotinia sclerotiorum*-induced disease development on canola and dry bean using machine learning algorithms. *Sci Rep* 12(1):1–10
26. Pane C, Manganiello G, Nicastro N, Carotenuto F (2022) Early detection of wild rocket tracheofusariosis using hyperspectral image-based machine learning. *Remote Sens* 14(1):84
27. Danilevicz MF, Bayer PE (2022) Machine learning for image analysis: leaf disease segmentation. In: *Plant bioinformatics*, pp. 429–449. Humana, New York, NY
28. Kecoglu I, Sirkeci M, Unlu MB, Sen A, Parlattan U, Guzelcimen F (2022) Quantification of salt stress in wheat leaves by Raman spectroscopy and machine learning. *bioRxiv*
29. Hayashi E, Amagai Y, Kozai T, Maruo T, Tsukagoshi S, Nakano A, Johkan M (2022) Variations in the growth of cotyledons and initial true leaves as affected by photosynthetic photon flux density at individual seedlings and nutrients. *Agronomy* 12(1):194
30. Barzin R, Lotfi H, Varco JJ, Bora GC (2022) Machine learning in evaluating multispectral active canopy sensor for prediction of corn leaf nitrogen concentration and yield. *Remote Sens* 14(1):120
31. Diaz-Gonzalez FA, Vuelvas J, Correa CA, Vallejo VE, Patino D (2022) Machine learning and remote sensing techniques applied to estimate soil indicators—review. *Ecol Indic* 135:108517
32. Banerjee I, Madhumathy P (2022) IoT based agricultural business model for estimating crop health management to reduce farmer distress using SVM and machine learning. In: *Internet of things and analytics for agriculture*, vol 3, pp 165–183. Springer, Singapore
33. Grinblat GL, Uzal LC, Larese MG, Granitto PM (2016) Deep learning for plant identification using vein morphological patterns. *Comput Electron Agric* 127:418–424
34. Singh H, Rehman TB, Gangadhar C, Anand R, Sindhwani N, Babu M (2021) Accuracy detection of coronary artery disease using machine learning algorithms. *Appl Nanosci*, pp 1–7
35. Srisongkram T, Waithong S, Thitimetharoch T, Weerapreeyakul N (2022) Machine learning and in vitro chemical screening of potential α -amylase and α -glucosidase inhibitors from Thai indigenous plants. *Nutrients* 14(2):267
36. Hounkpatin KO, Bossa AY, Yira Y, Igwe MA, Sinsin BA (2022) Assessment of the soil fertility status in Benin (West Africa)—digital soil mapping using machine learning. *Geoderma Reg* 28:e00444
37. Kumar M, Phukon SN, Paygude AC, Tyagi K, Singh H (2022) Mapping Phenological Functional Types (PhFT) in the Indian Eastern Himalayas using machine learning algorithm in Google Earth Engine. *Comput Geosci* 158:104982
38. Hao M, Chen S, Qian Y, Jiang D, Ding F (2022) Using machine learning to identify the potential marginal land suitable for giant silvergrass (*Miscanthus × giganteus*). *Energies* 15(2):591
39. Yu J, Wang J, Leblon B, Song Y (2022) Nitrogen estimation for wheat using UAV-based and satellite multispectral imagery, topographic metrics, leaf area index, plant height, soil moisture, and machine learning methods. *Nitrogen* 3(1):1–25
40. Sindhwani N, Maurya VP, Patel A, Yadav RK, Krishna S, Anand R (2022) Implementation of intelligent plantation system using virtual IoT. In: *Internet of things and its applications*, pp 305–322. Springer, Cham
41. Jadhav S, Garg B (2022) Comprehensive review on machine learning for plant disease identification and classification with image processing. In: *Proceedings of international conference on intelligent cyber-physical systems*, pp 247–262. Springer, Singapore
42. Kohli L, Saurabh M, Bhatia I, Sindhwani N, Vijn M (2021) Design and development of modular and multifunctional UAV with amphibious landing, processing and surround sense module. *Unmanned aerial vehicles for internet of things (IoT) concepts, techniques, and applications*, pp 207–230

43. Wilhelm RC, van Es HM, Buckley DH (2022) Predicting measures of soil health using the microbiome and supervised machine learning. *Soil Biol Biochem* 164:108472
44. Jindal S, Sharma M, Awasthi P, Sreejath S, Sharma M, Mannully CT, Laxmi A (2022) A negative feedback loop of TOR signaling balances growth and stress-response trade-offs in plants. *Cell Rep* 39(1):110631
45. Jiménez-Arias D, Hernández AE, Morales-Sierra S, García-García AL, García-Machado FJ, Luis JC, Borges AA (2022) Applying biostimulants to combat water deficit in crop plants: research and debate. *Agronomy* 12(3):571
46. Kundu A, Mishra S, Kundu P, Jogawat A, Vadassery J (2022) Piriformospora indica recruits host-derived putrescine for growth promotion in plants. *Plant Physiol* 188(4):2289–2307
47. Sun Y, Wu J, Shang X, Xue L, Ji G, Chang S, Niu J, Emaneghem B (2022) Screening of siderophore-producing bacteria and their effects on promoting the growth of plants. *Curr Microbiol* 79(5):1–12
48. Wei C, Jiao Q, Agathokleous E, Liu H, Li G, Zhang J, Fahad S, Jiang Y (2022) Hormetic effects of zinc on growth and antioxidant defense system of wheat plants. *Sci Total Environ* 807:150992
49. Parecido RJ, Soratto RP, Guidorizzi FV, Perdoná MJ, Gitari HI (2022) soil application of silicon enhances initial growth and nitrogen use efficiency of Arabica coffee plants. *J Plant Nutr* 45(7):1061–1071
50. Khalid U, Sher F, Noreen S, Lima EC, Rasheed T, Sehar S, Amami R (2022) Comparative effects of conventional and nano-enabled fertilizers on morphological and physiological attributes of *Caesalpinia bonducella* plants. *J Saudi Soc Agric Sci* 21(1):61–72

Comparative Analysis of Optimization Algorithms for Antenna Selection in MIMO Systems



Nidhi Sindhvani, Rohit Anand, G. Nageswara Rao, Sudhir Chauhan, Alka Chaudhary, Ankur Gupta, and Digvijay Pandey

Abstract Since past few years, there has been a huge demand for the applications involving enormous data rate. Between them, Multiple Input–Multiple Output (MIMO) schemes have been proved to have exceptional benefits in connection with spectral efficiency. Implementation of MIMO device is intricate because of the usage of a greater number of antennas. To address this key issue, antenna selection is one attractive approach to mitigate this requirement. The upcoming radio networks need basic perception of the design concepts and regulation processes to easily control the various resources. The policies related to the assignment of resources are very crucial as far as radio communication networks are concerned as they are directed toward the Quality of Service (QoS) required at the client level. This paper investigates the concept of choice of antennas and related topics for MIMO wireless networks.

N. Sindhvani · A. Chaudhary
AIIT, Amity University, Noida, Uttar Pradesh, India
e-mail: nidhiece15@gmail.com

A. Chaudhary
e-mail: achaudhary4@amity.edu

R. Anand (✉)
DSEU, G.B.Pant Okhla-1 Campus, Govt. of NCT of Delhi, New Delhi, India
e-mail: roh_anand@rediffmail.com

G. Nageswara Rao
EEE Department, Lakireddy Bali Reddy College of Engineering, Mylavaram, JNTUK Kakinada, Andhra Pradesh, India

S. Chauhan
AIAS, Amity University, Noida, Uttar Pradesh, India
e-mail: skchauhan@amity.edu

A. Gupta
Department of Computer Science and Engineering, Vaish College of Engineering, Rohtak, Haryana, India
e-mail: ankurdujana@gmail.com

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

In this regard, we have first proposed adaptive genetic algorithm (AGA). Secondly, antenna selection is implemented by using firefly algorithm with LTE scheduling, then we have explored the new optimization algorithm based on Adaptive Spectrum Matching (ASM), a Joint Optimization-based scheduling method for a successful Multiple Input–Multiple Output communication is considered, finally, their characteristics, advantages, performance metrics, limitations have been studied in detail, and research gaps have been identified. Moreover, to demonstrate the effectiveness of developed antenna selection schemes and to validate the results, the comparison of various antenna selection schemes is performed.

Keywords LTE · Bit error rate (BER) · Multiple Input–Multiple Output (MIMO) · Optimization algorithms

1 Introduction

Current wireless systems like cellular mobile phones, Bluetooth, mobile low earth orbit satellite (LOS), etc. all require very high data rate (>100 Mbps), lower delay, transmission reliability, and wider coverage [1–3]. So, we can say that future wireless services demand high data rates and good quality of service [4]. But, the limitations are fading, limited available spectrum, and battery life of wireless portable devices [5, 6]. Transmission reliability in wireless channel is a stimulating topic. Several factors like thermal noise, time varying multipath fading, and power and bandwidth limitations harshly degrade the quality of transmission. In response, wireless communication technology has improved greatly in terms of spectral efficiency to satisfy the increasing demand of capacity and data rate with various techniques. Among them, MIMO systems and OFDM are the prime techniques to enhance the bit rate of various wireless communications systems. Antenna selection is a creating examination territory in MIMO system in the late years. MIMO systems have the most important benefit of providing higher efficiency with no need for the extra transmitted power or bandwidth expansion. However, because many antennas are utilized, it has one major disadvantage: it necessitates the purchase of additional high-cost RF modules [7]. In an attempt to minimize the price of the various radio frequency modules, a technique called antenna subset selection is suggested to reduce the complications associated with the hardware while preserving various advantages [8–10].

1.1 *Need of Antenna Selection*

This refers to the usage of only the finest group of antennas, while the other antennas are generally not used, thus minimizing the count of RF chains [11]. To overcome the same, there is requirement of various techniques for MIMO systems to find the optimal subset selection [12–14]. Bearing all these points in mind, people working in

this field can undoubtedly understand antenna selection technique to reduce PAPR, intercarrier interference, and criteria to remove unnecessary hardware and complexity to enhance overall system performance for our next-generation wireless communication systems. Consequently, choosing a proper optimum antenna selection technique which fulfills targeted QOS, improved capacity and transmission reliability requirements are still an open issue. Because of their substantial futuristic applications, the MIMO systems are integral part of the upcoming telecommunication systems. Since the concerned systems are only wireless systems, the introduction of wireless communication and need of antenna selection have been explained in Sect. 1.

The remaining sections are organized as follows. Next section discusses a short review on the work related to the proposed research. Section 3 explains the different types of optimization algorithms for antenna selection in MIMO Systems. Section 4 provides the comparative analysis of various optimization algorithms used in antennas selection. At last, the entire work has been concluded in Sect. 5.

2 Literature Survey

The author in [15] talked about how the four antennas' deployment in traditional Alamouti STBC did not allow for full advantage of group-wise features. The most challenging task in the STBC system is determining status (used and unused bandwidth).

The authors in [16] proposed an effective scheduling algorithm for heterogeneous MIMO networks related to Simulated Annealing (SA) and Particle Swarm (PS) techniques. The main goal of this method was to simplify user scheduling while also increasing the sum rate. The downlink transmission was investigated with coordinated multi-user multi-cell MIMO.

By assuring a favorable compromise between the cell average and user spectral efficiency, the authors in [17] answered the problem of assignment of resources in MIMO. The Channel State Information (CSI) was shared across the nodes in order to reduce interference and maintain a power level that was consistent with the desired signal. To tackle some sequence convex sub-problems, a Successive Convex Approximation (SCA) technique was devised.

In MIMO broadcast networks, the authors in [18] proposed a novel way to handle the optimization problem that are not convex. The transmitter was used to manage the quantity of power distributed among the harvesting users, and the Majorization–Minimization (MM) approach was used to improve the optimization's performance. The gathered power and information transport were optimized jointly in this work.

The authors in [19] looked at several methods based on the choice of antennas in order to discover the optimal answer by looking at all of the options. Based on an extensive search, the mean rate of the ideal antenna selection was examined. The goal of this strategy was to use an optimal selection procedure to reduce the complexity of searching. In addition, the complexity of antenna selection was studied using the greedy searching method.

To lower the searching complications of MIMO communication systems, the authors in [20] proposed an efficient metaheuristic approach called Genetic Algorithm (GA). In this paper, the binary GA approach was combined with user and antenna scheduling to provide an acceptable sum rate with lesser complexity.

Optimization research for the choice of antennas in MIMO Two-Way Relay Networks was proposed by the authors in [21]. By taking into account the issues of power allocation and relay position, it was able to reduce the likelihood of an outage. Furthermore, the MIMO system's outage performance was examined without the use of sophisticated derivations, which was a benefit of this study.

3 Optimization Algorithm for Antenna Selection in MIMO Systems

3.1 Adaptive Genetic Algorithm (AGA)

The fundamental problem with this system is that as the number of transmitting and receiving antennas grows, so does the convolution of the hardware. In this chapter, an efficient optimum transmitting antenna subset choice technique is recommended with the help of the Adaptive Mutation Genetic Algorithm to solve this problem (AGA). The adaptive mutation process of the genetic algorithm [22, 23] is used to choose transmit antenna subsets in the MIMO-OFDM scheme at this stage. The fitness value for each mutation point is determined, and the best fitness-based mutation points are chosen based on that value. The mutation operation is carried out after the best mutation points have been collected.

3.2 Firefly Optimization Algorithm with Adaptive Scheduling (FFOAS)

In a typical MIMO communication system, there will be increased bandwidth losses and potentially more noisy signals. To achieve this, we developed quasi orthogonal space time block codes with a firefly optimization technique for antenna selection and a modified LTE system for scheduling, in addition to an adaptive scheduling system, to improve our communication performance while in transmission [24]. The Q-OSTBC methodology is described as having a high degree of versatility for transmission schemes. The Firefly optimization process is used to select the best Rayleigh channel utilization range. By giving the optimum fitness function, it reduces bandwidth losses. By allocating the time of transmission between transmission and reception, LTE-based scheduling is used to exploit throughput constraints. To reduce the transmission time, the suggested FFOAS technique is combined with a spectrum management scheme.

To provide a high level of flexibility in the transmission system Q-OSTBC approach is employed. In MIMO wireless communication, the Firefly optimization approach is utilized to increase the performance rate. By assigning time to communicate over transmitter and receiver, LTE scheduling is used to improve throughput constraints. Spectrum management is used to reduce the communication latency between the transmitter and receiver.

3.3 Adaptive Spectrum Matching Based Optimal Scheduling (ASMOS)

To anticipate the best channel from the available bands, an adaptive spectrum matching technique is applied. In addition, prioritizing on the basis of high-spectrum intensity ensures that data are sent to the receiver efficiently. The mistake probability rates are minimized by arranging available channels and data prioritization. In terms of average block error probability and bit error rate, this study compares the effectiveness of proposed optimal channel utilization to various modulation techniques such as three-dimensional complementary codes and linear network coding with quadrature phase shift keying.

3.4 Joint Optimization and Sub-band Expediency-Based Scheduling Technique (JO-SES)

The major goal of this research is to lower the BER and boost the efficiency in this communication. The Cellular Network (CN) is initially created with a set of nodes, after which the communication channel is initialized and its parameters are extracted. The channel's parameters are then extracted using the beamforming feature extraction technique [25]. The optimization is done to choose the optimum channel [26] for the gainful communication from the available channels. The Power Spectral Density (PSD) of the specified channel is then calculated using the channel bandwidth's subcarrier frequency. For channel scheduling, the Sub-Band Expediency-Based Scheduling (SES) technique [27–29] is suggested, which reduces communication delays during transmission.

4 Comparative Analysis of Antenna Selection in MIMO Systems

The simulation study is done for 4×4 MIMO schemes, which include four transmitting antennas and four receiving antennas. The implementations are carried out

Table 1 Performance metrics for optimization algorithms

Parameter	Expression	Explanation
BER	$BER = \frac{N_e}{T_{bs}}$	Number of bit errors (N_e) divided by the total number of bits transferred (T_{bs})
Average SNR	$\bar{\gamma} \triangleq \int_0^{\infty} \gamma p_{\gamma}(\gamma) d\gamma$	γ denotes the instantaneous SNR and $p_{\gamma}(\gamma)$ denotes the probability density function (PDF) of γ

Table 2 Advantages and shortcomings of each algorithm

Algorithm	Advantages	Shortcoming
AGA	High data rate and higher capacity with the adaptive mutation process of genetic algorithm, complexity reduction of 2%	Convergence in local optima, it can provide significant optimal result
FFOAS	At maximum SNR, BER reduces to almost 92%	Requirement of additional antenna at the network base station
ASMOS	At maximum SNR, BER reduces to almost 21%	More number of iteration required during optimization
JO-SES	At maximum SNR, BER reduces to almost 1.75%	Channel information should be known at the transmitter

in order to verify the SNR and BER CDF derivations. When contrasted with standard OSTBC and TAS/MRC-based MIMO systems, an improved level of performance rate is obtained using both arrangements of antenna selection optimization and scheduling approach as far as minimum slope of CDF, less BER, and reduced time complexity are concerned. The performance metrics [30–33] of optimization algorithms are tabulated in Table 1. Table 2 highlights the advantages and shortcomings of each algorithm. Table 3 represents the comparative analysis for all algorithms. Table 4 shows time complexity reduction of various algorithms.

Figure 1 shows comparative analysis of BER at minimum SNR (0 dB) for each algorithm, and Fig. 2 shows comparative analysis of BER at maximum SNR (15 dB) for each algorithm; and from the graphical analysis, it is observed that Joint Optimization and Sub-Band Expediency-Based Scheduling technique (JO-SES). Figure 3 shows time complexity reduction of each optimization algorithm. It has been found that time complexity reduction is also better for JO-SES as compared to ASMOS, FFOAS, and AGA.

5 Conclusion

Investigations on various optimization techniques for the choice of antennas in MIMO systems are reported in this paper. Research initiatives are mainly focused toward the capacity and average sum rate enhancement [34–37]. There are two main approaches

Table 3 Comparative analysis of BER and SNR for various algorithms

Algorithms	BER at minimum SNR (0 dB)	BER at maximum SNR (15 dB)
AGA	1.1E-1	6E-3
TAS/MRC-QAM	9E-1	5E-5
TAS/MRC-QPSK	2E-1	8E-5
OSTBC-QAM	9E-1	5E-2
OSTBC-OPSK	E-1	3E-5
FFOAS-QAM	8E-2	5E-5
FFOAS-QPSK	2E-3	8E-6
LNC	0.091	4.3E-5
3-DCC	0.089	2E-5
ASTBC	0.0865	1.4E-5
ASMOS	0.085	1.14E-5
JO-SES	0.080	1.12E-5

Table 4 Time complexity reduction of optimization algorithms

Algorithm	Time reduction complexity (%)
AGA	2
FFOAS	16.6
ASMOS	35.43
JO-SES	38.38

to meet this challenge: (i) BER minimization through antenna selection and (ii) scheduling optimization for throughput enhancement [38]. This paper deals with BER minimization problem by suggesting the new/improved antenna selection algorithms/schemes for the successful implementation of MIMO systems so that it may be applied in various types of future wireless communication networks and hence has a lot of future scope.

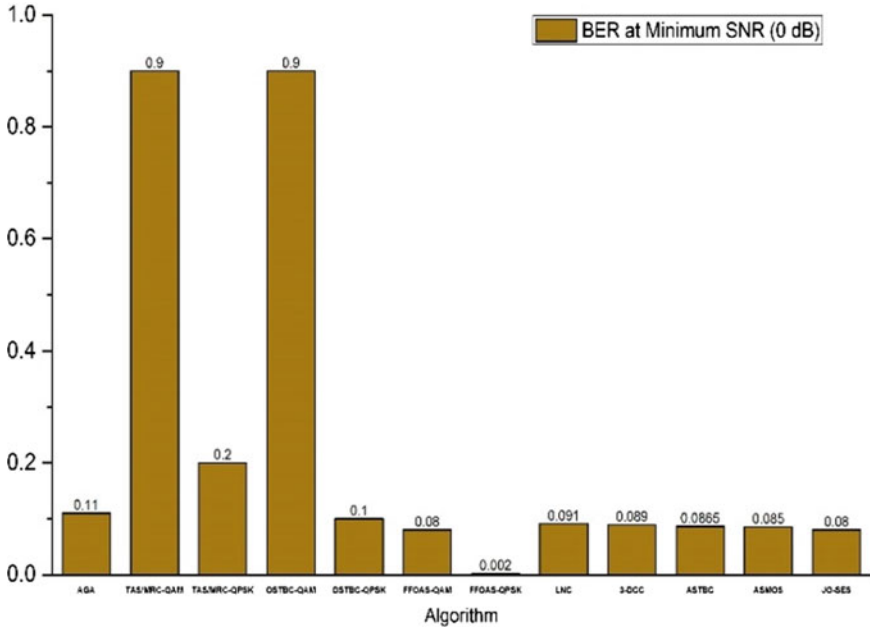


Fig. 1 Comparative analysis of BER at minimum SNR (0 dB) for each algorithm

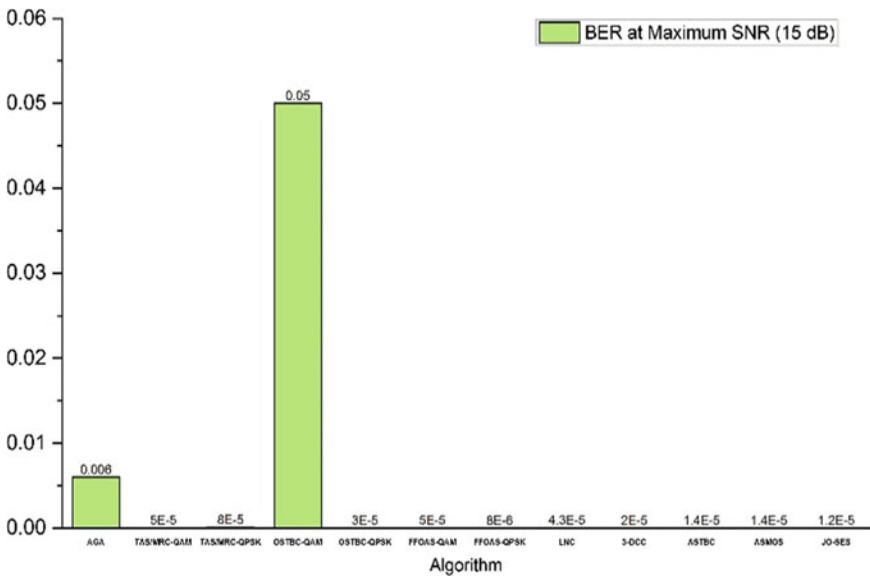
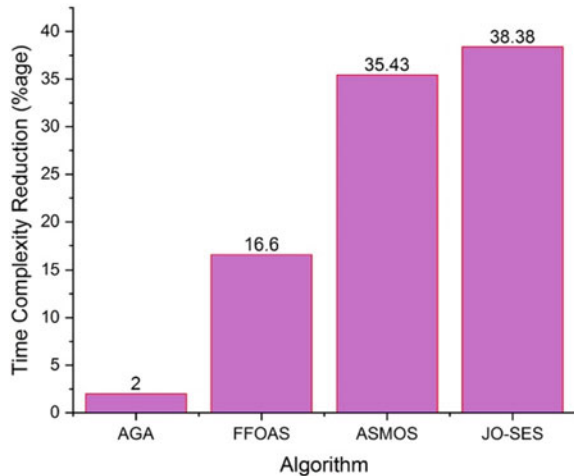


Fig. 2 Comparative analysis of BER at maximum SNR (15 dB) for each algorithm

Fig. 3 Time complexity reduction of each optimization algorithm



References

1. Juneja S, Juneja A, Anand R (2019) Reliability modeling for embedded system environment compared to available software reliability growth models. In: 2019 International conference on automation, computational and technology management (ICACTM), Apr 2019, pp 379–382. IEEE
2. Anand R, Arora S, Sindhwani N (2022) A miniaturized UWB antenna for high speed applications. In: 2022 International Conference on Computing, Communication and Power Technology (IC3P), Jan 2022, pp 264–267. IEEE
3. Dahiya A, Anand R, Sindhwani N, Deshwal D (2021) Design and construction of a low loss substrate integrated waveguide (SIW) for S band and C band applications. *Mapan* 36(2):355–363
4. Anand R, Chawla P (2020) A novel dual-wideband inscribed hexagonal fractal slotted microstrip antenna for C-and X-band applications. *Int J RF Microw Comput Aided Eng* 30(9):e22277
5. Garg P, Anand R (2011) Energy efficient data collection in wireless sensor network. *Dronacharya Res J* 3(1):41–45
6. Paliwal KK, Anand R, Garg P (2011) Energy efficient data collection in Wireless sensor network-a survey. In: International conference on advanced computing, communication and networks' II, pp 824–827
7. Cho YS, Kim J, Yang WY, Kang CG (2010) MIMO-OFDM wireless communication with MATLAB. Wiley (Asia) Pte. Ltd, ed
8. Jafarkhani H (2005) Space time coding theory and practice. Cambridge University Press
9. Jankiraman M (2004) Space time codes. Universal Personal Communication
10. Simon MK, Alouini MS (2004) Digital communication over fading channels. Wiley Interscience
11. Di Renzo M, Haas H, Ghrayeb A, Sugiura S, Hanzo L (2014) Spatial modulation for generalized MIMO: challenges, opportunities, and implementation. *Proc IEEE* 102:56–103
12. Sindhwani N, Singh M (2014) Comparison of adaptive mutation genetic algorithm and genetic algorithm for transmit antenna subset selection in MIMO-OFDM. *Int J Comput Appl* 97(22)
13. Sindhwani N (2017) Performance analysis of optimal scheduling based firefly algorithm in MIMO system. *Optimization* 2(12):19–26
14. Sindhwani N, Bhamrah MS (2017) An optimal scheduling and routing under adaptive spectrum-matching framework for MIMO systems. *Int J Electron* 104(7):1238–1253

15. Liu TH (2015) Analysis of the Alamouti STBC MIMO system with spatial division multiplexing over the Rayleigh fading channel. *IEEE Trans Wirel Commun* 14(9):5156–5170
16. Purmehdi H, Elliott RC, Krzymien WA (2016) Reduced-complexity user scheduling algorithms for coordinated heterogeneous MIMO networks. *IEEE Trans Veh Technol* 65:6184–6203
17. Mosleh S, Liu L, Zhang J (2016) Proportional-fair resource allocation for coordinated multi-point transmission in LTE-advanced. *IEEE Trans Wirel Commun* 15:5355–5367
18. Rubio J, Iserte AP, Palomar DP, Goldsmith A (2017) Joint optimization of power and data transfer in multiuser MIMO systems. *IEEE Trans Sig Process* 65:212–227
19. Jang S, Ahn M, Lee H, Lee I (2016) Antenna selection schemes in bidirectional full-duplex MIMO systems. *IEEE Trans Veh Technol* 65:10097–10100
20. Pastor-Pérez J, Palou FR, Femenias G (2016) Evolutionary algorithms for multiobjective optimization of frequency reuse schemes in CoMP-based MIMO-OFDMA networks. In: *International symposium on wireless communication systems (ISWCS)*, pp 235–241
21. Erdogan E, Afana A, Gucluoglu T, Ikki S (2016) Unified analysis and optimization study of antenna selection in two-way relay networks. *IEEE Wirel Commun Lett* 5:216–219
22. Sindhwani N, Bhamrah MS, Garg A, Kumar D (2017) Performance analysis of particle swarm optimization and genetic algorithm in MIMO systems. In: *2017 8th International conference on computing, communication and networking technologies (ICCCNT)*, July 2017, pp 1–6. IEEE
23. Anand R, Chawla P (2016) A review on the optimization techniques for bio-inspired antenna design. In: *2016 3rd International conference on computing for sustainable global development (INDIACom)*, Mar 2016, pp 2228–2233. IEEE
24. Sindhwani N, Singh M (2016) FFOAS: antenna selection for MIMO wireless communication system using firefly optimisation algorithm and scheduling. *Int J Wirel Mobile Comput* 10(1):48–55
25. Anand R, Sindhwani N, Dahiya A (2022) Design of a high directivity slotted fractal antenna for C-band, X-band and Ku-band applications. In: *2022 9th International Conference on Computing for Sustainable Global Development (INDIACom)*, Mar 2022, pp 727–730. IEEE
26. Meelu R, Anand R (2011) performance evaluation of cluster-based routing protocols used in heterogeneous wireless sensor networks. *Int J Inf Theory Knowl Manag* 4(1):227–231
27. Sindhwani N, Singh M (2020) A Joint Optimization based sub-band expediency scheduling technique for MIMO communication system. *Wirel Pers Commun* 115(3):2437–2455
28. Akila D, Bhaumik A, Duraisamy B, Suseendran G, Pal S (2021) Improved nature-inspired algorithms in cloud computing for load balancing. In: *Intelligent computing and innovation on data science*, pp 547–558. Springer, Singapore
29. Chawla P, Anand R (2017) Micro-switch design and its optimization using pattern search algorithm for applications in reconfigurable antenna. *Mod Antenna Syst*, 10
30. Tuan NA, Akila D, Pal S, Sarkar B, Tran TK (2021) Data optimization in IoT-assisted sensor networks on cloud platform
31. Gupta A et al (2019) Script classification at word level for a multilingual document. *Int J Adv Sci Technol* 28(20):1247–1252. Retrieved from <http://serisc.org/journals/index.php/IJAST/article/view/3835>
32. Anand R, Chawla P (2022) Bandwidth optimization of a novel slotted fractal antenna using modified lightning attachment procedure optimization. In: *Smart antennas*, pp 379–392. Springer, Cham
33. Sindhwani N, Rana A, Chaudhary A (2021) Breast cancer detection using machine learning algorithms. In: *2021 9th International conference on reliability, Infocom technologies and optimization (trends and future directions) (ICRITO)*, pp 1–5. IEEE
34. Dahiya A, Anand R, Sindhwani N, Kumar D (2022) A novel multi-band high-gain slotted fractal antenna using various substrates for X-band and Ku-band applications. *Mapan* 37(1):175–183
35. Meivel S, Sindhwani N, Valarmathi S, Dhivya G, Atchaya M, Anand R, Maurya S (2023) Design and method of 16.24 GHz microstrip network antenna using underwater wireless communication algorithm. In: *Cyber technologies and emerging sciences*, pp 363–371. Springer, Singapore

36. Anand R, Sindhwani N, Juneja S (2022) Cognitive internet of things, its applications, and its challenges: a survey. In: *Harnessing the internet of things (IoT) for a hyper-connected smart world*, pp 91–113. Apple Academic Press
37. Juneja S, Anand R (2018) Contrast enhancement of an image by DWT-SVD and DCT-SVD. In: *Data engineering and intelligent computing: proceedings of IC3T 2016*. Springer Singapore, pp 595–603
38. Meelu R, Anand R (2010, November) Energy efficiency of cluster-based routing protocols used in wireless sensor networks. In: *AIP conference proceedings*, vol 1324, no 1. American Institute of Physics, pp 109–113

Methodology for Classifying Objects in High-Resolution Optical Images Using Deep Learning Techniques



P. Lalitha Kumari, Santanu Das, B. Kannadasan, Niranjana Sampathila, C. Saravanakumar, Rohit Anand, and Ankur Gupta

Abstract The classification of objects that are present in the images or in the videos is being developed progressively to obtain good results because of the use of convolutional networks. In this work, we have used the convolutional networks for the detection of objects that are present in high-resolution satellite images. Tests were carried out on ships that are on the high seas and in the ports. This classification is useful for monitoring the coasts, as well as for analyzing the dynamics of the ships which can be applied in the search of ships. To cover this task of classifying ships in the spectral images, the use of high-resolution satellite images of coastal areas and with a large number of ships is used in order to build a set of images,

P. Lalitha Kumari

Department of Computer Science and Engineering, Malla Reddy Institute of Technology, Maisammaguda, Secunderabad, Andhra Pradesh, India

S. Das

Department of Biotechnology, Seshadripuram First Grade College, Yelahanka New Town, Bangalore, India

B. Kannadasan

Department of Civil Engineering, B.S. Abdur Rahman Crescent Institute of Science and Technology, Vandalur, Chennai, Tamil Nadu, India

N. Sampathila

Department of Biomedical Engineering, Manipal Institute of Technology, Manipal, Karnataka, India

e-mail: niranjana.s@manipal.edu

C. Saravanakumar

Department of Electronics and Communication Engineering, SRM Valliammai Engineering College, Kattankulathur, Tamilnadu, India

R. Anand (✉)

DSEU, G.B.Pant Okhla-1 Campus, New Delhi, India

e-mail: roh_anand@rediffmail.com

A. Gupta

Department of Computer Science and Engineering, Vaish College of Engineering, Rohtak, Haryana, India

e-mail: ankurdujana@gmail.com

containing images of the ships. In order to be used for training setting and testing of the convolutional network, a very particular configuration of the convolutional network caused by the particularity of high-resolution satellite images is presented. The methodology developed indicating the procedures performed is also presented in which a set of images containing 300 was built images of ships that are in the sea or are anchored in the ports. The results obtained in the classification using the convolutional networks are acceptable to be able to be used in different applications.

Keywords Convolutional networks · Satellite image · Classification · High resolution · Multispectral image

1 Introduction

Convolutional networks are a very interesting variation of the artificial neural networks. They are currently being widely used because of the availability of hardware that can be counted in these times. The best known convolutional networks are convolutional neural networks (CNNs) [1–3].

We find jobs in many applications, such as in the extraction of automatic features from multispectral images, where CNN is used to label and identify palm cultivation units, resulting in a map of characteristics [4]. We find works where CNN and virtual reality are combined using many data sources called multimodal data, where CNNs are used for image classification and virtual reality to create a 4D model [5]. In the classification of objects present in images and videos such as pets using CNN, there is obtained a classification level of 90% [6]. Works are also presented, where images that are found in large volumes of databases are retrieved using CNN. The results are better compared to conventional neural networks [7]. CNNs are also used in vehicle detection, using the videos produced by surveillance cameras, achieving acceptable results [8].

The visual interpretation for the mapping of land use and land cover is commonly done using temporary satellite data, for which a network based on SegNet has been implemented, where four different classes of ground cover have been classified: water bodies, forest lands, croplands, and buildings with a result of 0.84, considering as a good classifier [9].

Satellite systems provide heterogeneous data, which are used as inputs in the implementation of convolutional neural networks [10]. We also have many different works where the areas with the presence of water and ice are analyzed, through which a spectral analysis is carried out with a statistical approach by applying neural networks and convolutional networks [11, 12].

If it is required to work with more bands, we find the works where 13 spectral bands of the Sentinel-2 satellite are used, for which a new set of ten classes is used, with a total of 27,000 images that have been evaluated by deep convolutional neural networks (CNNs). This new dataset has achieved an accuracy of 98.57% [13, 14]. To improve the performance of convolutional networks, there is used a complement

with the classifiers based on random forest [15], where they have been tested on four different sets of data including spectral bands and transformations of principal components, after the transformation of the original images as input data considered for extraction [16–20].

2 Materials and Methods

2.1 *Satellite Image*

A satellite image is a representation of data in the form of a matrix formed by the optical instrument that is on-board in space missions. This data matrix will depend on the resolution and characteristics of the optical instrument. They are formed by spectral bands, so they take the name of multispectral images. The various bands are: the visible bands, red, green, and blue.

Another characteristic of satellite images is spatial resolution; this resolution is characterized by the equivalence of the pixel and its corresponding representation on the ground. This characteristic determines that optical instruments are classified into metric and sub-metric.

The images used are in the range of sub-metrics so that in the image, objects can be observed. The ability to distinguish them will depend on the experience of the evaluator and the conditions of image acquisition having a disadvantage of the coverage in each acquisition. Sub-metric images [21–23] have an approximate coverage of between 14 km of average width and approximately 90 km longer, compared to metric images that are 60 km wide and approximately 300 km long.

In Fig. 1, it can be seen that the images are in the combination of natural color, formed by the bands of the visual spectrum and ordered in the order of red, green, and blue.

2.2 *High-Resolution Satellite Image*

When we work with sub-metric images, a pixel corresponds to less than one meter on land, managing to observe the objects that on average have a dimension of 5 m, represented in the image between 6 and 7 pixels. By this characteristic, the image is called high-resolution image, because the image does not degrade as it becomes wide. Another important feature that the image has is its size, on average; an original image is between 3 gigabytes. The larger the length [24–26] of the image, larger the size in gigabytes. The high-resolution image is shown in Fig. 2.

In Fig. 2, we can see the result of performing the zoom process. We can also see that the image does not lose resolution and objects present in the image begin to



Fig. 1 Sub-metric satellite image viewed from its original size



Fig. 2 Sub-metric satellite image to view the realizer of the zoom process

appear. In the image, we can see that we can distinguish roads, parks, and vegetation areas grouped by green.

The image does not lose its resolution if we continue to perform the zoom process. When the zoom process is performed, objects within the image begin to discriminate as shown in Fig. 3.

In Fig. 3, it is obtained by zooming to the image, where we can see that we can distinguish the houses, as well as the sea and objects that are on its coasts. By increasing the zoom, the detail of the objects improves, as we can appreciate in Fig. 4.



Fig. 3 Sub-metric satellite image to view the realizer of the zoom process increased



Fig. 4 Sub-metric satellite image to distinguish objects when performing the zoom process

In Fig. 4, it can be seen that the detail of the objects in the image improves. Also by increasing the zoom in the image, it is not distorted. We can see the boats that are on the coast of the sea and the houses present in the image.

2.3 Convolutional Networks

The convolutional networks are a variant to the neural networks, with the characteristic of having many internal layers that perform a specific task to be able to improve

Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, 21, 28, 32)	896
leaky_re_lu_1 (LeakyReLU)	(None, 21, 28, 32)	0
max_pooling2d_1 (MaxPooling2D)	(None, 11, 14, 32)	0
dropout_1 (Dropout)	(None, 11, 14, 32)	0
flatten_1 (Flatten)	(None, 4928)	0
dense_1 (Dense)	(None, 32)	157728
leaky_re_lu_2 (LeakyReLU)	(None, 32)	0
dropout_2 (Dropout)	(None, 32)	0
dense_2 (Dense)	(None, 2)	66
Total params: 158,690		
Trainable params: 158,690		
Non-trainable params: 0		

Fig. 5 Convolutional network model created

in the classification. The convolutional network [27–29] is characterized by basic layers that are described below:

- Convolution.
- Pooling.
- Flatten.
- Dense.

The development mechanism uses Python programming language, with the Keras and TensorFlow libraries. Convolutional network model created is shown in Fig. 5.

2.4 Dataset Creation

The proposal indicates that we must create an image dataset, with as many images as possible. As a premise, we can indicate that increasing the number of images in the dataset tends to improve the performance.

The procedure for creating the image base corresponds to manual work. For this purpose, the greatest amount of high-resolution satellite images, containing information corresponding to the ships, is used. The job is to be able to detect and classify these objects in the original images. One of the characteristics of the investigation is to be able to perform the detection first on ships and then to continue with



Fig. 6 Images of the created dataset

other objects, achieving a multi-classifier, using the different techniques provided by Artificial Intelligence [30–32].

One recommendation is to increase the number of images in the dataset so that we can use the data growth technique, with the variants such as rotation and size enhancement. Images of the created dataset are shown in Fig. 6.

2.5 Proposed Methodology

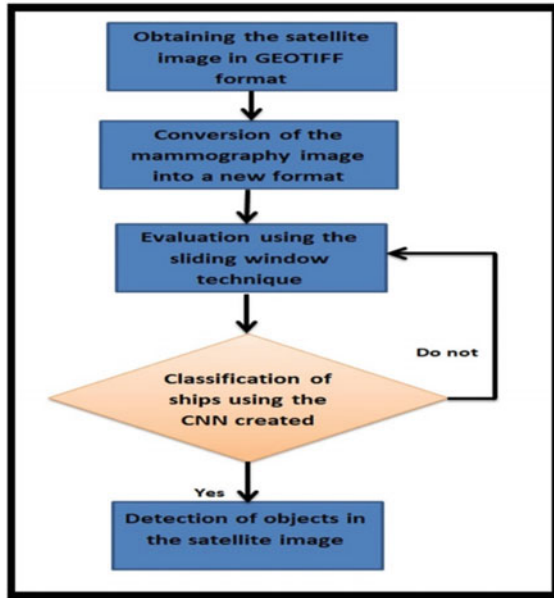
The proposed methodology is related to the reading and detection of ships present in the high-resolution satellite images. An image bank with tagged ships has been implemented. The first thing to consider is that the satellite image has the original format. The working format of satellite images is GEOTIFF. So, it is very difficult to open such files. The proposed procedure consists of converting the satellite image that is in 16-bit GEOTIFF format into an 8-bit BMP image. This conversion is necessary because the networks work with 8-bit images. With the images in BMP format, we proceed to crop and create the bank of images that correspond to the boats. With this image bank, the classification is carried out using convolutional networks [33–35]. The technique used to navigate the original image is “Sliding Window”. The flow chart of the proposal is shown in Fig. 7.

3 Results

The results obtained at the time of the detection of the ships, using the convolutional network created, using the images from our database created show that the accuracy is 0.94 that determines the network performance.

In Fig. 8, it can be seen that the convolutional network created allows to obtain a 94% performance in a first classification, using the 300 images of our dataset. For this analysis, 150 images for training are used and the other 150 images to perform the tests. This choice of the 150 images is made randomly, without using any selection criteria.

Fig. 7 Flow chart of the proposal



```
32/36 [=====>...] - ETA: 0s[ ] - 0s 244us/step  
36/36 [=====] - 0s 244us/step  
( 'Test loss:', 0.229321559270223)  
( 'Test accuracy:', 0.9444444444444444)
```

Fig. 8 Registration values in a normal situation

4 Conclusion

While most of the work is concerned with high-resolution satellite images, we have worked with special images, because the objects to be analyzed are immersed in a super image due to their large size and number of pixels. The biggest technical problem that can arise is related to the resolution of the image and its original format; hence, it is necessary to convert to a commercial format such as BMP at 8-bit resolution so that it is compatible with the libraries of the convolutional networks. Authors have performed the classification through the use of artificial intelligence through the implementation of convolutional networks in different types of images for different applications. To improve the classification, the structure of the convolutional network can be improved, increasing the number of layers. Among the recommended layers, different kinds of filters (such as the medium filter) are used to improve the resolution and presentation of the image [36]. Also, layers with edge detection filters are used to separate the objects. Finally, we can conclude that with the increase in the number of images in the dataset, the probability of improving the result of the classification

as well as improving the structure of the network increases through trial and error analysis.

References

1. Singh SK, Thakur RK, Kumar S, Anand R (2022, March) Deep learning and machine learning based facial emotion detection using CNN. In: 2022 9th international conference on computing for sustainable global development (INDIACom). IEEE, pp 530–535
2. Sindhwani N, Anand R, Shukla R, Yadav M, Yadav V (2021) Performance analysis of deep neural networks using computer vision. *EAI Endorsed Trans Ind Netw Intell Syst* 8(29):e3
3. Tiwari I, Juneja S, Juneja A, Anand R (2020) A statistical-oriented comparative analysis of various machine learning classifier algorithms. *J Nat Remedies* 21(3S1):139–144
4. Peña A, Bonet I, Manzur D, Góngora M, Caraffini F (2019, June) Validation of convolutional layers in deep learning models to identify patterns in multispectral images. In: 2019 14th Iberian conference on information systems and technologies (CISTI). IEEE, pp 1–6
5. Maiwald F, Bruschke J, Lehmann C, Niebling F (2019) A 4D information system for the exploration of multitemporal images and maps using photogrammetry, web technologies and VR/AR. *Virtual Archaeol Rev* 10(21):1–13
6. Weinstein BG (2018) Scene-specific convolutional neural networks for video-based biodiversity detection. *Methods Ecol Evol* 9(6):1435–1441
7. Sánchez Santiesteban S (2018) Recuperación de imágenes por contenido usando descriptores generados por Redes Neuronales Convolucionales. *Revista Cubana de Ciencias Informáticas* 12(4):78–90
8. Laura Riveros ER (2018) Detección de vehículos con aprendizaje profundo en Cámara de Vigilancia
9. Sathyanarayanan D, Anudeep DV, Das CAK, Bhanadarkar S, Uma D, Hebbar R, Raj KG (2020, December) A multiclass deep learning approach for LULC classification of multispectral satellite images. In: 2020 IEEE India geoscience and remote sensing symposium (InGARSS). IEEE, pp 102–105
10. Arndt J, Lunga D (2020) Sampling subjective polygons for patch-based deep learning land-use classification in satellite images. In: IGARSS 2020-2020 IEEE international geoscience and remote sensing symposium. IEEE, pp 1953–1956
11. Poliyapram V, Imamoglu N, Nakamura R (2019, July) Deep learning model for water/ice/land classification using large-scale medium resolution satellite images. In: IGARSS 2019–2019 IEEE international geoscience and remote sensing symposium. IEEE, pp 3884–3887
12. Rana A, Dhiman Y, Anand R (2022, January) Cough detection system using TinyML. In: 2022 international conference on computing, communication and power technology (IC3P). IEEE, pp 119–122
13. Helber P, Bischke B, Dengel A, Borth D (2019) Eurosat: a novel dataset and deep learning benchmark for land use and land cover classification. *IEEE J Sel Topics Appl Earth Observ Rem Sens* 12(7):2217–2226
14. Molina PC, Castro MP, Anjos CS (2020, March) Assessment of PCA and MNF influence in the VHR satellite image classifications. In: 2020 IEEE Latin American GRSS and ISPRS remote sensing conference (LAGIRS). IEEE, pp 143–148
15. Gupta A, Anand R, Pandey D, Sindhwani N, Wairya S, Pandey BK, Sharma M (2021) Prediction of breast cancer using extremely randomized clustering forests (ERCF) technique: prediction of breast cancer. *Int J Distrib Syst Technol (IJ DST)* 12(4):1–15
16. Lapini A, Fontanelli G, Pettinato S, Santi E, Paloscia S, Tapete D, Cigna F (2020, September) Application of deep learning to optical and SAR images for the classification of agricultural areas in Italy. In: IGARSS 2020–2020 IEEE international geoscience and remote sensing symposium. IEEE, pp 4163–4166

17. Antropov O, Rauste Y, Šćepanović S, Ignatenko V, Lönnqvist A, Praks J (2020, February) Classification of wide-area SAR mosaics: deep learning approach for Corine based mapping of Finland using multitemporal sentinel-1 data. In: IGARSS 2020–2020 IEEE international geoscience and remote sensing symposium. IEEE, pp 4283–4286
18. Shinde RC, Potnis AV, Durbha SS, Andugula P (2019, July) Compressive sensing based reconstruction and pixel-level classification of very high-resolution disaster satellite imagery using deep learning. In: IGARSS 2019–2019 IEEE international geoscience and remote sensing symposium. IEEE, pp 2639–2642
19. Deepa R, Velnath R, Guhan EH, Moorthy C, Gomathi P, Dinesh A (2021, October) Stability analysis of ball and beam system using PID controller. In: 2021 international conference on advancements in electrical, electronics, communication, computing and automation (ICAECA). IEEE, pp 1–4
20. Zhang L, Chen Z, Wang J, Huang Z (2018, December) Rocket image classification based on deep convolutional neural network. In: 2018 10th international conference on communications, circuits and systems (ICCCAS). IEEE, pp 383–386
21. Sreekanth N, Rama Devi J, Shukla A, Mohanty DK, Srinivas A, Rao GN, Gupta A (2022) Evaluation of estimation in software development using deep learning-modified neural network. *Appl Nanosci* 1–13
22. Anand R, Shrivastava G, Gupta S, Peng SL, Sindhvani N (2018) Audio watermarking with reduced number of random samples. In: Handbook of research on network forensics and analysis techniques. IGI Global, pp 372–394
23. Kumar R, Anand R, Kaushik G (2011) Image compression using wavelet method and SPIHT algorithm. *Digit Image Process* 3(2):75–79
24. Gupta A, Verma A, Kaushik D, Garg M (2020) Applying deep learning approach for brain tumor detection. *Mater Today Proc.* <https://doi.org/10.1016/j.matpr.2020.10.063>
25. Garg M, Gupta A, Kaushik D, Verma A (2020) Applying machine learning in IoT to build intelligent system for packet routing system. *Mater Today Proc.* <https://doi.org/10.1016/j.matpr.2020.09.539>
26. Verma A, Gupta A, Kaushik D, Garg M (2021) Performance enhancement of IOT based accident detection system by integration of edge detection. *Mater Today Proc.* <https://doi.org/10.1016/j.matpr.2021.01.468>
27. Gupta A, Garg M, Verma A, Kaushik D (2020) Implementing lossless compression during image processing by integrated approach. *Mater Today Proc.* <https://doi.org/10.1016/j.matpr.2020.10.052>
28. Gupta A (2019) Script classification at word level for a multilingual document. *Int J Adv Sci Technol* 28(20):1247–1252. Retrieved from <http://sersc.org/journals/index.php/IJAST/article/view/3835>
29. Singh H, Ramya D, Saravanakumar R, Sateesh N, Anand R, Singh S, Neelakandan S (2022) Artificial intelligence based quality of transmission predictive model for cognitive optical networks. *Optik* 257:168789
30. Gupta A (2020) An analysis of digital image compression technique in image processing. *Int J Adv Sci Technol* 28(20):1261–1265. Retrieved from <http://sersc.org/journals/index.php/IJAST/article/view/3837>
31. Kamalraj R, Neelakandan S, Kumar MR, Rao VCS, Anand R, Singh H (2021) Interpretable filter based convolutional neural network (IF-CNN) for glucose prediction and classification using PD-SS algorithm. *Measurement* 183:109804
32. Shukla R, Dubey G, Malik P, Sindhvani N, Anand R, Dahiya A, Yadav V (2021) Detecting crop health using machine learning techniques in smart agriculture system. *J Sci Ind Res (JSIR)* 80(08):699–706
33. Gupta M, Anand R (2011) Image compression using set of selected bit planes on basis of intensity variations. *Dronacharya Res J* 3(1):35–40
34. https://doi.org/10.1007/978-3-030-96308-8_87

35. Chawla P, Juneja A, Juneja S, Anand R (2020) Artificial intelligent systems in smart medical healthcare: current trends. *Int J Adv Sci Technol* 29(10):1476–1484
36. https://doi.org/10.1007/978-981-10-3223-3_57

PID Control and Estimation of the Attitude Model Applied to Geostationary Satellites



Radhey Shyam Meena, Rongeet Talukdar, B. Kannadasan, G. Nageswara Rao, Kumar A. Shukla , Rohit Anand, and Ankur Gupta

Abstract The estimation of systems is based on the analysis of data obtained after experimenting about the behavior of a real process in response to an external disturbance to find a scheme that allows the mathematical stages to be modeled. This is the importance of its in-depth study, since if it is carried out correctly, re-directing resources to other fields is one of the many benefits of its applications. This research work seeks to highlight the importance of correcting the displacements that satellites in orbit may experience, due to the influence of various agents, both of their own functioning and those generated by the environment in which they operate. That is why, the technique of systems estimation should be presented as a very useful option

R. S. Meena

Department of Industrial and Management Engineering, Indian Institute of Technology, Kanpur, India

R. Talukdar

Alma Mater Studiorum—Università di Bologna, Bologna, Italy

e-mail: rongeettalukdar07@gmail.com

B. Kannadasan

Department of Civil Engineering, B. S. Abdur Rahman Crescent Institute of Science and Technology, Vandalur, Chennai, Tamil Nadu, India

G. Nageswara Rao

EEE Department, Lakireddy Bali Reddy College of Engineering, Mylavaram, India

JNTUK, Kakinada, Andhra Pradesh, India

K. A. Shukla

School of Computer Science and Engineering, Vellore Institute of Technology, Tamil Nadu, Vellore, India

e-mail: anik912345699@gmail.com

R. Anand (✉)

DSEU, G. B. Pant Okhla-1 Campus, New Delhi, India

e-mail: roh_anand@rediffmail.com

A. Gupta

Department of Computer Science and Engineering, Vaish College of Engineering, Rohtak, Haryana, India

e-mail: ankurdujana@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

V. V. S. S. Chakravarthy et al. (eds.), *Advances in Signal Processing, Embedded*

Systems and IoT, Lecture Notes in Electrical Engineering 992,

https://doi.org/10.1007/978-981-19-8865-3_56

because of the analysis tools of the MATLAB program. Putting-in orbit of a team of these characteristics implies investment of strong amounts of time, money, and technology; likewise maintaining its initial orientation leads to the analysis of new parameters without mentioning the great problem generated in the orbit of the Earth, if for reasons other than those considered “normal,” this orientation varied and the implemented processes did not fulfill their function. The necessary driver for our project is a PID control since a control law can compensate with the integral and derivative actions. The necessary tool for the development of our controller is from MATLAB. With this tool, we can reach our goal, since it shows us the trajectory of the roots as well as the necessary graphs showing control characteristics such as the maximum overshoot, stabilization time, and percentage of system error, among others.

Keywords Satellite · Control · Geostationary · Attitude model · PID control · MATLAB programming

1 Introduction

In order to develop a scheme that allows the phases of a real process to be mathematically described, system estimation relies on analysis of data gathered from experiments on the behavior of the process in response to an external perturbation. This highlights the need of doing in-depth research into the topic, since one of the numerous advantages of its use is the ability to reallocate previously allocated resources to other areas of interest [1–5]. With the help of the analysis tools provided by the MATLAB program, this work aims to show how important it is to correct the displacements that satellites in orbit may experience due to the influence of various agents, both those inherent to their operation and those generated by the environment in which they operate [6–10]. Putting a team with these characteristics into orbit requires a significant investment of time, money, and technology. Similarly, keeping its initial orientation requires the analysis of new parameters, not to mention the huge problem caused in Earth’s orbit, if for reasons other than those considered “normal,” said orientation changes and the implemented processes fail to perform their function. Because our work requires a controller that can perform both integral and derivative operations, we have settled on the proportional, integral, and derivative (PID) control. Our controller was developed using MATLAB [11–15], which provided all of the essential development tools.

This research work will help us reach our goal since it provides us with crucial graphs displaying control properties like maximum overshoot, stabilization time, and system error.

2 Materials and Methods

In the development of the methodology, four main steps are presented. The overall design of a data center can be classified in four categories Tier I-IV each one presenting advantages and disadvantages related to power consumption and availability [16, 17]. In most cases, availability and safety issues yield redundant $N + 1$, $N + 2$ or $2N$ data center designs and this has a serious effect [18, 19] on power consumption. According to Fig. 1, a data center has the following main units.

2.1 Satellite Altitude

The main focus of our work is the study of systems identification methods, in addition to inquiring about the main control techniques, as well as the different signals that will be used to obtain the models. When it is necessary to know the behavior of a system under certain conditions and when faced with certain inputs [20], one can resort to experimentation on said system and the observation of its outputs. However, in many cases, the experimentation can be complex or even impossible to carry out, that makes it necessary to work with some type of representation that is close to reality, and which is known as a model.

The orientation of a satellite with respect to planet Earth is called attitude, since this is the focus of the ellipse, around which it rotates tirelessly (see Fig. 2). The coordinate system used to describe the position of the satellite’s spin axis is called the BAHN coordinates, which on Earth are governed by the latitude and longitude of our planet. For the attitude of a satellite to be ideal, it is necessary that these coordinates have a longitude of 180° and latitude of 0° , being practically impossible due to the external forces that act on it when it is in orbit. This is where the need arises

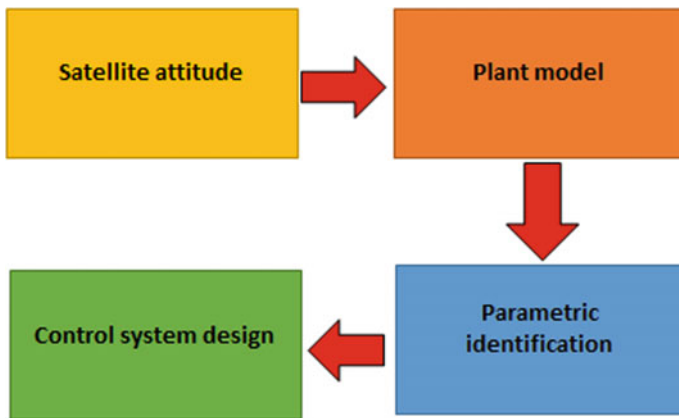
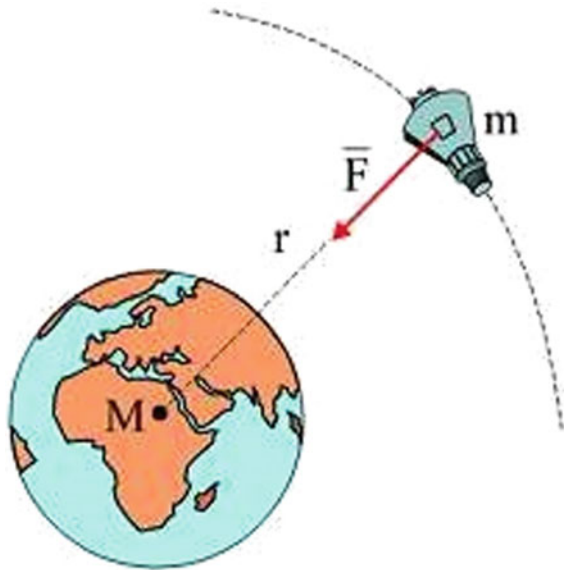


Fig. 1 Proposal block diagram

Fig. 2 Positioning a satellite

to implement a control method for this system, which is currently a fundamental piece for the development of humanity [21–23].

2.2 Plant Model

To start with the modeling of our virtual plant [24], we first need to have the differential equations which describe the motion of our system and we also need the values of all the time-invariant elements that belong to our system [25–27].

The system is based on a satellite in orbit, which will have attitude variations over time. It consists of three axes on which our satellite will rotate due to the disturbances represented with torques. The satellite will also have inertia with respect to each of the axes, and like any satellite, it will be at a certain height from the Earth’s surface orbiting at a respective angular velocity.

A basic satellite has the following moments of inertia:

$$I_x = 80[\text{kg m}^2] \quad (1)$$

$$I_y = 82[\text{kg m}^2] \quad (2)$$

$$I_z = 4[\text{kg m}^2] \quad (3)$$

With these values and for future abbreviations, we calculate:

$$\sigma_x = (I_y - I_z)/I_x = 0.975 \quad (4)$$

$$\sigma_y = (I_x - I_z)/I_y = 0.926 \quad (5)$$

$$\sigma_z = (I_y - I_x)/I_z = 0.5 \quad (6)$$

For a circular orbit with altitude 800 km, we have:

$$w_0 = 0.001038(\text{rad/seg}) \quad (7)$$

The disturbances that are expected for our system are around:

$$T_{dy} = 10^{-5}(\text{Nm}) \quad (8)$$

Our satellite system consists of a wheel damper control which consists of the following elements:

$$D = 0.002 \quad (9)$$

$$I_w = 1(\text{kg m}^2) \quad (10)$$

Here, D is the damping coefficient of the fluid in which the wheel is immersed. It is the moment of inertia of the wheel.

With all these data, we will proceed with independently analyzing the YB axis of the system, keeping the other two axes constant without dynamics. Simplified YB axis attitude dynamics equation for the plant design with respect to the tilt axis:

$$\begin{aligned} \theta(S) = (SI_w + D)[T_{dy}/I_y + S\theta(0) + \theta(0)] \\ + S(I_w/I_y)D\theta(0)I_w S^3 + D(I_w + 1)S^2 \\ 3I_w W^2 \sigma_y S + 3DW^2 \sigma_y \end{aligned} \quad (11)$$

Replacing the numerical values and assuming initial conditions equal to zero:

$$\theta(S) \frac{1.219 \times 10^{-7}}{(S^2 + 1.038 \times 10^{-5}S + 2.984 \times 10^{-6})} \quad (12)$$

The response of the open-loop system without considering the perturbation is shown in Fig. 3. A highly nonlinear response is observed.

The plant and the sensor add error to the data reading that we have as output from the entire system, the noise will be reproduced in Simulink (see Fig. 4). In this case, the sensor will be made up of white noise as input to a first-order transfer function that acts as a low-pass frequency filter [28–30] to color it to the noise with a

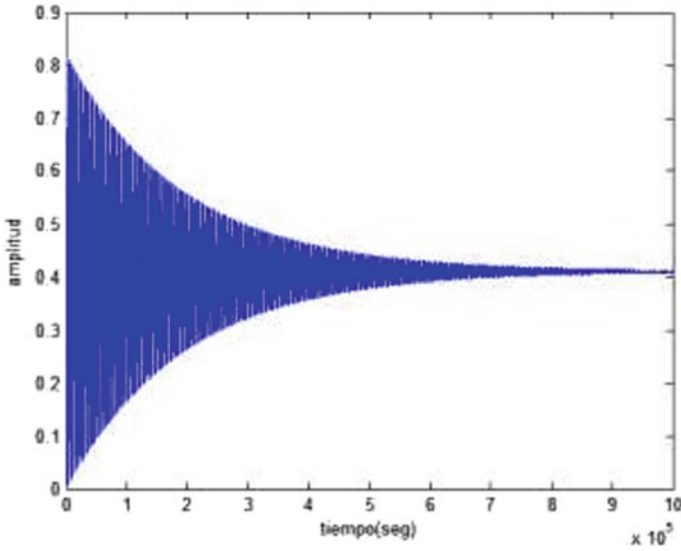


Fig. 3 Open-loop system response

cut-off frequency equal to that of the plant simulated previously and finally saturate its amplitude in a range from -0.01° to $+0.01^\circ$ since the actual sensor is accurate to 0.02.

The general transfer function for second-order systems when compared with the transfer function of the YB axis can solve for both the natural frequency and the relative damping factor (1).

$$G(s) = \frac{KW_n^2}{S^2 + 2\xi W_n S + W_n^2} \tag{13}$$

Comparing with the general polynomial form:

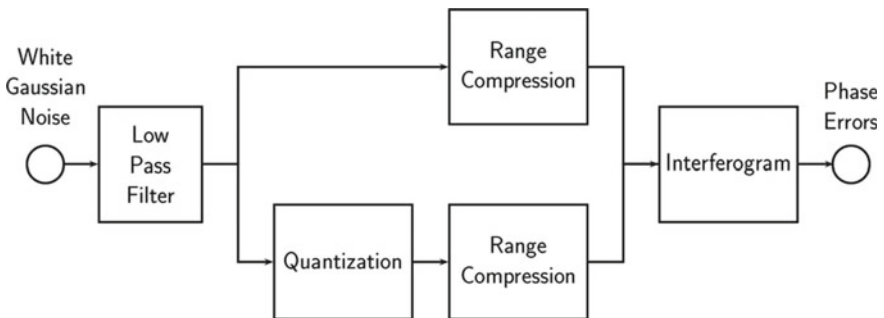


Fig. 4 Block diagram of the noisy estimation

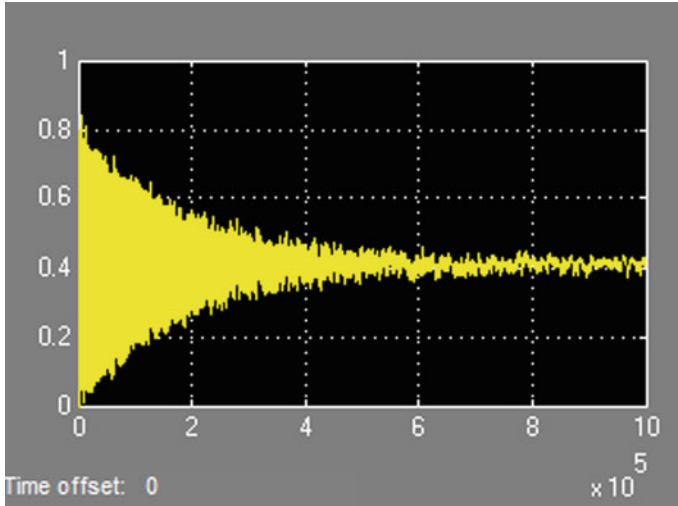


Fig. 5 System response to open loop with disturbance

$$\theta(s) \frac{1.219 \times 10^{-7}}{(s^2 + 1.038 \times 10^{-5} s + 2.984 \times 10^{-6})} \tag{14}$$

where we obtain the undamped angular frequency and the damping factor:

$$w_n = \sqrt{2.984 \times 10^{-6}} = 0.001727$$

$$\xi = 1.038 \times 10^{-5} / 2 \quad w_n = 0.003005$$

The equation of the first-order filter is:

$$F(s) = \frac{0.001727}{s + 0.001727} \tag{15}$$

The plant and the sensor add an error to the data reading that we have as an output from the entire system including the disturbance (see Fig. 5).

3 Results

3.1 Parametric Identification

Parametric identification with prediction error estimation is used to identify the transfer function and its order. The error prediction model structures are: ARX,

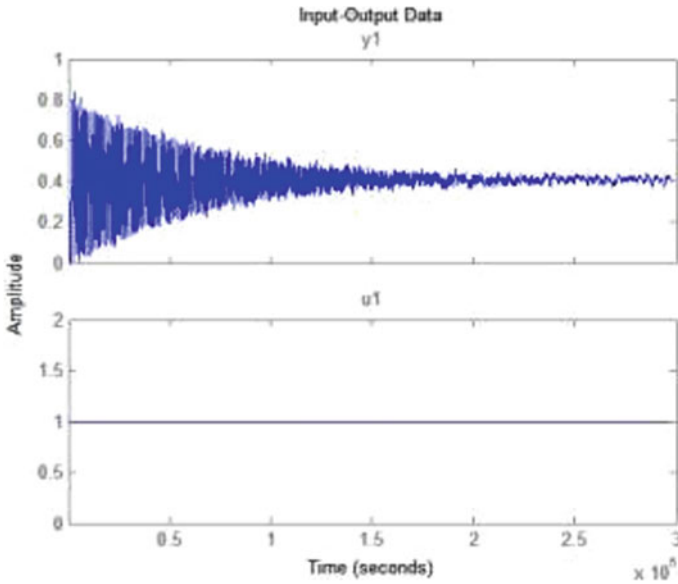


Fig. 6 Output and input signals with DATA

ARMAX, FIR, Error Output, and Box-Jenkins. To obtain the dynamic model from experiment, the input and output data is required, then proceed to plot (see Fig. 6). What is sought is to find the simplest model that has an adequate fit for the choice of the final structure of the model.

Working on a prototype model, we find the transfer function. MATLAB lines of code will allow obtaining the process model. MATLAB code returns the transfer function that has a far zero in the stable region.

We can look for an approximate transfer function that is a prototype as can be verified in the following equation:

$$G(s) = \frac{9.6188 \times 10^{-6}}{s^2 + 0.005236s + 2.395 \times 10^{-5}} \quad (16)$$

We verify the prototype compared with the plant of order 2 that has a finite zero. The answers are close to each other, which validates the transfer function (see Fig. 7).

3.2 Control System Design

Once we have carried out the nonparametric and parametric analysis and identified our system, the next step to follow is to set an objective that allows us to improve some aspect of the plant. Several aspects can be considered when choosing a controller

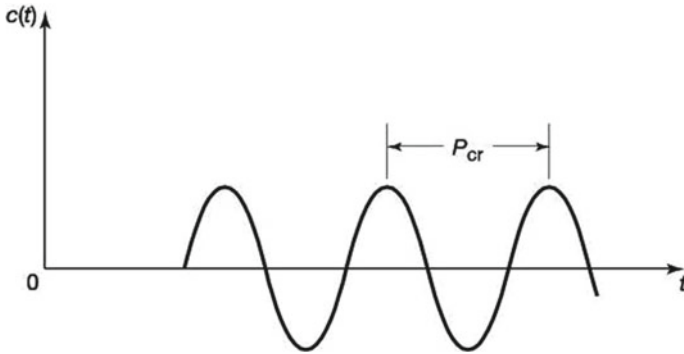


Fig. 7 Model response comparison

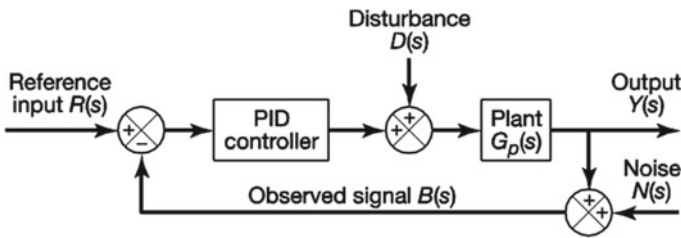


Fig. 8 Model response comparison

that optimizes the operation of a control system, among which we highlight stability, precision in steady state, response characteristics, and robustness. These parameters, depending on their application, give rise to analysis criteria that will be our basis for choosing the controller that best suits our identified plant. The choice of the controller must have a stable closed-loop behavior, that the effects of disturbances are minimized, that fast and smooth responses are obtained to changes in the set point and that the system be robust, that is, not very sensitive to changes in process conditions or due to errors. Figure 8 shows our control system where the design objective is to find a controller capable of solving the error in zero steady state and with minimal overshoot.

After analyzing all the points mentioned above, we propose the objective of our identified system: to reduce the stabilization time and reduce the percentage overshoot of 2%.

3.3 Result Controller PID

The values to be improved with the design of a controller: Stabilization Time: 196,000 (sec). We proceed to manipulate the trajectory of the roots until we obtain the desired

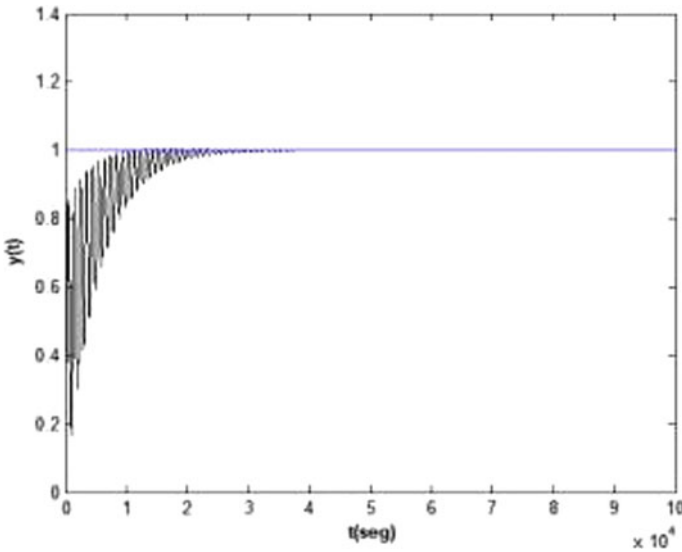


Fig. 9 Comparison of model responses

improvement in our system. With this built-in controller, a much faster stabilization time is obtained, which is necessary for a better performance in the functions of a satellite, thus improving the control of our system. In the resulting graph (see Fig. 7), we can see a great improvement in the stabilization time compared to the plant without a controller, from approximately 196,000 to 2570 s making it faster, its value of percentage over level at 78.5% which is lower although it is not close to the percentage level obtained with the identified model.

The results of the performance measurement with zero value derivative control action are shown in Fig. 9.

4 Conclusion

Performing identification in systems implies great costs due to the production stoppages that experimentation requires. Therefore, for academic purposes, it is helpful to work with a base mathematical model that represents the dynamics of the process. Through the tests carried out with the different parametric estimation models, we established that the autoregressive “ARMAX” model, a moving average with an external input of order $n_a = 2$, $n_b = 1$, $n_c = 2$ and $n_k = 1$ considered low and an approximation of 93.41% gives us the best satellite identification representation. It was shown that the identification process offers us an alternative for improvement by reducing our stabilization time and over percentage level by 98.68% and 24.52%, respectively. This allows us to make the attitude control system of a satellite more

efficient, showing that the application of this technique is not limited to industrial processes. Using formulas, it was determined that the ideal sampling period for our satellite system is 180 s, but taking into account, that increase in the sampling time results into less data to analyze, we decided to set this period to reduce to 500 s.

References

1. Wertz JR (1978) *Spacecraft attitude determination and control* (softback). Kluwer Academic Publishers
2. Davies (1970) W.D.T. *system identification for self-adaptive control*. Wiley-Interscience, London
3. Godfrey K (1993) *Perturbation signals for system identification*. Prentice-Hall
4. Ljung L (1987) *System identification, theory for the user*. Prentice Hall, Englewood Cliffs, N. J.
5. Kocaoğlu S, Kuşçu H (2013) Design and control of PID-controlled ball and beam system. In: *Unitech. international science conference Gabrovo*, pp 41–46
6. Rice RC (2010) *PID tuning guide—a best-practices approach to understanding and tuning PID controllers*
7. Mehedi IM, Al-Saggaf UM, Mansouri R, Bettayeb M (2019) Two degrees of freedom fractional controller design: application to the ball and beam system. *Measurement* 135:13–22
8. Ding M, Liu B, Wang L (2019) Position control for ball and beam system based on active disturbance rejection control. *Syst Sci Control Eng* 7(1):97–108
9. Ali AT, Ahmed AM, Almahdi HA, Taha OA, Naseraldeem A (2017) Design and implementation of ball and beam system using pid controller. *Automatic Control Inf Sci* 3(1):1–4
10. Aziz NS, Adnan R, Tajjudin M (2017) Design and evaluation of fuzzy PID controller for ball and beam system. In: *2017 IEEE 8th control and system graduate research colloquium (ICSGRC)*, pp 28–32
11. AG et al (2020) An analysis of digital image compression technique in image processing. *Int J Adv Sci Technol* 28(20):1261–1265. Retrieved from <http://sersc.org/journals/index.php/IJAST/article/view/3837>
12. AG et al. (2019) Script classification at word level for a multilingual document. *Int J Adv Sci Technol* 28(20):1247–1252. Retrieved from <http://sersc.org/journals/index.php/IJAST/article/view/3835>
13. Gupta A, Garg M, Verma A, Kaushik D (2020) Implementing lossless compression during image processing by integrated approach. *Mater Today Proc.* <https://doi.org/10.1016/j.matpr.2020.10.052>
14. Verma A, Gupta A, Kaushik D, Garg M (2021) Performance enhancement of IOT based accident detection system by integration of edge detection. *Mater Today Proc.* <https://doi.org/10.1016/j.matpr.2021.01.468>
15. Aggarwal B, Gupta A, Goyal D, Gupta P, Bansal B, Barak D (2021) A review on investigating the role of block-chain in cyber security. *Mater Today Proc.* <https://doi.org/10.1016/j.matpr.2021.10.124>
16. Kaushik D, Gupta A (2021) Ultra-secure transmissions for 5G-V2X communications. *Mater Today Proc.* <https://doi.org/10.1016/j.matpr.2020.12.130>
17. Garg M, Gupta A, Kaushik D, Verma A. (2020) Applying machine learning in IoT to build intelligent system for packet routing system. *Mater Today Proc.* <https://doi.org/10.1016/j.matpr.2020.09.539>
18. Gupta A, Verma A, Kaushik D, Garg M (2020) Applying deep learning approach for brain tumor detection. *Mater Today Proc.* <https://doi.org/10.1016/j.matpr.2020.10.063>

19. Gupta A, Kaushik D, Garg M, Verma A (2020) Machine learning model for breast cancer prediction. In: 2020 fourth international conference on I-SMAC (IoT in social, mobile, analytics and cloud) (I-SMAC), pp 472–477. <https://doi.org/10.1109/I-SMAC49090.2020.9243323>
20. Shukla KA, Ahamad S, Rao GN, Al-Asadi AJ, Gupta A, Kumbhkar M (2021) Artificial intelligence assisted IoT data intrusion detection. In: 2021 4th international conference on computing and communications technologies (ICCCCT), pp 330–335. <https://doi.org/10.1109/ICCCCT53315.2021.97111795>
21. Bansal R, Gupta A, Singh R, Nassa VK (2021) Role and impact of digital technologies in E-learning amidst COVID-19 pandemic. In: 2021 fourth international conference on computational intelligence and communication technologies (CCICT), pp 194–202. <https://doi.org/10.1109/CCICT53244.2021.00046>
22. Gupta A, Singh R, Nassa VK, Bansal R, Sharma P, Koti K (2021) Investigating application and challenges of big data analytics with clustering. In: 2021 international conference on advancements in electrical, electronics, communication, computing and automation (ICAECA), pp 1–6. <https://doi.org/10.1109/ICAECA52838.2021.9675483>
23. Sreekanth N, Rama Devi J, Shukla KA et al (2022) Evaluation of estimation in software development using deep learning-modified neural network. *Appl Nanosci*. <https://doi.org/10.1007/s13204-021-02204-9>
24. Sindhwani N, Maurya VP, Patel A, Yadav RK, Krishna S, Anand R (2022) Implementation of intelligent plantation system using virtual IoT. In: *Internet of things and its applications*. Springer, Cham, pp 305–322
25. Anand R, Singh B, Sindhwani N (2009) Speech perception and analysis of fluent digits' strings using level-by-level time alignment. *Int J Inf Technol Knowl Manag* 2(1):65–68
26. Chawla P, Juneja A, Juneja S, Anand R (2020) Artificial intelligent systems in smart medical healthcare: current trends. *Int J Adv Sci Technol* 29(10):1476–1484
27. Srivastava A, Gupta A, Anand R (2021) Optimized smart system for transportation using RFID technology. *Math Eng Sci Aerosp (MESA)* 12(4)
28. Goyal B, Dogra A, Khoond R, Gupta A, Anand R (2021) Infrared and visible image fusion for concealed weapon detection using transform and spatial domain filters. In: 2021 9th international conference on reliability, infocom technologies and optimization (trends and future directions)(ICRITO). IEEE, pp 1–4
29. Vyas G, Anand R, Holè KE, Implementation of advanced image compression using wavelet transform and SPHIT algorithm. *Int J Electron Electric Eng*. ISSN 0974-2174
30. Kumar R, Anand R, Kaushik G (2011) Image compression using wavelet method & SPIHT algorithm. *Digital Image Processing* 3(2):75–79

Comparative Analysis of Autoencoders and U-net-Based Image Steganography



Juhi Singh, Akshat Yadav, Ayushi Siddhu, and Shivani Sharma

Abstract In today's fast paced and developing era, there is lot of information in different forms of data, pictures, audios, videos, documents, etc., which needs to be kept secured. Steganography is defined as a method of hiding data inside another form of data. Image steganography is a sub-branch where the data to be hidden is encoded inside an image. In this paper, two of the deep learning methods to hide data inside image (full RGB image inside another RGB image) have been studied and compared. The methods, namely autoencoder and U-net-based image steganography, have been compared using metrics like PSNR and SSIM. Experimental analysis is provided to prove that although for the encoding part, both U-net and autoencoder perform well, but while decoding the actual secret image, the performance of U-net is better than the autoencoder architecture.

Keywords Autoencoders · U-net · Image steganography · PSNR · SSIM

1 Introduction

The word steganography has been derived from the Greek word '*steganographia*', a combination of the words '*steganós*' meaning concealed or hidden, and the word '*graphia*' meaning writing [1]. The advantage of using steganography over cryptography has the advantage that it is not clear to another person if some other data is

J. Singh (✉) · A. Yadav · A. Siddhu · S. Sharma
Amity School of Engineering and Technology, Amity University, Noida, India
e-mail: jsingh7@amity.edu

A. Yadav
e-mail: akshat.yadav1@s.amity.edu

A. Siddhu
e-mail: ayushi.siddhu@s.amity.edu

S. Sharma
e-mail: shivani.sharma44@s.amity.edu

hidden inside the data that is visible to the eyes. In today's fast paced and developing era, there is lot of information out there in different forms of data, pictures, audios, videos, documents, etc., and it is possible that information which is not supposed to be looked at by others is exposed or put in unintended hands. Data security has become one of the key issues today and it is important to protect the data. Steganography has proven to be one of the very useful methods to hide such information. In this paper, two techniques to hide data inside images have been compared using metrics like PSNR and SSIM. These are two prominent techniques that are used in order to assess the quality of the image, or more specifically the quality of the imperceptibility. Peak signal-to-noise ratio (PSNR) is a system that helps to check the similarities and the differences. In simple words, it is mean-squared error. Generally, the value of PSNR varies from 30 to 50, where higher the value is better, if the images have a significant difference, then the value can be as low as 15. PSNR is an easy and fast method to assess the quality of the images; however, in reality, it may differ with the human perception, and thus, we need structural similarity index measure (SSIM), it returns similarity index which is averaged over all the channels of the image, and the value ranges between 0 and 1 where 1 is considered to be the perfect fit. Both the methods have been deployed on the images in this paper and the result is used to bring out a comparative analysis.

2 Literature Survey

The first usage of steganographic methods can be traced to as back as 440 BC [1, 2], where secret message was written on top of a shaved scalp and was hidden as the hair grew back, the intended receiver could shave the head and get to know the secret message.

2.1 *Traditional Method*

One of the very old methods in image steganography has been the one where data is hidden inside the least significant bits (LSBs) of the image [3]. The secret information is converted into its binary equivalent, and then, the bits are put in the LSBs of the cover image and thus providing us with the final encoded image. Similarly, when the secret data is required, the LSBs are used to reconstruct the original message. One disadvantage of the LSB method is that it is not secure and is very vulnerable to techniques like steganalysis which are used for detection of steganography. Secret data could easily get into the wrongs hands when using this technique.

2.2 *Improvements over the LSB Method*

Techniques like Hash-LSB with RSA encryption [4] and Blow-fish encryption [5] have been used to make the data more secure. In the former method, Hash-LSB one, rather than encoding the secret information inside the LSB of the container data, the RSA algorithm is used to encrypt the secret data and a Hash function finds the positions suitable for hiding the secret information, and then, the secret message is encoded at those positions. According to this method, it hides 8 bits of the information in the order of 3, 3, 2 in the RGB pixel, meaning that 3 bits of the information are hidden in LSB of red pixel, 3 in LSB of green pixel and 2 in the green pixel, as change in the blue color is more susceptible to the human eye. Similarly, the data is decrypted after finding the positions of the encoded image LSB using the Hash function, finding the data in the order of 3, 3, 2 and then using RSA to retrieve the original message.

In a similar way, the Blow-fish algorithm is used for encrypting the secret message and then applying the LSB steganography method for data hiding.

2.3 *Newer Methods*

Considering the recent methods, attempts have been made to conserve the image statistics through the use of first-order and second-order statistics matching models, one of them is known as HUGO.

2.4 *Use of Deep Learning*

Attempts to use neural networks to for the purpose of hiding data by letting them select the suitable LSBs to place the information have been made successfully. Some have managed to use these networks to instead work on which bits of the encoded image to extract the information from.

In this paper, two of the deep learning methods have been studied and compared for hiding a full RBG image inside another RGB image. These methods are:

2.4.1 *Autoencoders*

Autoencoders are a type of artificial neural network (ANN) and a technique of unsupervised learning where the input is regenerated from the encoding [6]. They are used to compress the data into a lower dimension representation, and then reconstruct the original input from this representation, learning to use the required features and ignoring the noise or insignificant data.

They have a wide range of applications from detecting features, facial recognition, anomaly detection, data denoising, etc.

2.4.2 U-net

U-net is a type of convolution neural network (CNN) originally developed for the purpose of biomedical image segmentation. The network is made up of two layers known as the contracting and expansive path [7].

The contracting path performs repeated convolutions, each of which are followed by a rectified linear unit (ReLU) and a max pooling operation. During this phase, the feature information is increased along with the reduction of spatial information.

The expansion path combines these features and their respective spatial information with the help of up convolutions and concatenating the features from the contraction path.

3 Methodology

3.1 Dataset

The dataset is a collection of images gathered from various freely available sources like kaggle, unsplash, etc. The images are full RGB images of sizes with at least 300pixels of height and 300pixels of width for further adjustment during preprocessing.

3.2 Frameworks/Language

Python with the frameworks including tensorflow and pytorch is selected for the preprocessing, implementation and validation of the autoencoder and U-net-based models.

3.3 Autoencoder Architecture

In this paper, a three-layer architecture comprising of a preparation layer, an encoding layer and a decoding layer is used for training the model similar to what has been mentioned in the paper [8].

The preparation layer takes care of the preprocessing of the images, making sure the images are in the correct measurement and are of the same size and pass them to the encoding network.

The output of the preparation layer is the input of the encoding network, and it takes two images of the same size and attempts to hide the secret image inside the cover image. 5 convolution layers having 50 filters of 3×3 , 4×4 , 5×5 patches have been used.

The final decoding network takes in the input the output of the encoding network and attempts to reconstruct the original secret image from the encoded image.

The network learns by reducing the error, and the loss function used to train the autoencoder network is: $S - S'$

$$L(C, C', S, S') = ||C - C'|| + b||S - S'|| \quad (1)$$

where $C - C'$ are the difference between the original cover image and the encoded image, and the $S - S'$ is the difference between the original secret image and decoded secret image. b is the hyperparameter used to control the reconstruction quality of the secret image.

3.4 U-net Architecture

The U-net architecture unlike autoencoders comprises two layers, a hiding layer and a revealing layer. This architecture is similar to the work done in the paper [9].

The hiding network encodes the secret image inside the cover image. This is the contraction phase where both the images are concatenated into a 6-channel feature tensor followed by 4×4 convolution layer in the down-sampling process. It is then followed by an activation function (Leaky ReLU) and BN operation for speeding up the network training. After 7 such operations, we have a feature map of 2×2 with 512 feature channels.

The expansion phase or the revealing layer takes in the input the output of the hiding network and performs up-sampling operations each concatenated with the feature map from the respective contraction phase.

4 Workflow

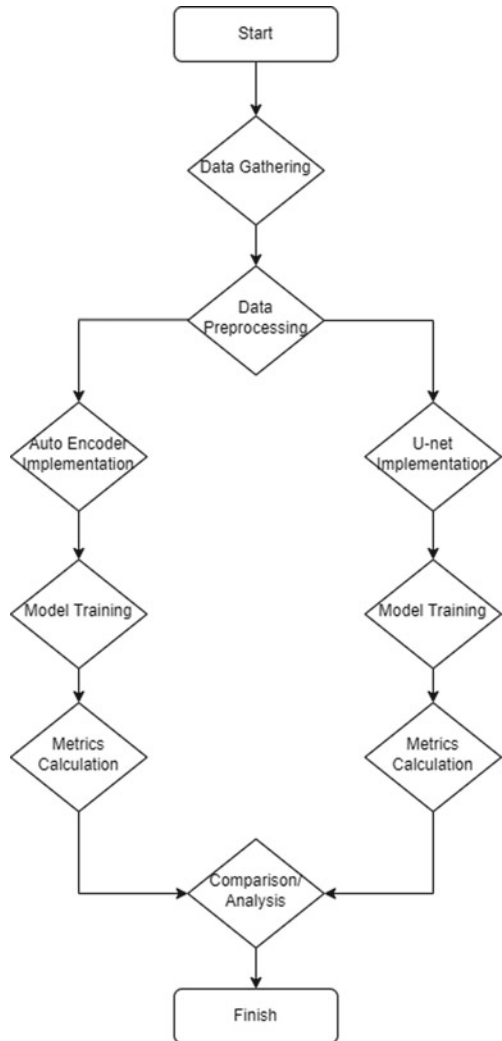
As is visible in the diagram, we begin with data gathering. In this paper, the image dataset we have used to train our models has been collected from various freely available online resources such as kaggle and unsplash.

After collecting the image dataset as per our requirements, we begin with the data preprocessing making sure that there are no discrepancies in the image format or corruption in the files, converting the data to appropriate sizes for model training.

Then comes the model implementation part, where we begin with the implementation of the autoencoder model according to the architecture mentioned earlier in the paper and then train the model using our image dataset.

After successful training, metrics (PSNR, SSIM) are calculated. Same is the process for the U-net model. After gathering the metrics, comparison and analysis are done (Fig. 1).

Fig. 1 Workflow



5 Results and Experiments

Figures 2 and 3 are the images acquired from the autoencoder and the U-net model after training of both the models on the same image dataset for 200 iterations.

Below are the calculated metrics, PSNR and SSIM for the images in Figs. 2 and 3.

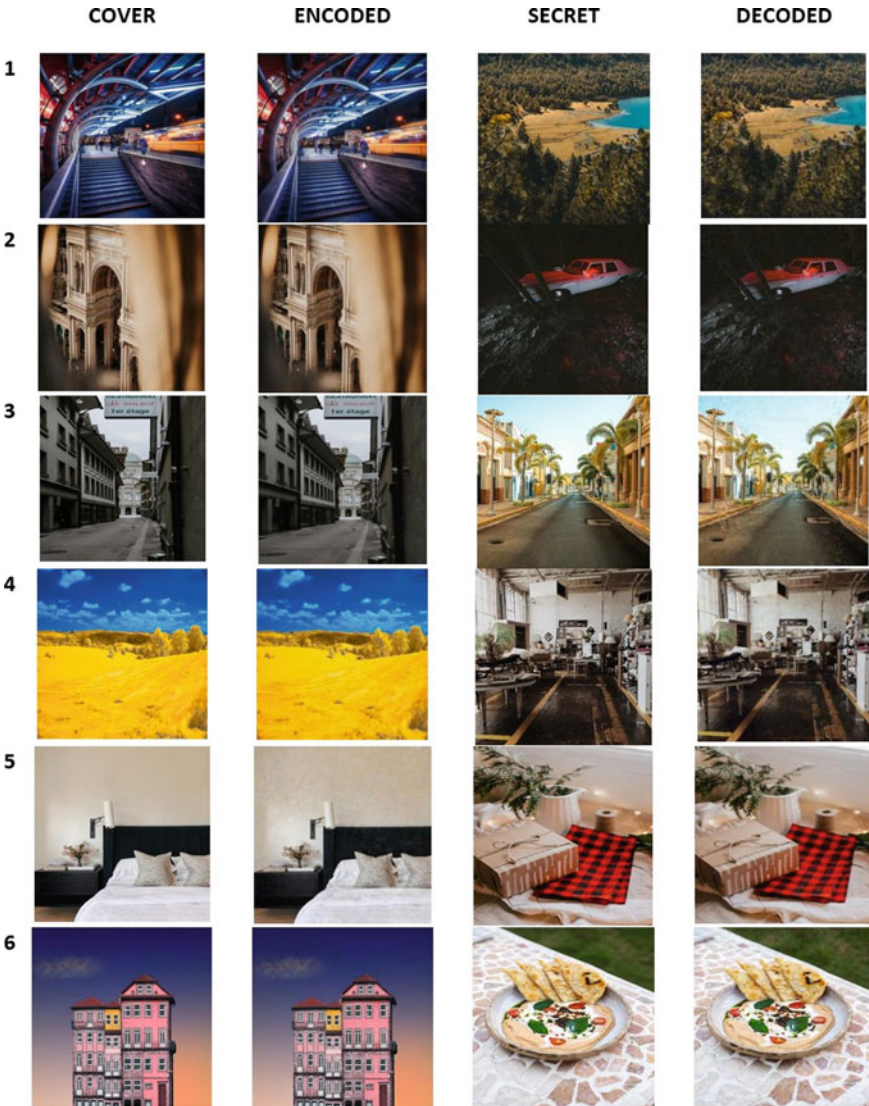


Fig. 2 U-net

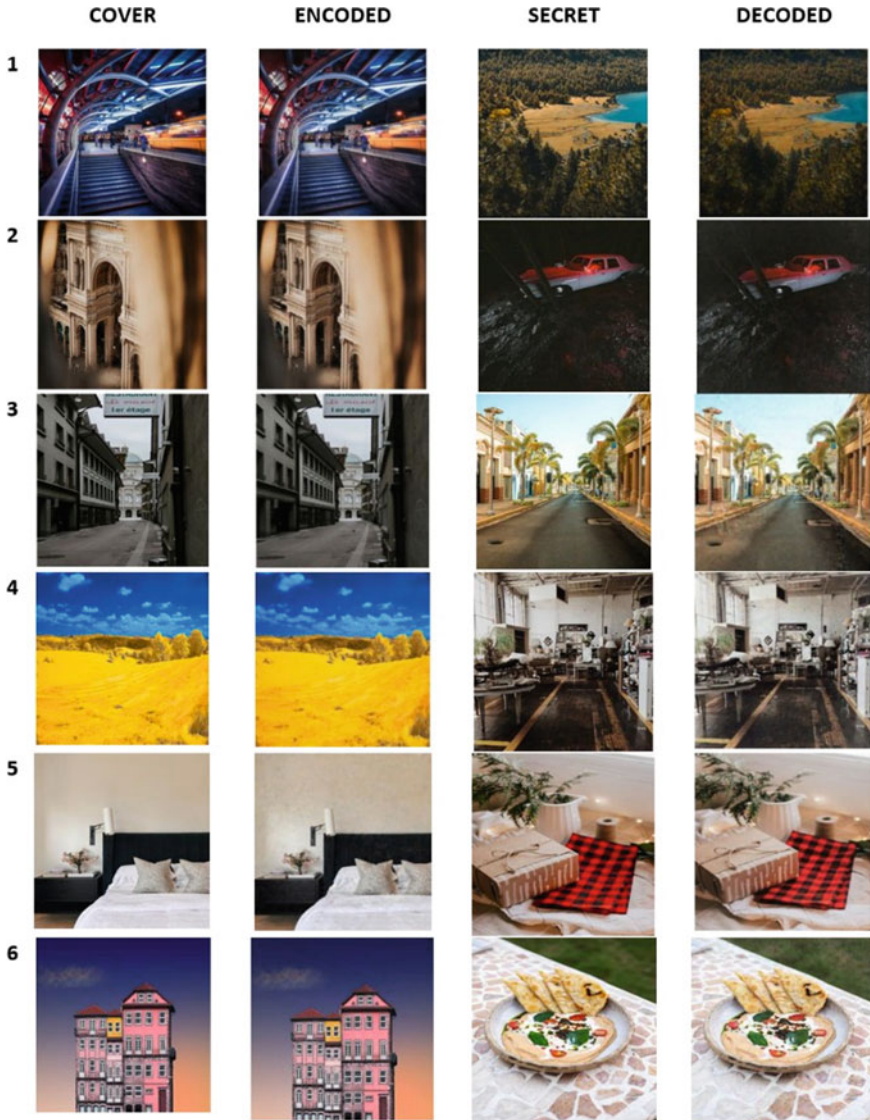


Fig. 3 Autoencoder

Table 1 depicts the structural similarity index measure and peak signal-to-noise ratio for the above mentioned images, specifically for the encoding part, the values are for both the models that are to be compared in the paper, namely U-net and autoencoder and it can be observed that for the encoding part, there is very minute difference between the two techniques. This is more clear with Figs. 4 and 5.

Table 1 PSNR and SSIM for the cover and encoded images

	U-net		Autoencoder	
	SSIM	PSNR	SSIM	PSNR
1	0.956	33.03	0.945	31.29
2	0.967	37.26	0.960	34.04
3	0.949	35.07	0.950	33.97
4	0.816	33.11	0.763	30.22
5	0.927	35.61	0.924	34.81
6	0.947	34.07	0.939	32.06

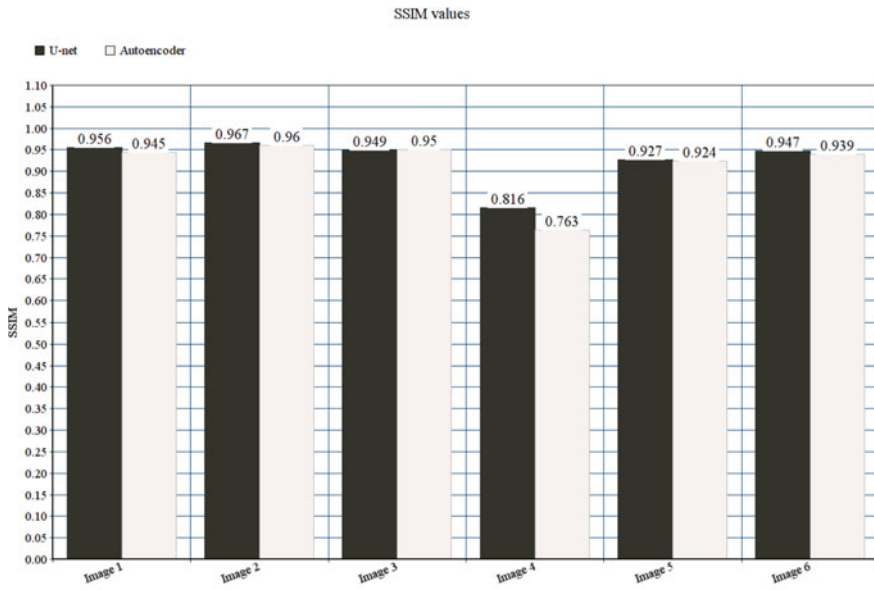


Fig. 4 SSIM values (cover versus encoded image)

Table 2 gives the structural similarity index measure and peak signal-to-noise ratio for the above mentioned images, specifically for the encoding part, the values are for both the models that are to be compared in the paper, namely U-net and autoencoder and it can be observed that for the encoding part, there is a slight difference between the two techniques and can be observed that the U-net model performs better than the autoencoder architecture. This is more clear with Figs. 6 and 7.

The average values are calculated for the PSNR and SSIM values and further conclusions are made regarding the two approaches.

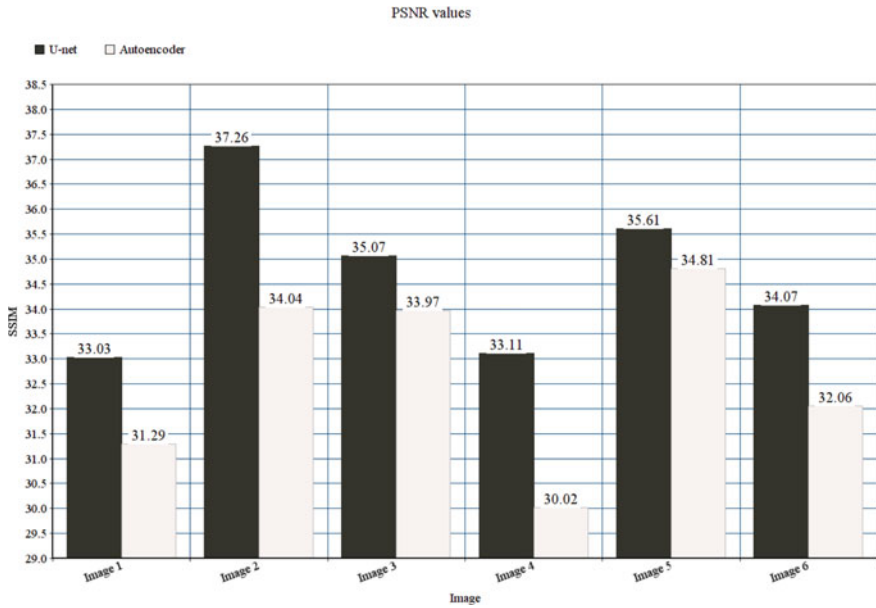


Fig. 5 PSNR values (cover versus encoded image)

Table 2 PSNR and SSIM for the secret and decoded images

	U-Net		Autoencoder	
	SSIM	PSNR	SSIM	PSNR
1	0.918	31.13	0.805	22.14
2	0.926	35.22	0.908	29.39
3	0.925	31.45	0.911	26.49
4	0.926	31.35	0.881	22.00
5	0.957	36.76	0.909	25.57
6	0.942	33.26	0.929	27.47

6 Conclusion

The tests were performed on newly acquired images from open sources websites like unsplash, kaggle, etc. For the U-net and the autoencoder models, the results are as follows:

Values for cover image—encoded image U-Net

$$\text{Average SSIM} = 0.927$$

$$\text{Average PSNR} = 34.69$$

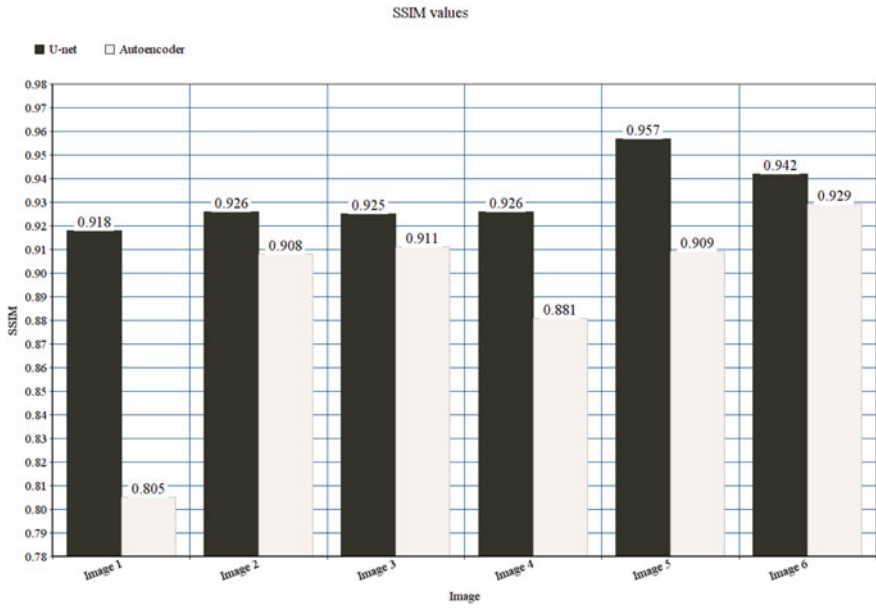


Fig. 6 SSIM values (secret versus decoded image)

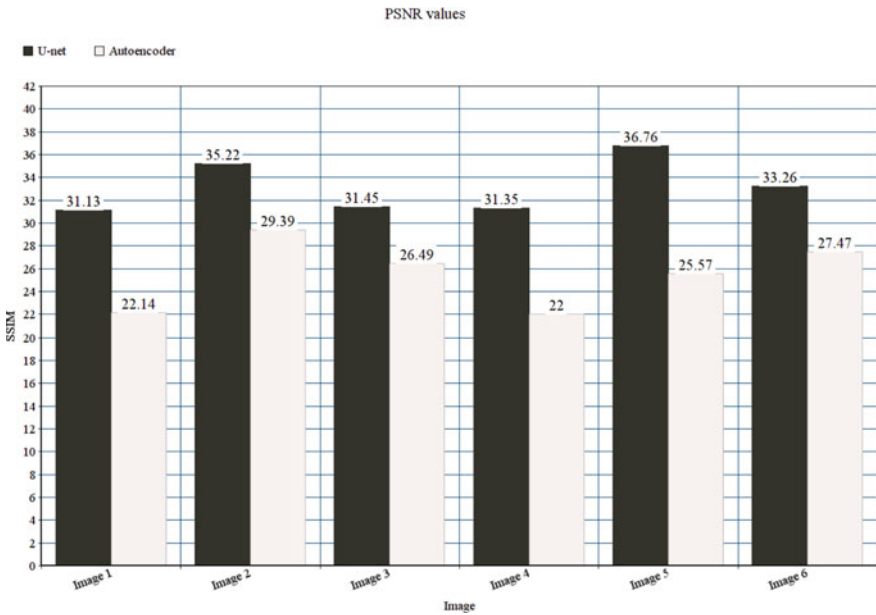


Fig. 7 PSNR values (secret versus decoded image)

Autoencoder

Average SSIM = 0.913

Average PSNR = 32.73

Values for secret image—decoded image U-Net

Average SSIM = 0.932

Average PSNR = 33.19

Autoencoder

Average SSIM = 0.890

Average PSNR = 25.51

It is evident from the results that although for the encoding part both the U-Net and autoencoder perform well, while decoding the actual secret image, U-Net performs surprisingly better than the autoencoder architecture.

For the purpose of hiding the secret image inside the cover image, U-Net performs 1.53% better than the autoencoder considering the SSIM, and 5.98% better than autoencoder considering the PSNR.

While decoding the secret image from the encoded image, U-Net performs 4.72% better than autoencoder for SSIM, and 30.1% better than autoencoder for PSNR.

It is to be noted that both the models were trained on the same image dataset and for the same number of iterations.

We can conclude that if less amount of data and resources are available, U-Net although with the architecture really similar to an autoencoder, or can even be called as a kind of encoder, can perform relatively better than a simple autoencoder for hiding an image inside another and decoding it later on when trained on a similar dataset and for the same number of iterations.

References

1. Moerland T, Steganography and steganalysis. Leiden Institute of Advanced Computing Science. www.liacs.nl/home/tmoerl/privtech.pdf
2. Petitcolas FAP, Anderson RJ, Kuhn MG (1999) Information hiding: a survey (PDF). Proc IEEE 87(7):1062–1078. CiteSeerX 10.1.1.333.9397. <https://doi.org/10.1109/5.771065>. Retrieved 02 Sept 2008
3. Morkel T, Elloff JHP, Olivier MS, An overview of image steganography
4. Shridevishetti, Anuja S (2015) A secure image steganography based on RSA algorithm and hash-LSB technique. Int J Eng Res Technol (IJERT) ICESMART—2015, 3(19)
5. Murad S, Gody A, Barakat T (2019) Enhanced security of symmetric encryption using combination of steganography with visual cryptography. Int J Eng Trends Technol, 65. <https://doi.org/10.14445/22315381/IJETT-V65P227>
6. Goodfellow I, Bengio Y, Courville A (2016) Deep learning. MIT Press. ISBN 978-0262035613

7. Ronneberger O, Fischer P, Brox T (2015) U-net: convolutional networks for biomedical image segmentation. [arXiv:1505.04597](https://arxiv.org/abs/1505.04597)
8. Baluja S (2017) Hiding images in plain sight: deep steganography. In: Neural information processing systems
9. Xintao D, Jia K, Li B, Guo D, Zhang E, Qin C (2019) Reversible image steganography scheme based on a u-net structure. IEEE Access. pp 1–1. <https://doi.org/10.1109/ACCESS.2019.2891247>

Challenges in VLSI Design for Efficient Energy Harvesting



Sanjay Kumar and Mansi Jhamb

Abstract Energy harvesting means extracting energy from the ambient environment to power stand-alone systems. Ambient energy sources are vibration, solar, radio frequency, wind, pressure and thermal, etc. Piezoelectric materials, photo voltaic cell and RF antennas are very common sources of energy harvesting. This technology plays a very important role where batteries are impractical such as body sensor network and inaccessible remote systems. For example, using the electronic circuits in an automobile, one can measure various parameters such as tire pressure and engine temperatures by placing sensors with battery in various hard to reach remote places inside the vehicle and diagnose the various automotive issues. Thus, replacing or extending life of a battery is one of the major problems. In medical branch, the doctors can monitor patients remotely with the help of various sensors that monitor health issues of the patients. This paper reviews the state of art energy harvesting models which provide high efficiency energy conversion (AC to DC and DC to DC) to regulate voltages or to charge batteries and super capacitors storage elements.

Keywords CMOS · Energy harvesting · Low-frequency · Rectifier · Power conversion efficiency · Piezoelectric · Near-field · Far-field

1 Introduction

The energy harvesting research can be divided two major areas. One is developing finest structures of energy harvesters and other is designing electronic circuits that

S. Kumar (✉)
USIC&T, GGSIPU, New Delhi, India
e-mail: sanjay.kumar2@dseu.ac.in

DSEU, G.B Pant Okhla 1 Campus (Formerly GBPEC), New Delhi, India

M. Jhamb
Department of Electronics & Communication Engineering, USIC&T, GGSIPU, Dwarka, New Delhi, India
e-mail: mansi.jhamb@ipu.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
V. V. S. S. Chakravarthy et al. (eds.), *Advances in Signal Processing, Embedded Systems and IoT*, Lecture Notes in Electrical Engineering 992,
https://doi.org/10.1007/978-981-19-8865-3_58

657

Table 1 Power harvested from different sources [1, 3]

Source of energy	Density (Power)	Technology	Advantages	Disadvantages
Solar energy	Indoor: 11.0 $\mu\text{W}/\text{cm}^2$ Outdoor: 10.0 mW/cm^2	Photovoltaic	Mature and large power density	Not implantable necessitates light and costly
Vibration energy	Human: 4 $\mu\text{W}/\text{cm}^2$ Industrial: 104.0 $\mu\text{W}/\text{cm}^2$	Piezoelectric Electro magnetic Electrostatic	High efficiency and implantable	Not available always material limitation
Thermal/heat energy	Human: 30 $\mu\text{W}/\text{cm}^2$ Industrial: 1.0–10 mW/cm^2	Thermo-electric Pyro-electric	Implantable and High power density	Not available always heat generates in excess
Radio energy	GSM: 0.10 $\mu\text{W}/\text{cm}^2$ WIFI: 1.0 W/cm^2	Antenna	Implantable available	Efficiency inversely proportional to distance and low density

have very good voltage conversion efficiency and power efficiency to store the generated charge. The combination of low current, low-threshold voltage requirement of the transistors, and the low power supply voltages allows energy harvesting technology to be utilized in various miniaturized low power CMOS electronic systems. Many kind of ambient energy sources are available around us such as solar, thermal, electrostatic, acoustic noise, human generated, nuclear power, wind, radio frequency, and mechanical vibration [1, 2]. Among these energy sources, electromagnetic, electrostatic, thermal, chemical, and mechanical-based energy harvesting systems are the ones typically used for small-scale power required by electronic systems. Power harvesting from different sources of energy is given in Table 1.

2 System Overview

Figure 1 shows the block diagram of energy harvesting system.

2.1 Ambient Energy Source Based on Vibration

Ambient energy source based on vibration can be found in many different forms. For example, a human while walking can produce kinetic energy based on vibration. Table 2 gives different types of vibration energy generated based on frequency and

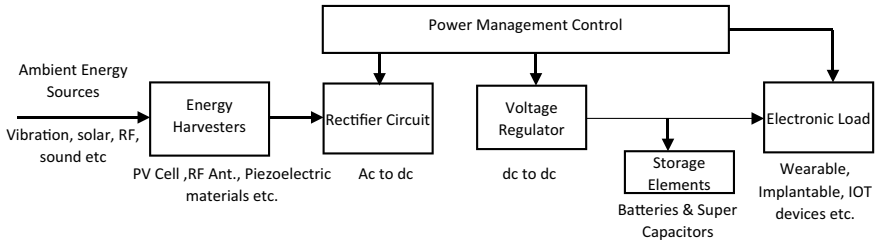


Fig. 1 Energy harvesting system

amplitude of acceleration at the fundamental frequency. As can be seen from Table 2, there are many forms of vibration based kinetic energy present around us.

The operating frequency of all devices is relatively low. Figure 2 shows the simplified model of a vibration translator [4].

From the Fig. 2, the power translated from vibration energy can be written as.

$$\text{Power} = \frac{1}{2}bez^2$$

where

be is representing the coefficient of an electrically generated damping,

Table 2 Maximum frequency and acceleration of vibrational energy sources [2, 4]

Vibration sources	Max. frequency (Hz)	Acceleration (m/s ²)
Clothes dryer	120	3.3
Microwave oven	120	2.19
Washing machine	110	0.4
Computer CD R/W	75	0.6
Car engine	200	1.2
Vehicles	5 ~ 2000	0.5 ~ 110
Kitchen blender	121	3.4
Refrigerator	240	0.1

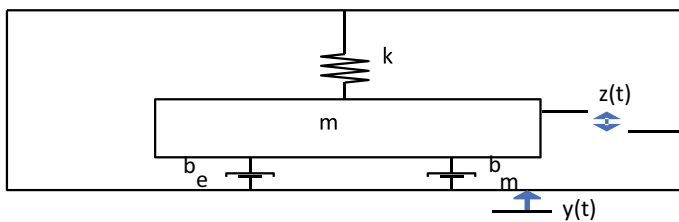


Fig. 2 Simplified model of a vibration t translator [4]

b_m is representing the coefficient of a mechanical damping,
 k is representing the coefficient of spring constant,
 m is representing the mass.

2.1.1 Piezoelectric Effect

The word “piezo” came from the Greek word meaning “pressure” [4]. The phenomenon of the piezoelectric effect was discovered by Jacques and Pierre Curie brothers in 1880, and demonstrated piezoelectric effect using a quartz and a tourmaline [4, 5]. Piezoelectric materials can be demonstrated by means of a piezo-effect when subjected to electrical force or mechanical bending or stretching. Such behavior will cause a piezoelectric material to undergo a change in electrical polarization. A direct piezoelectric effect is called a generator or a sensor effect, which converts mechanical energy into electrical energy [4]. In addition, an electrical potential can be applied causing a change in length or deformation of the shape of the piezoelectric material. This is called as the inverse piezoelectric effect or the actuator effect [5].

2.1.2 Piezoelectric Material

There are diverse types of natural or synthetic piezoelectric materials currently available for commercial applications. Typical natural piezoelectric materials are quartz, salt, cane sugar, tourmaline, etc. [5]. Examples of typical man-made or synthetic piezoelectric materials are lead zirconate titanate (PZT), barium titanate (BaTiO_3), polyvinylidene fluoride (PVDF), ZnO, etc. [5]. From industrial as well as research point of view, breakthrough of piezoelectric materials technology begins with the development of piezoelectric ceramics.

Table 3 lists the most commonly used piezoelectric materials. From the table, it can be easily seen that there is one particular material, which has superior properties compared to the others. The superior material is PZT. Usually, a PZT has the highest Curie temperature as well as the largest electromechanical coefficient. Figure 3 shows how the piezoelectric material acts with different applied forces or strains [5, 6]. In general, a piezoelectric crystal is in non-symmetrical form and is electrically neutral as well. Therefore, electrons in piezoelectric materials have perfectly balanced charges. However, when a piezoelectric material is squeezed or stretched, atoms inside of piezoelectric crystals are getting closer together or fall apart from each other. Such action causes upsetting of the balance of positive and negative charges. As a result, a piezoelectric material emits electrical charges. For commercial applications, two types of PZT materials are available. One is a hard type PZT, and the other is a soft type PZT [6]. Usually, a PZT ceramic is doped with either acceptor dopants or donor dopants. If a PZT is doped with an acceptor dopant, then it is called as a hard type PZT, and if it is doped with a donor dopant, then it is called as a soft type PZT. The difference between a hard and a soft type of PZT is whether a piezoelectric constant is higher or not. Since the domain wall

motion of a hard type of a PZT is blocked or disrupted by its impurities, it has a lower piezoelectric constant but has lower loss. On the other hand, a soft type of PZT has a higher piezoelectric constant but higher loss as well. This is due to the fact, that an acceptor dopant creates an oxygen vacancy, while a donor dopant creates a metal vacancy. When the polarization process is performed with strong electrical field applied to two electrodes, the directions of polarization are determined to be the axis 3.

The polarized piezoelectric material can be characterized by several coefficients [6]. The basic simplified form of the polarized piezoelectric ceramic with respect to the relationship of the electrical charge and the mechanical strain, which describe

Table 3 Characteristics of piezoelectric materials [4, 5]

Materials	Shapes or forms	d_{31} (m/V or C/N) ¹	$(\epsilon_{33}/\epsilon_0)^2$	k_{31}	TC(°C) ⁴
Quartz	Single crystal	2.4	4.5	–	–
P.Z.T	Sol–gel thin film	190 ~ 250	800 ~ 1100	–	–
P.Z.T	Polycrystalline	–191 ~ 321	1810 ~ 3200	0.31 ~ 0.43	234 ~ 351
PZT	Sputtered thin film	100	–	–	–
P.V.D.F	Film	23	11 ~ 14	0.14	82 ~ 104
ZnO	Thin film	11.5 ~ 12.5	11.8 ~ 12	–	–

d_{31} denotes the piezoelectric coefficient

$\epsilon_0 = 8.854 \times 10^{-12}$ F/m which is the permittivity of the empty space, ϵ is the dielectric constant

k_{31} denotes the electromechanical coupling coefficient

TC is the Curie temperature

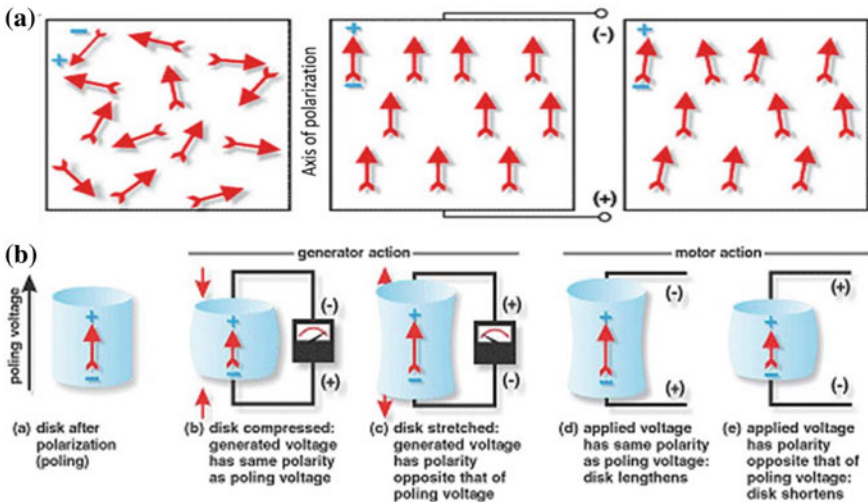


Fig. 3 a Polarizing (Poling) a piezoelectric ceramic. b Generator and motor (actuator) action of a piezoelectric material [6, 7]

the piezoelectric.

$$D = d * T + cT * E \quad (1)$$

$$S = SE * T + d * ET \quad (2)$$

where

D is the coefficient of an electrical flux density,

T is the coefficient of a mechanical stress,

E is the coefficient of an electric field,

S is the coefficient of a mechanical strain,

D is the coefficient of a piezoelectric charge,

ϵT is the coefficient of a permittivity,

SE is the coefficient of a compliance or elasticity.

These simplified Eqs. (1) and (2) can be applied to the small signal model. In addition, mechanical strain (S), electrical field (E), stress (T), and electrical flux density (D) are linear and the coefficients are constant. The coefficients are usually obtained from the piezoelectric material data sheet.

2.1.3 Piezoelectric Material as Energy Sources

A piezoelectric material can be used as a generator, which converts mechanical energy into electrical energy, as well as an actuator, which converts electrical energy into mechanical. Since this research is focused on energy harvesting from an electrical energy of a piezoelectric material, therefore, this section concentrates on the mechanism of transformation of mechanical energy into electrical energy in piezoelectric materials. Figure 5 shows various working (energy generator) modes of a piezoelectric material [6, 7]. As shown in Fig. 5, there are three general directions of force for the material to be working as a generator. Among these, the longitudinal compression mode produces the highest amount of electrical energy conversion and this direction is also called 33 mode (Figs. 4 and 6).

In addition to the direction of the force, there are two modes of operation in a piezoelectric material as an actuator: one is series and the other is parallel. The series and the parallel operations depend on the polarization and wiring shape or configuration of the piezoelectric material layer. If the voltage is applied over the entire piezoelectric layer, then it is called as series operation. If the voltage is applied on each layer of a piezoelectric material, then, it is called as a parallel operation.

2.1.4 Piezoelectric Transducer Modeling

A piezoelectric material can be used as a generator, which produces electrical energy by applying mechanical energy as well as an actuator or a motor, which produces

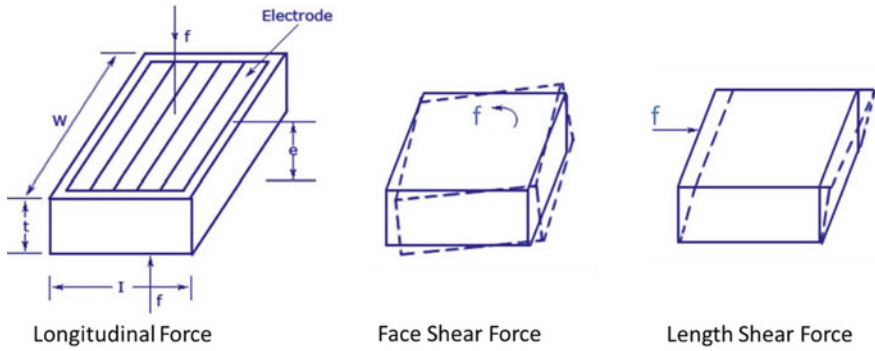


Fig. 4 Various working modes of a piezoelectric material [7]

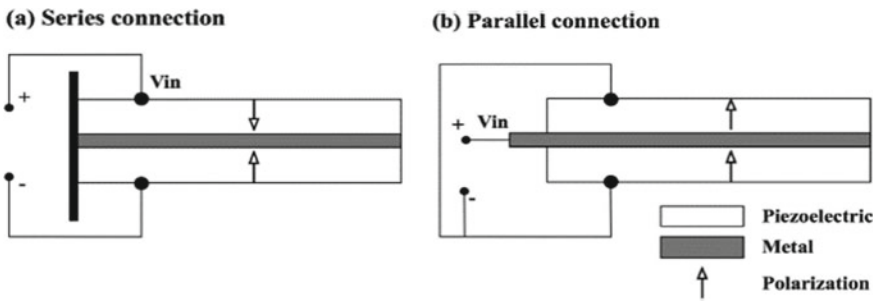
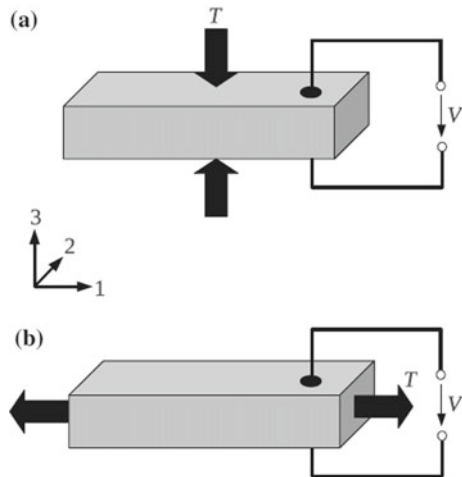


Fig. 5 Series and parallel operation of a piezoelectric material [7]

Fig. 6 Piezoelectric material operating in a 33 mode, b 31 mode



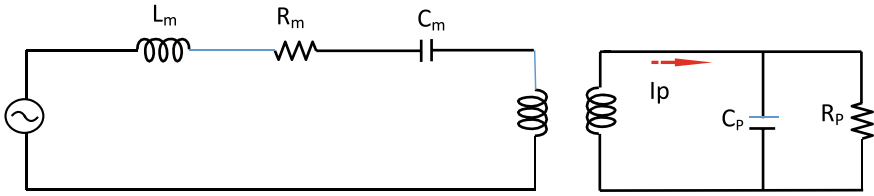


Fig. 7 Piezoelectric transducer modeling in a mechanical domain and an electrical domain [7, 8]

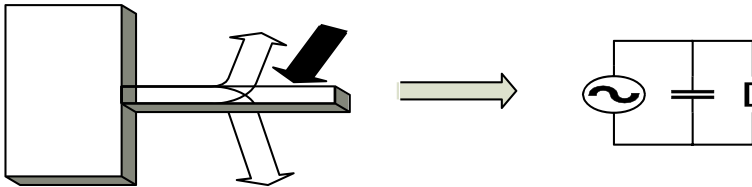


Fig. 8 Simplified electrical domain modeling at or close to resonance frequency [7, 8]

mechanical energy by applying electrical energy. Figures 7 and 8 show the models of a piezoelectric transducer in mechanical and electrical domains [7]. In Fig. 7, L_m denotes the mechanical mass, C_m represents the mechanical stiffness, and R_m denotes the mechanical losses. A transformer presented in the mechanical domain in Fig. 8 allows for conversion of the mechanical stress into current. C_p represents the parasitic capacitance of the piezoelectric material, and R_p denotes the internal loss in an electrical domain.

An electrical domain is transformed at or close to the resonance frequency of a piezoelectric material. In general, a typical power supply or a battery has a very low internal impedance. However, as shown in Figs. 7 and 8, the internal impedance of a piezoelectric transducer, R_p , has a very high value. As a result, the amount of output current produced by the piezoelectric transducer is limited by this high internal impedance, R_p . The common value of the output current produced by a piezoelectric transducer is in the micro-ampere range. In addition, because of this low current, the output voltage of a piezoelectric transducer is low as well. The output current of a piezoelectric transducer is proportional to the input vibration amplitude and can be expressed by the following equation which assumes that the input vibration is sinusoidal and therefore,

$$i_P = I_P \sin \omega P T \tag{3}$$

where i_P is equal to I_{AC} in Fig. 8, I_P is the peak current of the sinusoidal current source, $\omega P = 2\pi f_P$, and f_P is the excited frequency of the piezoelectric transducer [8].

2.2 Ambient Energy Source Based on Radio Frequency

Power can be harvested from abundantly available various sources of energy such as thermal, solar, wind, electromagnetic, and vibrational. Since electromagnetic energy is always available in our surrounding due to its numerous application, we can utilize it profoundly for harvesting purpose. With the help of antenna and rectifier, AC can be converted into DC very easily. However, the size of antenna, RF power loss with distance (free space power loss), the distance between the RF source and receiving antenna, frequency, antenna gain, and power consumption by rectifier are the major challenges. The power density of RF wave majorly changes with distance, hence, it is broadly divided by far-field and near-field range. Power density is weak and uniform in far-field and called Fraunhofer’s distance. However, in near-field, it is very strong and both electric and magnetic vectors are independent as shown in Fig. 9.

The Fraunhofer’s distance is defined as:

$$d_f = \frac{2D^2}{\lambda} \tag{4}$$

where

- d_f = Fraunhofer’s distance,
- D = diameter of the antenna,
- λ = wavelength.

Up to the distance of $\sqrt[3]{\frac{D}{2\lambda}}$ from radiating source the region is reactive; hence, electric and magnetic vectors are out of phase which is a big reason for energy distortion. In far away region, i.e., far-field region electric and magnetic vectors still

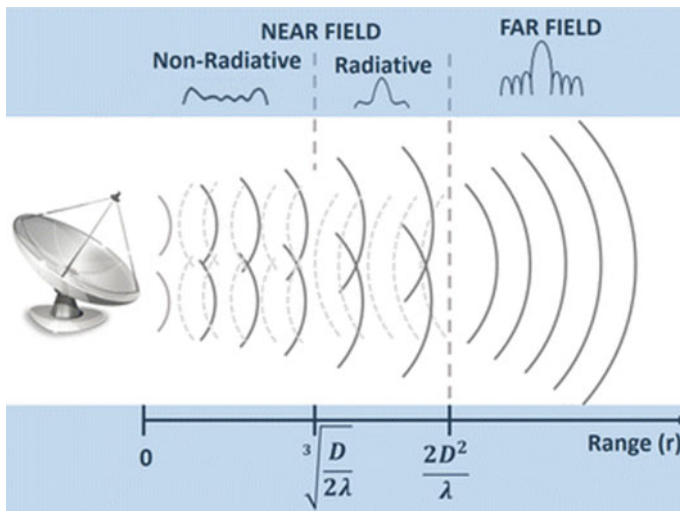


Fig. 9 Near-field and far-field regions [9]

vary but this region is radiative and called Fresnel region. The power in this region is given as:

$$P_R = \frac{P_T G_T G_R \lambda^2}{(4\Omega R^2)} \tag{5}$$

Also the received antenna gain can be calculated as [10]

$$P_R = \frac{P_T G_T G_R}{(4\pi d/\lambda)^2} \tag{6}$$

where

P_R is power at the receiver antenna,

G_R is receiver antenna gain relative to the isotropic source (dBi),

$$\lambda = \frac{C}{f} \tag{7}$$

$k = 2\pi/\lambda$ is the wave number.

From the above formula, the FSPL, P_L for far-field can be inferred as

$$P_L = \frac{P_T}{P_R} = \frac{(4\Omega R^2)}{G_T G_R \lambda^2} = \frac{(4\Omega f R^2)}{G_T G_R C^2} = \frac{4}{G_T G_R} (KR^2) \tag{8}$$

Or

$$P_L(\text{dB}) = 20 \log_{10} f + 20 \log_{10} R + 20 \log_{10} \frac{4\pi}{C} - G_T - G_R \tag{9}$$

Here, f in Mhz, R in Km, gain in dBi.

The above function becomes.

$$P_L(\text{dB}) = 20 \log_{10} f + 20 \log_{10} R + 32.44 - G_T - G_R \tag{10}$$

The factors such as reflection, diffraction, and absorption create difficulty in measuring the power at this range.

2.3 AC to DC Rectifier

Two types of conventional rectifier structures exit such as diode full bridge rectifier and voltage doubler as illustrated in Fig. 10. The drawback of a conventional structures is the high value of diode turn-on voltage. Since the charging current is the piezoelectric current, iP , where $iP = IP \sin(\omega Pt)$ is a sinusoidal signal, it needs to

charge a parasitic capacitor C_P before a energy transducer can transfer the energy to the output. As a result, it causes a problem of wasting energy by charging the parasitic capacitor, C_P .

In addition, due to the diode turn-on voltage, which is in the range of 0.5 ~ 0.7 V for a conventional diode, an energy harvester has to overcome this voltage to transfer the power to the output or a load.

The available power from a conventional full bridge rectifier as well as a voltage doubler can be written as following equations [8],

$$P_{RECT} = 4C_P V_{RE} (V_P - V_{RECT} - 2V_d) \tag{11}$$

$$P_{RECT} = C_P V_{RE} (2V_P - V_{RECT} - 2V_d) \tag{12}$$

where

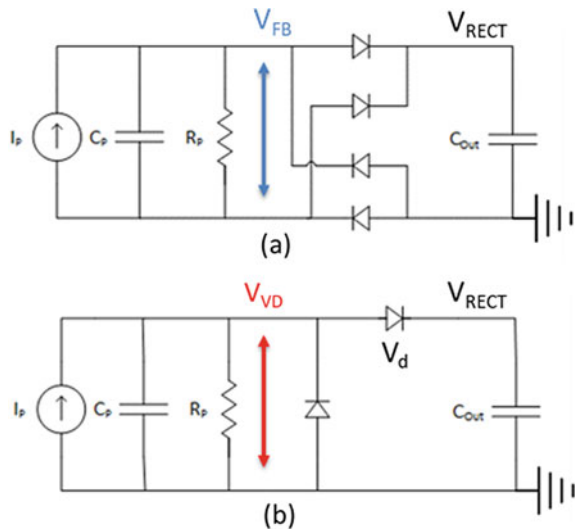
- V_P is the open-circuit voltage amplitude of energy harvester,
- f_P is the resonance frequency of energy harvester,
- V_d is the diode forward voltage drop,
- V_{RECT} is the output voltage of a rectifier,
- C_P is a parasitic capacitor.

The maximum power that can be harvested from those two topologies can be expressed by following Eqs. (7) and (8) [8]

$$P_{RECT, (MAX)} = (V_P - 2V_d) 2f_P \tag{13}$$

$$P_{RECT, (MAX)} = (V_P - 2V_d) 2f_P \tag{14}$$

Fig. 10 Conventional rectifiers. **a** bridge rectifier, **b** voltage doubler circuit [11]



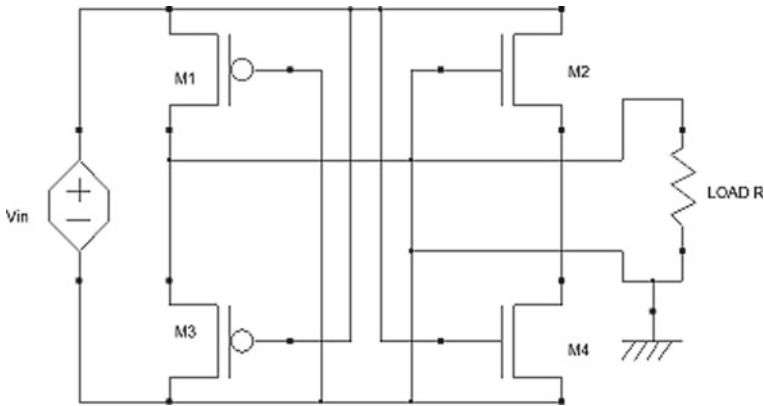


Fig. 11 Cross-coupled CMOS rectifier [9]

Rectifier build with CMOS using VLSI technology has resolved the problem of higher forward voltage drop of conventional diode rectifier. The most common configuration of CMOS rectifier is gate cross-coupled rectifier [9]. It has good power conversion efficiency but the most common problem is leakage power loss which need to be addressed with other configurations. Figure 11 shows a CMOS rectifier configuration using two PMOS and two NMOS transistors in cross-coupled configuration.

Both transistors rectify positive and negative AC input cycles. A very good power conversion efficiency can be achieved using suitable scalable technology.

3 Literature Review

3.1 Hassan Elahi [12]

Introduced the importance of self-powered or battery less web enable smart devices like processors, sensors, and communication hardware. Energy harvesting increases the efficiency and life time of these devices. This work emphasized the role of power management integrated circuits (PMIC) to minimize the power consumption of batteries which enhances the system's life span. Two important techniques are discussed, i.e., harvest-store-use and harvest-use energy harvesting architecture. This paper points out that energy storage elements are not required in harvest-use architecture in long-term use.

PMIC is primarily used to reduce the power consumption from the batteries. The efficiency, size, and cost are the factors to be considered in designing PMIC.

3.2 Venkat Subha Rao [13]

Various types of ultra-low power energy harvesting design techniques for IOT applications are discussed in this paper. It discussed about different level shifters circuits which can convert sub-threshold level signals to super threshold level signals. CMOS inverters buffer circuit is also fitted in output to improve the energy efficiency. The cross couple level shifter design generates large delays and area penalty. In current mirror type, high static current flows and output voltage is reduced. In current mirror with feedback transistor design, high static current developed. In level converter technique with pass transistor, the power consumption is high but speed is drastically increased. In single supply level shifter, the overall cost is reduced. Author has mainly focused on power dissipation and enhancing voltage level by using different energy harvesting design.

A comparative study for best design in terms of cost, efficiency, speed, and power dissipation needs to be addressed.

3.3 Nagraja [14]

This work has introduced two FinFET-based negative voltage rectifier for RF application using the topology of 4 T diode connected Si FinFET (F4TP) and 4 T cross-coupled Si FinFET (F2TP). As per the tabulated results [14], F2TP LVT Si FinFET negative rectifier shows better results at low RF input voltages. F2TP HVT negative rectifier achieves maximum power conversion efficiency of 92% at 915 MHz with $V_{ac}(RF)$ of + 0.4 V. The optimized parameter is $WD = WP = 75 \mu\text{m}$, $WT = WN = 5 \mu\text{m}$, $C1 = C2 = 5\text{PF}$, $CL = 10\text{PF}$, and $RL = 20 \text{K}\Omega$.

3.4 David Rivadeneira [15]

Author has presented the optimization and comparison of three active voltage rectifier (RF-DC) circuit design for energy harvesting systems. Design simulation is carried out on 90 nm CMOS technology with output resistance ranging from 50 to 500 K Ω . The different design is two-stage differential drive cross-coupled (DDCC) rectifier, DDCC rectifier cascaded with modified DDCC using body biasing technique for NMOS and PMOS and DDCC using 2 pairs of transistor. The input voltage and frequency for each design are 0.4 V at 950 MHz. The optimization results show that for all resistive loads in DDCC with and without body biasing, PCE% reaches 70%. Hence, circuits are good for full wave rectifier also. Design and simulation may be explored for other CMOS technology than 90 nm CMOS technology to achieve best results with more power conversion efficiency.

3.5 Xiaafei Li [16]

In this work, author proposed a CMOS passive rectifier using active bias tuning. This design allows an extended input range and achieves large power conversion efficiency. This design also compensates for voltage and temperature variation that leads to a better design for VHF operation. The rectifier design is fabricated in a 65 nm CMOS process. Conventional cross-connected rectifier achieves good efficiency in a very thin input power range. However, they can operate at high frequency and low input voltages. It achieves $PCE \geq 80\%$ for narrow range of inputs due to process variations and MOSFET threshold voltage. CMOS devices have threshold voltage of $V_{THN} = 0.46$ V and $V_{THP} = 0.41$ V.

3.6 Shuo Li [17]

In this paper, author has proposed the two-stage power management circuit. The proposed capacitive power management unit operates from tiny photovoltaic cells and consumes very low power. In this PMU design, a capacitive-based method is used. Also a photovoltaic energy harvester is used with capacitive power management unit on standard 0.18 μ m CMOS technology (Table 4).

4 Conclusion

The main challenge of power harvesting and power delivered by energy harvesters is very weak and unstable. Only ultra-low power devices can easily be coupled with them. Designing VLSI architecture, i.e., rectifiers and regulators for energy harvesting [18–21] with sub- μ A quiescent current consumption is challenging. Sometimes, it may happen that power to devices become intermittent or even stops completely. We should take required steps against sudden shutdown [22, 23]. Low power management units can provide solution in this case.

Table 4 Literature survey

S. No.	Publication and year	Topic	Software	Parameters				
				Width		Load (KΩ)	Output voltage	Power conversion efficiency (%)
				PMOS	NMOS			
A	MPDI Journal (2020) [12]	Energy harvesting toward self-powered IOT devices	Cadence –	–	–	2	–	70
B	(IJEAT) (2019) [13]	Different types of ultra-low power energy harvesting design techniques for IOT application	Cadence –	400 μm	200 μm	2	–	65
C	(ICCE).IEEE (2019) [14]	RF harvesting system for low power applications using FinFET	Cadence 20 nm FinFet	7.5 μm	7.5 μm	20	– 342 mV	92
D	(LASCAS) (2020) [15]	Optimization of active voltage rectifier/doubler designed for 90 nm technology	– 90 nm CMOS	12.2 μm	8.21 μm	50	816.8 mV	70
E	IEEE (2020) [16]	A VHF wide input range CMOS passive rectifier with active bias tuning	Cadence 65 nm HV	0.384 (chip area)	0.384 (chip area)	3	–7-7 dBm	64.4

References

1. Mahammad AH, Saad M, Salina AS, Aini H (2014) Energy harvesting for the implementable biomedical devices: issues and challenges. *Biomed Eng Online* 13:78
2. Roundy S, Leland ES, Baker J, Carleton E, Reilly E, Lai E, Otis B, Rabaey JM, Wright PK, Sundararajan V (2005) Improving power output for vibration-based energy scavengers. *IEEE Pervasive Comput* 4:28–36
3. Gyanenfra R, Fujii A (2016) Extracting knowledge from technological research papers in application of IoT. In: *Portland international conference on management of engineering and technology*, pp 2645–2651
4. Mitcheson PD, Green TC, Yeatman EM, Holmes AS (2004) Architectures for vibration-driven micro-power generators. *J Micro-electromech Syst* 13:429–440
5. Zhang Q, Gu L, Yang K, Halim MA, Rantz R, Roundy S (2016) Kinetic energy harvesting using improved eccentric rotor architecture for wearable sensors. *IEEE Sens*, pp 1–3

6. Roundy S (2005) On the effectiveness of vibration-based energy harvesting. *J Intell Mater Syst Struct* 16:809–823
7. Lee HJ, Zhang S, Bar-Cohen Y, Sherrit S (2014) High temperature, high power piezoelectric composite transducers. *Sensors* 14(8):14526–14552
8. Le T, Han J, von Jouanne A, Mayaram K, Fiez T (2006) Piezoelectric micro-power generation interface circuits. *IEEE J Solid-State Circ* 41(6):1411–1420
9. Tran L-G, Cha H-K, Park W-T (2017) RF power harvesting: a review on designing methodology and application. *Springer Open Access J:Micro Nano Syst Lett* 5(14). <https://doi.org/10.1186/s40486-017-0051-0>
10. Anand R, Chawla P (2020) A novel dual-wideband inscribed hexagonal fractal slotted microstrip antenna for C-and X-band applications. *Int J RF Microwave Comput Aided Eng* 30(9):e22277
11. Li S, Calhoun BH (2021) Sub-microAmp energy harvesting and power management units for self-powered IoT SoCs. In: 34th international conference on VLSI design and 2021 20th international conference on embedded systems (VLSID). ISSN: 2380-6923, <https://doi.org/10.1109/VLSID51830.2021.00011>
12. Elahi H, Munir K, Eugeni M, Atek S, Gaudenzi P (2020) Energy harvesting towards self-powered IOT devices. *MPDI J Energies*. *Energies* 13/5528. 103390/en13215528
13. Rao VS, Manchala J, Rajendra MD (2019) Different types of ultra-low power energy harvesting design techniques for IOT Application. *Int J Eng Adv Tech (IJEAT)* 9(1S5). ISSN: 2249-8958
14. Nagateja T, Srinivasulu A, Pal D (2019) RF harvesting system for low power applications using FinFET. In: Presented at IEEE international conference on consumer electronics (ICCE). ISSN: 2158-4001. <https://doi.org/10.1109/ICCE.2019.8661838>
15. Rivadeneira D, Villegas M, Procel LM, Lionel (2020) Optimisation of active voltage rectifier/doubler designed for 90nm technology. In: Presented at IEEE international conference on 11th Latin American symposium on circuits and systems (LASCAS). ISSN: 2473-4667. <https://doi.org/10.1109/LASCAS45839.2020.9069013>
16. Li X, Mao F, Lu Y, Martins RP (2020) A VHF wide input range CMOS passive rectifier with active bias tuning. *IEEE J Solid State Circ* 55(10):2629–2638. <https://ieeexplore.ieee.org/xpl/conhome/9052588/proceeding> ISSN:0018-9200, <https://doi.org/10.1109/JSSC.2020.3005814>
17. Li S, Calhoun BH (2021) Sub-microAmp energy harvesting and power management units for self-powered IoT SoCs. In: 34th international conference on VLSI design and 2021 20th international conference on embedded systems (VLSID). ISSN: 2380-6923, <https://doi.org/10.1109/VLSID51830.2021.00011>
18. Li S, Calhoun BH (2021) Sub-micro amp energy harvesting and power management units for self-powered IoT SoCs. In: 34th international conference on VLSI design and 2021 20th international conference on embedded systems (VLSID). ISSN: 2380-6923, <https://doi.org/10.1109/VLSID51830.2021.00011>
19. Paradiso JA, Starner T (2005) Energy scavenging for mobile and wireless electronics. *IEEE Pervasive Comput* 4(1):18–27
20. Roundy S, Wright PK, Rabaey J (2003) A study of low level vibrations as a power source for wireless sensor nodes. *Comput Commun* 26:1131–1144
21. Roundy S, Wright PK (2004) A piezoelectric vibration based generator for wireless electronics. *Smart Mater Struct* 13:1131–1142
22. Singh SK, Thakur RK, Kumar S, Anand R (2022, March) Deep learning and machine learning based facial emotion detection using CNN. In: 2022 9th international conference on computing for sustainable global development (INDIACom). IEEE, pp 530–535
23. Meelu R, Anand R (2010, November) Energy efficiency of cluster-based routing protocols used in wireless sensor networks. In: AIP conference proceedings, vol 1324, no 1. American Institute of Physics, pp 109–113

Area Efficient Design of In-Place RFFT Scaling for OFDM Applications



A. Padmavathi and G. L. Sumalata

Abstract An improvised 8-point butterfly unit is being proposed in this project, which is developed employing 4-point butterfly units to lessen complexity and storage demand. In-place RFFT designs are increasing in popularity because of their reduced complexity than pipeline topologies. We emphasized a conceptual design for developing a hardware efficient and low latency with reduced power consumption for in-place real-valued FFT in the traditional models. A butterfly block in an in-place way of implementation for real-valued fast Fourier transform (FFT) system conducts a succession of butterfly computation activity for each clock cycle. Higher butterfly block sizes make memory access conflict resolution more challenging. As a result, designers must examine the input information is getting flowed in the data path and memory accessibility of “in-place” real-valued FFT design tools in order to address a variety of performance challenges. Using this input, we introduce a new method for segregating the overall storage block into multiple smaller banks/segments (without influencing the size of storage memory) to avert memory management issues by interchanging information across internal banks at the same time. The real-valued rapid Fourier transform’s (RFFT) computation time has sped up dramatically in popularity and prompted excitement in this era because to its numerous research and applications DSP and surviving streams. As a consequence, as comparison to the prior strategy, the new methodology employs reduced area occupancy with shorter latency. Using the Xilinx ISE 14.7 tool, the synthesis and simulation outcomes are verified.

Keywords Fast fourier transform (FFT) · In-place computation · Butterfly unit—BFU · DSP · Real-valued FFT · Radix-r butterfly blocks

A. Padmavathi (✉) · G. L. Sumalata
Department of ECE, Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad,
India
e-mail: padmavathi.rs@gmail.com

G. L. Sumalata
e-mail: sumaprasadgl@gmail.com

1 Introduction

The discrete FT stands for “discrete Fourier transform” which is a technique for analyzing data, for transforming certain types of function sequences into other forms of representations. An approach for generating the discrete Fourier transform is the rapid Fourier transform of some sequence. The DFT takes a very long time to construct since it involves $N/2$ floating point multiplications. As i and “ k ” change, most of those multiplications are replicated. The FFT is a set of routines which are intended to minimize the number of repetitive processes. Each version of the FFT has its own set of attributes and benefits. The discrete Fourier transform, or DFT, is a transform that operates with a finite number of frequencies and a finite or discrete spatial signal. Many digital signal processing systems [13] use the discretized FT, which is accomplished using an efficient algorithm such as Cooley–Tukey Fourier algorithms. Such systems can process information in both the special domain and the frequency domain. Almost every application field of digital signal processing uses the FFT technique. FFT is extensively used on real-valued inputs in a various applications, including speech, audio, image, and multimedia processing. The rapid rise in biomedical field and the efficacious accomplishment of FFT of real-valued signals has managed to gain a tremendous interest recently due to its vast variety of applications in real-valued time-series analysis. Conjugate symmetry arises in the FFT of a real-valued signal, making half of the FFT outcomes redundant. Pipeline structure [1] (i.e., single-path delay feedback and multi-path delay commutator structures) and processor memory-based folded in-place design [12] implementations are the two basic structures of FFT designs. Some of these designs, in general, are composed of 3 functional units: (i) Butterfly CU (ii) memory block, and (iii) twiddle factor generating block. They are able to maintain a persistent data flow as a result of their consistency and improved BCU usage efficiency, MDC-based configurations are strongly suggested for FFT pipeline processing [6] above SDF-based structures. However, MDC-based structures have a somewhat higher memory demand. Inter or/and intra butterfly blocks and different stages of operations are folded in the BCU for low-complexity implementation in PM-based structures, which take use of in-place FFT. The decision between in-place systems that use a PM and a pipeline structure [9] is mostly determined by the target application’s space–time restrictions [7]. For effective implementation of FFT for various applications, numerous pipeline and PM-based structures were proposed. The RFFT [2] which implies real values FFT signals are employed in a variety of applications, including voice, audio, picture, and video processing.

1.1 FFT

In the area of audio and acoustics applications, “fast Fourier transform” is a crucial measurement design. It disintegrates a signal into its separate spectral components,

revealing the signal’s frequency. FFTs are employed in equipment and machines for malfunction diagnostic testing, quality management, and condition monitoring [1]. The discrete FT is a computational design method for communicating N-point length of complex sequence to its same N-point length of complex spectrum. Despite the fact that most FFT algorithms are designed to calculate the DFT of a complex sequence, the sequence to be converted is frequently real-valued in many applications. Using length-(N/2) complex FFT approach, the DFT of a length-N real-valued sequence may be calculated, which is commonly known. It is less generally recognized that by utilizing symmetries inside complex-valued algorithms, more efficient algorithms may be constructed (Fig. 1) [8].

The output sequence is decomposed into extremely small sub-segments using DIFFFT procedures. $X(k)$ is divided into even index and the converse which is odd digit index samples in this output sequence. The spectrum which is a domain of frequency, there the pattern is split in this procedure. DIFFFT uses a natural order input sequence. DFT should also be read backwards in time (Fig. 2) [8].

Fig. 1 General DIF FFT architecture

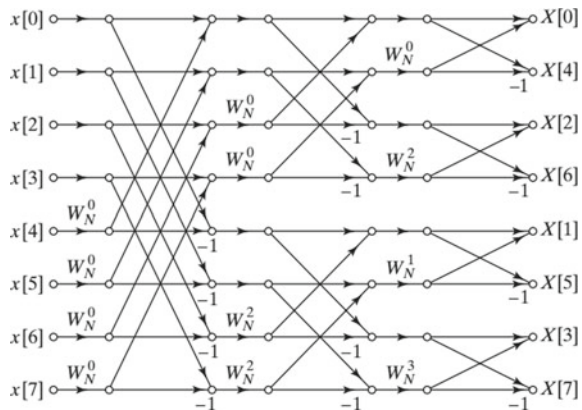
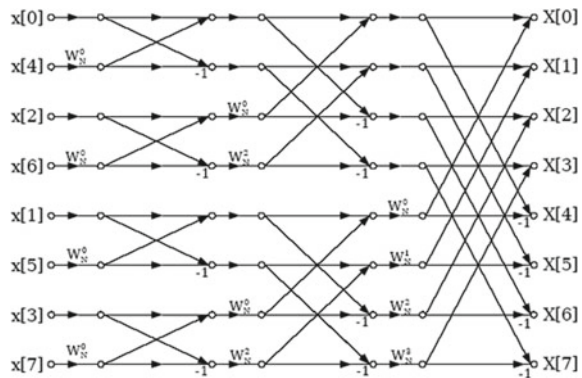


Fig. 2 General DIT FFT architecture



The fragmentation of the input pattern into smallest possible sub-segments is the premise of DITFFT algorithms. $x(n)$ is segregated into even and odd digit samples in this input signal. The time domain sequence is used to split the data. The input sequence of the DIT [3] FFT is bit inverted, whereas the output sequence is normal. The following is how the rest of the paper is organized: basic DFT overview followed by existing RFFT using radix-2 BFU in following part which is part II. Proposed architecture for RFFT using optimized radix-8 BFU in next part which is III. Comparison of existing with proposed along with schematics generated and output discussion in part IV. Concluded in part V.

2 Earlier Work

2.1 DFT Architecture

Let $x_0, x_1, x_2, \dots, x_n, x_{N-1}$ be input numbers. The N-point DFT can be derived by the below expression,

$$X_k = \sum_{n=0}^{N-1} x_n e^{-j \frac{2\pi k}{N} n}, k = 0, 1, \dots, N - 1 \tag{1}$$

Figure 3 [10] depicts the preexisting system for developing N-point in-place RFFTs [4]. A certain structure is composed of an arithmetic block (AU), a twiddle factors value storage block (TFSU), and a control unit along with data storing module. During each cycle, the AU receives an 8-sample block from the DSU, including a duo of twiddle factors (Fig. 3).

The current configuration for in-place method of calculating of RFFT is distinguished in the image above, which incorporates an arithmetic unit, data storage unit, twiddle factor storage unit, and control unit. The ALU units handle butterfly operations in FFT, while the data storage and twiddle factor storage unit's, respectively, record internal data and twiddle factor multipliers information. A control unit is crucial to be aware of this process. Figure 4 [5] shows a block level representation of the CU. It is constituting of counters and multiplexers, with the counter signifying the AND cell and the AC indicating the AND cell.

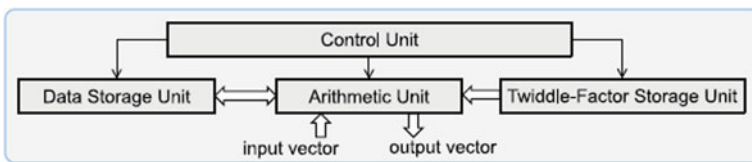


Fig. 3 Existing method of RFFT

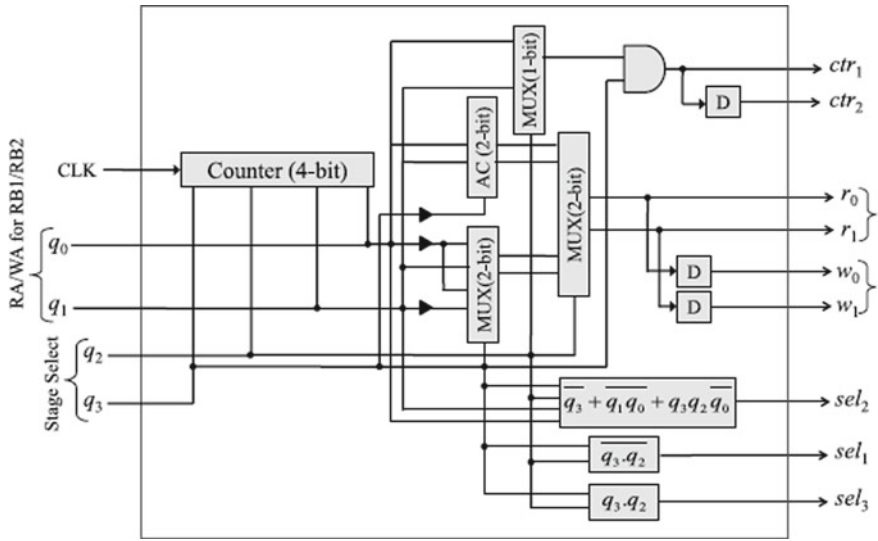
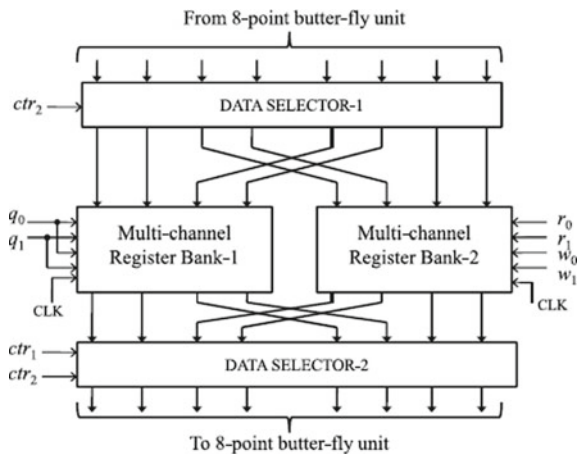


Fig. 4 Control unit (CU)

As several datasets from a single bank are required within a clock cycle, accessibility from that bank becomes complex and difficulty occurs. Adjusting samples from each data block in combinations before and after they are synchronously recorded and fetched in and out of the register banks can solve the accessing issue [11]. The storage unit makes use of two data-swapping circuits to handle bank conflicts and organize input-block samples (DS1 and DS2) (Fig. 5) [14].

According to previous research, the building of a TF memory unit in this structure for 32-point and 4-point blocks for butterfly operations that requires an 8-word

Fig. 5 Memory unit of existing RFFT structure



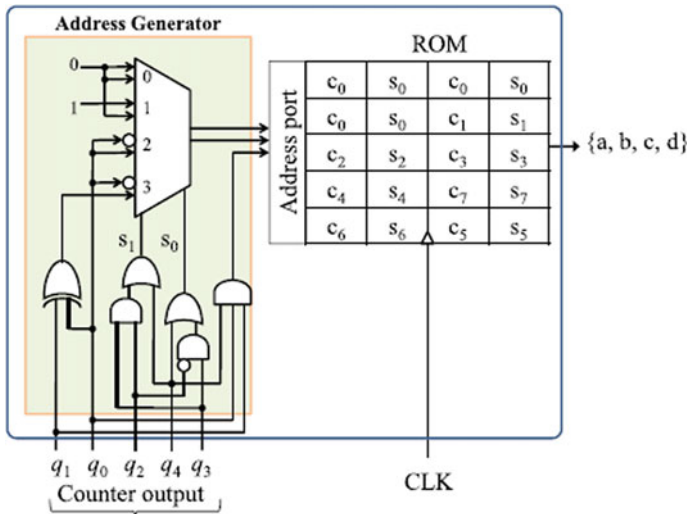


Fig. 6 Twiddle factor storage unit

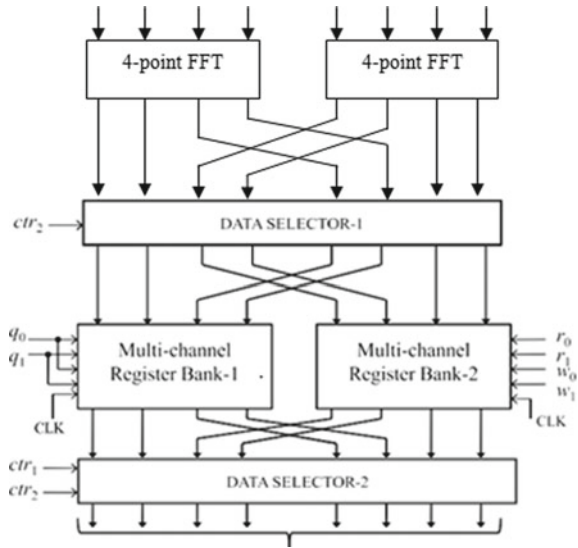
ROM look-up-table (LUT). However, we discovered that the TF memory maintains some unnecessary values, which can be removed such that minimize memory space. Figure 4 demonstrates a modified twiddle factor values storing block architecture for a 32-point real-valued fast FT [5] with an 8-block size. It uses one 5 4w ROM LUT, in this, “w” is “word length” of the real twiddle factor constants c_1 , s_1 , and c_1 , s_1 are the actual twiddle factor constants. The control bits s_1 , s_0 of the MUX are generated by encoding the 5 BF stages of a 32-point FFT with a 3-bit word q_4 , q_3 , q_2 for picking bottom address of each BF stage which is of 2 bit width (Fig. 6) [14].

3 Proposed Work

A novel method of designing from an 8-point input FFT unit is proposed in the proposed work, as illustrated below. To implement 32-point FFT, we utilize this design. The number of twiddle factors and design complexity will be somewhat increased if these methods are used.

A design methodology for developing power efficient and with reduced hardware complexity design for “in-place RFFT” is described in the existing technique. So, in this design, the 8-point FFT was split into two 4-point FFT multiple delay commutators in the suggested approach, as illustrated in Fig. 7. The suggested structure has a basic butterfly structure and requires less memory, resulting in a smaller footprint and faster response time. A few changes have been made to this twiddle factor unit in order to comply with the 4-point FFT, which requires less memory.

Fig. 7 Proposed memory unit with 4-point FFT



3.1 Memory/Storage Block

In this, a framework is designed for N number of points in-place DIT RFFT computing which is identical to that proposed at the block level. SU and BCU, on the other hand, are not designed in the same way. Figure 7 depicts the suggested structure’s SU design for RFFT using 32-point with a size of butterfly unit. This system is made up of two MCRBs and a couple of selection blocks for data inputs. “Multiple channel register banks” which are in short “MRCB” define single device that integrates several memory banks with shared address decoding circuitry. Register banks (B1, B2, B5, and B6) are implemented in the MCRB-1, whereas register banks (B1, B2, B5, and B6) are implemented in the MCRB-2 (B3, B4, B7, and B8). The internal construction of the MCRB is seen in Fig. 8 [14], which has four bank channels with a depth of four.

In order to obtain address for selection of twiddle factors, we use a 2 input decoder, and the 2 inputs of decoder be the write-address (w_0, w_1) and activate the clock signal CLK for a certain number of registers in a row-wise corresponding 4 separate register banks/blocks for writing input 1 unit of information $\times 1, \times 2, \times 3,$ and $\times 4$. The content of each bank’s four registers may be accessed via the multiplexer, which selects one of them on the basis of 2-bit read address (r_0, r_1). 4 data values are interpret from 4 registers of 1 MCRB throughout each clock period, culminating for one unit of 8 data being MCRB-1 and MCRB-2 were fetched to accomplish i -th stage BF process, and transitional results are returned to original positions over upcoming clock period change for the next stage of BF operations from $i = 0$ to n and n denotes number of stages.

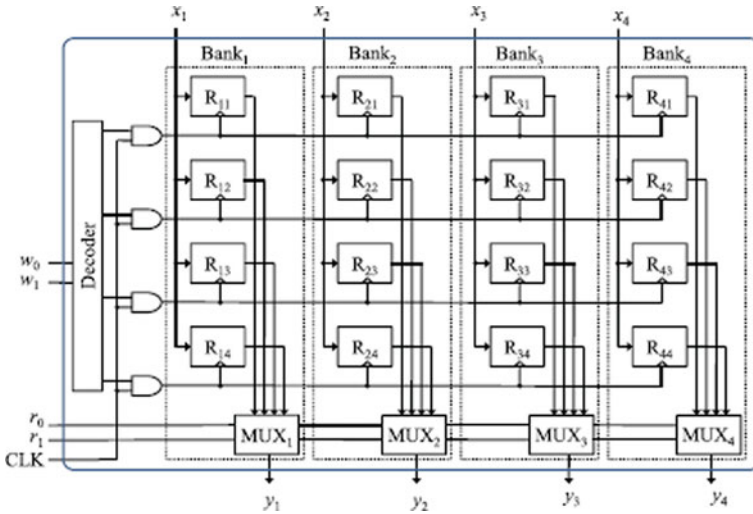


Fig. 8 Multi-channel register bank structure

The internal structure of data selector-1 and data selector-2 is given Fig. 9. Every input (Ii) and output (Oi) lines on DS-1 and DS-2 receives and delivers a two input samples from and to the BCU in order to process (013). The data switching procedure between DS-1 and DS-2 can be selected by the ctr2 input signal. Ctr1 is given 1 during the DS-1 changeover mechanism. As shown in Fig. 9, MUX1, MUX2, and MUX3, MUX4, and MUX5, and MUX6, swap data to negate DS-1’s swapping operation (b). The BCU design block size 4 is similar to that of the arithmetic unit, except for the inclusion of two multipliers involving complex data and 2 pairs of line-changers LC1 and LC2.

4 Experimental Results

Figure 10 is RTL schematic diagram of the implemented design. RTL refers to register transfer level. This is the schematic produced by the tool based on the code designed by the user as per the top module of the circuit.

Figure 11 shows the simulation results of waveforms for implemented design using test bench. In this, we observe that based on the clock signal and reset signal, for every clock signal, the input of FFT inputs is taken and control signals are generated which retrieves the twiddle factor values and the FFT output is generated which is the final output FFT process.

Table 1 describes the comparison of performance metrics like as resource consumption and latency between FFT algorithm implemented with Radix-2 (Existing) and Radix-8 using Radix-4 butterfly unit (Proposed). From these results,

Fig. 9 **a** First data block selector. **b** Second data block selector

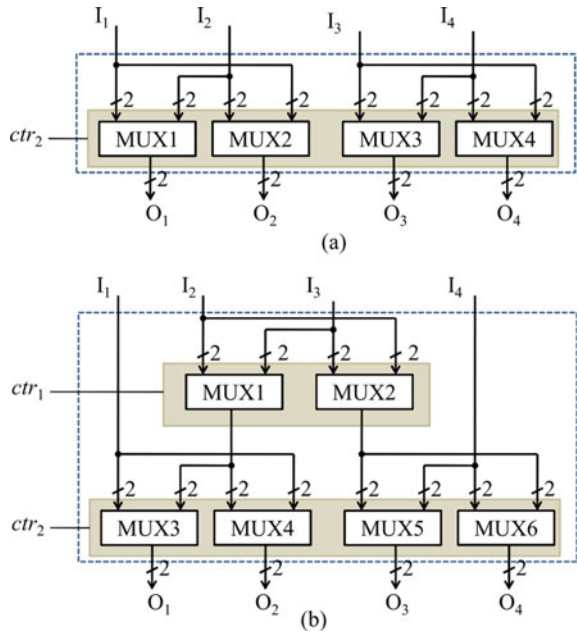
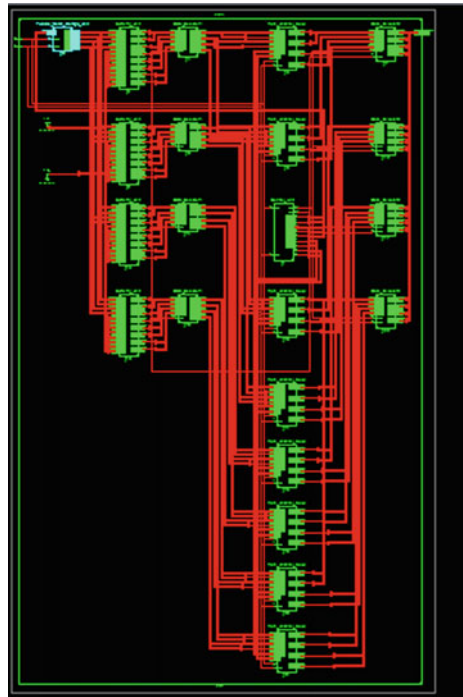


Fig. 10 RTL schematic of FFT



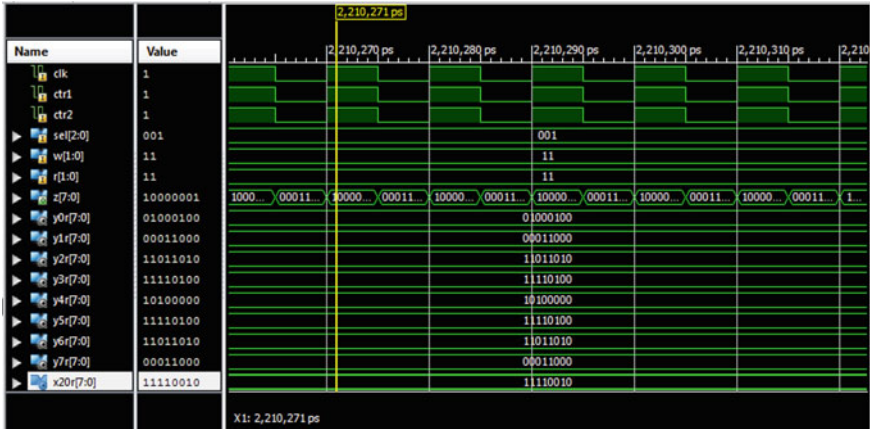


Fig. 11 Simulation results for FFT

Table 1 Evaluation table for existing vs proposed

Parameter	Existing FFT	Proposed FFT
Delay(ns)	39.458	19.62
Area(LUT's)	10,746	4489

we conclude that the delay is reduced in proposed method by using higher radix implementation.

The area efficiency is also increased in case of proposed FFT. This is due to the decreased complexity of higher radix butterfly architecture using Radix-4 implementation. Such that a trade-off between area and delay is established between existing FFT algorithm and proposed FFT algorithm.

5 Conclusion

This paper provides a design technique for creating a high-throughput, large-FFT-size in-place RFFT architecture. In existing design, the 8-point butterfly unit is proposed and given as input to storage device to perform storing and data-swapping operations which complexity is more. In proposed design, a novel 8-point FFT architecture implemented by using two 4-point FFT and design is implemented. The above results show the newly implemented design consumes less area and less delay when compare to existing design. The synthesis and simulation results have been verified by using Xilinx ISE 14.7 tool.

References

1. Chang Y-N, Parhi KK (2003) An efficient pipelined FFT architecture. *IEEE Trans. Circ Syst. II, Analog Digit Sig Proc* 50(6):322–325
2. A.Wang and A. P. Chandrakasan (2003) Energy-aware architectures for a real valued FFT implementation. In: *Proceeding of international symposium on low power electronics and design*, pp 360–365
3. Meher PK, Mohanty BK, Patel SK, Ganguly S, Srikanthan T (2015) Efficient VLSI architecture for decimation-in-time fast Fourier transform of real-valued data. *IEEE Trans Circ Syst I Reg Pap* 62(12):2836–2845
4. Sorensen H, Jones D, Heideman M, Burrus C (1987) Real-valued fast Fourier transform algorithms. *IEEE Trans Acoust Speech Sig Process ASSP-35*(6):849–863
5. Ayinala M, Lao Y, Parhi KK (2013) An in-place FFT architecture for real-valued signals. *IEEE Trans Circ Syst II Exp Briefs* 60(10):652–656
6. Wold EH, Despain AM (1984) Pipeline and parallel-pipeline FFT processors for VLSI implementations. *IEEE Trans Comput C-33*(5):414–426
7. Hsiao S-F, Shiue W-R (2000) Design of low-cost and high-throughput linear arrays for DFT computations: algorithms, architectures, and implementations. *IEEE Trans Circ Syst II Analog Digit Sig Process* 47(11):1188–1203
8. Garrido M, Parhi KK, Grajal J (2009) A pipelined FFT architecture for real-valued signals. *IEEE Trans Circ Syst I Reg Papers* 56(12):2634–2643
9. Ayinala M, Brown M, Parhi KK (2012) Pipelined parallel FFT architectures via folding transformation. *IEEE Trans Very Large Scale Integr (VLSI) Syst* 20(6):1068–1081
10. Chi H-F, Lai Z-H (2005) A cost-effective memory-based real-valued FFT and Hermitian symmetric IFFT processor for DMT-based wire-line transmission systems. In: *Proceeding IEEE Int Symp Circ Syst* 6:6006–6009
11. Johnson LG (1992) Conflict free memory addressing for dedicated FFT hardware. *IEEE Trans Circ Syst II Analog Digit Sig Proc* 39(5):312–316
12. Jo BG, Sunwoo MH (2005) New continuous-flow mixed radix (CFMR) FFT processor using novel in-place strategy. *IEEE Trans Circ Syst I, Reg Papers* 52(5):911–919
13. Kuang SR, Wang JP, Guo CY (2009) Modified booth multipliers with a regular partial product array. *IEEE Trans Circ Syst II Exp Briefs* 56(5):404–408
14. Mohanty BK, Meher PK (2019) Area-delay-energy efficient VLSI architecture for scalable in-place computation of FFT on real data. *IEEE Trans*

Design and Implementation of Comparator Using GDI Decoder



Sangeeta Singh, T. Akankhsa, Y. Vamshi, Shubham Munratiwar, and M. Venkata Jayanth

Abstract Mixed logic designs are given a higher priority because they provide a simpler mechanism for analyzing digital circuits. Similar logic expressions and logic relations are produced when mixed logic notation is used correctly. A circuit's actions can also be viewed using a mixed logic implementation. In this post, we looked at mixed logic architectures including pass transistors (DVL), transmission gates (TGL), and static CMOS. A 2:4-line decoder in CMOS technology takes 20 transistors, while with mixed logic, by using 14 transistors the same 2:4 decoder will be created. 4:16 line decoders are also built using these unique mixed logic topologies. GDI is a novel technique for developing low-power digital sequence circuit. This method allows for low-power consumption, transmission delay, and region of digital circuits whereas keeping a low logic uncertainty. Furthermore, this unique mixed logic topology is used to create 4:16 line decoders utilizing a 4:16 line mixed line decoders a 2-bit comparators was constructed using the mixed logic technique. All of the proposed circuits contain fewer transistors and are developing utilizing the gate diffusion input technique. These logics demonstrate that transistor count, consumption, and time may be related to the satisfactory level.

Keywords Line decoders · Mixed logics · Power delay optimization

1 Introduction

CMOS technology circuits are used to create the majority of logic gates inside an integrated circuit. A pull up networking for PMOS and a pull-down network for NMOS make up the Complementary Metal Oxide Semiconductor (CMOS) circuits, which offer good efficiency in comparison with the device fluctuations and noise. In CMOS circuit, inputs are connected only to transistor gate nodes, leading to fewer designs and cell-based circuit design and fabrication. The goal of Pass Transistor Logic (PTL) was to provide an alternative to CMOS logic. In comparison with CMOS

S. Singh (✉) · T. Akankhsa · Y. Vamshi · S. Munratiwar · M. V. Jayanth
Vardhaman College of Engineering, Hyderabad, Telangana, India
e-mail: sangeethasingh@vardhaman.org

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
V. V. S. S. Chakravarthy et al. (eds.), *Advances in Signal Processing, Embedded Systems and IoT*, Lecture Notes in Electrical Engineering 992,
https://doi.org/10.1007/978-981-19-8865-3_60

685

logic performance, power, and area, all improved by using PTL [1]. Pass transistor circuits are distinguished by the fact that the inputs are linked to the transistors' gate or source or drain diffusion connections.

Pass transistor circuits can be implemented in two different ways. Individual transistors, after known as PMOS or CMOS transistors are utilized in the first technique, but the transmission gate method uses a parallel combination of PMOS and NMOS transistors. The decoder is straight-forwarded a combinational design that takes and it turns it to a series of output signals. Line decoders are used widely in a wide range of the applications, including memory array address decoding and data demultiplexing, displays with Seven segments, and microchip/microcontroller-based frameworks [2–7]. Address decoders are crucial in SRAM Memory blocks since the power consumption and response time are substantially determined by the decoder design. A 4-16 decoder takes 4 inputs and outputs 16 min-terms, D0 through D15. The 16 complementary min-terms I0—I15 are produced using an inverted 4-16 decoder. Conventional CMOS circuitry, 4 input NOR or NAND gates are required to build 4 to 16 decoders. However, a pre-decoding technique can be utilized for a more efficient implementation, in which n basic inputs are pre-decoded within one of $2n$ pre-decoded segments that directly or indirectly contribute to next stage decoder. A non-inverting 4-16 decoder is constructed using the pre-decoding technique and needs two 2-4 inverting decoders, and 16 2 input NOR gates. Similarly, two 2-4 non-inverting decoders and the 16 2 input NAND gates are required to create an inverting 4-16 decoder. As a result, decoders with traditional CMOS logic need 20 transistors for 2 to 4 decoders and transistors count 104 for 4 to 16 decoders.

2 Existing Method

This section presents the existing technique available to design decoder circuits.

2.1 Transistor Topology

Designing a mixed logic system with four (TGL) or OR gates with 2 inverters and a total of 16 transistors would be required for a 2-4 decoder. Either of the two inverters can be omitted a decoder architecture of 14 transistors was created by merging both DVL of logic AND gates and TGL having similar design and selecting the desired control and propagation signals, which results in a decoder architecture of 14 transistors. Let A and B are the two inputs taken to the decoder, which results in four min-terms D0-D3. In order to get rid of inverter B, the first and the third min-terms, D0(AB) and D2(AB), are accomplished by utilizing A and B as the propagation signals to the DVL (AND) Gates. The TGL (AND) Gate is used to implement the min-terms D1(AB) and D3(AB), for both min-terms, B is the transmit signal. With this combination of gates and inputs, the B inverter can be removed, resulting in

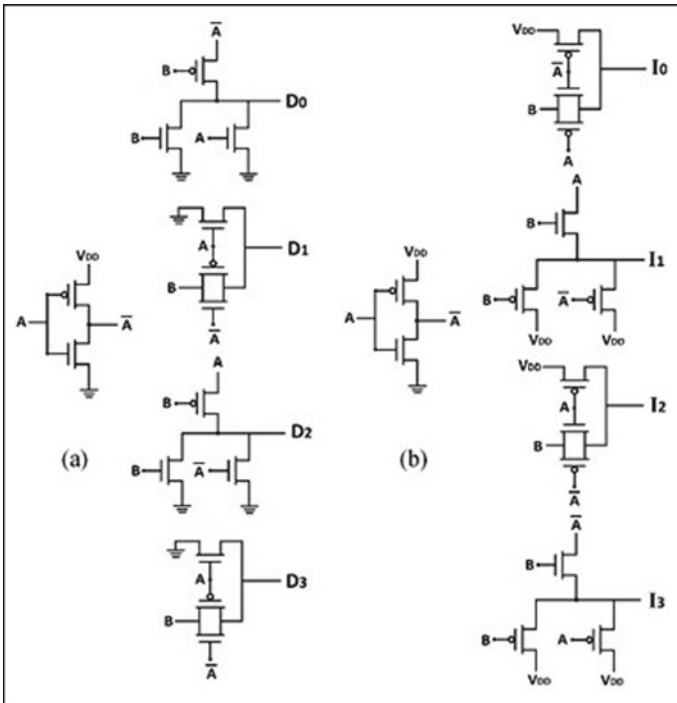


Fig. 1 2:4 Decoder using mixed logic

a decoder consisting of 14 transistors. Similarly, a 14 transistors topology using four TGL/DVL (OR) Gates and one inverter can be used to create a 2-4 inverting decoder. To implement min-terms I0 and I2, TGL (OR) Gates with B as propagating signal input a utilized, whereas DVL (OR) Gates with input A are used to implement min-terms I1 and I3. As the propagate signal, “2-4 LP” and “2-4 LPI” are new two low-power decoder circuits implemented in this segment, with LP signifying low power and I denoting inverting (Fig. 1).

3 Literature Survey

Different mixed logics are used to create line decoders in [8]. The construction of a comparator utilizing standard CMOS and the mixed logic methodologies was presented in this work, along with a comparative study of all their parameters both in terms of design and simulation outcomes. The 4:16 line decoder used fewer transistors in [9]. Using TGL and DVL gates they constructed a 4:16 line decoder to save space. A comparator with a lower average power consumption and a shorter time by utilizing this less transistor line decoder is desirable. Thus to minimize the size

of the circuit by utilizing this line decoder, and also using this 4:16 line decoder, an area-efficient comparator has been proposed in our work. The paper [10] proposed 4:16 line decoder to make comparison. They built 2-4 as well as 4-16 Line Decoders with TGL and (DVL) Gates to reduce space. But the work presented in this attempted to create a comparator with a lower average power consumption and a shorter latency utilizing this less transistor line decoder [11–13]. We can lower the size of the circuit by utilizing this line decoder, and we can also create an area efficient comparator by employing this 2:4 and 4:16 line decoder. By merging both logic designs and low-power comparator, the paper [14] suggested a high performance magnitude comparators. They provided a high-speed magnitude comparator in their article. They were created to boost comparator's speed. This can teach us how to make a high-speed, elevated comparator. We built the comparator with extremely efficient line decoders to improve speed with less region and time. This IS consists of an LBT-based Assisted Access License (LAA) [15] that includes hardware such as load-based gear (LBE) and frame-based equipment (FBE) that can be built to compete with Wi-Fi-based contact instruments for acceptable access on a shared channel. Calculate the performance of two recently suggested 3GPP medium access control (MAC) and Wi-Fi-based methods in a hybrid scenario with a variety of parameter combinations in this study. To construct address decoder [16] required to consider two things, first selecting the optimal circuit technique and second sizing of the transistor. The purpose of pass transistor logic [17] is to reduce the number of transistors required in DVL while maintaining logic speed. Instead of charging the nodes via VCC and then releasing them to GND, this logic transmits the charge between them.

4 Proposed Work

4.1 Application of Line Decoder

As a 2-Bit Comparator, a 4-16 low-power line decoders is used. A 2-Bit comparator was created utilizing a 4-16 line decoder and the mixed logic design method's pass transistor and transmission Gates. The low-power output. The OR gates are coupled to the 4:16 line decoders which are designed using three gates: a pass transistor and a transmission Gate. The input lines of the line decoder are A, B, C, and D, while the output lines are D0 to D15. Two 2:4-line decoders were used to create a 4:16 low-power line decoder. TGL and DVL mixed logic methods are used to design the line decoders.

Static CMOS logic was used to construct the 4:16 line decoder. The output lines of the 4-16 line decoder are taken as input of OR gate to perform a comparison operation as shown in Fig. 2. The output lines D2, D1, D6, D3, D7, and D11 are attached to the OR gate to acquire the comparator output of $A = B$. (out 17). To make $A = B$, D0, D5, D10, and D15 are connected to the second (OR) Gate (out 18). To get $A > B$ output, these are attached to the third (OR) Gate. To get output $A > B$, D8, D4, (D12),

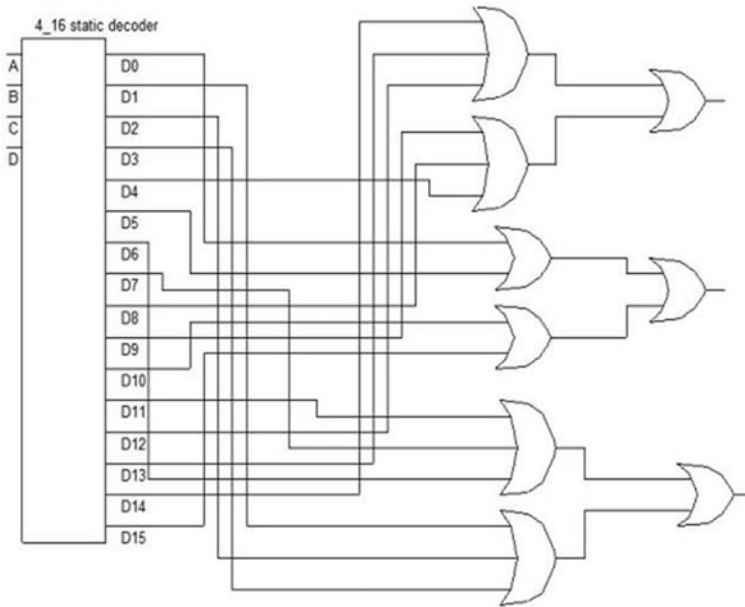


Fig. 2 2-bit comparator using CMOS 4:16 line decoder

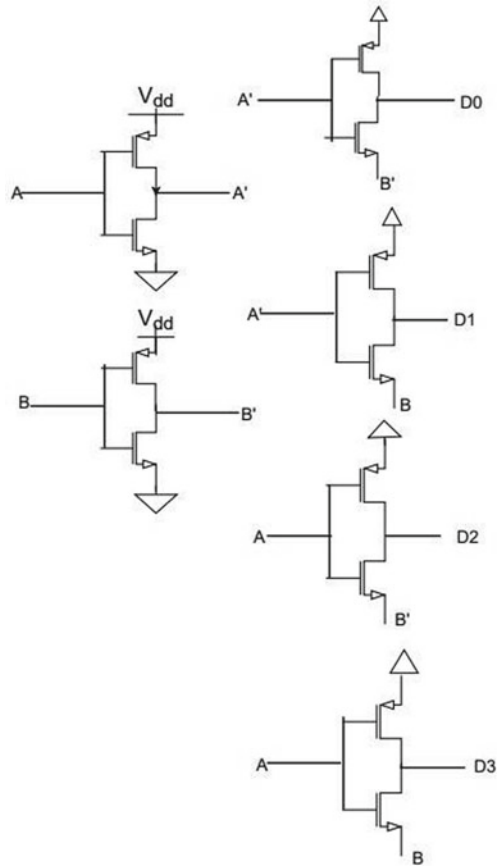
D9, D14, and D13 are connected to the third (OR) gate (out 20). The comparator can be constructed by using the gate diffusion input technique (GDI)—a new digital low-power sequential design circuits method is presented. These techniques allow digital circuits to use less power, have shorter propagation delays, and take up less space while keeping a low level of logic complexity.

4.2 12 Transistor 2:4 GDI Decoder

By using the GDI technology, the 2:4 decoders can be implemented with only 12 transistors as shown in Fig. 3, using this GDI technology we have reduced the count of the transistors, and this technique is also used to reduce the power consumption and propagation delay.

The circuits are implemented using Tanner tools and the corresponding results have been verified. Figure 4 presents the schematic of 2:4 decoder obtained using 14 transistors and the 2-bit comparator schematic is presented in Fig. 5 (Figs. 6 and 7; Table 1).

Fig. 3 12 Transistor 2:4 GDI decoder



5 Conclusion

The work presented here used a mixed logic strategy to design the comparator, which resulted in better performance measures such as lower power consumption and faster rise and fall times. A 4:16 static CMOS line decoder and a 4:16 mixed logic line decoder is used to accomplish this. By reducing the transistor count, we were able to reduce the area, power, and latency by using NAND, NOR, and NOT gates instead of OR and AND gates. In our suggested design, a line decoder using TGL and DVL with CMOS logic has been implemented and is compared it to a traditional CMOS design technique. Finally, the comparator is optimized for power that have been well integrated on the layout level. Comparator was implemented using 4-16 mixed logics line decoder and GDI decoder. Comparative analysis was done using tanner tool.

Fig. 4 14 Transistor 2–4 line decoder

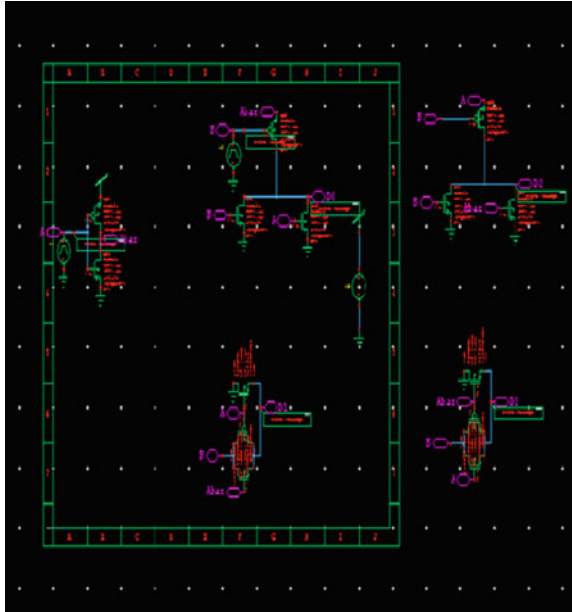


Fig. 5 2-bit comparator using mixed logic line decoder

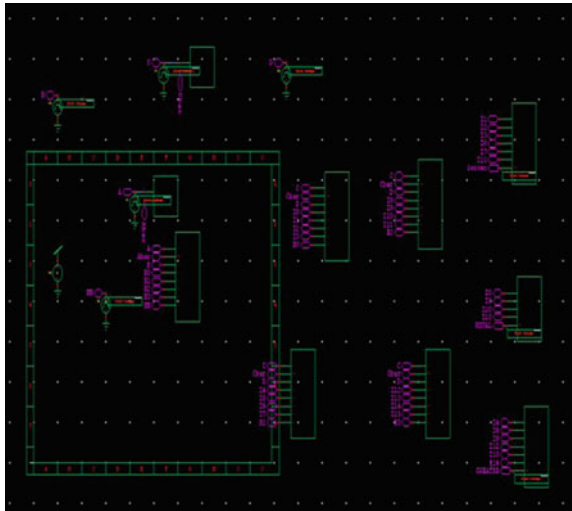


Fig. 6 2-4 Line decoder using GDI

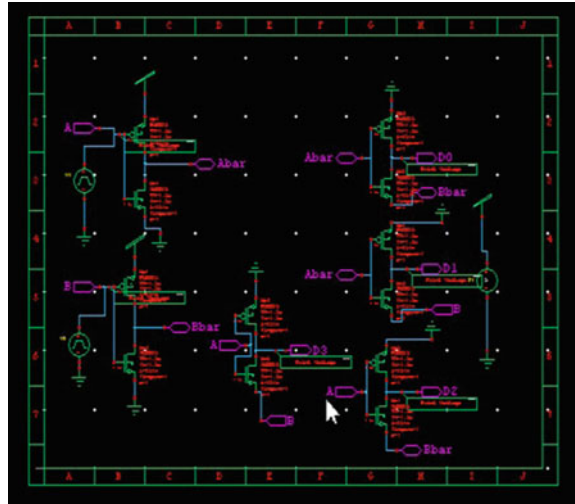


Fig. 7 2-bit comparator using GDI decoder



Table 1 Comparative analysis

Design	Number of Transistors	Time Delay (s)
2:4 mixed logic line decoder	14	0.78
2: GDI decoder	12	0.68
2-bit comparator with 4:16 mixed logic line decoder	199	3.43
2-bit comparator with 4:16 GDI decoder	153	1.20
2-bit comparator with 4:16 conventional decoder	226	4.11

References

1. Zimmermann R, Fichtner W et al (1997) Low-power logic styles: CMOS versus passtransistor logic. *IEEE J Solid-state Circuits* 32(7):1079–1090
2. Camus V, Schlachter J, Enz C (2016) A low-power carry cut-back approximate adder with fixed-point implementation and floating-point precision. In: 2016 53rd ACM/EDAC/IEEE Design Automation Conference (DAC), IEEE, pp 1–6
3. Schlachter J, Camus V, Enz C, Palem KV (2015) Automatic generation of inexact digital circuits by gate-level pruning. In: 2015 IEEE International Symposium on Circuits and Systems (ISCAS), IEEE, vol 57, pp 173–176
4. Camus V, Schlachter J, Enz C, Gautschi M, Gurkaynak FK (2016) Approximate 32-bit floating-point unit design with 53% power-area product reduction. In: ESSCIRC Conference 2016: 42nd European SolidState Circuits Conference, IEEE, pp 465–468
5. Prabu AV, Sateesh Kumar G (2019) Hybrid MAC based adaptive preamble technique to improve the lifetime in wireless sensor networks. *J Adv Res Dyn Control Syst* 11(1):240–249
6. Soumya N, Sai Kumar K, Raghava Rao K, Rooban S, Sampath Kumar P, Santhosh Kumar GN (2019) 4-bit multiplier design using cmos gates in electric VLSI. *Int J Recent Technol Eng* 8(2):1172–1177
7. Rooban S, Saifuddin S, Leelamadhuri S, Waajeed S (2019) Design of fir filter using wallace tree multiplier with kogge-stone adder. *Int J Innov Technol Exploring Eng* 8(6):92–96
8. Balobas D, Konofaos N (2016) Design of low-power high performance 2–4 and 4–16 mixed-logic line decoders. *IEEE Trans Circuits Syst II: Express Briefs* 64(2):176–180
9. Sharma A (2019) optimizing power and improving performance of 4-16 hybrid-logic line decoder using power gating technique. In: 2019 4th international conference on Recent Trends on Electronics, Information, Communication & Technology (RTEICT), IEEE, pp 510–513
10. Meena S, Augustin Jacob J, Poornalakshmi M, Preyanga S (2018) Design of low power, area efficient 2–4 mixed logic line decoder. In: 2018 second International Conference on Electronics, Communication and Aerospace Technology (ICECA), IEEE, pp 1536–1540
11. Lo JC (1993) Novel area-time efficient static cmos totally self-checking comparator. *IEEE J Solid-State Circuits* 28(2):165–168
12. Sharma G, Arora H, Chawla J, Ramzai J (2015) Comparative analysis of a 2-bit magnitude comparator using various high performance techniques. In: 2015 International Conference on Communications and Signal Processing (ICCSP), IEEE, pp 0079–0083
13. Rooban S, Lakshmi Swathi K, Monica Ch, Shivaramakrishna B (2019) An odd parity generator design using nanoelectronics. *Int J Eng Adv Technol* 8(4):597–601
14. Sharma G, Arora H, Chawla J, Ramzai J (2015) Comparative analysis of a 2-bit magnitude comparator using various high performance techniques. In: 2015 International Conference on Communications and Signal Processing (ICCSP), IEEE, pp. 0079–0083
15. Varakumari S, Prabu AV, Gopiram K, Venkatesan S (2017) Coexistence and fair access on the shared channel for lte-u and wifi. *J Adv Res Dyn Control Syst* 9(6):728–744
16. Mishra AK, Acharya DP, Patra PK (2014) Novel design technique of address Decoder for SRAM. In: 2014 IEEE international conference on advanced communications, control and computing technologies, IEEE, pp 1032–1035
17. Oklobdzija VG, Duchene B (1995) Passtransistor dual value logic for low-power CMOS. In: 1995 international symposium on VLSI technology, systems, and applications. Proceedings of Technical Papers, IEEE, pp 341–344

High Speed Efficient Three Operand Adder



N. Udaya Kumar, K. Bala Sindhuri, S. S. Harsha Varma, K. Asha Shaini, K. Sri Hari, and K. Sai Sowmya

Abstract Today's digital world is accompanied by digital electronics, and anything with low power consumption and fast response time is in high demand. The devices which perform basic arithmetic operations like addition, multiplication, and subtraction are used in many processors, digital filters, and digital communication. Adders are used in multipliers to produce the result by adding the partial products in the final stage and in many applications. Most partial multiplier products are added using two operand adders, although addition using three operand adders is faster. Therefore, an efficient high speed three operand adder is designed and comparison of the performance metrics of the designed adder with the existing adders using Xilinx Vivado 2017.2 software tool is performed.

Keywords Carry save Adder (CSA) · Ripple Carry Adder (RCA) · Linear Carry Select Adder (CSLA) · Square Root Carry Select Adder (SQRT CSLA) · Han-Carlson Adder (HCA) · Hybrid Han-Carlson Adder (HHCA) · High Speed Three operand Adder · Propagate and Generate Logic (PG logic)

N. Udaya Kumar · K. Bala Sindhuri · S. S. Harsha Varma · K. Asha Shaini · K. Sri Hari (✉) · K. Sai Sowmya
Department of E.C.E, S.R.K.R Engineering College, Bhimavaram, India
e-mail: kusumanchisrihari9@gmail.com

N. Udaya Kumar
e-mail: nuk@srkrec.ac.in

K. Bala Sindhuri
e-mail: kbsindhuri@srkrec.ac.in

S. S. Harsha Varma
e-mail: harshavarma5073@gmail.com

K. Asha Shaini
e-mail: kashashaini@gmail.com

K. Sai Sowmya
e-mail: kondasaisowmya@gmail.com

1 Introduction

Implementation of VLSI architectures in the software tools helps to realize the design efficiency. Furthermore, it helps to design the new architectures and test them to know if they have optimized performance or not based on the results generated like power consumption, delay, area, etc. Adders are employed in arithmetic logic units, or ALUs, included in many electronic devices and microprocessors. Additionally, they are employed in the other processor components, where they can be used to store programs or instructions and calculate the addresses of program counters and stack pointers to increment and decrement the addresses. Adders are also employed in operators and many kinds of operations by enhancing the performance of adders, multipliers can also employ adders to add the partial products produced while multiplying numbers. Improvements have been acquired in the performance metrics such as area, power consumption, and latency by incorporating the efficient adders. A survey of two operand adders, such as the Ripple Carry Adder (RCA), the Carry Select Adder (CSLA), the Square Root CSLA (SQRT CSLA), the Han-Carlson Adder (HCA), and the Hybrid Han-Carlson Adder (HHCA), is conducted to analyze the performance of these adders such that these adders can be incorporated in three operand adders to further improve the performance. A three-operand adder is used to add three binary values, such as a , b , and c , of any number of bits. A review of three-operand adders is also conducted; some of these adders use two stages of two operand adders while others only have a single three-operand adder. The new architecture of the three-operand adder is designed by modifying the techniques and the logics in the existing three operand adder, i.e., High Speed Efficient Three Operand Adder. In this paper, Sect. 2 deals with review of different adders. Section 3 deals with the study of three operand adders. Section 4 deals with the design of new VLSI architecture of High Speed Efficient Three Operand Adder (modified adder). Sections 5 and 6 deal with the results, discussion, and conclusion.

2 Review of Contemporary Adders

The two fundamental adders are Half Adder and Full Adder in which Half Adder adds two inputs A and B with a single binary digit each; and obtained outputs are sum (S) and carry (C). Similar to that, a full adder adds three single-digit binary inputs, A , B , and C , and obtained outputs are sum (S) and carry (C).

2.1 *Ripple Carry Adder (RCA)*

Ripple Carry Adder is an adder which is used to perform the addition of binary numbers of any number of bits. RCA contains multiple full adders cascaded in

parallel which can be used to add n-bit binary numbers [1]. The number of full adders in an RCA depends on the number of bits of binary numbers. The Ripple Carry Adder logically signifies that the carry is rippled at every stage since each full adder is cascaded, and the carry-out of each full adder serves as the carry-in of the immediately succeeding full adder [2]. This adder may experience a delay since each active full adder must wait for the completion of the previous addition before performing the next addition. The interval between the input and the occurrence of the carry signal is known as the carry propagation delay.

2.2 Linear Carry Select Adder (CSLA)

A carry select adder often consists of many Ripple Carry Adders. The n bits are linearly partitioned into groups of bits for n-bit binary addition. [1]. For example, to perform the 16-bit addition, 16-bits are divided into four sets of 4-bits. In this four 4-bit Ripple Carry Adders are used with the two stages of Ripple Carry Adders like one stage with carry, another stage without carry. In order to determine the correct sum and carry outputs, both stage results are computed simultaneously. The carry outputs are then determined using the preceding Ripple Carry Adder output [2]. Hence, the carry propagation delay is decreased [3]. But in terms of area and hardware, it has the disadvantage that two stages of Ripple Carry Adders are used which occupy more area but at the end only the output of the single stage is used to get the result [4].

2.3 Square Root CSLA (SQRT CSLA)

Square root CSLA is the improved version of CSLA, it has the similar operation as CSLA but instead of dividing the bits of a binary number linearly they are divided with different sequences [3]. For example, to perform the 16-bit addition, 16-bits are divided into 2, 2, 3, 4, and 5 bits. Then the 2-bit, 2-bit, 3-bit, 4-bit, and 5-bit Ripple Carry Adders in two stages are used the same as Linear CSLA, one stage with carry and another stage without carry. The variable bit size of Ripple Carry Adders in SQRT CSLA. It is designed to equalize the delay of the two stages [2].

2.4 Carry Save Adder (CSA)

To add three or more binary digits, a Carry Save Adder (CSA), a type of digital adder, is utilized. It differs from other adders in the way it generates both the sum and carry of the result concurrently and without any delay, and at the end, after acquiring the summation result, the carry bits are added to it [5]. In RCA, there are multiple full adders. Each full adder has to wait for the carry-out of the previous full adder to

continue the process [2]. But in CSA that is not the case. In this, all the sum outputs are obtained at a time without any delay and then for that all the carry bits are added. Carry is therefore recorded in the current stage and updated as an added value in the following stage. Due to its improved performance and simplicity as a three-operand adder, it is typically employed in multipliers.

2.5 *Han-Carlson Adder (HCA)*

The Han-Carlson Adder (HCA) is modified by Han and Carlson (1987). HCA is a family of networks of Brent-Kung and Kogge-Stone adders [6]. It has a number of logic levels and those levels are depending on the number of bits of binary number we are going to add. It is defined by an equation $[\log_2(n) + 1]$ where n is the number of bits. The logic levels consist of black and gray cells which perform the logic on generate and propagate [7]. Generate is nothing but it is AND operation of two bits of a binary number and propagate is nothing but XOR operation of two bits of a binary number [8]. The computation of the odd-numbered prefixes is done using one Brent-Kung stage at the beginning, followed by Kogge-Stone Stages, and then another Brent-Kung stage at the end [6]. In comparison with Kogge-Stone, it performs better with small adders. The Han-Carlson Adder is the better option when comparing the two parallel prefix adders [7] in terms of cost, area, and power. Due to its benefits, this adder is also utilized in the three operand adders [7].

2.6 *Hybrid Han-Carlson Adder (HHCA)*

The HCA has been modified to develop the Hybrid Han-Carlson Adder for further efficiency. It employs the Kogge-Stone Stages in the middle and the two Brent-Kung Stages at the adder's beginning and end [6]. The main reason to design the HHCA is to reduce the gate complexity. In HCA, the gate complexity is more and it has less delay [8]. But HHCA has a slightly more delay, when compared to HCA but the complexity is reduced [9]. Furthermore, we see the internal structure and the more details about the HCA and HHCA.

3 Three Operand Adders

In order to improve the performance of multipliers, three Operand Adders are employed as the basic unit to conduct the arithmetic operation in a variety of cryptographic applications. In multipliers, the addition of partial products is done at the final partial product stage more efficiently by using three operands rather than only two. To accomplish the three operand addition, many adders can be used, such as

the Carry Save Adder (CSA), Han Carl-Son Adder (HCA), and Hybrid Han-Carlson Adder (HHCA). One three operand adder or two operand adders can be used to perform the three operand addition. The three operand addition is carried out using these three adders. According to the survey, the Carry Save Adder is a popular and effective solution for three operand addition [5]. Parallel prefix adders reduce the path latency but increase the area, while the Han-Carlson Adder (HCA) is the fastest when compared to the Carry Save Adder (CSA) and both of these adders [9]. So based on the necessity of the consumer, a preferable adder is used. However, the HCA has the gate complexity in order to reduce the gate complexity by 10–18% Hybrid Han-Carlson Adder (HHCA) is used [9]. Figure 1 shows the three operand adder using Carry Save Adder. Figure 2 shows how the two operand adders with two stages are used to implement the three operand adder. Like HCA, HHCA is also used to implement the three operand adder. The operation of the HCA and HHCA is similar, the only difference is the PG logic stage arrangement. PG logic consists of black cells and gray cells [6, 7], it performs the operation on the generate and propagate obtained by doing AND and XOR operation on the input binary bits.

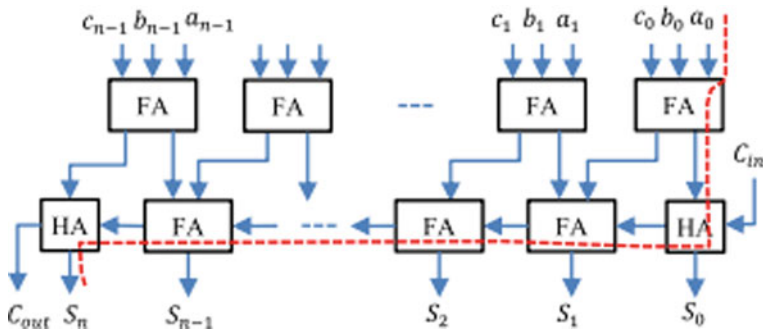
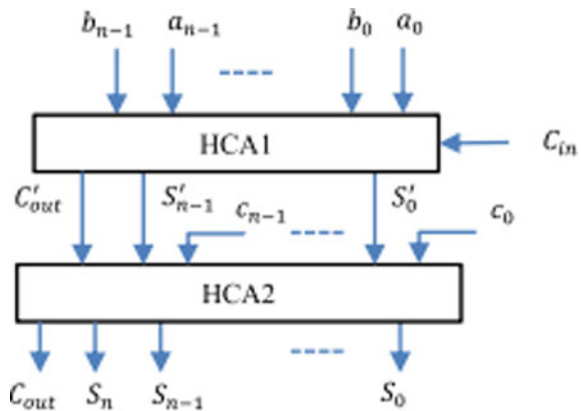


Fig. 1 Three operand adder using CSA

Fig. 2 Three operand adder using two stages of Han-Carlson Adder



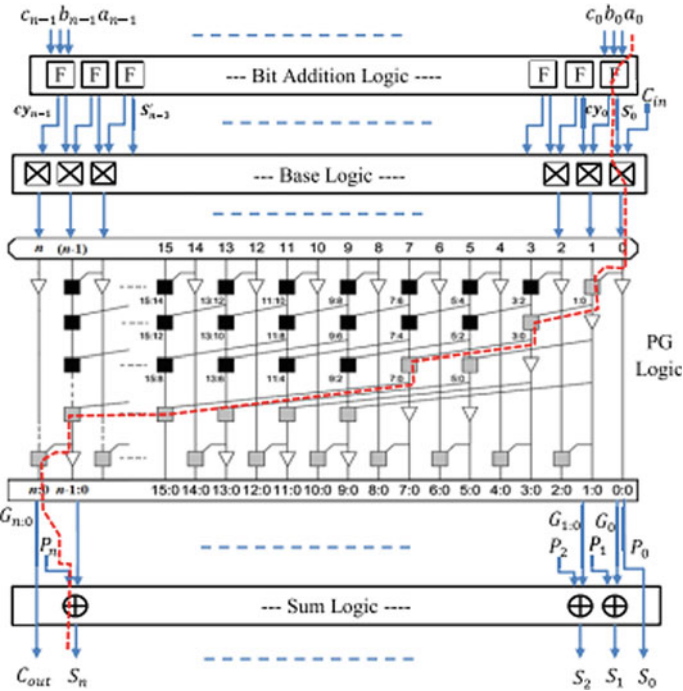


Fig. 3 VLSI architecture of high speed three operand adder

Figure 3 is the VLSI architecture of the three operand adder [9]. It has four stages.

The comprehensive adders are used in the first step (bit addition) to obtain the sum and carry bits from the three input bits a, b, and c. The outputs sum and carry bits along with carry-in (C_{in}) are given as input to the second stage that is base logic. Base logic stage is the stage where the generate and propagate are obtained. From Fig. 4, the generate and propagate are obtained by the AND and XOR operation of the first stage (bit addition logic) outputs. From the logical diagram in Fig. 4 for black cell and gray cell, the number of gates required for the black cell is more than gray cell. So, the gate complexity is more in HCA than the HHCA because of the greater number of black cells than gray cells.

The base logic stage generate and propagate are given to the third stage (PG logic). Then the PG logic stage performs an operation on generate and propagate with the various levels of black cell and gray cell and get the result of PG pairs like (P_0, G_0) , $(P_1, G_{1:0})$, $(P_2, G_{2:0})$, and so on. These are given as input to the final stage (sum logic) which is simply an XOR operation to get the final result of the three operand addition [9].

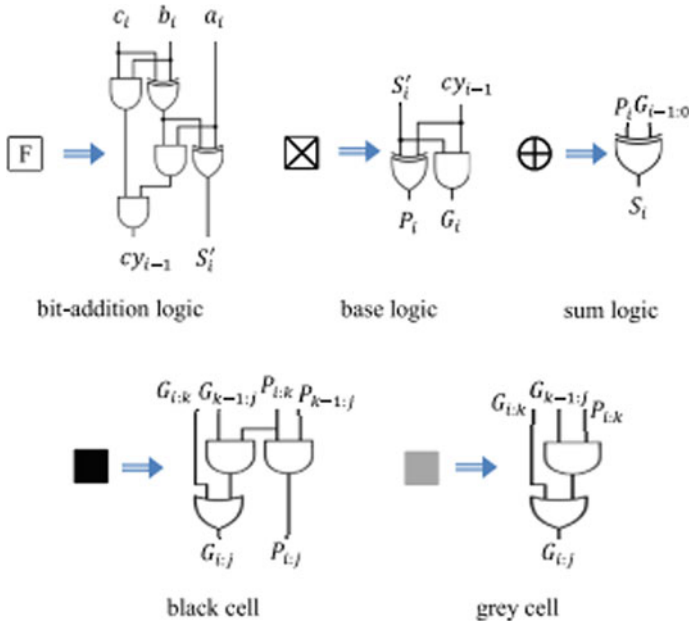


Fig. 4 Black cell, gray cell, base logic, addition logic, and bit addition logic logical diagram

4 Modified High Speed Efficient Three Operand Adder

The modified architecture of three operand adder is and shown in Fig. 5. In the three operand adder shown in Fig. 3 full adder used in the bit addition logic has the greater delay because the full adder logic contains the XOR gate in it which has the highest delay among all the gates and also in the sum logic stage XOR gates are used. The three operand adder’s delay is increased as a result of these bits’ addition logic and sum logic stages. The new architecture is designed by replacing the bit addition logic and sum logic with the modified bit addition logic and modified sum logic.

The modified full adder in Fig. 6 produces the sum and carry output using AND, OR gates rather than the common XOR gates, which further improves the performance because of employing AND, OR gates. Figure 7 is the modified XOR logic diagram, the XOR logic is implemented simply with the help of AND, OR, and NAND gates instead of using XOR gate directly. Then the delay of implementation XOR logic is greatly reduced.

5 Results and Discussion

The various VLSI architectures of the adders are designed by generating Verilog code using the VIVADO 2017.2 software tool. Simulation results for the modified

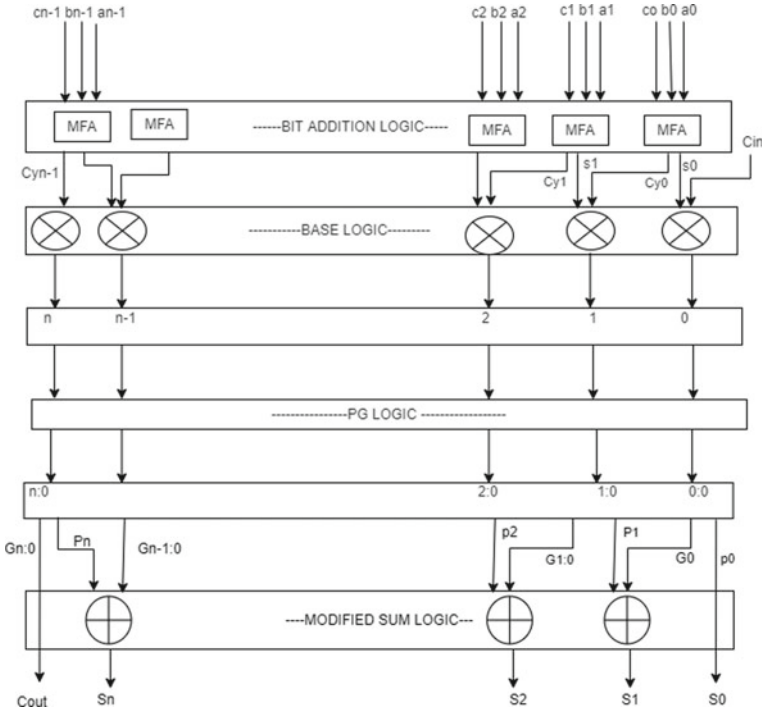


Fig. 5 Modified high speed efficient three operand adder

Fig. 6 Modified full adder (MFA)

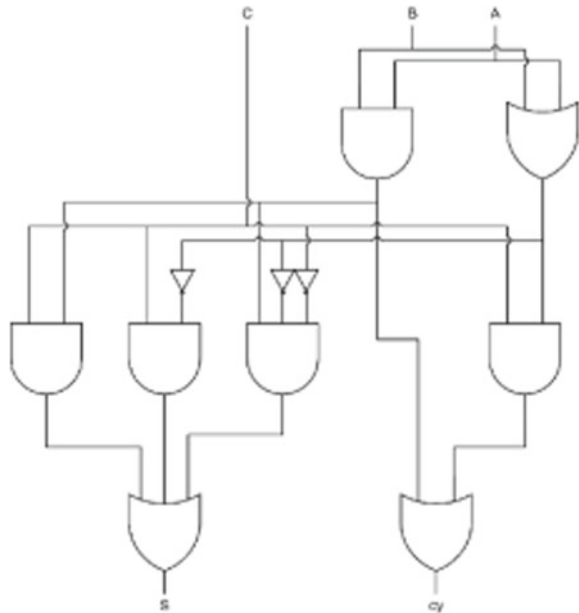
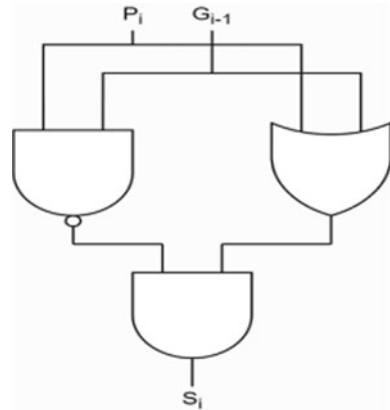


Fig. 7 Modified XOR logic diagram



adder are shown in Fig. 8. Verilog codes for the adders are written and the results obtained by using VIVADO 2017.2 such as power consumption and timing report after synthesizing the Verilog code are shown in Figs. 9 and 10.

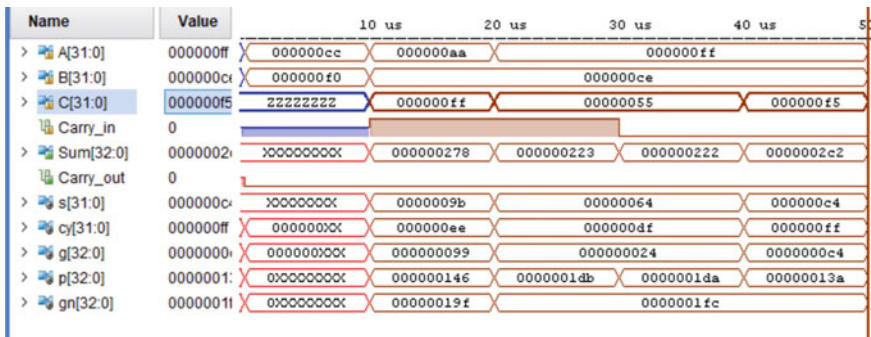
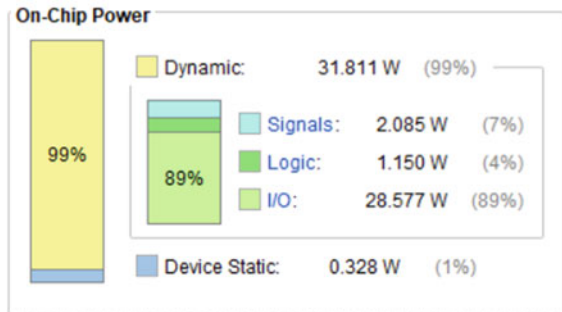


Fig. 8 Simulation results of modified adder

Fig. 9 Power consumption report of modified high speed efficient three operand adder



Timing Report

Slack: inf
 Source: C[1]
 (input port)
 Destination: Carry_out
 (output port)
 Path Group: (none)
 Path Type: Max at Slow Process Corner
 Data Path Delay: 7.127ns (logic 3.557ns (49.905%) route 3.570ns (50.095%))
 Logic Levels: 7 (IBUF=1 LUT5=1 LUT6=4 OBUF=1)

Location	Delay type	Incr(ns)	Path(ns)	Netlist Resource(s)
		0.000	0.000 r	C[1] (IN)
	net (fo=0)	0.000	0.000	C[1]
	IBUF (Prop_ibuf_I_O)	0.830	0.830 r	C_IBUF[1]_inst/0
	net (fo=5, unplaced)	0.584	1.413	C_IBUF[1]
	LUT6 (Prop_lut6_I5_O)	0.053	1.466 r	Sum_OBUF[2]_inst_i_2/0
	net (fo=5, unplaced)	0.692	2.158	p[2]
	LUT5 (Prop_lut5_I0_O)	0.053	2.211 r	Sum_OBUF[9]_inst_i_5/0
	net (fo=1, unplaced)	0.665	2.876	Sum_OBUF[9]_inst_i_5_n_0
	LUT6 (Prop_lut6_I0_O)	0.053	2.929 r	Sum_OBUF[9]_inst_i_3/0
	net (fo=4, unplaced)	0.688	3.617	gn[7]
	LUT6 (Prop_lut6_I1_O)	0.053	3.670 r	Sum_OBUF[32]_inst_i_5/0
	net (fo=3, unplaced)	0.358	4.028	Sum_OBUF[32]_inst_i_5_n_0
	LUT6 (Prop_lut6_I2_O)	0.053	4.081 r	Carry_out_OBUF_inst_i_1/0
	net (fo=1, unplaced)	0.584	4.665	Carry_out_OBUF
	OBUF (Prop_obuf_I_O)	2.462	7.127 r	Carry_out_OBUF_inst/0
	net (fo=0)	0.000	7.127	Carry_out
				r Carry_out (OUT)

Fig. 10 Timing report of modified high speed efficient three operand adder

The results of the modified three operand adder and the existing three operand adders are compared in Table 1. Compared to other adders, the modified adder has a shorter delay. Based on the simulation results and comparison with the three operand adders mentioned above, the performance of the Modified High Speed Efficient Three Operand Adder is evaluated and compared. The set up delay, hold delay; power consumption for the modified adder is reduced. Power delay product (PDP) is nothing but the product of power consumption and set up delay. Hence, the modified adder is the efficient one.

6 Conclusion

In this paper, the modified VLSI architecture of High Speed Efficient Three Operand Adder that is efficient in terms of power usage and setup time has been developed. The results are evaluated with the help of Xilinx VIVADO 2017.2 software tool. The results shows that the High speed efficient Three Operand Adder (Modified Adder) has the less delay and less power consumption capacity when compared to the other three operand adders like Hybrid Han-Carlson Adder and high speed three operand

Table 1 Comparison of three operand adders with the modified adder

Adder	Power (Watts)	Set up delay (ns)	Hold delay (ns)	Power-delay product (PDP)	Look up tables (LUT's)
Modified high speed efficient three operand	32.138	7.127	1.905	229.04	174
High speed three operand adder [9]	32.981	8.133	1.994	268.23	151
Hybrid Han-Carlson Three operand adder	33.237	9.571	1.994	318.11	165

adder. This work can be used as extension in multipliers, filters and in cryptographic applications etc. to improve their efficiency.

References

1. Udaya Kumar N, BalaSindhuri K, Teja KD, Satish DS (2017) Implementation and comparison of VLSI architectures of 16-bit carry select adder using Brent Kung adder. In: 2017 Innovations in power and advanced computing technologies (i-PACT), pp 1–7. <https://doi.org/10.1109/IPACT.2017.8244982>
2. Gaur N, Mehra A, Kumar P, Kallakuri S (2019) 16 Bit power efficient carry select adder. In: 2019 6th International conference on signal processing and integrated networks (SPIN), Noida, India, Mar 2019, pp 558–561
3. Bala Sindhuri K, Kalluru PV, Inty SP (2021) Design of area efficient VLSI architecture for carry select adder using logic optimization technique. *Comput Intell* 37: 1155–1165
4. Ramkumar B, Kittur HM (2012) Low-power and area-efficient carry select adder. *IEEE Trans Very Large-Scale Integr (VLSI) Syst* 20(2):371–375. <https://doi.org/10.1109/TVLSI.2010.2101621>
5. Kim T, Jao W, Tjiang S (1998) Arithmetic optimization using carry-save-adders. In: Proceedings 1998 design and automation conference. 35th DAC. (Cat. No.98CH36175), pp 433–438. <https://doi.org/10.1109/DAC.1998.724511>
6. Muthyala Sudhakar S, Chidambaram KP, Swartz lander EE (2012) Hybrid Han-Carlson adder. In: 2012 IEEE 55th International midwest symposium on circuits and systems (MWSCAS), pp 818–821. <https://doi.org/10.1109/MWSCAS.2012.6292146>
7. Lakshmanan AM, Othman M (2006) High-speed hybrid parallel-prefix carry-select adder using Ling’s algorithm. In: 2006 IEEE international conference on semiconductor electronics, pp 598–602. <https://doi.org/10.1109/SMELEC.2006.380702>
8. Neha V, Greeshma A (2019) Design and execution of enhanced carry increment adder using Han-Carlson and Kogge-Stone adder technique. In: Third international conference on electronics communication and aerospace technology (ICECA 2019), Mar 2019
9. Panda AK, Palisetty R, Ray KC (2020) High-speed area-efficient VLSI architecture of three-operand binary adder. *IEEE Trans Circuits Syst I Regul Pap* 67(11):3944–3953. <https://doi.org/10.1109/TCSI.2020.3016275>

Author Index

A

Akankhsa, T., 685
Akhendra Kumar, P., 27
Akshay Kumar, E., 485
Alam, Shahzad, 345
Anand, Rohit, 607, 619, 631
Anguera, Jaume, 309
Arora, Chirag, 301
Asha Shaini, K., 695
Athira, A. B., 431

B

Babulu, K., 11
Bala Sindhuri, K., 695
Bammidi, Deepa, 207
Banerjee, Rupa, 227
Barai, Ranjit Kumar, 543
Barot, Parth, 387
Bhagyashree, K., 407
Bhateja, Vikrant, 227, 255
Bhattacharya, Anagha, 543
Bodepu, Vinay Kumar, 207
Brahmanandam, P. S., 165

C

Chacko, Anu Mary, 431
Challa Ram, G., 215
Chaudhary, Alka, 607
Chauhan, Sudhir, 607
Chowdhury, Chandreyee, 497
Chudasama, Vipul, 387

D

Das, Santanu, 619
Deekshatulu, B. L., 145
Deepa, Bammidi, 195
Dejen, Arebu, 309
Desai, Nirav, 461
Devi, Chellaboyina Lalitha, 195
Dogra, Ayush, 255, 565, 579
Dube, R. R., 127

G

Geethalakshmi, B., 67
Goriparthi, M. Krishna, 67
Gorle, Gayatri, 207
Gowtham Chowdary, M., 183
Goyal, Bhawna, 565, 579
Gubbala, Srilakshmi, 81
Gude, Reshma, 207
Gupta, Ankur, 607, 619, 631
Gupta, Priyanka, 333
Gupta, Sparshi, 255
Gupta, Sumita, 103
Gupta, Vikas, 447

H

Harsha Varma, S. S., 695
Heera, W., 407
Hiremath, Shraddha, 407

I

Iyer, Nalini, 407

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

V. V. S. S. Chakravarthy et al. (eds.), *Advances in Signal Processing, Embedded Systems and IoT*, Lecture Notes in Electrical Engineering 992,
<https://doi.org/10.1007/978-981-19-8865-3>

J

Jalal, Anand Singh, 287
 Jayanth, M. Venkata, 685
 Jayasinghe, Jeevani, 309
 Jhamb, Mansi, 657
 Jha, Rakesh Kumar, 1
 Jhariya, D. C., 263
 Jidesh, P., 509
 Jothi Chitra, R., 365
 Jyothish Lal, G., 553

K

Kamaraju, M., 11
 Kamireddy, Durgarao, 115
 Kannadasan, B., 619, 631
 Karanam, Sreekanth, 115
 Kaur, Jaspreet, 579
 Kaur, Satwinder, 565
 Kaushik, Vandana Dixit, 439
 Kavitha, D. N. S. B., 135
 Keerthana, Beeram, 215
 Keshav, S., 553
 Khamitkar, Ravikant, 127
 Khan, Ajoy Kumar, 397
 Khan, Akbar, 145
 Khandelwal, Sumit Kumar, 173
 Khanna, Priyanka, 277
 Khan, Sameera, 419
 Kothapudi, Venkata Kishore, 93
 Koushik, Garapati Jaya Surya, 531
 Krishna Rao, E. V., 183
 Kumar, Anil, 519
 Kumar, Manoj, 475
 Kumar, Ravi, 475
 Kumar, Sanjay, 657
 Kumar, Santosh, 475
 Kumar, Vishal, 475

L

Lalitha Kumari, P., 619
 Lingampalli, Govardhan, 207
 Lin, Jerry Chun-Wei, 227

M

Madhu Kumar, S. D., 431
 Maithani, Mansi, 103
 Mallik, Manjarini, 497
 Mamatha, P., 183
 Manjusha, P., 237
 Meena, Radhey Shyam, 631
 Meher, Dibyasha, 103

Mishra, Megha, 419
 Mishra, Reetika, 227
 Mishra, Vishnu Kumar, 419
 Mohapatra, Pradyut, 173
 Mothilal, Bommali, 195
 Munratiwar, Shubham, 685

N

Nageswara Rao, G., 607, 631
 Nagwanshi, Kapil Kumar, 519
 Nandy, Arup, 115
 Narendra Babu, A., 165
 Nhu, Nguyen Gia, 255
 Nunna, Bala Ankaiah, 93

P

Padmavathi, A., 673
 Paliwal, Manish, 519
 Pallavi, CH., 51
 Panda, Niharika, 321
 Pandey, Digvijay, 607
 Pandey, Kiran, 375
 Paramkusham, Venkata Keshava Krishna, 155
 Patil, Rohan, 407
 Penumutchi, Bujjibabu, 81
 Praneetha, A., 165
 Prasad Reddy, P. V. G. D., 135
 Pratap Reddy, L., 145
 Premjith, B., 553
 Pushpa, K., 165

R

Raghavkrishna, B., 237
 Rajesh, Thota, 195
 Raju, G. S. N., 247
 Rajvee, Mohammad Hayath, 145
 Rajyalakshmi, V., 195
 Ramesh Varma, D., 215
 Rangasamy, Jothi, 485
 Ridwan, Murad, 309
 Rudrapal, Dwijen, 531

S

Sabbiseti, Venkatesh, 207
 Sahu, Mridu, 263, 277
 Sai Sowmya, K., 695
 Sampathila, Niranjana, 619
 Saravanakumar, C., 619
 Satyanarayana, R. V. S., 37
 Shahu, Ronit, 447

Shakeel, Mohd Raagib, 345
 Sharma, Dilip Kumar, 287
 Sharma, Himani, 1
 Sharma, Riya, 591
 Sharma, Shivani, 643
 Sharma, Vivek, 375
 Shrivastava, Anushree, 263
 Shukla, Kumar A., 631
 Shukla, Parag, 461
 Siddhu, Ayushi, 643
 Siddiqui, Taufeeque Ahmad, 345
 Sikander, Bilal, 287
 Sindhuja, Karri, 195
 Sindhwani, Nidhi, 591, 607
 Singh, Juhi, 643
 Singhla, Ria, 333
 Singh, Sangeeta, 685
 Singh, Sourabh, 255
 Singh, Vineeta, 439
 Sirohi, Simran, 475
 Sivarathinabala, M., 365
 Smitha, A., 509
 Snehitha, B., 183
 Somraj, Keerthi, 27
 Sreenivasulu, G., 51
 Sri Hari, K., 695
 Srilatha, G., 247
 Srinivas, K., 165
 Sruthi, G., 237
 Subhashini, T., 11
 Subramanyam, Avvaru, 37
 Sumalata, G. L., 673
 Sunny Dayal, P. A., 247
 Supriya, M., 321
 Suryaprakashrao, B., 237
 Sushma, Damodara Satya, 195

T

Talukdar, Rongeeet, 631
 Tamrakar, Ashish Kumar, 419
 Taran, Sachin, 155
 Terlapu, Sudheer Kumar, 215
 Thakur, Hardeo Kumar, 333
 Thakur, Palash, 447
 Thakur, Sanjay, 543
 Tikle, Shubham, 509
 Tirudeepak, Kasukurthi, 531
 Tiwari, Shikha, 375
 Travieso-Gonzalez, Carlos M., 227

U

Udaya Kumar, N., 695
 Uma, G., 165

V

Valiveti, Sharada, 387
 Vamshi, Y., 685
 Vamsikrishna, M., 237
 Vandana, 519
 Vanlalhrauaia, 397
 Vashisth, Rashmi, 591
 Venkata Rao, K., 135
 Venkata Subbarao, M., 215
 Venkateswararao, M., 237
 Verma, Bhupendra, 375
 Verma, Sheela, 419
 Verma, Siddharth, 255
 Virmani, Chaitanya, 333
 Visweswara Rao, J., 183

Y

Yadav, Akshat, 643
 Ykuntam, Yamini Devi, 81