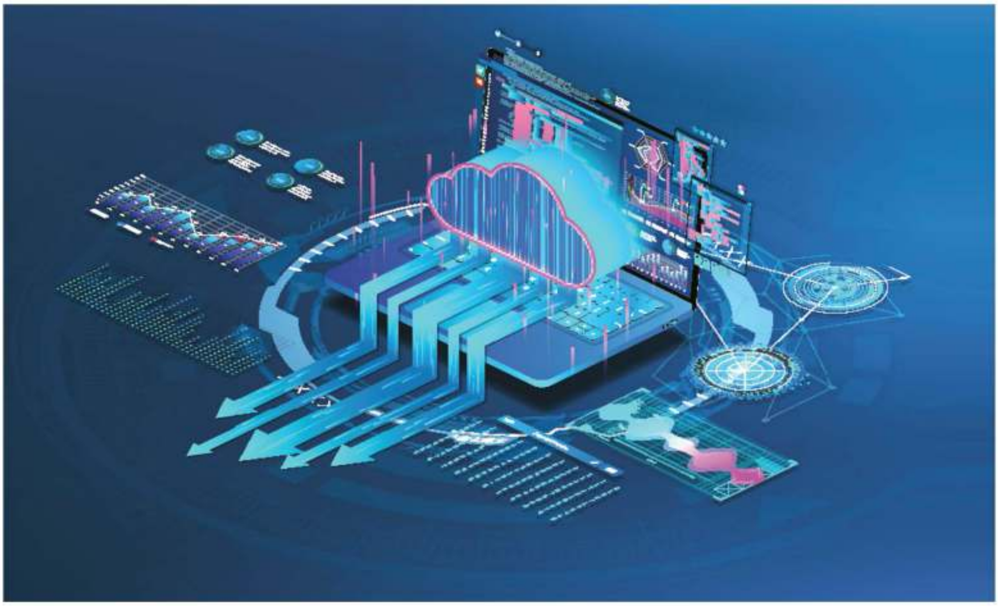
Research Notes on Computing & Communication Sciences

Applied Soft Computing Techniques

Theoretical Principles and Practical Applications



Samarjeet Borah | Ratna Raja Kumar Jambi

Sharifah Sakinah Syed Ahmad | Mahendra Prabhakar Deore

Editors

CRC Press
Taylor & Francis Group
APPLE ACADEMIC PRESS

Applied Soft Computing Techniques

Theoretical Principles and Practical Applications

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Research Notes on Computing and Communication Sciences Series

Applied Soft Computing Techniques

Theoretical Principles and Practical Applications

Edited by

Samarjeet Borah, PhD

Ratna Raja Kumar Jambi, PhD

Sharifah Sakinah Syed Ahmad, PhD

Mahendra Prabhakar Deore, PhD

AAP | APPLE ACADEMIC PRESS

First edition published 2026

Apple Academic Press Inc.
1265 Goldenrod Circle, NE,
Palm Bay, FL 32905 USA

760 Laurentian Drive, Unit 19,
Burlington, ON L7N 0A4, Canada

CRC Press
2385 NW Executive Center Drive,
Suite 320, Boca Raton FL 33431

4 Park Square, Milton Park,
Abingdon, Oxon, OX14 4RN UK

© 2026 by Apple Academic Press, Inc.

Apple Academic Press exclusively co-publishes with CRC Press, an imprint of Taylor & Francis Group, LLC

Reasonable efforts have been made to publish reliable data and information, but the authors, editors, and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The authors, editors, and publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged, please write and let us know so we may rectify in any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, access www.copyright.com or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. For works that are not available on CCC please contact mpkbookspermissions@tandf.co.uk

Trademark notice: Product or corporate names may be trademarks or registered trademarks and are used only for identification and explanation without intent to infringe.

For Product Safety Concerns and Information please contact our EU representative GPSR@taylorandfrancis.com Taylor & Francis Verlag GmbH, Kaufingerstraße 24, 80331 München, Germany

Library and Archives Canada Cataloguing in Publication

Title: Applied soft computing techniques : theoretical principles and practical applications / edited by Samarjeet Borah, PhD, Ratna Raja Kumar Jambi, PhD, Sharifah Sakinah Syed Ahmad, PhD, Mahendra Prabhakar Deore, PhD.
Names: Borah, Samarjeet, editor | Jambi, Ratna Raja Kumar, editor. | Syed Ahmad, Sharifah Sakinah, editor. | Deore, Mahendra Prabhakar, editor.
Description: First edition. | Includes bibliographical references and index.
Identifiers: Canadiana (print) 20240535332 | Canadiana (ebook) 20240535375 | ISBN 9781774918845 (hardcover) | ISBN 9781774918852 (softcover) | ISBN 9781003593133 (ebook)
Subjects: LCSH: Soft computing—Industrial applications.
Classification: LCC QA76.9.S63 A67 2025 | DDC 006.3—dc23

Library of Congress Cataloging-in-Publication Data

CIP data on file with US Library of Congress

ISBN: 978-1-77491-884-5 (hbk)
ISBN: 978-1-77491-885-2 (pbk)
ISBN: 978-1-00359-313-3 (ebk)

RESEARCH NOTES ON COMPUTING AND COMMUNICATION SCIENCES

About the Series

Editor-in-Chief:
Dr. Samarjeet Borah
Department of Computer Applications, Sikkim Manipal Institute of Technology,
Sikkim Manipal University (SMU), Majhitar, East Sikkim, India
Email: samarjeet.b@smit.smu.edu.in / samarjeetborah@gmail.com

Brief Description of the Series:

Computing can be defined as the practice in which computer technology is used to do a goal-oriented assignment. It covers design and development of hardware and software systems for various purposes. Computing devices are becoming an integral part of life now-a-days, including desktops, laptops, hand-held devices, smartphones, smart home appliances, etc. The evolution of the Internet of Things (IoT) has further enriched the same. The domain is ever growing and opening up many new endeavors, including cloud computing, social computing, ubiquitous computing, parallel computing, grid computing, etc.

In parallel with computing, another field has emerged that deals with the interconnection of devices. It is communication, and without which, the modern world cannot be thought of. It works with a basic purpose of transferring information from one place or person to another. This technology has a great influence in modern-day society. It influences business and society by making the interchange of ideas and facts more efficient. Communication technologies include the Internet, multimedia, e-mail, telephone, and other sound-based and video-based communication means.

This book series consists of both edited volumes as well as selected papers from various conferences. Volumes of the series contain the latest research findings in the field of communication engineering, computer science and engineering, and informatics. Therefore, the books cater to the needs of researchers and readers of a broader spectrum.

Coverage & Approach:

The series

- Covers a broad spectrum of research domains
- Presents on market-demanded product-based research works
- Discusses the latest developments in the field

The book series broadly considers contributions from the following fields:

- Artificial Intelligence and Expert Systems
- Big Data Analytics
- Broadband Convergence System and Integration Technologies
- Cellular and Mobile Communication
- Cloud Computing Technologies
- Computational Biology and Bioinformatics
- Computer and Information Security
- Computer Architecture
- Computer Graphics and Video Processing
- Control Systems
- Database Management Systems
- Data Mining
- Design Automation
- Digital Signal Processing
- GSM Communication
- High Performance Computing
- Human-Computer Interaction
- IoT and Blockchains
- Machine Learning
- Natural Language Processing
- Next Generation Communication Technologies
- Operating Systems & Networking
- Pervasive Computing and Cyber-Physical Systems
- Robotics and Automation
- Signal Processing
- Smart Internet of Everything
- SOC and System Platform Design Technologies
- Social Network Analysis
- Soft Computing

Types of Volumes:

This series presents recent developments in the domains of computing and communications. It will include mostly the current works and research findings, going on in various research labs, universities and institutions and may lead to development of market demanded products. It reports substantive results on a wide range of computational approaches applied to a wide range of problems. The series provides volumes having works with empirical studies, theoretical analysis or comparison to psychological phenomena. The series includes the following types of volumes:

- Conference Proceedings
- Authored Volumes
- Edited Volumes

Volumes from the series must be suitable as reference books for researchers, academicians, students, and industry professionals.

To propose suggestions for this book series, please contact the book series editor-in-chief. Book manuscripts should be minimum 250–500 pages per volume (11 point Times Roman in MS-Word with 1.5 line spacing).

Books and chapters in the series are included in Google Scholar and selectively in Scopus and possibly other related abstracting/indexing services.

Books in the *Research Notes on Computing and Communication Sciences* series

Applied Soft Computing: Techniques and Applications
Editors: Samarjeet Borah and Ranjit Panigrahi

Intelligent System Algorithms and Applications in Science and Technology
Editors: Sunil Pathak, Pramod Kumar Bhatt, Sanjay Kumar Singh, Ashutosh Tripathi, and Pankaj Kumar Pandey

Computing and Communications Engineering in Real-Time Application Development
Editors: B. K. Mishra, Samarjeet Borah, and Hemant Kasturiwale

Cyber-Physical Systems: Applications, Challenges, and Research Directions
Editors: Rajanikanth Aluvalu, Samarjeet Borah, and Lalit Garg

Applied Soft Computing Techniques: Theoretical Principles and Practical Applications
Editors: Samarjeet Borah, Ratna Raja Kumar Jambi, Sharifah Sakinah Syed Ahmad, and Mahendra Prabhakar Deore

Intelligent Computing and Analytics in Image Processing
Editors: Samarjeet Borah, Jambi Ratna Raja Kumar, Laxmi Bewoor, Rahul Joshi, and Mahendra Prabhakar Deore

ABOUT THE EDITORS

Samarjeet Borah, PhD

Professor and Head, Department of Computer Applications, SMIT, Sikkim Manipal University (SMU), Sikkim, India

Samarjeet Borah, PhD, is currently working as Professor and Head in the Department of Computer Applications, SMIT, Sikkim Manipal University (SMU), Sikkim, India. Dr. Borah has carried out various funded projects from AICTE (GoI), DST-CSRI (GoI), etc. He has published around 90 research papers in various journals and conferences and has organized workshops and conferences at the national and international level. Dr. Borah is involved with book volumes and journals of repute from Springer, IEEE, Inderscience, IGI Global, and Apple Academic Press as editor or guest editor. He is the Editor-in-Chief of the book series *Research Notes on Computing and Communication Sciences*, published Apple Academic Press, USA. Dr. Borah has served as a keynote as well as invited speaker at several international conferences. His areas of research are data mining, information security, data science, and machine learning.

Ratna Raja Kumar Jambi, PhD

Principal, Genba Sopanrao Moze College of Engineering, Balewadi, Pune, India

Ratna Raja Kumar Jambi, PhD, is currently working as Principal, Genba Sopanrao Moze College of Engineering, Balewadi, Pune, India. He has 17 years of teaching and research work. He has patents at both the national and international levels and has published papers on artificial intelligence and machine learning. He has also published various book chapters on applications of artificial intelligence and machine learning. He has supervised more than 40 students on their ME (CSE/IT) dissertations and more than 120 students on their BE (CSE/IT) dissertations. He has organized national and international conferences. He received an Innovative Leader Award from the World Education Summit & Awards in 2019 at New Delhi. He is continuously participating in social activities and engaging to uplift young students through "Unity Group of Happy NGO." He completed his PhD (CSE) degree at Maharishi University of Information Technology, Lucknow, Uttar Pradesh, India, and his Master of Technology (CSE) at Pondicherry Central University, India.

Sharifah Sakinah Syed Ahmad, PhD

Associate Professor, Department of Intelligent Computing & Analytics (ICA), Faculty of Information & Communication Technology, Universiti Teknikal Malaysia Melaka (UTeM), Malaysia

Sharifah Sakinah Syed Ahmad, PhD, is currently working as Associate Professor in the Department of Intelligent Computing & Analytics (ICA), Faculty of Information & Communication Technology, Universiti Teknikal Malaysia Melaka (UTeM). Her current research work is on the granular fuzzy rule-based system, evolutionary method, modeling, and data science. Dr. Ahmad has been a member of numerous program committees at national and international conferences in intelligent systems and soft computing, where she is involved in organizing and reviewing activities. She is an editorial board member of the IEEE Transactions on Systems, Man, and Cybernetics: Systems and a member of the Association for Computing Machinery (ACM). She has co-authored numerous journal publications, conference articles, and book chapters. She has also received several grants from various funding agencies, including the Ministry of Education, Malaysia, and the Ministry of Science, Technology, and Innovation, Malaysia. Dr. Ahmad has 15 years of experience in teaching. She was also involved in conducting several workshops related to improving data science and modeling.

Mahendra Prabhakar Deore, PhD

Assistant Professor, Computer Engineering Department at MKSSS's Cummins College of Engineering for Women, Pune, India.

Mahendra Prabhakar Deore, PhD, is currently serving as an Assistant Professor in the Computer Engineering Department at MKSSS's Cummins College of Engineering for Women, Pune, India. Having earned a doctorate in Computer Science and Engineering, Dr. Deore boasts an impressive academic history. With over 16 years of dedicated teaching experience at the undergraduate level, he possesses a profound reservoir of knowledge and expertise in this domain. Throughout the course of his academic journey, Dr. Deore has authored and published more than 20 SCI- and SCOPUS-indexed research papers. His primary areas of interest encompass big data, security, computer networks, and machine learning. He earned his Master of Technology degree at Bharati Vidyapeeth Deemed University College of Engineering, Dhankawadi, Pune, India.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

CONTENTS

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

CONTRIBUTORS

Bhaskar Adepu
Department of Information Technology, Kakatiya Institute of Technology & Science, Warangal, India

Nikunj Agarwal
Symbiosis Institute of Computer Studies and Research, Symbiosis International (Deemed University), Atur Centre, Gokhale Cross Road, Model Colony, Pune, Maharashtra, India

Mohammad Jahangir Alam
Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh

Hasmot Ali
Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh

Mallekedi Anand
School of Electronics and Electrical Engineering, Lovely Professional University, Punjab India

Gajendra Bamnote
Ram Meghe College of Engineering, Amravati, MS, India

Rohit Barshile
Computer Engineering Department Marathwada Mitra Mandal's Institute of Technology, Pune, India

Ashish Bhise
Department of Computer Engineering, MMIT, Lohegaon, Pune, India

Chaitanya S. Bhosale
Department of Computer Engineering, MMIT, Lohegaon, Pune, India

Abhijit Bomble
Computer Engineering Department Marathwada Mitra Mandal's Institute of Technology, Pune, India

Prabhu Chakkaravarthy
Department of Computer Science and Engineering, SIMATS School of Engineering, Saveetha University, Chennai, Tamil Nadu, India

Vikas V. Chavan
Department of Computer Engineering, Marathwada Mitra Mandal's Institute of Technology, Lohegaon, Pune, India

Subhashree Choudhury
Department of Electrical and Electronics Engineering, Siksha 'O' Anusandhan (Deemed to be University), Bhubaneswar, India

Karan Dabare
Department of Computer Engineering, Marathwada Mitra Mandal's Institute of Technology, Lohegaon, Pune, India

A. S. Das
J Systems and Control, Cuttack, India

Bishwa Ranjan Das
Department of Computer Application, MSCB University, Baripada, India

Sanju Das
Department of Computer Science, Assam University Silchar, Assam, India

Niladri Sekhar Dash
Linguistic Research Unit, Indian Statistical Institute, Kolkata, India

Taraprasanna Dash
Department of Electrical and Communication Engineering, Siksha ‘O’ Anusandhan (Deemed to be University), Bhubaneswar, India

Pranjali Deshmukh
Department of Computer Engineering, MMIT, Lohegaon, Pune, India

Yogesh Dongare
Faculty of Computer Engineering, Savitribai Phule Pune University, Pune, India

Aparna Gade
Department of CSE, Marathwada Mitra Mandal Institute of Technology (SPPU), Pune, India

Ranjeet Gaikwad
Department of Computer Engineering, Marathwada Mitra Mandal’s Institute of Technology, Pune, India

Naman Gupta
Symbiosis Institute of Computer Studies and Research, Symbiosis International (Deemed University), Atur Centre, Gokhale Cross Road, Model Colony, Pune, Maharashtra, India

Mrunal Jagtap
Computer Engineering Department Marathwada Mitra Mandal’s Institute of Technology, Pune, India

Anuj Jain
School of Electronics and Electrical Engineering, Lovely Professional University, Punjab India

Abhijeet Jadhav
Computer Engineering Department Marathwada Mitra Mandal’s Institute of Technology, Pune, India

Atharv Kaleker
Department of Computer Engineering, Marathwada Mitra Mandal’s Institute of Technology, Lohegaon, Pune, India

Janvi Kankhar
Department of CSE, Marathwada Mitra Mandal Institute of Technology (SPPU), Pune, India

Uma B. Karanje
Department of CSE, Marathwada Mitra Mandal Institute of Technology (SPPU), Pune, India

Gayatri Kinge
Department of CSE, Marathwada Mitra Mandal Institute of Technology (SPPU), Pune, India

R. N. Khobragade
Post Graduate Department of Computer Science & Engg., SGBAU, Amravati, Maharashtra, India (MMIT, PUNE)

Radhika Kokate
Department of Computer Engineering, Marathwada Mitra Mandal’s Institute of Technology, Pune, India

Suhas R. Kothavale
Department of Computer Engineering, Marathwada Mitra Mandal’s Institute of Technology, Lohegaon, Pune, India

G. Kulanthaivel
NITTTR, Chennai, India

Atharva Kulkarni
Department of Computer Engineering, Marathwada Mitra Mandal's Institute of Technology, Pune, India

Aakanksha Magaonkar
Department of CSE, Marathwada Mitra Mandal Institute of Technology (SPPU), Pune, India

Rohini D. Mahale
Department of Computer Engineering, Marathwada Mitra Mandal's Institute of Technology, Lohegaon, Pune, India

Kshyamasagar Mahanta
Department of Computer Application, Maharaja Sriram Chandra Bhanja Deo University, Odisha, India

Hima Bindu Maringanti
Department of Computer Application, Maharaja Sriram Chandra Bhanja Deo University, Odisha, India

Sarita Misra
Department of Electronics and Communication Engineering, GITA Autonomous College, Bhubaneswar, India

Jyoti Prakash Mishra
Gandhi Institute for Education and Technology, Baniatangi, Affiliate to Biju Patnaik University of Technology, Rourkela, Odisha, India

Sambit Kumar Mishra
Gandhi Institute for Education and Technology (Affiliated to Biju Patnaik University of Technology, Rourkela), Baniatangi, Odisha, India

A. Mohan
Department of Computer Science and Engineering, SIMATS School of Engineering, Saveetha University, Chennai, Tamil Nadu, India

Aditya More
Department of Computer Engineering, Marathwada Mitra Mandal's Institute of Technology Engg. College, Pune, India

S. K. Nayak
Department of Computer Science Engineering, GITA Autonomous College, Bhubaneswar, India

G. Palai
Department of Electronics and Communication Engineering, GITA Autonomous College, Bhubaneswar, India

Neha Panwar
Department of CS & IT, Manipal University Jaipur, Jaipur, India

Sushama L. Pawar
Vishwakarma Institute of Information Technology, Pune, India

G. Parthasarathy
Department of Computer Science and Engineering, Reva University Bangalore, Karnataka, India

Kshitij Patange
Department of Computer Engineering, Marathwada Mitra Mandal's Institute of Technology Engg. College, Pune, India

Varad Patil
Department of Computer Engineering, Marathwada Mitra Mandal's Institute of Technology, Pune, India

Shubham Phad
Department of Computer Engineering, Marathwada Mitra Mandal's Institute of Technology Engineering College, Pune, India

Zdzislaw Polkowski
The Karkonosze University of Applied Sciences in Jelenia Góra, Poland

Kiran Pothare
Computer Engineering Department Marathwada Mitra Mandal's Institute of Technology, Pune, India

Md. Mahfujur Rahman
Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh

Md. Rahmatul Kabir Rasel
Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh

Subhash G. Rathod
Department of Computer Engineering, Marathwada Mitra Mandal's Institute of Technology, Pune, India

Ammar Sakriwal
Department of Computer Engineering, Marathwada Mitra Mandal's Institute of Technology, Lohegaon, Pune, India

Mangesh D. Salunke
Department of Computer Engineering, Marathwada Mitra Mandal's Institute of Technology, Lohegaon, Pune, India

Soumya Ranjan Samal
Department of Electronics and Communication Engineering, Silicon Institute of Technology, Bhubaneswar, India

Komal Sharma
Department of Computer Engineering, Marathwada Mitra Mandal's Institute of Technology, Pune, India

V. Singh
Department of Computer Science Engineering, GITA Autonomous College, Bhubaneswar, India

P. Shivani
Department of Information Technology, Kakatiya Institute of Technology & Science, Warangal, India

Manoj Kumar Shukla
Department of Robotics and automation Symbiosis Institute of Technology, Pune, India

Juned A. Siddiqui
Medi-Caps University, Indore, India

Sarthak Surve
Department of Computer Engineering, Marathwada Mitra Mandal's Institute of Technology, Lohegaon, Pune, India

K. P. Swain
Department of Electronics and Communication Engineering, GITA Autonomous College, Bhubaneswar, India

Sarat Kumar Swain
Department of Electrical and Electronics Engineering, GITA Autonomous College, Bhubaneswar, India

Vilas M. Thakare
Post Graduate Department of Computer Science & Engineering, SGBAU, Amravati, India

Nilesh N. Thorat
Department of Computer Engineering, Marathwada Mitra Mandal's Institute of Technology, Lohegaon, Pune, India

V. Ulagamuthalvi
Sathyabama University, Chennai, India

Sakshee Vaidya
Medi-Caps University, Indore, India

P. Visu
Department of Computer Science and Engineering, Velammal Engineering College, Chennai, Tamil Nadu, India

K. H. Walse
Post Graduate Department of Computer Science & Engineering, SGBAU, Amravati, India

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

PREFACE

Soft computing techniques are always helpful in solving various problems. These techniques have the ability to handle complex, uncertain, and imprecise information. Keeping in view the same, we present *Applied Soft Computing Techniques: Theoretical Principles and Practical Applications*, which consists of 25 chapters. All the chapters are arranged in five sections, namely: biomedical applications, communication technologies, data analytics and applications, image processing, and natural language processing.

There are eight works in total from the biomedical applications domain, covering some of the critical healthcare issues such as cancer data analysis, depression and mental health analysis, heart disease detection, and so on. Communication systems frequently contend with intricate and ever-changing data, and soft computing methodologies can assist in tackling the obstacles associated with signal processing, network optimization, quality of service, and beyond.

In this book, four chapters have been included from the communications domain, which discuss some important issues on the same. Similarly, five chapters from the data analytics domain have been presented in this volume, discussing data compression, handling of large-scaled heterogeneous databases, visualization techniques, and so on. Soft computing techniques are widely used in image processing. The four chapters included in this book provide a discussion on human face recognition, casualty detection, traffic sign recognition, and soil features prediction using satellite imagery. The fifth section of the book includes some works on natural language processing.

Applied Soft Computing Techniques: Theoretical Principles and Practical Applications transcends beyond a mere book and assumes the role of a conduit that facilitates the amalgamation of theoretical principles and practical applications, the convergence of academic pursuits and industrial endeavors, and the symbiotic connection between an inquisitive intellect and the instrumental resources necessary to effectuate a substantive influence. Whether one is an erudite researcher in quest of inspiration, an eager student yearning for knowledge, or a seasoned professional embarking on a quest for solutions, this written work proffers an abundance of wisdom and serves as a launching pad for individual investigations. The editors hope readers will benefit from the works included in the book.

Finally, the editors would like to express their sincere gratitude to all the authors, reviewers, and the publisher Apple Academic Press, USA, for making this book possible.

INTRODUCTION

Applied soft computing pertains to the pragmatic application of diverse soft computing methods and approaches to resolve actuality problems and confront practical hurdles across a broad spectrum of domains and sectors. Soft computing techniques are particularly valuable in situations where traditional, rule-based approaches may be insufficient. It encompasses a collection of computational techniques that draw inspiration from the cognitive processes and decision-making abilities of human beings, frequently tackling information characterized by uncertainty, imprecision, and complexity.

In the realm of *Applied Soft Computing* professionals frequently engage in the following activities:

- Identify actual situations or undertakings that can derive advantages from the attributes of soft computing, such as its capacity to manage intricate data, conform to dynamic circumstances, and offer resolutions even in the presence of incomplete or imprecise information.
- Choose and tailor suitable soft computing methods or algorithms according to the inherent characteristics of the problem. For instance, when confronted with uncertain environments, fuzzy logic may prove to be a fitting choice for decision-making, whereas neural networks might display exceptional performance in tasks involving pattern recognition.
- Operationalize and amalgamate these methodologies into pragmatic systems or applications, it is frequently imperative to engage in programming, data preprocessing, model training, and the assimilation of other technologies.
- Evaluate the performance of the applied soft computing system and fine-tune it as needed to achieve the desired results.

Soft computing methodologies are highly advantageous in resolving practical issues owing to their capability to handle intricate, unpredictable, and indistinct circumstances that frequently typify real-world situations. A wide range of applications are there in the domains of biomedical and healthcare, communication and control, data analytics, image processing, natural language processing, and so on.

Soft computing methods have a crucial impact on the enhancement of healthcare, the acceleration of medical research, and the improvement of patient outcomes through the resolution of the inherent intricacy and unpredictability found in biomedical data and processes. These techniques facilitate more precise identification, treatment, and comprehension of biological and medical phenomena. Soft computing techniques have found numerous applications in the areas of medical image analysis, disease diagnosis and risk assessment, drug discovery and design, bioinformatics, patient monitoring and healthcare management, neuroscience and brain-computer interfaces, drug dosage optimization, biomedical signal processing, rehabilitation and assistive technologies, genomic data analysis, etc.

These techniques can be further employed in a multitude of manners to augment and ameliorate communication technologies. Communication systems frequently contend with intricate and ever-changing data, and soft computing methodologies can assist in tackling the obstacles associated with signal processing, network optimization, quality of service, and beyond.

Data analytics refers to the systematic process of examining, purifying, reformatting, and interpreting data to uncover meaningful insights, recurring patterns, prevailing trends, and valuable knowledge that can be utilized to facilitate decision-making. This endeavor necessitates the employment of a diverse array of techniques and tools to extract invaluable information from voluminous and intricate datasets. The realm of data analytics encompasses a broad spectrum of activities, encompassing data exploration, the application of descriptive statistics, data visualization, predictive modeling, and more. These undertakings are applied to data derived from various sources, such as commercial transactions, sensory input, social media, and scientific experiments, to acquire a more profound comprehension of the fundamental phenomena or to inform business strategies.

The application of soft computing techniques in the domain of image processing is constantly progressing, thereby offering novel resolutions to a diverse array of image-centric obstacles in a multitude of sectors, encompassing healthcare, entertainment, security, and scientific research. These methodologies facilitate a higher degree of precision and efficacy in the manipulation and examination of visual data, ultimately enriching our capacity to extract valuable insights from images.

Natural language processing (NLP) constitutes a branch of artificial intelligence (AI) which directs its attention toward the interaction taking place between computers and humans via natural language. It encompasses the creation of algorithms and models that empower computers to comprehend, decipher, and produce human language in a manner that is both purposeful and advantageous. NLP assumes a crucial function in serving as a connection between human communication and computer comprehension. Soft computing is highly advantageous in the domain of NLP, owing to the innate intricacies, uncertainties, and vagueness that are prevalent in human language. NLP, which pertains to the realm of AI, concentrates on the interplay between computers and natural human language, encompassing various tasks such as textual examination, vocal recognition, mechanical translation, sentiment analysis, and additional endeavors.

Finally, it can be said that the soft computing techniques offer a comprehensive and flexible set of tools to tackle the intricacy and uncertainty that are inherent in real-world challenges. These techniques enable systems to acquire knowledge from data, engage in logical reasoning like human cognition, and arrive at well-informed judgments. Consequently, this leads to heightened effectiveness, superior decision-making, and augmented problem-solving capabilities in diverse fields.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

PART I
BIOMEDICAL APPLICATIONS

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

CHAPTER 1

SOFT COMPUTING ALGORITHMS AND THEIR APPLICATION ON BREAST CANCER DATA CLASSIFICATION: AN EXPERIMENTAL ANALYSIS

JUNED A. SIDDIQUI and SAKSHEE VAIDYA

Medi-Caps University, Indore, Madhya Pradesh, India

ABSTRACT

Breast cancer is a disease which has high morbidity and mortality. Breast cells are the starting point for the development of breast cancer. Cancer starts when old, damaged cells start to multiply uncontrollably. Breast cancer has the highest incidence rate of any cancer in women and the second highest incidence rate of any cancer in the world. Analysis of the dataset of cancer patients can give us a more personalized and early approach to the treatment. Data mining techniques using Machine learning and artificial intelligence algorithms help us to classify, analyze, and also visualize this kind of data. This research is done to predict if it is possible to employ data mining methods for the classification and analysis of dataset for categories such as breast cancer cell data. The accuracy and the error of the applied algorithms, that is, artificial neural network (ANN) and k-nearest-neighbor algorithm (KNN) are also found. RStudio is the environment on which these algorithms are run using the R programing language. Two algorithms ANN and the KNN are applied giving an accuracy of 94.59%, an error of 6.4%, and 52% for KNN.

Applied Soft Computing Techniques: Theoretical Principles and Practical Applications. Samarjeet Borah, Ratna Raja Kumar Jambi, Sharifah Sakinah Syed Ahmed, & Mahendra Prabhakar Deore (Eds.)
© 2025 Apple Academic Press, Inc. Co-published with CRC Press (Taylor & Francis)

1.1 INTRODUCTION

Globally, cancer of the breast is a lethal illness and one of the main roots that cause female death. Normally, human cells grow and multiply, cell division forms new cells because the body requires them. New cells take their place when the existing cells become damaged or grow old or they die. The abnormal/damaged cells start to grow and multiply, when this process breaks down. Early detection might increase the probability of better treatment and viability. Usually, X-rays are able to detect the small lumps that are formed or even physical examination might help in the detection of the presence of lumps in the breast. It occurs majorly in women, but men are also at risk of getting breast cancer. The majority of breast lumps are mild (noncancerous) instead of malignant (cancerous). Normal tumors do not grow and spread through or grow into nearby tissues. When removed, most dangerous tumors do not grow back, but some might grow back. By using statistical methods, we can classify if the cancer is benign or malignant. The majority of the patients are unable to identify if their symptoms are cancerous or noncancerous. With the help of data mining techniques, we analyze the dataset and try to understand if the artificial intelligence algorithms are helpful in the classification of the dataset to provide ways, for better and early detection and treatments possible. Some of the previous works on this kind of dataset include "Implementation of data mining and methods for classification regarding breast cancer detection and prognosis".[2] Over the last decade, there has been continuous progress in cancer research.[3] Researchers used various methods, such as screening, to find out how far along the cancer is before signs show up. Furthermore, specialists have come up with novel techniques for predicting cancer treatment outcomes early on. New technologies make precise cancer prediction the most difficult and intriguing challenge for physicians. Within the discipline of biomedical, the rate at which fresh data is accumulated is staggeringly high.[4] This information is a valuable resource for health researchers. From past experiences, the process of gaining information and expertise from data is facilitated by machine learning. It helps find a hard-to-see pattern in a vast, noisy dataset. The WHO says the identification of cancer in the breast is the greatest significant challenge facing the field of current medical research.[4] Every year, the amount of reported cases grows. Survival rates and rates of relapse (cancer coming back after therapy) are two aspects of breast cancer. These are critical breast phenomena. Cancer behavior is inextricably linked to the patient's death. After lung cancer, cancer of the breast is the next underlying reason of mortality among women, cancer of the lungs is the top cause of death among men.[22] In comparison to the United States, the amount of newly diagnosed breast cancer patients in India [21] is smaller, but breast cancer fatalities are much higher. As a result, predicting breast cancer early is crucial. It is possible to forecast the onset of disease by analyzing indicators in the data. To get the desired accuracy of breast cancer predictions, this study compares several ML methods. Due to the success of its technique in predicting and classifying, machine learning is seeing increasing use in the medical arena, notably in medical diagnostics for identifying breast cancer, and is now often used in biomedical research.

The goal is to combine the information presented in several reviews and technical papers dealing with the detection and prognosis of breast cancer. It makes available a snapshot of the state of the art in breast cancer research by summing together the findings from a number of different datasets using data mining tools, breast cancer diagnoses and prognoses can be made better. However, with a large amount of data being generated every second, it is difficult to find information, or the extraction of information has become difficult. With the help of advanced tools such as R programing language and RStudio, it has become possible. For operational demonstration, the author[6] has done a case investigation looking at big data and a forecast model about fertility. "Breast cancer detection using ML," 2018[9] focuses on the fact that during a woman's lifetime, about 8% of females each year are affected by breast cancer, and the algorithms for machine learning can be applied to diagnosing and classifying breast cancer. As algorithms are subtypes due to their high degree of accuracy in the categorization process and robust diagnostic capabilities breast cancer. Hence, the authors suggest contrasting the two brand-new approaches, and a cross-validation evaluation of their accuracy. The accuracy obtained is 97.51 and 96.19% for the k-nearest-neighbor algorithm (KNN) and Naïve Bayes machine learning algorithms, respectively.[8] In Support vector machine (SVM)-based hybrid approach is considered the best approach. Mitigation of the issues through the structure of sentiment, as well as the principles for sentiment calculation, are mentioned. Methods such as Naïve Bayes (NB), the use of SVMs, and the maximum entropy (ME) method are used to fulfill the task of sentiment categorization in conjunction with other classifiers are posted. The research on early cancer detection[10] demonstrates five machine learning methods: SVM, KNN, K-SVM, the use of random forest tree, and the use of decision tree. with an accuracy of 97.20, 95.10, 96.50, 98.60, and 95.80% for breast cancer detection, respectively on the data archive for machine learning repositories to identify cancer in the breast. In Ref. [18], tools such as convolutional and recurrent neural networks are mentioned. The author also emphasized the importance of analyzing sentiments. Aspect-based sentiment analysis is the notified topic. The author[1] defines sentiment analysis and the need for it using a real-life pandemic, COVID-19. The proposed techniques include machine learning algorithms such as Nave Bayes. The author[13] is considering stock market data majorly in this. Data analysis and prediction are performed on this type of data using machine learning algorithms. It's possible that conventional data analytics[5] can't deal with the massive amounts of data involved here. The author sites some of the analyses that are implied, including sampling of data, data compression, density-based methods, grid-based methods, divide and conquer, progressive learning, distributed computing, and application of PCA. With all these methods in consideration, the author cites the analysis of large data to derive knowledge. With a large amount of data being generated every second,[6] it is difficult to find information or the extraction of information has become difficult. With the help of advanced tools, that is, R programing language and RStudio, it has become possible. For operational demonstration, the author has conducted a study by analyzing big data and a predictive model related to fertility.[17] In this paper, the authors discussed breast cancer's high morbidity and mortality. The authors also pointed out the absence of robust prognostic models, which makes it challenging for medical professionals to create a plan of action that has a chance of increasing patient survival. Consequently, the focus of this research is on constructing algorithms with the smallest possible error to improve precision. The research compares using a variety of datasets, four different algorithms (SVM, Logistic Regression, the Random Forest algorithm, and KNN) which are able to predict cancer of the breast. The purpose of this research is to make predictions on the effectiveness of the algorithms. The obtained SVM's accuracy (97.13%) is better than that of C4.5, Naïve Bayes, and k-NN are all models with an accuracy that ranges from 95.12% to 95.28%.[16] The objective is to increase the percentage of breast tumors that are diagnosed at an earlier stage, which will allow for better treatment and a reduction in breast cancer-related mortality. A Chaotic-Jaya online sequential extreme learning machine (OSELM) model for cancer classification has been discussed in.[23]

Due to its high morbidity and mortality rates, the risk of women developing breast cancer is significant. It is well known that due to the lack of reliable prognostic designs, it is challenging for medical professionals to create a plan of action that can add years to the life of a patient., hence it is urgent to develop a classification system with low error and high precision.

In this study, the breast cancer prognosis predicting algorithms artificial neural network (ANN) and KNN are compared on a self-created dataset. This manuscript is divided into sections comprising of theory, simulation, result and discussion, and also the conclusion.

1.2 BACKGROUND

1.2.1 MACHINE LEARNING AND DATA MINING

One of the branches of artificial intelligence is the field of machine learning, this enables software strategies to get better at accurately anticipating outcomes without having to be explicitly coded to do so. It is composed of three primary subtypes: learning under supervision, learning without supervision, and learning through reinforcement. Each subtype has its subtype. When there is already a framework of inputs and outputs being established, supervised learning is the method that is utilized. The naming of places and things is done accurately here. Unsupervised Learning is utilized by the system in situations in which the pattern is unknown to it and there is no appropriate structure present. The naming of the names is not done correctly. Reinforcement learning is utilized whenever there is communication between the system and an ever-changing environment. Machine learning algorithms are the suite of tools for creating and testing prediction, pattern recognition, and classification algorithms.

Data mining extracts knowledge from data. It helps us gather information which is relevant, useful, beneficial and economical for any organization. Machine learning and artificial intelligence algorithms are used to mine the information from the sea of data. We make use of a variety of approaches, each of which is optimized for a certain goal. Each approach has a set of benefits as well as some potential drawbacks. Some of the machine learning algorithms are discussed below-

A typical use of the data mining technique known as the “decision tree” is the early detection of breast cancer in its more treatable stages. A decision tree helps to analyze the data fully by forming multiple trees and their subtrees. It is mainly used for both classification and regression of the data. It is also possible to understand it as a particular kind of rule set, which is distinguished by the hierarchical structure of the rules that it contains.

SVM is a supervised machine learning algorithm. It focuses on finding a hyperplane and classifying from the assembled data. Classification of both linear and nonlinear data is performed using this statistical learning theory.

It seeks to avoid over-fitting of the data by maximizing the margin of hyperplane separation.

Naïve Bayes is one of the most effective probability-finding classification algorithms based on the Naïve Bayes theorem. Class membership for the specified tuple to the specified class is detected based on the maximum probability. Both a test set and a training set are included in the dataset. It is parametric.

1.2.2 EXPERIMENTAL PROTOCOL

The two algorithms that are applied to the self-created dataset are KNN and ANN. Both of these algorithms are applied and their accuracies and error are noted. Visualization of the created dataset is done with the help of plotting commands such as pie chart, histogram, density plot, and corrplot function of R language.

i. ANN: An ANN is a computer network created to simulate the way neurons in the brain work. It tries to copy the human neuron functioning and provide us with accurate results. It is comprised of interconnected information processing units. It is good at storing information on the entire network, has parallel processing ability, has fault tolerance, the ability to train the machine, and also the ability to work with inadequate knowledge. Adjusting the hidden layer, momentum, and learning rate maximizes value. The linear threshold gate separates inputs into two categories of benign (2) or malignant (4).

   Therefore, the output Y is in binary format. Equations have been used to mathematically characterize this function:

$$\text{Sum} = \sum_{i=1}^{N} I_i w_i \tag{1.1}$$

$$Y = f(\text{Sum}) \tag{1.2}$$

   Here, w_i.... w_n are the weight values of the inputs provided which are normalized. The sum represents the weighted sum and also is a threshold constant. I_i... I_n are the inputs, respectively. "f" is a linear step function at the threshold.

   KNN: KNN is an example of an algorithm for supervised learning. It works well with numeric variables and classification problems. It is meant to be used with a collection of information in which the datasets are divided into many groups to forecast the categorization of fresh sample data points. It is called a lazy learning model with local approximation. It assumes the new data point to be similar to the neighbor and derives the output. It looks for *k* neighbors and comes up with the prediction. *k* is kept as an odd number for calculating a convincing majority in this scenario, in which there are only two potential factions, that is, class benign (2) or malignant (4). Euclidean distance is used to find the nearest neighbor.

$$ED = \sqrt{[(X2 - X1)\,2 + (Y2 - Y1)\,2]} \tag{1.3}$$

where ED is the Euclidean distance between the two points whose distance we are calculating.

(X1, Y1) are the coordinates for the first point.

(X2, Y2) are the coordinates for the second point.

1.3 SIMULATION

1.3.1 SIMULATION ENVIRONMENT

We utilized RStudio, an integrated development environment (IDE) designed specifically used for statistical computing and data analysis. It is a popular option among those who work in data science, statisticians, and researchers because RStudio provides a comprehensive collection of tools and features that increase the productivity and efficiency of programers. This RStudio provides an intuitive interface that simplifies the development process and facilitates data exploration and visualization. Consequently, we use it for classification and clustering in data analysis. It is primarily accessible in two formats: RStudio Desktop and RStudio Server. However, in this work, RStudio Desktop is used. The same has been depicted in Figure 1.1.

Dataset: The self-created dataset consists of 50 rows of observations and 11 columns consisting of headings—patient identification (ID) number, Marginal adhesion, cell shape, clump thickness, cell size, epithelial, bare nuclei, bland chromatin, normal nucleoli, mitoses, and class. The class is a binary value, that is, 2 or 4, where 2 stands for benign type of cancer and 4 stands for cancer type. In the dataset that was made, there are no lost numbers. The dataset is created manually only to check if the algorithms work for these kinds of data, as in real-life situations, the data received will be similar but not the same as the same data.

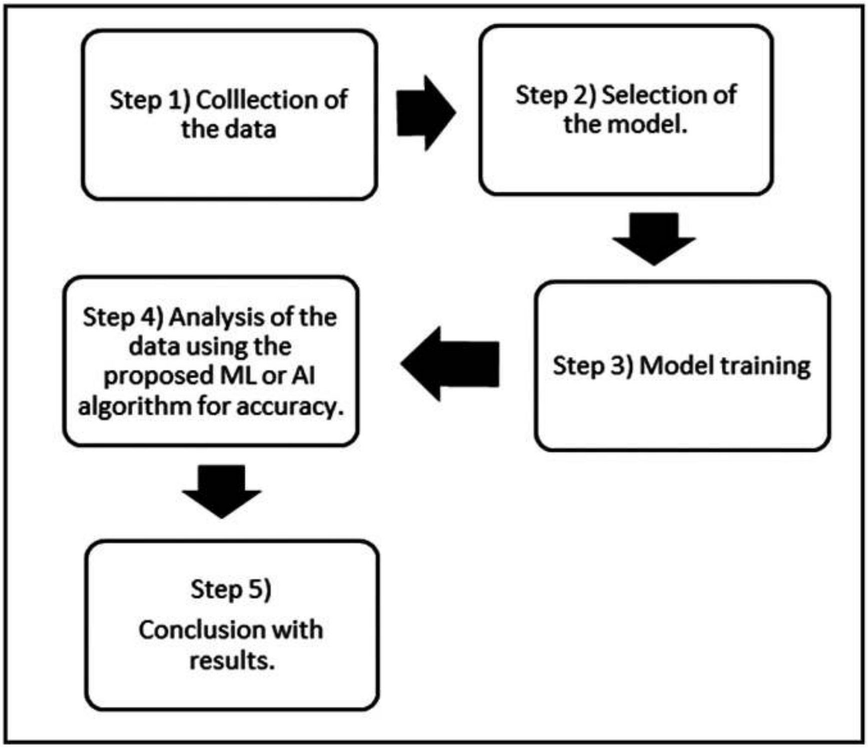


FIGURE 1.1 Flow diagram.

1.4 RESULT AND DISCUSSION

By applying algorithms such as ANN and KNN to our created dataset, we found that this data mining strategy may be helpful in the classification between benign and malignant types of breast cancer. Different visualization techniques were also used for the same. ANN gave us an accuracy of 94.59%, an error of 6.4% for classification and KNN gave us an accuracy of 52% for our created dataset.

Figure 1.2 shows that from the total of 50 observations taken, 46.3% of patients have a benign type (noncancerous) tumor and 53.7% of patients have a malignant type (cancerous) tumor.

The above density plot (Fig. 1.3) has the density of the number of the two different categories benign and malignant tumor vs cell size plotted on the same scale for the analyses of individual parameters in the two categories.

TABLE 1.1 Summary for Created Dataset.

Patient ID	Clump thick	Cell size	Cell shape	Mar adhesion	Epithelial	Bare nuclei	Bland chromatin	Normal nucleoli	Mitoses	Class
1000025	1	6	2	8	10	10	5	3	1	4
1016277	2	2	4	9	7	7	2	3	2	4
1017122	3	2	5	7	5	2	1	5	3	2
1033078	6	4	7	6	8	4	3	7	4	2
1048672	4	5	3	4	9	5	4	3	5	4
1056784	7	4	8	10	3	2	6	8	6	2
1079304	8	6	9	3	10	6	2	3	7	2
1103722	9	5	10	2	7	1	7	9	8	2
1110102	1	7	2	6	3	7	8	1	9	4
1116192	3	5	6	5	6	8	10	10	10	4

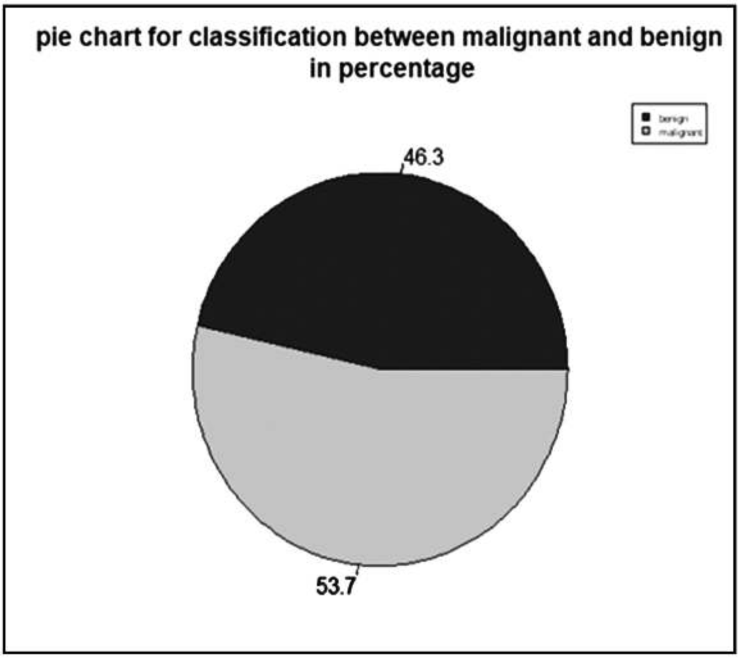


FIGURE 1.2 Pie chart for classification.

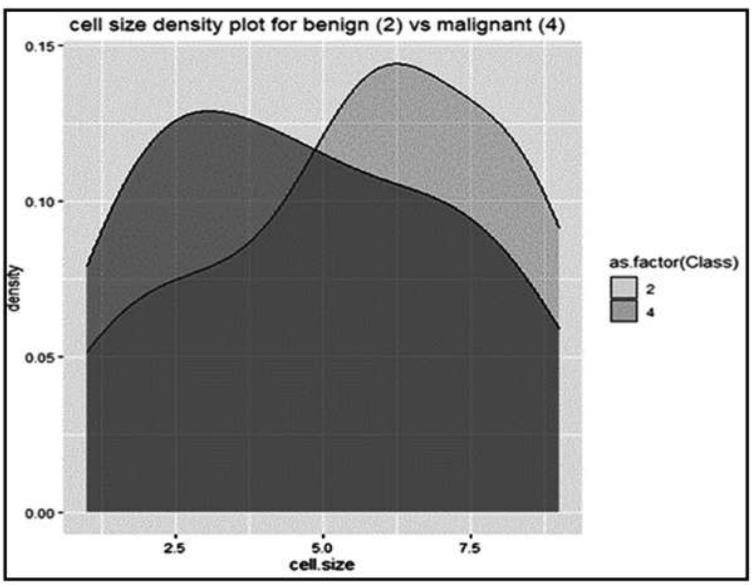


FIGURE 1.3 Density plot for density vs. cell size of benign and malignant tumor.

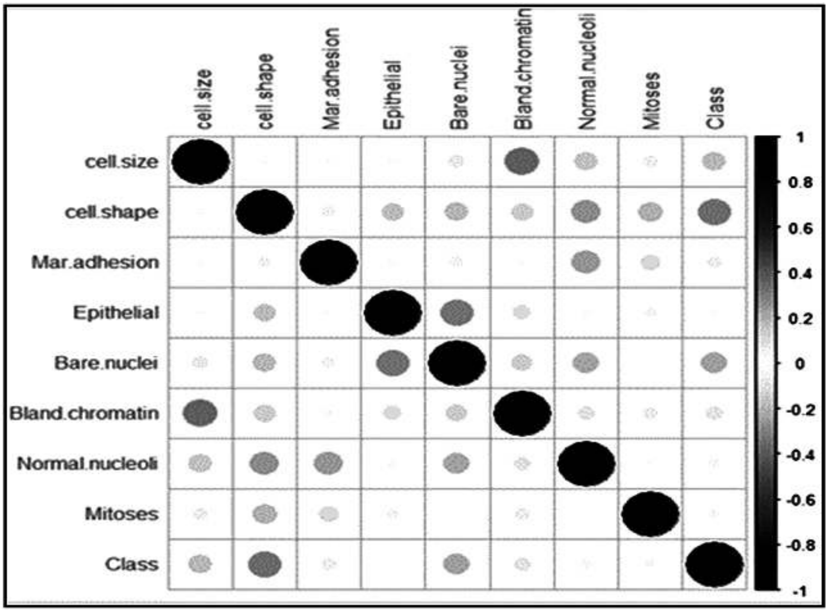


FIGURE 1.4 Circle correlation plot.

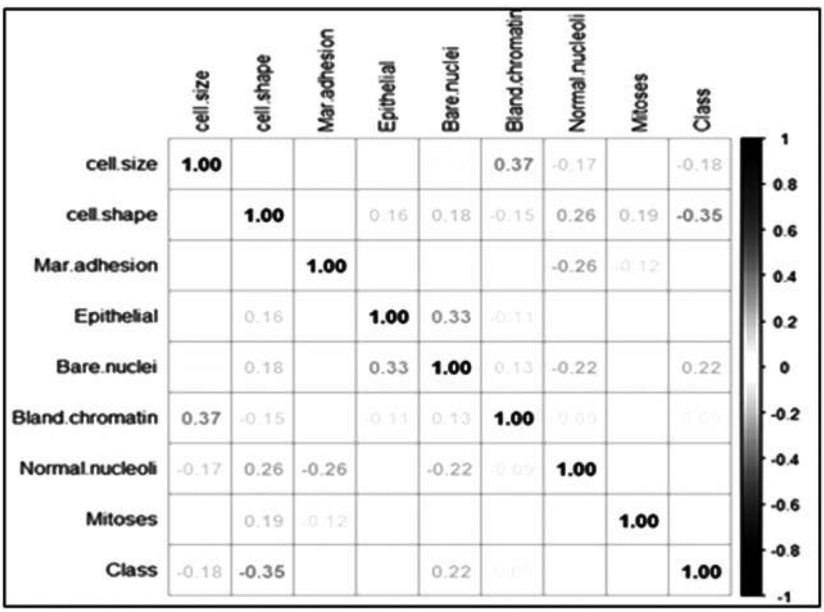


FIGURE 1.5 Number correlation plot.

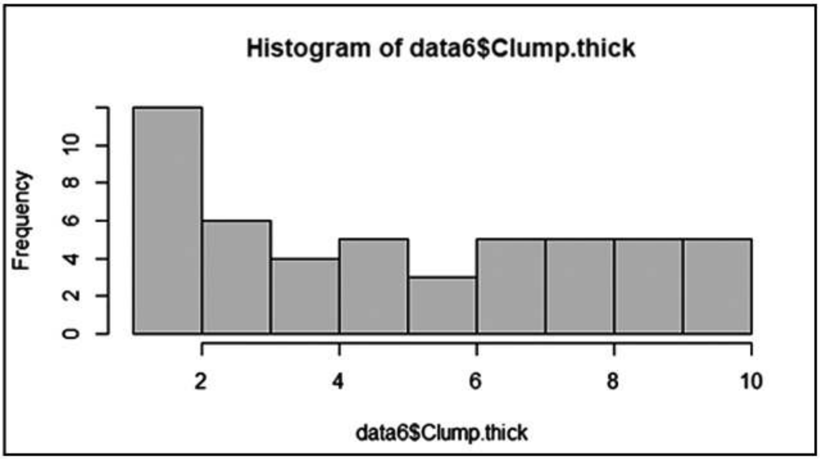


FIGURE 1.6 Histogram for clump thickness.

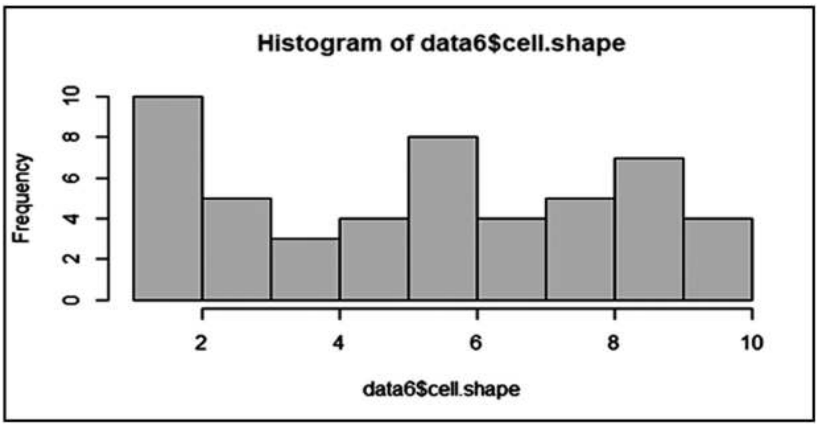


FIGURE 1.7 Histogram for cell shape.

Both the histograms are plotted for the individual parameter's categorization and analysis purpose. This will help in the analysis of individual categories as we are able to see the frequency of the number of patients in each parameter and this might also help in the pattern recognition process. Here data6 refers to the created dataset, which when imported to RStudio was renamed as data 6.

Figures 1.4 and 1.5 represent the circle and number correlation plots, respectively. The correlation plot is a table showing the correlation coefficients between the different variables/parameters. The diagonal of the table always contains 1 because the correlation between a variable and itself is always 1. The circle plot represents the correlation with the help of different size and color circles whereas the number plot presents the value in each box for correlation between parameters.

1.5 CONCLUSION

Accuracy varies with different datasets. However, because of variables such as earlier detection, customized therapy, and disease understanding breast cancer survival rates have increased and the amount of deaths attributable to the disease has steadily decreased. Our target is to build a precise and reliable classifier for the breast cancer cells dataset. By applying these algorithms, we found out that this type of analysis might prove to be beneficial, to classify the number of patients having different types of cancer in a given dataset. In this paper, with the created dataset, ANN has been a more efficacious algorithm to classify the breast cancer information database.

KEYWORDS

- **breast cancer**
- **data mining**
- **artificial intelligence algorithms**
- **R programing language**
- **RStudio**
- **ANN**
- **KNN**

REFERENCES

1. Medhat, W.; Hassan, A.; Korashy, H. N. D. Sentiment Analysis Algorithms and Applications: A Survey. *Ain Shams Eng. J.* **2014,** *5* (4), 1093–1113. DOI: 10.1016/j.asej.2014.04.011
2. Gupta, S.; Kumar, D.; Sharma, A. Data Mining Classification Techniques Applied for Breast Cancer Diagnosis and Prognosis. *Indian J. Comput. Sci. Eng. (IJCSE)* **2011,** *2* (2), 188–195.

3. Hanahan, D.; Weinberg, R. A. Hallmarks of Cancer: The Next Generation. *Cell* **2011,** *144* (5), 646–674. DOI: 10.1016/j.cell.2011.02.013
4. Marx, V. Biology: The Big Challenges of Big Data. *Nature* **2013,** *498* (7453), 255–260. DOI: 10.1038/498255a
5. Tsai; Chun-Wei, C.-F. L.; Chao, H.-C.; Vasilakos, A. V. Big Data Analytics: A Survey. *J. Big Data* **2015,** *2* (1), 1–32.
6. Shinde, P. P.; Oza, K. S.; Kamat, R. K. In *Big Data Predictive Analysis: Using R Analytical Tool*, International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Vol. 2017, 2017, 839–842. DOI: 10.1109/I-SMAC.2017.8058297
7. Gupta, M.; Gupta, B. In *A Comparative Study of Breast Cancer Diagnosis Using Supervised Machine Learning Techniques*, Second International Conference on Computing Methodologies and Communication (ICCMC), Vol. 2018, 2018, 997–1002. DOI: 10.1109/ICCMC.2018.8487537
8. Athindran, N.; Srivats, S. M.; Kamaleshwar, R. In *Comp. Anal*, Customer Sentiments Competing Brands Using Hybrid Model Approach 3rd International Conference on Inventive Computation Technologies (ICICT), Vol. 2018, 2018, 348–353.
9. Amrane, M.; Oukid, S.; Gagaoua, I.; Ensari, T. In *Breast Cancer Classification Using Machine Learning*, Electric Electronics, Computer Science, Biomedical Engineerings' Meeting (EBBT), Vol. 2018, 2018, 1–4. DOI: 10.1109/EBBT.2018.8391453
10. Awatramani, J.; Hasteer, N. In *Early Stage Detection of Malignant Cells: A Step Towards Better Life*, International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), Vol. 2019, 2019, 262–267. DOI: 10.1109/ICCCIS48478.2019.8974543
11. Saoud, H.; Ghadi, A.; Ghailani, M. In *Proposed Approach for Breast Cancer Diagnosis Using Machine Learning*, Proceedings of the 4th International Conference on Smart City Applications, 2019, pp 1–5. DOI: 10.1145/3368756.3369089
12. Awatramani, J.; Hasteer, N. In *Early Stage Detection of Malignant Cells: A Step Towards Better Life*, International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), Vol. 2019, 2019, 262–267. DOI: 10.1109/ICCCIS48478.2019.8974543
13. Pahwa, K.; Agarwal, N. In *Stock Market Analysis Using Supervised Machine Learning*. International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon), Vol. 2019, 2019, 197–200. DOI: 10.1109/COMITCon.2019.8862225
14. ACS Style Quick Guide. In The ACS Guide to Scholarly Communication. *ACS Guide to Scholarly Communication*; American Chemical Society, 2019, https://doi.org/doi: 10.1021/acsguide.40303.
15. Shamrat, F. M.; Mehedi, J.; Abu Raihan, Md.; Sazzadur Rahman, A. K. M.; Mahmud, I.; Akter, R.; others. An Analysis on Breast Disease Prediction Using Machine Learning Approaches. *Int. J. Sci. Technol. Res.* **2020,** *9* (02), 2450–2455.
16. Chowdhury, A. Breast Cancer Detection and Prediction Using Machine Learning, 2020.
17. Rawal, R. Breast Cancer Prediction Using Machine Learning. *J. Emerg. Technol. Innov. Res. (JETIR)* **2020,** *13* (24), 7.
18. Yang, J.; Yang, J. In *Aspect Based Sentiment Analysis with Self-Attention and Gated Convolutional Networks*, 11th International Conference on Software Engineering and Service Science (ICSESS), Vol. 2020; IEEE Publications, 2020, pp 146–149. DOI: 10.1109/ICSESS49938.2020.9237640
19. Hossen, Md. S.; Jony, A. H.; Tabassum, T.; Islam, Md. T.; Rahman, M. M.; Khatun, T. In *Hotel Review Analysis for the Prediction of Business Using Deep Learning Approach*,

International Conference on Artificial Intelligence and Smart Systems (ICAIS), Vol. 2021, 2021, pp 1489–1494. DOI: 10.1109/ICAIS50930.2021.9395757

20. What Is Cancer [Online]. n.d. https://www.nationalbreastcancer.org/what.
21. Breastcancer [Online]. n.d. http://www.who.int/cancer/detection/breastcancer/en/ (accessed Sep 20, 2016).
22. Womens-Health-Issue [Online]. n.d. http://www.stopcancerfund.org/pz-diet-habits-behaviors/lung-canceris-a-womens-health-issue/ (accessed Aug 20, 2016).
23. Debata, P.P.; Mohapatra, P.; Mishra, D.; Borah, S. A Chaotic-Jaya Optimized OSELM Model for Cancer Classification. In *Soft Computing Techniques and Applications. Advances in Intelligent Systems and Computing*; Borah, S., Pradhan, R., Dey, N., Gupta, P., Eds.;Springer: Singapore, 2021; Vol. 1248. https://doi.org/10.1007/978-981-15-7394-1_55

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

CHAPTER 2

REAL-TIME-BASED HEART PATIENT MONITORING SYSTEM: AN APPLICATION OF HEALTHCARE IOT

K. P. SWAIN[1], SOUMYA RANJAN SAMAL[2], SARITA MISRA[1], SARAT KUMAR SWAIN[3], TARAPRASANNA DASH[4], and SUBHASHREE CHOUDHURY[5]

[1]*Department of Electronics and Communication Engineering, GITA Autonomous College, Bhubaneswar, Odisha, India*

[2]*Department of Electronics and Communication Engineering, Silicon Institute of Technology, Bhubaneswar, Odisha, India*

[3]*Department of Electrical and Electronics Engineering, GITA Autonomous College, Bhubaneswar, Odisha, India*

[4]*Department of Electrical and Communication Engineering, Siksha 'O' Anusandhan (Deemed to be University), Bhubaneswar, Odisha, India*

[5]*Department of Electrical and Electronics Engineering, Siksha 'O' Anusandhan (Deemed to be University), Bhubaneswar, Odisha, India*

ABSTRACT

Currently, India has the highest cardiovascular disease (CVD) patients in the world. According to the year 2016 data, 54.5 million heart patients[1] are in India with a mortality rate of 1:4. According to WHO, India is accountable for one-fifth of this noncommunicable disease deaths.[2] Aside from medicinal treatment, the tracking of heart patients is very much necessary for the early management of the patient. Referring to the above information about CVD,

Applied Soft Computing Techniques: Theoretical Principles and Practical Applications. Samarjeet Borah, Ratna Raja Kumar Jambi, Sharifah Sakinah Syed Ahmed, & Mahendra Prabhakar Deore (Eds.)
© 2025 Apple Academic Press, Inc. Co-published with CRC Press (Taylor & Francis)

here an Internet of Things (IoT) application is proposed especially for tracking heart patients. This work is designed with Arduino Uno and a heartbeat sensor (Amped) along with a global system for mobile communications (GSM) and a global positioning system (GPS) module. The proposed system enables the patient concerned person with all details such as heart rate information and current position of the heart patient through a smartphone so that medical aid can be supplied immediately.

2.1 INTRODUCTION

Accurate heart rate monitoring is mostly possible in the medical center with the help of ECG machine where electrodes are attached to the human body with the help of wires and patches. However, the said method is a colossal process and needs technical knowledge to handle. Besides the ECG technique, there are two ways to measure the heart rate; one is manually checking the pulses in the wrist or neck and the other is the use of heartbeat sensors. Nowadays, using a heartbeat sensor is very handy as it comes in different shapes and sizes. These sensors are available now in smart watches, smartphones, chest straps, etc. where instead of mere reading, the related information can also be stored and distributed through the Internet using different IoT devices such as Arduino, Raspberry Pi, etc. The principle behind the heart rate sensor is photoplethysmography. According to this, the change in the volume of blood in an organ is measured by the change in the intensity of light passing through that organ.

2.2 LITERATURE SURVEY

Being a low-cost IoT device, Arduino is used in many applications such as medical, sensing, military, home automation, fault detection, and many more[3–10] with real-life implementation. In,[3] Arduino Uno is used to detect the fault in optical fiber line by sensing the received power from the transmitter end. An LDR with an operational amplifier (Opamp) is used as a sensor and the behavior of the circuit is simulated by Proteus integrated set of information systems (ISIS) simulation software. A web-based design is also proposed to monitor the date and time of the fault occurrence. The Arduino is used[4] for sending and receiving the binary information [on-off keying (OOK) signal] through visible light communication links. Two Arduino are implemented in the transmitting and receiving ends each where the light-emitting diode

(LED) is used for transmitting and a light-dependent resistor (LDR) is used for receiving the binary signal through the visible spectrum. Here, not only binary signal is communicated between two Arduino boards but also any electrical gadget can be controlled by sending the electrical control pulses. The Proteus ISIS simulation is realized in this work for validating the hardware. The image processing research is carried out[5] for the detection of motion via Arduino Uno. The MATLAB image processing toolbox is simulated through Proteus ISIS with other hardware like a webcam, LCD, and buzzer for realizing object detection. A remote sensing-based home automation system using Arduino Uno is realized in.[6] In the first part, the sensors are controlled through a smartphone manually and in the second part, the devices are controlled by an automatic management system. For the automatic management system, MATLAB graphical user interface (GUI) platform is designed with Arduino for easy and smart controlling operation. A home automation system for physically challenged people is proposed[7] using a brain-controlled interface (BCI). The human brain signal is analyzed by different electrodes which are placed on the head and converted into raw waveform and transmitted toward the Arduino board through Bluetooth module. Then Arduino triggers the required devices through various relays to perform the desired operation. A smart energy-saving system is proposed,[8] where Arduino Uno is used as pressure sensor via a piezoelectric sensor made of Rochelle salt. The work is validated experimentally and found useful in energy saving application in bridges and tunnels by turning on the lighting system when the vehicle enters the bridge and turned off when it crosses. The Arduino is proposed[9] for the measurement of urban air parameters such as CO, PM 2.5, CO_2, temperature, and humidity and further the said parameters are analyzed and forecasted. A combination of Arduino Uno, NANO, and XBee modules are implemented[10] to monitor the different parameters like heart rate, temperature, rumination, and humidity of cattle. By this, the farmers are profited by being watchful to spot any decline in cattle health.

Arduino is implemented in road condition detection device (RCDD)[11] in real-time detection with the help of an Android-based smartphone. Here, an Android user interface (AAP) is used to fetch the GPS location whenever an abnormal road condition is found. Then the information is spread to the nearby vehicle for improving the driver's safety and efficiency. A new Arduino-based potentiostat is demonstrated[12] to sense the hydrogen peroxide for the implementation of chronoamperometry and cyclic voltammetry functionality. In the agricultural plant monitoring system, Arduino is also used[13–16] for checking the various external factors like air, soil, light and other environmental parameters which can impact the harvests. Here, Ubidots IoT cloud platform is realized to collect the required parameters from Arduino and also the information is sent to the user's smartphone. To regulate the volumetric water contents in soil, Arduino is proposed[17] for an improved green irrigation system. Soil moisture is controlled by using Arduino[18] with other hardwires. In Ref. [19], the Arduino is used as a core control system of an obstacle avoidance wheeled robot equipped with infrared tracking, a motor driver, and power modules.

Pertaining to healthcare, Arduino plays a vital role. Several sensors[20] are incorporated for different parameters like body temperature, pulse oximeter, and airflow sensor to process the relevant information either locally or by using the cloud. An e-platform is implemented with MATLAB platform for reading the relevant data flawlessly. This is realized experimentally by considering 15 candidates between 21 and 52 years including both male and female. In Ref. [21], Arduino is employed to design a prototype for the diabetes patient community with the help of a Bluetooth-based glucometer and an Android mobile app. Smart healthcare monitoring systems are extensively used currently by using different sensing devices and processing devices with their associated hardware and software as discussed in.[22–26] The various vital parameters of a patient are monitored using different wearable devices where the related information is collected and processed through Arduino or Raspberry Pi, which are analyzed by the medical expert to provide the necessary aid immediately.[22, 23] Not only patient information but also room environment parameters[24, 25] of a patient are analyzed by medical specialists using sensors such as a heartbeat, body temperature, room temperature, CO, and CO_2 information in IoT environment. The heart rate and oxygen level of babies (0–18 months) are observed and analyzed using Arduino[26] by medical practitioners for any irregularities. The real-time information about patient health is processed and stored using Arduino and the ThingSpeak website[27,28] and different parameters are also analyzed using the cloud-based server with a mobile app.[29] By analyzing the breadth of a diabetic patient, the ketone level is monitored remotely using Arduino Uno and ESP 8266 Wi-Fi module.[30]

Aside from the patients, sports person's health conditions are monitored[31,32] by using different wearable devices and cloud servers. An IoT application comprising wearable sensors is described[31], especially for sports personnel. The health details with exercise records are effectively optimized by employing machine learning (ML) algorithms. Also, IoT application is implemented with the fog assistance wearable sensors networks[32] for athletic world health. Besides sports health monitoring, Arduino can also be used to draw the court lines via an automatic bot.[33]

The proposed system represents a groundbreaking integration of soft computing methodologies with IoT technology, aiming to revolutionize the way patient care is delivered in the context of cardiovascular disease (CVD). By harnessing the potential of soft computing techniques, such as artificial intelligence (AI), ML, and data analytics, the system enables patient caretakers to access a wealth of comprehensive information in real time.

One of the key features of the system is its ability to continuously monitor the patient's heart rate, providing caretakers with up-to-date and accurate data. This real-time heart rate information allows for proactive interventions and timely responses to any abnormalities or distress signals. By leveraging soft computing algorithms, the system can detect patterns and trends in the heart rate data, enabling early identification of potential risks or anomalies that might otherwise go unnoticed. This empowers caretakers to take swift action and provide necessary medical attention promptly, potentially preventing adverse events or complications.

Additionally, the proposed system incorporates precise geolocation capabilities through the integration of GPS technology. This means that caretakers can track the patient's current position in real-time, enhancing the efficiency of medical aid dispatch. In emergencies, knowing the exact location of the heart patient enables responders to reach the individual quickly and deliver targeted interventions. By optimizing the dispatch process through intelligent algorithms and data analytics, the system ensures that medical assistance reaches the patient in the shortest possible time, maximizing the chances of positive outcomes.

The fusion of soft computing techniques with IoT technology holds immense promise for enhancing patient care and improving overall health outcomes. By leveraging AI and ML algorithms, the system can continuously learn from the collected data, adapting and refining its predictive capabilities over time. This enables personalized and tailored interventions for each patient, as the system gains insights into individual patterns and responses. As a result, the system has the potential to optimize treatment plans, minimize risks, and improve the overall quality of care provided to CVD patients.

At last, the proposed system represents a significant advancement in healthcare technology by combining soft computing methodologies with IoT infrastructure. By providing comprehensive real-time heart rate information, precise geolocation tracking, and optimizing medical aid dispatch through intelligent algorithms, the system empowers caretakers to deliver immediate and targeted interventions. This promising fusion of technologies has the potential to revolutionize patient care, improve health outcomes, and contribute to the ongoing battle against CVDs.

Precisely, in this work, an emergency health tracking system is designed using GSM, GPS, and a heartbeat sensor with Arduino. Here Amped which is an Arduino-compatible heartbeat sensor is implemented and the data which are generated by the sensor is transmitted via the GSM900A GSM module. The GSM module is used to send a message and call the respective person to get the details. Also, a GPS module is interfaced with Arduino to send the exact location to the concerned person. A 16×2 I2C-based LCD is also connected to display the real-time heart rate information.

The entire research article comprises of five sections. The deep insights into the system overview have been projected in Section 2.2. Section 2.3 discusses the circuit description. The implementation and the working of the entire system have been thoroughly described in Section 2.4. Finally, the conclusion from the entire study has been drawn in Section 2.5.

2.3 SYSTEM OVERVIEW

To implement the proposed work, Arduino Uno with its related IDE is used as the prime component of the whole system. Aside from Arduino other hardware such as Amped heartbeat sensor, 16×2 LCD (I2C based), LED, push button, Buzzer, SIM900A GSM module, GPS module (PA6E-CAM), and jumper wires are required. A detailed description of each system component is presented in this section. Figure 2.1 shows the block diagram of the proposed system.

2.3.1 AMPED HEART RATE SENSOR

It is an optical heart rate sensor that is compatible with Arduino with plug-and-play facilities. It is very popular due to its amplification and noise cancelation properties for fast and reliable pulse reading. It comes with three connecting wires for power [voltage common collector (VCC)], ground (GND), and signal (S) which are to be connected to the analog pin of the Arduino. It is operated with either 3 or 5 V. It gives a noticeably accurate and reliable pulse reading.

2.3.2 16×2 LCD (I2C BASED)

LCD is used to display the heartbeat of a person locally with the help of the Arduino. Instead of a general LCD, an I2C-based LCD is implemented here to reduce the pin count while connected with Arduino. Primarily, a simple

LCD requires 6 pins (RS, RW, D7, D6, D5, and D4) while an I2C LCD needs only two pins [serial data line (SDA) and serial clock line (SCL)] to communicate with any microcontroller unit (MCU) thereby reducing four pins. The I2C circuitry is packed with the backside of the LCD module called the backpack. This type of LCD works with an input/output (I/O) expander (PCF8574) for communication via the I2C protocol. Here, PCF8574 provides remote I/O expansion for most of the MCU via two bidirectional lines (SDA and SCL). This device consists of an 8-bit quasi-bidirectional port and an I2C bus interface. The quasi-bidirectional lines act as input and output ports without the use of data direction control signals. It also consumes a very low current for operation and is capable of handling high-current drive devices. It also has an interrupt signal to inform the microcontroller about the incoming data available on its port. Here, the communication between I2C-based LCD and MCU is held at a 9600 baud rate.

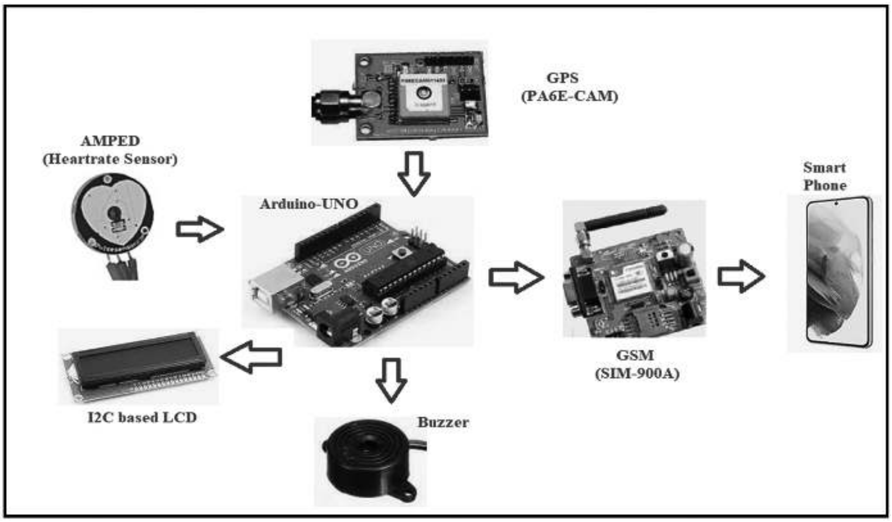


FIGURE 2.1 Block diagram of the proposed system having the major components such as Arduino uno, amped sensor, GPS and GPRS module, I2C-based LCD module, buzzer.

2.3.3 SIM900A GSM MODULE

This GSM module can accept any GSM network operator and also acts as a simple mobile phone. It is having RS 232 port through which it communicates with the mobile operator. It can also be operated through the universal serial bus (USB) port of any personal computer (PC) or any microcontroller directly. It can also be used in general packet radio service (GPRS) mode to connect to the Internet for data logging and control operation. It is operated by 12V DC and communicated serially through serial pins [receive (Rx) and transmit (Tx)].

2.3.4 GPS (PA6E-CAM) MODULE

Commonly, GPS is used to find the exact location in terms of latitude and longitude with the help of the satellite on the earth along with the exact coordinated universal time (UTC) time. Currently, the GPS module is used as the prime component in automobile tracing which collects the location coordinates from the corresponding satellite at some time interval with the date on a real-time basis. Scientifically, it sends the relevant information in the National Marine Electronics Association (NMEA) format. This format comprises a number of sentences, in which only one sentence is appropriate for user information. Usually, the sentence starts from $GPGGA which is the GPS fix data that contains the coordinate, time, and other useful information. Users can extract the location information by considering the commas available in the $GPGGA string. In general, $GPGGA strings are stored in the form of an array, then after two commas, latitude can be found, and then after two commas longitude is obtained. In the GPS, each satellite has a stable atomic clock on board for the dissemination of precise time data which can be received by a receiver. This system comprises four satellites in range for an accurate “lock” for the earth’s position.

2.4 CIRCUIT DESCRIPTION

The circuit diagram shows, a heartbeat sensor (Amped) for measuring the heart rate, a push button switch for making the call and sending the location data from GPS via a GSM module (GSM900A), and an I2C-based LCD (16×2) to display the heart rate are connected with Arduino Uno to make a complete healthcare tracking system for a heart patient. Here, Arduino acts as the main controller and processing unit for the whole system. It reads the heartrate information from Amped, observes the value for a specified range, sends the information to LCD, and also send a message and makes phone calls through GSM by using button manually and also sends the location information from the GPS module in case of any irregularity in heart rate. Figure 2.2 depicts the circuit diagram of the proposed system.

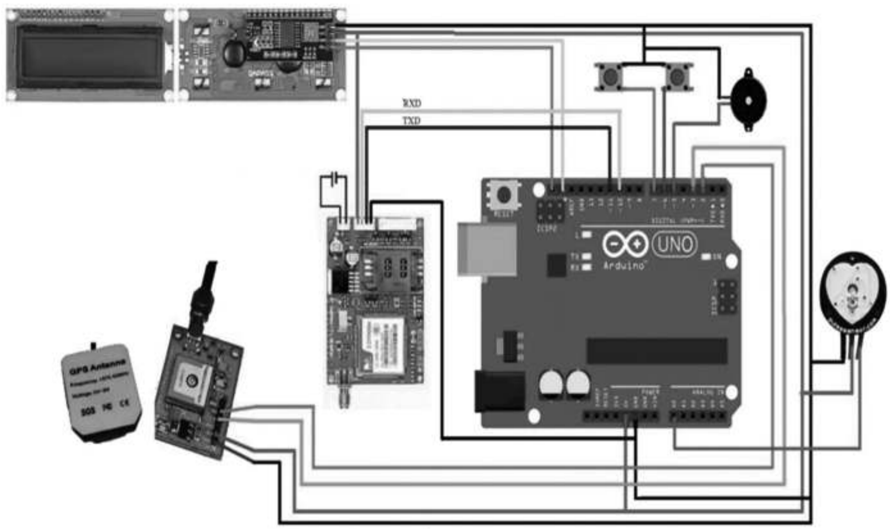


FIGURE 2.2 Circuit diagram of the proposed system where red and black color wires indicate VCC and GND respectively.

About the details of circuit connection, the Amped output is directly connected to the analog input (A0) pin of Arduino. The LCD is connected with Arduino via SCL and SDA pins of Arduino. Two push button switches are added for making a call and sending GPS location to the Arduino digital pins 6 and 7. A buzzer is also connected to Arduino digital pin 5 for an emergency alert. The Rx and Tx pins of the GSM module are connected to the digital pins 10 and 11 of Arduino, respectively. The GPS module is connected to Arduino via digital pins 2 and 3. A gist of the connection details is provided in Table 2.1.

2.5 LIBRARY FUNCTION, SENSORS VARIABLES WITH PROGRAM SEQUENCE

2.5.1 *HEADER FILE AND PINS VARIABLES*

This work includes four header files for smooth conductance of all the functionalities; SoftwareSerial.h, TinyGPS.h, Wire.h, LiquidCrystal_I2C.h.

SoftwareSerial.h: This header file is used to connect the GPS and GPRS module with Arduino through SoftwareSerial library file.[34] This library file allows the serial communication between GPS, GPRS module, and

Arduino microcontroller board. It has a limitation in that it cannot transmit and receive simultaneously although multiple software ports are activated. The speed of data transmission can go up to 115,200 bps. The advantage of using this library is; that it ensures the data exchange serially by any digital pins through software. This header file can only be used with Arduino IDE version 1.0 and above and only be realized through some specific version of Arduino board.

TABLE 2.1 The Connection Details Between Arduino Uno and Other Major Components Like Amped Heart Rate Sensor, GPS and GPRS Module, I2C LCD Module, and Push Button Switches.

	Arduino uno
Heartbeat sensor	
Signal (s)	Analog pin (A0)
VCC	VCC
GND	GND
16×2 LCD	
SCL	SCL
SDA	SDA
VCC	VCC
GND	GND
Push button	
Button 1	Analog pin 6
Button 2	Analog pin 7
Buzzer	Digital Pin 5
GSM module	
Rx	Digital pin 10
Tx	Digital pin 11
GPS module	
Rx	Digital pin 2
Tx	Digital pin 3

TinyGPS: It is the compact Arduino NMEA (GPS) parsing library which is used for parsing NMEA data streams to Arduino.[35] It provides compressed and user-friendly techniques for extracting the GPS-related information from the GPS. It converts the NEMA format GPS information into user understandable information. Some inbuilt functions like gps.encode(c), gps.

get_position(&latitude, &longitude, &age), gps.f_get_position(&flatitude, &flongitude, &age) are associated with this library.

Wire.h: This library is used for communication between Arduino and I2C devices through two hardware pins; SDA and SCL.[36] Internally, it uses a 32-byte buffer which simply discards the data if the size exceeds its limit. Some of the inbuilt functions include begin(), end(), requestFrom(), beginTransmission(), endTransmission(), etc.

LiquidCrystal_I2C.h: This library controls the display functionalities in IIC-based LCD devices in a very simple way.[37] It is also compatible with other AVR family microcontrollers. In the program, some variables are used for some specific purpose as shown in Table 2.2.

TABLE 2.2 The Details of Variable and Its Datatypes with their Preassign Value.

Variable name	Data types	Value
PulsePin	Int	0
Buzzer	Int	5
Call	Int	6
Location	Int	7
BPM	Volatile Int	
Signal	Volatile Int	
IBI	Volatile Int	600
Pulse	Volatile Boolean	False
QS	Volatile Boolean	False

2.6 PROGRAM FUNCTIONALITIES DIRECTION

Step 1 When the heart rate is detected from the sensor by Arduino, the variable QS becomes true and prints the heart rate value on LCD using the same variable and in the serial monitor using the variable BPM.

Step 2 When the heart rate goes higher than 120, then the buzzer is activated using the digitalwrite() function 3 times at some definite interval.

Step 3 At the same time, when the user presses the call button, gsmcall() is activated to call the respective person's mobile number.

Step 4 When the user presses the location button, the gpslocation() is activated to send the tracking location to the concerned mobile number. After that, the concerned person receives the longitude/latitude information through messaging [short messaging service (SMS)]. There is also a provision available to change the person mobile number by editing the gsmlocation ().

Step 5 At last, the quantified self-flag is reset for the use of another time.

2.7 IMPLEMENTATION

According to the circuit diagram, the heartbeat monitoring system is designed with Arduino Uno, a heartbeat sensor, buttons for making a call and sending GPS location, and 16×2 LCD to display the pulse data. Also, it includes a GSM and GPS module which helps in sending messages and initiates calls by interacting with buttons. Here, Arduino Uno is the major controlling unit that controls the whole process of the system such as reading pulses from the heartbeat sensor module, calculating heart rate, and sending this data to 16×2 LCD. The following table describes all the connection details among all the sensors and other modules with Arduino Uno.

Figure 2.3 represents the complete hardware implementation of the whole system along with a mobile phone operated at the 900 MHz range. Here, the GSM module is used to send the coordinates and make a call on a mobile phone via message. After finding any irregularity in the value of the pulse sensor the GPS sends the coordinates and, in the meantime, the buzzer is activated to alert the user. Then the user pushes the button to make the call and sent the location information. Arduino extracts the sensor information and sends it to a mobile phone using a GSM module via SMS.

Figure 2.4 shows the reading of heartrate from Amped and displaying at LCD at regular intervals of time. The value in the COM port is reflected at LCD which confirms the accurate reading from Amped. If the heartrate information is not within the specified range, then the GPS module sends the coordinate of the patient position for sending the SMS and calling the responsible person afterward which is shown in Figures 2.5–2.7 sequentially.

2.8 CONCLUSION

An open-source heart rate monitoring system is successfully implemented here using Arduino Uno as a prime component. The Arduino is used to track the heart stroke person immediately using GSM and GPS modules. Any abnormality in a patient's heart can be informed to the responsible person through the mobile phone with the exact location and other details. The whole system makes a suitable application for an IoT application targeted toward heart patients.

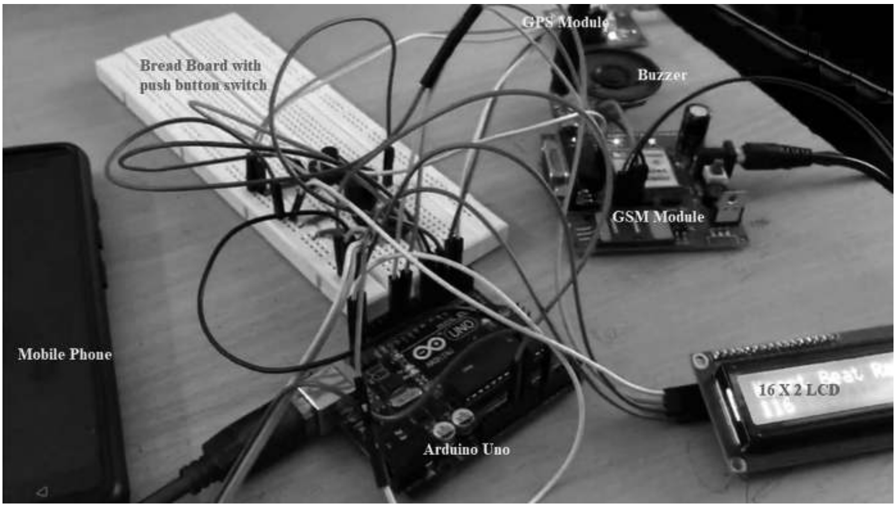


FIGURE 2.3 Hardware implementation of the complete system showing all hardwires such as Arduino Uno, GPS, and GSM module, and breadboard containing push button switches.

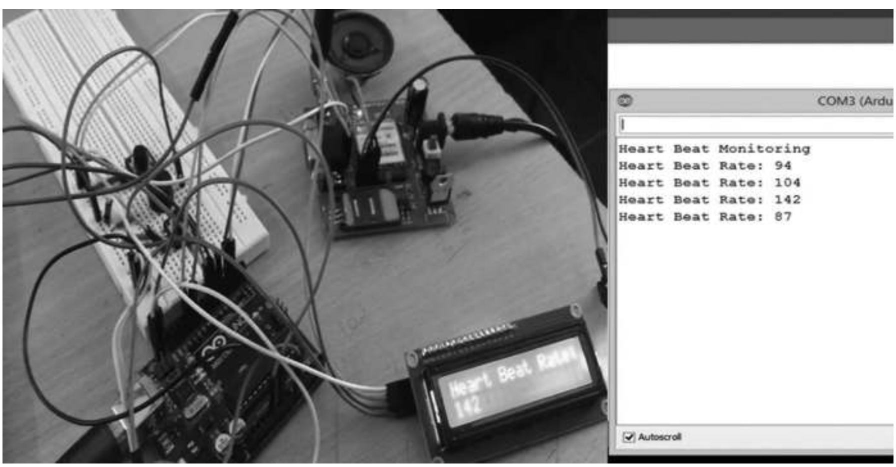


FIGURE 2.4 Display of heartbeat rate on LCD and also the display of heartbeat rate through COM3 port.

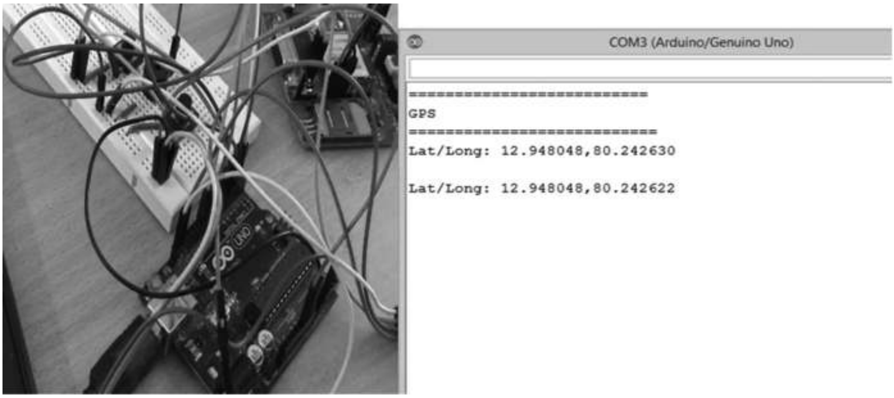


FIGURE 2.5 Collecting the location information through GPS Module on COM3 port.

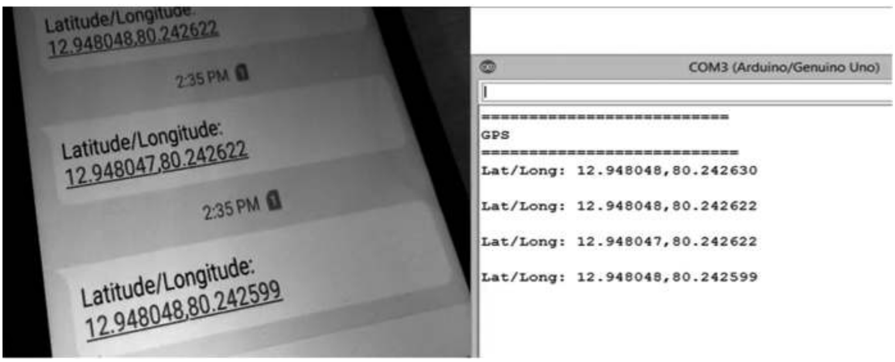


FIGURE 2.6 Receiving the location information (Lat/Long) by mobile phone through SMS.

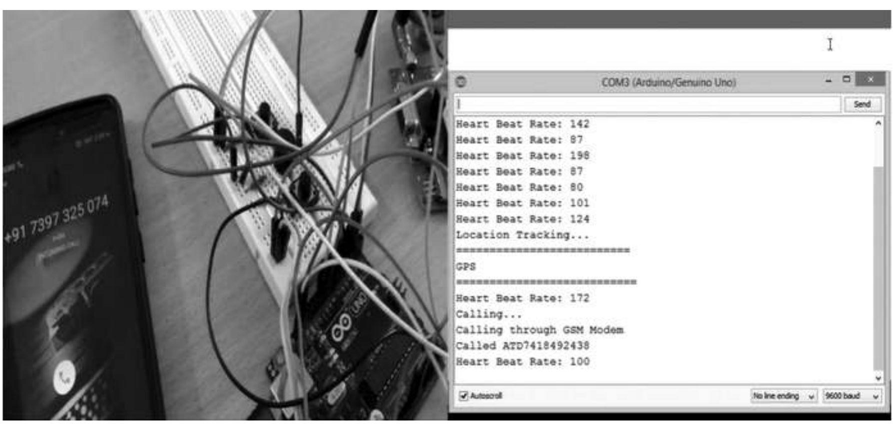


FIGURE 2.7 Sending the information and calling the person concerned when any irregularity of heart rate is found.

This work can be further improved by connecting more sensors related to healthcare such as breadth meter, pressure, force, airflow, oxygen, pulse oximetry, temperature sensors, etc. A web-based application can be designed for the ICU for continues monitoring the various health related parameters. In the web-based application, all the sensors which are interfaced with Arduino can transmit the relevant information to Thing Speak (a web server) through ESP8266 wirelessly for further action. Besides this, mobile apps can also be developed for the benefit of mankind.

The envisioned system harnesses the potential of soft computing methodologies to provide patient caretakers with comprehensive details such as real-time heart rate information and precise geolocation through a smartphone interface. By employing intelligent algorithms and data analytics, the proposed system may help to optimize the efficiency of medical aid dispatch, ensuring immediate and targeted interventions. This harmonious fusion of soft computing techniques with IoT technology holds significant promise in enhancing patient care and improving overall health outcomes.

KEYWORDS

- **Arduino**
- **heartbeat sensor**
- **GSM900A**
- **I2C-based LCD**
- **GPS**

REFERENCES

1. Abdul-Aziz, A. A.; Desikan, P.; Prabhakaran, D.; Schroeder, L. F. Tackling the Burden of Cardiovascular Diseases in India. *Circ.: Cardiovasc. Qual. Outcomes* **2014,** *12* (4).
2. Kumar, A. S.; Sinha, N. Cardiovascular Disease in India: A 360 Degree Overview. *Med. J. Armed Forces India.* **2020,** *76,* 1–3.
3. Swain, K. P.; Sahoo, J.; Prasad, M. V. S. V.; Palai, G. Fault Detection System in an Optical Fiber Using Arduino. *Int. J. Appl. Eng. Res.* **2015,** *10,* 31745–31749.
4. Swain, K. P.; Palai, G.; Sahoo, J.; Prasad, M. V. S. V.; Mohanty, M. N. Exploiting VLC Tchnique for Smart Home Automation Using Arduino. Advances in Intelligent Systems and Computing. In *Artificial Intelligence and Evolutionary Computations in Engineering Systems*, 2017; pp. 211–220.

5. Das, S. R.; Swain, K. P.; Sahu, S. In *Motion Sensing using MATLAB and Arduino*, International Conference on Applied Electromagnetics, Signal Processing and Communication (AESPC), IEEE: Bhubaneswar, India, 2018.
6. Iman, I. M.; Abu, S.; Sami, H. A. A.; Mohamed, S. S.; Majed, O. D. In *Designing and Implementation of Home Automation System Based on Remote Sensing Technique with Arduino Uno Microcontroller*, IEEE-GCC Conference and Exhibition (GCCCE) IEEE: Manama, Bahrain, 2017.
7. Babu, K. M. C.; Vardhini, P. A. H. In *Brain Computer Interface based Arduino Home Automation System for Physically Challenged*, International Conference on Intelligent Sustainable Systems (ICISS) IEEE: Thoothukudi, India, 2020.
8. Kaur, A.; Saini, S. S.; Singh, L.; Sharma, A.; Sidhu, E. In *Efficient Arduino UNO Driven Smart Highway/bridge/tunnel Lighting System Employing Rochelle Piezoelectric Sensor*, International Conference on Control, Computing, Communication and Materials (ICCCCM) IEEE: Allahabad, India, 2016.
9. Lobur, M.; Korpyljov, D.; Jaworsk, N.; Iwaniec, M.; Marikutsa, U. M. In *Arduino Based Ambient Air Pollution Sensing System*, International Conference on the Perspective Technologies and Methods in MEMS Design (MEMSTECH) IEEE: Lviv, Ukraine, 2020.
10. Swain, K. B; Mahato, S.; Patro, M.; Pattnayak, S. K. In *Cattle Health Monitoring System Using Arduino and LabVIEW for Early Detection of Diseases*, International Conference on Sensing, Signal Processing and Security (ICSSS) IEEE: Chennai, India, 2017.
11. Chen, S. Y.; Shih, A.; Hsiao, C. Y. In *Road Conditions Detection Using Arduino based Sensing Module and* Smartphone, International Conference on Consumer Electronics, IEEE: Taiwan, 2015.
12. Gao, W.; Luo, X.; Liu, Y.; Zhao, Y.; Cui, Y. Development of an Arduino-based Integrated System for Sensing of Hydrogen Peroxide. *Sens. Actuators Rep.* **2021,** *3,* 100045
13. Audrey, D. A. D.; Stanley; Tabaraka, K. S.; Lazaro, A.; Budiharto, W. Monitoring Mung Bean's Growth using Arduino. *Proc. Comput. Sci.* **2021,** *179,* 352–360.
14. Cui, X.; Kou, Z.; Shi, Y.; Yan, X.; Wang, Z.; Huang, H.; Dong, Y. In *Design and Experiment of Control System for Intelligent Agriculture based on Arduino*, International Information Technology and Artificial Intelligence Conference (ITAIC), IEEE: Chongqing, China, 2020.
15. Mittal, A.; Sarma, N. N.; Sriram, A.; Roy, Trisha.; Adhikari, S. In *Advanced Agriculture System Using GSM Technology*. International Conference on Communication and Signal Processing (ICCSP), IEEE: Chennai, India, 2018.
16. Sushanth, G.; Sujatha, S.; In *IOT Based Smart Agriculture System*. International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), IEEE: Chennai, India, 2018.
17. Zhu, H. H.; Huang, Y. X, Huang, H.; Garg, A.; Mei, G. X.; Song, H. H. Development and Evaluation of Arduino-Based Automatic Irrigation System for Regulation of Soil Moisture. *Int. J. Geosynth. Ground Eng.* **2022,** *8* (13).
18. P V S, Divya Dhatri.; Pachiyannan, M.; Rani K, J. S.; Pravallika, G. In *A Low-Cost Arduino based Automatic Irrigation System using Soil Moisture Sensor: Design and Analysis*, International Conference on Signal Processing and Communication (ICSPC) IEEE: Coimbatore, India, 2019.
19. Li, Z. F.; TaoLi, J. T.; FanLi, X.; Yang, Y. J.; Xiao, J.; WenXu, B. Intelligent Tracking Obstacle Avoidance Wheel Robot Based on Arduino. *Proc. Comput. Sci.* **2020,** *166,* 274–278.

20. Zainee, N. M.; Chellappan, K.; In *Emergency Clinic Multi-Sensor Continuous Monitoring Prototype Using e-Health Platform*, IEEE Confernce on Bbiomedical Engineering and Science: Malaysia, 2014.
21. Sabbir, A. S.; Bodroddoza, K. M.; Hye, A.; Ahmed, M. F.; Saha, S.; Ahmed, K. I. In *Prototyping Arduino and Android Based m-Health Solution for Diabetes Mellitus Patient*, International Conference on Medical Engineering, Health Informatics and Technology (MediTec), IEEE: Bangladesh, 2016.
22. Vippalapalli, V.; Ananthula, S. In *Internet of things (IoT) Based Smart Health Care System*. International Conference on Signal Processing, Communication, Power and Embedded System, IEEE: India, 2016.
23. Banka, S.; Madan, I.; Saranya, S. S. Smart Healthcare Monitoring using IoT. *Int. J. Appl. Eng. Res.,* **2018,** *13,* 11984–11989.
24. Islam, M. M.; Rahaman, A.; Islam, M. R. Development of Smart Healthcare Monitoring System in IoT Environment. *SN Comp. Sci.* **2020,** *1,* 1–11.
25. Valsalan, P.; Baomar, T. A. B.; Baabood, A. H. O. IoT Based Health Monitoring System. *J. Crit. Rev.* **2020,** *7,* 739–743.
26. Abdulameer, T. H.; Ibrahim, A. A.; Mohammed, A. H. Design of Health Care Monitoring System Based on Internet of Thing (IOT). *4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT),* IEEE: Turkey, 2020.
27. Hartalkar, A.; Kulkarni, V.; Nadar, A.; Johnraj, J.; Kulkarni, R. D. In *Design and Development of Real Time Patient Health Monitoring System using Internet of Thing*. 1st International Conference for Convergence in Engineering (ICCE), IEEE: India, 2020.
28. Shalini, V. B. In *Smart Health Care Monitoring System based on Internet of Things (IOT)*, International Conference on Artificial Intelligence and Smart Systems, IEEE: India, pp. 1449–1453, 2021.
29. Yeri, V.; Shubhangi, D. C. In *IoT based Real Time Health Monitoring*, International Conference on Inventive Research in Computing Applications, IEEE: India, 2020.
30. Rahman, R. A.; Aziz, N. S. A.; Kassim, M.; Yusof, M. I. IoT-based Personal Health Care Monitoring Device for Diabetic Patients. *Symposium on Computer Applications & Industrial Electronics,* IEEE Malaysia, 2017, pp. 168–173.
31. Huifeng, W.; Kadry; Raj, E. D. Continuous Health Monitoring of Sportsperson Using IoT Devices Based Wearable Technology. *Comput. Commun.* **2020,** *160,* 588–595.
32. Li, S.; Zhang, B.; Fei, P.; Shakeel, M.; Samuel, R. D. J. Computational Efficient Wearable Sensor Network Health Monitoring System for Sports Athletics Using IoT. *Aggress. Violent Behav.* **2020,** 101541.
33. Megalingam, R. K.; Sreevatsava, R. M.; Nidamanuru, H. S. V. S. K.; Gadde, L. In *Arduino based Automated Sports Court Drawing Bot*, International Conference on Electronics and Sustainable Communication Systems (ICESC) IEEE: Coimbatore, India, 2020.
34. https://docs.arduino.cc/learn/built-in-libraries/software-serial (accessed Sep 4, 2022).
35. http://arduiniana.org/libraries/tinygpsplus/ (accessed Sep 4, 2022).
36. https://www.arduino.cc/reference/en/language/functions/communication/wire/ (accessed Sep 4, 2022).
37. https://www.arduino.cc/reference/en/libraries/liquidcrystal-i2c/ (accessed Sep 4, 2022).

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

CHAPTER 3

SENTIMENT AND DEPRESSION ANALYSIS USING MACHINE LEARNING

K. P. SWAIN[1], SARITA MISRA[1], S. K. NAYAK[2], S. R. SAMAL[3], and G. PALAI[1]

[1]*Department of Electronics and Communication Engineering, GITA Autonomous College, Bhubaneswar, India*

[2]*Department of Computer Science Engineering, GITA Autonomous College, Bhubaneswar, India*

[3]*Department of Electronics and Communication Engineering, Silicon Institute of Technology, Bhubaneswar, India*

ABSTRACT

Depression and anxiety are the most health hazards nowadays for modern lifestyle which may lead to suicide. This can be prevented by proper treatment if it is detected well before. The prediction of such things is possible now with advent of data science. In this work, the corpus of blog posts in social media is collected to identify the depressed people. Here, the various text encoding schemes like natural language tool kit, BOW, term frequency–inverse document frequency are implemented to extract the features and applied to Numpy, Panda-like classifiers to differentiate between clinical and control subjects to specify the handler mindset.

3.1 INTRODUCTION

The social media platforms like Twitter, Facebook, and Instagram are popular in youth at the present time to express their views, expression, thoughts,

Applied Soft Computing Techniques: Theoretical Principles and Practical Applications. Samarjeet Borah, Ratna Raja Kumar Jambi, Sharifah Sakinah Syed Ahmed, & Mahendra Prabhakar Deore (Eds.)
© 2025 Apple Academic Press, Inc. Co-published with CRC Press (Taylor & Francis)

and so on these posts sometimes indicate their mind states which can be analyzed by different tools like machine learning and deep learning. In Ref. [1], sentiment analysis is performed by using Bengali text preprocessing and count vectorizer and result is predicted by using naïve Bayes with accuracy 86.7% among six machine learning (ML) models. The LSTM network[2] model along with perinatal depression screening is used to detect the depression level in the Edinburgh Postnatal Depression scale. The article[3] realizes the NLP to detect the depression level using emotion analysis. In this, support vector machine and naïve Bayes classifier are implemented to identify the same. The data mining[4] is applied on psychology area to find the depressive people who are engaged in the social network. A study is carried out[5] which are based on sentiment analysis approaches with their challenges and related future directions. The tweet sentiments are compared by machine learning approach in Ref. [6] during the COVID-19 era. The article in Ref. [7] presents a review on application of drugs for sentiment analysis by comparing various deep learning approaches. In Ref. [8], depression prediction is focused on employment field and determines the sadder person between two sexes of people. Various soft computing and machine learning techniques can be valuable in depression analysis by assisting in the processing and interpretation of complex and uncertain data.

Sentiment analysis has become one of the emerging research fields using natural language processing. Different algorithms using machine learning can be applied to a given paragraph to investigate the sentiment of a human being whether he/she is happy or unhappy. The test of the paragraph needs to be preprocessed first and then it will be subjected Tokenization by count vectorizer. Multinomial naïve Bayes very accurately analyzes the human's sentiment with an accuracy of 86.67%.[8] Depression causes mental health hazards which result in early age death. Depression leads to suicidal attempts and degrades the quality of day-to-day lives. Researchers, nowadays mainly focusing on emotions of human beings using artificial intelligence in the area of data mining. Huge amount of information in terms of texts and images are available from social media itself for sentiment as well as depression analysis. NL processing can be applied over Twitter data to investigate the severity of depression sufferers. This investigation is carried out by classifying individual tweets with support vector machine and naïve Bayes as normal or negative basing upon some list of words.[2] Automatic systems have been proposed to be designed using artificial intelligence for the proper diagnosis of depression. The speech is said to possess important features to investigate the level of depression. The selection of these important features in speech through manual design consumes labor and time too. The features obtained through deep learning procedure in neural networks show superior performance than manual handy features. Both features are combined together to effectively measure the human depression strategy.[9] The features through deep earning are obtained from raw speech data by applying deep convolutional neural networks to it. Then media robust extended local binary patterns descriptor is used to capture the state of art from the spectrogram of raw speech data. Then at the end joint fine-tuning layer is proposed to capture the complementary information from both manual and deep learning features. This method shows better performance in the analysis of depression in comparison with audio-based approaches. The lifestyle of a person degrades due to depression while he/she is unconscious about the fact that he is suffering from such a health hazard. The researcher predicts the level of depression for different genders on the basis of their satisfaction and dissatisfaction in the current working place.[10] The feedback of job satisfaction is explored in terms of different factors such as monthly income, last job, age, city, occupation, marital status, comments of colleagues, and so on. Different ML classifiers such as K neighbors, naïve Bayes, random forest classifier, random forest regression, factor analysis are used to predict stress level activities in employees indicating their job satisfaction or dissatisfaction. This is done through data analysis, data visualization, and different ML algorithms. The highest accuracy among all the ML algorithms predicts depression more effectively.

Sentiment analysis is carried out through a person's understanding in various circumstances. The entire process involves data analysis and classification using social media data in terms of text data, emoji's, and so on with various AIML techniques. The highest accuracy among all the ML algorithms predicts depression more effectively. Multiclass classifier with deep learning algorithms classifies the social media data in a more précised way comparison to binary or ternary classifier.[11] Cyber network provides unlimited horizon for the people to share their daily life activities, their remarks, and influence for sentiment analysis. Based upon their posting and comments in social media, their severity of stress can be analyzed. This huge database along with tweets can be accomplished for sentiment analysis. Different machines learning algorithms, BERT sentiment classifier, and a deep learning model are unused for analyzing the sentiments.[12] Latent Dirichl et al. location, which is basically an unsupervised machine learning method, is used to scan a group of documents, recognize the word and phrase patterns within that group and gather the similar expressions which most suitably describe the documents. This in turn relates the topic to the textual data and helps to signify the emotions of online platform user. Several depression patients share their feelings via social media during COVID-19 pandemic situation. Hence, real-time detection of stressor depression is done using big data analysis. The stress detection is done by taking into account demographic features and views of Twitter users on patient health questionnaire 9.[13] Five different ML techniques have been applied such as support vector machine, decision tree, naïve Bayes, random forest, and deep learning. Random forest technique indicates accurate result in detecting the stress. This model identifies the mental state condition of depression patients. This piece of work is dedicated to preventing the suicidal death rate.

The vital information is captured to take decisions during data analysis through gathering of data, cleaning of data, transformation and modeling of data, interpretation, and visualization of data. Data are collected from big data stores like social media platform. Data cleaning is needed to make it error free and to achieve that all the redundant words, unnecessary spaces should be removed. The collected cleaned data is subjected to data analysis step. The analyzed data will then be represented in terms of easy words, flow charts, and tables. Then at last the data are visualized in the form of flow charts and graphs. Reports are available in literature dealing with binary classification where human sentiment is classified as positive polarity and negative polarity.[14,15] Later, ternary classifier comes into play where the collected data are classified as positive, negative, and neutral.[16,17] Multiclass classifier classifies the data more precisely where data are classified into a number of subclasses based upon their polarity.[18,19] The sentiment WordNet is basically a dictionary which is used to interpret the sentiment of the sentences.[20] The emotion behind the tweet texts is classified as fear ness, sadness, worries, anxieties, joy, and surprise.[21] The level of depression is classified as normal, mild, borderline, moderate, and severe depression.[22] Reshma et al. divides the classes into depression, stress, and normal, relax, happy, and unhappy.[23] The feature extraction process extracts and combines some important features from the classified groups and in doing this amount of initial data reduces along with the redundant data. Term frequency–inverse document frequency (TF–IDF), term document matrix, linguistic inquiry, and word count (LIWC) are different feature extraction techniques which emphasizes on a particular word or group of words and their frequency of occurrence in the classified documents based upon which features are to be extracted.[24–26]

The process of training a system to take accurate decisions is termed ML. Different machine learning algorithms are proposed in literature to train the machine for making predictions with more accuracy.[27,28] Table 3.1 summarizes the level of accuracy using different ML algorithms. The prediction accuracy level of support vector machine (SVM), multinominal naïve Bayes (MNB), random forest (RF), ensemble vote classifier, gradient boost, and K-nearest neighbor was 74.18%, 77.89%, 81.04%, 85.09%, and 79.12%, respectively.[24,29,30] Govin Gaikward et al. have shown the accuracy level of SVM, naïve Bayes (NB), and KNN are 82%, 64%, and 73%, respectively.[31] Solakidis et al. improves the prediction accuracy of multinomial naïve Bayes, SVM, LOG to 92.2%, 93.1%, and 93.2%, respectively.[32] Chatterjee suggested the decision level of naïve Bayes to be 76.7%.[33] Jose et al. demonstrated that the accuracy level of Senti Word Net (SWN), naïve Bayes, HMM, and ensemble approach (EA) are 21.05%, 69.9%, 64.06%, and 71.46%, respectively.[34]

TABLE 3.1 Comparison of Different Machine Learning Algorithms.

Classifier	Accuracy	Dataset	Paper
SVM	74.18%	Twitter	[35]
NB, SVM, RF	74.2%,81.2%,72.5%	Twitter	[24]
MNB, SVM	78%, 79.7%	Twitter	[17]
SVM, RF	68.57%,84.6%	Facebook	[36]
MNB, RF, EV	77.89%, 81.04%,85.09%	Twitter	[29]
NB	76.6	Facebook	[33]
MNB, SVM, LOG	92.2%,93.1%,93.2%	Facebook	[32]
KNN, RF, NB, SVM	72%,82%,71%,79%	Twitter	[30]
SWN, NB, HMM, EA	21.05%,69.92%,64.06%,71.46%	Twitter	[34]
MNB, SVM	83%, 79%	Twitter	[3]

In the current research, a supervised model is proposed to detect the depressive words posted on Twitter which indicates whether the corresponding user is gloomy or not and the premature indications of depression.

3.2 PROPOSED WORK

Figures 3.1 and 3.2 show the steps required to execute the proposed work which comprises data collection using Twitter API, data preprocessing, data training, testing and evaluation, execution of algorithm on 26,000 classified tweets. Using Google Colab on Google Drive raw data are pulled from Twitter API's, that is, Tweepy. It is reusable library function to pull data from open source. In this case, from twitter.developer.com, around 26k tweets are extracted from Twitter API and these tweets contain a three-point ordinal scale, that is, Zero to four target variables.

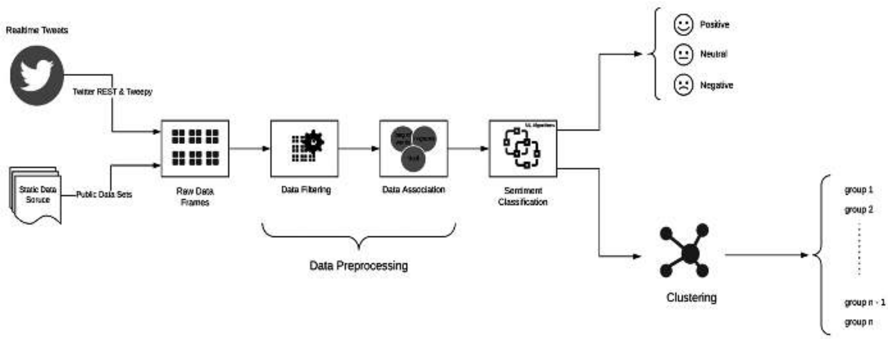


FIGURE 3.1 The schematic diagram of the proposed work.

By NLTK (natural language tool kit) with Pandas' library functions the raw data are fetched to rows and columns in JASON format. Here we will have text and ID in rows and columns to preprocess the data, respectively. Then the URLs and posts are eliminated which will be followed by string operations for handling the textual data. After that, data are pre-processed by removing Twitter handles, that is, @ user, links that is Http or https, punctuation, numbers, and special characters. These operations will help us to prepare the data to be fed to our machine learning model. In the data cleaning process, numeric values are removed, taking the tweets and converting it to tokens (tokens means the texts separated by space in text). Normalization is done in two ways one of such ways is stemming and in stemming we need to remove the affixes, prefixes, and suffixes, small words, that is, casing keeping all the words in lower case. Lastly, in the data visualization process, all positive and negative words are distinguished.

3.3 IMPLEMETATION

By fetching raw Twitter data from Twitter API, that is, Tweepy which is a reusable library function to pull data from open source and NLTK is applied to manipulate and preprocess the textual data. Then for methods like stemming, lemmatization for preparation of the data to be fed to machine learning model and also stop words are imported. It gives a list of stop words that need to be removed from the textual data that have imported in porter stemmer. The method like TF–IDF vectorizer is converting the tokens into numbers so that the data can now be encoded from text to numbers. Then, the data are visualized to search the positive and negative words where positive word prioritization with fonts, that is, the mostly used positive words are having larger fonts and depressive word prioritization with fonts, that is, the mostly used depressive words are having larger fonts which are discussed in data visualization section.

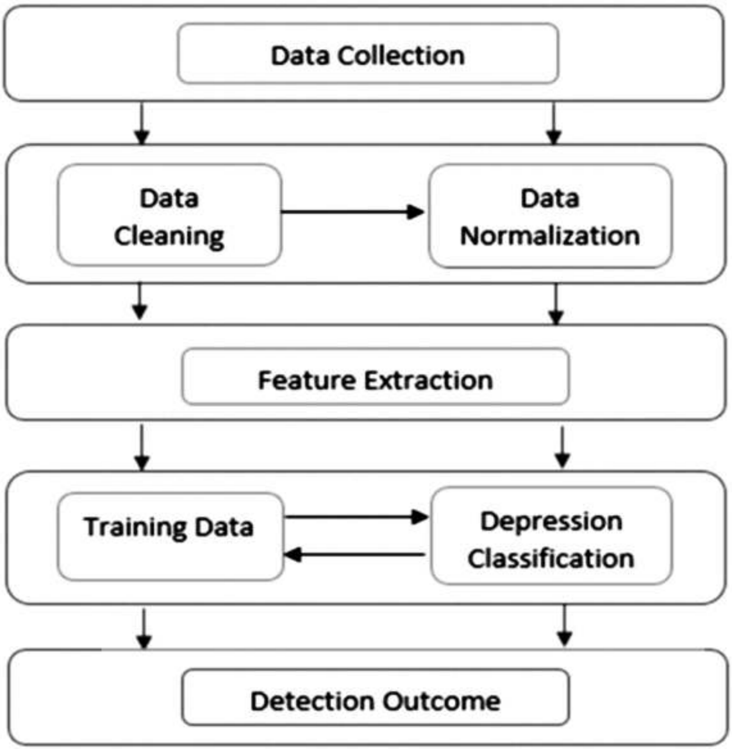


FIGURE 3.2 Workflow for prediction of suicidal tendency.

3.3.1 IMPORTING TWITTER DATA

Google Colab is a platform where a programmer can write and execute arbitrary Python code through the browser and is especially well suited to machine learning, data analysis, and execution. Here the data are pulled through API's from Tweepy as shown in Figure 3.3. Tweepy is reusable library function to pull data from open source.

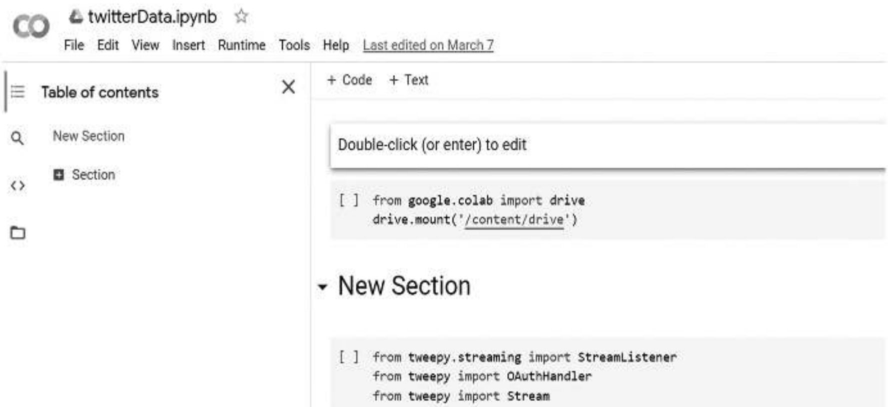


FIGURE 3.3 Importing Twitter data into the colab platform.

3.3.2 DATA EXTRACTION

NLTK is imported to download the latest version of natural language toolkit. Again, Panda library is also imported in order to fetch the raw data to rows and column in Jason format. Here, text and Id in rows and column are applied to preprocess the data as demonstrated in Figure 3.4.

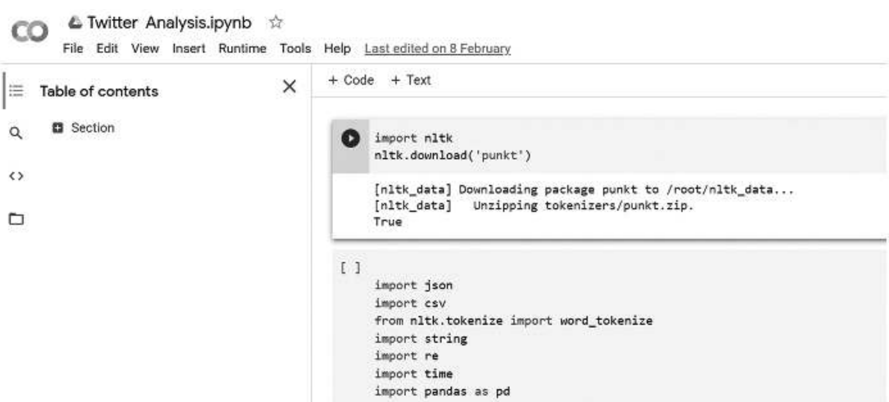


FIGURE 3.4 Importing requisite library.

3.3.3 DATA CLEANSING

This process removes the unnecessary parameters to build the appropriate model. The parameters include removal of stop words using the stop word module, removal of the column, the names, the Twitter handler, the links, punctuation and numbers, numeric removal, normalization (removing suffixes and prefixes), and array representation of target values (Figure 3.5).

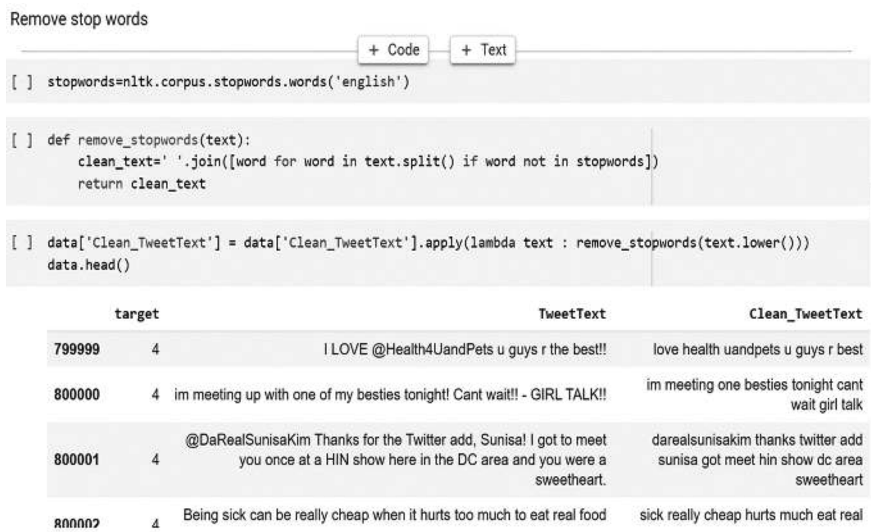


FIGURE 3.5 Stop words removal process.

3.3.4 DATA VISUALIZATION

Data visualization with max font is the most common words in the datasets for positive and depressive tweets with maximum no of usages.

Here, TF–IDF vectorizer is used for converting the tokens into numbers so that the data can now be encoded from text to number because the machine learning model actually work on numbers not text. Figures 3.6 and 3.7 represent the training of all the machine learning models to check the most accurate analysis for this work.

3.3.5 MODEL BUILDING

Here, four machine-learning classifiers algorithms are implemented like XGB classifier, random forest, logistic regression, and support vector machine to obtain the accuracy score as shown in Figure 3.8.

```
positive_words =' '.join([text for text in data['Clean_TweetText'][data['target'] == 4]])
wordcloud = WordCloud(width=800, height=500, random_state=21, max_font_size=110).generate(positive_words)

plt.figure(figsize=(10, 7))
plt.imshow(wordcloud, interpolation="bilinear")
plt.axis('off')
plt.show()
```

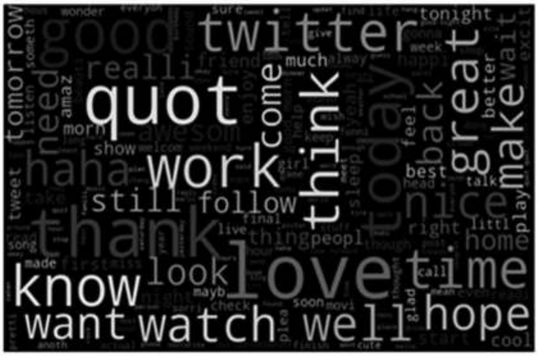


FIGURE 3.6 Positive words having larger fonts.

```
depressive_words =' '.join([text for text in data['Clean_TweetText'][data['target'] == 0]])
wordcloud = WordCloud(width=800, height=500, random_state=21, max_font_size=110).generate(depressive_words)

plt.figure(figsize=(10, 7))
plt.imshow(wordcloud, interpolation="bilinear")
plt.axis('off')
plt.show()
```

FIGURE 3.7 Depressive words having larger fonts.

The accuracy score of all four types of algorithms is displayed in Figures 3.9–3.12. The scores are 0.98193 for XGB, 0.95635 for random forest classifier, 0.96289 for logistic regression, and 0.96193 for SVM.

```
[ ] from xgboost import XGBClassifier
    import xgboost as xgb
    from sklearn.ensemble import RandomForestClassifier
    from sklearn.linear_model import LogisticRegression
    from sklearn import svm
    from sklearn.metrics import accuracy_score
    from sklearn.model_selection import train_test_split
    from sklearn.feature_extraction.text import CountVectorizer
```

```
[ ] count_vectorizer = CountVectorizer(stop_words='english')
    vectorizer = CountVectorizer()
    # tokenize and build vocab
    vectorizer.fit(data['Clean_TweetText'])
```

FIGURE 3.8 Classifiers like XGN, RF, LR, and SVM are used to learn the accuracy score.

```
xgbc = XGBClassifier(max_depth=6, n_estimators=1000, nthread= 3)
xgbc.fit(X_train,y_train)
prediction_xgb = xgbc.predict(X_test)
print(accuracy_score(prediction_xgb,y_test))

0.9619230769230769
```

FIGURE 3.9 Application of XGB classifier.

```
rf = RandomForestClassifier(n_estimators=10, random_state=2)
rf.fit(X_train,y_train)
prediction_rf = rf.predict(X_test)
print(accuracy_score(prediction_rf,y_test))

0.9563461538461538
```

FIGURE 3.10 Application of random forest classifier.

After training all the models, the Matplotlib library function is used for graphical representation of the classifiers to show the best accurate figures. As shown in Figure 3.13, it is concluded that the accuracy lies in the logistic regression with higher percentage accuracy.

```
lr = LogisticRegression()
lr.fit(X_train,y_train)
prediction_lr = lr.predict(X_test)
print(accuracy_score(prediction_lr,y_test))

0.9628846153846153
```

FIGURE 3.11 Application of logistic regression.

```
svc = svm.SVC()
svc.fit(X_train,y_train)
prediction_svc = svc.predict(X_test)
print(accuracy_score(prediction_svc,y_test))

0.9619230769230769
```

FIGURE 3.12 Application of support vector machine.

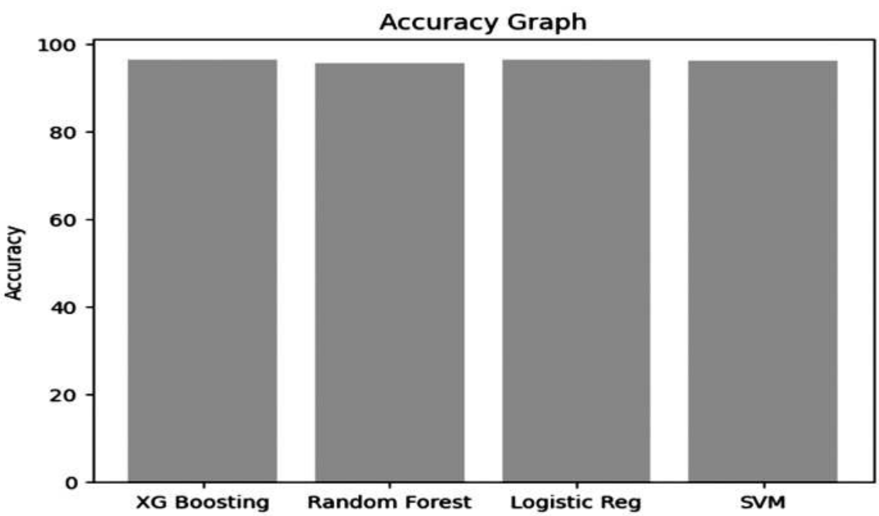


FIGURE 3.13 Accuracy representation of all the ML models.

Finally, a zip folder is created with the CSV content where all tweets, unzipping the suicide folder into Python, now final local drive is used for universal plug-and-play event for accepting incoming connections. Then, a test case is generated as shown in Figure 3.14 to predict the suicidal tendency.

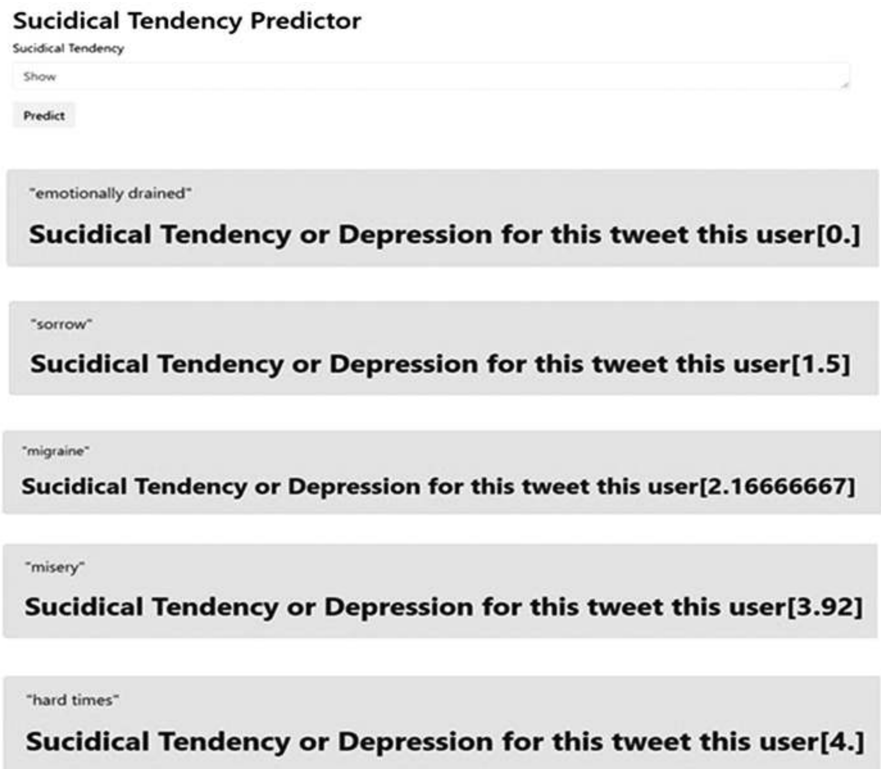


FIGURE 3.14 A Test case of suicide prediction.

3.4 CONCLUSION

In this work, five concurrent stages are initiated such as data collection, data preprocessing, training, data testing, and evaluation to predict the suicidal tendency of a Twitter handler. Commencing with data collection, the data are pulled from Tweepy. This eventually helped in fetching the raw data which is processed and preprocessed in a manner for machine learning usage hence called the vector format. In the preprocess, stop word is removed and lemmatization is performed to extract positive, negative, and neutral outcomes. Then the end result parameters are inserted into the predictor analyzer. While doing the prediction, baselining is done with criteria [0 < 2 < 4] where "0" denotes specifically to the negative side and "4" inclining toward the positivity.

KEYWORDS

- **depression**
- **social media**
- **neural network**
- **NLTK**
- **machine learning**

REFERENCES

1. Khan, M. R. H.; Afroz, U. S.; Masum, A. K. M.; Abujar; S.; Hossain S. A. H. In *Sentiment Analysis from Bengali Depression Dataset using Machine Learning*, 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), IEEE, Kharagpur, India, 2020.
2. Yong, C.; Bin, Z.; Weina, Z.; Wenjie, G.; Guangfu, S. In *Sentiment Analysis Based on Deep Learning and Its Application in Screening for Perinatal Depression*, Third International Conference on Data Science in Cyberspace (DSC), IEEE, Guangzhou, China, 2018.
3. Deshpande, M.; Rao, V. In *Depression Detection using Emotion Artificial Intelligence*, International Conference on Intelligent Sustainable Systems (ICISS), IEEE, Palladam, India, 2017.
4. Wang, X.; Zhang, C.; Ji, Y.; Sun, Li.; Wu, L.; Bao, Z.; A Depression Detection Model Based on Sentiment Analysis in Micro-blog Social Network. *Trends and Applications in Knowledge Discovery and Data Mining, Springer* **2013,** 201–213.
5. Birjali, M.; Kasri, M.; Hssane, A. B.; A Comprehensive Survey on Sentiment Analysis: Approaches, Challenges and Trends. *Knowledge-Based Syst.* **2021,** *226,* 107134.
6. Yao, Z.;, Yang, J.; Liu, J.; Keith, M.; Guan, C. H.; Comparing Tweet Sentiments in Megacities using Machine Learning Techniques: In the Midst of COVID-19. *Cities* **2021,** *116,* 103273.
7. Colón-Ruiz, C.; Segura-Bedmar, I. Comparing Deep Learning Architectures for Sentiment Analysis on Drug Reviews. *J. Biomed. Inform.* **2020,** *110,* 103539.
8. Khan, M. R. H.; Afroz, Umme, S. A.; Masum, A. K. M.; Abujar, S.; Hossain, S. A. In *Sentiment Analysis from Bengali Depression Dataset Using Machine Learning*, 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), *IEEE,* 2020, pp. 1-5.

9. He, L.; Cao, C. Automated Depression Analysis Using Convolutional Neural Networks from Speech. *J. Biomed. Inform*. **2018,** *83,* 103-111.
10. Moon, N.; Mariam, N. A.; Sharmin, S.; Islam, M. M.; Nur, F. N.; Debnath, N.; Machine Learning Approach to Predict the Depression in Job Sectors in Bangladesh. *Curr. Res. Behav. Sci*. **2021,** *2,* 100058.
11. Babu, N. V.; Kanaga, E. Sentiment Analysis in Social Media Data for Depression Detection Using Artificial Intelligence. A Review. *SN Comput. Sci.* 3, **2022,** *1,* 1–20.
12. Nijhawan, T.; Attigeri, G.; Ananthakrishna, T. Stress Detection Using Natural Language Processing and Machine Learning Over Social Interactions. *J. Big Data* **2022,** *9* (1), 1–24.
13. Angskun, J.; Tipprasert, S.; and Angskun, T. Big Data Analytics on Social Networks for Real-time Depression Detection. *J. Big Data* **2022,** *9,* 1–15.
14. Tanna, D.; Dudhane, M.; Sardar, A.; Deshpande, K.; Deshmukh, N. In *Sentiment Analysis on Social Media for Emotion Classification*, International Conference on Intelligent Computing and Control Systems (ICICCS 2020), pp. 911–915.
15. Nkomo, L. M.; Ndukwe, I. G.; Daniel, B. K. Social Network and Sentiment Analysis: Investigation of Students Perspectives on Lecture Recording. *IEEE Access* **2020**, *8,* 228693–228701.
16. Kundale, J. U.; Kulkarni, N. J. In *Language Independent Multi-class Sentiment Analysis*, 5th International Conference on Computing Communication Control and Automation (ICCUBEA), 2019, pp.1–7.
17. Arora, P. In *Mining Twitter Data for Depression Detection*, IEEE International Conference on Signal Processing and Communication (ICSC), 2019, pp. 186–189.
18. Bouzazi, M.; Ohtsuki, T. A Pattern-based Approach for Multi-class Sentiment Analysis in Twitter. *IEEE Access* **2017,** *5,* 20617–20639.
19. Rosa, R. L.; Ruggiero, W. V.; Rodrıguez, D. Z. A Knowledge- based Recommendation System that Includes Sentiment Analysis and Deep Learning. *IEEE Trans. Ind. Inf.* **2019,** *15,* 2124–2135.
20. Imran, A. S.; Daudpota, S. M.; Kastrati, Z.; Bharat, R. Cross-cultural Polarity and Emotion Detection Using Sentiment Analysis and Deep Learning on COVID-19 Related Tweets. *IEEE Access* **2020,** *8,* 181074–181090.
21. Studiawan, H.; Sohel, F.; Payne, C. Sentiment Analysis in a Forensic Timeline with Deep Learning. *IEEE Access* **2020,** *8,* 60664–60675.
22. Tariq, S.; Akhtar, N.; Afzal, H.; Khalid, Khalid, S.; Mufti, M. R.; Hussain, S.; Habib, A.; Ahmad, G. A Novel Co-training based Approach for the Classification of Mental Illnesses using Social Media Posts. *IEEE Access* **2019,** *7,* 166165–166172.
23. Baheti, R. R.; Kinariwala, S. Detection and Analysis of Stress Using Machine Learning Techniques. *Int. J. Eng. Adv. Technol. (IJEAT)* **2019,** *9,* (ISSN: 2249–8958).
24. Ruz, G. A.; Henriquez, P. A.; Mascareno, A. Sentiment Analysis of Twitter Data During Critical Events Through Bayesian Networks Classifiers. *Future Gener. Comput. Syst.* **2020,** *106,* 92–104.
25. Cheng, L. C.; Tsai, S. L. In *Deep Learning for Automated Sentiment Analysis of Social Media*, IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, 2019, pp. 1001–1004.
26. Rao, G.; Zhang, Y.; Zhang, L.; Cong, Q.; Feng, Z. MGL-CNN: A Hierarchical Posts Representations Model for Identifying Depressed Individuals in Online Forums. *IEEE Access* **2020,** *8,* 32395–32403.

27. Sethi, M.; Pande, S., Trar, P., Soni, P. In *Sentiment Identification in COVID-19 Specific Tweets*, International Conference on Electronics and Sustainable Communication Systems (*ICESC 2020*), pp. 509–516.
28. Tadessi, M. M.; Lin, H.; Xu, B.; Yang, L. Detection of Depression Related Posts in Reddit Social Media Forum. *IEEE Access* **2019,** *7,* 44883–44893.
29. Kumar, A.; Sharma, A.; Arora, A. In *Anxious Depression Prediction in Real-time Social Data*, International Conference on Advanced Engineering, Science, Management and Technology— 2019 (ICAESMT19).
30. Ahmad, S.; Asghar, M. Z; Alotaibi, F. M.; Awan, I. Detection and Classification of Social Media-based Extremist Affiliations Using Sentiment Analysis Techniques. *Human Centric. Comput. Inf. Sci*. **2019,** *24,* 1–23.
31. Gaikwad, G.; Joshi, D. In *Multiclass Mood Classification on Twitter Using Lexicon Dictionary and Machine Learning Algorithms*, International Conference on Inventive Computation Technologies (*ICICT*), 2016, pp. 1–6.
32. Solakidis, G. S.; Vavliakis, K. N.; Mitkas, P. A. In *Multilingual Sentiment Analysis Using Emoticons and Keywords*, IEEE/WIC/ACM International Joint Conferences on web Intelligence (WI) and Intelligent Agent Technologies (IAT), 2014, pp. 102–109.
33. Chatterjee, R.; Gupta, R. K.; Gupta, B.; Depression Detection from Social Media Posts Using Multinomial Naive Theorem. *IOP Conf. Ser. Mater. Sci. Eng*. **2021,** *1022,* 012095.
34. Jose, R.; Chooralil, V. S. In *Prediction of Election Result by Enhanced Sentiment Analysis on Twitter data Using Classifier Ensemble Approach*, International Conference on data Mining and Advanced Computing (SAPIENCE), 2016, pp. 64–67.
35. Chen, B; Cheng, L.; Chen, R.; Huang, Q.; Chen, Y. P. In *Deep Neural Networks for Multiclass Sentiment Classification*, IEEE 20th International Conference on High Performance Computing and Communications, IEEE 16th International Conference on Smart City, IEEE 4th International Conference on Data Science and Systems, 2018, pp. 854–859.
36. Katchapakirin, K.; Wongpatikaseree, K.; Yomaboot, P.; Kaewpitakkun, Y. In *Facebook Social Media for Depression Detection in the Thai community*, 15th International Joint Conference on Computer Science and Software Engineering (JCSSE), 2018, pp. 1–6.

CHAPTER 4

BRAIN TUMOR DETECTION USING DEEP CONVOLUTION NEURAL NETWORKS

BHASKAR ADEPU and P. SHIVANI

Department of Information Technology, Kakatiya Institute of Technology & Science, Warangal, Telangana, India

ABSTRACT

A brain tumor may be a growth of abnormal cells in our brain. Benign and malignant are the two types of brain tumors. Early recognition and characterization of brain tumors is a significant exploration space in the area of clinical imaging and as necessary aids in picking the majority of invaluable therapy ways to save patients' lives. Our objective in this paper is to create multicharacterization of brain tumors for the early conclusion aspiration utilizing deep convolutional neural networks (CNNs). Our contributions in this paper are (i) identification of brain tumor from input magnetic resonance imaging (MRI) pictures, that is, having a tumor or its normal, (ii) outline of brain tumor type, that is, glioma, meningioma, and the (iii) classification of the brain tumor grades as grade 2, grade 3, and grade 4. Our experimentation results show that by using CNNs we achieved about 90% accuracy in the prediction and classification of tumors.

4.1 INTRODUCTION

4.1.1 CANCER AND ITS IMPACT

This disease is perhaps the most well-known reason for death around the world. Cancer growth is a sickness that lamentably spreads in cells and filling

Applied Soft Computing Techniques: Theoretical Principles and Practical Applications. Samarjeet Borah, Ratna Raja Kumar Jambi, Sharifah Sakinah Syed Ahmed, & Mahendra Prabhakar Deore (Eds.)
© 2025 Apple Academic Press, Inc. Co-published with CRC Press (Taylor & Francis)

day by day in the body.[1,2] According to the WHO, the disease is the subsequent driving reason for death internationally, and in 2018, it caused roughly 9.6 million deaths. There are numerous kinds of diseases, for example, brain tumor growth, cellular breakdown in the lungs, and prostate disease. BC is a far-reaching infection that pounds the presence of women conventionally in the age social event of 25–50.[3] With the conceivable climb in the amount of BC cases in India, the agony coming to is upsetting.

4.1.2 BRAIN TUMOR

A brain tumor growth happens when unusual cells structure inside the mind as in Figure 4.1. There are two principal kinds of growths: dangerous (harmful) growths and harmless (nondestructive) cancers. These can be additionally delegated essential growths, what start inside the mind, and optional cancers, which are most commonly spread from growths situated external the cerebrum, known as cerebrum metastasis growths.[4]

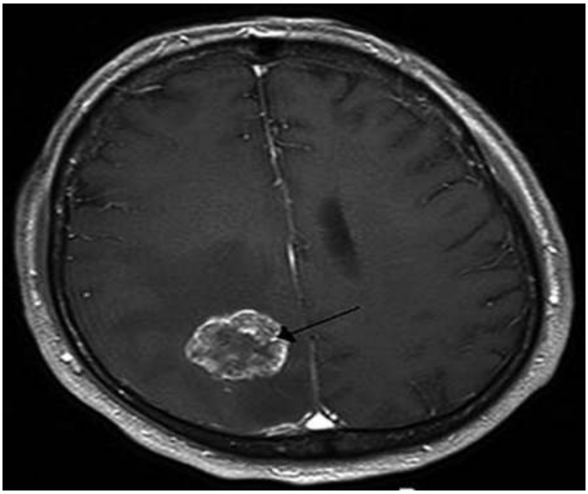


FIGURE 4.1 Brain tumor.

4.1.3 OBJECTIVE

The objective is to distinguish whether a tumor is favorable or of threatening in nature, as harmful tumors are dangerous and ought to be treated straightaway to diminish and forestall further inconveniences. So, it is a 2-fold order issue and can be settled by different artificial intelligence (AI) techniques.[5] It has appeared in the past that AI calculations perform better compared to human pathologists. A greater part of researchers have tracked down that clinical picture preparation utilizing AI gives more precise outcomes when contrasted with the target determination given by a pathologist.[11] An examination in Europe has been directed by Phillips in which a bunch of calculations alongside bosom pictures gave more exact identification. This finding is likewise proof that utilizing high-goal pictures and better calculations will improve the exhibition and precision of disease discovery.[10,13]

4.2 RELATED WORK

It is quick to evaluate cerebrum growth recreation plans with one variety of pics and datasets.[15] Various methodologies were dispersed and stepped forward to perceive and describe cerebrum cancer illness through the usage of big neural agencies with extraordinary fashions. The irksome task of clinical photograph evaluation and portrayal is to orchestrate histopathological pics reliant upon the improvement of the cells and complicated morphology and surface.[3] Various advanced systems have been proposed to determine the troublesome problem of photograph request, as an instance, substantial mastering fashions and preprepared significant neural groups.[6,7]

The brain is the maximum complex part of our head. Intracranial neoplasm or brain tumor is the extraordinary growth of cells within the mind. The symptoms of a tumor can be common complications and migraines.[9] Over the years it can even result in imaginative and prescient loss. At this second science is scarce approximately the origins and elements leading to this bizarre increase. Tumors are categorized on two bases: whether or not they are cancerous or not and their area of origin. These are without difficulty distinguishable and feature a sluggish growth charge.[10] Cancerous Tumors are called malignant. These are very competitive and may be existence threatening as these are hard to detect. When it involves detecting a tumor, doctors can opt for either an X-ray or magnetic resonance imaging (MRI).[11] All other tests fail to offer adequate information; hence MRI examinations are acceptable. The residences of magnetism and radio waves are used in an MRI exam to generate accurate images. MRIs are commonly prescribed by neurosurgeons because they provide them with enough information to detect even the tiniest anomalies.[12] However, as MRI makes use of magnetic waves, so it is far improper for sufferers with pacemakers and metallic implants. Now once we have the scanned image of the brain, it is essential to appropriately locate the tumor, its size, and its place. All these records are vital for the neurosurgeon to complete his analysis. This is wherein computerized image processing segmentation strategies and feature extraction approach; we can correctly stumble on the tumor.[4,6] Comes to assist. With the set of different in this task, the proposed gadget divides the entered picture into several slices and preprocessing takes area in parallel. This software program also runs on multicore surroundings for processing and extraction of each and every picture slice one at a time. We have investigated the exclusive entropy capabilities for tumor segmentation and its detection from diverse MRI images.[4] The extraordinary threshold values are obtained depending on the specific type of entropy. These threshold values are dependent on the exceptional entropy characteristic which in turn impacts the segmented effects. The segmented consequences depend upon the Shannon and Non-Shannon conduct at distinct examples of parametric selections.[8] The texture evaluation of medical pics is also performed which will get a higher accuracy. The satisfactory result is obtained from Havrda Charvat entropy this is higher than the alternative entropy functions utilized in the feel of detecting tumors and a good way to assist in advanced detection of the tumors and could provide a remedy to cure.[2]

4.3 EXISTING SYSTEM

Brain tumors category has been carried out the use of many gadget learning procedures and imaging modalities after some time. A system to categorize different grades of glioma without employing a binary classification for the excessive and coffee grades was introduced by Zacharaki et al. in 2009.[16] They used support vector machines (SVMs) and k-nearest neighbor (KNN). For multicategory, an exactness of 65% is obtained, and for binary type, 68%. Elhamzi et al.[3] developed a method to classify brain tumors' normal and abnormal images using discrete wavelet transformations (DWT) to separate capabilities, principal component analysis (PCA) to reduce elements, and then artificial neural networks (ANN) and KNN to organize images with an average precision of 77 and 68%, respectively. Cheng et al.[17] in 2015, developed a strategy to improve the overall classification of brain tumors by enhancing the tumor site through photo enlargement, followed by splitting into subregions. They extracted features using the depth histogram, the gray level co-prevalence matrix, and the bag of words (BOW) techniques and in the end accomplished a nice accuracy of 91.28% via the use of ring structure partition further to tumor area augmentation.

Brain tumor type the detection of system mastering strategies has recently been examined by scientists mainly throughout the last few years. The improvement in synthetic observation and deep mastering-based latest technology achieves a splendid effect inside the discipline of clinical photo review, specifically in relation to the issue of sickness analysis. Parallel to this, various studies on brain tumor prediction and multiclass classification using convolutional neural networks (CNN) were monitored. This phase is devoted to a multiclass brain tumor literature assessment using CNN. The studies in the literature may be found in a variety of formats. With 96.56% accuracy, their suggested version successfully classified brain tumors as meningioma, glioma, and pituitary. In any other examination, Mzoughi introduced a profound multiscale 3-D CNN version for brain tumor grading from volumetric 3-D MRI snapshots.

4.4 PROPOSED SYSTEM

In this work, we show the fundamental commitments and our proposed strategy utilized in the characterization task of brain tumor growth utilizing histopathological picture grouping. There have been various investigations on utilization of profound figuring out how to the order of bosom malignant growth from histopathological pictures. In our work, we utilized a freely accessible database which is called as Brain Tumor Classification (MRI). Our structure utilizes the accompanying advances as shown in Figure 4.2, in any case, to get the informational index and to move toward the minuscule biopsy pictures, after prehandling, we jumped the information haphazardly on 30% for the testing set and 70% for the training set. Then, we used information increase for the preparation dataset as it were. The increase included picture tasks, for example, resizing, turning, and reflecting pictures. After preparing the boundaries of the organization it orders the test set and execution measurements such as characterization precision and disarray lattices are figured. In the following section, we present the outcomes utilizing diverse assessment measurements.

In this work, as referenced earlier, we utilized the Brain Tumor Classification (MRI) Database dependent on picture level, which has two principal classes: benign and malignant.

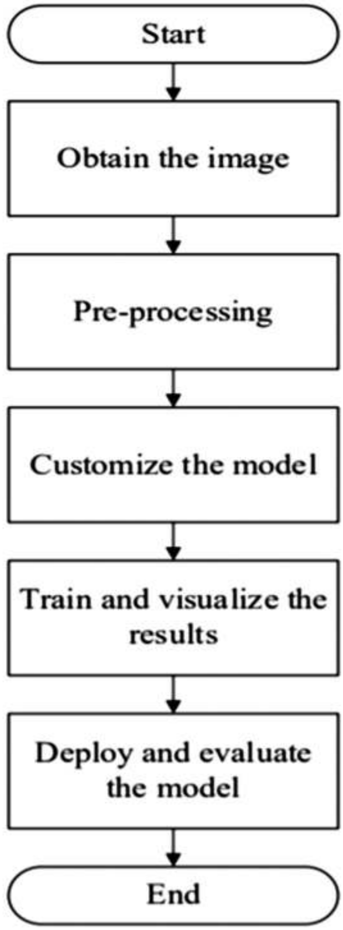


FIGURE 4.2 Brain tumor detection framework overview.

4.4.1 CONVOLUTIONAL NEURAL NETWORK FOR BRAIN TUMOR IMAGE CLASSIFICATION

The dataset was taken care of as a contribution to the CNN in application to the bosom disease order. After taking care of the information, we prepared the profound convolutional parts in the proposed design of CNN. We utilized Leaky ReLU nonlinearity for the convolutional layers. It can be characterized as

$$F(X) = \begin{cases} X, & if\ X > 0; \\ aX, & otherwise \end{cases} \tag{4.1}$$

After all is done, the CNN layer is communicated as

$$yj = f(bj) + \Sigma\, k\, ij + xi \tag{4.2}$$

As shown in the above equation, xi and yj are tending to the ith input guide and jth yield map, separately; bj addresses the inclination boundary of jth map; and bij shows the convolutional part utilized between ith and jth maps. After that, the max pooling layer (MPL) is used. In the MPL, each neuron participates in the yield map yi pools over s × s noncovering area in the data map, that is, xi. By and large, the MPL is portrayed as

$$yj\ i = max0 \leq m \leq s\ \{xj.\ s + mi\} \tag{4.3}$$

The SoftMax classifier has a direct relationship with the two layers. This classifier includes an equivalent number of yield classes to the yields. In our proposed design, we utilized hyperbolic tangent as a nonstraight convention in the availability of a multitude of layers. SoftMax works as an element of crushing, and K-dimensional info vector is restandardized, yielding in the scope[1,2] of genuine qualities. Numerically it is very well appeared as:

$$(z)\, j = e\ zj\ \Sigma\ e\ K\ zk\ k = 1, \text{ where } j = 1 \text{ to K} \tag{4.4}$$

Two kinds of mistakes delivered during the cycle of advancement of the machine learning (ML) strategy. Preparing mistakes is prompted during the preparation period of neural organization, although the speculation blunder is created in our proposed classifier during the testing time. In profound learning, preparation is frequently influenced by the overfitting (OF) and or underfitting (UF) issues. To diminish OF and UF issues in our presented CNN design, we put in application group standardization after every layer. The dropout is presented to the primarily associated layers. The total proposed design for the tumor detection arrangement is seen in Figure 4.3.

A profound learning engineering for automized cerebrum growth discovery that joins ideas of AI and picture arrangement. We have depicted diverse deep neural network (DNN) structures, particularly those adjusted to picture information like CNNs.

This utilized the marked (harmless/dangerous) input picture from the crude pixels and featured the visual examples, and afterward, used those examples to recognize nonmalignant and disease-containing tissue, working similar to advanced staining, which spotlights picture portions vital for indicative choices, with the assistance of a classifier organization as shown in Figure 4.3. The CNN was prepared utilizing 2480 harmless and 5429 threatening examples having a place with the red, green, and blue (RGB) shading model. Therefore, the proposed system depicted in Figure 4.3

provides an effective classification model for classifying brain tissue as being either benign or malignant.

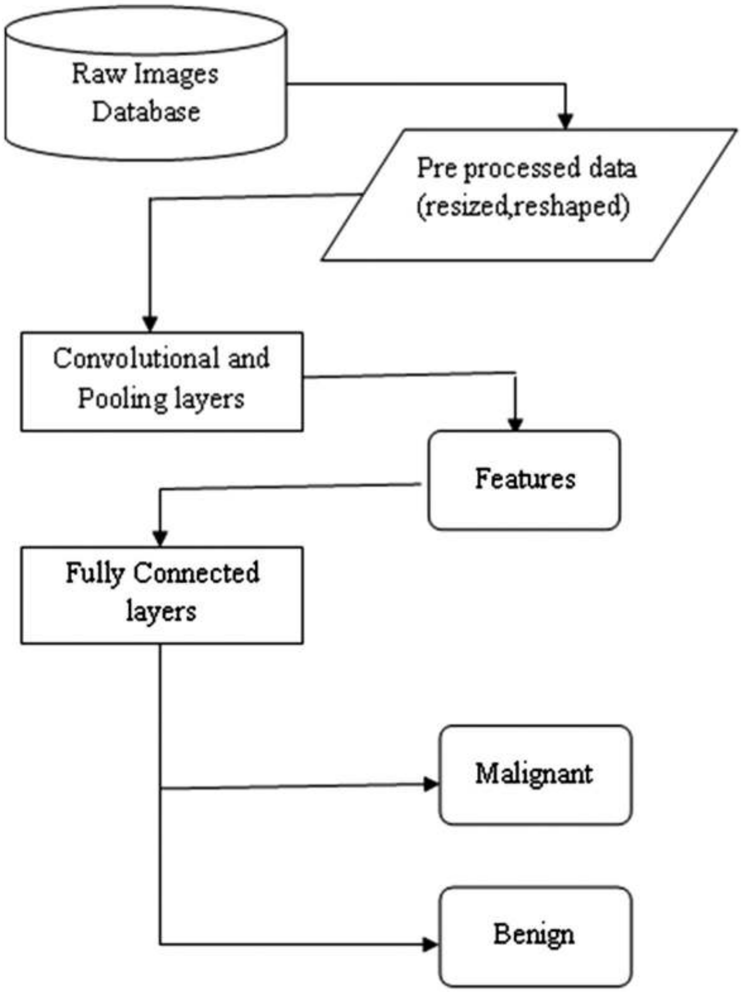


FIGURE 4.3 Data flow outline of the proposed method.

4.5 METHODOLOGY

Recently, the IT industry has witnessed a surge in the popularity of a certain set of skills called deep learning (DL). Calculations inspired by the operation of the human mind or brain, organizations fall within the DL, a subclass of the ML. These patterns are referred to as NN. It instructs the computer to carry out human actions naturally. Some of the models used in DL include ANN, autoencoders, recurrent neural network (RNN), and reinforcement learning. One model specifically, the CNN or ConvNets, has made a critical commitment to image and vision analysis.

CNNs, a subclass of DNN that is often employed for image analysis, are able to recognize and organize particular details in images. A few of their uses include image classification, computer vision, image and video recognition, image analysis for therapeutic reasons, and natural language processing. The word "convolution" in CNN represents the numerical capacity of convolution, which is an uncommon type of direct activity in which two capacities are raised to form a third capacity that communicates how the state of one capacity is altered by the other. In essence, two grids are multiplied to get a result that is used to eliminate highlights from the image.

4.5.1 CONVOLUTIONAL NEURAL NETWORK ARCHITECTURE

- A convolutional device separates and identify distinctive qualities of input picture for analysis and it is known as feature extraction.
- A fully integrated layer which uses the output of the convolutional layer to infer the class of image using the information gathered in earlier steps.

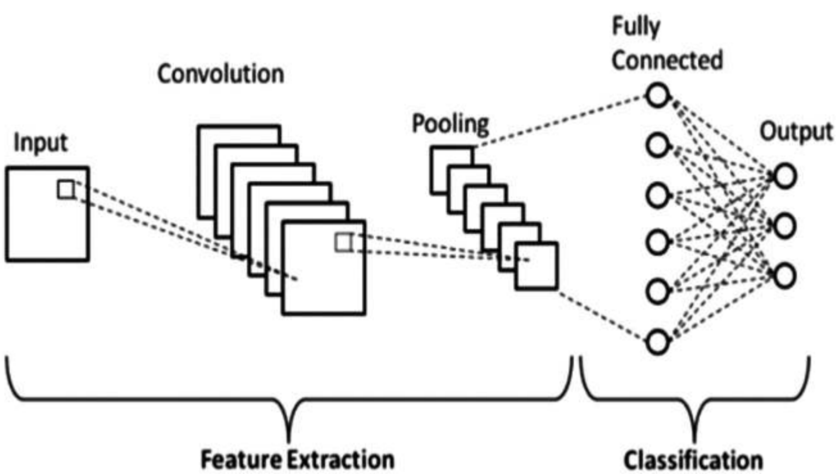


FIGURE 4.4 Convolution neural network architecture.

CNN consists of the following three layers:

1. Convolution layer (ConvLayer)
2. Pooling layer (PoolLayer)
3. Fully connected (FC) layer)

These layers together would result in a CNN framework, as seen in Figure 4.4. The activation function and dropout layer are two more crucial components that are described below. The specific layers in the CNN are displayed in Figure 4.5.

1. Convolution Layer

This layer is used for extracting the features from the input data.

Steps involved in the convolution layer:

Step 1: Align the element with the image correction (which by default uses a nine-pixel size).
Step 2: Multiply each image pixel by the highlight pixel for comparison.
Step 3: Add them all up.
Step 4: Divide by the complete number of pixels in the element.

2. Pooling Layer

The convolutional layer's attributes are still very significant. Whenever utilized straightforwardly, the preparation stage will be inclined to OF and will consume a large chunk of the day to finish. To resolve this issue, this layer utilizes down examining to pack the picture and diminish the quantity of boundaries. Normal pooling and Max pooling are two sorts of subsampling utilized in the writing. The element of component maps is brought down in the recommended model by utilizing the MaxPooling activity, which is easy to utilize and creates the best outcomes. In this layer, we shrivel the picture into a more modest size.Steps are listed as follows:

i. Pick a window size (normally 2 or 3).
ii. Pick a step (typically 2).
iii. Walk your window across your separated pictures.
iv. From every window, take the most extreme worth.

3. Fully Connected (FC) Layer

At the end, the FC layer is utilized. The target of this layer is to stuff the result of the former layer in light of the fact that the highlights should be one-layered (1D) information prior to preparing with the classifier. The result is fixed as the quantity of classes utilized when it is utilized as the last layer.

4. Dropout Layer

All the attributes are connected to the FC layer, the preparation dataset is prone to OF. OF is the process of a model performing so well at preparing information that it negatively impacts how it presents fresh information. To solve this problem, a dropout layer is used, which removes a few neurons from the CNN network during the preparation phase, leading to a more modest model. 30% of the hubs in the brain network exit arbitrarily after a dropout of 0.3.

5. Activation Function

Finally, one of the most extreme significant elements inside the CNN model is the initiation trademark. They are used to extract and imprecisely identify any kind of network variable-to-variable association, both simple and complex. During the basic phase of the network, it chooses which model data should fire ahead and which data should not. It gives the network nonlinearity.

The SoftMax, ReLU, Sigmoid, and tanH capacities are probably the more frequently used activation functions.

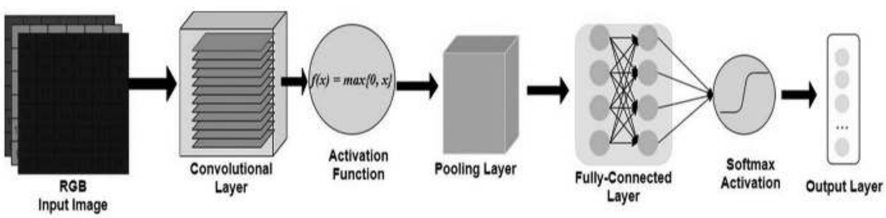


FIGURE 4.5 Layers used in the CNN architecture.

4.5.2 FLOWCHART

The suggested method's block design is shown in Figure 4.6, where our model starts by loading and extracting images and labels them using raw files in the dataset, then goes through preprocessing and augmentation steps after dividing the dataset into model train, test, and validate data sets. The proposed method is shown in Figure 4.6, followed by a description of the hyper-parameter setup, regularization strategies, and optimization algorithm. Lastly, computations for network performance and training are displayed.

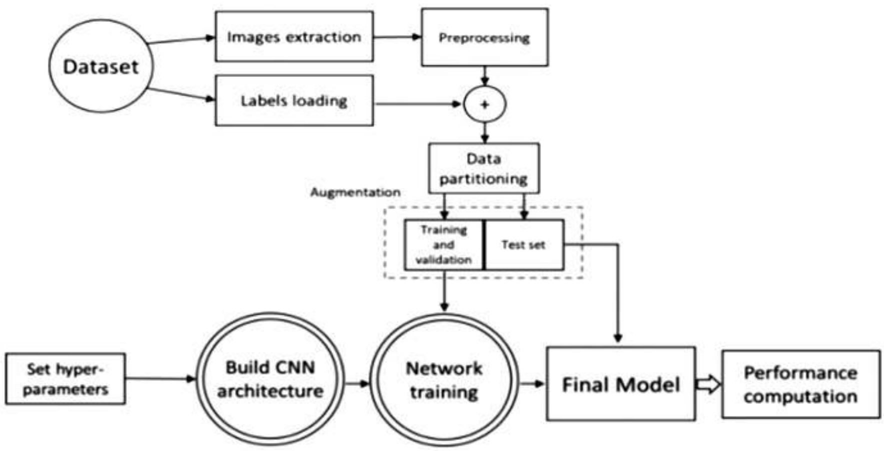


FIGURE 4.6 Proposed model flow diagram.

4.5.3 DATASET DESCRIPTION

In this postulation, we utilize two datasets that are publicly available, they are brain tumor multiclassification MRI dataset from the Kaggle and the repository of molecular brain neoplasia data (REMBRAND) from the Cancer Imaging Archive (CIA). The dataset Brain Tumur Multiclassification MRI has partitioned into two principal gatherings: benign tumors and Malignant tumors with various factors the photos are RGB pictures of objective 700 × 460 pixels in size with three channels. The REMBRAND dataset contains 130 patients with MR images with gliomas of grade 2, grade 3, and grade 4 as shown in Table 4.1.

For the tumor prediction task, 12,025 MR images were collected out of which there are 4500 are of benign tumors and 7525 images of malignant tumors. For the type of tumor prediction task, 7525 MR images are collected out of which 2147 are of glioma; 2582 of meningioma; and 2796 images are of pituitary tumor. For the grade prediction task, 4261 MR images are collected in which there are 1679 MR images of grade 2, 1234 MR images of grade 3, and 1679 MR images of grade 4.

TABLE 4.1 Number of Images in the Dataset.

Task type	Target value	Count of images in each category	Total no. of images
Tumor prediction	Benign	4500	12025
	Malignant	7525	
Tumor-type prediction	Glioma	2147	
	Meningioma	2582	7525
	Pituitary	2796	
Grade prediction	Grade 2	1679	
	Grade 3	1234	4261
	Grade 4	1348	

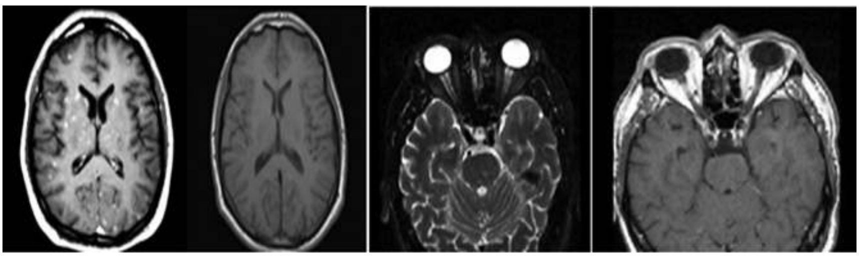


FIGURE 4.7 Images of benign.

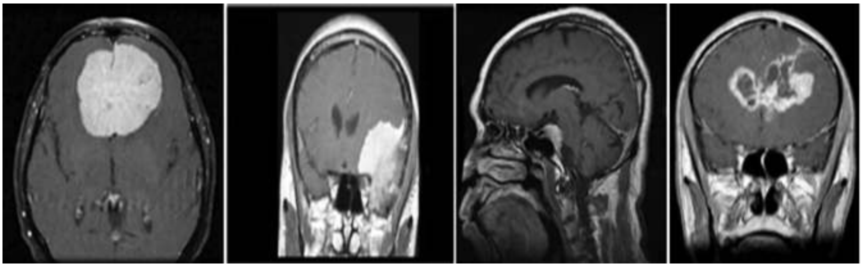


FIGURE 4.8 Images of malignant.

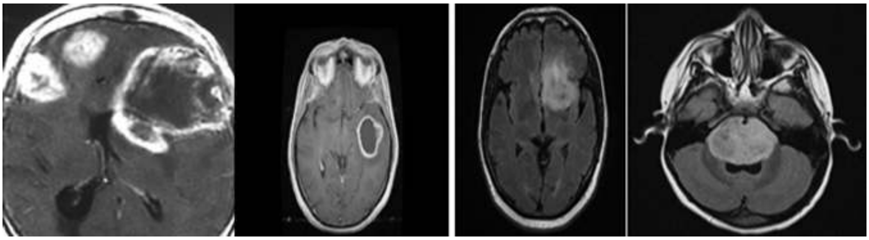


FIGURE 4.9 Images of glioma.

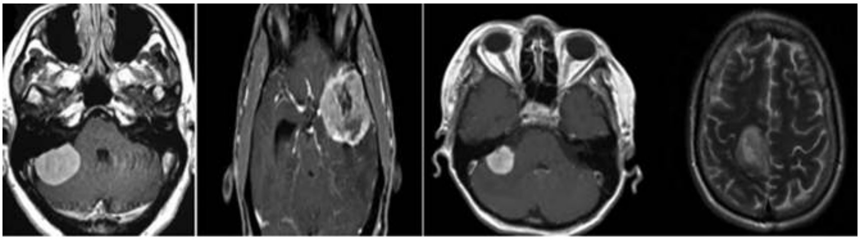


FIGURE 4.10 Images of meningioma.

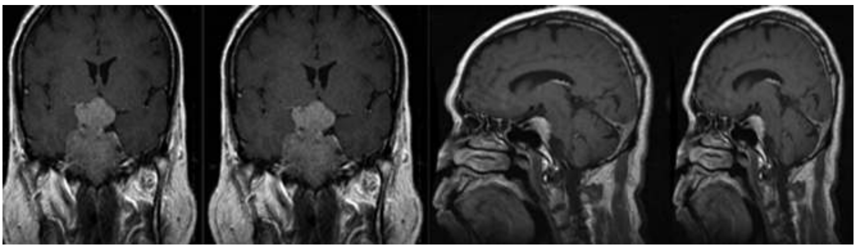


FIGURE 4.11 Images of pituitary.

4.5.4 CNN MODEL

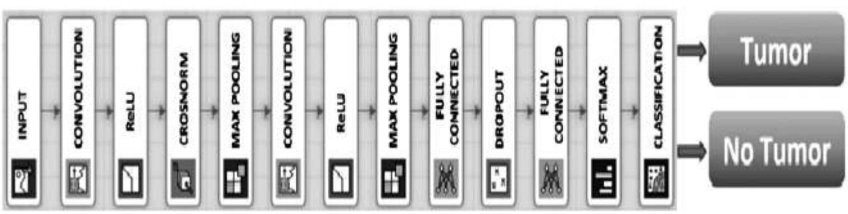


FIGURE 4.12 Architecture of CNN model for tumor prediction.

4.5.5 EVALUATION METRICS

- True Positive (T_P): It is an observation that is or is anticipated to be favorable.
- False Negative (F_N): Although a positive observation was made, a negative result was anticipated.
- True Negative (T_N): It is an observation that is unfavorable, and an unfavorable result is anticipated.
- False Positive (F_P): Although the observation was negative, a positive outcome was anticipated.

TABLE 4.2 CNN Architecture Details Used in Tumor Prediction.

Type of layer	Shape of output	Number of parameters
Convn2D_17	[None, 148, 148, 32]	[896]
Act_Fn_20	[None, 148, 148, 32]	[0]
MaxPool2D_17	[None, 74, 74, 32]	[0]
Convn2D_18	[None, 74, 74, 32]	[9248]
Act_Fn_21	[None, 74, 74, 32]	[0]
Batch Normalization_12	[None, 74, 74, 32]	[128]
MaxPool2D_18	[None, 36, 36, 32]	[0]
Convn2D_19	[None, 34, 34, 64]	[18496]
Act_Fn_21	[None, 34, 34, 64]	[0]
Batch Normalization_12	[None, 34, 34, 64]	[256]
MaxPool2D_19	[None, 17, 17, 64]	[0]
Flatten_9	[None, 18496]	[0]
Dense_18	[None, 64]	[1183808]
Act_Fn_23	[None, 64]	[0]
Batch Normalization_12	[None, 64]	[256]
Dropout_4	[None, 64]	[0]
Dense_19	[None, 1]	[65]
Act_Fn_24	[None, 1]	[0]

Total parameters: 1,213,153.

Act_Fn: activation function; Convn2D: 2D convolution.

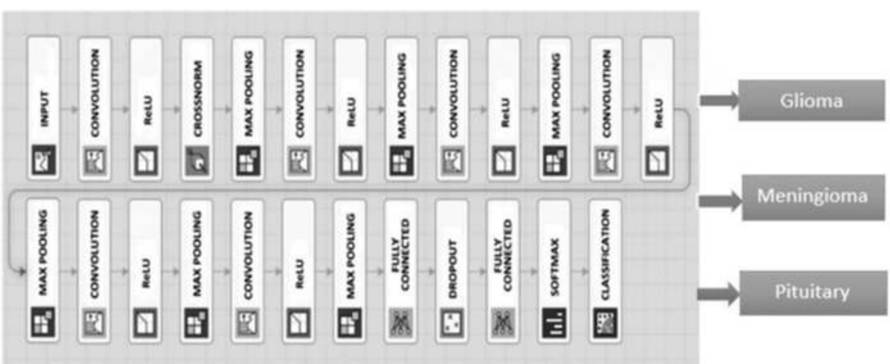


FIGURE 4.13 Architecture of CNN model for tumor-type prediction task.

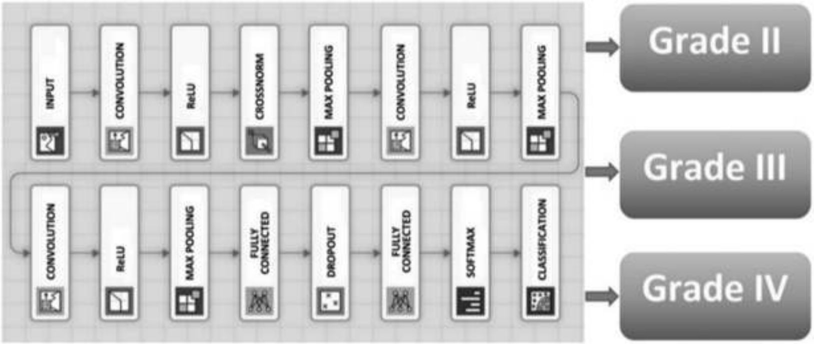


FIGURE 4.14 Architecture of CNN model for tumor grade prediction.

TABLE 4.3 CNN Architecture Details Used in Grade Prediction.

Type of layer	Shape of output	Number of parameters
Convn2D	[None, 148, 148, 32]	[896]
Act_Fn	[None, 148, 148, 32]	[0]
MaxPool2D	[None, 74, 74, 32]	[0]
Total parameters: 896		
Trainable parameters: 896		
Nontrainable parameters: 0		
Convn2D	[None, 148, 148, 32]	[896]
Act_Fn	[None, 148, 148, 32]	[0]
MaxPool2D	[None, 74, 74, 32]	[0]
Convn2D_1	[None, 74, 74, 32]	[9248]
Act_Fn_1	[None, 74, 74, 32]	[0]
Batch Normalization_1	[None, 74, 74, 32]	[128]
MaxPool2D_1	[None, 36, 36, 32]	[0]
Convn2D_2	[None, 34, 34, 64]	[18496]
Act_Fn_2	[None, 34, 34, 64]	[0]
Batch Normalization_2	[None, 34, 34, 64]	[256]
MaxPool2D_2	[None, 17, 17, 64]	[0]
Flatten_11	[None, 18496]	[0]
Dense_18	[None, 64]	[1183808]
Act_Fn_3	[None, 64]	[0]
Batch Normalization_	[None, 64]	[256]
Dropout_4	[None, 64]	[0]
Dense_19	[None, 1]	[65]
Act_Fn_3	[None, 1]	[0]
Total parameters: 1,213,153		
Trainable parameters: 1,212,833		
Nontrainable parameters: 320		
Act_Fn : Activation function		
Convn2D : 2D convolution		

TABLE 4.4 Contingency Matrix for Binary Classification and Multiclass Classification.

	PREDICTED classification				
	Classes	**C1**	**C2**	**C3**	**C4**
ACTUAL classification	**C1**	T_N	F_P	T_N	T_N
	C2	F_N	T_P	F_N	F_N
	C3	T_N	F_P	T_N	T_N
	C4	T_N	F_P	T_N	T_N

TABLE 4.5 Assessment Measurements.

Metrics used for evaluation	Equation
Accuracy	$\frac{TP+TN}{TP+FP+TN+FN}$
Precision, i.e., positive predictive value	$\frac{TP}{TP+FP}$
Sensitivity/recall, i.e., true positive rate	$\frac{TP}{TP+FN}$
Specificity, i.e., true negative rate	$\frac{TN}{FP+TN}$
F1 score	$\frac{2 * Sensitivity * Precision}{Sensitivity + Precision}$

4.6 RESULTS

TABLE 4.6 Learning Scheme of the CNN Model.

Classification task	Target values	Image count (in each group)	Images count	Training set (70%)	Testing set (30%)
Tumor prediction	Benign	4500	12025	8418	3607
	Malignant	7525			
Tumor-type prediction	Glioma	2147	7525	5268	2257
	Meningioma	2582			
	Pituitary	2796			
Grade prediction	Grade 2	1679	4261	2983	1278
	Grade 3	1234			
	Grade 4	1348			

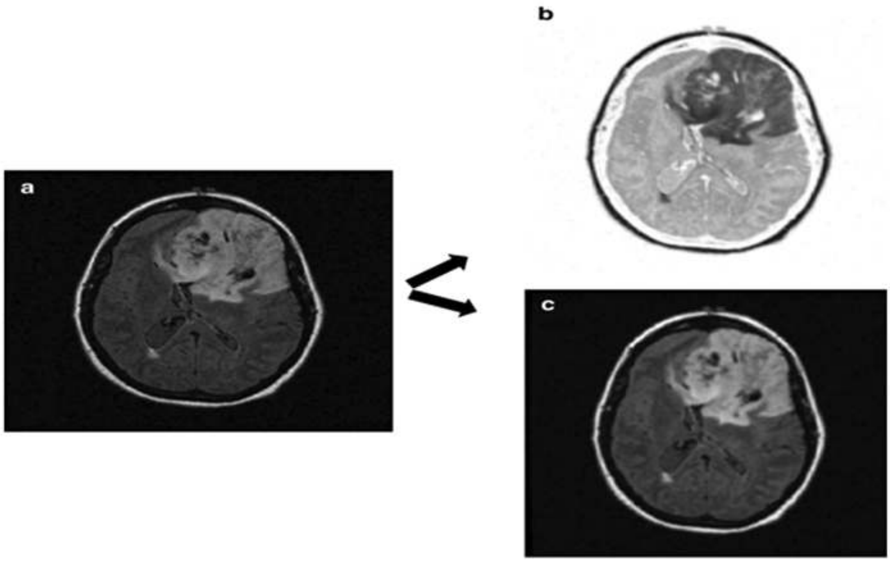


FIGURE 4.15 Resultant images.

Figure 4.15a is an MRI input, Figure 4.15b is the initiations in a particular channel, and Figure 4.15c is the most grounded enactment channel for the first convolution layer for classification I. Pixels of white in Figure 4.15c show solid actuations showing that this channel is firmly initiated at growth positions.

In recent years, picture categorization using CNNs has become increasingly popular in the identification of medical disorders. It is neither imaginable or practicable to develop a CNN model that is efficient and scalable be used in conjunction with other models to all categorization concerns should be resolved with satisfactory results. As a result, each kind of problem is given its own CNN model. The design and intricacy of the CNN model fluctuate in view of the sort of issue, inputs, and anticipated yields. Three distinct CNN models are used in this study for three different categorization goals. The first module is intended to identify brain tumors using MRI pictures as input. The 2nd module is used to predict the type of brain tumor and the third module is used to forecast grades of tumor. The first constructed CNN model achieves a highly satisfactory accuracy of 88.6% in detecting brain tumors.

Furthermore, the subdivision of brain tumor types is done with a 90.2% accuracy. Finally, the brain tumor grading was completed with 90.35%

TABLE 4.7 Performance Measures of the CNN Model.

Classification problem	Tumor groups	T_P	T_N	F_P	F_N	Accuracy (%)	Precision	Recall	F1 score	Support
Tumor prediction	Benign	1571	1229	334	26	88.6	0.82	0.98	0.90	1597
	Malignant	1229	1571	26	334	88.6	0.98	0.79	0.87	1563
Tumor type prediction	Glioma	490	1144	20	107	90.2	0.96	0.82	0.89	597
	Meningioma	559	1040	140	22	90.2	0.80	0.96	0.87	581
	Pituitary	534	1160	18	49	90.2	0.97	0.92	0.94	583
Grade prediction	Grade 2	322	653	31	52	90.35	0.91	0.86	0.89	374
	Grade 3	335	658	41	24	90.35	0.89	0.93	0.91	359
	Grade 4	299	703	30	26	90.35	0.91	0.92	0.91	325

accuracy. Evaluation metrics criteria such as the area under the receiver operating characteristic (ROC) curve, accuracy, specificity, sensitivity, and precision are used to assess or evaluate the suggested models' findings.

Loss is a function that computes the errors of our model. It evaluates how (or ineffectively) the model is doing. The loss will be enormous assuming the slip-ups are high, showing that the model is not performing great. In any case, the lower it is, the more our methodology works. A loss or cost function is used to calculate the loss. The model's purpose is to reduce the loss's value as much as possible. Accuracy is a lot easier to understand. To determine how effectively our model predicts, it compares model expectations to true attributes in rates.

The model performs significantly worse than expected in the vast majority of the data, with poor accuracy but substantial loss. The model produces small errors in the majority of the data, even though both accuracy and loss are low. Hitherto, if they are both high, it results in large errors in part of the data. The model will make slight errors on a tiny subset of the data if accuracy is high and loss is low, which is an optimal scenario.

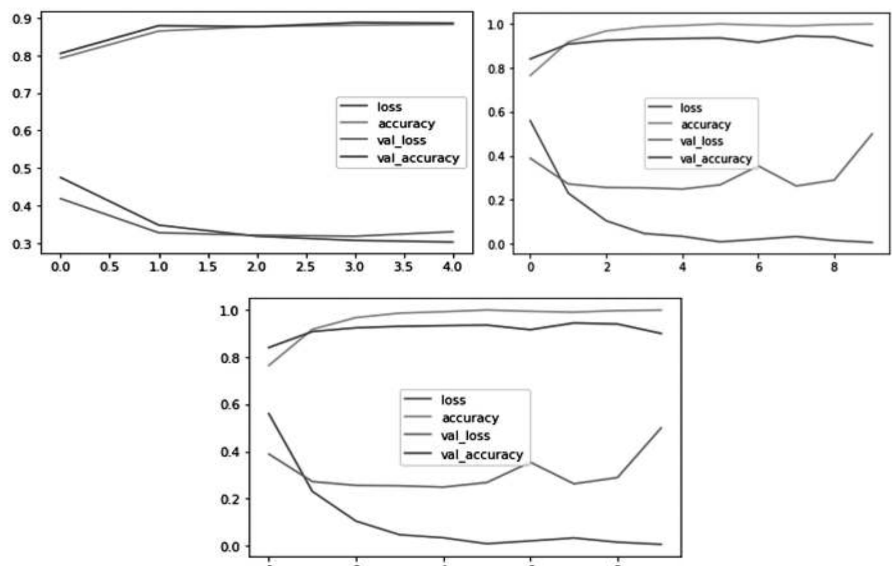


FIGURE 4.16 Accuracy and loss values for task 1, task 2, and task 3.

- There are two target classes: "benign" and "malignant," that is, "malignant" means they are cancer tumors, and "benign" means noncancerous tumors.
- The model predicted 1229 patients had the presence of tumor.
- Out of those 3160 cases, the model predicted correctly 1255 times as "malignant" and 1905 times as "benign."
- In reality, 1563 patients in the sample have the disease, and 1597 patients do not.

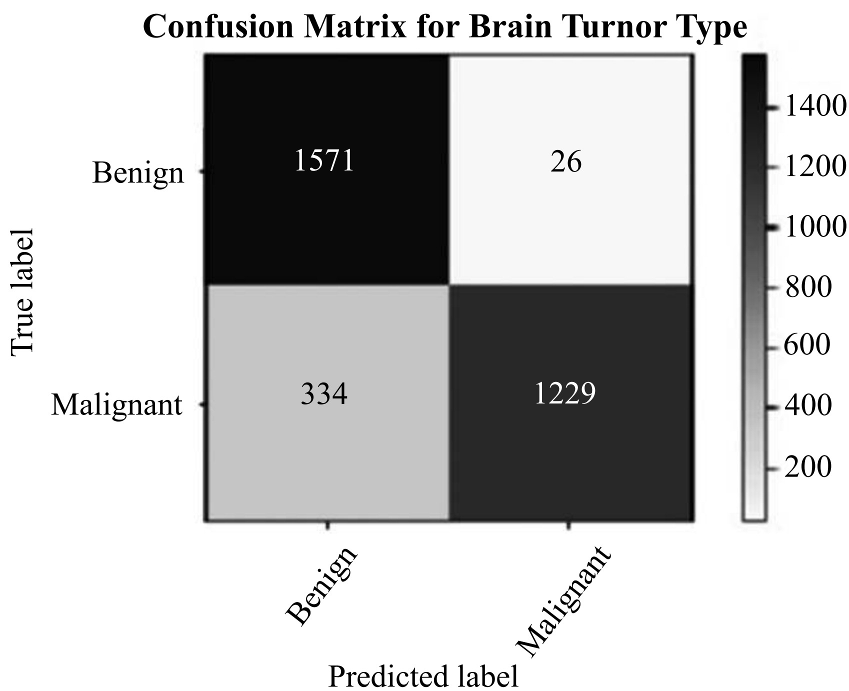


FIGURE 4.17 Confusion matrix for tumor prediction.

- There are three possible predicted classes: glioma (C1), meningioma (C2), and pituitary (C3); these are the type of malignant tumors.
- A total of 490 patients were examined for the existence of glioma tumor, 559 patients for the presence of meningioma, and 534 patients for the presence of pituitary tumor, according to the classifier.

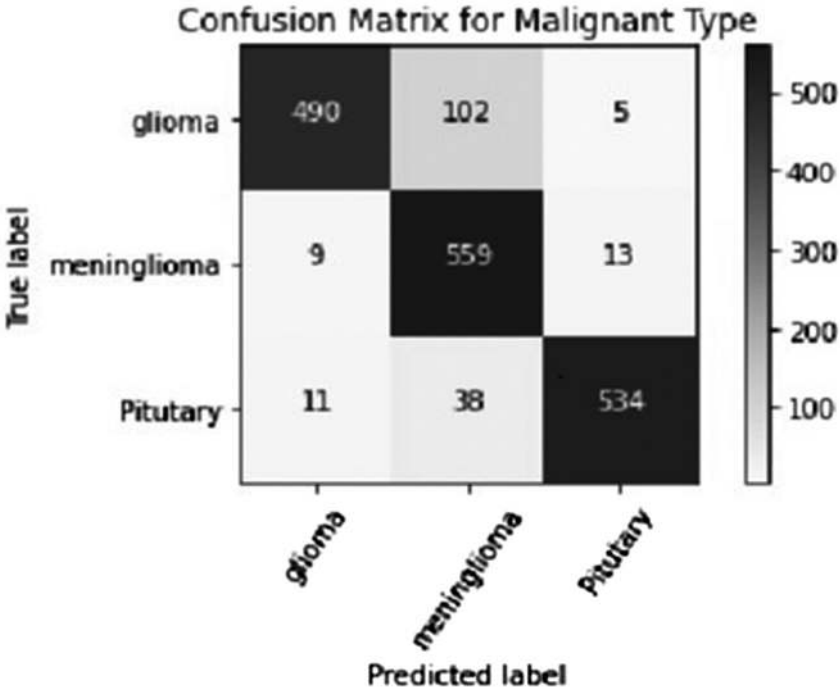


FIGURE 4.18 Confusion matrix for tumor-type prediction.

- There are three possible predicted classes: grade 2, grade 3, and grade 4.
- A total of 269 patients were examined for the existence of grade-2 tumors, 283 patients for the presence of grade-3 tumors, and 260 patients for the presence of grade-4 tumors, according to the classifier.

4.7 CONCLUSION

This study uses CNN models to give a multiclass prediction of brain tumors for fast diagnostic cause. Using openly available medical image datasets, three viable CNN models are discovered for three different brain tumor prediction tasks. The accuracy rate for detecting a brain tumor is 87.3%. Furthermore, the accuracy of brain Magnetic Resonance categorization into glioma, meningioma, and pituitary is 92.66%. Finally, with an accuracy of 90.75%, glioma brain tumors are categorized as grade 2, grade 3, or grade 4. The suggested CNN models are trained & evaluated utilizing a huge enough no. of radiology pictures. The outcome acquired utilizing the suggested CNN models and comparisons with state-of-the-art CNN models illustrate the efficiency of the CNN models developed among the given methodology.

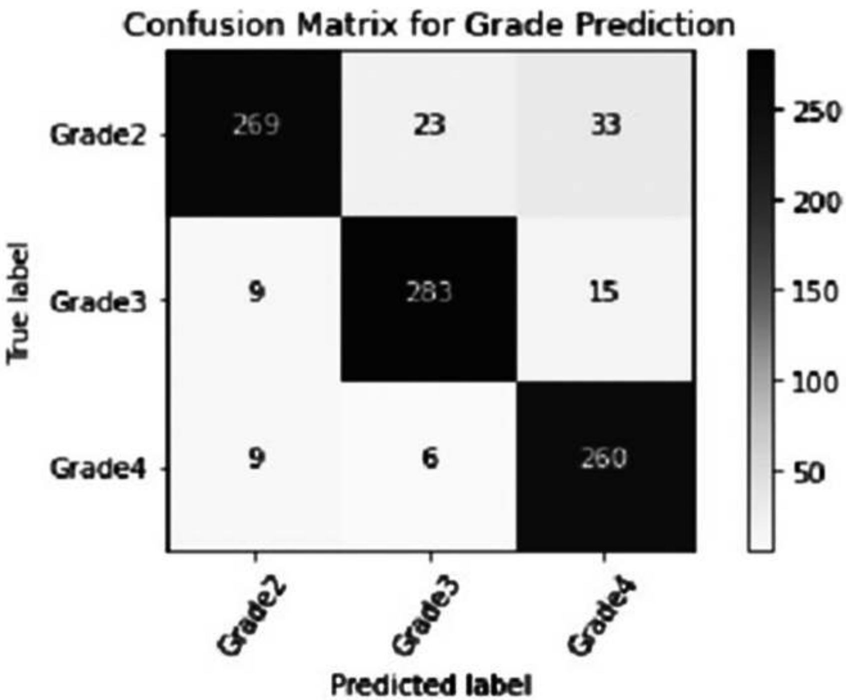


FIGURE 4.19 Confusion matrix for grade prediction.

KEYWORDS

- **deep neural networks**
- **brain tumor**
- **machine learning**
- **prediction**
- **classification**

REFERENCES

1. Kleihues, P.; Burger, P. C.; Scheithauer, B. W. Histological Typing of Tumours of the Central Nervous System. **2012**.
2. Emrah, I. Multi-Classification of Brain Tumor MRI Images Using Deep Convolutional Neural Network with Fully Optimized Framework. **2021**.
3. Ayadi, W.; Elhamzi, W.; Charfi, I.; Atri, M. Deep CNN for Brain Tumor Classification. **2021**.
4. Mohammed, B. A.; Al-Ani, M. S. An Efficient Approach to Diagnose Brain Tumors Through Deep CNN. **2020**.
5. Khawaldeh, S.; Pervaiz, U.; Rafiq, A.; Alkhawaldeh, R. S.; Noninvasive Grading of Glioma Tumor Using Magnetic Resonance Imaging with Convolutional Neural Networks. **2017**.
6. Scarpace, L.; Flanders, A.; Jain, R.; Mikkelsen, T.; Andrews, D.; Data from Rembrandt the Cancer Imaging Archive. **2015**.
7. Alanazi, M. F.; Ali, M. U.; Hussain, S. J.; Zafar, A.; Mohatram, M.; Irfan, M.; AlRuwaili, R.; Alruwaili, M.; Ali, N. H.; Albarrak, A. M. Brain Tumour/Mass Classification Framework Using Magnetic-Resonance-Imaging-Based Isolated and Developed Transfer Deep-Learning Model. **2022**.
8. Sultan, H. H.; Salem, N. M.; Al-Atabany, W., Multi-Classification of Brain Tumor Images Using Deep Neural Network [Online]. https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8723045
9. Abd El Kader, I.; Xu, G.; Shuai, Z.; Saminu, S.; Javaid, I.; Salim Ahmad, I. Differential Deep Convolutional Neural Network Model for Brain Tumor Classification. **2021**.
10. Banzato, T.; Cherubini, G. B.; Atzori, M.; Zotti, A. A Methodological Approach for Deep Learning to Distinguish Between Meningiomas and Gliomas on Canine MR-images. **2018**.
11. Talo, M.; Yildirim, O.; Baloglu, U. B.; Aydin, G.; Acharya, U. R. Convolutional Neural Networks for Multi-class Brain Disease Detection Using MRI Images. **2019**.
12. Cinar, A.; Yildirim, M. Detection of Tumors on brain MRI Images Using the Hybrid Convolutional Neural Network Architecture. **2020**.
13. Khan, H. A.; Jue, W.; Mushtaq, M.; Mushtaq, M. U. Brain Tumor Classification in MRI Image Using Convolutional Neural Network. **2020**.
14. Anaraki, K.; Ayati, M.; Kazemi, F. Magnetic Resonance Imagingbased Brain Tumor Grades Classification and Grading via Convolutional Neural Networks and Genetic Algorithms. **2019**.
15. Banzato, T.; Cherubini, G. B.; Atzori, M.; Zotti, A. Development of a Deep Convolutional Neural Network to Predict Grading of Canine Meningiomas From Magnetic Resonance Images. **2018**.
16. Zacharaki, E. I.; Wang, S.; Chawla, S. Classification of Brain Tumor Type and Grade using MRI Texture and Shape in a Machine Learning Scheme. **2009**.
17. Cheng, J.; Huang, W.; Cao, S.; Yang, R.; Yang, W.; Yun, Z. Enhanced Performance of Brain Tumor Classification Via Tumor Region Augmentation and Partition. **2015**.

CHAPTER 5

SMART HEALTHCARE APPROACH USING MACHINE LEARNING FOR BREAST CANCER DIAGNOSIS AND PREDICTION

NEHA PANWAR and D. P. SHARMA

Department of CS & IT, Manipal University Jaipur, Jaipur, India

ABSTRACT

Smart healthcare is the demand of today's world that helps in many health-related issues to cured timely and also helps to develop new approaches in the medical field for managing medical data which is broad and complex. With the help of a smart healthcare approach, it is easy to analyze and predict diseases like breast cancer with better accuracy and early prediction of having presence in the human body. Machine learning, as great field of artificial intelligence, plays a parental role to develop smart healthcare systems. Machines need to learn from input that plays a vital role, because by crunching large volume of data, it helps medical professionals to provide medical solutions to patient-based customized individual characteristics. Here in this study, we find how breast cancer disease prediction is done for diagnosis with a large dataset. In this study, supervised classifier support vector machines (SVM) are used for the classification Wisconsin breast cancer dataset for prediction and diagnosis. In this research study, the classification problem is solved using supervised algorithms logistic regression, SVM, k-NN, Naïve Bayes, and decision tree. A comparative study of accuracies is generated by implementing these supervised classifiers in classification problems. The study provides the result that SVM gives

Applied Soft Computing Techniques: Theoretical Principles and Practical Applications. Samarjeet Borah, Ratna Raja Kumar Jambi, Sharifah Sakinah Syed Ahmed, & Mahendra Prabhakar Deore (Eds.)
© 2025 Apple Academic Press, Inc. Co-published with CRC Press (Taylor & Francis)

higher results of accuracy with all dataset. This study concluded with better result by implementing SVM classifier approximately up to 99% that is improved than previous research accuracies for breast cancer classification problems. The comparison of performance among all classifiers is measured with a confusion matrix by calculating performance measuring factors, that is, accuracy, sensitivity, specificity, and precision.

5.1 INTRODUCTION

This study is intended to have look toward breast cancer in women. This study focused on the diagnosis and prediction problem of breast tumor with the help of automated models to support smart healthcare. Although there are different methods available that are used for the identification of the presence of breast cancer at early age so that detection and diagnosis of breast cancer can be done at prior stage and that helps to decrease the death rates. Earlier used procedural or traditional methods used for diagnosis of breast cancer are tricky and take so much time because that based on fine-needle tailored cancer therapies.[1] By considering tumor type or category decisions are taken based on current population-based therapies. This method is a human decision based in which knowledge base is past experience study of tumor category and patient which is not efficient way due to the dissimilarity found in patient tumor that can lead to the opposite outcomes for same type of tumor.[3] Due to this reason, an expert advice and automated systems are necessary for prediction or diagnosis of cancer. For this early prediction, different medical tests are performed to find the cancerous cell these are by using mammograms, MRI (magnetic resonance imaging), X-ray, US (ultrasound), and biopsy to detect cancer-affected cells along with this there can be possibility for some test or procedures for further screening of the cancer testing which is the conventional way of cancer diagnosis.[4] There are two types of cell malignant (having cancer) and benign (no having cancer) that are actually masses or tumors presented in breast tissue examined as abnormal lymph. If this lymph is found this is tested with fine-needle aspirant biopsy for the malignancy and benign. As the advancement in technology and to avoid the undertaking of needless treatment machine learning is used to classify the cells that are of the malignant and benign category with the procedure of detection of different feature value in the complex and huge dataset and acknowledged to get classification of data points for having benign or malignancy cellular properties.[5] Breast cancer is very well known and highly discovered cancer category that is the reason for high rate of mortality due to which deaths are increasing vastly. For cancer cure, the most important early step that has to be taken is the detection of cancer at a prior stage of life and diagnosis. The process involves examining the breast tissue as malignant and benign are called tumor categorization. An automated diagnosing system is needed because traditional methods take long hours to diagnose and detect breast cancer manually. An automatic diagnosis system provides the best accuracy that results in the medical domain and generates a smart healthcare approach.[6] Soft computing plays a significant role in breast cancer diagnosis and prediction by providing intelligent techniques to analyze complex medical data and assist healthcare professionals in making accurate decisions.[7] Soft computing is a subfield of artificial intelligence that focuses on developing computational models inspired by human-like thinking and reasoning processes. It encompasses various methods such as fuzzy logic, neural networks, genetic algorithms, and machine learning. For identification of breast cancer, different approaches and methods used in which very well-known is physically examine of cell tissues in human body that include lymph nodes, skin, nipple discharge, etc., which is traditional and time-consuming as well as painful methodology. Another way is to use data acquired from images of breast and features of cell structure represented in digital form to process by different machine learning models. This digital data are made up of digital images known digital mammograms or can be called electronic images that are helpful in categorization of breast tissues as malignant or benign. Another one is to use US that help in finding of calcification in breast tissue and identifying the structure of lymph nodes. Along with these such methods other are MRI, histopathology, fine needle, and biopsy. In all these approaches, digital images are widely used for the prediction of breast cancer. Machine learning performs well and is a conventional hand-crafted technique that helps to select the most important features in dataset. In machine learning, the deeper analyzer tool for a huge and complex data structure is deep learning which is used for presenting fine results. This study provides different supervised machine learning algorithms (LR, support vector machines [SVM], k-NN, NB, DT) comparative analysis for prediction and detection of breast cancer in which SVM is selected and implemented for classifier modeling that provides greater accuracy than other models in the supervised category.

5.2 LITERATURE REVIEW

For the analysis process of data of extracting meaningful features and relativity of research study with the outcomes for stating the research problem successfully

TABLE 5.1 Literature Review.

Reference	Year	Purpose	Method/Result
[1]	2018	Breast cancer classification and prediction	In the study, SVM are used with three different kernel functions (1) linear, (2) polynomial, and (3) radial Achieved the highest accuracy with using RF of 99.90%, MLP of 99.80%, and k-NN with 99.10%
[2]	2018	Diagnosis and prognosis of breast cancer	Wisconsin breast cancer dataset are classified using ANN, k-NN, DT, and SVM The highest overall accuracy achieved by SVM 97.20%, feature selection accuracy by SVM 98%
[3]	2016	WBC dataset breast cancer risk prediction and diagnosis	Support vector machines, decision tree (C4.5), NB, and k-NN are implemented with weka tool and SVM provide the highest accuracy 97.13%
[4]	2006	Wisconsin dataset for diagnosis and prognosis are classified using AI	SVM implemented and for Wisconsin prognosis and diagnosis accuracy is achieved to 96.91% and 90%
[5]	2017	SVM implemented for prediction of cancer with different types of kernels	GA + RBF SVM perform better with accuracy 98.28%
[6]	2019	Breast cancer prediction using machine learning algorithms in comparative approach	SVM achieved highest accuracy of 97.20%
[7]	2017	Analysis of CAD system by different classifiers of machine learning	SVM is used extensively for breast tissue classification
[8]	2017	Compare between preprocessing techniques to find the best one that can increase accuracy for mammogram mass dataset	Highest accuracy achieved 97.37%
[9]	2015	Survey of ML techniques for modeling cancer prognosis	SVM outperform best for cancer prognosis classification
[10]	2018	Breast cancer diagnosis by using SVM-based ensemble learning algorithm	Highest accuracy achieved variance 97.89%
[11]	2023	Breast cancer diagnosis and prediction	Highest accuracy achieved to 70%

with the accomplishment of comparative review, here a review of literature is presented for the papers used for study analysis with a deep lookout.

5.3 MACHINE LEARNING

The amount of dye uptake machine learning is a very efficient method used for healthcare domain to diagnose at an early stage of life of having disease in which there are some unknown regularities and trends that can be easily identified. Machine learning helps to identify these trends to analyze error-prone data in a vast dataset or from empirical data to acquire knowledge from it for diagnosis or prediction. The conceptual unit of machine learning is an algorithm that makes it possible to supervise first and then make it as human decision-maker for the result by analyzing the various data sets for the acquisition of knowledge to take decision without intervention of human in prediction.

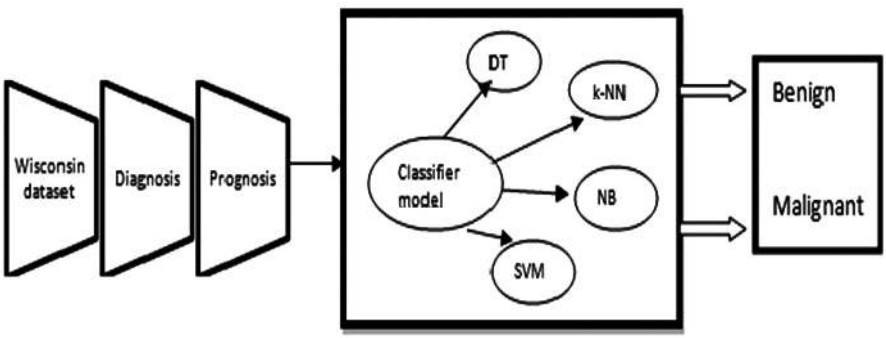


FIGURE 5.1 Machine learning classifier.

Classification of malignant and benign tumor is studied with the help of data gathered by images of tumor for checking growth points from different angle it perceive at different stages of age. Machine learning techniques categorized in three groups: (1) supervised, (2) unsupervised learning, and (3) reinforcement.[8] In supervised learning, a machine is trained to learn and then based on learning a testing is done to get known output corresponding to given input. On the other hand, unsupervised learning is to make machine self-capable or trained to categorize given input data in which there are no known class or output variables only input is given and machine is trained based on some common properties of features.[9] Reinforcement is a combination of supervised and unsupervised learning in which there are some known data labels and some unknown data labels are given based on which it train machine for known data labels.

5.3.1 SUPPORT VECTOR MACHINES

Experiments were under the supervised category the very popular and very important technique is SVM for classify procedure and regression for data. For maximizing the accuracy of prediction, it is necessary to avoid over fitting of data which is very smoothly covered by SVM. The procedure used by SVM is using hypothesis space of linear functions that helps to map lower level dimensional space to higher level feature space. SVM use vector representation of input and these vectors are used for classification with the help of a hyper-plan of two or more up to n-dimensional space. There can be multiple hyper-plans to separate the classes but to get optimal hyper-plan by introducing the maximum margin classifier hyper-plan that is a linear classifier called linear support vector machine (L-SVM). Here, eq 5.1 is for calculating the maximum marginal classifier.

$$\text{Margin} \equiv \arg\min \mathrm{d}(x) = \arg\min \frac{|x \cdot w + b|}{\sqrt{\sum_{i=1}^{d} w_i^2}} x \epsilon D \tag{5.1}$$

The purpose of using the maximum marginal classifier is for removing boundaries that are nonlinear by using kernel trick and then separating data with the help of using the concept of hyper-plan. For generating feature space kernel trick is used for mapping it to high dimensional feature space in the case when it is nonlinear and inseparable. The equation for Kernel is eq 5.1 and the equation for separating classifier is eq 5.2. There are different types of kernels available, that is, polynomial, GRB (Gaussian radial basis), ERB (exponential radial basis), and multilayer perceptron. SVM is very successful for classification that can be improved with the help of a better choice of kernel.

$$\begin{aligned} &K(X,Y) =_{\varphi} (X) \cdot_{\varphi} (Y) \\ &\text{If } Y_i = +1; \quad w \cdot X_i + b \geq 1 \end{aligned} \tag{5.2}$$

$$\text{If } Y_i = -1; \quad w \cdot X_i + b \leq 1 \tag{5.3}$$

$$\text{For all } i; \quad Y_i(w_i + b) \geq 1$$

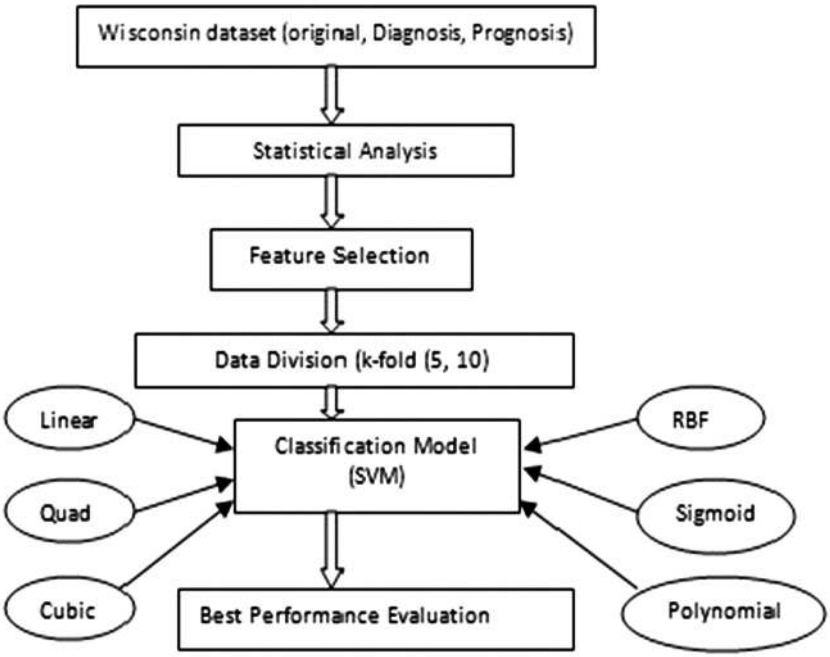


FIGURE 5.2 SVM modeling for classification flow.

5.3.2 GRID SEARCH

Grid search (GS) is used to optimize hyper-parameter of SVM. There are different hyper-parameter, that is, *C*, degree, gamma coefficient; these are defined by lower bound, upper bound, and number of steps.[2] GS is helpful in prevention of over-fitting with the help of cross-validation. SVM is implemented with four types of kernels each one with different settings of parameters: (1) linear kernel have one parameter *C* (the penalty), (2) RBF, and (3) Sigmoid kernels have two parameters with *C* and gamma, (4) polynomial kernel has three parameter *C*, gamma, and degree setting.[10] Due to variety of parameters in SVM modeling, the best combination formed to get better accuracy result.

5.3.3 FEATURE SCALING (FS)

Feature scaling (FS) is the part of data preprocessing that intends to clean the data to understand how the values are correlated to each other and impact the prediction accuracies for model. Here for applying data preprocessing and cleaning in datasets, first of all, null values eliminated along with data tape conversion performed for converting of object type to numeric. FS is used to keep the same domain range between 0 and 1. With eqs 5.3 and 5.4, FS can be done for x_{train} and x_{test}.

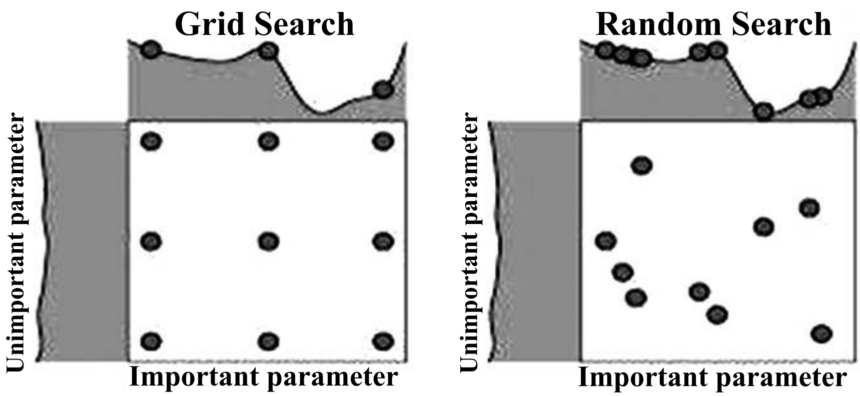


FIGURE 5.3 Grid search.

$$x_{\text{train scaled}} = \frac{x_{\text{train}} - x_{\text{train min}}}{x_{\text{train max}} - x_{\text{train min}}} \tag{5.4}$$

$$x_{\text{test scaled}} = \frac{x_{\text{test}} - x_{\text{test min}}}{x_{\text{test max}} - x_{\text{test min}}} \tag{5.5}$$

5.4 EXPERIMENTAL ENVIRONMENT

For classification, the most common used algorithm is SVM based on maximum boundary which use kernel methods. It finds in high-dimensional space a classification hyper-plane so that classification error can be minimized. The experiment environment in this study is about to apply a supervised learning algorithm for classification to find out the better result in accuracy on available WBC (Wisconsin breast cancer) dataset for (WDBC) diagnosis and (WPBC) prognosis prediction. In this study, Python language interface is used and supervised learning algorithms (LR, k-NN, NB, SVM, and DT) are applied. All three datasets are analyzed by dividing it into training and testing set.

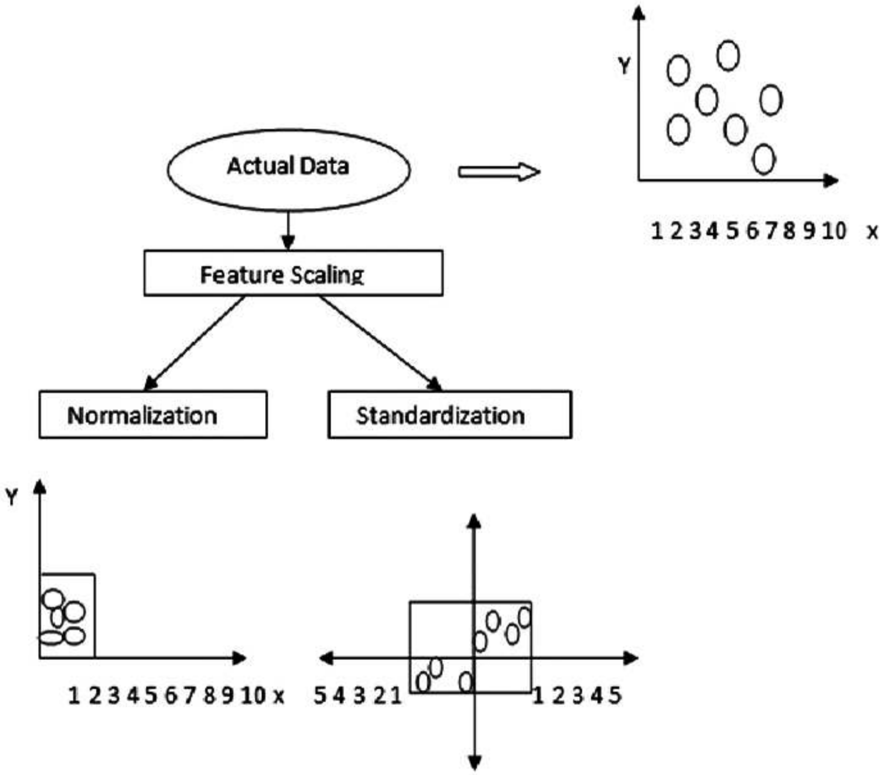


FIGURE 5.4 Feature scaling.

5.4.1 METHODOLOGY

After studying the dataset, it is very important of what types of algorithms and techniques are going to be implemented based on features whether are categorical or continuous. In this case before reaching to classification, there are some important steps required that comes under data preprocessing which is used for better fitting of data, that is, selection of features, reduction of dimension, and extraction of features. Here, the selection of features is followed by dimensionality reduction that select the subset of feature from existing large data set of original features by eliminating ineffective and unwanted feature from input set that reduce data noise and enhance the reliability of getting better results.[11] In feature extraction, new feature are generated from the existing or available features. Filter approach for feature selection process is an evaluation method for attribute and subset for calculating relevance among the features, in attribute relevancy each attribute is tested for having relevant mapping to target feature and assigned by weight according to degree of relevancy. Compared to filter method, wrapper method is very slow and a time-taking process that focuses on only feature subset relevancy or dependencies among them, but very good to provide optimal feature subset.[12] Embedded method takes less time than the wrapper and use SVM approach for selecting features.

The classification of breast cancer data for malignancy or benign starts with the collection of data that are taken from the UCI repository, preprocess data to remove noise and empty or null values for variable in input data and scale in a range of [0,1] that helps to remove or avoid outliers form data which is called scaling. Normalization is implemented for scaling the data set values to a range from 0 to 1 which is called data preparation technique to use a common scale without losing any information In this, there is a need to identify the upper and lower bound and with no outlier. Then, for further preprocessing step includes avoiding over-fitting that can lead data for low accuracy result and if data is over–fitted, then machine learning try to get avoid noise represents the data points that are not true properties in our data. These procedures introduce a loss function called residual sum of squares or RSS. These coefficients minimize loss function by adjusting the coefficient based on training data. There is different form of regularization known as L1 and L2. Ridge and Lasso are the variation techniques for regularization. L1 form is Lasso; it uses modulus as penalty for penalizing the high coefficient. Here an introduction of tuning parameter that is used to define the flexibility of penalty in our model that is called shrinkage factor lambda. Lasso is an equation of summation of modulus of coefficient less than or equal to s that is constant exist for lambda.

$$|\beta_1||\beta_2| \leq s$$

These are produced by standard least squares and expressed by sum of square of coefficient where all the points lie within circle given:

$$\beta_1^{\,2}\beta_2^{\,2} \leq s$$

5.5 RESULT ANALYSIS

The result in this research is evaluated by deriving a table used to measure the performance of model called confusion matrix. It is a graphical representation in form of a table with four combinations of predicted and actual values.

To obtain accuracy, precision, recall, specificity, sensitivity, $F1$ score, and AUC–ROC curves, a confusion matrix is used and evaluated for determining the performance of model for which there are four elements: true positive, false positive, false negative, and true negative (TP, FP, FN, TN). Sensitivity is a true-positive rate that are correctly classified positive as malignant cases where specificity is negative rate of correctly classified for benign cases. Accuracy is represented by ratio of true positive classified data points and true-negative data points out of total data points or we can say that it is for correct accuracy in model prediction. Precision is the correct positive label out of all classified positive data points. This research study compares three Wisconsin breast cancer dataset for diagnosis and prognosis. Using of default parameters is not implemented here besides this a range for domain is used for hyper-parameters (C, degree, gamma, and coefficient) and four different categories of kernels are used (linear, Sigmoid, polynomial, and RBF). Here, Table 5.2 is presenting accuracies for different classifier. By using GS, it is easy to find optimal parameter choice but also it is time consuming process, if there is large set of parameters and with high-dimensional dataset. Here, accuracies can be measured for all three datasets, where in the supervised machine learning algorithm, SVM achieve highest accuracies of 99% and NB performs lowest accuracies. By using GS for parameter optimization, accuracies achieved highest for WBC original dataset with SVM kernel radial bias to 93.65%.

TABLE 5.2 True Positive, False Positive, Precision, Recall, Specificity F1 Score Accuracy for Classifiers.

S.No.	Classifier	TP	FP	Precision	Sensitivity	Specificity	F1 score
1	SVM	0.99	0.99	0.99	1.00	0.98	0.99
2	k-NN	0.97	0.96	0.96	0.98	0.94	0.96
3	LR	0.93	0.92	0.92	0.98	0.83	0.90
4	DT	0.92	0.92	0.92	0.93	0.90	0.89
5	NB	0.47	0.57	0.44	0.97	0.21	0.04

TABLE 5.3 Grid Search.

Dataset	Grid search (%)
Wisconsin original	93.65
Diagnosis	92.98
Prognosis	74.35

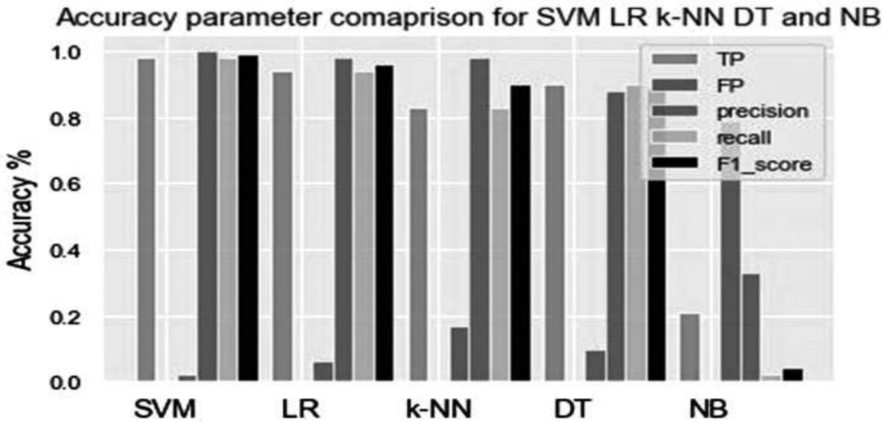


FIGURE 5.5 Efficiency and classification comparison for SVM, LR, k-NN, and NB.

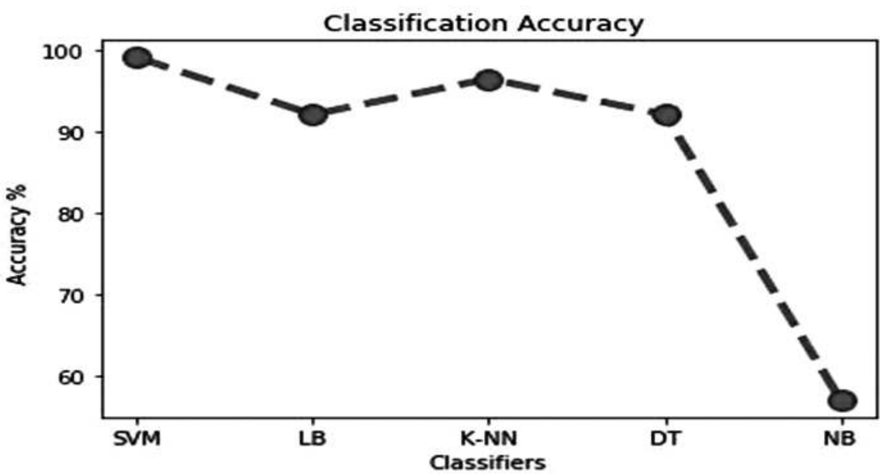


FIGURE 5.6 Classification accuracy.

5.6 CONCLUSION

In this paper, supervised machine learning models are compared for breast cancer prediction for finding out the best accuracy in terms of different accuracy measurement parameters. In all the models, SVM identified the best one for breast cancer models fall under the category for prognosis and diagnosis. Here single SVM and SVM with GS for different kernel combinations were examined for breast cancer prediction. GS helps to automate the selection of hyper-parameter and reduces time for classification with enhanced accuracy. For better comparison, different scaled datasets are used for accuracy measurement.

This experiment helps to understand the prediction performance and accuracy of SVM and other machine learning model for supervised environment. This research helps us to use SVM for classification problem for breast cancer prediction in future by assembling with other methodologies. Future study will lead to find the best assemble method of SVM for classification and prediction for breast cancer. Type of dataset is also a crucial point of study that becomes another area of research for future which also leads to study different available breast cancer dataset with SVM for better accuracy.

KEYWORDS

- **breast cancer**
- **classification**
- **SVM (support vector machines)**
- **smart healthcare**
- **Wisconsin dataset**

REFERENCES

1. Saygılı, A. Classification and Diagnostic Prediction of Breast Cancers via Different Classifiers. *ISVOS J.* **2018,** *2* (2), 48–56.
2. Yue, W.; Wang, Z.; Chen, H.; Payne, A.; Liu, X. Machine Learning with Applications in Breast Cancer Diagnosis and Prognosis. *Designs* **2018,** *2*, 13. https://doi.org/10.3390/designs2020013.
3. Asri, H.; Mousannif, H.; Moatassime, H. A.; Noel, T. In *Using Machine Learning Algorithms for Breast Cancer Risk Prediction and Diagnosis*, The 6th International Symposium on Frontiers in Ambient and Mobile Systems (FAMS 2016); *Procedia Comput. Sci.* 2016; vol *83*, pp 1064–1069.
4. Zafiropoulos, E.; Maglogiannis, I.; Anagnostopoulos, I.; In *Artificial Intelligence Applications and Innovations*; International Federation for Information Processing (IFIP): Springer, 2006; vol *204*, pp 500–507.
5. Huang, M. W.; Chen, C. W.; Lin, W. C.; Ke, S. W.; Tsai, C. F. SVM and SVM Ensembles in Breast Cancer Prediction. *PloS One* **2017,** *12* (1), e0161501. DOI: 10.1371/journal.pone.016150.

6. Yadav, A.; Jamir, I.; Jain, R.; Sohani, M. Comparative Study of Machine Learning Algorithms for Breast Cancer Prediction - A Review. *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol.* **2019,** 979–985.
7. Yassin, N. I. R.; Omran, S.; Houby, E. M. F. E., Hemat, A. Machine Techniques for Breast Cancer Computer Aided Diagnosis using Different Image Modalities: A Systematic Review; Elsevier. *Comput. Learning Methods Prog. Biomed.* **2018,** *156*, 25–45.
8. Thomas, T.; Pradhan, N.; Dhaka, V. S. In *Comparative Analysis to Predict Breast Cancer using Machine Learning Algorithms: A Survey*, *International Conference on Inventive Computation Technologies (ICICT),* 2020; pp 192–196.
9. Kourou, K.; Exarchos, T. P.; Exarchos, K. P.; Karamouzis, M. V.; Fotiadis, D. I. Machine Learning Applications in Cancer Prognosis and Prediction. *Comput. Struct. Biotechnol. J.* **2015,** *13*, 8–17. ISSN 2001-0370;
10. Wang, H.; Zheng, B.; Yoon, S. W.; Ko, H. A Support Vector Machine-Based Ensemble Algorithm for Breast Cancer Diagnosis. *Eur. J. Operat. Res.* **2017,** *267*, 687–699.
11. Planey, K.; Sugimoto, M.; Hikichi, S.; Takada, M.; Toi, M. Machine Learning Techniques for Breast Cancer Diagnosis and Treatment: A Narrative Review. *Ann. Breast Surg.* **2023,** *7*, 7. doi.org/10.21037/abs-21-63
12. Lambin, P.; Roelofs, E.; Reymen, B.; Velazquez, E. R.; Buijsen, J.; Zegers, C. M.; Carvalho, S.; Leijenaar, R. T.; Nalbantov, G.; Oberije, C.; Scott Marshall, M.; Hoebers, F.; Troost, E. G.; van Stiphout, R. G.; van Elmpt, W.; van der Weijden, T.; Boersma, L.; Valentini, V.; Dekker, A. Rapid Learning Health Care in Oncology - An Approach Towards Decision Support Systems Enabling Customised Radiotherapy. *Radiother. Oncol.* **2013,** *109* (1), 159–64.

CHAPTER 6

SEDONA: A MENTAL HEALTH TRACKER

SUBHASH RATHOD, KOMAL SHARMA, ATHARVA KULKARNI, RANJEET GAIKWAD, and VARAD PATIL

Department of Computer Engineering, Marathwada Mitra Mandal's Institute of Technology, Pune, India

ABSTRACT

Mental health is a crucial aspect of overall well-being, and maintaining good mental health is essential for leading a happy and fulfilling life. Mental health issues, such as anxiety, depression, and stress, can have a significant impact on an individual's quality of life, affecting their relationships, work, and overall happiness. Our project aims to develop an Android application that helps them to stay on top of their mental health and stay informed about their mental state. Sedona is an Android application that helps users track their mental health by enabling them to log their moods, journal their feelings, and access articles and resources that relate to how they are feeling. With Sedona, users can keep track of their mental health over time, visualize their mood trends on a graph, and get insights that help them make informed decisions about their well-being. Sedona is a valuable tool for anyone looking to prioritize their mental health and stay on top of their emotional well-being. Innovative approaches like Sedona offer hope for addressing this critical public health concern as mental health issues continue to rise globally.

6.1 INTRODUCTION

An individual's overall wellness depends significantly on their mental health, which may also greatly impact their quality of life. However, identifying and

Applied Soft Computing Techniques: Theoretical Principles and Practical Applications. Samarjeet Borah, Ratna Raja Kumar Jambi, Sharifah Sakinah Syed Ahmed, & Mahendra Prabhakar Deore (Eds.)
© 2025 Apple Academic Press, Inc. Co-published with CRC Press (Taylor & Francis)

treating mental health disorders can be difficult, particularly considering the stigma and dearth of resources related to mental health problems. Numerous detrimental effects, such as lower productivity, harmed relationships, and even physical illness, can result from poor mental health. One in four individuals worldwide suffers from common mental health diseases, according to the World Health Organization (WHO), and due to the severity of the problem, many health researchers decided to focus on studying this discipline.[8] Nearly 66% of individuals with a recognized mental health problem never seek assistance from a licensed therapist. If stress and anxiety start to interfere with daily activities, there may be a major problem. In the long run, this could make the problem worse. The use of machine learning (ML) and deep learning techniques for improving the early diagnosis and treatment of mental health conditions has gained increasing attention in recent years.

Since automated applications and virtual assistants are perceived as being nonjudgmental and nonopinionated, people are more likely to express themselves when speaking with them. This could be a significant therapeutic advantage because effective treatment of mental health disorders requires the early detection of mental health issues.[7]

To solve this problem, we created Sedona. An Android application called Sedona for tracking mental health—giving users the ability to keep track of their mental well-being by recording their moods, keeping a log of their emotions, and reading articles that are relevant to their present mood. By giving customers a user-friendly platform to efficiently track and manage their mental health, Sedona offers a distinctive and accessible approach to mental health management.

Databases are used by the application to store and manage user data. These databases include the main application database, a resource database that contains articles linked to each mood, a journal database that contains journal entries, and a mood database that keeps track of the user's moods.

These datasets are utilized by Sedona to give users personalized, real-time insights regarding their mental health. This information helps users comprehend their emotional patterns, determine potential triggers, and make healthy changes to their lives.

This research paper's main goal is to evaluate Sedona's effectiveness in fostering mental health and well-being. The paper will give a thorough description of the application, covering both its technical details and overall operation. Sedona takes a more advanced approach to tracking mental health, giving users a convenient, personalized platform for keeping tabs on their mental well-being. Since Sedona gives users the tools to effectively manage their mental health, it has the potential to increase mental health and well-being.

The stigma surrounding mental health illnesses, which prevents people from getting treatment and causes them to suffer in silence, is another factor that inspired the development of this project. Mental illnesses can degrade the quality of life, cause physical illnesses, and reduce productivity, happiness, and fulfillment.

The goal of this project is to implement an Android application that aids in tracking moods and keeping a journal to enhance users' mental health. Applications and virtual bots offer hope for addressing this crucial public health issue as mental health disorders continue to develop internationally. An advanced mental healthcare application with built-in calming games, a chatbot, in-depth analysis, and round-the-clock therapist support can be created using this technology.

6.2 RELATED WORK

6.2.1 PROPOSED IDEA

As shown in Figure 6.1, the Sedona mental health tracking application offers use cases including mood logging, journal writing, mood graph construction, and so on. The application was created using a client–server architecture, allowing users to interact with it via a user-friendly interface while the server handles tasks like data storage, analysis, and suggestion. The components of the system architecture are as follows:

1. User interface (UI):
   - The UI component gives users a simple and engaging interface via which they may enter their feelings and diary entries. Users can browse suggested articles based on their current mood and investigate mood graphs.
   - Mood and journal entry handler: This component accepts and processes user input for mood and journal entries.
2. Application server:
   - It verifies the input and records it, together with pertinent metadata like the date and the degree of the emotion, in the mood database and journal database, respectively.
   - Mood graph generator: The mood graph generator component uses data visualization techniques to produce graphical representations of mood patterns over time by retrieving mood data from the mood database.

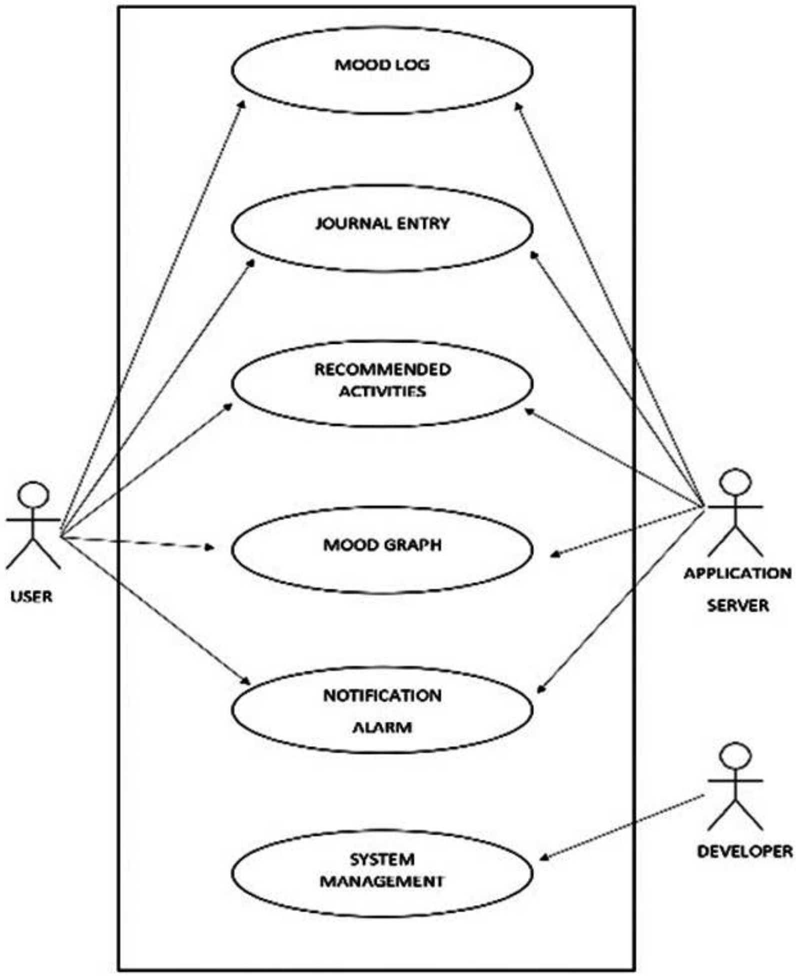


FIGURE 6.1 Use case diagram.

3. Databases:
   - Mood database: This database keeps track of the user's registered moods and their intensity. It keeps a record of the mood data in chronological order to make mood tracking and analysis possible.
   - Journal database: The user-provided journal entries are kept in the journal database. Users can evaluate their thoughts and experiences throughout time because each entry specifies the date and the entry's content.

- Resources database: Each article in the resources database has been labeled with pertinent metadata, such as mood, topic, and content. The Article Recommender component uses these articles as a resource pool to make suggestions based on the user's mood.

Overall, the system architecture of Sedona combines user interaction, data storage, analysis, and recommendation functionalities to provide users with an intuitive and comprehensive mental health tracking experience.

6.2.2 PROPOSED MODELING

The Android application was created in a way that will assist the user in monitoring their mental health. Users of the Sedona application can keep a daily mood journal and track their mood by recording journal entries. In accordance with the user's present mood, the program also suggests related articles, enabling them to learn more about their emotional condition and look for tools to enhance their mental health. Since it is a demanding undertaking, many academics have been working on tracking and detecting mental health for a long time. Keeping track of user emotions and assisting with emotion management is a difficult endeavor. Sedona uses a mood graph to show the user's mood over the last seven days, which is one of its main features. This function gives the user a visual picture of how their mood has evolved over time, allowing them to spot trends and adjust to enhance their mental health.

The internal architecture of the system is depicted in Figure 6.2's system architecture. This involves several actions, including mood logging, journal entry, article recommendation, and mood graph generation. Sedona promises to empower users to take control of their mental health and lead happier, more satisfying lives by offering customized resources and visual representations of mood data. The features of the program that are given below provide consumers with a better look at the Sedona functionalities that are available and how they may utilize them.

1. Mood tracking: One of the core features of Sedona is mood tracking. Users can log their moods daily using a simple interface, allowing them to monitor their emotional state over time. The app also provides a mood graph that displays the user's mood over the past seven days, helping users to identify patterns, and make changes to improve their emotional well-being and keep their mental health in check.

2. Journaling: Sedona also offers a journaling feature, allowing users to record their thoughts and feelings in a private, secure environment. This feature provides users with a space to reflect on their emotions and experiences, helping them to gain insight into their mental health and work through difficult emotions.

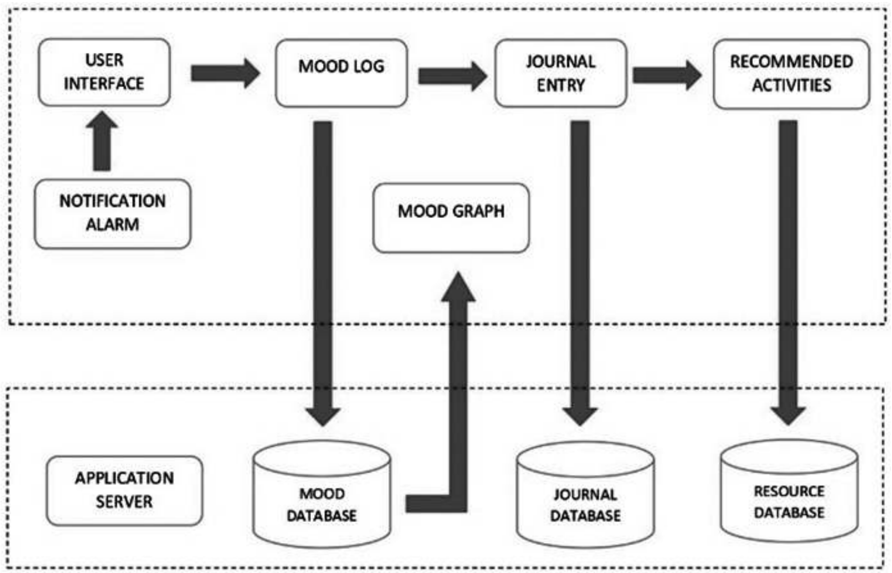


FIGURE 6.2 System architecture.

3. Related articles based on the user's current mood: Sedona provides personalized articles and resources to help users learn more about their emotional state and seek out resources to improve their mental health. This feature provides users with access to a wealth of information and tools to help them manage their mental health effectively.
4. Secure data storage: Sedona employs four databases to store data efficiently and securely. The main application database stores the user's journal entries and mood data, while the resource database stores related articles based on the user's current mood. Additionally, the journal database stores entries for the user's personal journal, and the mood database stores the user's mood history. This ensures that user data are kept safe and secure, protecting their privacy and confidentiality.

5. User-friendly interface: The user-friendly interface of Sedona makes it simple for users to access and utilize all its capabilities. Users may quickly and simply keep track of their mood and write journal entries thanks to the app's straightforward interface and easy navigation. The software includes a built-in dark mode to conserve resources and reduce eye strain.

The proposed Android application system is intended to offer an improved approach to mental health tracking, providing users with a comprehensive tool to help them monitor and improve their emotional well-being.

6.3 RELATED WORK

In Ref. [1], 88,126 user reviews are analyzed for sentiment using ML, and then the reviews are thematically analyzed to evaluate 103 mental health apps available on Google Play and the App Store. Using five classifiers that were developed using supervised ML techniques are used to compare their performance. The $F1$ score of the top-performing classifier, which was used to predict the sentiment polarity of reviews, was 89.42%. It also provides design suggestions on how to address the highlighted drawbacks to increase the efficiency of mental health apps.

Again in Ref. [2], research results have been described and discussed their implications for developing emotional wellness-promoting mobile applications. The research also carried out a feature analysis to investigate how these apps support the preparation, collection, reflection, and action stages of self-tracking. It was found that while mood-tracking applications offer a huge amount of functionality for the stages of collecting and reflection, they fall far short in the stages of planning and implementation.

In Ref. [3], U Srinivasulu, A Dharun, and Aditya Thota have used ML techniques to examine the stress patterns of working people and identify the variables that have the biggest impact on their stress levels.

In Ref. [4], the work aims to review the literature on DL algorithm used in outcome research on mental health. The authors first give a succinct introduction to contemporary DL approaches. The literature pertaining to DL applications in mental health outcomes is reviewed and challenges in DL are discussed.

In Ref. [5], the signal processing approach and the machine algorithms for mental health tracking are discussed. Signal processing is used for mapping from data to representations of behaviors and mental states and ML algorithms are used for drawing the inferences.

In Ref. [6], the authors develop an advanced artificial intelligence conversational agent that is completely functional and acts as a real-time therapist, analyzing the user's emotions at each stage and offering suitable responses and feedback.

In her work, Norah Saleh Alghamdi[6] has designed a user-friendly application using smart devices and text analytical tools. It provides quick pieces of advice, breathing exercises that are animated, and text-based therapy from licensed psychologists. The SVM fared the best when compared to other ML classifiers, getting a score of 79.81 percentage points on the text analytical tool.

Recently, various businesses, including the field of mental health, have paid a lot of attention to AI techniques.[8] Advanced AI methods and ML algorithms have made it possible to provide individualized care that is focused on offering emotional support tailored to a particular person. By evaluating the algorithms and parameters used in each system, numerous systems have been analyzed for mental health monitoring in this study, including virtual counseling, precision therapy, and diagnostic systems.

Teles et al.[9] reviewed an app used to help people with depression for this research. The results of this investigation showed that a few apps, including chatbots, online treatment, educational materials, mood monitors, testing, and self-help, are being used more frequently.

In Ref. [10], the study shows that people with mental illnesses worry about a variety of interconnected traits. To enhance the user experience, which will consequently enhance user adherence, these elements should be carefully reviewed within the design phase all through the development cycle of the cognitive behavior therapy apps in the healthcare industry.

6.4 RESULTS AND DISCUSSION

In this section, the results of our application have been discussed. Following are the results of our Android application. Figure 6.3 shows the welcome page of our application. It shows the functionalities offered by the application and the terms and conditions.

In Figure 6.4, we can see the journal entries added by the user to the Android application. It shows the title of the journal entry, which further shows the journal entry when we click on it. These data can be used by the user to assess feelings and thoughts. It can help the user to keep their mental health in check.

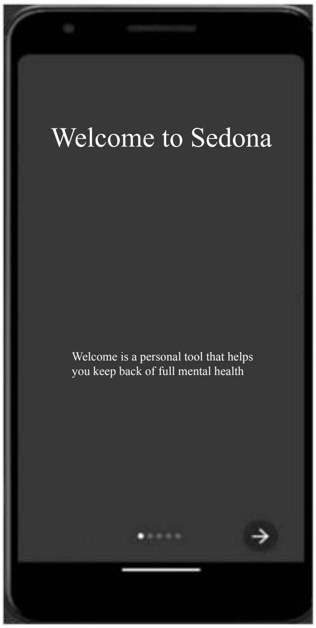


FIGURE 6.3 Welcome page.

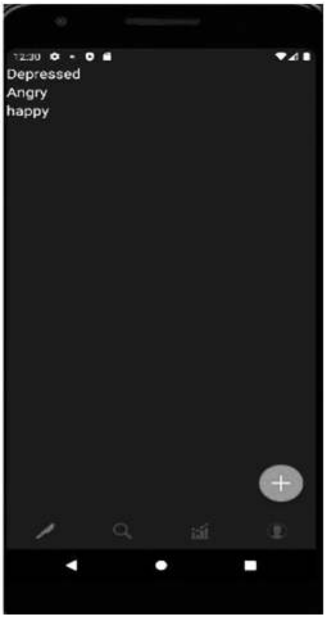


FIGURE 6.4 Mood logs.

In Figure 6.5, we can see the activities recommended by the Android application based on the mood logged last by the user. It recommends the user articles, songs, videos, and so on, based on their mood logs so the user can relax. These recommendations are made with the help of the resource database which consists of the various resources along with the related moods.

FIGURE 6.5 Resource recommendation.

In Figure 6.6, we can see the mood graph generated based on the mood logs by the user. The application accesses the mood database and plots the graph based on the previous seven-day entries by the user. It helps the user to keep track of their emotions and manage their mental health.

Our application helps users to navigate and manage their emotions by offering features like mood logs, journal entries, and mood graphs. With the ability of mood tracking, it enables the users to understand their mental state and helps them work through their issues. The app offers a user-friendly and easy-to-navigate interface.

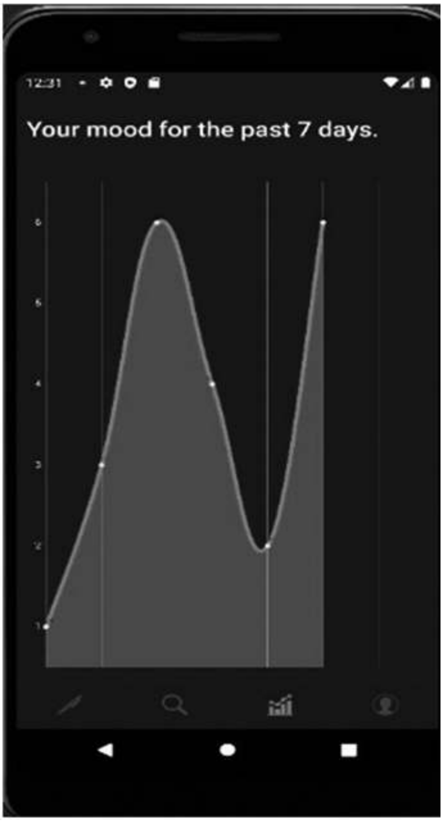


FIGURE 6.6 Mood graph.

6.5 CONCLUSION

In conclusion, the Sedona application has been successfully created to fulfill the requirement for a comprehensive, user-friendly, and accessible tracking tool for mental health. Sedona has given people a useful tool for tracking their mental health and encouraging self-care through the deployment of its core features, which include mood tracking, journaling, relevant articles, and mood graphs. The application's databases guarantee effective data management, and its user-friendly interface makes it simple for users to access and use all the application's functions. Sedona's aims have been implemented successfully, as evidenced by our peers' praise for it. There is enormous potential for the Sedona application to develop in the future. The incorporation of a chatbot would make the user experience more interactive and make it easier for consumers to get support and direction right away. Users who require professional assistance would benefit from an additional level of support if therapy services were included in the application. Furthermore, thorough mental health analysis might offer users tailored suggestions and resources based on their requirements. In conclusion, Sedona represents a significant step forward in addressing the need for accessible and comprehensive mental health tracking tools.

KEYWORDS

- **chatbot**
- **mental health**
- **monitor**
- **Sedona**
- **application**
- **mood**

REFERENCES

1. Caldeira, C. In *Mobile Apps for Mood Tracking: An Analysis of Features and User Reviews*, AMIA Annual Symposium Proceedings. AMIA Symposium, 2018. https://pubmed.ncbi.nlm.nih.gov/29854114/
2. Srinivasulu Reddy, U.; et al. In *Machine Learning Techniques for Stress Prediction in Working Employees*, IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), 2018. https://ieeexplore.ieee.org/document/8782395
3. Su, C.; Xu, Z.; Pathak, J.; Wang, F. Deep Learning in Mental Health Outcome Research: A Scoping Review. *Nature News*, April 22, 2020. https://www.nature.com/articles/s41398-020-0780-3
4. Bone, D., et al. Signal Processing and machine learning for Mental Health Research and Clinical Applications. *IEEE Signal Process. Mag.* **2017,** *34* (5), 196–195.. https://ieeexplore.ieee.org/document/8026204/
5. Moulya, S.; et al. Mental Health Assist and Diagnosis Conversational Interface using Logistic Regression Model for Emotion and Sentiment Analysis. *J. Phys. Conf. Ser.* **2021,** *2161* (1), 012039. https://iopscience.iop.org/article/10.1088/1742-6596/2161/1/012039

6. Alghamdi, N. S. Monitoring Mental Health Using Smart Devices with Text Analytical, Tool, 6th International Conference on Control, Decision and Information Technologies (CoDIT), 2019. https://ieeexplore.ieee.org/document/8820381/
7. Mody, V. In *Mental Health Monitoring System using Artificial Intelligence: A Review*, IEEE 5th International Conference for Convergence in Technology (I2CT), 2019. https://ieeexplore.ieee.org/document/9033652
8. Priya, A.; et al. Predicting Anxiety, Depression and Stress in Modern Life using Machine Learning Algorithms. *Procedia Comput. Sci.* **2020,** *167*, 1258–1267. https://www.sciencedirect.com/science/article/pii/S1877050920309091
9. Teles, A.; et al. In *Mobile Mental Health: A Review of Applications for Depression Assistance*, 2019 IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS), 2019. https://ieeexplore.ieee.org/document/878740
10. Thach, K. S. In *User's perception on mental health applications: a qualitative analysis of user reviews*, 2018 5th NAFOSTED Conference on Information and Computer Science (NICS), 2018. https://ieeexplore.ieee.org/document/8606901/

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

CHAPTER 7

MACHINE LEARNING-BASED HEART DISEASE PREDICTION SYSTEM

MRUNAL JAGTAP, KIRAN POTHARE, ABHIJEET JADHAV, ABHIJIT BOMBLE, ROHIT BARSHILE, and RADHIKA KOKATE

Computer Engineering Department, Marathwada Mitramandal's Institute of Technology, Pune, India

ABSTRACT

One of the main reasons for unexpected death or the main cause of mortality globally is heart syndrome. A person now loses their life due to heart disease every minute, making heart disease deaths one of the greatest challenges in today's society. Heart function is impacted by heart syndrome. According to a survey conducted by the World Health Organization, heart disease claimed the lives of 18 million individuals. The paucity of resources makes early disease prediction extremely difficult. When used in the healthcare industry, machine learning has the potential to diagnose diseases or syndromes accurately and early. Medical parameter attributes must be present in datasets. The datasets are analyzed in Python using the Random Forest Machine Learning Algorithm. In this study, a trustworthy machine learning algorithm is used to forecast cardiac illness or syndrome. An algorithm reads a CSV file of patient record data that it has obtained from surveys or hospitals. Following dataset reading, the procedure is carried out, and the effective heart attack level or syndrome level is generated.

7.1 INTRODUCTION

People in the current world do not even have time for themselves. To build fortune, they do not truly care about their health. Fast food is preferred

Applied Soft Computing Techniques: Theoretical Principles and Practical Applications. Samarjeet Borah, Ratna Raja Kumar Jambi, Sharifah Sakinah Syed Ahmed, & Mahendra Prabhakar Deore (Eds.)
© 2025 Apple Academic Press, Inc. Co-published with CRC Press (Taylor & Francis)

over nutritious homemade cuisine, which is bad for our general health and especially for the heart. Heart disease is caused by a variety of risk factors, which doctors classify into two groups.

Age, sex, and family history are examples of risk variables that cannot be modified. Risk factors include things like high blood pressure, smoking, inactivity, and high cholesterol. Additionally, some customers cannot afford or use some technology, such as CT scans and electrocardiograms, which are essential for identifying coronary heart disease. 17 million individuals have died because of the cause alone. Employees with cardiovascular disease were responsible for 25–35% of a company's yearly medical costs. To reduce the financial and physical costs of heart disease for individuals, organizations, and society as a whole, early identification is crucial. By 2030, the WHO predicts that there will be 23.6 million CVD-related deaths worldwide, with heart disease and stroke being the main contributors.

Then you start to consider how machine learning techniques differ from conventional techniques. For example, conventional techniques require a lot of data and people to undergo various tests, which are time- and money-consuming.[5] As a result, some people may assume that we can wait to perform those tests, which could result in a person losing their life due to a disease that was detected too late. On the alternative hand, machine learning approaches for detecting heart disease are highly effective, affordable, and very important for people with modern lifestyles. As a result, it is now possible to detect heart disease early using risk factors and machine learning algorithms. It can forecast whether a person will be overweight or have heart disease in the future based on their lifestyle, eating habits, and exercise routine. And if he is at risk for developing heart disease, he may start taking preventative measures right away techniques, and we are able to identify concealed patterns that may be used to clinical diagnostics.

Data mining is therefore essential in the medical sector, as evidenced by research carried out over the preceding few decades.[3] Many indicators should be considered when diabetes, elevated blood pressure, excessive cholesterol, and an irregular pulse are risk factors for heart disease. The findings in predicting heart disease are frequently affected by incomplete medical data.

Machine learning plays a crucial role in the medical sector. With the aid of machine learning, we can understand, identify, and forecast numerous illnesses. Using data mining and machine learning techniques, there has recently been an increase in interest in forecasting the chance of contracting specific diseases. Previously published studies use data mining techniques to predict disease. Numerous studies have tried to forecast the possibility of illness development in the future, but reliable results are not yet available. The main goal of this research is to accurately forecast the potential for human cardiovascular diseases.

7.2 EXPERIMENTAL METHODS AND MATERIALS

7.2.1 DECISION TREE

When trying to handle problems with regression and classification, decision tree supervised learning techniques are often utilized.[1] It has internal nodes for record properties, branches for decision-making, and leaf nodes for classification results at the conclusion of each branch. This classifier is tree structured. A decision tree has two nodes: a decision node and a leaf node. To make decisions, decision nodes are employed, On the other hand, a leaf node is the outcome of a choice and contains no more branches. Run tests and make judgements depending on the dataset's characteristics.

- Decision Tree Types many well-known decision tree algorithms, including the ones listed below, are built on "Hunt's algorithm," it was created in Psychology in the 1960s to simulate human learning.
- ID3: The creation of ID3, also known as the "Iterative Dichotomiser 3," is ascribed to Ross Quinlan. Entropy and information gain are used by this technique as dealings to assess potential splits.
- C4.5: This approach is regarded as a more recent version of "Quinlan's ID3" algorithm. To be able to judge divide values within "decision trees," it might employ information gain or gain ratios.
- CART: Leo Breiman coined the term "classification and regression trees" (CART), which is an acronym for it. Gini impurity is often used by this method to choose the best characteristic to split on. The frequency of misclassification for a randomly chosen property is quantified by the Gini impurity. A lower number is preferable when using Gini impurity to evaluate.

The decision tree's leaf nodes depict the program's results, while the trunks reflect the guidelines with each main component the functions, they govern an adaptive approach to supervising ML which can be used for issues related to regression and classification. Some of the most potent algorithms is this one. Additionally, among the greatest potent machine learning approaches, random forests, is used to train on various subsets of the training data.

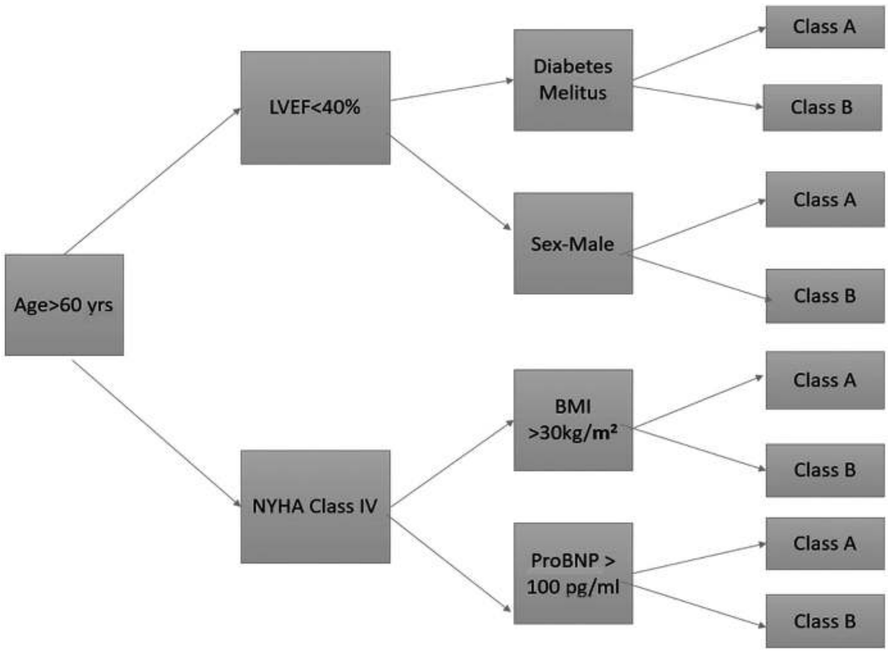


FIGURE 7.1 Decision tree.

A decision tree is what it is called since it origins from the origin node and expands like a tree. The tree is created using the "classification and regression tree method" (CART algorithm). Depending on the answer (yes/no), the decision tree creates a single query and splits it into subtrees.

7.2.2 RANDOM FOREST

Recommended Machine Learning Algorithms Supervised learning approaches are related to random forests. You can use it to solve ML problems related to "classification and regression." It is predicated on the concept of "ensemble learning," a technique aimed at addressing complicated issues and combining a significant number of classifiers to enhance model performance.

Both regression and classification tasks can be handled by an ensemble technique called random forest. This technique uses a series of decision trees and a technique called bootstrapping and aggregation (also known as bagging). This approach's main tenet is to mix together several decision trees in order to obtain the desired outcome rather than depending just on one.

For random forests, the primary learning paradigm is multiple decision trees. Rows and characteristics are chosen at random from the dataset to provide a sample dataset for each model. The bootstrap refers to this section.

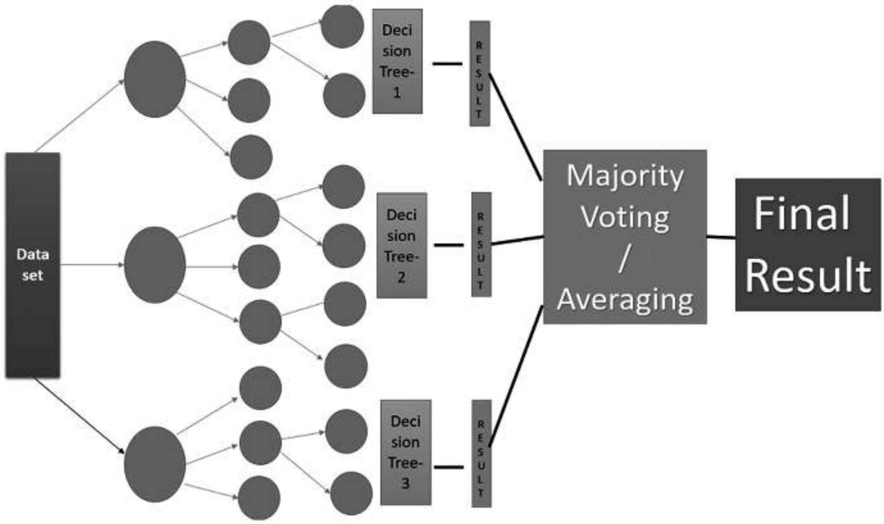


FIGURE 7.2 Random forest.

The variance is decreased since each decision tree has a wide range of variations, but when you mix them all at once, the conclusion depends on many decision trees rather than just one because each decision tree was effectively trained on that specific data sample. In a classification challenge, the winner is the classifier who receives the most votes.[1] The end outcome in regression problems is the average of all outputs. "Random Forest," equally its term shows, is a classifier that increases the predictability of a dataset by averaging numerous decision trees that have been applied to various subsections of the dataset.

Using predictions from each decision tree instead of just one, Random Forest predicts outcomes depending on the majority of votes in the forecast. The accuracy and overfitting issues are reduced when there are more trees in the forest. As its name suggests, a random forest is a type of classifier that enhances predictability by averaging many decision trees that have been used to analyze various parts of the dataset.[8] Using predictions from each decision tree instead of just one, Random Forest predicts outcomes depending on the majority of votes in the forecast. The accuracy and overfitting issues are reduced when there are more trees in the forest.

7.2.3 XG-BOOST

When training machine learning models, the distributed gradient boosting toolkit XGBoost was created to be quick and scalable. Methods of ensemble learning combine predictions from various weak models to get predictions that are more accurate. Due to its ability to deal with big datasets, XGBoost, or extreme gradient booster, is one of the most well-known and popular machine learning methods and provides innovative results in various machine learning tasks such as classification and regression. The effective treatment of missing data is a key aspect of XGBoost. This enables you to manage actual data with missing values without requiring any planning.[2] Additionally, XGBoost offers the ability to run parallel computations, which may be applied to model training.

Kaggle tournaments, recommendation engines, and click-through rate forecasting are just a few uses for XGBoost. Additionally, the model's many adjustable characteristics make it incredibly adaptable and simple to speed-optimize. Researchers at the University of Washington developed the idea of Extreme Gradient Boosting (XGBoost). This C++ package of his helped improve gradient reinforcement training methods.[6]

By minimizing the loss of the objective function on the dataset, XGBoost is trained. As a result, the selection of the loss function is a crucial hyper parameter that, like any deep learning neural network, is greatly influenced by the kind of issue it is intended to solve.

XGBoost may be used for a variety of purposes, such as Kaggle competitions, recommendation systems, and click-through rate prediction. It also allows for the fine-tuning of many model parameters, which makes it very versatile and facilitates speed optimization.

7.2.4 A MULTILAYER PERCEPTRON (MLP)

Multi-layer recognition, or MLP, is the term used for this. By using important and closely related levels, each input dimension may be converted into a target dimension. A multi-layer neural network is referred to as multi-layer perception. Neurons are coupled in a neural network such that some of their inputs and outputs are linked.[5]

A multi-layer perceptron includes neurons (or nodes) in each of its input and output layers, and any number of its hidden levels may have any number of nodes.

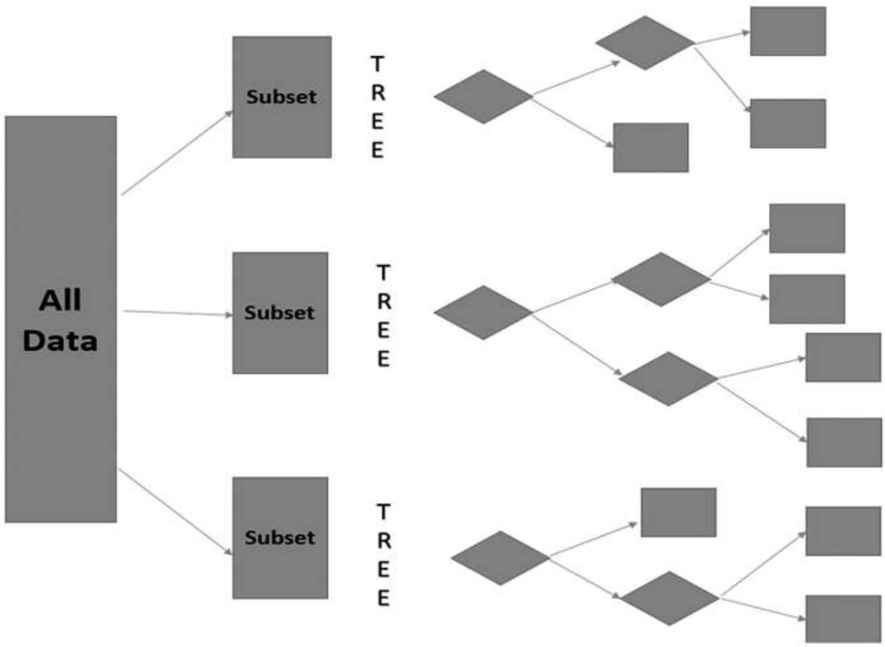


FIGURE 7.3 XGBoost.

Three inputs provide three input nodes and three nodes in the hidden layer in the multilayer perceptron design that was just explained. Since there are two outputs from the output layer, there are also two output nodes. As can be seen in the figure above, each of the three nodes in the hidden layer receives its results from its input layer nodes, which in turn analyzes the data already passing it to the output layer.[8] Inputs are received by the nodes of the input layer and forwarded for processing. In multilayer cognition, each node uses a sigmoidal activation function. Real data are converted to integers between 0 and 1 via a sigmoidal activation function.

Artificial neural feedforward networks with many inputs and outputs are called multilayer perceptrons (MLPs). MLPs are characterized by directed networks that connect input nodes at different input node levels and are connected between input and output levels. MLP trains the network via backpropagation. MLP is a deep learning technique.

The multilayer perceptron diagram above shows that there are three inputs that connect to three input nodes as well as three nodes in the "hidden layer." There are two nodes per output since the output layer has a pair of outputs. In the image, there are two nodes per output on the resulting level since each node has two outputs. And the hidden layers similarly process the data before sending it to the output layer node. Nodes at the input layer take in data and transmit it elsewhere for processing.

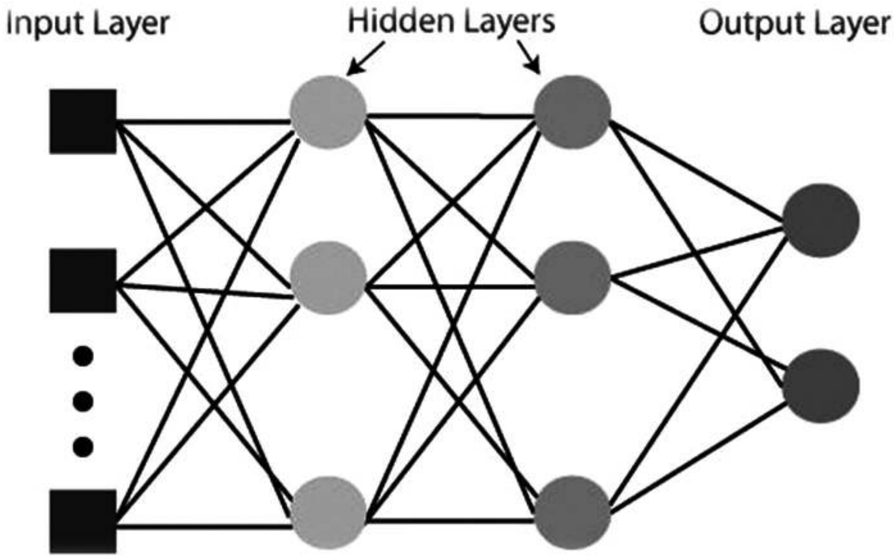


FIGURE 7.4 Multilayer perceptron (MLP).

7.2.5 SUPPORT VECTOR MACHINE

Support vector machines (SVMs) are among the most extensively accepted techniques for supervised learning used to solve regression and sorting issues. However, the majority of its applications involve classification using machine learning issues.

The SVM technique aims to provide optimum lines or decision boundaries that can categorize the n-dimensional space, allowing for rapid classification of fresh data points in the future. The best way around this restriction is through hyperplanes.

Both classification and regression are carried out using supervised machine learning techniques termed support vector machines (SVM). We often refer to regression problems, but they might be better described as classification problems. The SVM method identifies hyperplanes in N-dimensional space that accurately categorizes data points. The hyperplane size depends on the number of features. One way to visualize the hyperplane is to represent it as a line containing only two input features. The hyperplane will be a 2D plane if the input consists of three features. Anything having more than two traits is difficult to conceive of.

SVM comes in two varieties:

- Linear SVM: Frequently referred to as linearly separable data, “linear SVM” is utilized for data that can be broken down into dual groups by a single straight line. Linear SVM classifiers are utilized for such data.
- Nonlinear SVM: “Nonlinearly” divided data is carried using it, meaning that if a dataset cannot stay categorized using a straight line, it is “nonlinear” data and a “nonlinear SVM” classifier is being utilized.

Finding the ideal decision boundaries or lines to divide n-dimensional space into groups is the aim of the SVM approach, which enables following data points to be promptly classified into the appropriate category. A hyperplane is the perfect decision boundary. By utilizing SVM to choose extreme vectors and points, hyperplanes are built. The procedure is referred to as a “support vector machine” because of this peculiar circumstance.

7.3 RESULTS AND DISCUSSION

The 70,000 rows and 12 features in the original dataset were reduced to around 59,000 rows and 11 features after cleaning and preprocessing. To enhance model performance, outliers were eliminated as all characteristics were categorical. In this study, Decision Trees, Hierarchical Recognition, Random Forests, and XGBoost Classifiers were employed. F1 score, recall, accuracy, precision, and area under the ROC curve were the performance indicators employed in this investigation. 80% of the dataset was used to train the model, while 20% was used to test it.

Hyper parameter optimization was automated with the “GridSearchCV” method. “GridSearchCV” needs an estimator, a set of hyper parameters to search, and a scoring technique in order to identify the set of hyper parameters that maximizes the scoring method. This method, which is part of the “scikit-learn module,” employs “k-fold cross-validation” to evaluate the effectiveness of different sets of hyper parameters.

Following hyper parameter tuning of the cardiovascular disease dataset, many ML, including “MLP, RF, decision trees, and XGBoost,” were employed to detect the existence of cardiovascular illness. The stadium was the winner thanks to the multilayer perceptron (MLP) technique, which also had the greatest cross-validation accuracy (87.28%), extraordinary “recall, precision, F1 score, and AUC scores” (84.85, 88.70, 86.71, and 0.95). All classifiers had accuracy percentages greater than 86.5%. By using GridSearchCV to modify the “random forest algorithm's” hyper parameters, the accuracy was augmented from 86.48% to 86.90% by 0.5%. Similar to this, the XGBoost algorithm’s accuracy increased through hyper parameter optimization from 86.4% to 87.02%, a gain of 0.6%.

TABLE 7.1 Algorithm Comparison.

Model	Accuracy		Precision		“Recall”		“F1-score”	
	“Without CV”	“CV”	“Without CV”	“CV”	“Without CV”	“CV”	“Without CV”	“CV”
Random forest	87.85	76.2	89.03	88.70	82.95	84.85	85.88	86.71
Decision tree	76.92	81.02	88.52	89.42	83.46	83.43	85.91	86.32
XG-boost	86.53	8.37	90.10	89.11	81.17	81.61	85.40	85.32
MLP	72.05	82.4	89.62	88.93	82.11	83.57	86.30	86.16
SVM	82.23	87.36	78.36	78.41	79.25	78.36	81.38	83.25

An ROC (Receiver Operating Characteristic) curve is used to graphically depict the performance of a binary classifier. For various categorization criteria measurements occur across the “true positive rate” (TPR) and “false positive rate” (FPR).

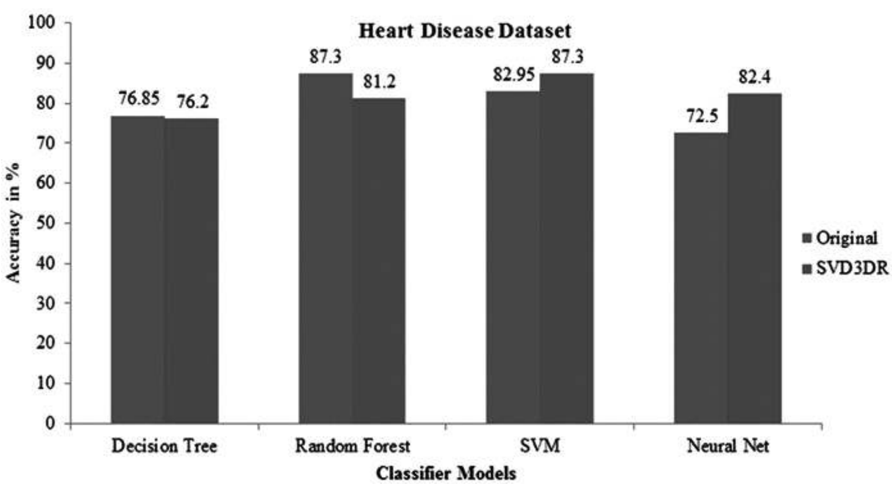


FIGURE 7.5 Chart for algorithm comparison.

AUC, a scalar statistic, assesses a classifier's overall performance as well as its sensitivity and specificity. All models had high AUCs over 0.9. The XGBoost, the greatest AUC of 0.95, is shared by the Random Forest and Multilayer Perceptron (MLP) models. In total 14 pertinent qualities, including those predicted to affect a person's risk of acquiring heart disease, are chosen from the 76 features in the dataset in order to evaluate the system. When these characteristics are considered, the system the innovator obtained is less effective. For efficiency, attribute selection is employed. The model with the highest accuracy should be evaluated using n criteria. Due to their close correlation, several dataset characteristics are excluded.[1] It is considerably less efficient to take into account all attributes in the dataset.

We discovered that the random forest method beats other algorithms in terms of accuracy after using machine learning techniques for both training and testing. The values based on the TP, TN, FP, and FN quantities were calculated using this equation. The best structure is found to be a random structure after extensive investigation and comparison. The confusion matrix is used to determine accuracy for each approach.

7.4 CONCLUSION

This study's primary objective was to categorize heart disorders using various models and actual information. We employed a K-mode clustering technique to forecast illness onset using a dataset of cardiac patients. Blood pressure readings, both diastolic and systolic measurements, were separated into 10-interval bins, while the age features were translated to years and divided into 5-year bins. The data was then divided by gender to account for certain traits and variations in the development of cardiac disease in both genders.

Statistics show that the MLP model was 87.23% accurate. These findings imply that K-mode clustering may reliably predict heart illness and that this method may be helpful in the development of heart disease diagnostics and targeted therapeutics. This study employed his 70,000-item Kaggle cardiovascular illness dataset, and all algorithms were created using Google Colab. As previously indicated, all approaches had an accuracy of 86.37%, with the multilayer perceptron having the highest accuracy and the decision tree having the lowest.

7.4.1 LIMITATIONS

There are several restrictions to consider despite the promising outcomes. The study may not apply to other demographics or patient groups because just one dataset was used. In this analysis, only a small number of demographic and clinical criteria were included, as well as additional possible contributors to risk for heart disease, including dietary habits and genetic predisposition, were also disregarded. Furthermore, the dataset kept for testing has not been used to evaluate model performance. This might be a sign of the model's ability to generalize to fresh, untested data. Finally, the results' interpretability and capacity to justify the clustering generated by this approach were not assessed.

KEYWORDS

- **machine learning**
- **heart disease**
- **prediction**
- **random forest**
- **cardiac illness**

REFERENCES

1. Gandhi, M.; Singh, S. N. In *Predictions in Heart Disease using Techniques of Data Mining*, 2015 International Conference on Futuristic Trends on Computational Analysis and Knowledge Management (ABLAZE), 2015. DOI: 10.1109/ABLAZE.2015.7154917
2. Thomas, J.; Theresa Princy, R. In *Human Heart Disease Prediction System using Data Mining Techniques*, International Conference on Circuit, Power and Computing Technologies (ICCPCT), 2016. DOI: 10.1109/ICCPCT.2016.7530265.
3. Bharti, S.; Singh, S. N. In *Analytical Study of Heart Disease Prediction Comparing with Different Algorithms*, International Conference on Computing, Communication & Automation, May 2015. ISBN Information: INSPEC Accession Number: 15274875. DOI: 10.1109/CCAA.2015.7148347
4. Purushottam; Saxena, K.; Sharma, R. In *Efficient Heart Disease Prediction System using Decision Tree*, International Conference on Computing, Communication & Automation, 2015. DOI: 10.1109/CCAA.2015.7148346.

5. Palaniyappan, S.; Awang, R. In *Intelligent Heart Disease Prediction using Data Mining Techniques*, IEEE/ACS International Conference on Computer Systems and Applications, Aug 2008. DOI: 10.1109/AICCSA.2008.4493524.
6. Sharma, H.; Rizvi, M. A.; In *Prediction of Heart Disease using Machine Learning Algorithms: A Survey*, 2019 1st International Conference on Innovations in Information and Communication Technology (ICIICT), Aug 2017. DOI: 10.1109/ICIICT1.2019.8741465
7. Hazra, A.; Mandal, S. K.; Gupta, A.; Mukherjee, A.; Mukherjee, A. Heart Disease Diagnosis and Prediction Using Machine Learning and Data Mining Techniques: A Review. *Adv. Comput. Sci. Technol.* **2017,** *10* (7), 2137–2159. ISSN 0973-6107 © Research India Publication.
8. Krishnaiah, V.; Narsimha, G.; Subhash Chandra, N. Heart Disease Prediction System using Data Mining Techniques and Intelligent Fuzzy Approach: A Review. *Int. J. Comput. Appl.* **2016,** *136* (2), 0975–8887.
9. Kaur, R.; Kaur, P. A Review-Heart Disease Forecasting Pattern using Various Data Mining Techniques. *Int. J. Bio-Sci. Bio-Technol.* **2016,** *8* (4), 139–148. DOI: 10.14257/ijbsbt.2016.8.4.16
10. Vijayashree, J.; SrimanNarayanaIyengar, N. C. Heart Disease Prediction System Using Data Mining and Hybrid Intelligent Techniques: A Review. *Int. J. Bio-Sci. Bio-Technol.* **2016,** *8* (4), 139–148. DOI:10.14257/ijbsbt.2016.8.4.16

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

CHAPTER 8

BREAST CANCER RISK DETECTION

UMA B. KARANJE, CHAITANYA S. BHOSALE, PRANJALI DESHMUKH, MAYURI BHUTKAR, SRUSHTI RAUT, and GAYATRI GAIKWAD

Department of Computer Engineering, MMIT, Lohegaon, Pune, India

ABSTRACT

Compared to other cancers, breast cancer is currently one of the top concerns for women. Breast cancer is a condition that develops when a woman's breast cells multiply uncontrolled. To detect breast cancer and change in the cells, we can use different algorithms and methods like Support Vector Machine (SVM), Decision Tree, Random Forest, and Convolutional Neural Network (CNN), in which convolutional Neural Network (CNN) algorithm was utilized mostly because it provides greater accuracy compared to other methods. One of the best methods for classifying images in deep learning is the convolutional neural network (CNN).[13] Through image processing, this research improves the automatic identification of breast cancer using MRI scans. To extract the appropriate areas that are relevant from the MRI images using this approach, segmentation is combined with several image processing procedures. Results from various research studies show that, in comparison to other methods, the accuracy rate for CNN has increased.

8.1 INTRODUCTION

As per the research, the second most prevalent cancer in women and the one with the greatest risk of death is breast cancer. The conclusions of a survey of death rate in India due to breast cancer from 2018 to 2022 is shown in Table 8.1 below.

Applied Soft Computing Techniques: Theoretical Principles and Practical Applications. Samarjeet Borah, Ratna Raja Kumar Jambi, Sharifah Sakinah Syed Ahmed, & Mahendra Prabhakar Deore (Eds.)
© 2025 Apple Academic Press, Inc. Co-published with CRC Press (Taylor & Francis)

TABLE 8.1 Death Rate.

Year	Death rate
2018	87,090
2019	41,760
2020	>76,000
2021	7,89,202
2022	8,08,558

In 2021, the Survival Rate was 90% and in 2022 the USH has a 90–95% Breast Cancer Survival Rate whereas India has a 66%. The Survival Rate is increased because of early detection of breast cancer. Because breast tumor patients have an average survival rate of over 90% over 5 years, early detection of the condition is vital for greatly lowering fatality rates in females. After all, the tumor was discovered when it was still very treatable. Breast cancer can be identified through biopsies and imaging testing.

Mammography, ultrasound, and magnetic resonance imaging (MRI) are the three imaging modalities that are most frequently used to diagnose breast cancer. Mammography is recognized as the most reliable method for cancer screening among these modalities shown in Figure 8.1. Despite being a common screening procedure, mammography is not as successful for women under the age of 40 who have thick breasts, are less responsive to the early tumor, or do not detect any cancer.

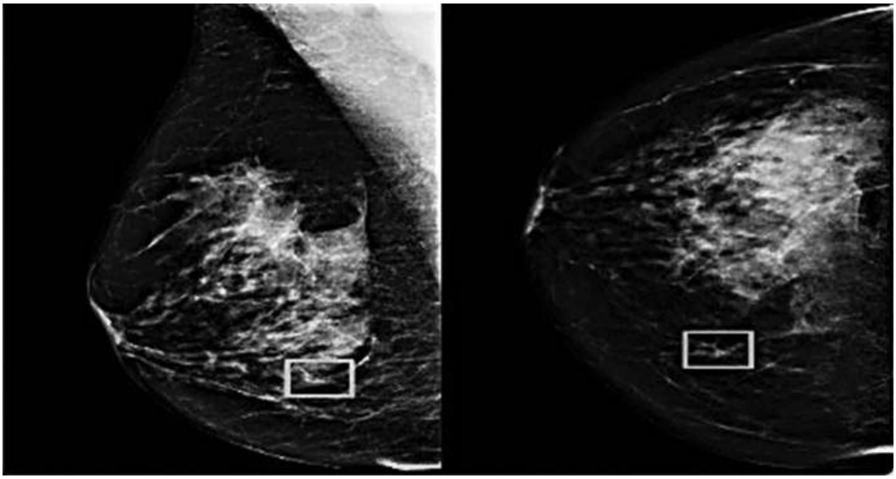


FIGURE 8.1 Sample mammography images.

8.2 LITERATURE SURVEY

In this paper, Nguyen et al.[1] use the Convolutional Neural Network also called CNN. Convolutional neural networks (CNN) are a particular kind of deep learning that excels at classification, image, and speech recognition. From the public BreakHis dataset, they classify and identify breast cancer photos using CNN. There are four benign subclasses and four malignant subclasses among the 7909 histopathological pictures of breast cancer (BC) in this collection.[1]

With open-source access to artificial intelligence-based software, the goal of this project is to create a system that can aid doctors in the identification of breast cancer. The suggested method, which classifies breast cancer (benign/malignant) based on histopathological pictures, was developed utilizing an open-source data set. In this case, the images produced by staining biopsy specimens collected from breast tissues with hematoxylin and eosin were analyzed using the Keras library and convolutional neural networks from deep learning techniques. The proposed system provides promising predictions in the classification of breast cancer, and it can be utilized for clinical decision support in the classification, according to experimental data from the study.[2]

This study's authors suggested the stacking classifier, an ensemble method that effectively distinguishes between benign and malignant tumors by combining various classification algorithms. This study makes use of the "Wisconsin Diagnosis Breast Cancer" dataset, often known as WDBC. Over the dataset, they used various classification approaches, tuning the parameters to increase accuracy. In this work, CART, Random Forest, Logistic Regression, K-Nearest Neighbors, and Support Vector Machine are used as classification algorithms.[3]

The "Wisconsin Diagnosis Breast Cancer" dataset, generally known as WDBC, is used in this study. The dataset used for this experiment consists of features collected from digitized images of FNA (Fine-needle aspiration) examinations done on breast lumps on cells from breast cancer patients and cells from healthy individuals. To assess the prediction findings, they employed the classification techniques Support Vector Machine, Decision Tree, Logistic Regression, K-Nearest Neighbor, Naive Bayes, Random Forest, and Neural Network-based algorithm-Multilayer Perceptron. To determine the top performing algorithm, the resulting findings were also contrasted with ensemble-based learning techniques as Gradient Boost, XGBoost, and Adaboost classifiers.[4]

Weka software has employed classification and clustering algorithms on the breast cancer dataset to forecast the nature of breast cancer in advance of additional therapy. Data preprocessing, classification, regression, clustering, association rules, and visualization are all performed on a dataset using the WEKA tool. Actual cancer patients' data sets have been gathered for the UCI Machine Learning Repository. The evaluation of the breast cancer database involved the analysis of several machine learning computations. Used were the three classification algorithms. Among the clustering techniques, simple k-means is the most suitable route for the study of the patients.[5]

This study used two well- known ensemble machine literacy styles to examine a bone cancer dataset and read the development of bone cancer. Bone cancer was prognosticated using Extreme Gradient Boosting (XGBoost) and Random Forest. For this exploration, an aggregate of 275 exemplifications with 12 features were used. In this disquisition, delicacy results were 73.63 using XGBoost and 74.73 exercising the Random Forest fashion. The study is based on "Wisconsin opinion bone Cancer" (generally known as WDBC) dataset.[6]

This study employs neural networks and different types of data mining to detect the presence of breasts. Breast cancer occurs when cells in the breast begin to grow or multiply out of control. It then spreads from the breast tissue. Cancers that arise from these cells are usually visible on a mammogram or can be felt as a lump. WEKA software was used in this study. Data preprocessing, classification, regression, clustering, association rules, and visualization are all performed on the dataset using WEKA tools.[7]

Traditional breast cancer image classification techniques call for the manual extraction of features from medical pictures. This requires expert medical knowledge in addition to being labor- and time-intensive and challenging to extract high-quality features. As a result, the paper suggests a computer-based strategy for classifying and detecting breast cancer images using feature fusion and convolutional neural networks. In this study, two convolutional neural networks with distinct structural pretraining are used. The convolutional neural network is then used to automatically extract feature characteristics, fuse the features from the two extracted structures, and then classify the fused features.[8]

This study shows that automated machine learning (AutoML) systems are helpful data science assistants designed to scan data for novel features, select appropriate supervised learning models, and optimize their parameters. For this purpose, Tree-based Pipeline Optimization Tool (TPOT) was developed using strongly typed genetic programming (GP) to recommend an optimized analysis pipeline for the data scientist's prediction problem. However, like other AutoML systems, TPOT may reach computational resource limits when working on big data such as whole genome expression data.[9]

8.3 MOTIVATION

Recent years have seen a rise in the understanding among many cancer researchers that this disease may be a systemic illness nearly from its start. Additionally, it is becoming more and more obvious that most malignancies are already advanced when they are initially brought to the attention of the clinician by the patient. The patient's obvious lesion may be eliminated, but its seeds have advanced in development or the immune system's ability to fight them has significantly decreased, rendering them ineffective. Table 8.2 shows the age-specific risk of breast cancer.

TABLE 8.2 Age-Wise Risk of Breast Cancer.

Age	Percentage
20–29	2.4
30–39	19.6
40–49	68.6
50–59	109.5
60–69	108.0
70–79	82.7
80–89	13.2
90–100	0.8

8.4 PROPOSED SYSTEM

The techniques recommended will help determine the difference between benign and malignant cancer more quickly. Despite being a sophisticated yet challenging classifier, CNN can automatically extract significant features without the need for preprocessing. It is more effective since it filters the crucial variables and is versatile enough to perform incredibly well with picture data.

The CNN algorithm, picture prior processing, and Feature Extraction are a few of the processes in the above process, which end with the detection of breast cancer. Figure 8.2 displays the main components that make up our application.

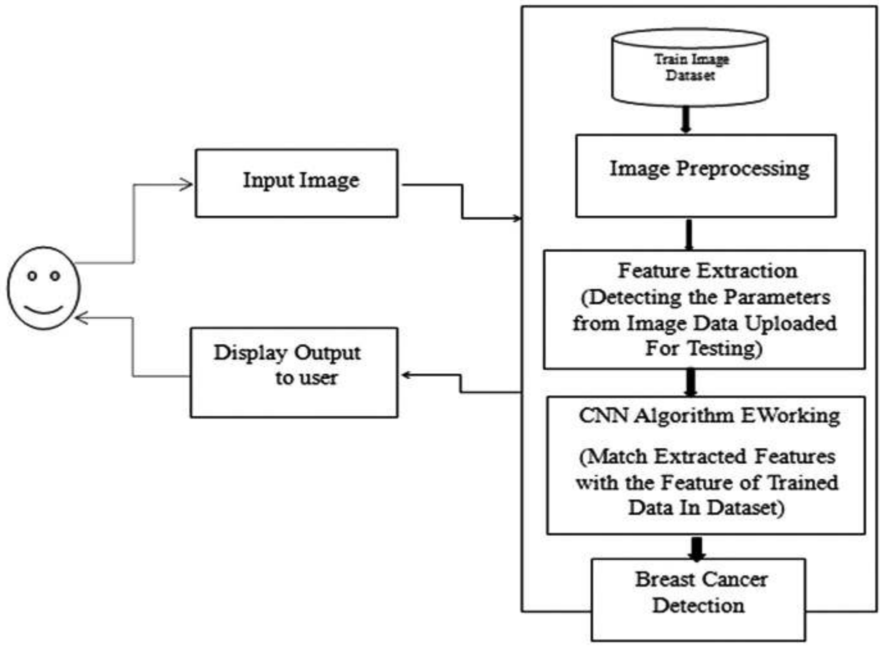


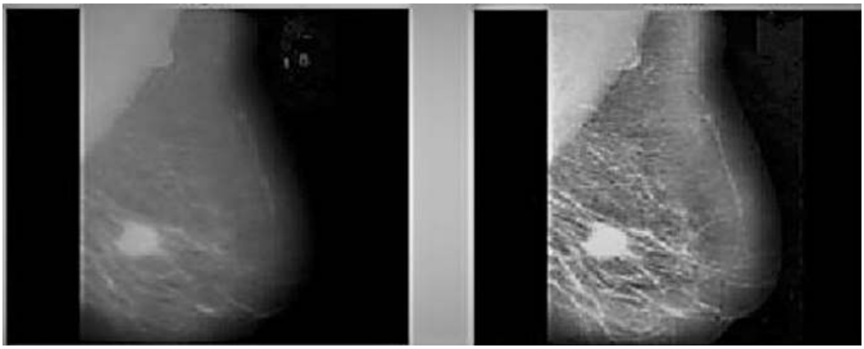
FIGURE 8.2 System architecture.

8.4.1 DATA PREPROCESSING

The main objective of preprocessing is to enhance the picture quality so that it is prepared for subsequent processing by eliminating or minimizing additional and irrelevant background elements from the mammography images as shown in Figures 8.3a and 8.3b. Medical pictures including examinations can be challenging to understand. Filtering techniques have been used to enhance image quality, reduce noise, and keep an image's edges.[9] Preprocessing is therefore necessary to raise the quality of input image. The mammography will be ready for the following two processes, that is, segmentation and extraction of characteristics.

8.4.1.1 BACKGROUND ERASURE

The illustration was first binarized with a threshold of 0.1, and the associated parts were subsequently put together in decreasing order to extract the biggest blob, which is the female breast profile but is made up of parietal muscle.



(a) (b)

FIGURE 8.3 Image preprocessing: (a) is an original mammogram and (b) is an image after median filtering.

8.4.1.2 PECTORAL MUSCLE SUPPRESSION

The next phase involved the use of a modified region-growing approach to shrink the pectoral muscle. An approach for segmenting imagery is referred to as "seeded region growth." It acts in two ways: first, by choosing a pixel's locational value, and second, by determining a seed point. The seed point can be either individually or adaptively chosen.

8.4.2 EXTRACTING FEATURES

The most accurate method for recognizing breast cancer is through computer-aided extraction of characteristics from mammography pictures. Geometric, Texture, and Gradient features make up each of the three basic types of a feature.

8.4.2.1 GEOGRAPHICAL FEATURES

The geometric parameters of the range of interest (ROI) have been defined by geometric features. It is shown as a group of pixels in a picture.

8.4.2.2 TEXTURE DETAILS

The term "texture" can be extremely confusing yet frequently used to refer to how a (woven) cloth feels or looks to the human eye. Everyone has a unique understanding of what texture is and how it works. Although there is no official definition of texture, this description intuitively offers measurements of qualities like smoothness, coarseness, and regularity. Texture characteristics are an effort to describe the differences in grayscale between neighboring pixels in the image.

8.4.2.3 GRADIENT FEATURES

The local image values serve as the basis for the gradient image. A greater value in the gradient image would correlate to an edge in the original image. CNN is a potent computational image processing approach. The components that appear in a picture are identified using this approach.

8.4.2.4 CLASSIFICATION BY USING CNN ALGORITHM

Classification using a CNN (Convolutional Neural Network) algorithm in breast cancer involves developing a neural network model to acquire knowledge automatically and classify breast cancer images or features. CNNs are powerful deep learning models that excel at image classification tasks, making them well suited for analyzing breast cancer data.

Here is a step-by-step process for classifying breast cancer using a CNN algorithm:

- Dataset Preparation: Gather a dataset of breast cancer images, which should be labeled with their corresponding class or category (e.g., malignant, or benign). Ensure that the dataset is diverse, representative, and properly annotated.
- Data preprocessing: Make sure the breast cancer photos are in a format that will work for the CNN model during training. To improve the dataset size and variability, this may entail scaling the photographs, standardizing the pixel values, and enhancing the data by applying transformations like rotation, flipping, or zooming.
- Model Architecture Design: Create the CNN model's architecture. To extract hierarchical information from the images, a stack of convolutional layers is often followed by pooling layers. After that, a layer or layers with full connections are added for classification. VGGNet, ResNet, or InceptionNet are a few common CNN architectures that are employed in the categorization of breast cancer.
- Divide the dataset into training and validation sets for the model. Input the training data into the CNN model, and then use backpropagation

and optimization algorithms (such as stochastic gradient descent) to iteratively update the model's weights. Keep track of how the model performs on the validation set and make any necessary adjustments to the hyper parameters (such as learning rate and batch size) to increase the model's precision.

- Analyze the trained CNN model using a different testing dataset that was not utilized during training after the training is over. Examine parameters like accuracy, precision, recall, and F1-score to judge how well the model categorizes breast cancer.
- Fine-tuning and Optimization: To further enhance the model's accuracy and generalization skills, fine-tuning it by changing the hyper parameters, altering the architecture, or adopting strategies like regularization or dropout.
- Use the trained CNN model to categorize fresh, unused breast cancer photos or samples in a real-world setting. Keep an eye on the model's performance and collect input to iteratively enhance it as needed.

It is worth noting that training a CNN model for breast cancer classification typically requires a significant amount of labeled data, computational resources, and expertise in deep learning. Collaborating with medical professionals and experts in the field can help ensure accurate interpretation and clinical relevance of the model's predictions.

8.5 RESULTS

By observing today's system which provides results in good manner but still there is some drawback, to improve today's system by adding some new features in an efficient way, and make our system more user friendly, we have implemented our system with CNN algorithm by following the stages below:

CNN implementation steps:

- Step 1: Convolution Operation (Filter image).
- Step 1(b): ReLU Layer.
- Step 2: Pooling (used max pooling function).
- Step 3: Flattening (Convert Matrix into 1D Array).
- Step 4: Full Connection.
- Step 4(b): Dense ().
- Step 4(c): Optimizer ().
- Step 4(d): Compile ().

FIGURE 8.4 Input image.

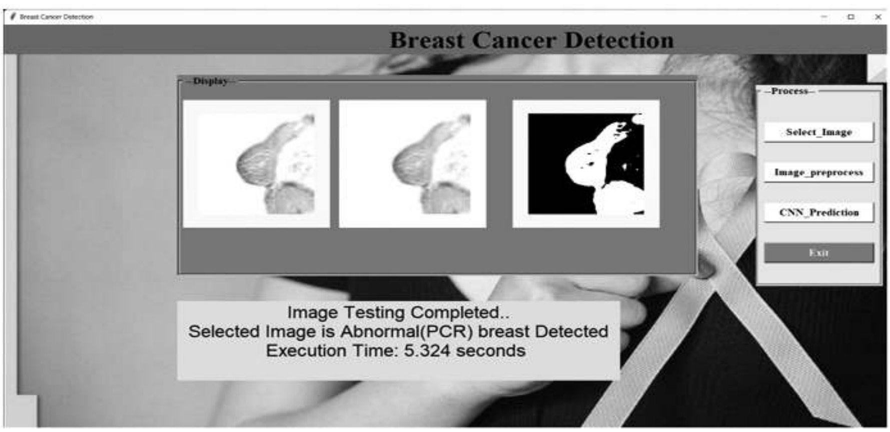


FIGURE 8.5 Detection of breast cancer.

8.6 ADVANTAGES

- Early identification of breast cancer can significantly increase the prognosis and survival prospects by encouraging patients to receive therapeutic therapy as soon as feasible.

- Early identification of breast cancer is crucial because it frequently means the patient will have access to treatments that are far less aggressive yet retaining remarkably effective outcomes.
- Early Detection: Risk detection methods, such as mammograms, genetic testing, and clinical evaluations, can help identify breast cancer risk factors at an early stage. This allows for timely intervention and treatment, potentially improving outcomes and survival rates.
- Personalized Risk Assessment: These methods can provide individualized risk assessments based on factors like family history, genetic predisposition, and lifestyle choices. This personalized information can help both individuals and healthcare professionals make educated decisions regarding screening, prevention strategies, and treatment plans.
- Improved Survival Rates: Detecting breast cancer risks early can lead to more successful treatment outcomes. Identifying high-risk individuals may prompt more frequent screenings, monitoring, and preventive measures, potentially reducing the mortality rate associated with breast cancer.
- Psychological Empowerment: Knowing one's breast cancer risk status can empower individuals to take proactive steps toward risk reduction and early detection. It allows them to engage in lifestyle modifications, regular screenings, and informed decision-making about potential preventive measures, fostering a sense of control over their health.

8.7 LIMITATIONS

- Most of the study subjects have benign tissues, which is a problem for mammography analysis. For the creation of a model that distinguishes malignancy, this supported an unbalanced data structure.[9]
- Because of a lack of experience or due to technical difficulties, radiologists may choose to disregard some lesions. Due to incomplete mammography data or to avoid potential legal issues in the event of a false positive diagnosis, the examinees may occasionally be put into a category for further recall for further examination. These circumstances might prevent an early intervention.[9]
- Despite the enormous advancements in breast cancer detection technologies, there are still some restrictions and difficulties that

researchers and practitioners must overcome. Here are a few typical restrictions related to breast cancer detection: Positive and negative false results: Methods for finding breast cancer, such mammography, or biopsies, may result in false positives or false negatives, any of which could overlook the existence of cancer. These mistakes may cause patients to worry needlessly or postpone diagnoses.

- Sensitivity to Breast Density: Mammography, the most commonly used screening tool, may be less sensitive in women with dense breast tissue. Dense breasts can mask CANVER, making them more difficult to detect. Additional screening techniques or alternative imaging methods may be required for women with dense breasts.
- When more testing is necessary, there may be waiting and worry periods.
- There may also be instances of over diagnosis.
- False Positives and Overdiagnosis: Some risk detection methods, like mammograms, can produce false-positive results, leading to unnecessary anxiety, additional testing, and even invasive procedures. Overdiagnosis is another concern, where a cancerous condition that may not have caused harm during a person's lifetime is identified, leading to overtreatment.
- False Negatives: Risk detection methods may also produce false-negative results, providing a false sense of security. This can lead to delayed diagnosis and treatment, potentially compromising outcomes.
- Cost and Accessibility: Certain risk detection methods, such as genetic testing or advanced imaging techniques, can be expensive. Not all individuals have access to these tests due to financial constraints, limited healthcare resources, or geographical factors, resulting in disparities in risk detection and management.
- Psychological Burden: Learning about one's increased risk of breast cancer can cause significant emotional distress, anxiety, and fear. This psychological burden may impact the quality of life and well-being of individuals, requiring additional support and counseling.
- Ethical and Privacy Concerns: Genetic testing for breast cancer risk may reveal information about an individual's genetic makeup and potential risks for other health conditions. This raises ethical dilemmas regarding privacy, discrimination, and the potential for psychological harm resulting from the disclosure of sensitive information.
- It is crucial to remember that the benefits and drawbacks listed above are broad considerations. The specific benefits and drawbacks of breast cancer risk detection may vary based on individual circumstances,

such as age, family history, and overall health. It is recommended to consult with healthcare professionals to determine the most appropriate screening strategies based on individual risk factors.

8.8 APPLICATIONS

- Three-dimensional mammography, sometimes referred to as breast tomosynthesis, is one technical advancement.
- The development of imaging technology has enabled progress in early detection and screening.
- One technological advancement is 3-D mammography, also called breast tomosynthesis.
- Early detection: Breast cancer risk detection methods such as mammography, clinical breast examination, and breast self-examination play an important role in the detection of early breast cancer. Early detection allows for timely intervention, increases the likelihood of successful treatment, and improves overall prognosis.
- Personalized Screening Guidelines: Breast cancer risk assessment tools, such as the Gail model or the Tyrer-Cuzick model, help estimate an individual's risk of developing breast cancer. These risk assessment models consider factors such as age, family history, reproductive history, and genetic predisposition. The results can guide healthcare providers in recommending personalized screening guidelines, including the frequency and type of screenings.
- Genetic Testing: Genetic tests, such as the BRCA1 and BRCA2 genetic tests, can identify specific genetic mutations associated with an increased risk of breast cancer. Individuals who carry these mutations can take proactive measures to manage their risk, such as increased surveillance or preventive surgeries like prophylactic mastectomy or oophorectomy.
- Prevention and Risk Reduction: Breast cancer risk detection helps identify high-risk individuals who may benefit from risk reduction strategies. For example, people with a strong family history of breast cancer or certain genetic mutations may be advised to undergo risk-reducing interventions such as chemoprevention (use of drugs to reduce risk) and lifestyle changes (e.g., maintaining a healthy weight, exercising regularly, and limiting alcohol consumption).
- Counseling and Education: Breast cancer risk detection provides an opportunity for healthcare professionals to offer counseling and

education to individuals at risk. This includes discussing the significance of risk factors, explaining the screening process, addressing concerns and anxieties, and providing information on available support services and resources.
- Research and Development: Breast cancer risk detection plays a vital role in research and development efforts aimed at improving screening techniques, risk assessment models, and treatment options. By studying risk factors and their impact on breast cancer development, researchers can identify new biomarkers, refine risk prediction models, and develop more targeted and personalized approaches to breast cancer prevention and treatment.

Overall, breast cancer risk detection has far-reaching applications in early detection, personalized screening, risk reduction, counseling, and research. It empowers individuals and healthcare professionals to take proactive steps in managing breast cancer risk and improving patient outcomes.

8.9 RISK FACTOR

- Cancer in the family history of the breast.
- Alcohol, animal fat, and cigarette use are all bad habits.
- Obesity and previous experiences with body cancer.

8.10 SYMPTOMS AND PREVENTION

8.10.1 SYMPTOMS

The following are the symptoms of breast cancer, and some real signs are shown in Figure 8.6.

- Breast or armpit lump that has recently appeared.
- Swelling or thickening of the breast region.
- Skin inflammation or dimples on the breasts.
- Flaking skin or redness in the breasts or around the nipple.
- Pulling in the nipple area or soreness there.
- Besides breast milk, other nipple discharges include blood.
- A change in breast size or shape.
- Pain can occur in any area of the chest.

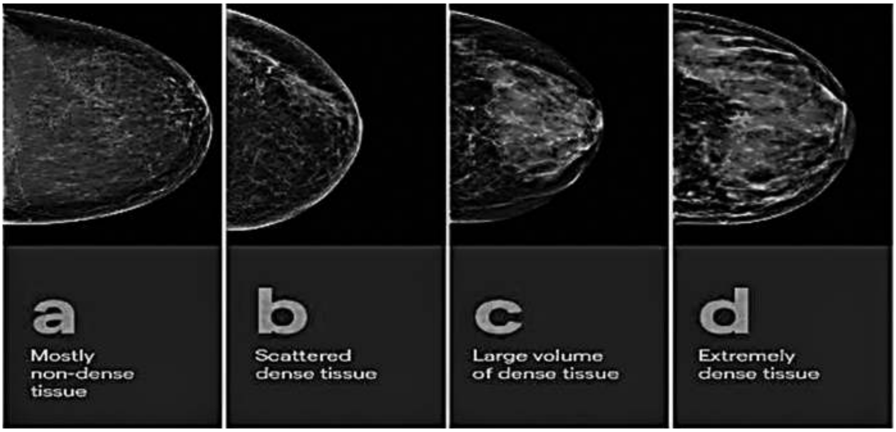


FIGURE 8.6 Real sign.

8.10.2 PREVENTION

- Get regular exercise.
- Limit using hormone replacement therapy.
- Keep your weight under control.
- Eat well-balanced meals.
- Keep an eye out for any changes to your breasts.
- Limit your exposure to pollutants.
- Avoid using tobacco products.
- Get a mammogram.

8.11 CONCLUSIONS

A reliable technique for detecting breast cancer is created in this study to help pathologists to do their jobs. The suggested approach automatically separates the image's regions into cancerous and noncancerous areas. If picture quality is kept constant throughout laboratories and difficulties with intra- and inter-observation variability among specialists are reduced, accuracy can still be raised.

It has been found that using deep learning and convolutional neural networks to classify breast cancer picture data produces superior outcomes for the linear breast data that is taken from the photos. For the most part, CNN uses the image dataset. The previous study likewise concluded that CNN performs better than machine learning methods.

KEYWORDS

- **breast cancer**
- **MRI image**
- **CNN**
- **filtering**
- **edge detection**

REFERENCES

1. Nguyen, P. T.; Nguyen, T. T.; Nguyen, N. C.; Le, T. T. In *Multiclass Breast Cancer Classification Using Convolutional Neural Network*, 2019 International Symposium on Electrical and Electronics Engineering (ISEE), 2019. DOI: 10.1109/isee2.2019.8920916
2. Arslan, A. K.; Yasar, S.; Colak, C. In *Breast Cancer Classification using a Constructed Convolutional Neural Network on the Basis of Histopathological Images by an Interactive Web-Based Interface*, 2019 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), 2019. DOI: 10.1109/ismsit.2019.8932942
3. Basunia, M. R.; Basunia, M. R.; Pervin, I. A.; Pervin, I. A.; Al Mahmud, M.; Al Mahmud, M.; … Arifuzzaman, M. In *On Predicting and Analyzing Breast Cancer using Data Mining Approach*, 2020 IEEE Region 10 Symposium (TENSYMP), 2020. DOI: 10.1109/tensymp50017.2020.923
4. Bhangu, K. S.; Sandhu, J. K.; Sapra, L. In *Improving Diagnostic Accuracy for Breast Cancer using Prediction-Based Approaches*, 2020 Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC), 2020. DOI: 10.1109/pdgc50313.2020.9315815
5. Padhi, T.; Kumar, P. In *Breast Cancer Analysis Using WEKA*, 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence), 2019. DOI: 10.1109/confluence.2019.8776911
6. Kabiraj, S.; Raihan, M.; Alvi, N.; Afrin, M.; Akter, L.; Sohagi, S. A.; Podder, E. In *Breast Cancer Risk Prediction using XGBoost and Random Forest Algorithm*, 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 2020. DOI: 10.1109/icccnt49239.2020.9225451
7. Nagpure, R.; Chandak, S.; Pathak, N. In *Breast Cancer Detection using Neural Network Mammogram*, 2020 International Conference on Convergence to Digital World - Quo Vadis (ICCDW), 2020. DOI: 10.1109/iccdw45521.2020.9318635
8. Li, M. In *Research on the Detection Method of Breast Cancer Deep Convolutional Neural Network Based on Computer Aid*, 2021 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC), 2021. DOI: 10.1109/ipec51340.2021.9421338

9. Le, T. T.; Fu, W.; Moore, J. H. Scaling Tree-Based Automated Machine Learning to Biomedical Big Data with a Feature Set Selector. *Bioinformatics* **2019**. DOI: 10.1093/bioinformatics/btz470
10. Lu, H. C.; Loh, E. W.; Huang, C. C. In *The Classification of Mammogram Using Convolutional Neural Network with Specific Image Preprocessing for Breast Cancer Detection*, 2nd International Conference on Artificial Intelligence and Big Data, 2019. 978-1-7281-0831-5/19/$31.00 ©2019.
11. Al-Shamlan. H. In *Feature Extraction Values for Breast Cancer Mammography Images*, International Conference on Bioinformatics and Biomedical Technology, Apr 2010.
12. Ramani, R.; Suthanthira Vanitha, N.; Valarmathy, S. The Pre-Processing Techniques for Breast Cancer Detection in Mammography Images. *Int. J. Image Graph. Signal Process.* **2013.**
13. Krizhevsky, A.; Sutskever, I.;Hinton, G. E. ImageNet Classification with Deep Convolutional Neural Networks. In *Advances in Neural Information Processing Systems*, 2012; pp 1097–1105.
14. Nallamala, S. H.; Mishra, P.; Koneru, S. V. Breast Cancer Detection using Machine Learning Approaches. *Int. J. Recent Technol. Eng.* 2019, *7* (5), 478–481.
15. Hu, Q.; Whitney, H. M.; Giger, M. L. A Deep Learning Methodology for Improved Breast Cancer Diagnosis using Multi-Parametric MRI. *Sci. Rep.* 2020, *10*, 1–11.
16. Nilashi, M.; Ahmadi, N.; Samad, S.; Shahmoradi, L.; Ahmadi, H.; Ibrahim, O. Disease Diagnosis Using Machine Learning Techniques: A Review and Classification. *J. Soft Comput. Decis. Support Syst.* **2020,** *7* (1), 19–30.
17. Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R. L.; Torre, L. A.; Jemal, A. Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries," CA. *Cancer J. Clin.* **2018,** *68* (6), 394–424.
18. Boroumandzadeh, M.; Parvinnia, E. ‘Automated Classification of BI-RADS in Textual Mammography Reports. *Turkish J. Electr. Eng. Comput. Sci.* **2021,** *29* (2), 632–647.
19. Cifci, M. A.; Aslan, Z. In *Deep Learning Algorithms for Diagnosis of Breast Cancer with Maximum Likelihood Estimation*, Proceedings of the International Conferences on Computer Science and Application, Springer, 2020; pp 486–502.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

PART II
COMMUNICATION TECHNOLOGIES

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

CHAPTER 9

A COMPREHENSIVE STUDY ON GRAPHS, DESIGN PROCESS, AND MAPPING TECHNIQUES BASED ON EVOLUTIONARY ALGORITHM FOR NETWORK ON CHIP ARCHITECTURE

KALIPRASANNA SWAIN[1], A. S. DAS[2], SARITA MISRA[1], V SINGH[3], TARAPRASANN DAS[4], and SUBHASHREE CHOUDHURY[5]

[1]*Department of Electronics and Communication Engineering, GITA Autonomous College, Bhubaneswar, India*

[2]*J Systems and Control, Cuttack, India*

[3]*Department of Computer Science Engineering, GITA Autonomous College, Bhubaneswar, India*

[4]*Department of Electrical and Communication Engineering, Siksha 'O' Anusandhan (Deemed to be University), Bhubaneswar, India*

[5]*Department of Electrical and Electronics Engineering, Siksha 'O' Anusandhan (Deemed to be University), Bhubaneswar, India*

ABSTRACT

Continuous demand for high-performance embedded systems driving the system on chip (SoC) market in an uphill trend. High performance requires more processing power and capacity. To achieve more processing power from a single chip, SoC designers are putting more processing elements (PE) onto a single silicon substrate. Putting a greater number of PEs on a single substrate

Applied Soft Computing Techniques: Theoretical Principles and Practical Applications. Samarjeet Borah, Ratna Raja Kumar Jambi, Sharifah Sakinah Syed Ahmed, & Mahendra Prabhakar Deore (Eds.)
© 2025 Apple Academic Press, Inc. Co-published with CRC Press (Taylor & Francis)

requires very high-speed bus and deadlock-free scheduling which are getting more and more complicated day by day. To address this situation, network on chip architecture is adopted in recent years. Network on chip (NoC) in contrast with SoC utilizes an on-chip interconnection network. On-chip interconnection networks facilitate the connection of a greater number of PEs. The on-chip network routing facilitates deadlock-free scheduling and parallel processing. One of the important aspects of NoC is optimal mapping. Optimally mapping the participating cores means low bandwidth requirement, low latency, and hence greater throughput. Optimization of communication bandwidth cost for on-chip networks or bandwidth cost optimization for network on chip is the prime factor for consideration. Soft computing plays a significant role in the design of network on chip (NoC) architectures. NoC is an on-chip communication infrastructure that provides efficient and scalable communication between various components in a system-on-chip (SoC) design. Soft computing techniques, which include fuzzy logic, neural networks, and evolutionary algorithms, can be employed to enhance different aspects of NoC architecture design.

In view of this, this review is presented to discuss the advantages of NoC over SoC technology, recent issues in NoC, types of graphs, mapping of NoC through numerous evolutionary algorithms, NoC design process, and NoC platform. An attempt has been made to comprehensively discuss every aspect of NoC regarding various optimization techniques for the process of mapping. Finally, a genetic algorithm is proposed here for the optimization of NoC performance by employing butterfly fat tree topology.

9.1 INTRODUCTION

A system is said to be efficient if all its subsystems work efficiently. Modern microprocessors and multicore chips are getting more complicated every day. An increase in complications directly affects communication between cores. Inefficient communication between cores results in an inefficient subsystem, which leads to an inefficient system. To increase the efficiency of multicore chips system on chip (SoC) concept was introduced. Due to bus-based architecture, SoC suffers scalability. Network on chip (NoC) is a concept that addresses the shortcomings of SoC.

System on Chip: System on chip (SoC) is a specific purpose subsystem in an embedded system or a larger system. A typical SoC may contain:

- A microcontroller, microprocessor, DSP core(s)

- Some SoC may have more than one processing element; they are called MP SoC
- Some memory blocks including ROM, RAM, EPROM, and Flash memory
- Clock generators or PLL
- Peripherals like power-on reset, real-time timers, counters etc.
- External interfaces like USB, I2C, SPI, USART etc.
- Analog interfaces ADCs and DACs
- Power conditioners and regulators.

Network on Chip: Network on chip (NoC) is the recent trend in the industry. This technique utilizes the concept of the network instead of the bus for communication between different cores of the subsystem. NoC is an on-chip communication infrastructure that interconnects different intellectual properties (IP)s of the subsystem. Interconnection topologies used in NoC are industry-standard high-performance computing (HPC) interconnection topologies. Interconnection networks overcome the problem of scalability. But nonoptimal positioning of nodes in a network degrades the overall performance. Optimal positioning depends on the communication pattern of nodes. Highly communicating nodes should be placed as neighbors for better performance.

The entire article has been categorized into the following sections. Section 9.2 highlights the research gap in NoC design. In Section 9.3, various types of graphs are discussed in detail. Section 9.4 throws light on the process of mapping. In Section 9.5, a detailed study concerning the NoC platform basics, NoC design process, and various optimization techniques for NoC mapping has been projected. The conclusion from the entire review of literature is presented in Section 9.6.

9.2 RESEARCH ISSUES IN NOC DESIGN

NoC is a recent trend. Several groups from industries to educational institutions are conducting research in different dimensions. These dimensions can be classified into the following categories.

9.2.1 COMMUNICATION INFRASTRUCTURE

It has already been stated NoC is an on-chip interconnection network. Hence it should have a communication infrastructure. This includes network topology, router architecture, buffer optimization, link design, clocking, floor planning, and layout.

9.2.2 COMMUNICATION PARADIGM

NoC infrastructure is a packet switch network. Hence, methods for routing policies, switching techniques, congestion control, power and thermal management, fault tolerance, and reliability should be defined.

9.2.3 EVALUATION FRAMEWORK

A properly designed NoC should have minimum latency, energy/power consumption, and bandwidth of the network. For the validation and simulation of an NoC evaluation framework plays an important role.

9.2.4 APPLICATION MAPPING

Application mapping is the fourth important aspect in the NoC design process. Application mapping means placing different interacting tasks (nodes) in suitable positions, such that overall performance is best. Application mapping influences the overall performance of the NoC. There exists a lot of work on all aspects of NoC design except for application mapping. Hence application mapping has scope for more work.

9.3 GRAPHS

9.3.1 TASK GRAPH

A task graph is a fully connected directed graph. Each node in the task graph is a single unique task that needs to be performed for the successful run of the whole application. Task graph G(V,E) is a directed graph where V represents the set of all nodes present in the graph and E represents the set of all edges. Each edge in the task graph represents a communication between the cores connected by that edge. And the weight of the edge represents the communication bandwidth requirement between nodes connected by the edge.[1] It is mathematically represented as $G(V,E)\ \forall\ v\_i \in V\ \exists\ e\_i,j \in E\ \forall\ i \neq j$. Figure 9.1 shows various types of task graphs.

9.3.2 TOPOLOGY GRAPH

Topology represents a scheme for interconnection. A topology graph is a graph that represents a topology using graph theory terminology. A topology graph is a fully connected graph representing target topology and architecture. Every node in the topology graph represents a router. Depending on the architecture and topology chosen routers can be the same or different. In 2Dmesh topology, all routers are the same. A topology graph is represented by T (U,F)$\forall$ ui$\in$U and f i,j$\in$F $\forall$ i$\neq$ j where U is the set of all vertices in the graph and F is the set of all edges in the topology graph. f i,j is termed bandwidth bwi,j which is the bandwidth needed between nodes u_i and u_j.[1]

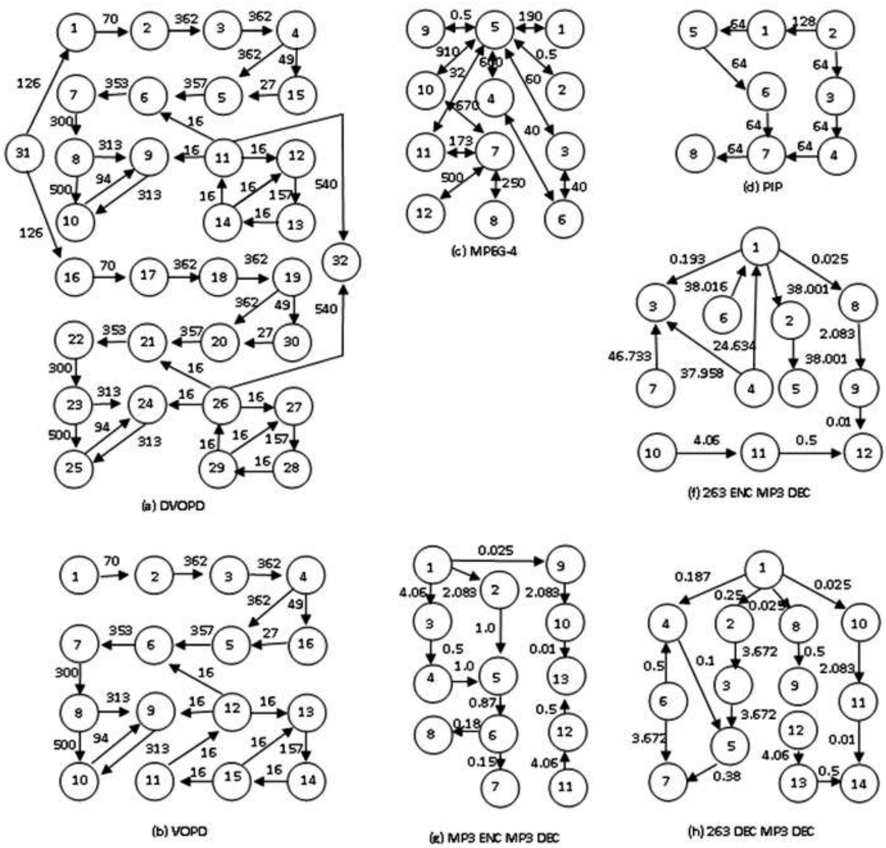


FIGURE 9.1 The exemplary task graphs[2] which are implemented previously.

9.4 MAPPING

Performance topology is such that every node in the task graph is mapped to a unique node in the topology graph.

A mapping between (V, E) and T (U, F) is defined as $map(vi) = u_j \ \forall\ v_i \in V, \exists\ u_j \in U$ and $|V| \leq |U|$. In other words, a map of a core vi to router ui is possible when the number of cores is less than or equal to the number of routers 2Dmesh or number of leaf connectors in a tree-based topology graph. The quality of mapping is defined by the total communication cost of the NoC. The communication flow between 2 cores can be defined as dk, $k = 1,2,3,\ldots,|E|$. If v_i is mapped to u_i and v_j is mapped to u_j then the weight of $e_i, j \in E$ is called comm i,j. Set of all commodities $D = \{dk\ |dk = commi,j$ for k = 1,2,3, $|E|\ ei,j \in E\}$. Maximum bandwidth requirement between 2 routers u_i and u_j is given by

$$x_{i,j}^{k} = \begin{cases} value(d_k)\, if\ f_{i,j} \in path \\ 0 \end{cases} \tag{9.1}$$

Constraint or Objective: Mapping of task graph G to topology graph T should be such that the total bandwidth consumption of all the cores is at minimum. This implies that highly communicating cores should be placed nearer to each other.

Mathematically,

$$cost = \sum_{i=1}^{|E|} bw_{i,j}\ X\ distance(u_i u_j) \ \text{ should be minimum} \tag{9.2}$$

where $distance\ (u_i, u_j)$ is the function that returns the distance between source and sink routers. The unit of distance is hops for mesh and related architecture and cycle count for tree-based architecture.

9.5 NoC MAPPING

The literature survey focuses on different mapping techniques. More attention is given to transformative heuristics such as Ant colony-based optimization (ACO), particle swarm optimization (PSO), and GA-based optimization. Any work related to the mapping of NoC cores cannot be complete without mentioning "Bandwidth Constrained Mapping of Cores onto NoC Architectures" by Srinivasan Murali, Giovanni De Micheli.[1] This research paper introduced NMAP, a constructive heuristic with an iterative improvement algorithm. NMAP is considered a de-facto standard by industries. This paper also proposed a 2DMesh topology for NoC. The

NMAP[1] algorithm starts mapping with the most demanding core to the freest neighbors router in the mesh. After mapping is complete, it optimizes the placement by trying to switch cores. In the end, an optimal map is produced. A modification to address multipath routing is also provided in the same paper. This research work is mostly based on the paper, “A survey on application mapping strategies for Network on Chip design” by P.K Sahu and S. Chattopadhyay, published in Journal of Systems Architecture, Elsevier, volume 59 page (6076).[2] Figure 9.2 shows a 2Dmesh internal that shows routers as the crossbar and connected to routers.

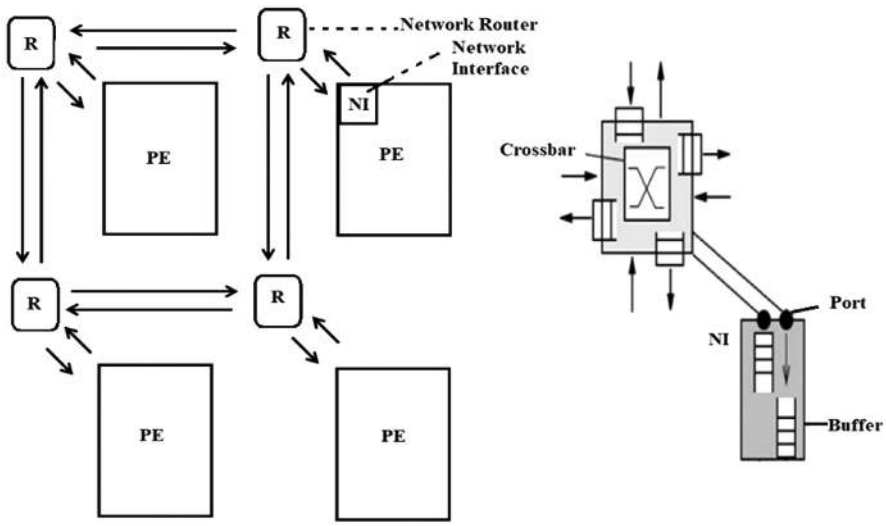


FIGURE 9.2 2D mesh router and NI details.

Source: Adapted from Ref. [3]

9.5.1 NoC PLATFORM BASICS

NoCs are packet-switched multi-hop networks. The packets are traveled from a source core to sink core travel by hopping through several routers. The connectivity and layout of the components are mostly determined by network topologies. Generally, three topologies are mostly used such as grid-based topology or regular network topology, irregular or custom structure-based topology, and tree-based topology. These topologies exploit different characteristics of the application’s traffic patterns such as locality of traffic.[3] Figure 9.2 shows a few details of 2Dmesh interconnection. The network interface (NI) part of the core takes care of communication intricacies.

9.5.2 NOC DESIGN PROCESS

Network on chip (NoC) follows a systematic process from concept to manufacturing. The main objective of NoC design is the reuse of IP core.

9.5.3 MAPPING ALGORITHM

Mapping is the process of placing IP cores of a task graph to a topology graph. Both the task graph and topology graph are fully connected graphs. The mapping of cores to a topology of NoC is NP-hard.[4] NP-hard problems cannot be solved in polynomial time. Hence, there is no polynomial-time algorithm to solve the mapping problem. Authors Carvalho et al.[5] did an extensive study on mapping methods. The authors conclude that dynamic mapping is a partial view and static mapping assumes a full description of the task. There are two possibilities when a core mapping can be done.

9.5.3.1 DYNAMIC MAPPING

Multimedia, networks, or any streaming applications have different processing demands at different times. Until execution, it cannot be decided what the demand will be. This in turn implies that the tasks should be scheduled dynamically. If the cores are mapped at run time that is while the task is executed then it is called dynamic mapping. In the case of dynamic mapping, performance is evaluated and processors are assigned as per the demand by the mapping algorithm. Though this scheme is good it has serious computational overhead, the system has to keep track of all the cores, the total performance of the system, and mapping needs. The authors in ref. [5] suggest using greedy algorithms for dynamic mapping, to reduce the computation overhead.

9.5.3.2 STATIC MAPPING

This process assigns cores before the task is executed, that is at design time. It is done after the task and communication infrastructure are decided. Static mapping tries to place cores onto target topology in such a way that the communication cost is always minimum. The static mapping algorithm is run once at design time which removes the computation overhead at execution time.[6] Static mapping is only considered for this research work. Static mapping (Figure 9.2) can be categorized into 2 categories.[2]

9.5.3.2.1 Exact Mapping

Exact mapping considers mathematical programming. The most primitive and naïve approach would be testing all the n! possibilities (where n is the number of cores) and checking which is the best solution. A still better way is using ILP or MILP for this purpose.

9.5.3.2.2 Search-Based Mapping

Search-based algorithm employs the searching technique to search for an optimal solution in the search space while satisfying the constraints imposed.

Jingcao Hu and Radu Marculescu[7] discussed such a method and claimed it to be better. It is claimed to achieve 60.4% more energy-efficient mapping for video and audio applications. Harmanani and Farah[8] proposed a simulated annealing-based application mapping technique for 2Dmesh. The authors claim their technique minimizes blocking while increasing bandwidth throughput. Lu et al.[9] proposed cluster-based simulated annealing for mapping onto 2Dmesh. As per the authors, clustering exploits locality, connectivity, and distance of the cores. Hence, cluster-aware simulated annealing can dynamically be applied within clusters. This in turn results in a good mapping of cores. Authors Orsila et al.[10] expressed their method using simulated annealing saves 50% time with only 0.3% performance loss. The authors pointed out that cyclic graphs are harder than acyclic graphs. Transformative heuristics transform some random solution into a better solution. The most common examples are evolutionary techniques, such as Genetic Algorithm (GA),[11] Particle Swarm Optimization (PSO),[12] Ant Colony Optimization (ACO),[13] Bee Colony,[14] etc.

A. Genetic Algorithm

Ozcan Ozturk and Dilek Demirbas[15] described an approach for heterogeneous NoC mapping using a Genetic Algorithm and compared it with the ILP result. The authors in their research work considered performance, energy, temperature, area, and communication bandwidth as a constraint. In their work the authors computed computation cost and communication cost, and compared them with ILP results. As per the authors, this is the first work that explores the possibility of employing evolutionary computational techniques for optimally placing the heterogeneous nodes in an NoC.

P. Mesidis and L.S. Indrusiak[3] discussed the use of genetic algorithm-based mapping for hard real-time applications. In their work the authors considered "hard real-time applications running over multicore processors based on wormhole Networks on Chip (NoCs)."

B. Particle Swarm Optimization (PSO)

PSO belongs to a swarm intelligence-based algorithm. In the case of swarm intelligence, individual particles in the swarm communicate among themselves. In this process for every swarm, a local best and a global best are calculated. These pieces of information are transmitted to all particles. Particles use this information and align themselves according to the current best. At the end of every iteration, the best values are checked and updated accordingly.[12] In contrast to GA swarm-based algorithms are superior as intelligence is built into every particle in the swarm. In the case of GA, the hypercube that ultimately provides the best solution moves in the search space. The governing factor for movement is the chance of getting a better solution in each step of crossover and mutation.[11]

PSO-based works are done in reference[16] by Sahu et al. The authors proposed the PSMAP algorithm in this paper and compared the result with standard results. The authors showed, the result obtained is better than earlier results. In their work the authors used different numbers of participating particles in swarms. Their result shows after a certain population size of swarm no perceivable effect is observed. In reference[17] the authors proposed a discrete PSO (DPSO)-based algorithm for MoT. A DPSO is running multiple PSO simultaneously with the deterministic generation of the initial population. In this paper, the authors have compared their results and other standard results available including ILP.

C. Ant Colony Optimization (ACO)

Ant Colony Optimization (ACO) is an emulation of biological ants' behavior. In a particular sense, ants optimally search for their food. While ants search for food, they put pheromone (a chemical substance) on the trail. The strength of the pheromone signifies how viable the path to the food[13] is. The same concept is adopted in the ACO algorithm. The artificially simulated ants put a weight on the solution they found. When all the ants finished their search the best result is obtained.

ACO-based work was done by Wang et al.[4] In their work the authors focused on 2Dmesh and XY routing. The authors also expressed mapping problem is NP-Hard, and such problems are better solved using heuristic methods. The authors show resulting mapping performs very good in comparison to the random mapping method.

D. Bee Colony

In a bee colony, worker bees go for searching of foods (Honey). When an individual bee finds a food source, it returns to the hive and starts dancing (wiggling and turning).[14,18] This dancing pattern has encoded messages about the source of food in terms of its distance and viability. Then some other worker bees follow the instruction and go for verification. The most viable solution is judged when most of the worker bees perform the same dance.

L. Pan et al.[19] used a chaotic enhanced bee colony method for multi-objective application mapping. As per the authors mapping is NP complete. Heuristic algorithms can better approximate the solution. The constraints considered by the authors are power consumption and wire delay. The authors introduced chaos to prevent premature convergence of the algorithm. For chaos introduction to prevent premature convergence problem the authors used one-dimensional chaos mechanism. The authors conclude Enhanced Chaos Discrete Artificial Bee Colony performs better than PSO and GA-based methods in both run time and mapping result.

9.6 IMPLEMENTATION OF GENETIC ALGORITHM ON BUTTERFLY FAT TREE TOPOLOGY

BFT is a tree-based NoC topology. It is a recursive inter-connection network for NoC. Pande et al. [20] proposed this to be an alternative inter-connection network for NoC. This architecture is proposed to address shortcomings of SoC design. The authors in reference[21] described and compared some other interconnection networks. It is found that under uniform traffic BFT provides a lower throughput. Localized traffic is restricted in a subtree; hence, other communications are not hampered. Sahu et al.[22] have done extensive work on mapping onto BFT topology.

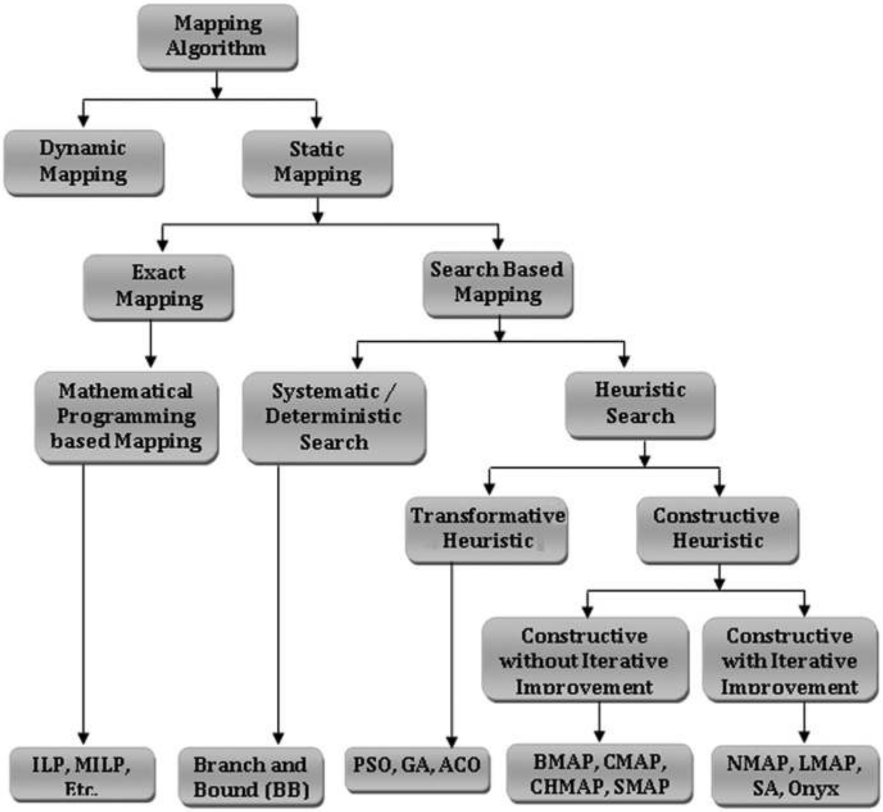


FIGURE 9.3 Classification of mapping algorithms.[2]

9.6.1 BFT STRUCTURE

In a 16 core BFT, the circles represent cores, and the rectangles represent routers. It has three levels where the lowest level is the leaf router level. All the cores are connected to Leaf Routers. Each Leaf Router connects to 4 cores in the child level and connects to 2 higher level routers.

In BFT each router has 2 types of connection, 4 child connections and 2 parent connections.[21] This proposes for N cores; there are $\lceil N/4 \rceil$ leaf routers. In any level there are 1/2 the number of routers of level below. Hence, the total number of routers S in a BFT with N cores will be:

$$S = \frac{N}{4} + \left(\frac{1}{2} \times \frac{N}{4}\right) + \left(\frac{1}{2} \times \frac{1}{2} \times \frac{N}{4}\right) + \left(\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \frac{N}{4}\right) + \ldots + \left(\frac{1}{2}\right)^{levels} \times \left(\frac{N}{4}\right)$$

or $$S = \frac{N}{4}\sum_{i=0}^{levels}\left(\frac{1}{2}\right)^{level}$$

or $$S = \left(\frac{N}{4}\right)\left(\frac{1-\left(\frac{1}{2}\right)^{levels}}{1-\frac{1}{2}}\right)$$

If number of cores (N) is arbitrarily large then level is arbitrarily large. Applying the limits we get

$$S = \lim_{level\to\infty}\frac{N}{4}\left(\frac{1-\left(\frac{1}{2}\right)^{levels}}{1-\frac{1}{2}}\right) = \frac{N}{2} \tag{9.3}$$

From the above, it is clear if the number of cores (N) is very large then the number of routers (S) comes close to $N/2$. Here some properties of the BFT interconnection network are presented. In the network, cores are placed at leaves and routers at vertices. For N number of cores ($N = 2^i$ where i = 3,4,5,6…),

Diameter: $2\times\left\lceil\left(\frac{log_2 N}{2}\right)\right\rceil - 2$

Bisection width: $N\times 0.5^{[\frac{log_2 N}{2}]}$ for i is even

$(\frac{N}{2})\times 0.5^{[\frac{log_2 N}{2}]}$ for i is odd

Routers needed: $(\frac{N}{2})\times(1-0.5)^{[\frac{log_2 N}{2}]}$

Node degree: 6 for non-root, 4 for root routers.

9.6.2 SIMULATION RESULT

For any source core to destination core there exist multiple paths. If both the cores are connected to the same Leaf Router, then there are no issues. But if the source and destination routers are on different leaf routers, then the packet has to travel to least common ancestor (LCA) of both the leaf routers. From LCA there exists a unique path from source to destination router.

In case of BFT the distance is calculated as number of cycles (cycle count). Cycle count represents number of cycles required to transport a message packet from source core to destination core. As per the router design every router takes 2 cycles to receive and transmit the packet. This means in the first cycle the router receives the data packet and in the next cycle it forwards to the next router. But if a router has only 2 connections, then it acts as a FIFO. This implies that the packet is received and forwarded in the same cycle. Hence for FIFO cycle count is 1.

By considering the environment setup (Intel core i3, 2.10 GHz, 8th generation, 2GB RAM, GNU C compiler), the distance function is modified according to the routing algorithm. For the determination of a path a helper function is developed to calculate LCA. Once LCA is determined then the number of routers from source to destination for shortest path can be determined and cycle count can be calculated. For optimization purpose cycle count is multiplied with bandwidth requirement between corresponding cores.

Our findings are presented in Table 9.1. And, it is compared with the result of PSMAP. It shows that our result is consistent with other published results. Only in the case of 263DecoderMp3decoder our result is higher (4.7%) than the published results. In the case of DVOPD our result is better than KLMAP and Simple PSO. A normalized cost graph is shown in Figure 9.4.

TABLE 9.1 Communication Cost Comparison for BFT.

Task graph	Communication cost (Cycle x bandwidth)			Normalized to our result	
	Our result	KLMAP	Simple PSO	KLMAP	Simple PSO
PIP	1536.0	1536.0	1536.0	1.0	1.0
MWD	3264.0	3264.0	3264.0	1.0	1.0
MPEG-4	10,144.0	10,144.0	10,144.0	1.0	1.0
MP3ENCDEC	41.712	41.782	41.782	1.0	1.0
263enc-mp3dec	561.876	561.876	561.876	1.0	1.0
263dec-mp3dec	44.180	42.180	42.180	0.955	0.955
VOPD	10,498.0	10,498.0	10,498.0	1.0	1.0
DVOPD	24,806.0	24,982.0	25,126.0	1.007	1.013

From Figure 9.5, it can be observed that the curve is some exponential function of number of cores. It is also found out the curve fits nicely to the function:

$$CPU\,Time = a \times e^{b \times e^{(c \times cores)}}$$

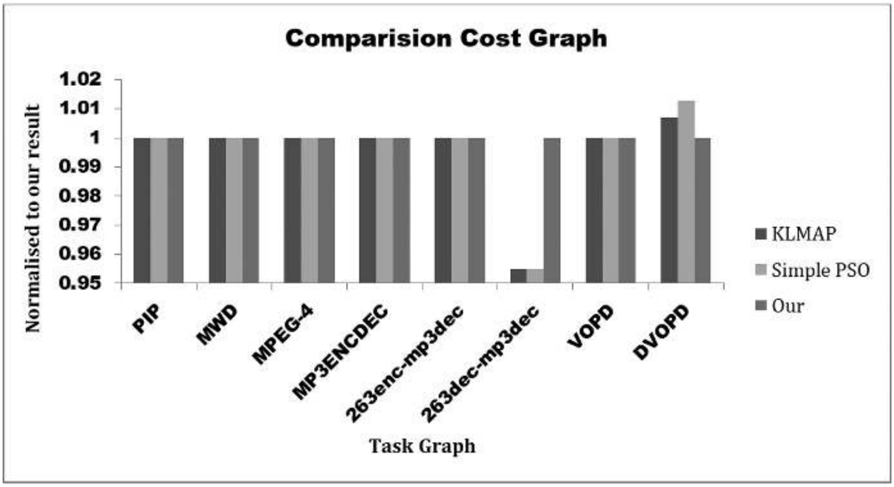


FIGURE 9.4 Normalized cost for BFT of our work with KLMAP and simple PSO.

TABLE 9.2 Number of Cores, Crossover Rate, Mutation Rate, and Result and CPU Time.

Task graph	Number of cores	Population	Crossover rate (Cr)	Mutation rate (Mr)	Result (cycle x bandwidth)	Time (second)
PIP	8	128	0.1	0.1	1536.0	0.288
MWD	12	128	0.1	0.4	3264.0	1.224
MPEG-4	12	128	0.1	0.5	10,144.0	1.459
263enc-mp3dec	12	128	0.7	0.3	561.876	1.373
MP3ENCDEC	13	128	0.1	0.1	41.782	1.518
263dec-mp3dec	14	128	0.1	0.1	44.180	2.073
VOPD	16	128	0.1	0.2	10,498.0	3.168
DVOPD	32	128	0.8	0.5	24,806.0	69.932
G1	32	128	0.1	0.7	61,528.267	537.478
G2	128	128	0.1	0.3	98,5922.125	3230.297

9.7 CONCLUSION

This review article focusses majorly on the NoC technology and the application of evolutionary techniques in NoC mapping. The issues related to SoC technology and the merits of NoC over it have been discussed clearly. Further, the recent problems associated with NoC have been projected. Different types of graphs have been also elaborated that is required for the mapping process of NoC. The article also throws light on NoC design process and NoC platform. Importance of mapping process and the mapping of NoC through numerous evolutionary algorithms has been broadly highlighted. An attempt has been made to comprehensively discuss every aspect of NoC regarding various optimization techniques for the process of mapping. At last, BFT topology is implemented for mapping onto NoC with our mapping, and it is found that our method is produceing a comparable result with the previous result.

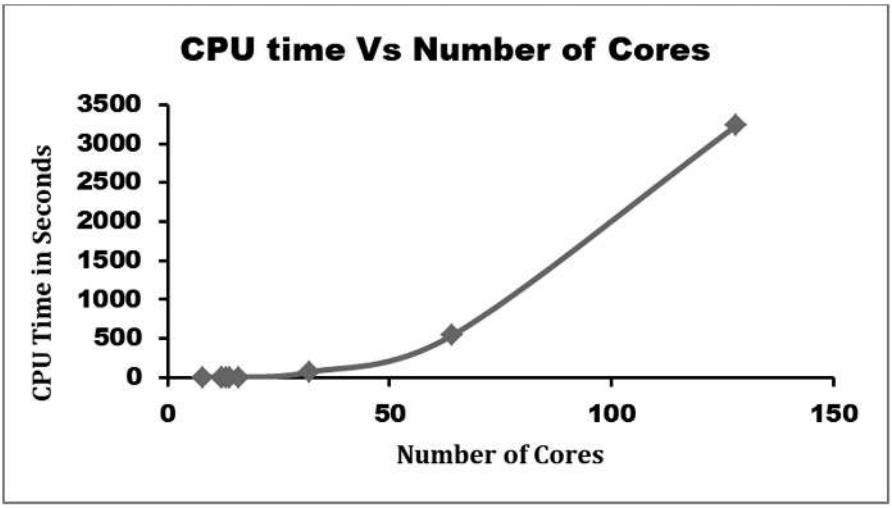


FIGURE 9.5 Relation of CPU time with number of cores.

KEYWORDS

- **network on chip (NoC)**
- **system on chip (SoC)**
- **mapping**
- **optimization**
- **processing elements (PE)**
- **scheduling**
- **butterfly fat tree**

REFERENCES

1. Murali, S.; Micheli, G. De. In *Bandwidth-Constrained Mapping of Cores Onto NoC Architectures*, Proceedings of the Conference on Design, Automation and Test in Europe, 2004, 2, Washington, DC, USA, pp. 20896–20902.
2. Sahu, P. K.; Chattopadhyay, S. A Survey on Application Mapping Strategies for Network-on-Chip Design. *J. Syst. Archit.* **2013,** *59,* 60–76.
3. Mesidis, P.; Indrusiak, L. S. In *Genetic Mapping of Hard Real-time Applications onto NoC Based MP SoCs; A First Approach*, 2011 6th International Workshop on Reconfigurable Communication-centric Systems-on-Chip (ReCoSoC), 2011, pp. 1–6.
4. Amin, W.; Hussain, F.; Anjum, S. iHPSA: An Improved Bio-inspired Hybrid Optimization Algorithm for Task Mapping in Network on Chip. *Microprocess. Microsyst.* **2022,** *90,* 104493.
5. Carvalho, E.; Marcon, C.; Calazans, N.; Moraes, F. In *Evaluation of Static and Dynamic Task Mapping Algorithms in NoC-based MPSoCs*, International Symposium on System-on-Chip, 2009, pp. 087–090.
6. Kadri, N.; Koudil, M. Multi-objective Biogeography-based Optimization and Reinforcement Learning Hybridization for Network-on Chip Reliability Improvement. *J. Parallel Distr. Com.* **2022,** *161,* 20–36.
7. Hu, J.; Marculescu, R. In *Energy-aware Mapping for Tile-based NoC Architectures Under Performance Constraints*, Proceedings of the Asia and South Pacific Design Automation Conference, New York, NY, USA, 2003, pp. 233–239.
8. Harmanani, H. M.; Farah, R. In *A method for Efficient Mapping and Reliable Routing for NoC Architectures with Minimum Bandwidth and Area,* 2008 Joint 6th International IEEE Northeast Workshop on Circuits and Systems and TAISA Conference, NEWCAS-TAISA, 2008, pp. 29–32.
9. Lu, Z.; Xia, L.; Jantsch, A. In *Cluster-based Simulated Annealing for Mapping Cores Onto 2D Mesh Networks on Chip*, Proceedings of the 2008 11th IEEE Workshop on Design and Diagnostics of Electronic Circuits and Systems, Washington, DC, USA, 2008, pp. 1–6.
10. Orsila, H.; Salminen, E.; Hamalainen, T. D. In *Parameterizing Simulated Annealing for distributing Kahn Process Networks on Multiprocessor SoCs*, International Symposium on System-on-Chip SOC, 2009, pp. 019–026.
11. Ozturk, O.; Demirbas, D. Heterogeneous Network-on-chip Design Through Evolutionary Computing. *Int. J. Electron.,* **2010,** *97,* 1139–1161.
12. Pascal, M.; Onwuchekwa, D.; Obermaisser, R. Adaptive Scheduling for Time-Triggered Network-on-Chip-Based Multi-Core Architecture Using Genetic Algorithm. *Electronics* **2021,** 49.
13. Raman, S.; krishnan, G.; Pati, D. Energy Consumption and Performance Comparison of DE Optimization and PSO-based IP-Core Mapping Technique for 2D and 3D Network-on-Chip. *Semiconduct. Sci. Technol.* **2021,** *36,* 085015.
14. Fang, J.; Tingwen Y.; Zelin, W. Improved ant Colony Algorithm Based on Task Scale in Network on Chip (NoC) Mapping. *Electronics* **2019,** *9.*
15. Amin, W.; et al. Performance Evaluation of Application Mapping Approaches for Network-on-chip Designs. *IEEE Access* **2020,** *8,* 63607–63631.
16. Sahu, P. K.; Venkatesh, P.; Gollapalli, S.; Chattopadhyay, S. Application Mapping onto Mesh Structured Network-on-Chip Using Particle Swarm Optimization. **2011,** 335–336.

17. Sahu, P. K.; Sharma, A.; Chattopadhyay, S. In *Application Mapping Onto Mesh-of-Tree Based Network-on-Chip Using Discrete Particle Swarm Optimization*. 2012 International Symposium on Electronic System Design (ISED), 2012, pp. 172–176.
18. Teodorović, D.; et al. Bee Colony Optimization and its Applications. *Handbook of AI-based Metaheuristics*; 2021, pp 301–322.
19. Pan, L.; Li, Z.; Ling, X. NoC Multi-Objects Mapping Based On Enhanced Chaos Discrete Artificial Bee Colony Algorithm. *Int. J. Future Comput. Commun.* **2012,** 116–120.
20. Pande, P. P.; Grecu, C.; Ivanov, A.; Saleh, R. In *Design of a Switch for Network on Chip Applications*, Proceedings of the 2003 International Symposium on Circuits and Systems*,* 2003. ISCAS '03 vol. 5, pp. V–217–V–220.
21. Pande, P. P.; Grecu, C.; Jones, M.; Ivanov, A.; Saleh, R. Performance Evaluation and Design Trade-offs for Network-on-chip Interconnect Architectures*. IEEE Trans. Comput.* **2005,** *54,* 1025–1040.
22. Sahu, P. K.; Shah, N.; Manna, K.; Chattopadhyay, S. An Application Mapping Technique for Butterfly-Fat-Tree Network-on-Chip. **2011,** 383–386.

CHAPTER 10

INTEGRATING IOT DEVICES TO RENEWABLE ENERGY SYSTEMS FOR INCREASING ENERGY EFFICIENCY

V. ULAGAMUTHALVI[1], and G. KULANTHAIVEL[2]

[1]*Sathyabama University, Chennai, Tamil Nadu, India*

[2]*NITTTR, Chennai, Tamil Nadu, India*

ABSTRACT

The main objective of this paper is to increase the energy efficiency of Internet of Things (IoT) devices incorporated into smart environment applications like smart buildings, smart grids, smart cities, etc. In this paper, one of the famous and cost effective renewable energy sources (RESs), such as solar photovoltaic (PV), which generates a sufficient amount of energy from the sunlight, is used. The term sufficient determines whether the energy supply fulfills the demand. The particle swarm optimization (PSO) algorithm optimizes the power demand and supply. The optimal power generation is used to increase energy efficiency in the IoT-based renewable energy framework (IoT-REF), which is implemented in this paper. Many advanced IoT devices in the framework will get the energy directly from the power grid, the battery charged by the solar PV, and indirectly from the solar PV. All the IoT devices interlinked with the smart environment need energy to perform their functionalities. The proposed IoT-REF is simulated in MATLAB software, and the performance is verified and evaluated by comparing its result with the earlier approaches.

Applied Soft Computing Techniques: Theoretical Principles and Practical Applications. Samarjeet Borah, Ratna Raja Kumar Jambi, Sharifah Sakinah Syed Ahmed, & Mahendra Prabhakar Deore (Eds.)
© 2025 Apple Academic Press, Inc. Co-published with CRC Press (Taylor & Francis)

10.1 INTRODUCTION

Renewable energy is energy that does not exhaust after use, such as solar or wind power. It is the energy acquired from processes or sources which are continuously replenished. Different forms of renewable energies are obtained directly or indirectly from wind, Sun, or heat produced deep within the Earth. Renewable sources have dual advantages such as clean energy and green energy. At times there is some minor difference between these two types. In solar energy, sunlight is one of our earth's most accessible and available energy resources. The total amount of energy that reaches the Earth's surface per hour is more than the Earth's energy requirements for a year. Though a good energy source, it will vary with time. It does not create pollution like other forms of energy and is fast-growing worldwide and making tremendous contributions to producing electricity. Turbines are used for power generation. There are many forms of systems used to convert wind energy to electricity. Figure 10.1 shows the implementation process of renewable energy sources (RESs) in smart buildings.

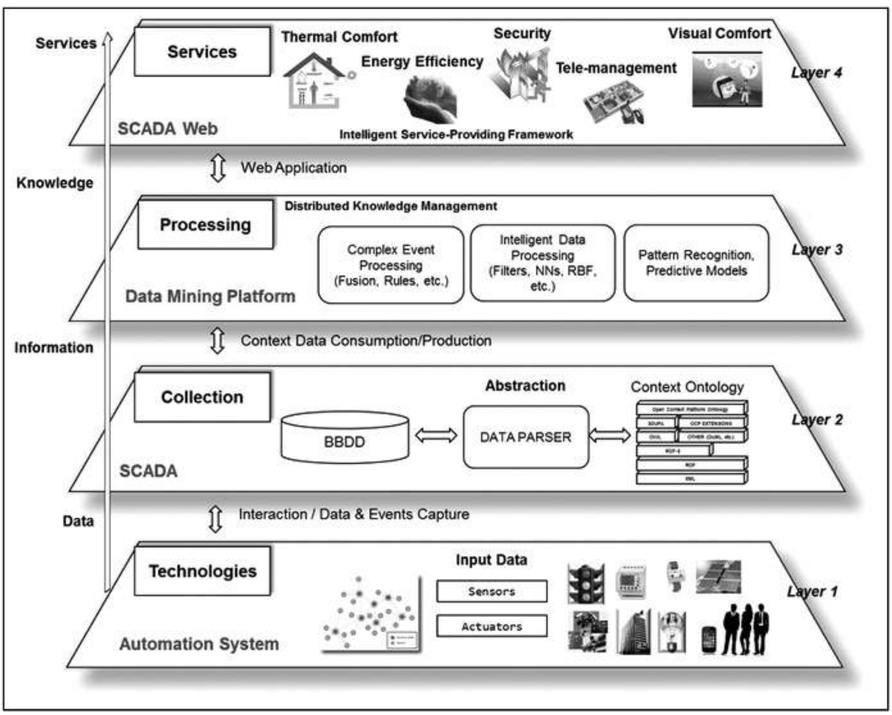


FIGURE 10.1 RES implementation model for smart buildings.

The processes that depend on nonrenewable energy are mining, using lots of machinery, pipelines, and transport. However, renewable energy dispenses the energy as mentioned earlier activities. Most fossil fuels are complex in substance and also expensive to process. Renewable energy is more sustainable and does not deplete. Hence, fewer carbon emissions do not emit greenhouse gases or create pollution using renewable energies. Compared to renewable energies, coal power plants produce around 2.2 pounds of CO_2 per kilowatt-hour of electricity.

In contrast, renewable energies such as solar panels and windmills create none. Our world is fast decarbonizing the environment and moving to energy sources that do not cause pollution. Burning fossil fuels to produce electricity is more harmful than warming the climate. It contaminates the air and water, and it affects humans. For example, coal power stations emit high CO_2 and N_2O (two leading greenhouse gases) directly entering the atmosphere. In addition, it also emits lead, sulfur dioxide, mercury, and dangerous metals that cause health problems such as breathing difficulties to death. However, renewable sources do not create pollution or contamination risks in the air and water. Renewable energy helps maintain clean air free from pollution experienced by the world in recent times. Over the last decade, the growth of renewable sources such as wind and solar power was the cheapest natural source in various parts of the world. The United Arab Emirates (UAE)—known for its arid and plain land and sunny climate—could produce electricity at a meager cost.

Once used, nonrenewable energy cannot be regenerated. Nonrenewable energies are majorly fossil fuels produced globally, as they are affordable. Generally, nonrenewable energies are created by organic materials compressed and heated for a long time and turned into natural gas and crude oil. One of the natural gases used in real-time heating and lighting applications is methane gas. It needs some organic chemical products to generate, but it's costly. Hydrocarbons are colorless and highly flammable, primarily consisting of ethane and methane. Fossil fuel is used for generating electricity, cooking, and heating and as a fuel for vehicles. Coal is a deposit of sediment that consists predominantly of readily combustible carbon. It is heated to make steam, piped at a higher pressure in the turbine, rotating and producing electricity. It is burned to heat the water to generate electricity by oil to convert into steam. Oil is burned under high pressure to create hot exhaust gases that rotate a turbine to produce electricity.

10.1.1 REASON FOR RENEWABLE ENERGY SYSTEM

Time consumption is a primary drawback of nonrenewable energy. Searching for oil, mining coal, installing and building oil rigs and drills, and transporting natural gases through pipes take a lot of time and physical work. As nonrenewable sources take billions of years to form, they gradually deplete from the face of the Earth. Fossil fuels are hazardous and emit gases such as carbon monoxide that cause respiratory problems in humans. Resources such as coal, natural gas, and oil emit heavy amounts of carbon dioxide while burning. It is costly and consumes a lot of space. A considerable fuel reserve must be maintained to run power stations without interruption.

An IoT network can be described as a group of interconnected devices that communicate with other devices without human intervention, such as cars, smart devices, and wearables. The source of renewable energies with advanced sensing and communication technologies resulted in sustainable processes in modern energy systems. Integrating actuators and sensors enables the system's connection and energy-related data exchange between resources, computer systems, and end-users in a network.[1] IoT devices' main aim is to simplify the process in various fields and improve efficiency. Sustainability is a major issue due to overpopulation, but the development of IoT technologies gradually increases sustainability.[2] Efficiency is a major issue in utilizing renewable energy. This is true for methods that rely on solar and wind energy resources.

Implementation of IoT devices helps to improve efficiency in electricity production. With the help of automated controls, the energy obtained from renewable sources can be utilized efficiently to the optimum. An IoT device helps identify where power can be optimally generated. The equipment can be adjusted according to the weather conditions to obtain the maximum output. One of the main benefits of IoT devices is they generate data in real time, which helps to reduce wastage, if any. It is used to monitor wind turbines' effect that results in turbines' inefficient working. With the help of IoT sensors, we can change the direction of wind turbines. Likewise, solar panels can also be altered to get maximum efficiency. These changes are done either manually or automatically to enhance electricity production. The power plants can also use this IoT device and produce maximum output. Some benefits of IoT devices in renewable energies are automation, cost-efficiency, grid management, residential systems, and distributed systems. To get the maximum benefits of renewable energy, governments, and other organizations have now shifted to using IoT devices and sensor systems from the conventional methods of power generation.

10.2 LITERATURE REVIEW

In this paper, the author has aimed to understand the various issues and challenges in obtaining energy efficiency in IoT-based sustainable environment. For that, a detailed literature survey is carried out on earlier research methodologies focused on enhancing energy efficiency using RESs. For example, E.Oró et al. have reviewed renewable energy systems used for energy efficiency.[3] The renewable energy system avoids carbon dioxide emissions and increases energy efficiency in the data center. The author's main aim is to find suitable renewable energy resources and provide a better design than the existing ones.

Tarhan[4] presented a detailed review of renewable energy systems and their impact on the economy, society, and society. The results show that the existing implementation provided a successful output using renewable energy, which eliminates the problems in energy production and related financial struggles. Abdmouleh et al.[5] have studied various methodologies in implementing renewable energy systems and various sectors and technologies. The summary shows that optimization methodologies based on renewable energy systems are more efficient than conventional methods. Kannan and Vakeesan[6] stated that the energy demand is increasing in proportion to the rise in the population. So, there was an urgency for a reliable and lasting RES to meet population demands. As other energy sources such as fossil fuel, thermal, and hydro are limited, the author has recommended solar energy as an alternative. Compared with other sources of renewable energies, solar energy is abundant in its availability, cost effective, and easy to access.

Herez et al.[7] reviewed that the most popular RES is solar energy. Moreover, it is mainly used for heat generation and electricity production. The solar energy production method is increased and used to reduce fossil fuel usage. Because during the production of fossil fuel, it emits chlorofluorocarbon (CFC). It results in global warming and ozone depletion. To prevent these issues, various solar energy systems and collectors are used. Kar et al.[8] reviewed the potential of solar energy in India in terms of its development, capacity, production, and trade status. However, solar energy has many problems in grid connectivity, such as evacuation infrastructure, capacity utilization, initial investment, consumer awareness, and acceptance. Therefore, some methods are suggested to generate pure energy to tackle these issues, such as attracting investments, easy access to financing, developing transmission and distribution infrastructure, etc. In the research paper, Langer et al.[9] have selected the radial-flux permanent-magnet synchronous motor (PMSM) generator for study. This generator is mainly used to produce low-power wind energy. Compared to other generators, PMSM has many advantages such as efficiency, reliability, energy density, small size, and lightweight. Along with PMSM, the author used four MTTP systems to produce wind energy.

Sitharthan et al.[10] stated that every country could implement any renewable resource for electricity production from nature due to an increase in energy production, which increases the energy supply. Wind energy plays a significant role among various renewable energy resources, fulfilling 28% of energy requirements. However, although wind energy is a freely available source, only a few countries show interest in implementing it. Of these, India is one country that implements wind power generation. Also, the author has surveyed aspects such as advanced technology, future opportunities, and government initiatives for power consumption. Kulkarni and Anil[11] stated that carbon dioxide emission causes many adverse effects on the environment and industries. To limit the emission of CO_2, most countries prefer RESs for power production. Examples are wind energy, biomass energy, Solar energy, etc. Of these, wind energy as a source of renewable one is growing fast. It converts kinetic energy into electricity. Choi et al.[12] discussed that different renewable energy resources are adopted nowadays. For example, wind power and photovoltaic (PV) power generation are maintained using IoT devices. It performs various tasks such as monitoring, diagnosing, and predicting using IoT devices such as Arduino, Raspberry Pi, and the LoRa network.

Kumar et al.[13] reviewed that solar energy has good potential in various kinds of renewable energy. To convert the sunlight into a PV electricity system is used. It can be installed in any place. First, however, to access the real-time performance and control the PV system, the IoT device is installed. It helps to monitor, diagnose faults, and generate and maintain records. Santos[14] states that IoT devices are used for various grid applications because of their acquisition, processing, transmission, and storage features. It is used in multiple places and fields such as homes/buildings, healthcare, agriculture, and cities to perform as a sensor, receiver, and sender. Khyam et al.[15] reviewed that different types of renewable energy are becoming more popular in the last few decades. Wind energy is one of the most popularly used types of renewable energy. As wind turbines and control centers are located in different places, the IoT System is used to monitor and analyze the performance of controllers and turbines. Alhmoud and AI-Zoubi[16] stated that RESs are essential in power production. IoT devices play a significant role. IoT, especially in wind energy, transfers the information between the turbine and controller. It reduces wind energy cost and risk in real-time with the help of the permanent magnet synchronous generator (PMSG). Hossein Motlagh et al.[18] reviewed the implementation of IoT devices in the energy sector to increase energy efficiency and its problems on privacy and security in the energy sector. To avoid this problem and to get the output, blockchain is used.

Himeur et al.[19] stated that energy needs have increased in the last decade. Buildings consume around 40% of the total energy produced worldwide. This consumption rate will increase by 1.3% annually from 2018 to 2050 in Organization for Economic Cooperation and Development (OECD) countries such as the USA, Australia, etc. In non-OCED countries such as China, Russia, etc., the increase will be around 2%. Pylsy et al.[20] reviewed that the rise in population will increase the usage of electricity and the rate of energy consumption. Zia et al.[17] state that PV and wind energy are highly used RESs. It increases the need for electricity and reduces greenhouse gas emissions. Eltamaly et al.[21] stated that renewable energy production has increased due to the disadvantage of fossil fuel production. A small hybrid renewable energy system (SHRES) is used to maintain energy production. This SHRES was performed with the help of an IoT device. Syed et al.[22] reviewed the fundamental properties of the IoT device used in smart cities for energy efficiency. The problems and risks are also discussed. Privacy and security, IoT, and AI devices are used to manage the smart city challenges.

10.2.1 LIMITATIONS AND MOTIVATION

The above study shows that IoT devices require a continuous and sustained energy supply for effective functioning. Several research works are discussed, showing the optimized usage of energy available to power IoT devices. These IoT devices connected to a network are optimized through machine learning algorithms and frameworks. The developments in IoT technologies have enabled IoT devices' usage in smart buildings, smart cities, tech-hubs, etc. The works discussed above show us the possible requirement of energy resources and their impact on the environment. So, there is a need to adopt methods that generate energy without affecting the environment of renewable energy resources. However, renewable energy resources have some limitations in terms of energy production. The energy requirements of IoT-based environments cannot be fulfilled through renewable energy resources alone. Several researchers have employed renewable energy systems to power IoT systems that lack efficiency and functioning. There is a need to implement renewable energy as much as possible to reduce the energy consumed by nonrenewable resources.

Energy efficiency is one of the significant challenges and is highly required to run everyday life. Thus, every country has aimed to increase energy production under various strategies. One of the strategies is energy generated from natural resources. One of the popular RESs is solar PV systems. It requires one-time installation costs and does not require more maintenance and production costs. Thus, the solar PV system is considered by everyone in the energy industry.

10.3 PROBLEM FORMULATION

The energy management strategies are more critical in the IoT environment as the available energy is less and the time required for the devices to function is infinite. However, specific optimization methods can be followed to improve efficiency and reduce the energy requirement for the devices. The energy requirement of the system needs to be calculated to optimize the whole system. The optimization algorithms use parameters such as time duration, thermal throttle, energy consumption, and other electrical parameters for better optimization. A mathematical model is proposed that calculates the following parameters needed for efficient optimization. Before implementing the optimal system, the entire functionality and its logical structure must be understood. Figure 10.2 shows the electricity flow architecture in home and other office environments. The components in Figure 10.2 are individual nodes that either receive or transfer energy. The energy is first generated in the RES, connected to the home appliances and the energy storage system (ESS). It does not take any input but gives output to the connected nodes. The primary grid is also connected to both nodes. The main grid gets the supply from the central lane, but the RES gets it through renewable resources. The power obtained through RES may not be consistent. Thus, an uninterrupted energy source such as the main grid is required for ongoing energy input. The home appliances are connected to three nodes from which it gets input. From the connected nodes, the home appliances receive an uninterrupted supply. If any node fails to supply, the other nodes fulfill the energy requirement. RES and ESS are just alternatives to the primary grid, as they cannot power them over a long period.

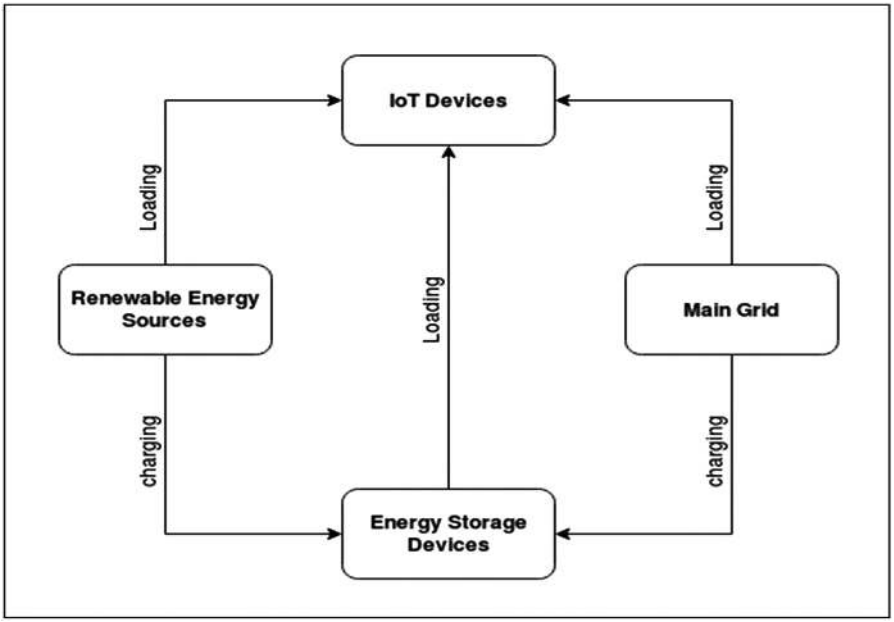


FIGURE 10.2 Electricity flow architecture.

10.3.1 RENEWABLE ENERGY SOURCE (RES)

In this model, a PV system is considered the RES. The power output of the RES be P_{RES} That produces power (in kW) at time τ as in eq 10.1.

$$P_{\text{RES}}(\tau) = \text{GHI}(\tau).S.\eta^{\text{RES}} \forall \tau 0 \le \tau \le 24 \tag{10.1}$$

The GHI represents the global horizontal irradiation in the place of solar panels. The energy produced per meter is measured in $\left(\frac{kW}{m^2}\right)$. The area of the solar PV is measured in m^2. The efficiency in energy conversion in a solar PV system is η^{RES}. Let the time taken be t and the duration of time taken be Δt. The energy generated by the PV system is $E_{\text{RES}}(t)$ can be shown at eq 10.2.

$$E_{\text{RES}}(t) = P_{\text{RES}}(\tau) \cdot \Delta t \forall t 1 \le t \le T \tag{10.2}$$

The real-time here is τ in the time t. Figure 10.2 shows that the energy generated from renewable sources is used by home appliances and to charge the ESS. Thus, it can also be written as

$$E_{\text{RES}}(t) = E_{\text{RES}}^{\text{load}}(t) + E_{\text{RES}}^{\text{charge}}(t) \forall t 1 \le t \le T \tag{10.3}$$

where the $E_{\text{RES}}^{\text{load}}(t)$ is the amount of energy used in the home in time t; $E_{\text{RES}}^{\text{charge}}(t)$ is the amount of energy used for charging the ESS in time t. From the equations, the $E_{\text{RES}}^{\text{load}}(t)$ and $E_{\text{RES}}^{\text{charge}}(t)$ should satisfy the constraints are

$$0 \le E_{\text{RES}}^{\text{load}}(t) \le \text{GHI}(\tau) \cdot S \cdot \eta^{\text{RES}} \cdot \Delta t \forall t\, 1 \le t \le T \tag{10.4}$$

$$0 \le E_{\text{RES}}^{\text{charge}}(t) \le \text{GHI}(\tau) \cdot S \cdot \eta^{\text{RES}} \cdot \le \Delta t \forall t 1 \le t \le T \tag{10.5}$$

10.3.2 ENERGY STORAGE SYSTEM (ESS)

The ESS can act as a substitute for both the renewable source and primary grid at the time of any other issue. It can conserve energy from both a renewable resource and the main grid. This energy can also be provided at a higher price. The parameters dealing with ESS are shown below. In general, ESS has two major roles to play, (1) Supplying energy to the home and (2) Selling the surplus energy outside. It also stores the surplus energy of both the RES and the main grid. The following formulas are used for estimating both the measures:

$$E_{\text{ESS}}^{\text{discharge}}(t) = \left(E_{\text{ESS}}^{\text{load}}(t) + E_{\text{ESS}}^{\text{selling}}(t)\right) \cdot \left(1 - \text{mode}^{\text{ESS}}(t)\right) \tag{10.6}$$

$$E_{\text{ESS}}^{\text{charge}}(t) = \left(E_{\text{RES}}^{\text{charge}}(t) + E_{\text{MG}}^{\text{charge}}(t)\right) \cdot \text{mode}^{\text{ESS}}(t) \tag{10.7}$$

Here, the energy is extracted from ESS within the time t is $E_{\text{ESS}}^{\text{discharge}}(t)$. In eq 10.7, $E_{\text{ESS}}^{\text{charge}}(t)$ represents the stored energy in ESS at time $t \cdot E_{\text{ESS}}^{\text{load}}(t)$ is the energy used in the home and $E_{\text{ESS}}^{\text{selling}}(t)$ denotes the energy sold outside at time t; $E_{\text{RES}}^{\text{charge}}(t)$ is the amount of energy given by RES to ESS and $E_{\text{MG}}^{\text{charge}}(t)$ is the amount of energy given by the primary grid to ESS. ESS stores and supplies power but cannot consume it. However, while storing the energy, there can be energy loss. It has only two modes that can be considered as a binary value. It is expressed in the following form as in eq 10.8.

$$\text{mode}^{\text{ESS}}(t) = \begin{cases} 1 \text{ if ESS is charged} \\ 0 \text{ if ESS is discharged} \end{cases} \tag{10.8}$$

The ESS energy level is assumed as $E_{\text{ESS}}^{\text{level}}(t)$ for time t. Here, $\forall t\, 1 \le t \le T$, the formula can be written as eq 10.9.

$$E_{\text{ESS}}^{\text{level}}(t) = E_{\text{ESS}}^{\text{level}}(t)(t-1) + E_{\text{ESS}}^{\text{charge}}(t) \cdot \eta^{\text{ESS}} - E_{\text{ESS}}^{\text{Discharge}} \frac{(t)}{\eta^{\text{ESS}}} \tag{10.9}$$

Here, the energy efficiency of ESS is η^{ESS}, included in the formula because some energy loss occurs during the charging and discharging of energy, and it is the round trip efficiency of the system. The following constraints should be satisfied by the ESS:

a) The rate of charging and discharging in ESS should not exceed $\frac{Ch_{\text{rate}}}{Dh_{\text{rate}}}$. Only a certain amount of energy could be extracted or given to the ESS with the time t within the duration Δt.
b) The ESS energy level should lie between the EL_{min} and EL_{max}.

From the above-said constraints, along with $\forall t\, 1 \le t \le T$, the following conditions can be added:

$$0 \le E_{\text{ESS}}^{\text{discharge}}(t) = E_{\text{ESS}}^{\text{load}}(t) + E_{\text{ESS}}^{\text{selling}}(t) \le Dh_{\text{rate}} \cdot \Delta t \tag{10.10}$$

$$0 \le E_{\text{ESS}}^{\text{charge}}(t) = E_{\text{RES}}^{\text{charge}}(t) + E_{\text{MG}}^{\text{charge}}(t) \le Ch_{\text{rate}} \cdot \Delta t \tag{10.11}$$

$$EL_{\text{min}} \le E_{\text{ESS}}^{\text{level}}(t) \le EL_{\text{max}} \tag{10.12}$$

$$0 \le E_{\text{ESS}}^{\text{load}}(t) \le Dh_{\text{rate}} \cdot \Delta t \tag{10.13}$$

$$0 \le E_{\text{ESS}}^{\text{selling}}(t) \le Dh_{\text{rate}} \cdot \Delta t \tag{10.14}$$

$$0 \le E_{\text{RES}}^{\text{charge}}(t) \le Ch_{\text{rate}} \cdot \Delta t \tag{10.15}$$

$$0 \le E_{\text{MG}}^{\text{charge}}(t) \le Dh_{\text{rate}} \cdot \Delta t \tag{10.16}$$

The calculation is only done for a day and is not for all the following days. So, the energy level should return to the initial state when the day ends. This constraint can be shown in eq 10.17.

$$E_{\text{ESS}}^{\text{level}}(T) = EL_0 \tag{10.17}$$

The energy to be sold can only be obtained from ESS. The energy generated at RES needs to be stored before selling. The constraints for the variable $E_{\text{RES}}^{\text{charge}}(t)$ is in eqs 10.5 and 10.15. If more energy is generated in RES, the excess energy is stored in ESS and sold.

10.3.3 HOME DEVICES

The home devices are of two types, namely, (1) movable M, (2) immovable N. The movable devices are $M = \{a_1, a_2, a_3, \ldots, a_m\}$. These devices can be operated at any time, and the operation time for the devices can be changed to lower price slots to save costs. The immovable devices $N = \{a_1, a_2, a_3, \ldots, a_n\}$ can be operated only within a certain time frame and is predefined. The energy consumption of the devices is $E_{\text{total}}^{\text{appliances}}(t)$, is the energy consumption of the *M* movable set $E_N(t)$, shown in the proceeding equations along with

$$\forall B_i \in N, \forall a_i \in M, \forall t 1 \leq t \leq T.$$

$$E_{\text{total}}^{\text{appliances}}(t) = E_N(t) + E_M(t) \tag{10.18}$$

$$E_N(t) = \sum_{i=1}^{n} \text{Power}_{\text{rate}}(b_i) \times O(b_i, t) \times \Delta t \tag{10.19}$$

$$E_M(t) = \sum_{i=1}^{m} \text{Power}_{\text{rate}}(a_i) \times O(a_i, t) \times \Delta t \tag{10.20}$$

Here, the power ratings of a_i, b_i devices are $\text{Power}_{\text{rate}}(a_i)$ and $\text{Power}_{\text{rate}}(b_i)$ respectively. The producers fix that. The status of the devices a_i, b_i can be seen through the binary variables $O(a_i, t)$ and $O(b_i, t)$.

$$O(a_i, t) = \begin{cases} 1 \text{ if shiftable device } a_i \text{ is ON} \\ 0 \text{ if shiftable device } a_i \text{ is OFF} \end{cases} \tag{10.21}$$

$$O(b_i, t) = \begin{cases} 1 \text{ if shiftable device } b_i \text{ is ON} \\ 0 \text{ if shiftable device } b_i \text{ is OFF} \end{cases} \tag{10.22}$$

The $E_N(t)$ has a fixed value as the $O(b_i, t)$ is fixed. In a whole day of $T = 24$ hours, the energy consumption of all the devices in the system can be calculated by eq 10.23.

$$\sum_{t=1}^{T} E_{\text{total}}^{\text{appliances}}(t) = \sum_{t=1}^{T} E_N(t) + \sum_{t=1}^{T} E_{(t)}. \tag{10.23}$$

If the movable devices cost much less, the demand for energy remains unchanged as its operational time does not change. The home devices are provided with three different sources of energy which are $E_{\text{RES}}^{\text{load}}$, $E_{\text{ESS}}^{\text{load}}$, and $E_{\text{MG}}^{\text{load}}$ in time *t*. The following formula can be written as $\forall t 1 \leq t \leq T$.

$$E_{\text{total}}^{\text{appliances}}(t) = E_{\text{RES}}^{\text{load}}(t) + E_{\text{ESS}}^{\text{load}}(t) + E_{\text{MG}}^{\text{load}}(t) \tag{10.24}$$

$$=> E_{\text{MG}}^{\text{load}}(t) = E_{\text{total}}^{\text{appliances}}(t) - E_{\text{RES}}^{\text{load}}(t) - E_{\text{ESS}}^{\text{load}}(t) \tag{10.25}$$

From 10.18, we have

$$=> E_{\text{MG}}^{\text{load}}(t) = E_N(t) + E_M(t) - E_{\text{RES}}^{\text{load}}(t) - E_{\text{ESS}}^{\text{load}}(t) \tag{10.26}$$

The main grid $E_{\text{MG}}^{\text{load}}(t) \geq 0$ always provides energy. If the home demand is "0," it represents that it does not consume any energy. Thus, the following constraint:

$$0 \leq E_{\text{RES}}^{\text{load}}(t) + E_{\text{ESS}}^{\text{load}}(t) \leq E_N(t) + E_M(t) = E_{\text{total}}^{\text{appliances}}(t) \tag{10.27}$$

10.3.4 LOAD DEMAND AND COST FUNCTION

The energy from the main grid is the load demand over a certain period. The cost of energy used for a day can be calculated from the load demand or the price of the energy used from the main grid. From eq 10.23, it can be concluded that there are various kinds of tariffs on electricity for the home. Day-ahead pricing (DAP) is used in this paper to collect charges on an hourly basis and remind every hour. The prices are forecast a day. It is also assumed that the energy produced by both the RES and ESS is more than the energy given by the primary grid, $E_{\mathrm{LD}}(t)$ includes $E_{\mathrm{MG}}^{\mathrm{load}}(t)$ and $E_{\mathrm{MG}}^{\mathrm{charge}}(t)$ are shown the Figure 10.2. In time, $\forall t 1 \le t \le T$, the Equation can be written as eq 10.28.

$$E_{\mathrm{LD}}(t) = E_{\mathrm{MG}}^{\mathrm{load}}(t) + E_{\mathrm{MG}}^{\mathrm{charge}}(t) \tag{10.28}$$

From eq 10.26, it is seen that

$$E_{\mathrm{LD}}(t) = E_N(t) + E_M(t) + E_{\mathrm{MG}}^{\mathrm{charge}}(t) - E_{\mathrm{RES}}^{\mathrm{load}}(t) - E_{\mathrm{ESS}}^{\mathrm{load}}(t) \tag{10.29}$$

The price that is got by selling energy outside is $E_{\mathrm{ESS}}^{\mathrm{selling}}(t)$ in time t. The cost of the energy to be paid in time t is $EC(t)$, which is shown in eq 10.30.

$$EC(t) = E_{\mathrm{LD}}(t) \times P_{\mathrm{MG}}(t) - E_{\mathrm{ESS}}^{\mathrm{selling}}(t) \times P_{\mathrm{sell}}(t) \tag{10.30}$$

Here, $P_{\mathrm{MG}}(t)$ is the price paid for the energy given by the main grid in time t. The provider fixes the price $P_{\mathrm{sell}}(t)$ in time t. The users decide this. The total cost to be paid for a day is C_{day}, then

$$\begin{aligned} C_{\mathrm{day}} &= \sum\nolimits_{t=1}^{T} EC(t) \\ &= \sum\nolimits_{t=1}^{T} \left(E_{\mathrm{LD}}(t) \times P_{\mathrm{MG}}(t) - E_{\mathrm{ESS}}^{\mathrm{selling}}(t) \times P_{\mathrm{sell}}(t) \right) \end{aligned} \tag{10.31}$$

From eq 10.29, the formula can be written as

$$C_{\mathrm{day}} = \sum\nolimits_{t=1}^{T} \left(\left(E_N(t) + E_M(t) + E_{\mathrm{MG}}^{\mathrm{charge}}(t) - E_{\mathrm{RES}}^{\mathrm{load}}(t) - E_{\mathrm{ESS}}^{\mathrm{load}}(t) \right) - E_{\mathrm{ESS}}^{\mathrm{selling}}(t) \times P_{\mathrm{sell}}(t) \right) \tag{10.32}$$

The following equation shows the objective function value (OFV), which aims to reduce the total energy cost (TEC) for a day is

$$\begin{aligned} &\min\left(C_{\mathrm{day}}\right) = \min \\ &\left(\sum\nolimits_{t=1}^{T} \left(\left(E_N(t) + E_M(t) + E_{\mathrm{MG}}^{\mathrm{charge}}(t) - E_{\mathrm{RES}}^{\mathrm{load}}(t) - E_{\mathrm{ESS}}^{\mathrm{load}}(t) \right) - E_{\mathrm{ESS}}^{\mathrm{selling}}(t) \times P_{\mathrm{sell}}(t) \right) \right) \end{aligned} \tag{10.33}$$

$$\min\left(\sum\nolimits_{t=1}^{T} (E_N(t) + \sum\nolimits_{i=1}^{m} \mathrm{Power}_{\mathrm{rate}}(a_i) \times O(a_i, t) \times \Delta t + E_{\mathrm{MG}}^{\mathrm{charge}}(t) - E_{\mathrm{RES}}^{\mathrm{load}}(t) - E_{\mathrm{ESS}}^{\mathrm{load}}(t) - \alpha \times E_{\mathrm{ESS}}^{\mathrm{selling}}(t)) \times P_{\mathrm{MG}}(t) \right) \tag{10.34}$$

Pea-to-average Ratio (PAR)

The ratio between peak load demands and the average load demand is called the PAR.

$$\mathrm{PAR} = \frac{\max\left(E_{\mathrm{LD}}(t)\right)}{\frac{1}{T}\sum_{t=1}^{T} E_{\mathrm{LD}}(t)} \tag{10.35}$$

10.4 PROPOSED PSO ALGORITHM

One popular optimization algorithm is the particle swarm optimization (PSO) algorithm. It is one of the computational algorithms used to imitate organisms' functionalities and solve persistent and complex optimization problems, as illustrated in Figure 10.3. Initially, the mass of particles produced is randomly distributed. These particles are present in the problem's search space. The issue needs to be optimized through the algorithm. After each iteration, all the particles move to different positions within the search space to find a better solution for the problem. A formula is derived from calculating the movements of the particles. It considers its position, where it currently relies, and its velocity to determine the next position. Usually, the new position and velocity are taken as $x_i(t + 1)$ and $v_i(t + 1)$ for each particle i over $t + 1$ times iteration, which can be written as in eqs 10.36 and 10.37.

$$v_i(t+1) = \omega \cdot v_i(t) + C_l \cdot r_l \cdot \left(lb_i(t) - x_i(t)\right) + C_g \cdot r_g \cdot \left(gb(t) - x_i(t)\right) \tag{10.36}$$

$$x_i(t+1) = x_i(t) + v_i(t+1) \tag{10.37}$$

where	ω	-	Inertia weight (constant)
	$v_i(t)$	-	velocity of a particle at iteration t
	C_i	-	acceleration coefficient
	r_l	-	random number between 0 and 1
	$lb_i(t)$	-	best position of the particle over iteration t.
	$x_i(t)$	-	position of the particle at iteration t
	C_g	-	acceleration coefficients (constant)
	r_g	-	random number distributed between 0 and 1.
	$gb(t)$	-	best position for iteration t.

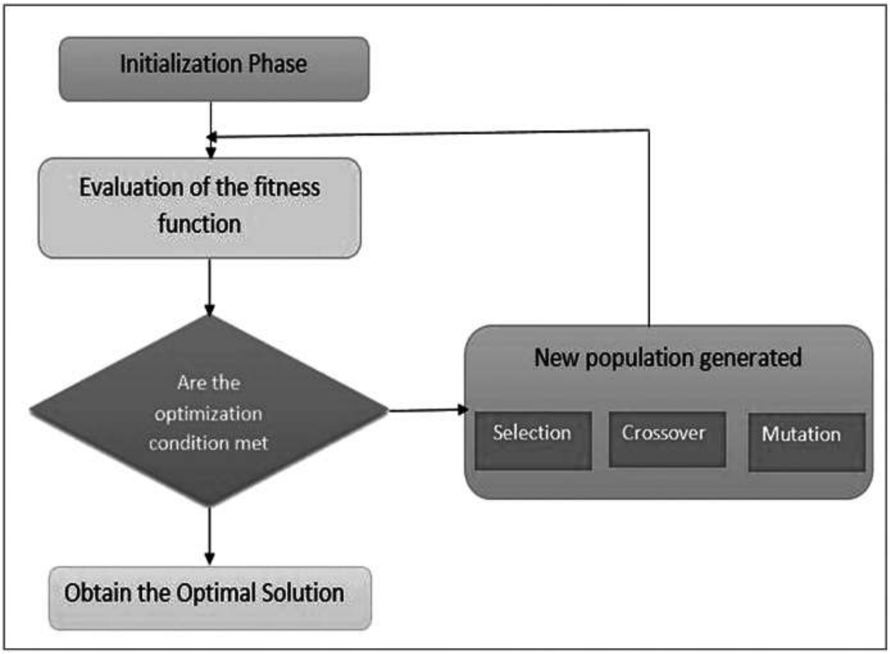


FIGURE 10.3 Architecture of PSO.

There are two kinds of variables in the proposed model. They are $E_{\mathrm{MG}}^{\mathrm{charge}}$ and $E_{\mathrm{ESS}}^{\mathrm{load}}$. The binary variables are taken as 0 and 1 with the $O(a_i,t)$. The proposed algorithm performs well in the continuous problems state, and in terms of binary variables, the binary PSO is used for better results. The same is applied to the proposed model, and the results are tabulated in the following sections.

10.5 SIMULATIONS AND DISCUSSIONS

The energy used for several household appliances is simulated for one day hourly. Here, a day consists of 24 h and is considered the time slots. Timeslot 1 indicates the time slot from 0 a.m. to 1 a.m., and time slot 2 denotes from 1 a.m. to 2 a.m., time slot 3 is from 2 a.m. to a.m., and so on. The smart home energy management system (SHEMS) is evaluated by single-objective optimization. It is used for controlling and scheduling every appliance. RESs and energy storage systems (ESSs) help to optimize the energy cost. The

TEC is evaluated based on the daily energy utilized. This program can run on Intel® Core™ i7-8700 CPU@ 320GHz and RAM size of 16GB with Windows 10pro.

The proposed simulation includes the input parameters for the home appliances such as washing machines, air conditioners, personal computers, television, lights, etc.

Table 10.1 lists all the appliances that are used in SHEMS. Here, the appliances are categorized into shiftable and nonshiftable. The operating time can be changed to time slots of low price in the shiftable appliances. On the contrary, it is impossible to change the time slots in nonshiftable appliances. Both types of appliances could not be interrupted at the time of operation. The power rating, daily usage, and the operation's time length for all the appliances are shown in Table 10.4. The power rating is expressed in kW, whereas the daily usage is expressed in hours.

TABLE 10.1 Home Appliances.

Load Type	Appliances	Power rating (kW)	Daily usage (Hours)	Start time
	Washing machine	0.8	5	-
	Air conditioner	1.3	10	-
Shiftable	Clothes dryer	0.7	4	-
	Dry washer	0.2	3	-
	Personal computers	0.2	18	7 A.M.
Non_shiftable	Security cameras	0.1	24	0 A.M.
Non_shiftable	Security cameras	0.1	24	0 A.M.
	Microwave oven	0.5	7	3 P.M.
	Refrigerator	0.9	20	2 A.M.
	Television	0.2	8	4 P.M.
	Lights	0.1	6	6 P.M.

Figure 10.4 shows the cost variance from one time slot to another based on the DAP signal. To compute $P_{RES}(\tau)$, solar irradiance is analyzed in 24 time slots, as shown in Figures 10.5 and 10.6 displays the RES energy that is generated by the PV system.

The PV system is used to generate electricity for the RES. The RES depends upon the solar generator, cells, and irradiation for energy conversion efficiency. The energy with the same amount is generated by configuring RES. Like RES, ESS also uses the same configuration.

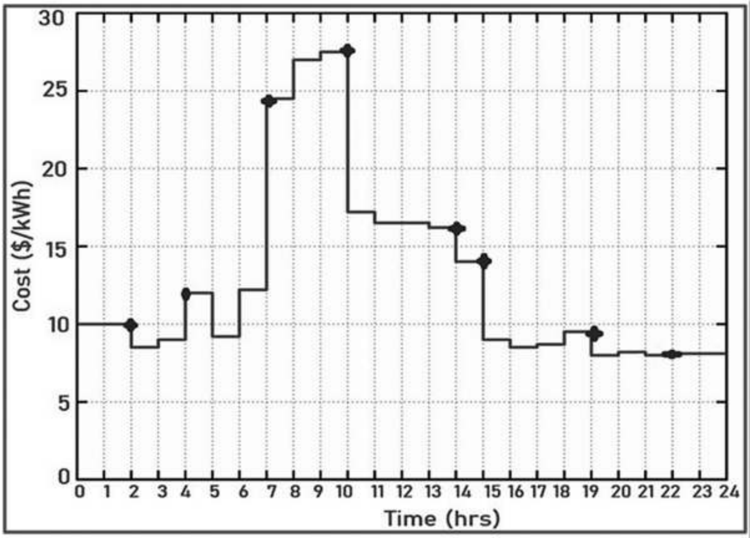


FIGURE 10.4 Cost variance with DAP signal.

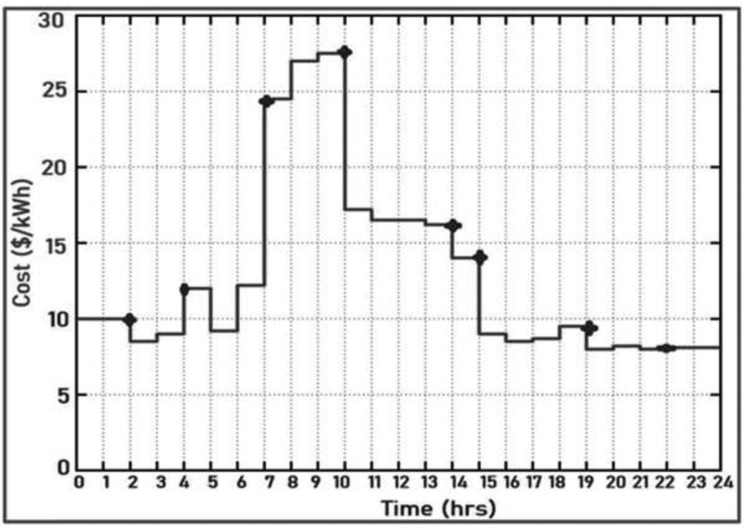


FIGURE 10.5 P_{RES} (τ) by solar irradiance.

Table 10.2 lists the ESS parameters. Ahmad et al.[23] proposed a new plan for both RES and ESS. The ESS is charged with the energy of 30%, which is taken from RES for every time slot. The rest of the energy is used for the appliances available in the home. The PV system is the only possible way to charge ESS in the daytime. The home load has a high price from t_{20} to t_{24} time slots while considering the energy in ESS. It is important to have a flexible plan for both RES and ESS.

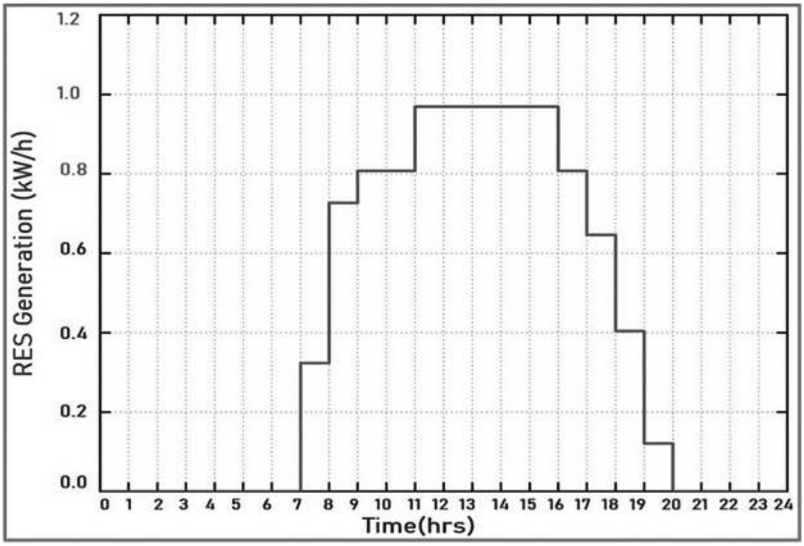


FIGURE 10.6 RES energy through solar PV.

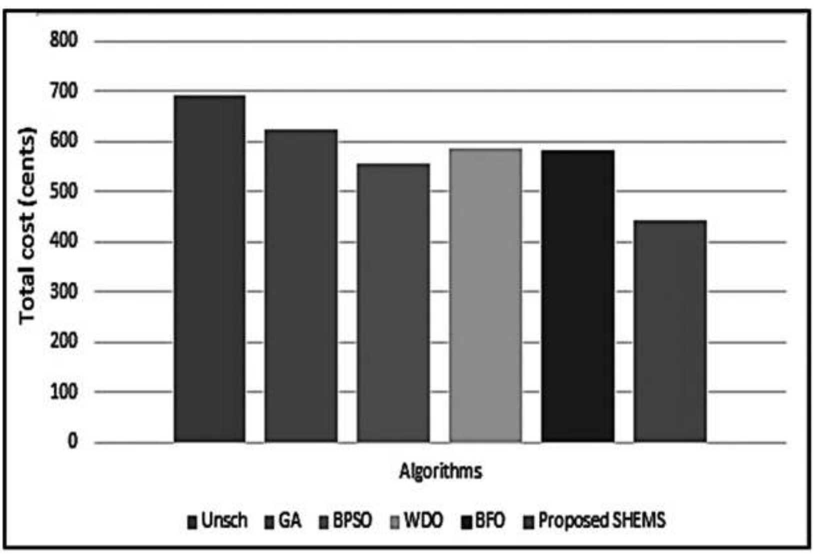


FIGURE 10.7 TEC of several algorithms.

TABLE 10.2 The ESS Input Parameters for Simulation.

η^{ESS} (%)	Ch_{rate}(kW)	Dh_{rate}(kWh)	EL_{min}(kWh)	EL_{max}(kWh)
95%	0.3	0.5	0.5	3

This paper considers only a single-objective optimization model for reducing the TEC for a day.

$$P_{\text{sell}(t)} = P_{\text{MG}}(t)\forall t \tag{10.38}$$

Figure 10.7 shows the comparison of the TEC of Ahmad et al.[23] and the proposed system. It shows that our proposed SHEMS utilizes the energy in both the primary grid and ESS.

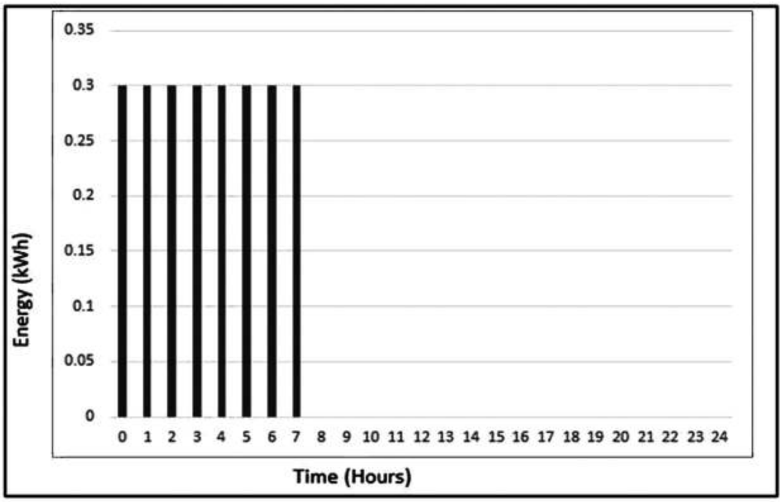


FIGURE 10.8 Main grid storage in ESS.

Figure 10.8 shows that the main grid's energy is stored in ESS hourly. It utilizes the energy at a low price from 0 a.m. to 2 a.m. This low-price energy will be used for home appliances for the time slots from 8 a.m. to 10 a.m. as shown in Figure 10.9. Or else, this cheap energy can be sold outside at the timeslots of from 7 a.m. to 8 a.m. and from 10 a.m. to 11 a.m., as shown in Figure 10.10. Figure 10.11 shows that the energy from RES is distributed to the home appliances. RES is used to generate energy at a high price. It is better to utilize the home energy rather than store it in ESS. If the RES energy is stored in ESS, discharge happens. Then the energy will be lost because of its round trip. The energy from RES should be more than the energy required for the home appliances. Hence, the proposed SHEMS focused on storing ESS from the energy of RES.

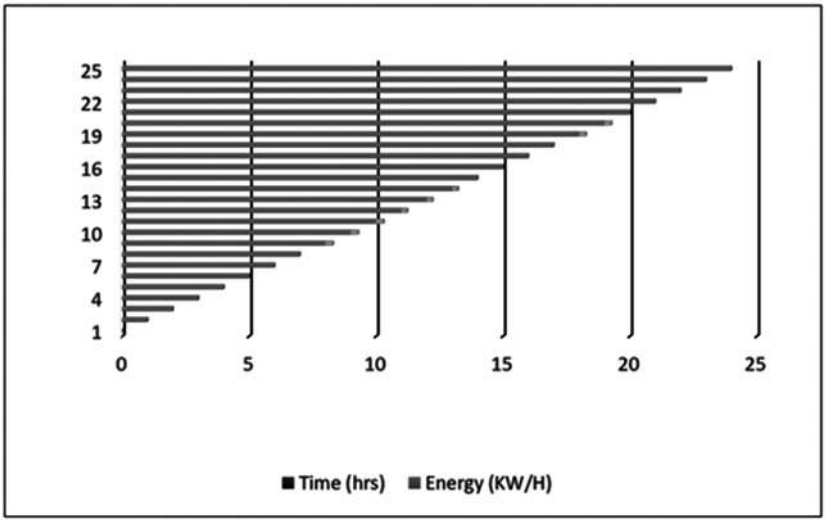


FIGURE 10.9 ESS energy use for home.

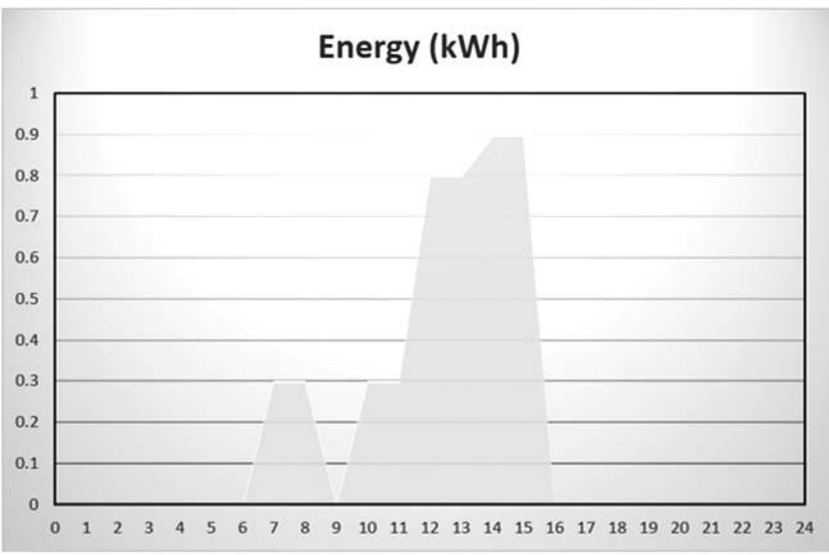


FIGURE 10.10 Selling of ESS energy.

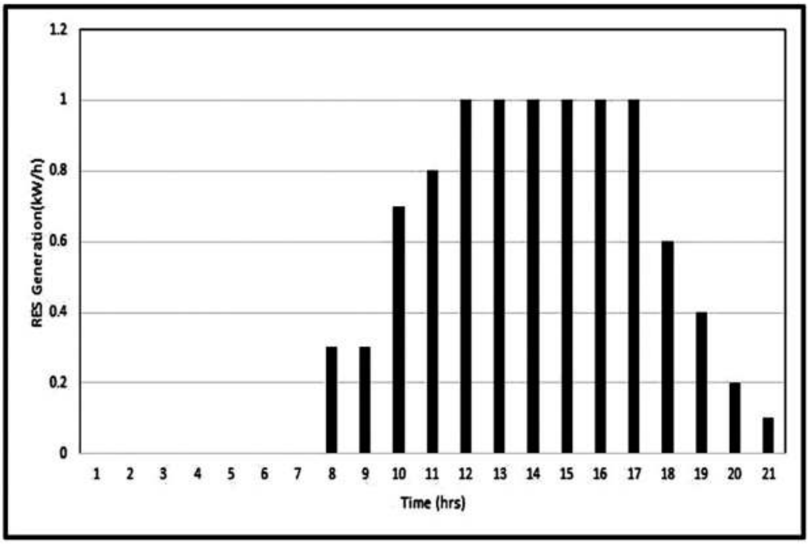


FIGURE 10.11 Energy storage—ESS from RES For home load.

Table 10.3 shows the scheduling time of each shiftable appliance. The proposed SHEMS is used to schedule every appliance to utilize the energy at a low price.

TABLE 10.3 Scheduling the Shiftable Appliances.

Load Type	Appliances	Daily usage (hours)	Start time
Shiftable	Washing machine	5	7 P.M.
	Air conditioner	10	2 P.M.
	Clothes dryer	4	7 P.M.
	Water heater	8	4 P.M.
	Dry washer	3	7 P.M.

Figure 10.12 shows the load demand from the main grid hourly. The red line displays the energy requirement needed for home appliances, and the blue line indicates the system's load demand from the grid. The time slots from 7 p.m. to 9 p.m. need more energy for using the home appliances. It needs less energy from 0 a.m. to 7 a.m. and has a lot of energy from 7 a.m. to 8 a.m. The proposed HEMS does not require any energy from the time slots 11 a.m. to 1 a.m. Both RES and ESS have

enough energy for using the home appliances. Figure 10.13 compares the PAR (PAR) for the different algorithms. This proposed system focuses on reducing the TEC and increasing the PAR. Figure 10.13 shows that there is a trade-off between TEC and PAR. So, it is necessary to balance these TEC and PAR.

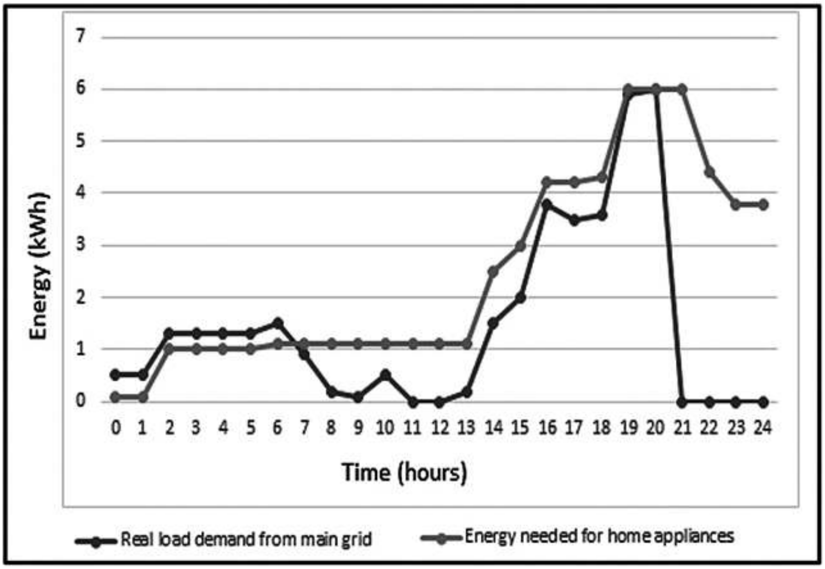


FIGURE 10.12 Load demand from the main grid.

10.5.1 PROPOSED SHEMS WITH DIFFERENT ESS

To reduce the TEC, it is important to consider Ch_{rate} / Dh_{rate} and the ESS capacity.

Figure 10.14 and Table 10.4 show that the average TEC with various Ch_{rate} / Dh_{rate} and the ESS capacity. It is concluded that ESS has no limit for storing energy. The capacity of ESS stores 4 kWh to 6.5 kWh. If Ch_{rate} / Dh_{rat} is large then more energy can be stored in the grid. It stores the energy at the low price time slots and utilizes it for time slots of high price. The Ch_{rate} / Dh_{ra-t} increases with the decrease of PAR, as shown in Table 10.4. When the Ch_{rate} / Dh_{rat} increases, the average TEC decreases. The ESS must store energy with a sufficient capacity.

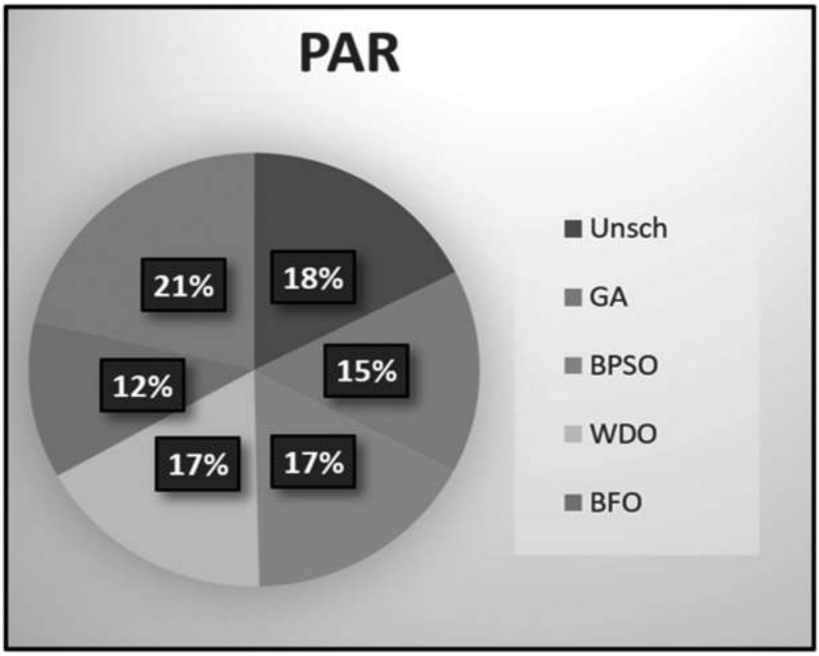


FIGURE 10.13 PAR value comparison.

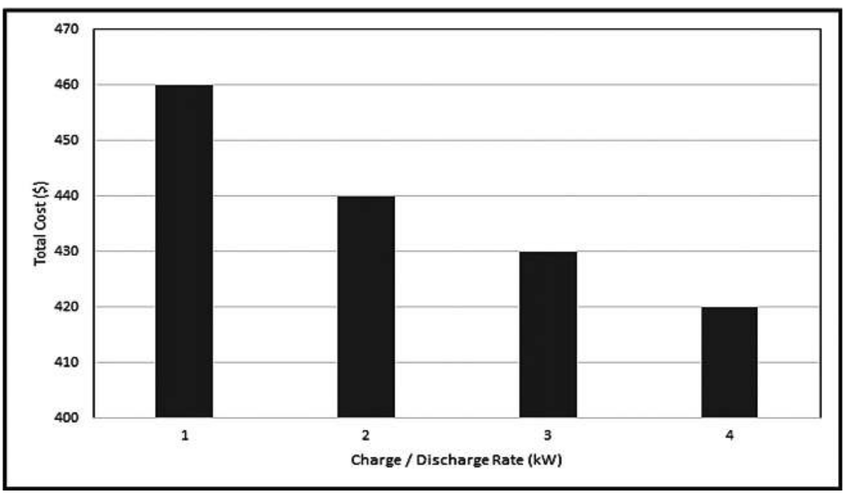


FIGURE 10.14 Average TEC with varying $Ch_{\text{rate}}/Dh_{\text{rate}}$.

TABLE 10.4 Average TEC with Different $Ch_{\text{rate}} / Dh_{\text{rate}}$.

Ch_{rate} / Dh_{rate}	ESS capacity	Average cost	Average PAR
0.0 kW	-	465.84 cents	2.93
0.3 kW	3 kWh	446.78 cents	2.91
0.6 kW	3 kWh no limit	432.61 cents 427.74 cents	2.84 2.73
0.9 kW	3 kWh no limit	426 cents 408.69 cents	2.82 2.66

10.5.2 THE PROPOSED SYSTEM WITH THE SMALLER SELLING PRICE

The proposed system focuses on the smaller selling price and the TEC. If the energy needed for home appliances is high, the energy should not be sent to the outside. The ESS has 0.9 kW of $Ch_{\text{rate}} / Dh_{\text{rate}}$ to store RES energy.

Table 10.5 describes the average TEC with different selling prices. It displays the TEC for $\alpha = 1$, 0.9, and 0.8. The selling prices for $\alpha = 1$, 0.9, and 0.8 are represented by the following equations:

$$P_{\text{sell}}(t) = P_{\text{MG}}(t) \tag{10.39}$$

$$P_{\text{sell}}(t) = 0.9 \times P_{\text{MG}}(t) \tag{10.40}$$

$$P_{\text{sell}}(t) = 0.8 \times P_{\text{MG}}(t)\ \forall t \tag{10.41}$$

If there is a decrease in the selling price, the average TEC will also be increased. The energy of RES is shown in Figure 10.15 by 24 time slots. Figure 10.16 shows the selling energy by doubling the solar cells hourly. The energy received by home appliances from the ESS load is displayed in Figure 10.17. Figure 10.18 shows the energy supplied by main grid that charges ESS with the solar cells.

TABLE 10.5 Average TEC with Different Selling Prices.

α	Max energy cost	Min energy cost	Average energy cost
1	304.01 cents	298.76 cents	302.66 cents
0.9	312.86 cents	307.07 cents	310.44 cents
0.8	323.53 cents	318.08 cents	319.57 cents

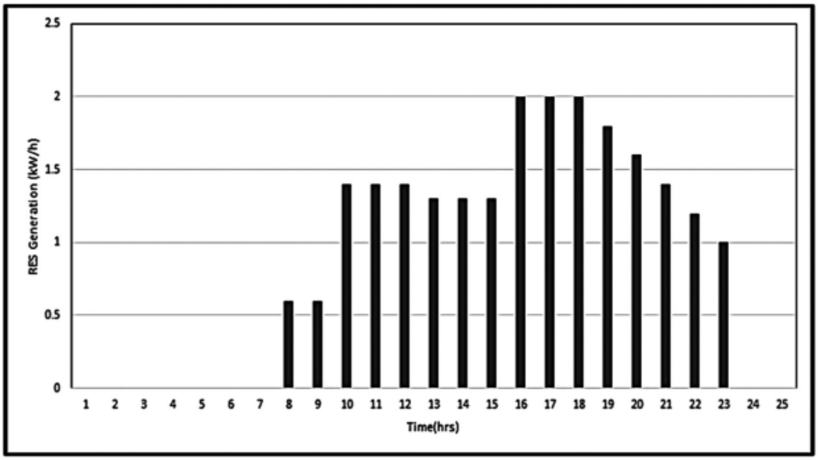


FIGURE 10.15 Rate of generation of ES energy.

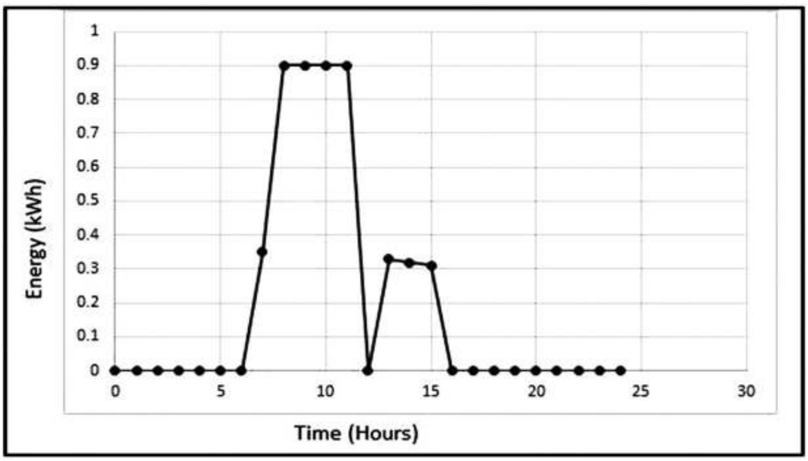


FIGURE 10.16 Rate of generation of selling energy.

10.6 CONCLUSION

The following work shows an implementation renewable energy model in IoT-based environments to improve energy efficiency. Several works before have considerably optimized the energy usage of IoT devices, but the energy produced to power all these systems is mainly from nonrenewable resources. This paper's new framework is designed to limit the usage of

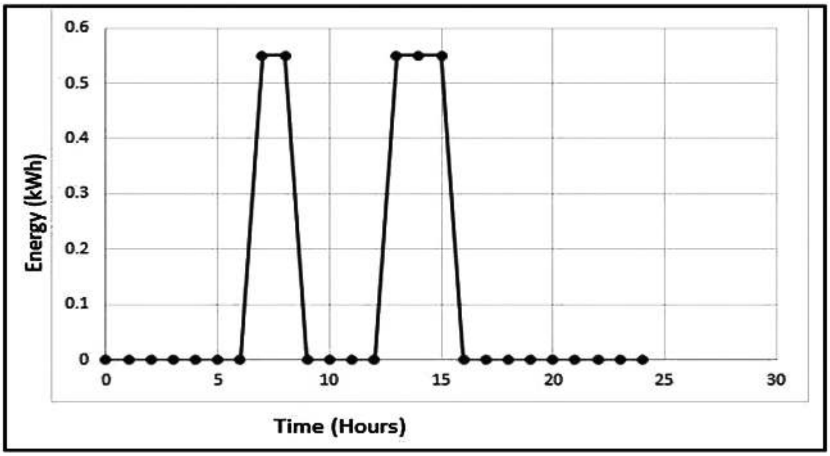


FIGURE 10.17 Generation of ESS energy for home load.

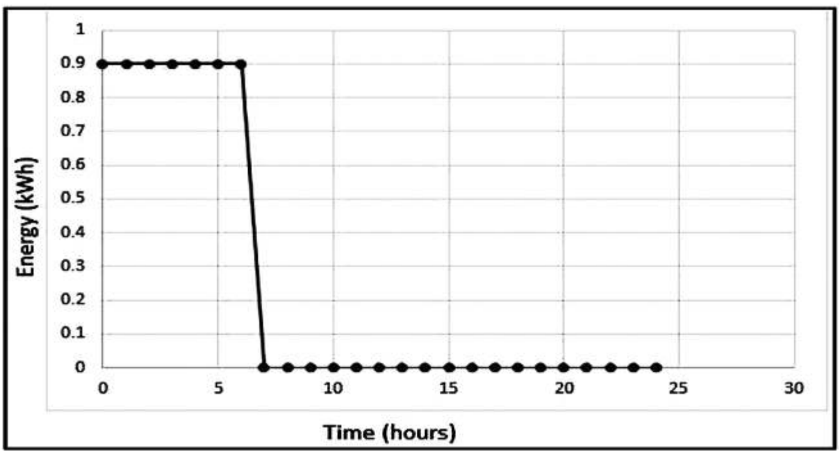


FIGURE 10.18 Amount of energy stored in ESS from the primary grid.

nonrenewable resources and create a renewable IoT-based environment. The proposed framework consists of both a renewable source and a primary grid. An intermediate system manages the energy produced from both systems to provide a sustainable and efficient energy supply. This framework tries to balance the renewable resources' efficiency and stability of the nonrenewable resource. The proposed framework uses a PSO algorithm that optimizes the IoT devices' energy usage and balances the energy demand with the supply. The functioning of this framework is discussed in detail, and the energy requirement of the system is formulated. The PSO algorithm takes advantage of the proposed framework and optimizes the energy usage of IoT devices. The results have identified that the energy generated using the Solar-PV system is more than the demand. It also reduces the energy production cost and provides additional income to the owner since the energy generation is high.

10.7 FUTURE WORK

The importance of implementing renewable energy systems has been discussed in detail, and a new framework employing renewable energy systems is given in this paper. The energy optimization is done through the PSO algorithm that improves energy efficiency but is examined in a software simulation. So, the proposed model can be implemented on a large scale in smart cities, industries, and buildings, and its performance can be verified. However, the proposed model needs to be further optimized using machine learning algorithms for different needs. The proposed framework is designed in such a way to adopt future developments in IoT technologies.

KEYWORDS

- **smart buildings**
- **smart environment**
- **IoT-devices**
- **Internet of Things**
- **energy efficiency**
- **energy optimization**

REFERENCES

1. Khatua, P. K.; Ramachandaramurthy, V. K.; Kasinathan, P.; Yong, J. Y.; Pasupuleti, J.; Rajagopalan, A. Application and Assessment of Internet of Things Toward the Sustainability of Energy Systems: Challenges and Issues. *Sustain. Cities Soc.* **2020,** *53,* 101957.

2. Nižetić, S.; Šolić, P.; López-de-Ipiña González-de-Artaza, D.; Patrono, L. Internet of Things (IoT): Opportunities, Issues and Challenges Towards a Smart and Sustainable Future. *J. Clean Prod.* **2020**, *274*,122877. doi: 10.1016/j.jclepro.2020.122877. Epub 2020 Jul 19. PMID: 32834567; PMCID: PMC7368922
3. Oró, E.; Depoorter, V.; Garcia, A.; Salom, J. Energy Efficiency and Renewable Energy Integration in Data Centres. Strategies and Modeling Review. *Renew. Sustain. Energy Rev.* **2015,** *42,* 429–445.
4. Tarhan, M. Renewable Energy Cooperatives: A Review of Demonstrated Impacts and Limitations. *J. Entrep. Organ. Divers.* **2015,** *4* (1), 104–120.
5. Abdmouleh, Z.; Alammari, R. A.; Gastli, A. Review of Policies Encouraging Renewable Energy Integration & Best Practices. *Renew. Sustain. Energy Rev.* **2015,** *45,* 249–262.
6. Kannan, N.; Vakeesan, D. Solar Energy for Future World:-A Review. *Renew. Sustain. Energy Rev.* **2016,** *62,* 1092–1105.
7. Herez, A.; Ramadan, M.; Abdulhay, B.; Khaled, M. In *Short Review on Solar Energy Systems*, AIP Conference Proceedings, (Vol. 1758, No. 1, p. 030041). AIP Publishing LLC, 2016.
8. Kar, S. K.; Sharma, A.; Roy, B. Solar Energy Market Developments in India. *Renew. Sustain. Energy Rev.* **2016,** *62,* 121–133.
9. Langer, K.; Decker, T.; Menrad, K. Public Participation in Wind Energy Projects Located in Germany: Which form of Participation is the Key to Acceptance? *Renew. Energy* **2017,** *112,* 63–73.
10. Sitharthan, R.; Swaminathan, J. N.; Parthasarathy, T. In *Exploration of Wind Energy in India: A short Review*, *2018 National Power Engineering Conference (NPEC),* (pp. 1–5). IEEE, 2018.
11. Kulkarni, S. H.; Anil, T. R. Renewable Energy in India—Barriers to Wind Energy. *Strateg. Plann. Energy Environ.* **2018,** *38* (2), 40–69.
12. Choi, C. S.; Jeong, J. D.; Lee, I. W.; Park, W. K. In *LoRa based Renewable Energy Monitoring System with Open IoT Platform*, 2018 International Conference on Electronics, Information, and Communication (ICEIC), (pp. 1–2). IEEE, 2018.
13. Kumar, N. M.; Atluri, K.; Palaparthi, S. In *Internet of Things (IoT) in Photovoltaic Systems*, 2018 National Power Engineering Conference (NPEC) (pp. 1–4). IEEE, 2018.
14. Santos, P. M.; Rodrigues, J. G.; Cruz, S. B.; Lourenço, T.; d'Orey, P. M.; Luis, Y.; Barros, J. PortoLivingLab: An IoT-based Sensing Platform for Smart Cities. *IEEE Int. Things J.* **2018,** *5* (2), 523–532.
15. Khyam, M. O.; Li, X.; Pesch, D. Sensor Fusion and State Estimation of IoT Enabled Wind Energy Conversion System. *Sensors* **2019,** *19* (7), 1566.
16. Alhmoud, L.; Al-Zoubi, H. IoT Applications in Wind Energy Conversion Systems. *Open Eng.* **2019,** *9* (1), 490–499.
17. Zia, M. F.; Benbouzid, M.; Elbouchikhi, E.; Muyeen, S. M.; Techato, K.; Guerrero, J. M. Microgrid Transactive Energy: Review, Architectures, Distributed Ledger Technologies, and Market Analysis. *Ieee Access* **2020,** *8,* 19410–19432.
18. Hossein Motlagh, N.; Mohammadrezaei, M.; Hunt, J.; Zakeri, B. Internet of Things (IoT) and the Energy Sector. *Energies* **2020,** *13* (2), 494.
19. Himeur, Y.; Alsalemi, A.; Bensaali, F.; Amira, A. Robust Event-based Non-intrusive Appliance Recognition Using Multi-scale Wavelet Packet Tree and Ensemble Bagging Tree. *Appl. Energy* **2020,** *267,* 114877.

20. Pylsy, P.; Lylykangas, K.; Kurnitski, J. Buildings' Energy Efficiency Measures Effect on CO2 Emissions in Combined Heating, Cooling and Electricity Production. *Renew. Sustain. Energy Rev.* **2020,** *134,* 110299.
21. Eltamaly, A. M.; Alotaibi, M. A.; Alolah, A. I.; Ahmed, M. A. IoT-Based Hybrid Renewable Energy System for Smart Campus. *Sustainability* **2021,** *13* (15), 8555.
22. Syed, A. S.; Sierra-Sosa, D.; Kumar, A.; Elmaghraby, A. IoT in Smart Cities: A Survey of Technologies, Practices and Challenges. *Smart Cities* **2021,** *4* (2), 429–475.
23. Ahmad, A.; Khan, A.; Javaid, N.; Hussain, H. M., Abdul, W.; Almogren, A.; Alamri, A.; Azim Niaz, I. An Optimized Home Energy Management System with Integrated Renewable Energy and Storage Resources. *Energies* **2017,** *10* (4), 549.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

CHAPTER 11

ARCHITECTURES OF COMPUTER NETWORKS WITH STANDARD

MANGESH D. SALUNKE, NILESH N. THORAT, SUHAS R. KOTHAVALE, ROHINI D. MAHALE, RADHIKA M. KOKATE, and VIKAS V CHAVAN

Department of Computer Engineering, Marathwada Mitra Mandal's Institute of Technology, Lohegaon, Pune, Maharashtra, India

ABSTRACT

The hierarchical design of the World Wide Web's protocol stack has a rectangular shape. Although improvements are frequently made at the bottom and topmost levels of the hourglass, the lower layers' protocols seem to have "ossified" at the top. Researchers suggest a model of abstraction for investigating the development of protocol stacks. Network standard is built on a few ideas regarding multilayered network designs and their growth in a market-based ecosystem where protocols compete with one another at the same layer and gain value from the applications running on higher layers. Evo-Arch robustly and from generic initial conditions generates an hourglass architecture resembling the Internet architecture. It also gives a logical justification for how certain protocols, such as transmission control protocol (TCP) or Internet protocol (IP), were able to endure for a far longer period.

11.1 INTRODUCTION

Networking architecture and industry norms are important aspects of the implementation of wire-free sensor networks (WSNs). Due to the severe

Applied Soft Computing Techniques: Theoretical Principles and Practical Applications. Samarjeet Borah, Ratna Raja Kumar Jambi, Sharifah Sakinah Syed Ahmed, & Mahendra Prabhakar Deore (Eds.)
© 2025 Apple Academic Press, Inc. Co-published with CRC Press (Taylor & Francis)

resource limits placed on vertices with sensors, the network arrangement has an essential effect on the use of electricity and, as a result, the efficiency and longevity of the whole thing. On the contrary, an Internet of sensors is composed of numerous sensor nodes that collaborate to carry out a detecting task and are dispersed throughout a sensing range. For the performance of many different connected operations and management tasks, such as synchronization, transformation, media access control, routing, gathering of data, node location, and security for the network, a set of protocols for networking is necessary. Nevertheless, considering that they ignore the power consumption, interpretation, and storage limitations of sensors, the present network needs for traditional wireless communication, including cellphone towers and broadband mobile networks [mobile ad hoc networks (MANETs)], cannot easily be turned into sensor communication. Contrarily, the bulk of sensor assemblies are user-specific with different application requirements. These factors need the development of an entirely novel set of networking standards that not only considers the limited resources of nodes that sense but also distinct networking systems, needs. To make the protocol design for WSNs easier, it is crucial to build a procedure tower. The basic ideas regarding network designs and WSNs software layer are introduced in this chapter. We begin by discussing the layout of sensors and common networking topologies, then we will talk about how networks of sensors are categorized, and at last, we will talk about a protocol stack for sensor networks.

11.2 EXPERIMENTAL METHODS AND WIRELESS SENSOR NETWORK ARCHITECTURES

In this segment, we first designate simple design principles for WSNs initially discussing the structure of a sensor node.

11.2.1 SENSOR NODES STRUCTURE

The four primary components of a sensor node usually exist as shown in Figure 11.1, and are the detecting device, the unit that processes the data, the link to the Internet unit, and the power supply unit. Analog-to-digital converters (ADCs) including several sensors are frequent parts of the sensory unit.

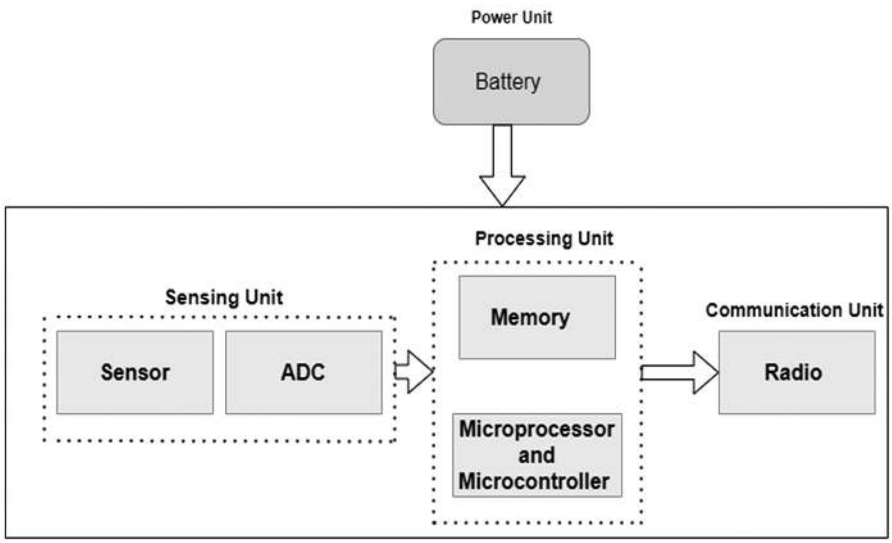


FIGURE 11.1 Architecture for sensor node.

Based on known physical events, the sensors produce analog signals. The ADCs transform analog data converting analog signals into digital signals and feeding those to the computational unit. Usually, the computational component consists of a device called a microcontroller or microprocessor with storage. Such as the StrongARM processor from Intel with the AVR microprocessor, supplying the sensor node with sophisticated management. A short-range radio which communicates with and receives signals via the radio networks serves as a transmission device. A type of battery that serves as the energy source in the power system. In addition to the specific functions, a sensor node may also be outfitted with other components. The use of a global positioning system (GPS) might prove necessary in certain applications which require location data for networks. During various sensing processes, a drive may be required to move the nodes that contain sensors. The idea is to put all these components together into a tiny device with inexpensive manufacturing and consumption of electricity.

11.2.2 NETWORK ARCHITECTURES

As demonstrated in Figure 11.2, the network of devices usually consists of several sink data nodes or sink stations that are situated near/within the sensing zone, as well as a significant number of nodes for sensors that are sparsely distributed across the area of attention. The sink transmits requests or orders to the device nodes in the perceiving expanse so that they can cooperate to finish the procedure of sensing and deliver what was experienced to the proper sink node. Additionally, the sink node serves as an entry point to outside networks like the globe's Internet Web. It collects information retrieved from the nodes that transmit sensors, does certain elementary computations on it, and finally sends the relevant data (or the processed data) to customers that requested it or used the knowledge via the worldwide web. Every node that collects sensors can transmit information over vast distances using just a single hop, which results in the one-hop network design depicted in Figure 11.3. Despite this, energy usage for communication over long distances is expensive. The dynamism used for transmission in networks of sensors is substantially greater than the energy used for sensing than computing. The dynamism prerequisite to transmit a single bit of data to a receiver a 100 m away, for instance, is equivalent to the amount of energy required to carry out 3000 instructions.[2] It may take about 1000 in addition to 10,000 times more energy to go through one bit than it does to transmit it via an Internet connection.[3,4]

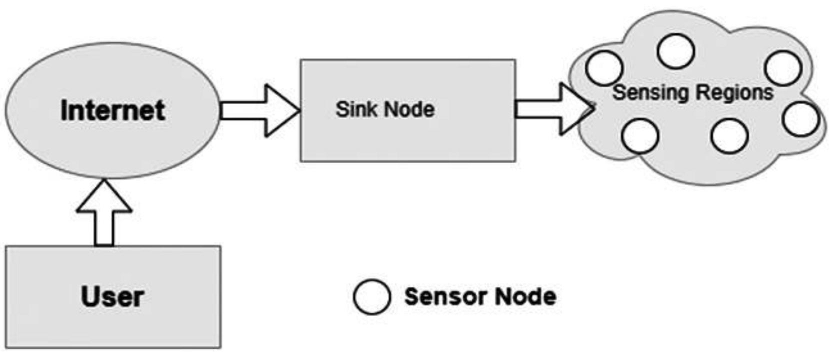


FIGURE 11.2 Network architecture (sensor).

Additionally, most of the energy that is used for exchange is used for transferring, and the essential broadcast influence rises proportionately with the distance traversed. Consequently, it tends to reduce traffic volume and transmit frequency to improve dynamism effectiveness and lengthen the lifespan of the system. In this type of usage, multihop short-distance communications are highly recommended. Multihop short-distance communications are strongly recommended in this kind of usage. Most sensor networks have nodes for sensors that are widely distributed and neighbor nodes that are closer to one another, making short-distance communications practical. The sensor node uses one or more intermediate nodes to transfer the information that it has perceived toward the sink during multihop interaction that may conserve power on transmission. The next two sections (Sections 11.2.2.1 and 11.2.2.2) provide an overview of both the flat and hierarchical[5] forms of multihop network structure.

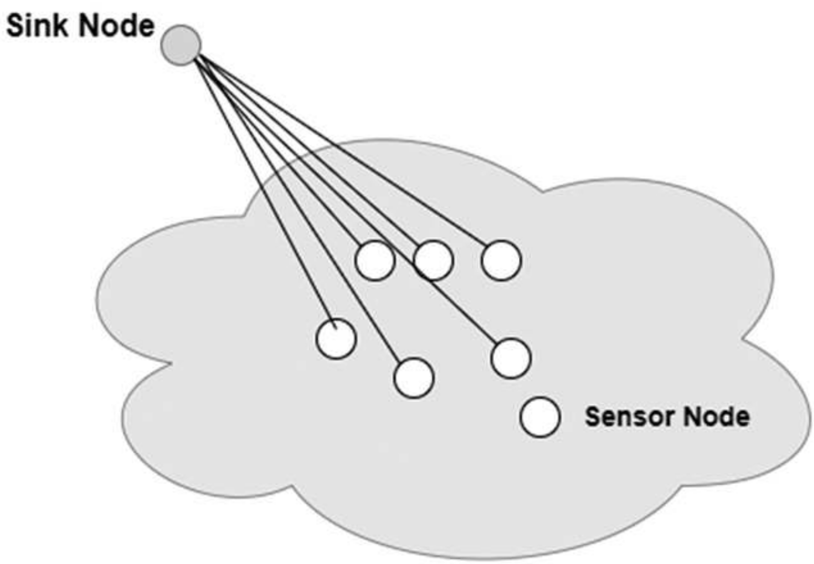


FIGURE 11.3 Network architecture for single hop.

11.2.2.1 NETWORK ARCHITECTURE (FLAT NODE)

Every node in a flat network performs the same function in detecting the job, and all device knobs are peer systems. It is impractical to provide each node in a sensor network with a unique universal identification number due to the sheer quantity of nodes that contain sensors. Due to this, data collection is typically carried out using data-centric routing, in which the information sink sends questions through pouring to every sensor node in the sensory zone and only receives responses from the sensor nodes that have information that matches the request. Every sensor node uses its peer nodes as relays to interact with the sink through a multihop connection. Figure 11.4 depicts the conventional design of a flat network.

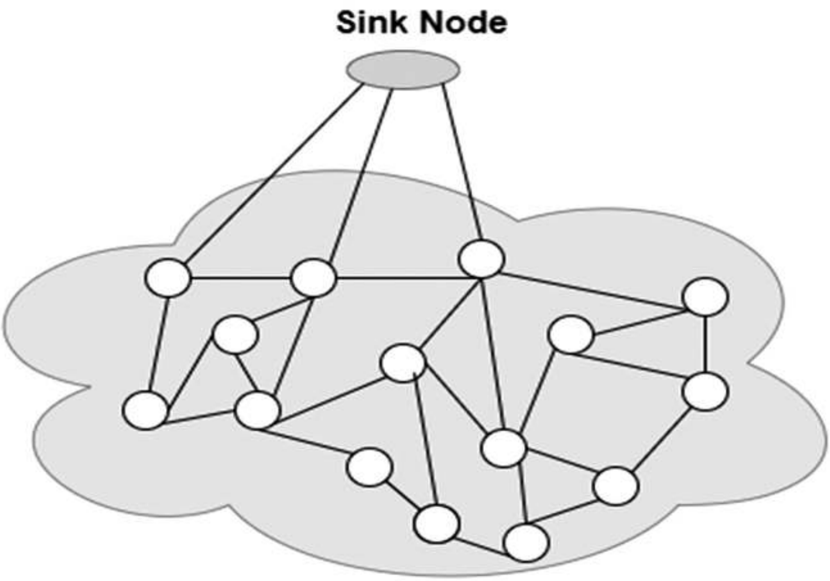


FIGURE 11.4 Network architecture (flat).

11.2.2.2 NETWORK ARCHITECTURE (HIERARCHICAL)

The highest points of the clusters serve as intermediaries in the detail's delivery to the drains as the other cluster members offer their information to them in a hierarchy architecture, which groups nodes with sensors into classifications. Whereas an element having more power can be chosen as the focal point of the cluster for processing the information received collected by the cluster mates and communicating the results of that processing to the source. It is possible to use a node with fewer assets to carry out the detection job and deliver the information gathered directly to the head of the cluster. In addition to lowering the number of resources needed for connection, this approach may help optimize the volume of information as well as increase adaptability as the structure expands.

According to the similar transmission speed between all sensor-related nodes, they are clustered to share the visitor load among each sensor node equally.

To reduce the amount of knowledge delivered to the sinks and improve the structure's power usage, gathering information can also be performed at the group management level.

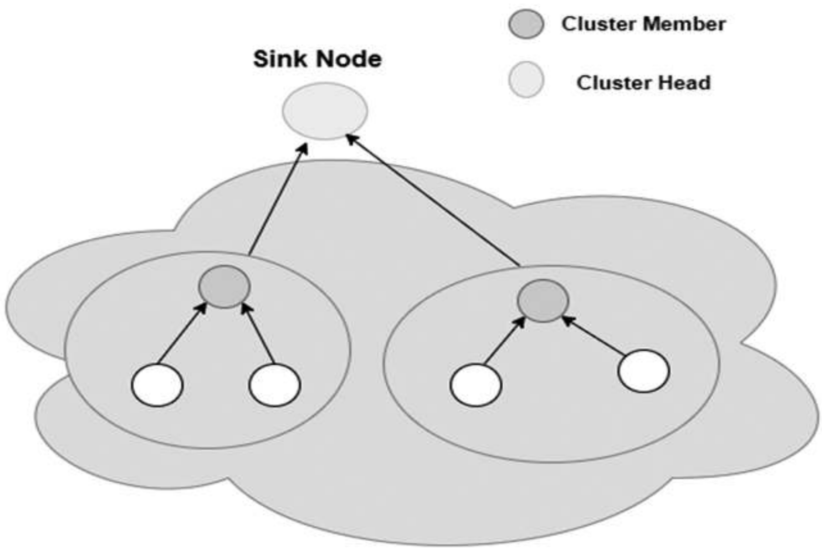


FIGURE 11.5 Cluster network architecture (single hop).

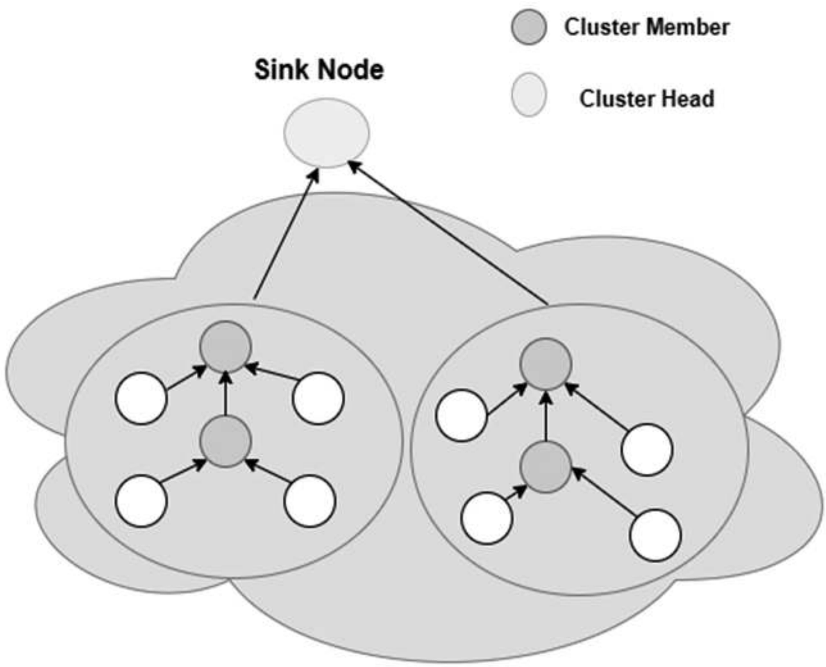


FIGURE 11.6 Cluster network architectures (multihop).

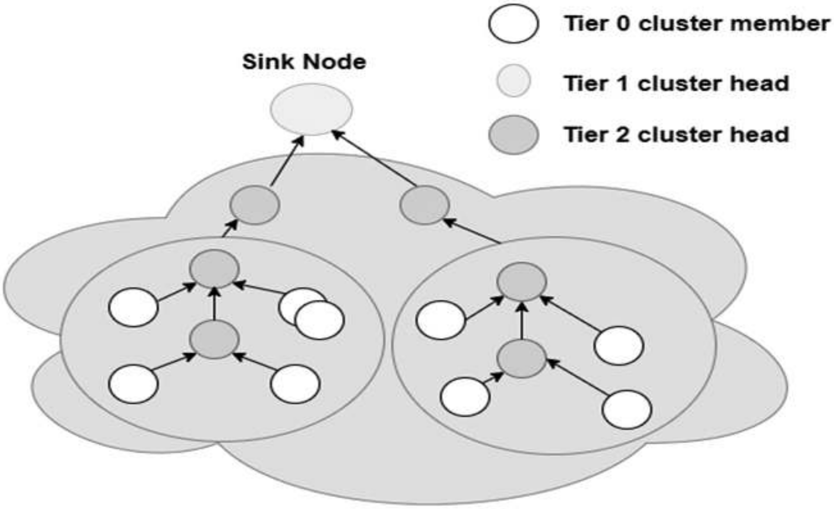


FIGURE 11.7 Cluster network architectures (multitier).

11.3 WIFI CONNECTIVITY SENSOR NETWORKING CATEGORIES

There are WSNs tailored to functions. Due to its specific usage, an arrangement of sensors frequently has a wide range of characteristics. Networks of wireless sensors can be grouped into various groups according to their wide range of characteristics.

- *Mobile and static networks*: According to the mobile nature of the sensing nodes, a network of sensors may be stable or dynamic. A fixed sensor network, as is the case in several industries, has immovable sensor nodes throughout. However, to complete a sensing task, some sensor applications need mobile nodes. Mobile wireless sensor networks, which use independently managed animals as sensors, are a common application.[14] Mobile sensor network design must consider the mobility impact, which makes execution more difficult than static sensor networks, which are easy to monitor and execute.
- *Networks with and without determinism:* The placement of the sensor nodes in an array of sensors determines whether it is predictive or unexpected. Once installed as well as scheduled, the placements of the nodes that collect data in a predictable network of sensors remain permanent. Utilizing this kind of connection is only possible in certain

rare cases wherein organized deployment is practical. However, it might be challenging to install sensor nodes in an organized way due to the adverse or severe conditions that exist in most locations. On the contrary, despite any previous design or technology, sensor nodes are distributed randomly. Unquestionably, unpredictable technologies can adapt along with expandable, although they additionally necessitate more intricate mechanisms for management.

- *Mobile-sink network and static-sink networks*: In a sensor system, an information source might be stationary or moving. A static sink network has a source or sinks that are static and fixedly placed near or inside a sensing zone. Each sensor distributes the data it has composed to the sink(s). A permanent sink clearly improves network operations easier to manage, but it might result in the hotspots effect.[5] The quantity of information that the device knobs must send dramatically increases as their proximity to the information sink gets closer. Due to such nature, the sensors on the nodes closest to the knowledge sink usually die first, leading to connectivity segmentation and on rare occasions even interference with normal communications. In a "mobile sink" arrangement, the sink(s) roam across the intuiting area collecting information from sensor nodes, reducing a load of traffic burden on the nodes that collect statistics and minimizing the hotspot impact of the entire system.
- *Multiple-sink and single-sink networks*: One sink or several drains can be present in a network of sensors. In a single-sink system, the sensor region's closest or largest sink is present. Every sensor node sends the records it has composed to this sink. A multilink link may contain several sinks spread out across several locations near or around the experience zone. Sensor arrays may guide their documents to the contiguous sink for the purpose of successfully regulating the flow of data capacity on the monitoring nodes and lessen the hotspot impact in an Internet connection.
- *Networks with single hop and several hops*: Sensor networks can be categorized as single-hop or multiloop depending on how many hops there are amongst a sensor node in addition to the records sink. In a single-hop system, all the detectors' communicate their experienced information straight to the source, simplifying network management. Simpler to put into operation. However, this calls for the distribution of knowledge across great distances, which is inflated in terms

of mandatory hardware and electrical consumption. The hubs that are farthest away from the information route will fail considerably quicker compared to those that are near it. Additionally, when the network grows larger, the total traffic volume can rise quickly. This would lead to more collisions that would result in higher energy use and shipment expectancy. In a multiloop system, antenna nodes use wireless communications that are short-range to guide the documents they have perceived to sink through one or more intermediary nodes. The data must be routed and forwarded via a multiloop path by each intermediary node. In addition, data aggregating can be carried out at an intermediary node to remove redundant data, this can lower overall network traffic and boost the network's energy consumption. A single-hop network typically has a simplified network design and becomes simpler to manage as a result. It is appropriate for use in applications requiring small sensing regions and sparsely placed sensor nodes. A broader variety of applications is possible with multiloop systems, but control complexity is higher.

- *Networks that are self-reliable and those that are not*: According to just how dependable the nodes that sense is a network of sensors may be self-configuring or not. The ability to organize a network is lacking in the sensor nodes in nonself-configurable networks. They must instead depend on a centralized regulator to succeed each antenna node and gather data from them. As a result, only small-scale networks are appropriate for this type of network. While collaborating to carry out a detecting task, the majority of most sensor networks allow nodes that sense to autonomously organize and preserve their connection to the Internet. This level of confidence in the system makes it proper for significant systems to handle difficult recognizing tasks.
- *A system that is both uniform and varied*: How heterogeneous or homogeneous a sensor network is will depend on the fact that every sensor node has the same features.[15] Each node with sensors connected to an identical connection has equal access to power, processing speed, and capacity. On the contrary, a network that is diverse has certain sophisticated nodes for monitoring that are more capable than standard detecting cells in terms of communication and computing power. In this case, a network might assign additional processing and communication duties to such advanced nodes to boost their use of energy and, as a result, expand their lifespan.

11.4 WIRELESS SENSOR COMMUNICATION PARADIGM ARCHITECTURE

As shown in Figure 11.8, the “physical layer”, which includes the data connection layer, the “network layer’s “transport layer,” and then the “application layer” makes up the protocol’s foundation for WSNs. Several application-layer abilities can be used to create applications for different sensor networks. The transportation layer oversees giving the application layers safe data transfer. Managing information transmission utilizing the transport layers is a task performed at the network layer. Data multiplication, communication framework delivery, including receiving data, multimedia entry, and handling mistakes are all specifically handled by the data link layer. The physical layer of a signal is in authority for creating frequencies, modifying messages, sending and receiving signals, transmitting and receiving signals, encrypting information, and other activities involved in sending and receiving signals through a physical communications channel.

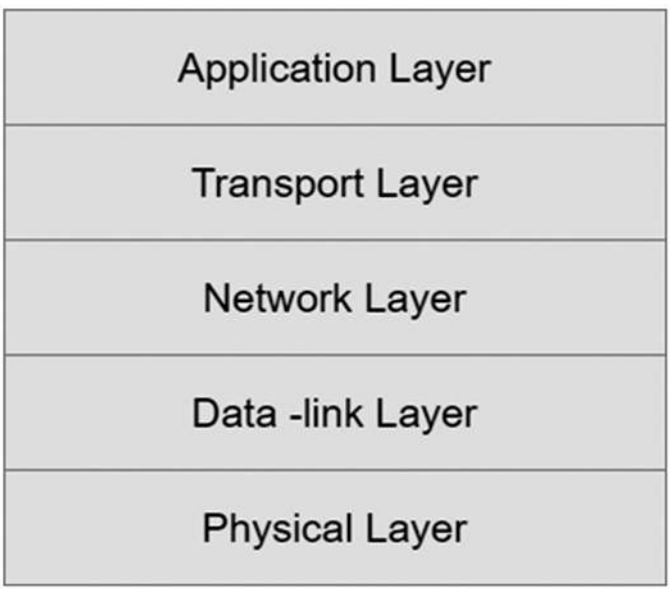


FIGURE 11.8 Protocol-layered stack.

The communication framework, on the opposite end of the spectrum, is composed of several operational layers that span every single layer,[16] including layers for strength, connections, and managing tasks. The energy administration carrier is responsible for managing the energy that is used by a sensor node for gathering, analyzing, exchange, and receiving. It does this by employing efficient power control techniques at various communication layers. In this scenario, a sensor node may switch off its wireless transmitter if there is not any data to send or acquire at the medium access control (MAC) level. Depending on which of its contemporaries has the most power left, any sensor node at the network's surface may choose to make its subsequent hop to the descent across the node in question. When a network's design changes as a result of the establishing of new nodes in a failure of existing nodes, the relocation of present nodes, etc., a communication preservation plane is in charge of rearranging sensor nodes to establish and sustain connections to the resource. A task administrative transport is in charge of allocating tasks across sensor nodes in a detecting domain to reduce energy consumption and increase the network's general service life. They are however not all required to complete the work because nodes for sensors are often widely distributed throughout sensing territories and are interchangeable for carrying out a sensing activity.

11.4.1 APPLICATION LAYER

The application layered has many distinct application-layer technologies that are utilized for sensor network features with the value demand dissemination, node localization, time synchronization, and protection of the system. As an illustration, the sensor authority regulation (SMP)[1] is an application-layer handling practice that permits programs to carry out a number of tasks, such as providing location-related knowledge, coordinating data between proceeding nodes of sensors, which are organizing sensors nodes, and applying for the current state of sensors. The user-friendly techniques for asking, answering, and gathering answers to inquiries are provided by the sensor inquiry along with data broadcasting program (SQDDP). Sensors look up in addition to assignment technologies (SQTL) are used to create a gateway for wireless sensor networks.[17,22]

11.4.2 TRANSPORT LAYER

The responsibility for assuring dependable end-to-end data transmission between the monitoring networks and the sink(s) usually lies on the communication protocol layer. Traditional protocols for transport cannot be used in sensor networks without modification due to the power, computation, and memory constraints of sensor nodes. By way of example, the traditional "transport control protocol (TCP)" methods for end-to-end retransmission-based detection of errors including window-based relief of congestion are not suitable for connected sensors since they are inefficient at managing bandwidth. In contrast, sensor interconnections are often utilized for specific purposes. An antenna system is typically developed for a particular sensing application, such as habitat management, tracking inventories, or combat intelligence. Programs may have different dependability needs, resulting in an important effect on how transport-layer procedures are designed. In networks of sensors, transmission of information occurs primarily in either the upstream or downstream directions. The preceding sensor nodes transmit the information they have sensed to the sink(s), while the corresponding downstream information provided by the sink(s), including those, that passes through the sink(s) to the origin sensor the nodes as requests, instructions, and program packages are supplied from the sink(s) to the parent sensor nodes. There may be different criteria for the data's trustworthiness. Because sensed data tends to be related or redundant to some extent, data flows in the upward instructions, for instance, are loss resistant.[23] To the sensor nodes that often demand 100% dependable transportation, the data streams in the downstream are instead inquiries, instructions, including program binaries.

11.4.3 NETWORK LAYER

However, this level oversees sending the knowledge received by the first node of sensing to the proper information sink(s). A monitoring network's detecting zone has sensors that are placed there to search for an unusual occurrence. The detected event or data must be delivered to the information sink. In principle, the choice between a multihop short-range connection and a single-hop long-range wireless communication can be used by the first station to communicate information detected to the sink immediately. However, long-range wireless communication is expensive for sensor nodes in terms of both energy consumption and installation complexity. Contrarily, multihop short-range interactions may significantly reduce the electrical power consumption of sensor nodes while effectively minimizing the effects of channel loss and signal transmission. The extensive deployment of data collection nodes and the proximity of neighboring nodes make multihop communication over short distances possible in sensor networks. In this scenario, the starting node must use a method of routing to choose an energy-efficient multihop route from the node to the source to send the information it has sensed there.[24]

However, typical wireless network routing strategies are not suitable for sensor networks since they neglect energy saving. In addition, multihop connections that have multiple-to-one connections combined significantly increase the amount of traffic connected to transit. There is an increased likelihood of network traffic congestion, collisions, loss, delay, and other issues when information migrates deeper into the sink. The sensor nodes closer to the sink, frequently within a few hops, will drastically reduce the functional lifetime of the network since they will lose more packets and consume more energy. The energy limits placed on sensor nodes as well as the unique transport pattern must be considered to successfully plan network layers such as firewalls and routing protocols. Numerous routing protocols have been created as a result of substantial research in this field to accommodate the various application scenarios that sensor networks may manage.[25]

11.4.4 DATA LINK LAYER

The data link layer, which enables trustworthy point-to-point or point-to-multipoint communications, is responsible for information stream multiplexing, information frame production and identification, media access, as well as error correction. The MAC function is one of the most crucial elements of the digital connection layers. The shared channel or media is judiciously distributed among several sensor nodes to maximize network efficiency in terms of power consumption, bandwidth utilization, and transmission latency. However, as they do not account for the unique characteristics of sensor networks, particularly the energy restriction, MAC standards for ordinary wireless networks cannot be used in networks of sensors without modification. Energy efficiency is only somewhat relevant because base stations have no power restrictions and mobile customers can recharge their phone power sources. In MANETs, mobile nodes are equipped with portable rechargeable battery-powered devices. Since the major goal of sensor networks is to conserve electricity to prolong the network's lifespan, conventional MAC protocol specifications are inappropriate for these networks. As a result, extensive research has been done on MAC and several MAC protocols. The data link layer also plays a significant role in regulating transmission faults. Often, the sensor network is installed in challenging conditions with erratic wireless connectivity. In this incident, the fault mechanism becomes essential and critical for creating connection dependability or trustworthy data flow. The two primary methods for error prevention were forward correction of errors (FEC). By resending lost data sessions or transmissions, ARQ encourages the proper transfer of information. It goes without saying that this leads to significant energy consumption and repetition costs, making it inappropriate for sensor networks. To increase network dependability, FEC employs error control codes during data transport, complicating the decoding and encoding procedures and requiring more computing power from sensors. FEC can be used to significantly reduce the bit error rate (BER) of the channel of communication for any certain broadcast strength. Because of the power constraints of sensor nodes, FEC is still a very effective approach to error correction in sensor networks. When creating an FEC approach, it is essential to use the right error control code since the right error control code can produce a virtuous coding improvement and a sizable reduction in BER. It is also important to consider the additional processing power required for encoding and decoding. To have a reliable, environmentally friendly, with low-complexity FEC approach, a trade-off involving the extra computational capacity and the corresponding coding advantage must be regulated.

11.4.5 PHYSICAL LAYER

Bits obtained from the information at the connecting surface must be transformed into indications that can be transmitted via the communication medium by the physical layer. Designed for this goal, it must be arranged with a variety of associated issues, such as signal identification and modulation, the development of carrier frequencies, data encryption, and the selection of the transmission medium and frequencies. The component architecture and the numerous mechanical and electrical connections must also be taken into consideration. The selection of medium and frequency has a big impact on how sensor nodes interact with one another. Use of broadcasting and the industrial, scientific, and medical (ISM) frequencies, which are frequently license-free, is one option. The unrestricted use, wide spectrum, as well as accessibility of the ISM frequencies are their key benefits.[18] However, some communication technologies, such as wireless telephone organizations besides wireless local area networks (WLANs), currently employ the ISM frequencies. The ISM frequencies' main advantages are their unrestricted use, broad spectrum, and accessibility.[18] However, several forms of communication currently use ISM frequencies, including WLANs and wireless telephone services. However, a compact, inexpensive, and ultralow frequency broadcaster is necessary for wireless sensor networks. These reasons have led to suggestions for the implementation of the ISM bands at 433 and 917 MHz in North Europe and the US, respectively. Many projects to develop hardware for antenna knobs have used transistor rate [radio frequency (RF)] circuits, such as the m AMPS project[19] and in,[20] wherever the antenna node uses a single-channel RF transmitter that works at 916 MHz, to radio, optically or infrared media may be a choice for instance, the "Smart Dust project"[21] used optical transmission to send data. Nevertheless, because together requires the transmitter and the recipient to be in the visual range to communicate, their application is rather limited.[22]

11.5 SUMMARY

A network's structure has an important effect on how much energy sources are consumed and, consequently, how long a wireless sensor network can operate. Multichip short-distance communications have become popular in networks of sensors due to the dynamism restriction in the knobs and the distinctive multiple to unique transportation configuration. In multihop networks, an ordered network design built around clustering can balance flow or load, enhance adaptability, and decrease energy usage for communications. On the contrary, to carry out various networking management and oversight tasks, networks of sensors need an entirely novel set of networking protocols. Such protocols must consider both the specific to the application characteristics of sensors network as well as the constraints on resources. The subsequent sections of this book can be understood with the aid from this section, that introduced basic principles on network structure and protocols stacks for networks of sensors.

KEYWORDS

- **architecture**
- **computer networks**
- **protocol stacks**
- **EvoArch**
- **TCP**

REFERENCES

1. Akyildiz, I. F. A Survey on Sensor Networks. *IEEE Commun. Mag.* **2002,** *40* (8), 102–114.
2. Pottie, G.; Kaiser, W. Wireless Integrated Sensor Networks (WINS). *Commun. ACM* **2000,** *43* (5), 51–58.
3. Hill, J.; Szewcyk, R.; Woo, A.; Culler, D.; Hollar, S.; Pister, K. In *System Architecture Directions for Networked Sensors*, Proceedings of 9th International Conference on 32 Network Architectures and Protocol Stack Architectural Support for Programming Languages and Operating Systems (ASPLoS IX), Cambridge, MA, Nov. 2000, pp. 93–104.
4. Merrill, W. M.; Sohrabi, K.; Girod, L.; Elson, J.; Newberg, F.; Kaiser, W. In *Open Standard Development Platforms for Distributed Sensor Networks*, Proceedings of SPI—Unattended Ground Sensor Technologies and Applications IV (AeroSense 2002), vol. 4743, Orlando, FL, Apr. 2002, pp. 327–337.
5. Al–Karaki, J. N.; Kamal, A. E. Routing Techniques in Wireless Sensor Networks: A Survey. *IEEE Wirel. Commun.* **2004,** *11* (6), 6–28.
6. Rajagopalan, R.; Varshney, P. Data - Aggregation Techniques in Sensor Networks: A Survey. *IEEE Commun. Surv. Tut.* **2006,** *8* (4), 48–63.
7. Abbasi, A.; Younis, M. A Survey on Clustering Algorithms for Wireless Sensor Networks. *Com. Commun.* **2007,** *30* (14–15), 2826–2841.
8. Gupta, G.; Younis, M. In *Load-Balanced Clustering of Wireless Sensor Networks*, Proceedings of 2003 IEEE International Conference on Communications (ICC' 03), Anchorage, AK, May 2003, pp. 1848–1852.
9. Bandyopadhyay, S.; Coyle, E. J. In *An Energy Efficient Hierarchical Clustering Algorithm for Wireless Sensor Networks*, Proceedings of IEEE INFOCOM' 03, 3, San Francisco, Mar.–Apr. 2003, pp. 1713–1723.
10. Heinzelman, W.; Chandrakasan, A.; Balakrishnan, H. An application - Specifi c Protocol Architecture for Wireless Microsensor Networks. *IEEE Trans. Wirel. Commun.* **2002,** *1* (4), 660–670.
11. Younis, O.; Fahmy, S. Heed: A hybrid, energy-efficient, Distributed Clustering Approach for ad-hoc Sensor Networks. *IEEE Trans. Mob. Comput.* **2004,** *3* (4), 366–379.
12. Wang, P.; LI, C.; Zheng, J. In *Distributed Minimum-Cost Clustering Protocol for Underwater Sensor Networks (UWSNs)*, Proceedings of 2007 IEEE International Conference on Communications (ICC ' 07), Glasgow, UK, June 2007, pp. 3510–3515.
13. Banerjee, S.; Khuller, S. In *A Clustering Scheme for Hierarchical Control in Multi-Hop Wireless Networks*, Proceedings of IEEE INFOCOM ' 01, Anchorage, AK, Apr. 2001, pp. 1028–1037.
14. Li, Y.; Panwar, S. S.; Mao, S.; A Wireless Biosensor Network Using Autonomously Controlled Animals. *IEEE Netw.* **2006,** *20* (3), 6–11.
15. Nakayama, H.; Ansari, N.; Jamalipour, A.; Nemoto, Y.; Kato, N. Fault-Resilient Sensing in Wireless Sensor Networks. *Com. Commun.* **2007,** *30* (11), 2375–2384.
16. Ruiz, L. B.; Nogueira, J. M.; Loureira, A. A. F. Sensor Network Management, *SMART DUST: Sensor Network Applications, Architecture, and Design (edited)*, CRC Press: Boca Raton, FL, 2006.
17. Shen, C.; Srisathapornphat, C.; Jaikaeo, C. Sensor Information Networking Architecture and Applications. *IEEE Pers. Commun.* **2001,** *8* (4), 52–59.
18. Su, W.; Ö. B. Akan, Cayirci, E. *Communication Protocols for Sensor Networks*, Wireless Sensor Networks (edited), Kluwer Academic Publishers: Norwell, MA, 2004.

19. Shih, E.; Cho, S.; Ickes, N.; Min, R.; Sinha, A.; Wang, A.; Chandrakasan, A. In *Physical Layer Driven Protocol and Algorithm Design for Energy-Effi cient Wireless Sensor Networks*, Proceedings of ACM Mobicom ' 01, Rome, Italy, July 2001, pp. 272–286.
20. Woo, A.; Culler, D. In *A transmission Control Scheme for Media Access in Sensor Networks*, Proceedings of ACM Mobicom ' 01, Rome, Italy, July 2001, pp. 221–235.
21. Kahn, J. M.; Katz, R. H.; Pister, K. S. J. In *Next Century Challenges: Mobile networking for Smart Dust*, Proceedings of ACM Mobicom '99, Washington, DC, 1999, pp. 271 – 278.
22. Zhao, F.; Guibas, L. *Wireless Sensor Networks: An Information Processing Approach*. Morgan Kaufmann Publishers: San Francisco, CA, 2004. Wang, L.; Li, J.; Wang, Y.; Zhao, L.; Jiang, Q. Adsorption Capability for Congo Red on Nanocrystal HneMFe204 (M=Mn, Fe, Co, Ni) Spinel Ferrites. *Chem. Eng. J.* **2010**, doi: 10.1016/j.cej.-2011.10.088
23. Salunke, M.; Kabra, R.; Kumar, A. Layered Architecture for DoS Attack Detection System by Combined Approach of Naive Bayes and Improved K-means Clustering Algorithm. *Int. Res. J. Eng. Technol.* **2015,** *2* (3), 372–377.
24. Mangesh, D. S.; Kumar, P. A Proposed Methodology to Mitigate the Ransomware Attack. *In Recent Trends in Intensive Computing*, IOS Press, 2021; pp 16–21.
25. Salunke, M. D.; Kabra, R. Denial-of-service Attack Detection. *Int. J. Innov. Res. Adv. Eng.* **2014,** *1* (11), 16–20.

CHAPTER 12

DRONE-ENABLED PRECISION AGRICULTURE: REVOLUTIONIZING CROP MANAGEMENT AND YIELD OPTIMIZATION

ROHINI MAHALE[1], KSHITIJ PATANGE[1], SHUBHAM PHAD[1], ADITYA MORE[1], and VIKAS CHAVAN[1]

[1]Department of Computer Engineering, Marathwada Mitra Mandal's Institute of Technology College of Engineering, Pune, Maharashtra, India

ABSTRACT

Autonomous vehicles are autos that can drive and navigate on their own without a driver. They monitor and analyze their environment using sensors, cameras, and artificial intelligence, then act accordingly. Increasing safety, reducing traffic congestion, and enhancing the effectiveness of transportation networks are the objectives of autonomous cars. There are still technical, legal, and ethical challenges to be addressed, but the development of autonomous vehicles has the potential to revolutionize the way we travel.

Quadcopters are increasingly being used for aerial surveillance, reconnaissance, and inspection due to their adaptability to challenging and dangerous conditions. Manually operating these quadcopters, however, may be difficult, especially when there are distractions or unforeseen occurrences present. There is rising interest in creating autonomous quadcopters that can navigate and carry out activities on their own to address these issues and boost the effectiveness of quadcopter operations.[1]

This research paper presents the design and development of an autonomous quadcopter using an Arduino Uno micro-controller board. This quadcopter is

Applied Soft Computing Techniques: Theoretical Principles and Practical Applications. Samarjeet Borah, Ratna Raja Kumar Jambi, Sharifah Sakinah Syed Ahmed, & Mahendra Prabhakar Deore (Eds.)
© 2025 Apple Academic Press, Inc. Co-published with CRC Press (Taylor & Francis)

equipped with ESP32 cam model for crop monitoring. For fully autonomous operation on a quadcopter With the Arduino Integrated Development Environment (IDE) and an exclusive software component every Arduino Uno is controlled.

12.1 INTRODUCTION

Increasing food demand and the scarcity of resources such as water and land are two issues that the important industry of agriculture must deal with. Farmers are utilizing cutting-edge technology such as drones for crop monitoring and analysis, or unmanned aerial vehicles (UAVs), to overcome these difficulties. Due to its stability, maneuverability, and simplicity of use, quadcopters have become one of the most widely used forms of drones. Quadcopters can gather information on crop health, growth, and yield using imaging and sensing technology including high-resolution cameras and multispectral sensors, giving farmers invaluable field-level insights. With the use of this technology, resources such as water and fertilizer can be used more precisely and effectively while also enabling farmers to make more educated crop management decisions.

Two difficulties that the significant sector of agriculture must address are the rise in global food demand and the shortage of resources such as water and land. To get around these challenges, farmers are turning to cutting-edge technology such as drones for crop monitoring and analysis or UAVs. Quadcopters have grown to be one of the most popular types of drones because of their stability, maneuverability, and ease of usage.[2] Using image and sensing technologies, such as high-resolution cameras and multispectral sensors, quadcopters may collect data on crop health, growth, and yield, providing farmers with crucial field-level insights. This technology enables farmers to make more informed crop management decisions while also allowing for the more accurate and effective use of resources such as water and fertilizer.

In this work, we have made a quadcopter that can fly autonomously over a given path and perform crop monitoring, with use of OpenCV.

12.2 EXPERIMENTAL METHODS AND MATERIALS

12.2.1 HARDWARE COMPONENTS

Sensors and computer platforms are necessary for UAVs. The sensors are typically mounted on computing devices such as the Arduino Uno. In our project, we have used Arduino Uno, four brushless DC electric motor (BLDC) motors (2200 kV each), four ESC's (30A each), ESP32 cam model, etc. We can also use some sensors like GPS receivers and IMU to control the platforms.

Now, just taking pictures may be used to gather crucial agricultural information. Even from a great distance, a visible-light camera can capture a picture with clarity and detail. Additionally, by employing multiple camera types, such a multispectral camera, the needed information may be gathered.[3] To confirm the topography of rice fields and crop development, multithermal cameras are also used. When employing a laser to map in three-dimensions (3D), a light detection and ranging (LiDAR) device spins 360. LiDAR is a crucial sensor for landscape mapping and reconnaissance.

12.2.2 CONTROL OF AGRICULTURAL UAVS

Agricultural UAVs do not idle while in flight and use a battery with a restricted capacity. As a result, several studies on control systems that are utilized to maximize agricultural efficiency have been carried out. The supervision of farming UAVs requires the employment of flight technological advances, including aspect and elevation regulations, navigational devices, obstacle detection and avoidance systems, decision-making and judgment, and large-scale supervision. These are other electronic devices for GCS communication of data. We shall now discuss quadcopters, which are well-known UAV.[4]

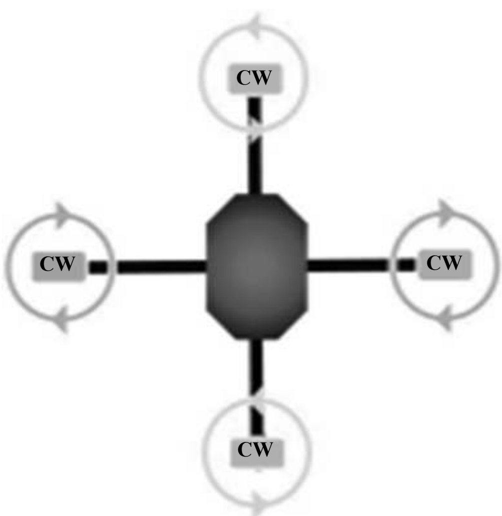


FIGURE 12.1 Quadcopter plus-type model.

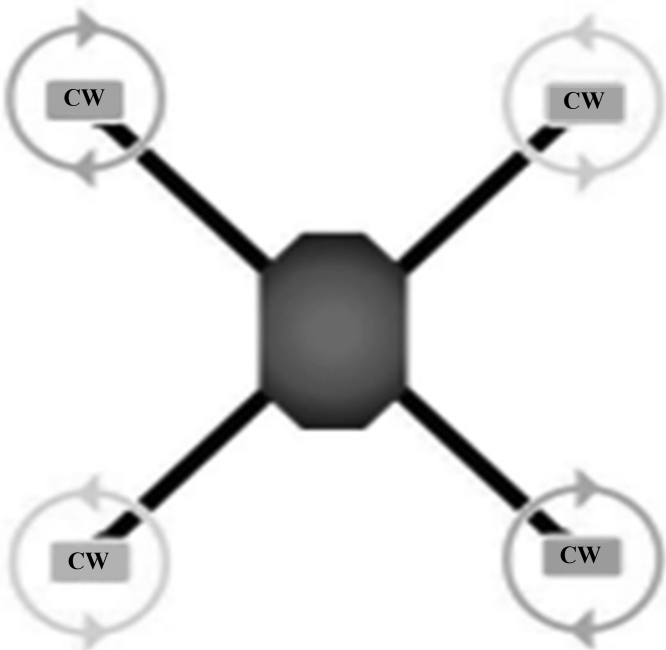


FIGURE 12.2 Quadcopter cross-type.

12.2.3 HARDWARE CONFIGURATIONS

We had four ESC's of 30A, that needed to be used with our model. For that, we had to calibrate it to test on what voltage and current it can operate. For that, we used Arduino IDE and made connections to calibrate it as per our needs.

Also, we performed testing of our ESP32 cam model by running the camera web server. The ESP32 cam model was of the Wrover model which needed to be connected appropriately with Arduino Uno to test its working.[5]

12.3 DISCUSSION

12.3.1 IMPACT OF OPERATING PARAMETERS

12.3.1.1 SPEED

The operating speed of an autonomous vehicle may have a considerable influence on passenger comfort, energy efficiency, and safety. Higher speeds might result in more collisions, a shorter driving distance, and passenger discomfort from vibration and noise. Lower speeds, on the other hand, can result in longer travel times and might not be ideal for some uses, such long-distance travel.

12.3.1.2 ACCELERATION AND DECELERATION

An autonomous vehicle's ride quality, energy efficiency, and safety are all impacted by its acceleration and deceleration speeds. Passengers may feel uncomfortable from sudden acceleration and deceleration, which can also restrict range and increase the possibility of accidents. On the contrary, sluggish acceleration and deceleration may lengthen travel times and make it harder for the car to keep up with traffic.[6]

12.3.1.3 ROUTE PLANNING

Algorithms for route planning are used by automated vehicles to find their way to their destination. The precision of these algorithms determines the effectiveness and safety of the vehicle's functioning. To achieve optimal performance, route planning algorithms must take into account elements such as the state of the roads, traffic volume, and the fuel consumption of the vehicle.

12.3.1.4 COMMUNICATION WITH OTHER VEHICLES AND INFRASTRUCTURE

To increase efficiency and safety, automated cars may connect with other vehicles and infrastructure. Vehicle-to-vehicle communication, for instance, can warn drivers about possible dangers or traffic jams, while vehicle-to-infrastructure communication can improve traffic flow by modifying traffic lights depending on the current volume of traffic.

12.3.1.5 ENVIRONMENTAL FACTORS

The weather and road conditions, for example, can have a big impact on how safely and effectively autonomous cars operate. Snow, ice, rain, and fog can impair sight and the traction of the vehicle, and uneven or broken road surfaces can be uncomfortable for drivers and passengers and raise the possibility of accidents.[7]

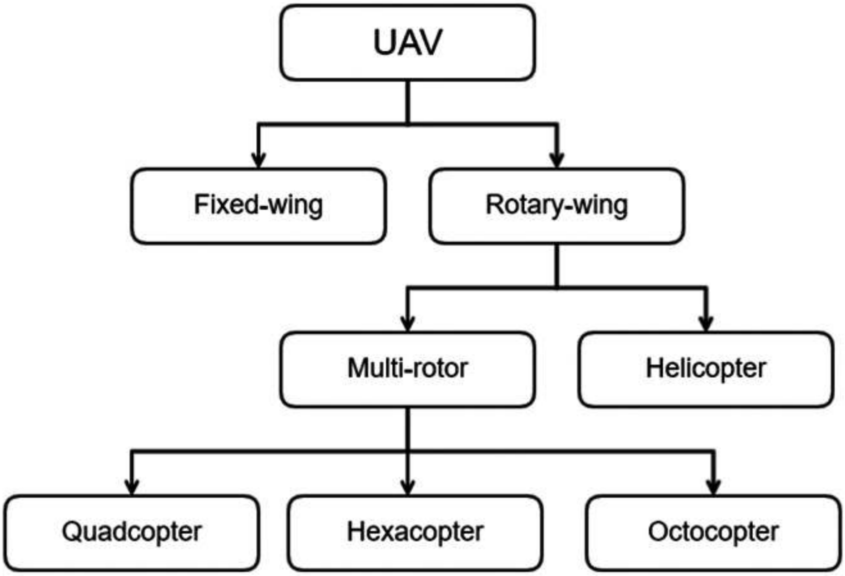


FIGURE 12.3 Types of drones.

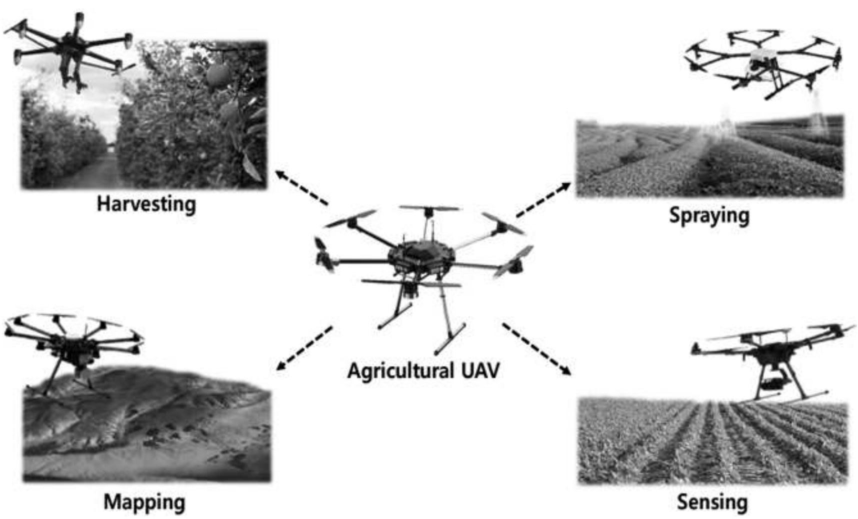


FIGURE 12.4 Usage of UVAs for agricultural UAVs.

12.3.2 MATERIALS AND METHODS

An explanation of the research area and actual information gathering. Aerial shots as on-the-ground measurements for this research were collected in large peach and citrus orchards totaling 90 and 80 ha, accordingly, in the southern Spanish cities of Cordoba (37°48 0 N, 4°48 0 W) and Seville (37°20 0 N, 5°50 0 W), in 2007 and 2008. With an average annual rainfall of 600 mm, the region has a climate that is similar to the Mediterranean, with most of the rain falling between the months of autumn and spring. The *Prunus persica* (L.) Batsch orchard research The *P. persica* (L.) Batsch orchard research was conducted in peach and nectarine plantations, respectively, using the peach variety Babygold 8 and the nectarine variety Sweet Lady. The pears as well as nectarine trees were planted in 1993 on dense alluvial soil at intervals of 5 m by 93.3 m and 6 m by 93.3 m, respectively, in lines pointing the other way.[8]

Eight blocks of data, with trees varying in LAI from 2 to 4.2, height between 2.2 and 4 m, and horizontal crown diameters of 1.4–3.5 m, were chosen to replicate a peach orchard. The experiment, which occurred in the citrus orchard, used the citrus sinensis L. Osb cv. Navelina, exquisite oranges, and *Citrus clementina* Hort. ex Tan. Cv. Orornules plants. In 1997, rows of orange and mandarin trees facing east to west were established, each one reaching a height of 7–9 m (476 trees per acre). Plants overall 16 areas ranged in size from 1.5 to 4 m, in straight head thicknesses of 0.8–4.5 m, and LAIs of 1–4.[9]

$$fIPAR = \frac{I_0 - T_c}{I_0} \tag{12.1}$$

$$fAPAR = \frac{(I_0 + R_s) - (T_c + R_c)}{I_0} \tag{12.2}$$

Trees with horizontal crown diameters of 1.4–3.5 m, LAIs of 2–4.2, and heights of 2.2–4 m were discovered in eight separate sections away from the peach orchard. *C. clementina* Hort. ex Tan. Cv. Orornules trees, *Citrus sinensis* L. Osb cv. Navelina, magnificent oranges, and additional species were used in the study in the citrus garden. At heights of 7–9 m (476 trees per acre), mandarin and orange tree arrays pointing the westward direction were first planted in 1997. A total of 16 blocks were selected, each containing trees that ranged in elevation from 1.5 to 4 m, straight crown diameter from 0.8 to 4.5 m, and LAIs of 1–4. The terms “intercepted photosynthetically active radiation (PAR)” and “absorbed PAR” are frequently employed indiscriminately in the field of science. The equation $Rs - Rc$, where Rs is the PAR flux density reflected by the soil and Rc is the PAR flux density reflected by the plant canopy, represents the variance across receiving and redirected PAR.[10] The variation is insignificant for the entire hue greenish canopy (Daughtry et al. 1992). In this study, fIPAR and IPAR were simultaneously used. Multiple airborne missions took place. The amount of sunlight that the orchard canopy in Cambridge, UK, absorbed over the course of the day was measured using a "ceptometer" (SunScan Canopy Analysis System, Delta-T Devices Ltd). The equipment consists of a portable 1-m-long sensor for the flux of PAR transmitted through the canopy and a beam fraction sensor (BFS) for the measurement of PAR incident on the canopy. Two light-emitting parts make up each BFS, one of which the shade ring can shield from the sun.[11] This enables the separation of the direct and diffuse parts of PAR. As one might expect, a tree's ability to intercept PAR is influenced by its surroundings and other nearby trees. The crowns of tree crops that are planted in rows are also close together. The fIPAR field experiments were conducted under each of the four core plants within every study region. SPRAYING UAVs may employ fewer chemicals and remain more efficient than wide- or speed-area in the field of sprayers.

The danger of employee illness and pollution of the environment are correlated with the amount of pesticides used per hectare of land for farming. The employing of UAVs can reduce the need for pesticides. This method allows for the large-scale decontamination of each 0.5 ha, up to 50 ha per day, with just 10 minutes of labor. UAV studies seek to do away with the need for work as a result.

Citrus groves were sprayed by a UAV at various altitudes to determine the greatest level of prophylactic intervention. Researched ways to automatically plan and carry out the best flights using the MSP430. One microchip was installed in the UAV to increase the efficiency of the cleanup process.[12]

12.3.3 ROLE OF UAV IN PRECISION PEST MANAGEMENT

Using remote sensing technology, precision pest management can be utilized to monitor crops and locate pest-affected areas. Control measures, such as pesticide spraying, will then be taken in response to disease prevention. Both technologies should be put on the UAV to do this.

Additionally, fertilizer and pesticides can be sprayed on agricultural areas using the UAV.[9] The UAV has a crucial function that allows it to spray fertilizer and pesticides with good speed and accuracy. The following are the primary UAV spraying components:

- Chemical container
- Hall-flow detector
- Tiny diaphragmatic pumps
- Pressurized nozzles
- spraying the controller
- Small diaphragmatic pump
- Land mapping interpreting software

To spray fertilizers or pesticides, a sprayer is mounted to the UAV. It can be crushed into droplets under pressure and injected via the tip of the nozzle. Using the spray motor, the correct amount of pressure is generated for dispensing the fluid. Utilizing the Hall-flow sensor, the spraying controller turns on the sprayer's nozzle while determining the system's fluid flow. UAVs used for spraying can vary in speed, payload, and the quantity of spraying nozzles they have. The spraying of pesticides and fertilizers using UAVs is more effective than using conventional methods. Human interaction with dangerous gases is decreased. There is a finite quantity of human power needed. The UAV cuts down on both costs and time.[13]

The use of remote sensing technology for pest identification has been thoroughly studied. Pest detection in various types of crops is shown in Tables 12.3–12.6, and the observations are investigated using satellite, ground-based, manned aircraft, and UAV-captured spectral images. The technical specifications of the UAV used to photograph crops in various agricultural areas and locales are further detailed in Table 12.1.

Machine learning (ML) and remote sensing tools play a key role in crop monitoring and pest surveillance. Early warning systems identify crops and pest-affected areas using remote sensing applications that give precise and economic data in different agricultural areas with suitable spectral, temporal, and geographical resolutions. The disease-affected area may be discovered by integrating efficient ML models with cutting-edge mobile applications and high-resolution UAV satellite photos of data; however, monitoring wider areas is difficult.[14]

The hybrid model system was developed using RGB–UAV aerial images from the fields in the Republic of Benin and the Democratic Republic of the Congo. It combines an object detection model (Retina Net) with a bespoke classifier for sickness categorization and banana localization. The results of several tests using performance indicators show that the RGB–UAV mixed model accurately classifies objects and can distinguish between healthy and unhealthy crops with a 99.4% accuracy rate. Therefore, the support systems

TABLE 12.1 Observations of Crops in Various Agricultural Areas.

References	Crop name	Types of UAV	Camera	No. of rotors	Observations
Author1 2021 [17]	Tree1	Multirotor [17]	RGB, multispectral [17]	4	Visual inspection of images [17]
Author 2 [8]	Tree2	Multirotor [8]	Multispectral [8]	6	Visual images [8]
Author 3 021 [9]	Tree3	Multirotor [9]	Multispectral [9]	6	Visual images [9]
Author 4, 2021[20]	Tree4	Multirotor [20]	Hyperspectral [20]	4	Visual inspection of images [20]
Author 5 [21]	Tree5	Multirotor [21]	Hyperspectral [21]	4	Damage assessments [21]
Author 6 [22]	Tree6	Multirotor [22]	Hyperspectral [22]	4	Visual images [22]
Author 7 [23]	Tree7	Multirotor [23]	Hyperspectral [23]	4	Visual inspection of images [23]
Author 8 [24]	Tree8	Multirotor [24]	Multispectral [24]	6	Damage assessments [24]
Author 9 2021	Tree9	Multirotor [21]	Hyperspectral [21]	4	Visual images [21]
Author 10 [25]	Tree10	Multirotor [25]	Hyperspectral [25]	4	Disease monitoring [25]
Author 11 (2010)	Tree12	Multirotor (2010)	Multispectral (2010)	6	Visual images (2010)
Author 12 [28]	Tree13	Multirotor [28]	Hyperspectral [28]	4	Visual images [28]
Author 13 [29]	Tree14	Multirotor [29]	RGB, Nutisectal [29]	4	Visual images [29]
Author 14 [30]	Tree15	Multirotor [30]	Hyperspectral [30]	4	Damage assessments [30]
Author 15 [31]	Tree16	Multirotor [31]	Hyperspectral [31]	4	Visual images [31]
Author 16 [33]	Tree16	Multirotor [33]	Multispectral [33]	6	Visual images [33]
Author 17 [3]	Tree17	Multirotor-om [3]	RGB [3]	4	Vision-based monitoring [3]
Author 18 (5)	Tree18	Multirotor (5)	Multispectral (5)	6	Visual images; degree of severity (5)
Author 19 [36]	Tree19	Multirotor [36]	Hyperspectral [36]	4	Visual inspection of images [36]
Author 20 [37]	Tree20	Multirotor- quadcopter [37]	RGB [37]	4	Pest segmentation and classification [37]
Author 21 [38]	Tree21	Multirotor- quadcopter [38]	RGB [38]	4	Visual Images, [38] degree of severity
Author 22 [39]	Tree22	Multirotor- quadcopter [39]	RGB [39]	4	Visual inspection of images [39]
Author 23 [33]	Tree23	Multirotor [33]	Multispectral [33]	6	Visual images [33]

TABLE 12.2 Ground-based Mapping Techniques.

References	Crop name	Parameters		
		Camera	**Pest name**	**Observations**
Ref. 1 2021 [80]	Crop1	Hyperspectral	*Tetranychus urticae*	visual inspection of the leaves
Ref. 2 2019 [81]	Crop2	Multispectral	Banks grass mite spotted spider mite	Damage assessments
Ref. 3 [82,83]	Crop3	Hyperspectral	Soybean aphid	Arthropod counts
Ref. 4 2018 [43]	Crop4	Multispectral; hyperspectral	Beet Cyst Nematode	Visual images
Ref.5 2018 [84]	Crop5	Multispectral	Two-spotted spider	Controlled infestations
Ref. 6 2017 [85]	Crop6	Hyperspectral	Striped stem borer	Damage assessments
Ref. 7 2017 [86]	Crop7	Hyperspectral	Two-spotted spiders	Damage assessments
Ref. 8 [87–89]	Crop8	Hyperspectral	Sugarcane thrips	Arthropod counts; damage assessments
Ref. 9 2012 [90]	Crop9	Multispectral	Russian wheat aphid	Visual inspections
Ref. 10 2009 [92]	Crop10	Hyperspectral	Spider mite	Arthropod counts; damage assessments
Fraulo et al., 2009 [93]	Crop11	Hyperspectral	Two-spotted spiders	Arthropod counts
Ref. 11 2008 [94]	Crop12	Hyperspectral	Corn leaf aphid	Arthropod counts,
Ref. 112007 [95]	Crop13	Hyperspectral	Leaf miner	Damage assessments
Ref. 12 2002 [65]	Crop14	Multispectral	Soybean cyst Nematode	Visual inspection of images
Ref. 13 1996 [70]	Crop15	Multispectral	Silverleaf whitefly	Visual inspections

made possible by this method hold great promise for spreading dangerous banana illnesses to Africa.

It is crucial to keep track of pests and illnesses to effectively remediate affected regions. When external variables are considered together with the vegetation index, the accuracy level of crops harmed by insects and pests is increased.

In addition, numerous pests, and diseases that harm crop growth may have symptoms that are comparable. Understanding the development phase is therefore helpful in comprehending crop changes brought on by insect and pest infestation. Using Landsat-8 satellite photos (bitemporal), a method is created by combining environmental factors with crop growth, testing with the effects of picture performance classification, and identifying the crops that are afflicted by pests and diseases.[15]

An acceptable reliability of 82.6% was shown by the combined model of environmental factors and chronological growth markers. Furthermore, it performed better when assessing damage in winter wheat harvests using images from the Landsat-8 satellite. To satisfy the current demands of precision farming, it is also necessary to increase the precision of sophisticated algorithms by merging multitemporal remotely sensed data with multisource, which gives a full spatial distribution of crop pests and diseases. The various ground-based mapping techniques are contrasted.

It was possible to accurately measure the cumulative abundance of A. glycine in soybean fields using spectral sensors with an infrared range and a 50-nm sensor bandwidth. Multispectral sensors are used to simulate ground-based hyperspectral data to find A. glycine on soybean. When compared to manual aphid counts and possible pest scouting in soybean and crop production systems, this method is less complicated and expensive.[16]

Since a few decades ago, most agricultural sectors have used RS technology for precision agriculture in a variety of applications, including crop monitoring, yield prediction, and pest management. Additionally, these methods are applied to nutritional deficits and plant stress. In a range of crops and fields, RS technologies can successfully identify pests and insects. Figure 12.10 displays the typical utilization of various RS Platform variants.[17]

Precision accuracy is increasingly crucial for the economic growth of the agricultural sector since it allows for accurate monitoring of crop quality and pest-infested crop monitoring. Additionally, the figure below illustrates the precision accuracy rate achieved using RS technologies in the sphere of agriculture.

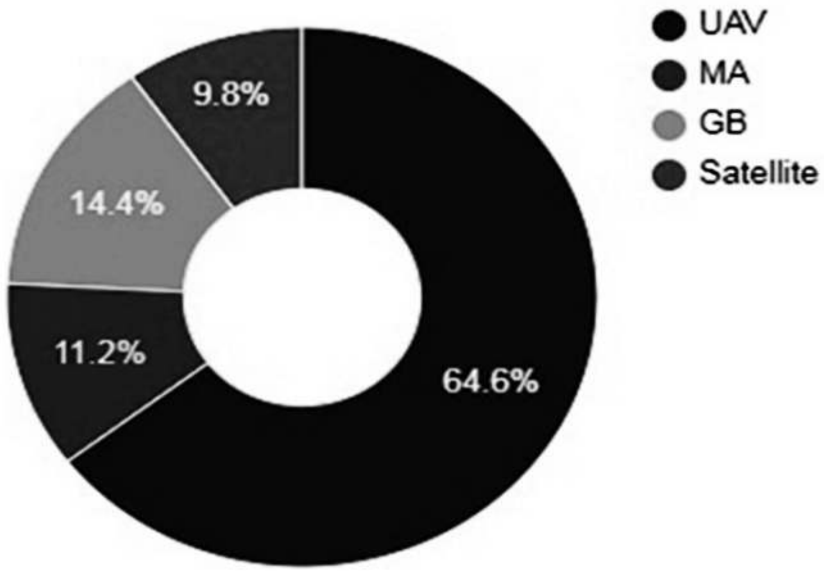


FIGURE 12.5 Average usage of RS platforms.

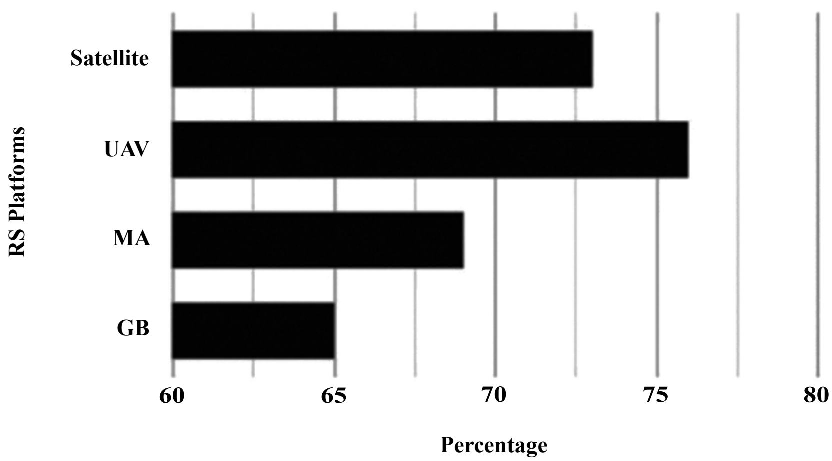


FIGURE 12.6 Precision accuracy rate.

12.3.3.1 THE USE OF DRONES IN INDIAN AGRICULTURE

Drones are currently being promoted in India largely as an automated spray device that may be used to spray pesticides, herbicides, and other chemicals over crops to eliminate the health risks associated with hand spraying and to save time, resources, and human labor.

According to the Indian Agriculture Ministry, each acre will cost between Rs 350 and Rs 450 for the use of a drone with a 10-kg me payload capacity. The figure is predicated on the notion that a drone with numerous batteries will be used for at least six hours every day, covering roughly 30 acres of farmland.[18]

The following provides specific spray operations and analytics related to drones.

These were extracted from an Indian Farmers Fertilizer Cooperative Limited (IFFCO) Kisan information-sharing webinar in which Dr. Shankar Goenka, an expert, discussed the specifics of drone operations based on his company's experiences launching drone sales and offering spray services (on a rent-per-acre basis).

Dr. Goenka claims that drones can evenly apply spray to plants of all heights, including sugarcane, mango orchards, and other taller plants. The input costs of spray products (pesticides/weedicides) are anticipated to drop by roughly 25–30% because of automated operations. Using nano urea as an example, a farmer's input costs are reduced by 25% if savings are realized. Each container used to spray 1 acre costs INR 240. An additional 80–90% reduction in water is also achieved thanks to the relatively small droplet sizes of 50 μm as contrasted to the manual spray droplet size of 500 μm. A person can typically only manually cover 3–4 acres per day, but a drone can spray an acre in around 5–7 min.[19]

A drone with a 10-L capacity has an 11-L tank, among other things. As a result, it may hold 10 L of water and 1 L of spray material, such as a weedicide, pesticide, germicide, or fungicide. A drone alone weighs between 12 and 13 kg. And, thus, a fully loaded 10-gallon capacity drone can fly with up to 25–27 kg of weight. The drone can fly up to 120 feet in the air, but its manufacturers have frozen it at 30 feet to address security issues. It has a flying range of 5 km, but it is frozen at 500 m. In the future, researchers are working to outfit drones with the ability to spray seeds, powdered material, and pellets.[20]

From the standpoint of farmer safety and health issues resulting from spraying and other activities in the field, drones present an appropriate answer. Over 58000 farmer deaths from snakebite have been reported in India. Almost 0.3 million farmers suffer from respiratory ailments because of spraying pesticides and other toxins. Drones have the potential to significantly minimize farmer suffering since they eliminate the need for farmers to penetrate crops deeply for spraying reasons. Regarding drone economics, as noted in the webinar, a typical drone costs between INR 0.7 and 0.9 million. The current cost of renting a drone for spraying is about INR 600–700/acre.

12.4 CONCLUSION

Automated cars are developing quickly and have the potential to completely change how we now travel. However, several operating criteria affect how safely and effectively these vehicles operate. These variables include route design, communication with other vehicles and infrastructure, speed, acceleration, and deceleration, as well as environmental factors. The performance of autonomous cars is significantly impacted by each of these parameters, thus each one must be carefully considered to achieve maximum performance. It is crucial to continuously check and adjust these settings as technology develops to guarantee that automated cars are both effective and safe. By doing this, we can fully use automated driving technology and usher in a new era of transportation. Additionally, research has been done to create precision control algorithms that will increase crop control accuracy.

KEYWORDS

- **automated vehicles**
- **operating parameters**
- **route planning**
- **communication**
- **vehicle-to-vehicle**
- **vehicle-to-infrastructure**
- **safety**

REFERENCES

1. Wang, X.; Zeng, Y.; Yang, Y. Optimization of Operating Parameters for Automated Vehicles. *IEEE Trans. Intell. Transp. Syst.* **2021,** *22* (6), 3841–3852.
2. Bhadani, R.; Gerdes, J. C. Optimal Speed Profiles for Automated Vehicles in Urban Environments. *Transp. Res. Part C: Emerg. Technol.* **2016,** *71,* 249–267.

3. Kim, J.; Choi, B.; Lee, J. Analysis of Acceleration and Deceleration for Improving Energy Efficiency of Autonomous Electric Vehicles. *Energies* **2019,** *12* (7), 1382.
4. Zeng, Y.; Wang, X.; Yang, Y. A Survey on Automated Vehicles from the Perspective of Energy Efficiency. *IEEE Trans. Intell. Transp. Syst.* **2019,** *20* (7), 2539–2553.
5. Ahn, S. H.; Choi, S. H.; Lee, J. W. Development of a Route Optimization Algorithm for Automated Vehicles Considering Traffic Congestion and Energy Efficiency. *Energies* **2021,** *14* (7), 2041.
6. Kusano, K. D.; Gabler, H. C. Use of Vehicle-to-vehicle Communication and Platooning to Address the Intersection Collision Problem. *IEEE Trans. Intell. Transp. Syst.* **2016,** *17* (8), 2224–2233.
7. Maja, M. S.; Soleymani, S. Adaptive Speed Control for Automated Vehicles Based on Local Traffic Density Estimation using Vehicle-to-vehicle Communication. *Transp. Res. Part C: Emerg. Technol.* **2018,** *95,* 388–404.
8. Zhou, X.; Li, K.; Li, L. Autonomous Vehicle Path Planning Under Varying Road Conditions. *Transp. Res. Part C: Emerg. Technol.* **2020,** *114,* 183–200.
9. Shao, F.; Hu, X.; Xie, L. Optimal Energy Management of Automated Vehicles Considering Traffic Conditions and Battery Degradation. *Transp. Res. Part C: Emerg. Technol.* **2018,** *97,* 168–183.
10. Bäckström, C.; Solyom, S. A Review of Automated Vehicle Guidance and Control Systems. *Ann. Rev. Control* **2018,** *46,* 163–176.
11. Chen, Q.; Liu, C.; Wang, Z. Optimization of Operating Parameters for Electric Autonomous Vehicles in Shared Urban Transportation System. *IEEE Trans. Intell. Transp. Syst.* **2020,** *21* (6), 2411–2420.
12. Yang, C.; Shu, C.; Wu, X. Adaptive Fuzzy Control of Autonomous Electric Vehicles for Passenger Comfort Enhancement. *IEEE Trans. Intell. Transp. Syst.* **2018,** *19* (2), 473–483.
13. Fan, X.; Peng, H.; Zhang, J. Energy Optimization for Automated Vehicles at Signalized Intersections Based on Vehicular Communications. *IEEE Trans. Intell. Transp. Syst.* **2020,** *22* (1), 91–103.
14. Leng, B.; Rakha, H. A. An Energy-Efficient Vehicle Routing Problem for Automated Vehicles in the Era of Shared Mobility. *Transp. Res. Part C: Emerg. Technol.* **2019,** *106,* 241–256.
15. Song, W.; Gao, H.; Gao, F. Intelligent Route Planning and Energy Management for Autonomous Electric Vehicles: A Review. *J. Mod. Transp.* **2018,** *26* (4), 301–312.
16. Li, J.; Li, J.; Li, K. A Framework for the Energy-efficient Path Planning of Autonomous Electric Vehicles. *Energies* **2019,** *12* (20), 3901.
17. Xing, Y.; Chen, X.; Li, K. Energy-Efficient Route Planning for Autonomous Electric Vehicles: A Mathematical Programming Approach. *Transp. Res. Part C: Emerg. Technol.* **2019,** *98,* 57–79.
18. Liu, J.; Wang, X. Optimization of Battery Capacity for Autonomous Electric Vehicles Considering Environmental Impact. *Energies* **2021,** *14* (8), 2082.
19. Rios-Torres, J.; Jaller, M.; Barrios, M. A. Optimizing the Operating Speed of Autonomous Vehicles on Highways: A Study of the Impact of Speed on Energy Consumption and Travel Time. *Transp. Res. Part C: Emerg. Technol.* **2018,** *97,* 151–167.
20. Lee, S.; Kim, E. Development of a Simulation Model for Evaluating the Energy Efficiency of Autonomous Vehicles. *Energies* **2019,** *12* (21), 4093.

PART III
DATA ANALYTICS AND APPLICATIONS

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

CHAPTER 13

MACHINE LEARNING-BASED INTRUSION DETECTION TECHNIQUES: A CONCISE STUDY

KSHYAMASAGAR MAHANTA and HIMA BINDU MARINGANTI

Department of Computer Application, Maharaja Sriram Chandra Bhanja Deo University, Odisha, India

ABSTRACT

An intrusion detection system, abbreviated as IDS, examines and detects suspicious activity in a network. Hence, IDS is significant for securing networks as there is a rapid sharing of confidential data over the internet by many organizations and people. In this paper, several machine learning (ML) approaches for intrusion detection (ID) will be reviewed. Various feature selection and classification algorithms have been explored in the context of ML for ID. A comparative analysis is also presented showcasing the various merits and demerits of the ML algorithms considered for the study. Various drawbacks of ML techniques for ID are discussed. ML techniques that are considered in this analysis include Naïve Bayes (NB), K-nearest neighbor (K-NN), Random Forest (RF), support vector machine (SVM), logistic regression (LR), multilayer perceptron (MLP), Classification and Regression Trees (CART), ExtraTree Classifier (ETC), Information Gain Ratio (IGR), principal component analysis (PCA), correlation-based feature selection (CFS), minimum redundancy and maximum relevance feature selection methods, and a few more. All the ML algorithms considered in this study have their advantages and disadvantages. From the analysis, it is found that most of the ML algorithms show acceptable accuracy for ID

Applied Soft Computing Techniques: Theoretical Principles and Practical Applications. Samarjeet Borah, Ratna Raja Kumar Jambi, Sharifah Sakinah Syed Ahmed, & Mahendra Prabhakar Deore (Eds.)
© 2025 Apple Academic Press, Inc. Co-published with CRC Press (Taylor & Francis)

classification. However, there is a scope for improvement in detecting the zero-day attacks. These studies have also provided suggestions for improving the system further. The common and important specificities to be considered are optimized and listed so that this will help the network security expert to design IDSs effectively and efficiently.

13.1 INTRODUCTION

An intrusion detection system (IDS) is a software or hardware-based application that functions to observe a computer system for any potential breaches in security protocols or malicious behavior. A network IDS is significant to secure a network as it permits detecting malicious traffic and responding to it accordingly. Soft computing techniques are always found helpful in resolving various real-world issues such as intrusion detection (ID). These methods include machine learning (ML), genetic algorithms, fuzzy logic, and artificial neural networks (ANNs). Numerous ML techniques are being utilized to develop various types of IDS to secure the network. A multistage cross-layer-based IDS has been introduced in this paper[1] to detect malicious activities. This study proposes a two-stage cross-layer-based IDS. The first stage involves five dedicated sniffers (DS) that gather data from the network and MAC layers. This data is then inputted into individual DS-mounted Random Forest (RF) classifiers, which generate correctly classified instances (CCIs). The CCIs are then transmitted to a supernode (SN), which is the second stage of the IDS. In the SN, a sliding window algorithm (SWA) is applied to each CCI obtained from the different DSs. This method calculates an accumulated measure of fluctuation (AMoF), and an iterative linear regression (ILR) method is applied to the AMoF. A crucial step in the identification process involves the selection of a detection threshold, which serves to differentiate between benign nodes and those that are malicious. This method has been tested using two varied mobility models, namely, Gaussian Markov models (GM) and random waypoint (RWP). Detection rates explored 98% for high-node velocity cases. Whereas they go down to 90% for low-node velocity cases. However, the drawbacks to this IDS include a false positive rate (FPR). In addition, the detection process has also been affected in the initial steps of the fitted-slope computation.[2] The probable solution can be based on filtering these initial steps. An alternative way is to use a highly complicated method based on adaptive feature selection at individual reporting time to enhance the performance, thereby affording effective differentiation amongst malicious and benign nodes during the initial step of the fixed slope process. A general study with a detection framework can be found in Ref [3].

In accordance with this, a framework comprising all the needed steps for reliable and accurate network IDS validation has been introduced. Last, the introduced framework has been tested by a comparison of semisupervised and unsupervised methods against supervised methods by the use of the recently developed dataset UGR'16.[4] The outcomes explored that further work has to be done to introduce standard and common frameworks to evaluate network IDS solutions based on ML through comparison. Various IDs have been classified through numerous ML algorithms.[5] It enhances the computer's learning process on the basis of experiences without being programmed. The study performed a comparison of several ML algorithms such as J48, principal component analysis (PCA), support vector machine (SVM), Decision Tree (DT), LR, modified K-means, and ANN for IDS. For testing, RF, linear discriminant analysis (LDA), and Classification and Regression Trees (CART) have been used. Outcomes revealed the outstanding performance of RF with an accuracy of 99.65% in comparison to other methods. Yet, the study needs an overall review of ML algorithms to afford an efficient solution for IDS by considering a real-time dataset.[6]

This study reviews various traditional ML algorithms for ID. This study also discusses the efficacy of numerous ML algorithms and their performance in ID. Future recommendations are also presented. Figure 13.1 shows the block diagram of IDS.

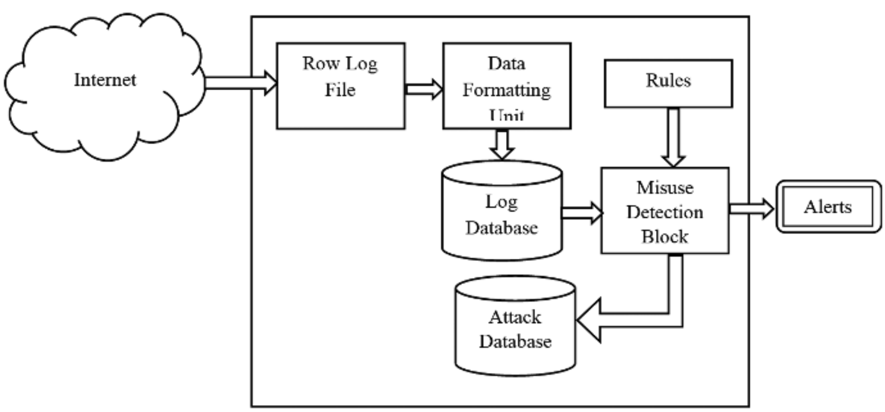


FIGURE 13.1 Block diagram of an IDS.

Figure 13.1 consists of various blocks that are briefly discussed. Data arriving from the internet is collected, gathered, and redirected to the Log File (LF). Then, the data gathered in the LF is classified by data formatting unit (DFU) per numerous packet header fields. This formatted data is stored in the Log Database (LB). Subsequently, the misuse detection block detects the known attacks. Various computer attacks possess fixed signatures that can be utilized to detect the specific attack. This block uses predefined rules for comparing the obtained data packet header (DPH). When the pattern matches, the IDS detects it as an intrusion and sends an alert signal to the administrator. This discourse concerns the broad procedure of the IDS whereby the identified LB entries, which are indicative of attacks, are subsequently conserved in the attack database.

13.1.1 OBJECTIVES

This study offers significant contributions which are outlined as follows:

- To undertake a comprehensive examination of diverse ML methodologies that can be employed in the detection of intrusions.
- To comparatively analyze the merits and demerits of various ML techniques for detecting intrusion.
- To recommend future directions to enhance the study in this domain.

13.1.2 CHAPTER ORGANIZATION

This study is structured in a subsequent manner. Section 13.1 explores the fundamental ideas about IDS and ML. Following this, Section 13.2 discusses the review of traditional research associated with the use of various ML techniques that are used for ID. Subsequently, Section 13.3 comparatively analyzes the merits and demerits of all the ML algorithms considered in this study for ID. The overall idea of the review is summarized in Section 13.4.

13.2 ID USING ML TECHNIQUES

IDS is an effective tool for detecting adverse attempts to gain unauthorized access to a network or system, mainly via the Internet. It detects various kinds of malicious activities by monitoring incoming as well as outgoing traffic.[7]

These malicious activities have to be detected as they jeopardize system security. The primary objective of IDS is to detect any unauthorized access or misuse within the host network. Thus, a security model has been presented based on ML.[8] It is termed Intrusion Detection Tree (IntruDTree). The security features have been ranked in accordance with the significance of modeling. Subsequently, a generalized ID model that is tree-based has been constructed on the basis of selected significant features. Following this, the test data has been utilized for model validation. This technique is efficient with respect to accuracy and prediction by reducing the overfitting and computational complexity. The efficiency of the proposed model has to be assessed by gathering huge datasets with security features of high dimension in IoT security, thereby measuring its efficacy at the application level in cybersecurity so that the performance can be further improved. Similarly, this study[9] introduced a new framework named multistage optimized network IDS on the basis of ML that minimizes the computational complexity. Concurrently, this technique sustains its performance in detection. Initially, the effect of oversampling methodologies on the training sample size of the model is reviewed. Then, the proper training size to detect intrusion efficiently has been determined. Various hyper-parameter optimization methodologies have been examined to improve the performance of network IDS and confirm its robustness and effectiveness. The proposed methodology has been evaluated using two recent traditional ID datasets named UNSW-NB 2015 and the CICIDS 2017 dataset. Several evaluation metrics such as precision, accuracy, false alarm rate (FAR), and recall have been utilized to evaluate the performance of the proposed system. The results explored that information gain-based feature selection (IGBFS) accomplished effective detection accuracy in comparison to the traditional method. This study presented some recommendations to be implemented in the future. Deep learning (DL) classifiers can be utilized as they perform effectively on high-dimensional and nonlinear datasets. The influence of integrating unsupervised and supervised ML techniques has to be examined and this will confirm its significance to detect new attacks. Likewise, a symmetrical combination has been proposed that comprises ML techniques and a knowledge-based strategy to construct a hierarchical IDS to assist in detecting the severity of zero-day attacks and existing kinds.[10] The results explored that the proposed technique outperformed other traditional techniques with respect to recall, FAR, and precision. This approach has to be extended to achieve dynamicity with respect to various new network attacks and make correct decisions for retraining the accessible prognostication models.[12]

Employing new methodologies is significant to enhance the performance of IDS in recent data networks due to the increase in cyber-attacks. An interactive three-dimensional strategy has been introduced[12] to analyze the network IDS datasets in addition to ML outcomes. The study intended to afford a visual data representation from the network IDS dataset in a way that exposes a geometric association of the data records among varied network traffic. Various ways for ML misclassification have also been examined from a visual viewpoint which explored the efficiency of the proposed approach to identify the clusters and patterns in data. Though the proposed system showed valuable results, a few areas are to be considered in the future. Other ML techniques and dimensionality reduction must be focussed. In addition, network traffic has to be visualized in real time to permit users to analyze the incoming network traffic. In addition, a focus has been given to enhance the IDS interpretability.[13] The proposed framework generated interpretation outcomes which were found to be reliable with the features of particular attacks. The outcomes have also been found to be intuitive. Moreover, the proposed system has various demerits that are to be rectified to enhance its efficacy. The proposed technique has to be employed on many datasets. The practical application of the proposed solution in real time has not yet been executed. This study has to explore various sophisticated attacks. Correspondingly, an ML technique for detecting intrusion information has been proposed.[14] Elman neural network (ENN) merits and noise removal of SVM are completely utilized. Later, it combined both to find a solution for the safety risks of information system ID to assure the information system security. In addition, various challenges are faced by IDS technology. The ID speed has to be enhanced to satisfy the network communication requirements. The false positives (FP) and false negatives (FN) have to be minimized in the IDS to enhance accuracy and safety. The IDS's interactive performance has to be enhanced to assure system safety. On the other hand, various deep reinforcement learning (DRL) algorithms have been explored for ID by utilizing a labeled dataset.[15] This study also explored the various ways to execute supervised learning on the basis of the proposed framework. The proposed system has been analyzed to validate its efficiency. From the results, it has been found that the classifier attained from the proposed framework is faster than traditional models. Thus, the proposed method outperformed other techniques that proved the outstanding performance of the introduced methodology. The application of the new DRL algorithms has to be examined, particularly in the context of multiagent as well as adversarial models. This can be applied to the ID issues.[16] This analysis is yet to be done. Hence it is a drawback.

13.2.1 SIGNIFICANCE OF ML FOR ID

Securing the network is important due to rapid development in the usage of cyber networks and various applications being run on it. Hence, IDS is becoming a vital tool to afford high security in the network environment. There are various IDS challenges such as false detection, misinterpretation, deficiency in real-time response, and so on. At present, the computing power of machines led to the development of various technologies such as ML. The ML classifiers have the ability to boost the system's robustness and accuracy. Due to these merits, they are significant and are being employed to report several attacks for enhancing security. Various ML algorithms have been comparatively analyzed.[17] These algorithms comprise naïve Bayes (NB), RF, SVM, and K-nearest neighbor (K-NN). The comparative analysis revealed that RF had been found to be efficient than all other algorithms that are considered for comparison. This RF effectively determined the classification and prediction of varied kinds of data packets as if it is an encrypted or normal packet. In addition, it determined the encrypted malicious or normal malicious packet. This study concluded that DL algorithms can be applied in the near future to enhance the accuracy and performance of classification and recognition of varied packet types that are transferred over a network.[18] Various existing studies used several methods for ID. Accordingly, an anomaly-based IDS system has been proposed that integrated (hybrid) two supervised ML algorithms, namely CART and RF.[19] Here, RF has been utilized for feature selection, whereas CART has been utilized for classification by finding various kinds of attacks. Results showed the efficacy of the introduced method than traditional methods. Yet, oversampling is to be performed to enhance the training instances of backdoor, analysis, shellcode, DOS, as well as worm attacks. Accomplishing oversampling will have various positive impacts on the performance of the model. The dataset utilized in this study will be balanced. This will increase the proposed model's performance. Furthermore, various DL and ML algorithms have been executed as IDS in the study.[20] A dataset named Coburg Intrusion Detection Data Sets (CIDDS)[21] has been utilized as a standard dataset. Under-sampling has been performed to balance the datasets in a way that makes every class possess a similar count of data. The results explored the effective performance of the proposed model in independent testing and tenfold cross-validation. Moreover, convolutional neural network (CNN) and CART with embedding accomplished slightly effective performance on the utilized dataset in comparison to the traditional models. The outcomes exhibit the efficacy of both techniques in ID.

As ID is significant, the study[22] introduced a feed-forward neural network (FFNN). The method proposed has been subject to analysis through comparison to previous works. The outcomes revealed the effective performance of the proposed technique with an accuracy of 99% and FPR less than 0.5. ID on the CICID 2017 dataset might be further enhanced. Additionally, it should be noted that Supervised Learning possesses the capability to solely identify intrusions that exist within the given dataset. Whereas, unsupervised learning can detect new attacks. So, there is a need to adopt unsupervised learning techniques. Additionally, an IDS that utilized PCA along with RF has been explored.[23] Both of these techniques worked for a specific purpose. PCA afforded granularity in data. The RF assisted in the classification of the nodes to predict the attacks. The attained outcomes state the effectiveness of the proposed methodology with respect to accuracy in comparison to other methodologies, namely NB, DT, and SVM. Similarly, various ML classifiers have been presented to detect their performance in ID. The used ML classifiers include K-NN, multilayer perceptron (MLP), LR, NB, SVM, RF, DT, and ExtraTree Classifier (ETC). The performance has been validated on the NSL-KDD dataset.[24] Initially, the dataset has been pre-processed. The model has been trained and tested on the basis of significant attributes. The experimental outcomes explored that every classifier produces promising outcomes for the DoS attack. The study has to further focus on a viewpoint of employing optimization methodologies to offer efficient IDS. On the other hand, the paper[25] employed a DL strategy to classify the network traffic data in IDS. In the early study, long short-term memory (LSTM) has been utilized for data classification. Followed by this, LSTM and CNN layers have been integrated. NSL-KDD dataset has been used. The proposed CNN-LSTM has been compared with the traditional methods. The results show the efficient performance of the proposed method with respect to minimum FPR and maximum accuracy than the traditional methods for detecting intrusion.[26]

13.2.2 ML TECHNIQUES FOR ID: FEATURE SELECTION AND CLASSIFICATION

Feature selection utilized in IDS assists in minimizing the classification time. Thus, efficient classification of ID can be accomplished. An IDS has been designed for the network through the use of feature selection and classification techniques. This is done by reviewing the combinations of major popular feature selection methodologies as well as classifiers.[27] The study afforded a strategy regarding the exact feature selection methodologies to be integrated with a classifier to accomplish accurate network ID. Various feature selection methodologies were focussed such as correlation-based feature selection (CFS), principal component analysis (PCA), information gain ratio (IGR), as well as minimum redundancy and maximum relevance feature selection methods. Moreover, the classifiers that have been considered include NB, SVM, NN, K-NN, and DT. The results showed that maximum accuracy is accomplished by the combination of IGR with K-NN. From the results, it has been concluded that the IGR and K-NN can together be utilized to design an efficient IDS. Various ML algorithms have been analyzed on IDS for using cybersecurity data.[28] Several popular ML classification methodologies have been employed to resolve it. The used ML classification algorithms comprised NB, DT, RF, Bayesian Network (BN), Random Tree (RT), ANN, and Decision Table (DT) to afford intelligent services in cybersecurity, specifically for ID. Lastly, the efficiency is tested by undertaking several experiments on cybersecurity datasets. It comprises various cyber-attack categories. The efficiency is evaluated by computing various performance metrics such as recall, precision, accuracy, and F1-score for these ML-based IDS models. It has been found that the RF-based IDS model showed consistent performance in comparison to other classifiers for ID. The cybersecurity dataset has to be extended in the near future. The study also plans to adopt a data-driven IDS for automated security services. Additionally, a supervised ML model has been developed.[29] It has the ability to classify the hidden network traffic on the basis of the learning achieved from the viewed traffic. ANN and SVM have been utilized to determine the efficient classifier with maximum success rate and accuracy. The examination of the outcomes exhibits that the model constructed using wrapper feature selection and ANN performed better than the traditional models. The detection rate has been found to be 94.02%. Detecting a zero-day or new attack remains a challenge. This happens due to the high FPR of the conventional systems. Similarly, software software-defined network (SDN) has been described as a platform to execute network IDS with DL or ML strategies outstretches of traditional review works.[30]

MLs have been utilized for misuse, hybrid, or anomaly detection strategy with a clear analysis of their ability to detect attacks.[31] Multiple classifiers have been taken into account. It consists of various ML algorithms within it while detecting intrusions and learning. A classifier set has been integrated to afford a common result for ID. A stacked ensemble learning model for network ID has been introduced by utilizing RF and gradient boosting machine (GBM) algorithms.[32] The introduced technique combined the features of RF and GBM classifiers. It also afforded efficient performance in comparison to other ensemble techniques such as boosting and bagging. Numerous ML techniques for ID have used stacked ensemble learning classifiers. However further work needs to be done by the analysis of probabilistic, DT, rule induction-based and nonprobabilistic classification algorithms and these have to be combined as an ensemble to accomplish better performance.[33]Numerous ML algorithms (DT, Bagging Tree (BT), XGBoost, RF, Bayes Net (BN), NB, SVM, EM, AdaBoost, K-means, and DBSCAN) have been reviewed by training as well as testing[34] on the dataset with respect to Area under Curve, Matthews correlation coefficient (MCC), and accuracy. The analytical outcomes reveal that the tree-based techniques accomplished high accuracy of 97% and XGBoost ranked the topmost by performing better than other methods and expectation maximization (EM) accomplished high accuracy. Yet, the study has to be extended by considering unsupervised techniques.[35]

13.2.3 EFFICIENCY OF VARIOUS ML TECHNIQUES FOR ID

Various traditional systems have used several ML techniques for ID. To enhance the IoT security, the study[36] have explored various intelligent methodologies. These methods have been applied for securing the computer networks (CN) and specifically in ID to accomplish effective recognition rates in ID. Yet, FPR seems to be a problem in many studies. This study explored the FPR issues in various other papers. Though few of the methodologies minimize FPR, they fail in terms of classification and training time. On the contrary, few methods perform inversely. This means the FPR is stabilized. Whereas the training cost and the testing cost gets increased. Thus, real-time implementation for ID is essential. All these issues have to be rectified to minimize the workload of security analysts. Thus, the article[37] proposed an Active Learning IDS-ML algorithm that employed active learning for ID. The algorithm has been introduced to adapt to certain threats against ML algorithms. It has been concluded that selective unlabeled data sampling for a human expert to perform classification can lead to effective labeling on huge datasets. This revealed a resilient strategy for ML. The methodology resolved issues associated with data evasion and tampering. Various ML algorithms have also been considered to evaluate its efficiency in ID.[38] These methods include RF, Decision Table, NB, BN, J48, and Random Tree (RT). The study used KDD databases that comprised certain attacks such as R2L, DOS, PROBE, and U2R. The results explored that all the selected ML classifiers had the ability to construct their training models within sufficient time, whereas MLP has been found to be the only technique that failed to accomplish it. Moreover, the average accuracy rate and True Positive (TP) were insufficient for ID. FPR and FN must also be taken into account. Additionally, a complexity and statistical analysis of the CIDDS-001 dataset has been explored.[39] Unsupervised and supervised ML techniques have also been used to examine the dataset complexity with respect to renowned performance metrics. The outcomes explored that DT-, RF-, K-NN-, DL-based classifiers and NBs have been utilized to develop an effective network IDS. The study concluded the suitability of the utilized dataset for assessing the anomaly-based network IDS. Moreover, this dataset has to be compared with the traditional benchmarking datasets in the near future. Various classification algorithms have been evaluated to find an efficient classifier capable of showing high classification accuracy with minimum execution time.[40] A UNSW-NB15 dataset has been applied to validate the efficacy of varied detection algorithms by the use of Apache Spark. The detection algorithm has to classify the incoming network traffic as an attack or normal by considering all the features of every network traffic pattern. The analytical outcome explored that the RF classifier showed efficient performance with respect to sensitivity, accuracy, execution time, and specificity. Various feature selection methods have to be considered to minimize the feature counts utilized for detection.[41] Likewise, an adaptive ensemble learning model has been introduced that has the ability to incorporate the merits of all the algorithms for varied data detection kinds, thereby accomplishing optimal outcomes via ensemble learning.[42] The ensemble learning has a merit that integrates the predictions of various base estimators for enhancing robustness and generalizability over an estimator. NSL-KDD dataset has been utilized. Few typical algorithms such as RF, deep neural network (DNN) and DT have been utilized for training the model. Adaptive voting, as well as multitree algorithm, have also been introduced to enhance the impact of ID. The analytical outcomes revealed that the Multi-tree algorithm showed 84.2% accuracy and adaptive voting system accomplished 85.2% accuracy.[43]

13.3 COMPARISON OF DIFFERENT ML METHODS FOR INTRUSION DETECTION

Table 13.1 shows various ML techniques, utilized by the traditional systems for ID to enhance system security.

TABLE 13.1 Various ML Techniques for IDS—A Comparative Analysis.

Sl. no.	Author	ML technique	Merits/demerits
1.	[7]	A security model has been proposed on the basis of machine learning (ML) named Intrusion Detection Tree—IntruDTree.	The proposed technique showed efficient outcomes with respect to accuracy and prediction by reducing the overfitting and computational complexity. The methodology has to be applied to a huge dataset to find its extreme efficacy, which is a drawback.
2.	[8]	The study introduced a new framework named multistage optimized network IDS on the basis of ML.	The introduced framework minimizes the computational complexity and sustains its performance in detection. The study has not examined the influence of integrating unsupervised and supervised ML techniques as this is needed to confirm its significance for detecting new attacks.
3.	[9]	A symmetrical combination has been proposed that comprises ML techniques and knowledge-based strategy.	The presented technique outperformed other traditional techniques with respect to recall, FAR, and precision. This approach requires a further extension for accomplishing dynamicity with respect to various new network attacks and making necessary decisions for retraining the accessible prognostication models.
4.	[11]	An interactive three-dimensional strategy has been introduced to analyze the network IDS datasets.	The discussed approach is efficient in identifying the clusters and patterns in data. The study lacks real-time implementation.
5.	[14]	Various deep reinforcement learning (DRL) algorithms have been presented to ID by utilizing a labeled dataset.	The classifier attained from the proposed framework is faster than traditional models. The study has not examined the application of new DRL algorithms, particularly in the context of multi-agent as well as adversarial models, as it is more applicable to the ID issues.
6.	[16]	NB, RF, K-NN and SVM	RF effectively determined the classification and prediction of varied kinds of data packets as if it is an encrypted or normal packet. In addition, it determined the encrypted malicious or normal malicious packet.

TABLE 13.1 *(Continued)*

Sl. no.	Author	ML technique	Merits/demerits
7.	[18]	Classification and Regression Trees (CART)-Random Forest (Hybrid)	Oversampling is yet to be performed to enhance the training instances of Backdoor, analysis, shellcode, DOS as well as worm attacks.
8.	[21]	The study introduced a feed-forward Neural Network (FFNN).	Unsupervised learning techniques must be used in the study since they are essential for finding zero-day attacks.

From Table 13.1, various ML techniques have been used by traditional systems such as IntruDTree, multistage optimized network IDS, knowledge-based strategy, interactive three-dimensional strategy, DRL algorithms, NB, RF, K-NN, SVM, CART-RF (hybrid), and FFNN. All these techniques have their own advantages and disadvantages. This study encountered the most common demerits. They are as follows. Several studies lack real-time implementation. The techniques have not been applied to large datasets. They have also not considered unsupervised ML as it is significant to detect new attacks. Thus, all these are needed to effectively detect intrusion and enhance system security.

13.4 RESULTS AND DISCUSSION

Here, a thorough summary of the notable studies' primary contributions is provided. The study proposed a two-stage cross-layer-based IDS. This study has selected a detection threshold for partitioning the malicious and normal nodes. The accuracy of this implementation afforded 98%. However, it has been affected by FPR as well as detection process in the initial stages of computing fitted slope.[1] Thus, the study provided a suggestion to overcome this by adaptive feature selection process[3] at individual reporting time so as to enhance the performance as well as afford better differentiation amongst malicious and benign nodes at the initial fitted slope stages. Additionally, examining the influence of integrating unsupervised and supervised ML methods might confirm paramount significance in this area for novel attack detection. Hence, this paper[41] considered NSL-KDD dataset and examined the existing issues in the ID area and introduced an adaptive ensemble learning model. It adjusted the training data and constructed multiple DTs (Decision Trees) for building the Multi-Tree algorithm. To enhance the overall detection impact, the paper selected various base classifiers that incorporated RF, DT, DNN, and KNN. The validation of the system was carried out using the NSL-KDD dataset. It revealed the accuracy of the proposed Multi-Tree as 84.2% and adaptive model showed 85.2% as accuracy. The ensemble model proposed in this study has been found to be more efficient than the traditional systems in terms of enhancing the accuracy of detection. Data analysis has also been carried out that exhibited the significance of data feature quality to find the effect in detection. Furthermore, the study have to enhance the preprocessing and feature selection of ID to accomplish efficient outcomes.[26]

Most of the existing studies intended to minimize the computational complexity. Accordingly, the article[7] exhibited a security model based on ML named IntruDTree. The importance in modeling have been taken into account to rank the security features. The outcomes explored that the proposed security model minimized the computational complexity. This study provided ways for further studies that could evaluate the efficiency of the proposed model by gathering huge datasets possessing high security feature dimensions in IoT and evaluating its efficacy at the application level in cybersecurity. Similarly, a multi-stage optimized network IDS has been employed to mainly minimize the computational complexity thereby detecting attacks.[8] This paper researches the influence of oversampling methods on the sample size of the training model and finds the appropriate sample size for training. Besides, varied ML hyper-parameter optimization methods have been examined to improve the performance of network IDSs. The analytical outcomes revealed that the proposed system minimized the computational complexity with 99% as accuracy for attack detection.

Existing research[4] utilized ML algorithms namely CART, RF, and LDA for detecting intrusions. This study has been tested through the use of KDD-CUP dataset. The implementation of the introduced models for ID explored that RF showed efficient accuracy of 99%. However, the study have to comprehensively review ML algorithms for providing effective solution by taking into account the real-time dataset. On the other hand, the cybersecurity experts have to optimize the decisions in accordance with the model assessment.[12] To resolve this, the study introduced SHapley Additive exPlanations (SHAP) and integrates global as well as local explanations for enhancing the IDSs interpretation. This proposed framework can assist the cybersecurity staffs possess a clear understanding about IDSs. The study suggested that using a real-time dataset will help to find the efficiency of the proposed algorithm in a better way.

Traditional paper used MLP on modern dataset to detect attacks.[21] CICIDS2017 comprise of intrusion attacks as well as traffics which indicate the present network usage. These steps have been detailed from data analysis to neural network (NN) tuning. Subsequently, existing studies on this dataset have been examined and their performance have been compared with the empirical outcomes. The results explored the efficacy of the proposed system in detecting attacks than the NN used by the previous studies.[5]

From the analysis, it is found that most of the studies provided common suggestions to be considered in future to increase the system efficiency. The present study found that these suggestions have been explored by many studies through analysing the existing systems. It is suggested to use real-time dataset, unsupervised ML algorithms, considering huge datasets and taking various features into account. Thus, this study have explored the significant suggestions to be considered in future for accurate ID.

13.5 CONCLUSION

In this study, various ML algorithms used by the traditional systems for ID have been analyzed. Various feature selection and classification algorithms for ID have been specifically reviewed that include NB, RF, K-NN, SVM, LR, MLP, CART, ETC, IGR, and CFS. It is found that all the ML algorithms possess their own merits and demerits. The existing methods were able to detect intrusions with a better accuracy. These studies have also provided suggestions for further improvement. Most of the studies suggested common specificities. Due to the rapid development in technology, new attacks are emerging and an effective design of ID is needed to detect zero-day attacks. Hence, it is significant to consider all the suggestions provided by the existing study for accurate ID. These studies mainly explored to consider unsupervised ML algorithms, high dimensional security features and implementation of ML algorithms on large dataset. Feature selection is the significant process needed to enhance the accuracy of the classification process that directly impacts the ID. On the other hand, real-time implementation helps to increase the system accuracy and considering huge datasets will make the system more effective. In addition, unsupervised ML is also significant to detect zero-day attacks. Hence, this study explores the common and significant measures to be considered for increasing the efficiency of the system in ID and concludes that all these suggestions will assist a network security expert to increase the system efficiency and select the suitable structure during the design of IDS.

KEYWORDS

- **secure network**
- **security**
- **machine learning techniques**
- **zero-day attacks**
- **intrusion detection**
- **redundancy**
- **relevance**

REFERENCES

1. Amouri, A.; Alaparthy, V. T.; Morgera, S. D. A Machine Learning Based Intrusion Detection System for Mobile Internet of Things. *Sensors* (*Switzerland*) **2020,** *20* (2). https://doi.org/10.3390/s20020461.
2. Khraisat, A.; Gondal, I.; Vamplew, P.; Kamruzzaman, J. Survey of Intrusion Detection Systems: Techniques, Datasets and Challenges. *Cybersecurity* **2019,** *2* (1). https://doi.org/10.1186/s42400-019-0038-7.
3. Panigrahi, R.; Borah, S.; Bhoi, A. K.; Mallick, P. K. In *Intrusion Detection Systems (IDS)—An Overview with a Generalized Framework*,Cognitive Informatics and soft Computing, pp. 107–117. Springer, Singapore 2020, https://doi.org/10.1007/978-981-15-1451-7_11
4. Magán-Carrión, R.; Urda, D.; Díaz-Cano, I.; Dorronsoro, B. Towards a Reliable Comparison and Evaluation of Network Intrusion Detection Systems Based on Machine Learning Approaches. *Appl. Sci.* **2020,** *10* (5), 1–21. https://doi.org/10.3390/app10051775.
5. Saranya, T.; Sridevi, S.; Deisy, C.; Chung, T. D.; Khan, M. K. A. A. Performance Analysis of Machine Learning Algorithms in Intrusion Detection System: A Review. *Procedia Comput. Sci.* **2020,** *171* (2019), 1251–1260. https://doi.org/10.1016/j.procs.2020.04.133.
6. Hindy, H.; Bayne, E.; Bures, M.; Atkinson, R.; Tachtatzis, C.; Bellekens, X. Machine Learning Based IoT Intrusion Detection System: An MQTT Case Study (MQTT-IoT-IDS2020 Dataset). *Lect. Notes Netw. Syst.* **2021,** *180,* 73–84. https://doi.org/10.1007/978-3-030-64758-2_6.
7. Ambhore, P. B. Intrusion Detection System for Intranet Security. **2014,** *4* (7), 626–631.
8. Sarker, I. H.; Abushark, Y. B.; Alsolami, F.; Khan, A. I. IntruDTree: A Machine Learning Based Cyber Security Intrusion Detection Model. *Symmetry (Basel)*. **2020,** *12* (5), 1–15. https://doi.org/10.3390/SYM12050754.
9. Injadat, M.; Moubayed, A.; Nassif, A. B.; Shami, A. Multi-Stage Optimized Machine Learning Framework for Network Intrusion Detection. *IEEE Trans. Netw. Serv. Manag.* **2021,** *18* (2), 1803–1816. https://doi.org/10.1109/TNSM.2020.3014929.
10. Sarnovsky, M.; Paralic, J. Hierarchical Intrusion Detection Using Machine Learning and Knowledge Model. *Symmetry (Basel)*. **2020,** *12* (2), 1–14. https://doi.org/10.3390/sym12020203.
11. Sahu, A.; Mao, Z.; Davis, K.; Goulart, A. E. In *Data Processing and Model Selection for Machine Learning-Based Network Intrusion Detection*, 2020 IEEE Int. Work. Tech. Comm. Commun. Qual. Reliab. CQR 2020; **2020**. https://doi.org/10.1109/CQR47547.2020.9101394.
12. Zong, W.; Chow, Y. W.; Susilo, W. Interactive Three-Dimensional Visualization of Network Intrusion Detection Data for Machine Learning. *Futur. Gener. Comput. Syst.* **2020,** *102,* 292–306. https://doi.org/10.1016/j.future.2019.07.045.
13. Wang, M.; Zheng, K.; Yang, Y.; Wang, X. An Explainable Machine Learning Framework for Intrusion Detection Systems. *IEEE Access* **2020,** *8,* 73127–73141. https://doi.org/10.1109/ACCESS.2020.2988359.
14. Fang, W.; Tan, X.; Wilbur, D. Application of Intrusion Detection Technology in Network Safety Based on Machine Learning. *Saf. Sci.* **2020,** *124* (January), 104604. https://doi.org/10.1016/j.ssci.2020.104604.
15. Lopez-Martin, M.; Carro, B.; Sanchez-Esguevillas, A. Application of Deep Reinforcement Learning to Intrusion Detection for Supervised Problems. *Expert Syst. Appl.* **2020,** *141,* 112963. https://doi.org/10.1016/j.eswa.2019.112963.

16. Andresini, G.; Appice, A.; Mauro, N. Di; Loglisci, C.; Malerba, D. Multi-Channel Deep Feature Learning for Intrusion Detection. *IEEE Access* **2020,** *8,* 53346–53359. https://doi.org/10.1109/ACCESS.2020.2980937.
17. Sumaiya Thaseen, I.; Poorva, B.; Ushasree, P. S. Network Intrusion Detection Using Machine Learning Techniques. *Int. Conf. Emerg. Trends Inf. Technol. Eng. ic-ETITE 2020* **2020,** 1–7. https://doi.org/10.1109/ic-ETITE47903.2020.148.
18. Meryem, A.; Ouahidi, B. EL. Hybrid Intrusion Detection System Using Machine Learning. *Netw. Secur.* **2020,** *2020* (5), 8–19. https://doi.org/10.1016/S1353-4858(20)30056-8.
19. Chkirbene, Z.; Eltanbouly, S.; Bashendy, M.; Alnaimi, N.; Erbad, A. In *Hybrid Machine Learning for Network Anomaly Intrusion Detection*, 2020 IEEE Int. Conf. Informatics, IoT, Enabling Technol. ICIoT 2020; **2020,** pp 163–170. https://doi.org/10.1109/ICIoT48696.2020.9089575.
20. Thapa, N.; Liu, Z.; Kc, D. B.; Gokaraju, B.; Roy, K. Comparison of Machine Learning and Deep Learning Models for Network Intrusion Detection Systems. *Futur. Internet* **2020,** *12* (10), 1–16. https://doi.org/10.3390/fi12100167.
21. Abdulhammed, R.; Faezipour, M.; Abuzneid, A.; Abumallouh, A. Deep and Machine Learning Approaches for Anomaly-Based Intrusion Detection of Imbalanced Network Traffic. *IEEE Sensors Lett.* **2019,** *3* (1), 1. https://doi.org/10.1109/LSENS.2018.2879990.
22. Rosay, A.; Carlier, F.; Leroux, P. In *Feed-Forward Neural Network for Network Intrusion Detection*, IEEE Veh. Technol. Conf. May, *2020*; **2020**. https://doi.org/10.1109/VTC2020-Spring48590.2020.9129472.
23. Waskle, S.; Parashar, L.; Singh, U. In *Intrusion Detection System Using PCA with Random Forest Approach*, Proc. Int. Conf. Electron. Sustain. Commun. Syst. ICESC 2020 **2020,** No. Icesc, 803–808. https://doi.org/10.1109/ICESC48915.2020.9155656.
24. Abrar, I.; Ayub, Z.; Masoodi, F.; Bamhdi, A. M. In *A Machine Learning Approach for Intrusion Detection System on NSL-KDD Dataset*, International Conference on Smart Electronics and Communication (ICOSEC), 2020, pp. 919–924, https://doi.org/10.1109/ICOSEC49089.2020.9215232.
25. Hsu, C. M.; Azhari, M. Z.; Hsieh, H. Y.; Prakosa, S. W.; Leu, J. S. Robust Network Intrusion Detection Scheme Using Long-Short Term Memory Based Convolutional Neural Networks. *Mob. Networks Appl.* **2021,** *26* (3), 1137–1144. https://doi.org/10.1007/s11036-020-01623-2.
26. Çavuşoğlu, Ü. A New Hybrid Approach for Intrusion Detection Using Machine Learning Methods. *Appl. Intell.* **2019,** *49* (7), 2735–2761. https://doi.org/10.1007/s10489-018-01408-x.
27. Biswas, Saroj. Intrusion Detection Using Machine Learning: A Comparison Study. *Int. J. Pure Appl. Math*. **2018,** *118,* 101–114.
28. Alqahtani, H.; Sarker, I. H.; Kalim, A.; Minhaz Hossain, S. M.; Ikhlaq, S.; Hossain, S. *Cyber Intrusion Detection Using Machine Learning Classification Techniques*. Springer: Singapore, 2020; Vol. 1235 CCIS. https://doi.org/10.1007/978-981-15-6648-6_10.
29. Taher, K. A.; Mohammed Yasin Jisan, B.; Rahman, M. M. Network Intrusion Detection Using Supervised Machine Learning Technique with Feature Selection. *1st Int. Conf. Robot. Electr. Signal Process. Tech. ICREST 2019* **2019,** 643–646. https://doi.org/10.1109/ICREST.2019.8644161.
30. Sultana, N.; Chilamkurti, N.; Peng, W.; Alhadad, R. Survey on SDN Based Network Intrusion Detection System Using Machine Learning Approaches. *Peer-to-Peer Netw. Appl.* **2019,** *12* (2), 493–501. https://doi.org/10.1007/s12083-017-0630-0.

31. Mishra, P.; Varadharajan, V.; Tupakula, U.; Pilli, E. S. A Detailed Investigation and Analysis of Using Machine Learning Techniques for Intrusion Detection. *IEEE Commun. Surv. Tutorials* **2019,** *21* (1), 686–728. https://doi.org/10.1109/COMST.2018.2847722.
32. Rajadurai, H.; Gandhi, U. D. A Stacked Ensemble Learning Model for Intrusion Detection in Wireless Network. *Neural Comput. Appl.* **2020,** *5*. https://doi.org/10.1007/s00521-020-04986-5.
33. Othman, S. M.; Ba-Alwi, F. M.; Alsohybe, N. T.; Al-Hashida, A. Y. Intrusion Detection Model Using Machine Learning Algorithm on Big Data Environment. *J. Big Data* **2018,** *5* (1). https://doi.org/10.1186/s40537-018-0145-4.
34. Liu, J.; Kantarci, B.; Adams, C. Machine Learning-Driven Intrusion Detection for Contiki-NG-Based IoT Networks Exposed to NSL-KDD Dataset. *WiseML 2020 - Proc. 2nd ACM Work. Wirel. Secur. Mach. Learn.* **2020,** 25–30. https://doi.org/10.1145/3395352.3402621.
35. Abdulhammed, R.; Faezipour, M.; Abuzneid, A.; Alessa, A. In *Enhancing Wireless Intrusion Detection Using Machine Learning Classification with Reduced Attribute Sets*, 2018 14th Int. Wirel. Commun. Mob. Comput. Conf. IWCMC, 2018; **2018,** pp 524–529. https://doi.org/10.1109/IWCMC.2018.8450479.
36. da Costa, K. A. P.; Papa, J. P.; Lisboa, C. O.; Munoz, R.; de Albuquerque, V. H. C. Internet of Things: A Survey on Machine Learning-Based Intrusion Detection Approaches. *Comput. Netw.* **2019,** *151,* 147–157. https://doi.org/10.1016/j.comnet.2019.01.023.
37. McElwee, S. In *Active Learning Intrusion Detection Using K-Means Clustering Selection*, *Conf. Proc.-IEEE SOUTHEASTCON*, 2017. https://doi.org/10.1109/SECON.2017.7925383.
38. Alsahli, M. S.; Almasri, M. M.; Al-Akhras, M.; Al-Issa, A. I.; Alawairdhi, M. Evaluation of Machine Learning Algorithms for Intrusion Detection System in WSN. *Int. J. Adv. Comput. Sci. Appl.* **2021,** *12* (5), 617–626. ttps://doi.org/10.14569/IJACSA.2021.0120574.
39. Verma, A.; Ranga, V. On Evaluation of Network Intrusion Detection Systems: Statistical Analysis of CIDDS-001 Dataset using Machine Learning Techniques. *Pertanika J. Sci. Technol.* **2018,** *26,* 1307–1332.
40. Belouch, M.; El Hadaj, S.; Idlianmiad, M. Performance Evaluation of Intrusion Detection Based on Machine Learning Using Apache Spark. *Procedia Comput. Sci.* **2018,** *127,* 1–6. https://doi.org/10.1016/j.procs.2018.01.091.
41. Park, S. T.; Li, G.; Hong, J. C. A Study on Smart Factory-Based Ambient Intelligence Context-Aware Intrusion Detection System Using Machine Learning. *J. Ambient Intell. Humaniz. Comput.* **2020,** *11* (4), 1405–1412. https://doi.org/10.1007/s12652-018-0998-6.
42. Gao, X.; Shan, C.; Hu, C.; Niu, Z.; Liu, Z. An Adaptive Ensemble Machine Learning Model for Intrusion Detection. *IEEE Access* **2019,** *7,* 82512–82521. https://doi.org/10.1109/ACCESS.2019.2923640.
43. Alhajjar, E.; Maxwell, P.; Bastian, N. Adversarial Machine Learning in Network Intrusion Detection Systems. *Expert Syst. Appl.* **2021,** *186,* 1–25. https://doi.org/10.1016/j.eswa.2021.115782.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

CHAPTER 14

DATA COMPRESSION FOR ACHIEVING COST-EFFICIENT AND SECURE DATA STORAGE OVER PUBLIC CLOUD: A PROPOSED MODEL

SUBHASH G. RATHOD[1], R. N. KHOBRAGADE[1], VILAS M. THAKARE[1], K. H. WALSE[1], GAJENDRA BAMNOTE[2], and SUSHAMA L. PAWAR[3]

[1]*Post Graduate Department of Computer Science & Engineering, SGBAU (MMIT, PUNE), Amravati, MS, India*

[2]*Ram Meghe College of Engineering, Amravati, Maharashtra, India*

[3]*VIIT, Pune (Infosys Pvt. Ltd. Pune), India*

ABSTRACT

Every day, millions of digital images are being generated and stored in the cloud. To free up space and give one another access to their confidential information from everywhere on any gadget, thousands of users have switched to the internet for hosting their confidential information. However, the privacy and security of private information can only be based on the dependability of the company that provides the cloud service. Most of the public cloud service providers do not assure data security and privacy. To guarantee data privacy and security, one must move to a private cloud. Also, higher data storage cost, access restriction, and data privacy are the major concerns in cloud platforms. Work in this regard has been achieved but with a minimalistic approach. This paper studies the technical challenges that come with constructing a cloud-based image processing system. We have explored various image processing tasks such as compression, which helps

Applied Soft Computing Techniques: Theoretical Principles and Practical Applications. Samarjeet Borah, Ratna Raja Kumar Jambi, Sharifah Sakinah Syed Ahmed, & Mahendra Prabhakar Deore (Eds.)
© 2025 Apple Academic Press, Inc. Co-published with CRC Press (Taylor & Francis)

to minimize storage cost, and fragmentation which helps to store images in chunks and provide extra security layers. To assess the system's efficacy, we ran a number of comprehensive research projects.

14.1 INTRODUCTION

In today's world, cloud is the major trend for data storage in a distributed way to overcome usage costs. The typical system uses heavy charges for both data usage and storage in their physical locations. Cloud computing is an approach that stores the data virtually in the storage system. This helps the user to avoid storage space and they can use it at anytime from anywhere. The cloud also helps to work efficiently for certain data access from multiple stations. The complexity of data storage is decreasing as researchers work on it. Services given by various organizations for data storage are easy with the cloud.[1,2,6]

Cloud computing is a system where loosely coupled data is used but the organization should know the correctness of the use; data correctness is the highest priority for any organization. In a system, different authorized users have been allowed to access cloud storage, but everyone has some restrictions.[13–15] The system should work efficiently but the system administrator is responsible for access control. Personal data security is a major concern while storing or transferring the data. Every activity should be passed through a secure manner as it is open storage and open to access for any users from the cloud. A restriction has been provided with the responsibility and give and take of data for a particular use; to restrict data usage, several methods have been introduced that cloud administrators can use. Restrictions can be imposed based on user role.[3,4]

Data storing capacity of the cloud provides great significance to users. The user will not care about data storage locations physically and the capacity or expenditure required for that. In the system, the user has no concern with the data-storing hardware requirements to execute the operation. Well-known examples of cloud systems are Amazon and Azure cloud. The users are free from the responsibility of local users to store the data. The overhead of data security is completely removed once the system is deployed over the cloud. In the user's aspect, even if he uses the cloud security model for data security priority should be higher.[6,9] In a cloud the security concern is because of the external attacks on the cloud data which leads to adulterating the content of data and violates the integrity of cloud storage. As the data correctness is not provided by most of the cloud individual users should worry about data integrity.[5,11] In response to the problem of security, some mechanism should be there to tackle the problems of security and provide meaningful architecture. In a mechanism, the data should be free from the risk of interactions with any cloud users. In response, the method has included the cryptography mechanism with significant modifications.[20–22]

Data over cloud can be accessed by any user with appropriate permission. Many cloud service providers offer data backup but do not ensure the security of data. Cloud-based services offer third-party storage; these third-party storage companies will not guarantee data privacy and security if one uses those services for free, also they can share one's personal data for financial gain. In the public cloud, organizations do not have complete command over their data. To get full access, one must purchase private cloud. In terms of privacy of data, data owners must be assured of data privacy and data access control.[12,15]

Various modern cloud platforms can save businesses money by providing free services such as storage. Free cloud storage allows users to store a limited amount of data so as new data is generated, one must upgrade storage space and might have to pay hefty fees.[16,21]

To achieve minimum storage cost, we are compressing the data before uploading it over the cloud; data compression is done in such a way that the quality of digital media will not degrade. Here as per the result evaluation, we have saved up to 60% of storage cost. To provide security to the data we are fragmenting data into chunks and then encrypting those chunks using an asymmetric key encryption algorithm that is RSA. The Owners private key will encrypt the generated chunks and the owner's public key will decrypt the chunks. Here we are providing two layers of security by fragmenting a single file into chunks and then encrypting those chunks. Our test findings demonstrate that the recommended approach delivers excellent safety at a minimal cost of maintenance.[8–10]

14.2 RELATED WORKS

a. Fast access to the cancellation and secure transfer of data across untrusted networks
b. Secure data sharing

   It is a novel conceptual approach that reinforces data confidentiality, transmission, and sharing, and its architecture consists of a data owner, multiple end users, and two storage providers, because they may be incredibly prone to outside hostile assaults; this writer made the assumption that storage suppliers are unreliable in this instance. It is also believed that user-to-user communication paths could be vulnerable to a man-in-the-middle attack.[22,23,28]

The data owner uploads data into a set of fragments, while transferring data it gets encrypted and the author uses symmetric encryption. The data transfer is (AON) type that is all or nothing type. It means that all the fragments are required to decrypt in the encrypted fragments. These fragments are then stored in two different cloud storage. Encryption keys are shared among respective users but not to cloud providers. Users can then access outsourced data from two different clouds and decrypt the fragments with the key[7,26,27]

c. Fast access revocation

Fast access revocation adopted the Mix Slice approach. This approach re-encrypts the random fragments of the outsourced data, and a new key is generated for re-encryption. This new key is then distributed to only those users who have access to the resource, after this access to the resource gets revoked for only those users who possess old keys; in this way, users with old keys will not be able to reconstruct the data.[8,9,17]

Protected storage in the cloud with owner-side integrated cloud-side access management combination. Cloud storage has three entities: data owners, data users, and the cloud provider.

- *Data owner*

Data proprietors are responsible for data publishing over the cloud and recompense for the source depletion and publisher of files.

- *Data users*

These users are authenticated by cloud service providers, and these are users who download the data uploaded by owners.

- *Cloud provider*

It provides services such as user authentication and resource storage and sharing. In this scenario, the cloud is not publicly accessible as users must be authenticated before downloading the data.

Here data owners are not always online so in order to restrict resource consumption, the system provides an option for data owners to specify a

TABLE 14.1 Gap Analysis on the Existing Methodologies.

Sl. no.	Title	Methodology	Gap analysis
1	"Secure Data Sharing with Fast Access Revocation through Untrusted Clouds (IEEE 2019)"	The author proposed a plan for document distribution among data owners and other consumers without breaking the security and privacy barriers, the system is designed to store data over untrusted cloud storage providers. All or Nothing (AON) methodology is used to transform data, before uploading data over the cloud it gets fragmented and distributed on two separate clouds by this way no cloud service provider has users' entire data so the dishonest or untrusted cloud service provider cannot read the data, not only the fragments are generated it also encrypt the newly generated fragments.	The author does not used any third-party auditing scheme. No storage optimization.
2	"SecACS: Enabling lightweight secure auditable cloud storage with data dynamics. (Elsevier 2020)"	The author proposed a "Secure Auditable Cloud Storage (SecACS) scheme" supportive document dynamics. The author used lightweight cryptographic operations to cut down computational time. The author preferred symmetric key encryption algorithm. Evaluated results demonstrations that the proposed organization is faster in relationships of outsourcing the data and data integrity verification.	Although certain open-source cloud-based storage platforms fail to protect customer files, researcher research demonstrates that the system is designed for preserving consistent data and demonstrates that the technology is probably safe.
3	"An ID-Based Privacy-Preserving Integrity Verification of Shared Data Over Untrusted Cloud (IEEE 2020)"	Focuses on the security issues while auditing the shared data. In this paper, every user is provided with special ID and data is shared based on that ID. System allows to create a group of users where they can share data within the group. All users in the group can update, delete and insert new block into the data, misbehaving user can be identified easily, cloud system admin can easily revoke the rights of misbehaving users, while auditing the data cloud maintains the privacy of data from the public verifier.	Research of the technology shows that it performs accurate data confirmation, however personal files on cloud servers are not safeguarded.

TABLE 14.1 *(Continued)*

Sl. no.	Title	Methodology	Gap analysis
4	"A Secure Data Dynamics and Public Auditing Scheme for Cloud Storage. (IEEE 2020)"	In this study, an independent monitoring methodology is suggested as a secure public reporting architecture. The paper's writer mainly concentrates on the system's dependability in terms of protecting private information and upholding data consistency. The author employed the RSA algorithm for key encoding, the 256-bit AES technique for encrypting information, and SHA for integrity of data verification.	No provision for storage cost minimization
5	"Improving Security Data Access Control for Multi-Authority Cloud Storage (IEEE 2019)"	Proposed paper focuses on collusion attacks, the author proposed a data access control protocol for multi-authorized cloud storage called "improving security data access control scheme" (ISDAC-MACS)	Takes too much time to complete the owner's process and no provision to check data integrity.
6	"Privacy-Preserving Public Auditing Scheme for Data Confidentiality and Accountability in Cloud Storage (IEEE 2019)"	Author suggests an unrestricted reviewing framework for maintaining data concealment, here users are dependent on third-party auditor for auditing. Author proposed special log called attestation where hash value is generated using user pseudonym to preserve users' privacy	No provision for storage cost minimization

maximum number of downloads for specific users.[29,31] This is implemented by using partially outsourced protocol (POP) and fully outsourced protocol. In partially outsourced protocol owner challenges user verification before giving access to the data and in fully outsourced protocol cloud service provider challenges of user verification. Framework CP-ABE is a public key encryption scheme for access control.[18–20]

14.3 DESIGN AND IMPLEMENTATION

We proposed a constructive and de-coupled data-sharing framework. It provides privacy to user's data by restricting other users to access data. Only owner-approved users can access the data. Our main goal is to provide data security and maintain minimum storage costs. We are using a data fragmentation methodology rather than storing the whole file on a cloud server. We are dividing the file into several chunks and the file gets reconstructed on demand if the owner approves a user request to access the file.

14.3.1 OWNERS PROCESS

The person who has ownership of the information oversees publishing it to the server in the cloud and may share it with other users upon permission. Data owners can define a set of rules as to how different users can access the data depending on their roles. Uploaded data goes through three processes: A document initially gets condensed after which it is divided up into multiple blocks. Metadata concerning the created chunks is then recorded in a log file, and it is subsequently encrypted and kept on an online server.

14.3.1.1 COMPRESSION

This module reduces storage size by half taking pictures as input and compressing them using the Discrete Cosine Transform (DCT) algorithm.

We employ DCT due to it is straightforward to compute, segmented (you may create many DCTs of rows and columns), that has "power integration" characteristics.[12,30] DCT performance is best described as:

- Image 2RLE reads the image and performs the DCT, using the simulation (Q-Matrix derived from the standard JPEG matrix obtained from the visual and visual test) and encoding using Run Length Encoding.

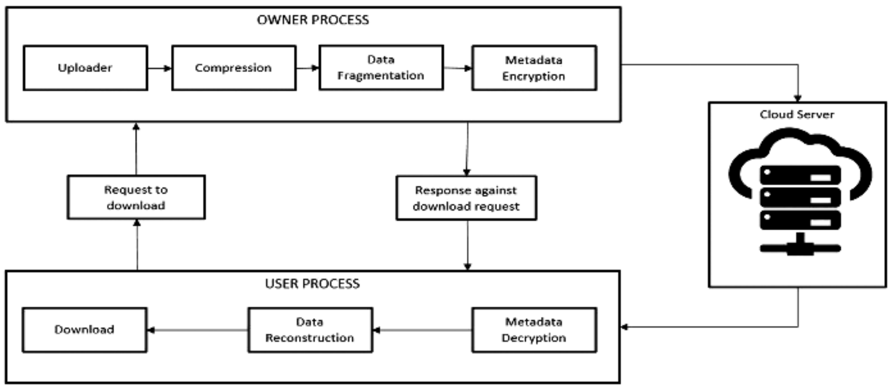


FIGURE 14.1 Proposed system.

- The encoded data is saved into your request for the smaller bits of the original image.
- RLE2image reads previously saved text file and then decodes the encoded bits to construct new image with reduced physical storage size to disk.

Algorithm: Data Compression

Data: String path (Original Image Path)

Result: String Path' (Compressed Image path)

1. ImageCodecInfo jpgEncoder = GetEncoder(ImageFormat.Jpeg);
2. System.Drawing.Imaging.Encoder QualityEncoder = System. Drawing.Imaging.Encoder.Quality;
3. EncoderParameters myEncoderParameters = new EncoderParameters(1);
4. EncoderParameter myEncoderParameter = new EncoderParamete r(QualityEncoder, quality);
5. myEncoderParameters.Param[0] = myEncoderParameter;
6. bmp1.Save(Path', jpgEncoder, myEncoderParameters);

14.3.1.2 CHUNK GENERATION (FRAGMENTATION)

This module fragments the image into pieces, after fragmenting the whole file into pieces information about created pieces is stored in a special log file. The file contains all the data associated with newly generated fragments. This log file is also stored over the cloud.

Algorithm: Data Fragmentation

Data: Image I

Result: byte[] Fragmented image, File log

1. Read(I)
2. int image_size = fileByteArray.length
3. byte[] subchunksize = image_size / number_of_fragments
4. Save subchunksize
5. Save subchunksize in log file

14.3.1.3 METADATA ENCRYPTION

We use AES to encrypt a metadata file instead of DES

- *AES* is much faster than DES almost 3 times faster.
- Triple-DES is easier to crack than AES.

Algorithm: "Log (Metadata) Encryption"

Statistics: Fragmented Arrangement fs, key

Result: Encrypted Fragmented Sequence byte[] cipherBytes

1. foreach fragment in fs
2. final Cipher cipher = Cipher.getInstance(this.algorithm);
3. cipher.init(Cipher.ENCRYPT_MODE, key);
4. cipherBytes = cipher.doFinal(fragment);
5. Write(cipherBytes, path);
6. End foreach

14.3.2 USER PROCESS

14.3.2.1 DATA RECONSTRUCTION

The information reconstruction component oversees putting together photographs from various fragments that the individual in question originally uploaded onto the internet. This component locates the actual locations of the broken chunks and then uses a file with metadata to integrate them before sending the specified images to the individual who requested it.

Algorithm: Data Defragmentation

Data: byte[] Fragmented image I'

Result: Image I

1. foreach fragment in I'
2. byte temp = Read(I[index])
3. I.BufferWrite(temp)
4. Write I
5. End
6. return I

14.3.2.2 DECRYPTION

We require a series of pieces prior to reconstructing an image file. A metadata file in compressed form is assembled and contains knowledge about the components. The encrypted information file is thus decrypted by cryptographic techniques and then sent to this information reconstructing mechanism.

Algorithm: Log Decryption

Data: Encrypted Fragmented Sequence Metadata File, Cipher key

Result: Plain Fragmented Sequence Metadata File F

1. Cipher cipher = Cipher.getInstance(RSA);
2. cipher.init(Cipher.DECRYPT_MODE, key);
3. decryptedBytes = cipher.doFinal(cipherBytes);
4. return decryptedBytes;

14.4 RESULTS AND ANALYSIS

14.4.1 DATASET AND COMPRESSION ANALYSIS

For image compression implementation we can use any dataset like ImageNet which is a subset of Kaggle dataset, we do not need a labeled dataset as we are not doing any object detection here, we are only reducing the current image physical size by more than 50% without visually reducing image quality.

We are using lossy compression, considering the image size is more than 5 MB most of our images will be in JPEG format so we are using Transform Coding which is a most used method for compression. Initially, we were considering two standard compression methods either DCT or DFT but in the end, we zeroed down to the following DCT reasons.

In compressed image techniques including JPEG, DCT is preferred above DFT due to the fact that is a real transformation that yields simply one real integer per point in the data. In comparison, a DFT yields an intricate number that needs twice as much information to store.

- *DCT is modest and quicker than DFT*

"DCT"—a compressed file technique that is most frequently employed. The DCT, which belongs to a category of discrete cosine encodes (see discrete cosine inversion), is an example of Fourier-related transformation. It is usually referred to as "DCT-II" in this context. It is the least effective type of compression for images.

- *DCT with quality factor*

We did not use plain DCT rather we edited existing algorithm and added Quality Factor because we need to control the quality of the image and compressing image with a certain quality factor is a way to make sure we are not losing image quality.

14.4.2 EFFICIENCY ANALYSIS

We first measure the effectiveness of the suggested architecture in relation to the time required to complete the task of data compression and fragmentation because both tasks are heavy and consume time as per the size of the image. While compressing we found the middle ground in terms of image quality of the compressed file and size of the compressed file. We are compressing images with a quality factor of 0.2f which not only maintains image quality but also reduces storage size by more than 50%. As shown in Figures 14.2 and 14.3, there is no difference in terms of visual quality.

FIGURE 14.2 Original image storage efficiency analysis.

FIGURE 14.3 Compressed image with QF of 0.2 f.

As shown in notation Table 14.1, storage allocation model is a “mathematical model” that is used to calculate statistics transfer, data per application and storage cost, and storage allocation cost in cloud computing.

TABLE 14.2 Notation of Storage Model.

St	Allocated storage
Tr	Data transfers from storage to web platform
DB_{size}	Database size in GB
DB_{usage}	Database accessed by per user
DB_{req}	Number of monthly storage requests in database for per user
$PerReq_{cost}$	Cost of each request from user
St_{avg}	Average storage cost
St_{cost}	Total storage cost
Tr_{cost}	Total transfer cost
Req_{cost}	Total request cost

$$St_{avg} = \frac{\sum St_{cost}\ X\ Tr_{cost}\ X\ Req_{cost}}{Total\ GB\ stored} \tag{14.1}$$

$$St_{cost} = \sum DB_{size}\ X\ St \tag{14.2}$$

$$Tr_{cost} = \sum Tr\ X\ DB_{size}\ X\ DB_{usage} \tag{14.3}$$

$$Req_{cost} = \sum St\ X\ DB_{req}\ X\ PerReq_{cost} \tag{14.4}$$

Our goal is to archive lower storage costs by reducing the data size by almost half without compromising the data or data quality. Data size reduction can reduce the transfer cost over the communication network. Minimizing data size can also maximize the storage allocation space.

Table 14.2 and Figure 14.3 show the comparative analysis of image compression with variant quality factor and time required for compression, as the quality factor increases image storage size increases, but the time required for compression decreases.

TABLE 14.3 Image Compression Time with Superiority Influence.

Image tag	Novel size	Compacted size	Superiority factor	Time to compress
butterfly.jpg	6.34 MB	196 KB	0.1f	865 ms
	6.34 MB	306 KB	0.2f	783 ms
	6.34 MB	434 KB	0.3f	797 ms
	6.34 MB	551 KB	0.4f	791 ms

The variation among the disc space sizes of the initial picture and the compressed version is displayed in Table 14.3 and Figure 14.4. Statistics indicate that the researchers were able to reduce the size of the images by over half of the initial size.

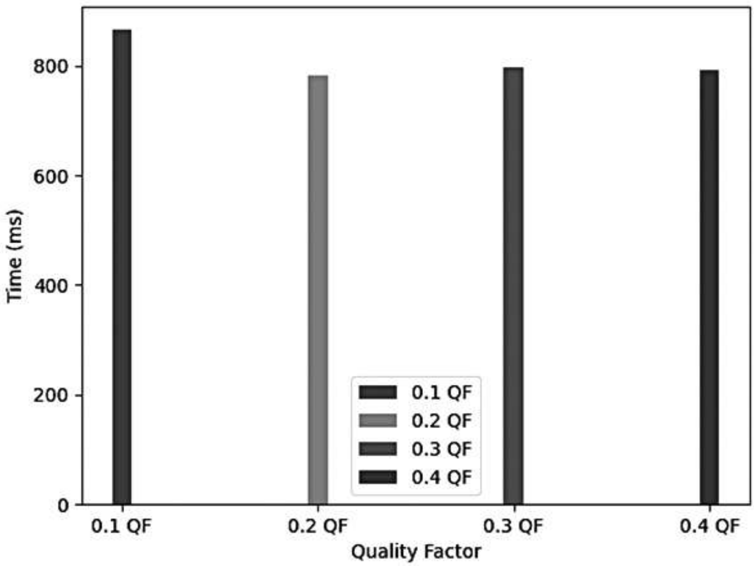


FIGURE 4.4 Image compression time.

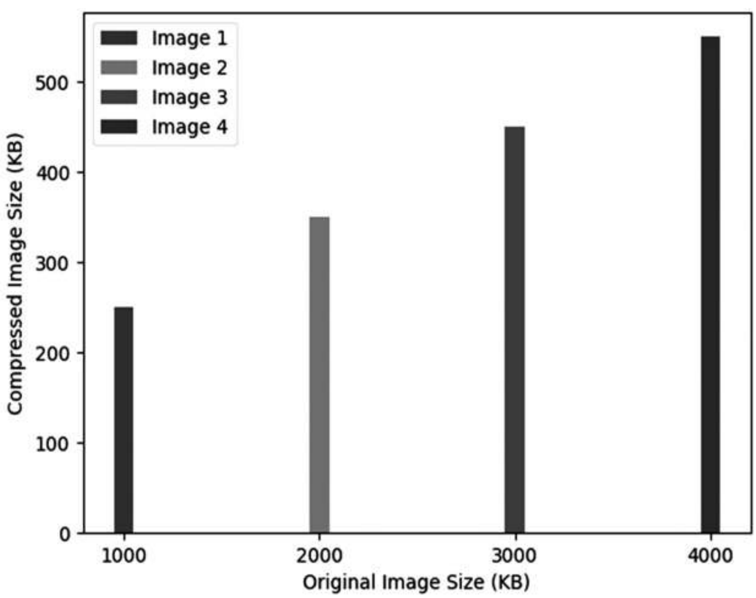


FIGURE 14.5 Image compression time.

TABLE 14.4 Image Compression Time.

Image	Novel size (KB)	Compressed Size (KB)
Image 1	1000	250
Image 2	2000	350
Image 3	3000	450
Image 4	4000	550

Table 14.3 and Figure 14.5 show the time required to fragment the image. Here the size of one fragmented chunk depends on the size of the image.

$$\text{Image chunk size} = \text{Image size} \div \text{Number of chunks} \tag{14.4}$$

Here size of image 1 is 250 KB then the size of one chunk if we must split the image into eight equal chunks, then 250/8 = 31.5

TABLE 14.5 Image Fragmentation Time.

Image	Size (KB)	Time (ms)
Image 1	250	80
Image 2	350	90
Image 3	450	100
Image 4	550	150

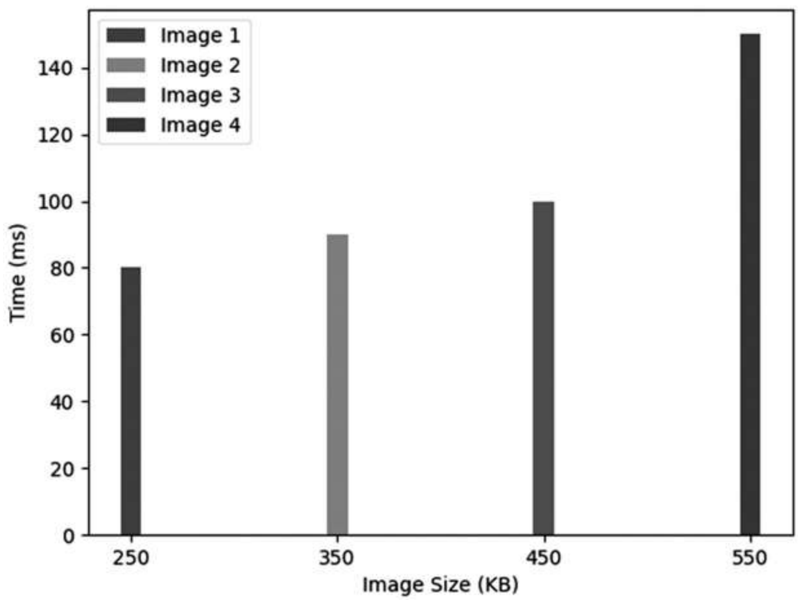


FIGURE 14.6 Image splitting time.

TABLE 14.6 Image Reconstruction Time.

Image	Size (KB)	Time (ms)
Image 1	250	80
Image 2	350	90
Image 3	450	100
Image 4	550	170

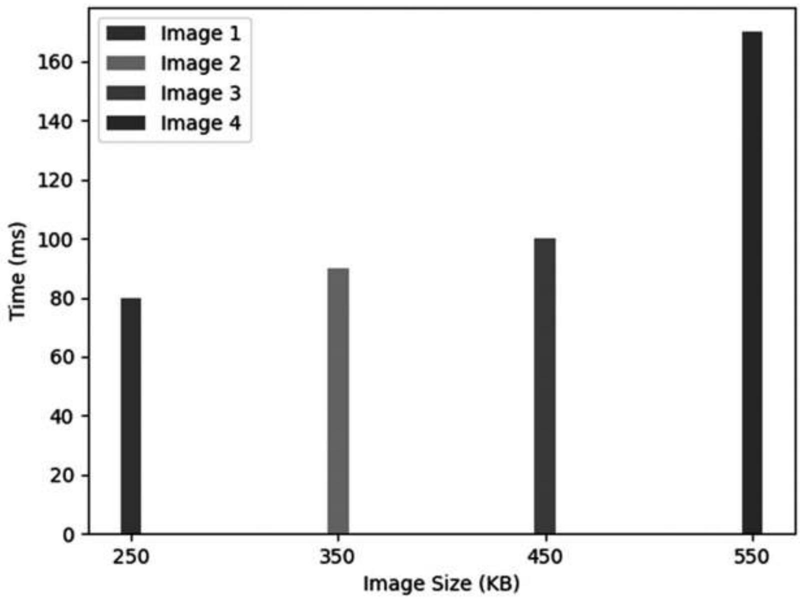


FIGURE 14.7 Image reconstruction time.

14.4.3 CORRECTNESS ANALYSIS

Utilizing the peak signal-to-noise ratio (PSNR) measurement, information precision is assessed. PSNR values are calculated in logarithmic (dB) scale. Logarithmic scale is a metric used to quantity the quality of any image reconstructed, with respect to its original image. The threshold value of good image quality is 20 dB and higher PSNR value indicates that the image is of higher quality. We have compressed the image with different quality factors and fragmented it after that and then reconstructed the same image. We found that the decrease in compression quality factor is directly proportional to lower PSNR value.

TABLE 14.7 Quality Factor and PSNR Value.

Original image name	Quality factor	PSNR value
Image 1	0.1	30.554785018837
	0.2	31.489643361814
	0.3	32.727103347367
	0.4	33.748332381816

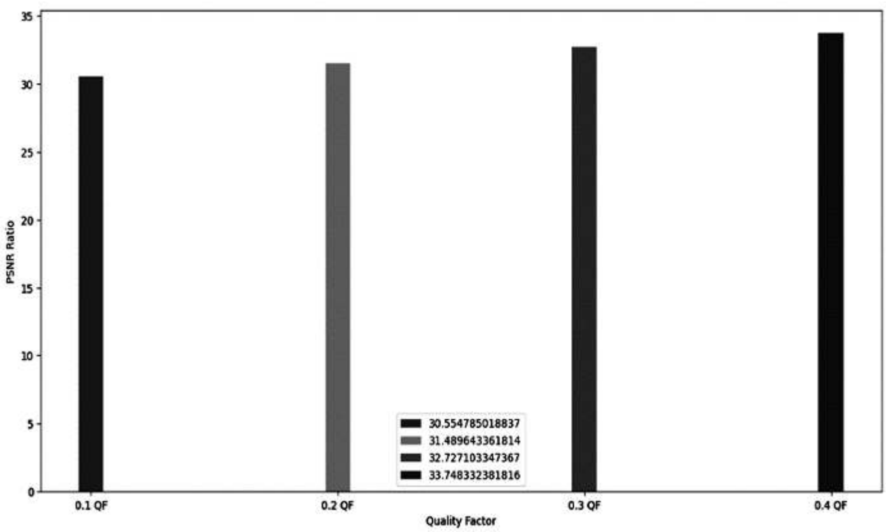


FIGURE 14.8 Quality factor and PSNR value.

TABLE 14.8 Effectiveness Characteristic Assessment for Different Structures.

	[1]	[2]	[3]	Proposed
Auditability	-	✓	✓	✓
Documents confidentiality	✓	✓	✓	✓
Confidentiality stabilizing	✓	✓	✓	✓
Parting of callings	✓	✓	-	✓
Data security	✓	-	-	✓
Low storage cost	-	-	-	✓

To assess the effectiveness of the approach, we first contrast three existing structures with the suggested plan, as given in Table 14.1. This contrast demonstrates that the suggested approach provides every detail monitoring functionality.

14.5 CONCLUSIONS

We have outlined a framework in this work that focuses on secure data storage at a low cost while utilizing approaches from multiple fields. With this architecture, data owners can take advantage of data security and privacy with minimal communication costs. Data owners can also provide access restrictions over their data, that is only owner-approved users can access documents. Information breakdown, which essentially divides information into numerous parts and stores metadata regarding the fractured bits in a log file that is later encrypted, is a strategy for achieving the security of information. To rebuild information, this log file is employed. Result analysis shows that we were able to achieve effectiveness and robustness while compressing, we managed to reduce image storage size by more than 50% and not compromise image quality. Data correctness is analyzed by evaluating reconstructed images in terms of PSNR value. The PSNR value of the reconstructed image is more than 20 dB which indicates that the image quality of the reconstructed image is maintained. Adding another security layer and providing data integrity on top of the proposed assembly is our significant forthcoming effort to be followed.

ACKNOWLEDGMENT

We are thankful to Dr. Ratnashil N. Khobragade who guided us throughout the journey of this research. We would like to express our gratitude to Dr. Vilas M. Thakare, who guided us throughout this project. We also want to express our appreciation to our friends and family for their encouragement, profound knowledge of the subject, cooperation, and assistance. We would like to express our gratitude to the technical as well as workers of the University of Amaravati's department for their assistance. I would also like to show my deep appreciation to the supervisors Dr. Gajendra Bamnote, Dr. K. H. Walse, and Dr. S. V. Dudul who helped to finalize the paper.

We are thankful to the Honorable Pro-VC and VC of Sant Gadge Baba Amravati University for their direct and indirect support during the research work.

KEYWORDS

- **storage cost**
- **cloud security**
- **data privacy**
- **data fragmentation**
- **data compression**

REFERENCES

1. Kapusta, K.; Qiu, H.; Memmi, G. In *Secure Data Sharing with Fast Access Revocation through Untrusted Clouds*, 10th IFIP International Conference on New Technologies, Mobility, and Security (NTMS), 2019; IEEE, June 2019; pp 1–5.
2. Li, L.; Liu, J. SecACS: Enabling Lightweight Secure Auditable Cloud Storage with Data Dynamics. *J. Inf. Secur. Appl.* **2020,** *54,* 102545.
3. Rabaninejad, R.; Sedaghat, S. M.; Attari, M. A.; Aref, M. R. In *An ID-based Privacy-Preserving Integrity Verification of Shared Data Over Untrusted Cloud*, 25th International Computer Conference, Computer Society of IRAN (CSICC), January 2020; IEEE, 2020; pp 1–6.
4. Singh, P.; Saroj, S. K. In A Secure Data Dynamics and Public Auditing Scheme for Cloud Storage, 6th International Conference on Advanced Computing and Communication Systems (ICACCS), March 2020; IEEE, 2020; pp 695–700.
5. Wang, J.; Wu, K.; Ye, C.; Xia, X.; Ouyang, F. In *Improving Security Data Access Control for Multi-Authority Cloud Storage*, IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/SustainCom), December 2019; IEEE, 2019; pp 608–613.
6. Yang, Z.; Wang, W.; Huang, Y.; Li, X. Privacy-Preserving Public Auditing Scheme for Data Confidentiality and Accountability in Cloud Storage. *Chin. J. Electron.* **2019,** *28* (1), 179–187.
7. Chen, F.; Meng, F.; Xiang, T.; Dai, H.; Li, J.; Qin, J. Towards Usable Cloud Storage Auditing. *IEEE Trans. Parallel Distrib. Syst.* **2020,** *31* (11), 2605–2617.
8. Zhang, L.; Cui, Y.; Mu, Y. Improving Security and Privacy Attribute-based Data Sharing in Cloud Computing. *IEEE Syst. J.* **2019,** *14* (1), 387–397.
9. Mohanaprakash, T. A.; Andrews, J. In *Novel Privacy Preserving System for Cloud Data security using Signature Hashing Algorithm*, International Carnahan Conference on Security Technology (ICCST), October 2019; IEEE, 2019; pp 1–6.
10. Tao, Y.; Xu, P.; Jin, H. Secure Data Sharing and Search for Cloud-Edge-Collaborative Storage, *IEEE Access.* **2019,** *8* (2020) 15963–15972.

11. Ding, W.; Hu, R.; Yan, Z.; Qian, X.; Deng, R. H.; Yang, L. T.; Dong, M. An Extended Framework of Privacy-Preserving Computation with Flexible Access Control. IEEE Trans. Netw. Serv. Manag. **2019,** *17* (2), 918–930.
12. Su, H.; Jung, C. Perceptual Enhancement of Low Light Images based on Two-Step Noise Suppression. *IEEE Access* **2018,** *6,* 7005–7018.
13. Azeez, N. A.; Van der Vyver, C. In *Security Challenges and Suggested Solutions for e-Health Information in Modern Society*, 5th EAI International Conference on IoT Technologies for HealthCare; Springer International Publishing, 2020; pp 57–72.
14. Wei, J.; Liu, W.; Hu, X. Secure and Efficient Attribute-Based Access Control for Multiauthority Cloud Storage. *IEEE Syst. J.* **2016,** *12* (2), 1731–1742.
15. Ranjan, R.; Rana, O.; Nepal, S.; Yousif, M.; James, P.; Wen, Z.; ... Dustdar, S. The Next Grand Challenges: Integrating the Internet of Things and Data Science. *IEEE Cloud Comput.* **2018,** *5* (3), 12–26.
16. Xue, K.; Chen, W.; Li, W.; Hong, J.; Hong, P. Combining Data Owner-Side and Cloud-Side Access Control for Encrypted Cloud Storage. *IEEE Trans. Inform. Forensic. Secur.* **2018,** *13* (8), 2062–2074.
17. Ning, J.; Cao, Z.; Dong, X.; Liang, K.; Wei, L.; Choo, K. K. R. CryptCloud $^+ $+: Secure and Expressive Data Access Control for Cloud Storage. *IEEE Trans. Serv. Comput.* **2018,** *14* (1), 111–124.
18. Wang, C.; Wang, Q.; Ren, K.; Cao, N.; Lou, W. Toward Secure and Dependable Storage Services in Cloud Computing. IEEE Trans. Serv. Comput. **2011,** *5* (2), 220–232.
19. Wang, C.; Wang, Q.; Ren, K.; Cao, N.; Lou, W. Toward Secure and Dependable Storage Services in Cloud Computing. *IEEE Trans. Serv. Comput.* **2011,** *5* (2), 220–232.
20. Kumar, S. P.; Subramanian, R. An Efficient and Secure Protocol for Ensuring Data Storage Security in Cloud Computing. *Int. J. Comput. Sci. Issues (IJCSI)* **2011,** *8* (6), 261.
21. Wang, C.; Zhang, B.; Ren, K.; Roveda, J. M. Privacy-Assured Outsourcing of Image Reconstruction Service in Cloud. *IEEE Trans. Emerg. Top. Comput.* **2013,** *1* (1), 166–177.
22. Wang, C.; Chow, S. S.; Wang, Q.; Ren, K.; Lou, W. Privacy-Preserving Public Auditing for Secure Cloud Storage. *IEEE Trans. Comput.* **2011,** *62* (2), 362–375.
23. Yang, K.; Jia, X. An Efficient and Secure Dynamic Auditing Protocol for Data Storage in Cloud Computing. *IEEE Tran. Parallel Distrib. Syst.* **2012,** *24* (9), 1717–1726.
24. Wang, B.; Li, B.; Li, H. Oruta: Privacy-Preserving Public Auditing for Shared Data in the Cloud. *IEEE Transac. Cloud Compu.* **2014,** *2* (1), 43–56.
25. Swathi, R.; Subha, T. In *Enhancing Data Storage Security In Cloud Using Certificateless Public Auditing*, 2nd International Conference on Computing and Communications Technologies (ICCCT); February, 2017; IEEE, 2017; pp 348–352.
26. Rathod, S.; Khobragade, R. N.; Thakare, V. M.; Walse, K. H.; Pawar, S. In *Lightweight Auditable Secure Cloud Storage With Privacy Enabled Data Storage Optimization*, 2022 IEEE International Conference on Blockchain and Distributed Systems Security (ICBDS), Pune, India, 2022; pp 1–6, doi: 10.1109/ICBDS53701.2022.9935980.
27. Rathod, S.; Khobragade, R. N.; Thakare, V. M.; Walse, K. H.; Pawar, S. In *Model for Efficient Data Storage on Public Cloud*, 2022 IEEE International Conference on Blockchain and Distributed Systems Security (ICBDS), Pune, India, 2022; pp 1–5, doi: 10.1109/ICBDS53701.2022.9935881.
28. Rathod, S. G. Security for Shared Data over Public Cloud for Maintaining Privacy. *Math. Stat. Eng. Appl.* **2022,** *71* (4), 7167–7173.

29. Salunke, M.; Kabra, R.; Kumar, A. Layered Architecture for DoS Attack Detection System by Combined Approach of Naive Bayes and Improved K-means Clustering Algorithm. *Int. Res. J. Eng. Technol.* **2015,** *2* (3), 372–377.
30. Mangesh, D. S.; Kumar, P. A Proposed Methodology to Mitigate the Ransomware Attack. In *Recent Trends in Intensive Computing*; IOS Press, 2021; pp 16–21.
31. Salunke, M. D.; Kabra, R. Denial-of-Service Attack Detection. *Int. J. Innov. Res. Adv. Eng.* **2014,** *1* (11), 16–20.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

CHAPTER 15

ACCESSIBILITY OF DATA LINKED TO LARGE SCALED HETEROGENEOUS DATABASES USING GENERALIZED APPROACH: A SOFT COMPUTING-BASED APPROACH

JYOTI PRAKASH MISHRA[1], ZDZISLAW POLKOWSKI[2], and SAMBIT KUMAR MISHRA[1]

[1] Gandhi Institute for Education and Technology, Baniatangi, Affiliate to Biju Patnaik University of Technology, Rourkela, Odisha, India

[2] The Karkonosze University of Applied Sciences in Jelenia Góra, Poland

ABSTRACT

In broad sense, optimizing queries in general can be a fundamental aspect linked to large scaled heterogeneous databases. Particularly, while replicating as well as performing storing options in databases associated with several disjoint locations, it is really required to focus on the suitability of sites during query execution. Likely, the process of choosing the most efficient query evaluation plans from the prescribed strategies should be feasible toward processing certain queries, particularly in complex situations. Assume that a large scaled database system is being provisioned with a finite number of resources in a way that computational capabilities of various resources can be shared by the users. Also the users complete the execution by the desired heterogeneous system. The assigned tasks are linked with provisioned parallel tasks. In such a situation, it is really difficult to decompose the

Applied Soft Computing Techniques: Theoretical Principles and Practical Applications. Samarjeet Borah, Ratna Raja Kumar Jambi, Sharifah Sakinah Syed Ahmed, & Mahendra Prabhakar Deore (Eds.)
© 2025 Apple Academic Press, Inc. Co-published with CRC Press (Taylor & Francis)

independent relationship and further to allocate in the host sites. Considering the data dependency within the parallel tasks into account, the total execution time associated with the queries depends on the maximum execution time of the parallel tasks. Accordingly, the particle swarm optimization (PSO) algorithm can be implemented prioritizing the fitness parameters accounting the required memory while executing the queries. PSO is considered one of the techniques used in soft computing. It is often used alongside other soft computing methods to solve complex optimization problems. PSO can be used in conjunction with fuzzy logic, neural networks, or other soft computing techniques to improve the performance and efficiency of problem-solving tasks. The time required to execute the queries along with the cost can be well taken toward measuring the performance of queries particularly in the allocated locations. The algorithm can act on the basis of swarm of early responses. In fact, it can act like the evolutionary computing mechanisms to obtain the optimized response by moving the early responses in successive repetitions.

15.1 INTRODUCTION

In general, during formulation of the problems linked to allocation of resources, it is required to prioritize the locations of the host systems along with the allocation of resources and query responses. The computing capabilities with the large scaled systems permit linking with the large pool of shared computing resources. Somehow, these resources are provisioned with the services and obtained the permission to be plugged with other databases of concern. Also being linked with the complex tasks, the computing resources can be managed along with the deployed applications. While managing the large scaled data, it is also needed for frequent update and consistency must be preserved at each and every scale. In specific perspective, provisioning the high availability as well as scalability can be a challenge as far as large scaled databases are concerned. The systems can be scaled up by incorporating computing resources to the processing elements and by that these can permit the systems to manage the large sized clusters. Sometimes, the systems with high scalability features may respond to the changes by enhancing their processing capabilities. In fact, the applications deployed within the large scaled systems may utilize the resources managing the resources to handle the loads. In this connection, the complete systems can depend on the ability toward scaling as well as be provisioned to manage the processing abilities as cited in Figure 15.1. The challenges linked with the deployed database systems in such situations can be attributed to obtain optimality. Accordingly, the large scaled database systems will maintain consistency within the databases provisioning high scalability and consistency. While provisioning the resources linked with the physical machines it is observed that each processing element is being processed with independent operating system. The physical resources can be managed being provisioned with the specific mechanisms. Sometimes, databases of concern can be partially replicated to enhance the performance of databases. The primary copy of data in such a situation may be referred as master replica which can be accessible through secondary replica. However, the replication may be implemented within the database systems across distributed machines. Also the synchronous replication may be used to keep the secondary data in proper state with the primary replica.

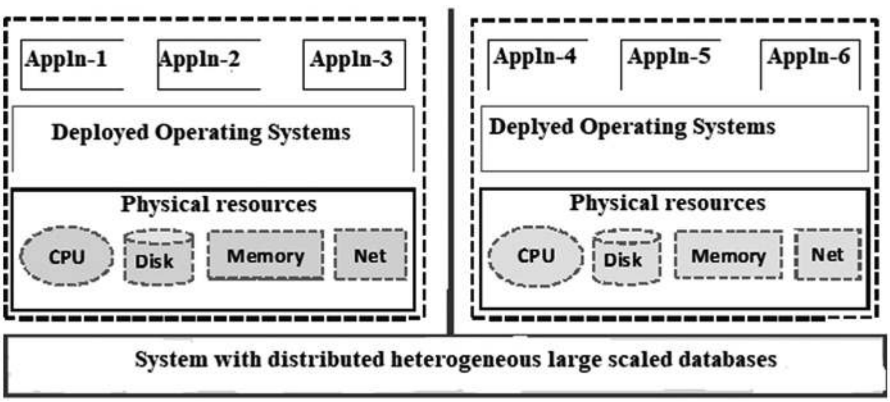


FIGURE 15.1 System applicability with distributed heterogeneous databases.

15.2 REVIEW OF LITERATURE

In this section, we are trying to focus on some of the notable research works on our domain of interest.

Shivle et al.[1] in their work focused on development of query plans on the basis of directed acyclic graph. They also prioritized on min–min algorithm toward sub-query scheduling linked to query plans.

Naijing et al.[2] during their study discussed on grid database based on the combination of databases with grid where the traditional database queries and their optimization techniques may not meet the requirements of grid databases provisioned with dynamic features.

Soliman et al.[3] in their work concentrated on the containerization as well as deployment of query evaluation plans. They focused on the layered system architecture and presented the mechanisms to containerize the queries. The containerized query may have the potential of achieving scalability for large data sets.

Tang et al.[4] in their research focused on the development of light-weight software container technologies. They have also concentrated on the performance of container technology for data-intensive applications.

Wilson et al.[5] in their work prioritized on the parameterized mapping toward transformation of output by a specific mechanism. Somehow, they observed unsuitability in the mapping process and induced proper distribution to obtain greater flexibility.

Rajaraman et al.[6] in their work focused on natural language processing to observe the generalized features as well as generation techniques. They tried to concentrate on frequency-inverse document frequency which is the ratio of how frequently a word shows up in a document to how often it shows up in the whole corpus of documents.

Wang et al.[7] in their work prioritize on the Latent Dirichlet Allocation technique toward transformation of corpus of documents into document-topic mappings. This technique is being proved to be general purpose enough to be useful in many document classification tasks.

Gligorijevi´c et al.[8] in their work discussed the usefulness of large-scale data management as well as semantics toward obtaining the combined and matched data. As such in the large scaled databases, the information characterizes with heterogeneity being provisioned with computing resources and storage.

Dong et al.[9] discussed integration of large scaled data in their research observed the steps to align the schema inclusive of data fusion. Generally toward integrating the large scaled data it is required to link up with the original data source and the role of the schema is to determine the matching linked with the attributes. The schema mapping is based on the source schema to determine the semantic relationships between the types of schemes.

Abd El-Ghafar et al.[10] in their work focused on mechanisms of partitioning the record sets to comply with the identified records lined to distinct entities. Basically, it is performed within blocking, pairwise matching as well as clustering. Also it follows the strategy toward scaling records linkage to large data sets.

Knight et al.[11] in their work prioritized the specific architectural concepts inclusive of control flow, data flow, event handler, and associated parameters. In fact, the control flow component is the core of the mechanism that acts as a manager defining the order of execution for the tasks. Also the data flow component can be used toward extraction as well as transformation of data.

Garcia Molina et al.[12] focused on complexities and enhancement of data as well as business requirements linked with information systems. They observed the consequences of usage of multiple rational data sources. As it is understood that the relational databases in general can be based on a standardized data model, the consistencies along with schema normalization must be prevailed. In their work they have also prioritized the standardized data model to permit the common users to transfer their skills.

Messina et al.[13] prioritized on NoSQL systems that support specific semi-structured data in their work. Accordingly, they have also focused on specific Internet-based applications that are equally important toward maintaining consistency on data. In their study they observed that mostly the NoSQL systems may emphasize on the flexibility of the data model to support data consistency in similar platforms.

Cruz et al.[14] discussed on syntactic heterogeneity implemented in several platforms in their work. They observed that mostly relational database systems can be provisioned with structured queries and can be implemented in several ways without any constraints.

Vathy-Fogarassy et al.[15] quantified the problems associated with data integration possessing the semantic heterogeneity. In fact, this type of heterogeneity can occur if the meaning of data may be dissimilar in several contexts of the similar domain. Also it has been observed that the semantic heterogeneity at the schema level can be occurred due to several factors especially by synonyms and the problem with synonyms may occur when the same entity may be named differently in different databases.

Noy et al.[16] during their study found that maintaining uniform access on data may not require data integration as well as migration of data. In fact, it may not create any issues related to integration application. They observed that the semantic heterogeneity of source databases should be more focused while performing uniform data access.

Sousa et al.[17] in their work prioritized the key factors toward obtaining solution for the database integration problems while linking the data semantic features and matching the semantic attributes with entities within the heterogeneous databases. In their work, they have also intended for data integration provisioning single system image for end users in a distributed environment.

Junior et al.[18] in their work discussed the heterogeneous database framework along with its independence and autonomy. They also focused on the design and implementation of information sharing between heterogeneous databases of network database system of many similar functionalities linked to specific data models.

Oleg Milder et al.[19] in their research focused on information sharing linked with the heterogeneous databases and shared among distributed and independent database systems. In their work they have also compared the workflow of the various systems prioritizing sharing the information.

Ashkan Tashk et al.[20] discussed on specific frames linked to distributed databases in their research considering the implementation mechanisms of the heterogeneous database systems. As the access linked to multiple heterogeneous database systems prevailed with transparency within the communication platforms, the database of every linked information system can be the independent entity with relative independence and autonomy. Also realization of data sharing of multiple heterogeneous databases is equally important especially in specific systems where the usage of web service middle layer is required to realize the connection of multiple heterogeneous data sources. Sometimes, due to specific reasons, it may not be desirable to share the important data even if for practical usages, but implementing particular mechanisms the web servers can connect the complete database systems. So practically while solving the problems within the middle tier as well as the heterogeneous data sources, the solution must be achieved provisioning the limited range query in Web Services. Several works on similar domain can be found in reference.[21]

15.3 STEP-WISE REPRESENTATION

The approach can be represented in the following steps:

Step 1: Consider n number of processes linked to n relations of linked query plans

Step 2: Prioritize the resources, r of the similar site and choose the resources

Step 3: Define the decision parameter u^n where $u^n = 1$ is assigned to task n allocated to resource r

Step 4: Prioritize the execution time of the process, t^n based on the resources.

Step 5: Determine the execution cost of process, e^n based on resources, r and query plans.

15.3.1 APPLICATION OF PARTICLE SWARM OPTIMIZATION

Generally, the particles associated with the swarm can move around the search space observing the optimum solution along with managing their positions as per inertia and social experiences. Initially, the particles are initialized from the solution space provisioned with the position and velocity. Then it is required to estimate the fitness parameter of each particle according to the fitness function. It is also required to update the individual best solution along with the global best solution. After updating the velocity as well as position of each particle, it is needed to prioritize again the inertial and cognition parametric constraints and continue the operation till optimality occurred.

15.3.2 ACCESSIBILITY OF HETEROGENEOUS DATA WITH PERFORMANCE

In general, the heterogeneous data can be provisioned with high variability of data types as well as with specific formats. Sometimes, the ambiguity on the same may be due to missing values along with high data redundancy. Accordingly, it may be difficult to integrate heterogeneous data to cope up with business intelligence. While analyzing practically, it may be observed that the heterogeneous data generated from cloud can be easily provisioned through various data acquisition devices acquiring multi-dimensional facilities. Being featured with large scaled data, the massive data acquisition equipments can be easily distributed, not only linked with the currently acquired data, but also the historical data within a certain time frame. Of course there is strong correlation between time and space. Each and every data acquisition device provisioned should be acquired with specific time stamp. Sometimes, within the datasets acquired by acquisition devices, very few amount of data may be valuable at specific time.

15.4 PROBLEM FORMULATION

A database server being a part of multi-point cluster probably can have the exclusive access toward large scaled data during processing. In fact, as shown in Figure 2, the data associated with the clusters can be provisioned locally. The databases within the servers can be segregated as distributed along with mutually partitioned within the clusters. The specific cluster responsible for the partition can easily cache the data provisioning the concurrency control over data. This mechanism may lead to higher performance as well as scalability as there may be fewer requirements toward coordinating the clusters.

TABLE 15.1 Deployed Database Servers with Response Time.

Sl. No.	Deployed database servers	Response time (m.sec.)
1	15	0.82
2	35	0.44
3	40	0.35
4	50	0.29

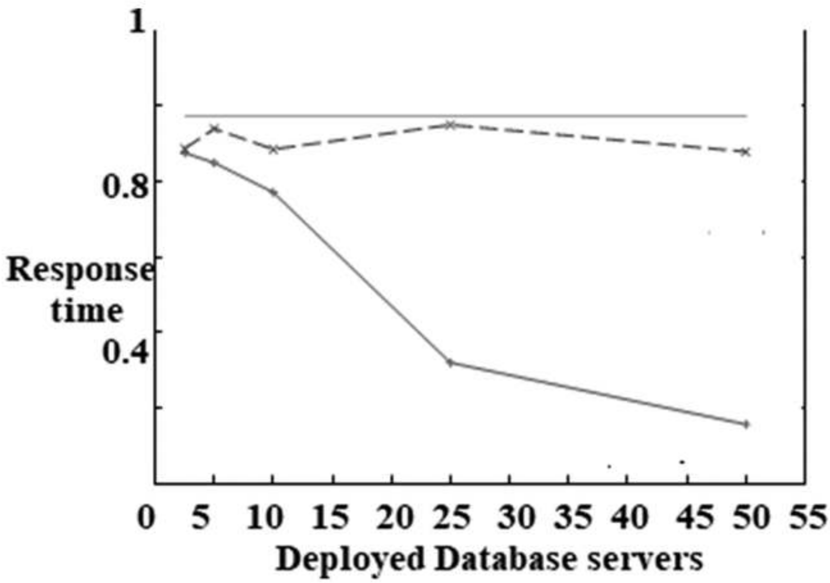


FIGURE 15.2 Deployed database servers with response time.

Somehow, concentrating on better partitioning may not be so easy, but it may require proper reorganization associated with the entire databases of concern. Every cluster in such a situation can access to the storage locations, i.e., shared global storage implementing the storage area networks as reflected in Figure 15.3. In that the shared data may respond to the existing system load by deploying required clusters. By the way, the data-sharing mechanisms can be complex toward implementation as it requires distributed concurrency control along with distributed cache coherence.

TABLE 15.2 Deployed Database Servers with Respect to Data Accessibility Rate.

Sl. No.	Deployed database servers	Data accessibility rate(%)
1	19	0.56
2	29	0.55
3	35	0.53
4	40	0.51

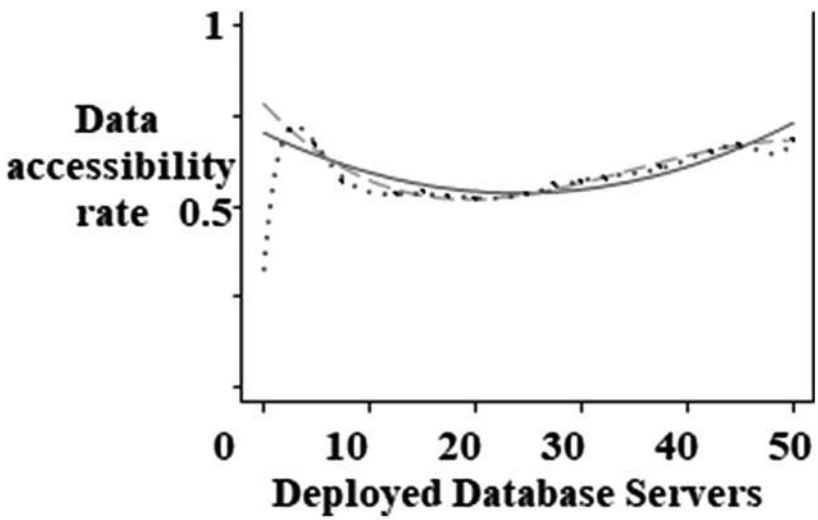


FIGURE 15.3 Deployed database servers with respect to data accessibility rate.

15.5 DISCUSSION AND FUTURE DIRECTION

It is understood that each and every metadata operation initiated by the database system can be treated as atomic operations and each may require locking and linking techniques in connection to the database kernels. Somehow, the distributed database systems may not permit the arbitrary client transactions. The complexity linked with query plans of distributed database systems should be resolved, and the systems should provide distributed concurrency control toward accessing the data and distributed cache coherence as well.

15.6 CONCLUSION

The concurrency control must be preserved while sharing the data through the deployed database servers and the locks on the shared data must be accessible to database servers. Any server within the distributed database domain may request suitable distributed lock to achieve better performance. Accordingly, achieving better performance may minimize the number of distributed database lock requests. As the distributed database systems have been provisioned with distributed caches, it is required to enhance the security measures and maintain the cache coherency.

KEYWORDS

- **large scaled database**
- **heterogeneity**
- **query plans**
- **containerization**
- **data accessibility**

REFERENCES

1. Shivle, S.; Siege, H. J. In *Mapping of Subtasks with Multiple Versions in a Heterogeneous Ad Hoc Grid Environment*, Third International Symposium on Parallel and distributed computing, 2004, pp 380– 387.
2. Naijing, H.; Yingying, W.; Liang, Z. In *Dynamic Optimization of Subquery Processing in Grid Database*, Third International Conference on Natural Computation, 2007.
3. Soliman, M. A.; Antova, L.; Raghavan, V.; El-Helw, A.; et al. In *Orca: A Modular Query Optimizer Architecture for Big Data*, ACM SIGMOD International Conference on Management of Data, 2014; pp 337–348.
4. Tang, X.; Zhang, Z.; Wang, M.; Wang, Y. In *Performance Evaluation of Light-Weighted Virtualization for PaaS in Clouds*, International Conference on Algorithms and Architectures for Parallel Processing, 2014; pp. 415–428.
5. Wilson, A.; Ghahramani, Z. Copula Processes. In *Advances in Neural Information Processing Systems*; 2010; pp 2460–2468.
6. Rajaraman, A.; Ullman, J, D. Data Mining. In *Mining of Massive Datasets*; Cambridge University: Press, 2011; pp 1–17. ISBN 9781139058452. http://dx.doi.org/10.1017/CBO9781139058452.002. Cambridge Books Online.

7. Wang, C.; Blei, D. M. In *Collaborative Topic Modeling for Recommending Scientific Articles,* Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; ACM, 2011; pp 448–456.
8. Gligorijevi´c, V.; Malod-Dognin, N.; Przßulj, N. Integrative Methods for Analysing Big Data in Precision Medicine. *Proteomics* **2016,** *16* (5), 741–758, DOI: 10.1002/pmic.201500396.
9. Dong, X.; Srivastava, D. Big Data Integration. *Proc. VLDB Endowment* **2013,** *6* (11), 1188–1189.
10. Abd El-Ghafar, R.; Gheith, M.; El-Bastawissy, A.; Nasr, E. In *Record Linkage Approaches in Big Data: A State Of Art Study*, 13th International Computer Engineering Conference (ICENCO), 2017.
11. Knight, B.; Veerman, E.; Moss, J. M.; Davis, M.; Rock, C. *Professional Microsoft SQL Server 2012 Integration Services*; John Wiley & Sons, Inc., 2012.
12. Garcia Molina, H.; Ullman, J. D.; Widom, J. *Database Systems: The Complete Book*; Pearson Prentice Hall, 2009.
13. Messina, A.; Storniolo, P.; Urso, A. In *Keep It Simple, Fast and Scalable: A Multi-Model NoSQL DBMS as an (eb)XML-overSOAP service*, 30th International Conference on Advanced Information Networking and Applications Workshops, Palermo, Italy, 2016.
14. Cruz, I.; Xiao, H. The Role of Ontologies in Data Integration. *J. Eng. Intel. Syst.* **2005,** *13* (12).
15. Vathy-Fogarassy, A.; Hugyák, T. Uniform Data Access Platform for SQL and NoSQL Database Systems. Inform. Syst. **2017,** *69*.
16. Noy, N. F. Semantic Integration: A Survey of Ontology-Based Approaches. *SIGMOD* **2004,** *33* (4), 65–70.
17. Sousa, M.; Pires, R.; Del-Moral-Hernandez, E. SOMprocessor: A High Throughput FPGA-Based Architecture for Implementing Self-Organizing Maps and Its Application to Video Processing. *Neural Netw.* **2020,** *125,* 349–362.
18. Junior, P. O.; Conte, S.; D 'Addona, D. M.; et al. An Improved Impedance-Based Damage Classification Using Self-Organizing Maps. *Procedia CIRP* **2020,** *88,* 330–334.
19. Milder, O.; Tarasov, D.; Tyagunov, A. The Artificial Neural Network Structure Selection Algorithm in the Direct Task of Spectral Reflection Prediction. *WSEAS Trans. Syst. Cont.* **2019,** *14* (9), 65–70.
20. Tashk, A.; Herp, J.; Nadimi, E. Automatic Segmentation of Colorectal Polyps based on a Novel and Innovative Convolutional Neural Network Approach, *WSEAS Trans. Syst. Cont.* **2019,** *14,* (47), 384–391.
21. Borah, S.; Mishra, S. K.; Mishra, B. K.; Balas, V. E.; Polkowski, Z. *Advances in Data Science and Management, Lecture Notes on Data Engineering and Communications Technologies*; Springer, 2022, DOI https://doi.org/10.1007/978-981-16-5685-9

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

CHAPTER 16

VISUALIZATION AND COMPARATIVE SIMULATION OF PATHFINDING, SEARCHING, AND SORTING ALGORITHMS

SUBHASH RATHOD, ABHINAV MISHRA, AKSHAY PATIL, ASHUTOSH DHANAWADE, and GAURAV DALVI

Department of Computer Engineering, MMIT Lohgaon, Maharashtra, India

ABSTRACT

This chapter aims to provide learners with a comprehensive understanding of algorithms by leveraging the power of comparative simulation and visualization. It is believed that visualization is a highly effective tool for improving learners' understanding of complex algorithms. By rendering algorithms in a visual format, learners can see how these algorithms behave in real-time and comprehend their inner workings. The proposed application enables learners to compare the performance of different algorithms and assess their strengths and weaknesses. They can observe how different algorithms handle different data sets, helping them to gain a better understanding of the advantages and limitations of each algorithm. The Visualizer e-Learning application supports a variety of algorithms, including Dijkstra's algorithm, DFS, BFS, Binary Search, and more. Learners can use this application to observe the differences between algorithms, such as how one algorithm may be more efficient than another under certain conditions. For instance, the comparative simulation feature can be used to compare the performance of Dijkstra's algorithm and A* algorithm in finding the shortest path between

Applied Soft Computing Techniques: Theoretical Principles and Practical Applications. Samarjeet Borah, Ratna Raja Kumar Jambi, Sharifah Sakinah Syed Ahmed, & Mahendra Prabhakar Deore (Eds.)
© 2025 Apple Academic Press, Inc. Co-published with CRC Press (Taylor & Francis)

two points on a map. By observing how each algorithm navigates the map, learners can identify the strengths and weaknesses of each algorithm and determine which one is better suited for a particular problem. Moreover, the comparative simulation feature can be used to compare different sorting algorithms, such as Quick Sort and Merge Sort. Learners can visualize how each algorithm performs when sorting a large data set, and they can compare the speed and efficiency of each algorithm. In conclusion, our Algorithm Visualizer e-Learning application is a powerful tool for learners of all levels who want to improve their understanding of algorithms. By leveraging the benefits of comparative simulation and visualization, we aim to make algorithmic concepts more accessible and enjoyable for learners.

16.1 INTRODUCTION

In the realm of computer science, algorithms are a fundamental area of study that is often referred to as "Design and Analysis of Algorithms." Proficiency in computational thinking and programming is required to excel in this subject. However, the complex nature of algorithms makes it a challenging subject to comprehend, even for experienced programmers. Traditional methods of teaching algorithms can be tedious and lack the necessary tools to facilitate efficient and effective learning. Hence, a visualization-centric approach to algorithm learning is vital to bridge this gap.

Our extensive research and in-depth discussions with leading experts have identified a pressing need for a robust, user-friendly application with dynamic visualization features. To this end, we are embarking on a project that will utilize various trending tools such as ReactJs, its framework, and libraries to design and deliver this much-needed tool. Our goal is to create an exceptional algorithm visualization platform that promotes efficient and effective learning.

The implementation of this software will drastically enhance students' algorithm comprehension capabilities by providing them with a dynamic visualization of the algorithms at work, enabling them to better understand the underlying concepts. By creating a visual representation of algorithms, we aim to revolutionize the way students learn and perceive algorithms, paving the way for a more comprehensive and effective mode of teaching.

Many prominent experts and educators support the algorithm visualization approach, stating that a picture is worth a thousand words. Extensive research has been conducted to support this claim, highlighting the significant positive impact of visualization in understanding complex concepts such as algorithms. As a result, our project aims to harness the power of visualization to provide learners with an effective and efficient means of comprehending algorithms.

In conclusion, we are committed to developing an algorithm visualization platform that will serve as a powerful tool for learners of all levels. By leveraging the benefits of visualization and comparative simulation, we are confident that our application will help to make algorithmic concepts more accessible, enjoyable, and comprehensible for all learners.

16.2 LITERATURE REVIEW

Prabhakar et al.[1] developed an algorithm visualizer using the Tkinter Module in Python. The visualizer was a desktop application that could be easily downloaded from their user-friendly website.

There have been several notable contributions in the field of algorithm visualization by researchers and developers. For instance, Yadav et al.[2] introduced a path-finding visualizer that employs A*, Dijkstra, and DFS algorithms to compute the shortest route between the start and end nodes. This tool provides a visual representation of the process and helps users develop an intuitive understanding of these algorithms.

Similarly, Goswami et al.[3] created an innovative algorithm visualization tool that features a variety of path-finding, sorting, and CPU scheduling algorithms. Their research study revealed that 60% of students preferred learning algorithms through visualization rather than relying on traditional textbooks. This study highlights the importance of interactive and visually engaging learning tools in helping students grasp complex algorithms more effectively.

Senger et al.[4] developed a visualizer to visualize BFS and Dijkstra's Algorithm and provided a comparison and analysis between the two algorithms. Nagaria et al.[5] emphasized the need for an interactive teaching tool that can assist students who learn best visually and provide a mechanism for them to submit their path-finding formulas and receive visual feedback.

Ghandge et al.[6] created AlgoAssist, a platform with integrated features that helps students improve their coding abilities, allows teachers to assess students' work, and emphasizes "Algorithm Visualization" to help students comprehend the flow and operations of algorithms. The platform integrates lab functionality into a single platform, making it easier for professors and students to use.

Jain[7] created a Sorting Algorithms Visualizer that can be used to visualize various sorting algorithms such as Insertion Sort, Bubble Sort, and Selection Sort. This application allows users to generate a random array for visualization and control the speed of sorting and size of the array.

Abedalrahim et al.[8] developed a desktop-based application using Java that involves the simulation of various sorting algorithms and can be used on a standalone PC only.

16.3 PROPOSED SYSTEM

The research paper proposes a system for visualizing and simulating searching and sorting algorithms, with a focus on comparative analysis. The system is designed to provide a user-friendly interface that allows users to compare the performance of different algorithms under various conditions.

The system includes a set of predefined data sets, each with a specific size and distribution, which can be used to evaluate the algorithms. Users can also input their own data sets to test the algorithms under custom conditions.

The system offers a variety of visualization tools to help users understand the behavior of each algorithm. For example, users can view the sorting process as it happens, with each step displayed on a graph.

The system supports a range of searching and sorting algorithms, including linear search, binary search, bubble sort, insertion sort, merge sort, quicksort, and heapsort. Users can choose which algorithms to compare and run them on the same data set to see how they perform relative to one another.

The proposed system is a dynamic web-based application that provides users with a platform to visualize and simulate searching and sorting algorithms. The system is made using ReactJs, HTML, CSS, and JavaScript, and includes a variety of visualization tools to help users understand the behavior of each algorithm.

The system will have the following features:

1. Auto-Generation and User Input: The system will allow users to auto-generate data or enter their own data for visualization.
2. Side-by-Side Comparison: The application allows users to run two algorithms simultaneously for a side-by-side comparison. This feature enables users to compare the performance of different algorithms under various conditions.
3. Algorithm Selection: The system will include the following searching algorithms:

   a. Linear Search
   b. Binary Search
   c. Depth First Search
   d. Breadth First Search

4. The following sorting algorithms will also be included:
   a. Bubble Sort
   b. Selection Sort
   c. Insertion Sort
   d. Quick Sort
   e. Heap Sort
   f. Merge Sort

5. Visualization Tools: The system will provide a variety of visualization tools to help users understand the behavior of each algorithm. Users can choose from several types of visualizations such as bar graphs, line graphs, or pie charts to display the results of each algorithm.
6. Performance Metrics: The system will provide detailed performance metrics for each algorithm. Users can view the number of comparisons, swaps, and other operations performed by each algorithm, as well as the time and space complexities of each algorithm.
7. Data Set and Algorithm Configuration Saving: The application will have a feature for users to save their data sets and algorithm configurations for future use. Users can save their results and compare them with previous runs to track their progress over time.
8. User Interface: The system will have a modern and intuitive user interface that is easy to navigate. Users will be able to choose from several color schemes and themes to customize the appearance of the application.

The proposed system will be an excellent tool for researchers, students, and developers to evaluate the performance of searching and sorting algorithms and gain a deeper understanding of their behavior. The application's user-friendly interface and visualization tools make it easy for users to compare algorithms and analyze their performance.

The proposed web application aims to provide a platform for comparative simulation and visualization of various pathfinding algorithms. Pathfinding is a common problem in computer science, and there are several algorithms available to solve it. However, it can be difficult to understand and compare these algorithms without visual aids. This web application aims to bridge that gap by providing users with interactive simulations and visualizations of different pathfinding algorithms.

The website will allow users to select from a variety of pathfinding algorithms, such as Dijkstra's algorithm, A* algorithm, and other variations. Users can create maps with different obstacles and terrain types to test the algorithms' performance in a 2D space. The application will simulate the pathfinding algorithm selected by the user on the map created by them. It will visualize the algorithm's progress and show the path found by the algorithm. Users can also compare the performance of different algorithms side by side on the same page.

Customization options will be available, allowing users to customize different parameters of the algorithms, such as the heuristic function, to see how it affects the algorithm's performance. The website will also provide learning resources to help users understand the algorithms' underlying principles, including tutorials and articles.

Overall, the website aims to provide a user-friendly platform for comparative simulation and visualization of pathfinding algorithms, using modern web development technologies and tools to ensure a responsive and interactive experience for users.

16.3.1 SYSTEM ARCHITECTURE

The visual representation of the major components of our application, depicted in Figure 16.1, provides an at-a-glance overview of the system's architecture, highlighting the core functionalities that have been carefully designed to provide an intuitive and user-friendly experience.

The three main components, Pathfinding Algorithms, Searching Algorithms, and Sorting Algorithms, have been thoughtfully designed to offer users an intuitive and engaging way to explore the fascinating world of algorithms.

The Pathfinding Algorithms component allows users to interact with a grid-based visualization where they can create a maze and set their starting and ending points. By selecting an algorithm, users can watch as the system employs a series of techniques to find the shortest path between these two points. The algorithm can be customized by adjusting the speed, adding barriers, or clearing the board. Additionally, the system provides an interactive feature where users can add a bomb to the maze as a barrier, making the visualization even more dynamic and exciting.

The Searching Algorithms component provides a comprehensive view of various algorithms used in searching, such as Linear Search and Binary

Search. The system generates blocks that represent an array's size, and each block contains a specific value. Users can interact with these blocks and observe how each algorithm works to find a particular value within the array. To make the visualization more interactive, the system also offers an option to drag the blocks to reorder them, providing a more hands-on experience.

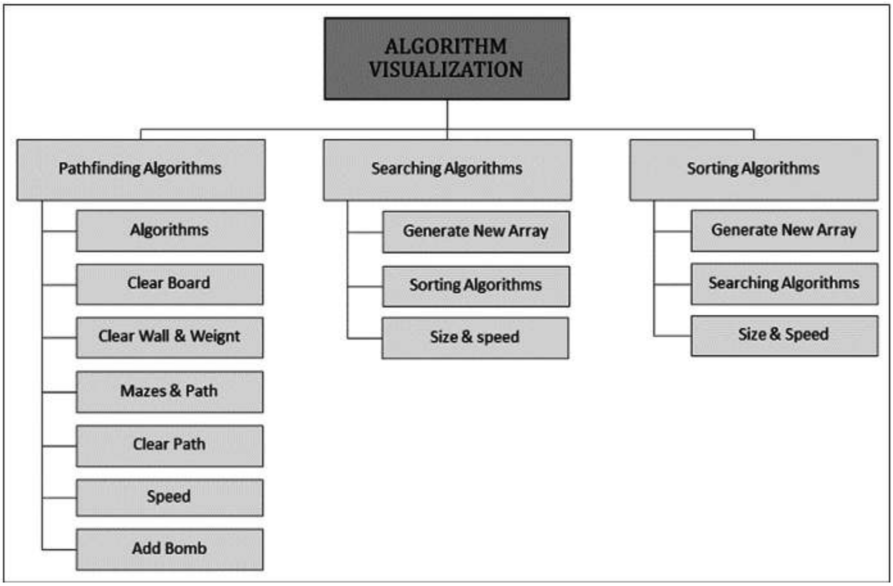


FIGURE 16.1 System architecture.

For DFS and BFS algorithms, the system generates circles to represent nodes and lines that represent edges of a graph or tree. Users can interact with these circles and lines by adding new nodes and connecting them with lines to form edges. This interactive feature allows users to gain a deeper understanding of how these algorithms work to traverse the graph or tree and discover all of its nodes.

Finally, the Sorting Algorithms component offers an excellent way for users to observe how different sorting algorithms work. The system generates bars that represent an array's size, and each bar's height indicates its value. Users can generate random arrays or insert their own arrays to visualize how each algorithm sorts the array. With this visualization, users can gain a deeper understanding of how algorithms such as Merge Sort, Quick Sort, and Bubble Sort work to arrange an array's elements in an ascending or descending order.

Overall, the system's components offer a diverse range of interactive features that allow users to learn and explore algorithms in an engaging and accessible way. By providing a user-friendly interface and interactive elements, users can gain a deeper understanding of how algorithms work and their applications in real-world scenarios.

16.3.2 ADDITIONAL FEATURES

Upon examining numerous similar projects, it was observed that certain characteristics are absent, which, if integrated, can augment the visualizer's allure, user-friendliness, and overall user experience. Our application endeavors to overcome these shortcomings by incorporating the following additional features:

Firstly, a superior and captivating User Interface (UI) is incorporated to enhance the overall User Experience (UX) of the application. Furthermore, the application is designed to be responsive, facilitating a seamless and excellent user experience on both desktop and mobile devices, thereby ensuring accessibility for all users.

To facilitate better visual comparisons, two algorithms can be executed simultaneously on the screen. Furthermore, the application analyzes algorithms' performance by conducting a comprehensive analysis of factors such as run times, complexity, memory consumption, and comparing them to other algorithms to determine their efficacy and performance.

The application also features a slider that enables users to adjust the visualization speed, thus facilitating better comprehension of the algorithm's operations. Furthermore, pause and play controls of the visualization are integrated into the application, enabling users to analyze each step of the algorithm in detail.

Lastly, the application includes a dark mode option, which not only adds aesthetic value but also reduces eye strain and enhances visual appeal in low-light conditions.

16.4 RESULT

The implementation of our project has yielded significant results, which we are excited to share with you. Following are the results achieved after implementation of the project.

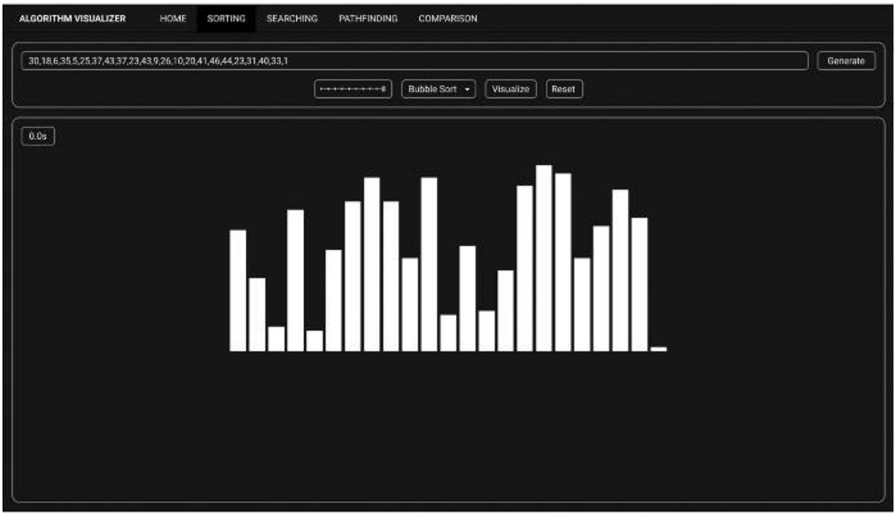


FIGURE 16.2 Sorting visualization page.

The Sorting Visualization Page, as shown in Figure16.2, comprises the following elements:

- Array Text field that displays the Array to be visualized.
- Generate Button, which generates a random array of random length when clicked.
- Speed Slider, used to set the speed of Visualization.
- Algorithm Drop Down, used to select the Algorithm to be Visualized.
- Visualize Button, which initiates the visualization.
- Reset Button, used to reset the visualization.
- Timer, that counts the time required for visualization to be completed.
- Visualization Area, comprising Bars that represent the values of the array.

In Figure 16.3, we can see the Sorting Visualization Comparison Page, which includes the following elements:

- Array Text field: shows the array to be visualized.
- Generate Button: generates a random array of random length when clicked.
- Speed Slider: sets the speed of visualization.
- Algorithm Drop Down: selects the type of algorithm to be visualized.
- Visualize Button: starts the visualization.
- Reset Button: resets the visualization.

- Two Timers: count the time required for visualization of respective algorithms to be completed.
- Two Drop Downs: select two algorithms of the selected type to be visualized and compared.
- Two Visualization Areas: consist of bars that represent the values of the array.

FIGURE 16.3 Sorting visualization comparison page.

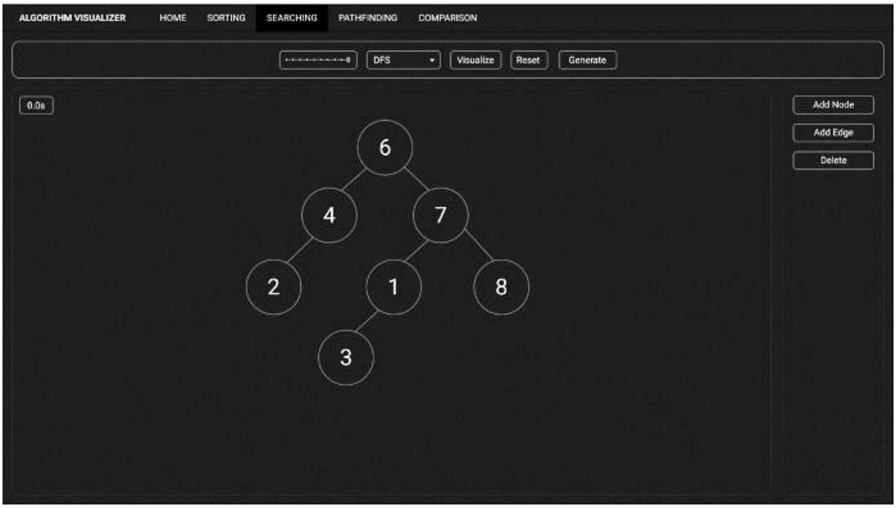


FIGURE 16.4 Searching visualization page.

The Searching Visualization Page, as shown in Figure 16.4, comprises the following elements:

- Add Node Button to add a new Node.
- Add Edge Button to add a new Edge.
- Delete Button to delete an edge or a node.
- Generate Button, which generates a random tree when clicked.
- Speed Slider, used to set the speed of Visualization.
- Algorithm Drop Down, used to select the Algorithm to be Visualized.
- Visualize Button, which initiates the visualization.
- Reset Button, used to reset the visualization.
- Timer, that counts the time required for visualization to be completed.
- Visualization Area, consisting of Nodes and Edges that represent a Tree.

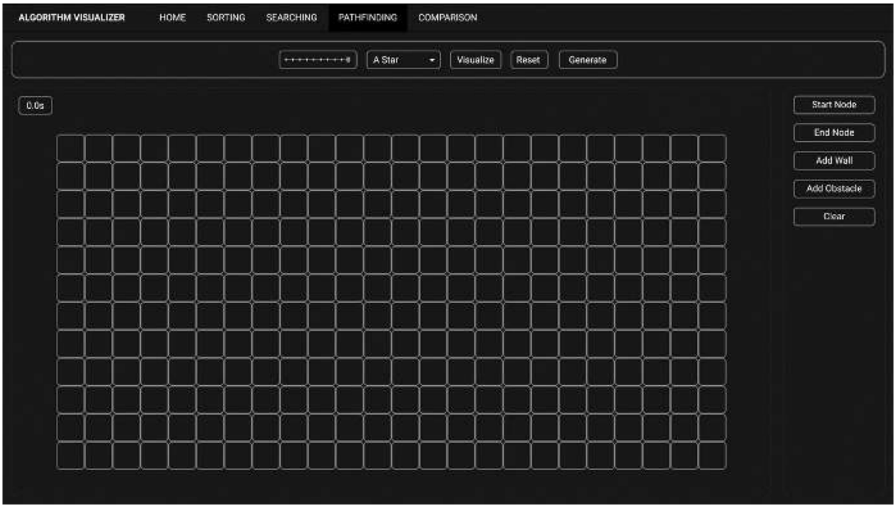


FIGURE 16.5 Pathfinding visualization page.

The Pathfinding Visualization Page, as shown in Figure 16.5, comprises the following elements:

- Start Node Button to set the start node in the Grid.
- End Node Button to set the end node in the Grid.
- Add Wall Button to add walls to the Grid.
- Add Obstacle Button to add obstacles in the Grid.
- Generate Button, which generates maze when clicked.
- Speed Slider, used to set the speed of Visualization.

- Algorithm Drop Down, used to select the Algorithm to be Visualized.
- Visualize Button, which initiates the visualization.
- Reset Button, used to reset the visualization.
- Timer, that counts the time required for visualization to be completed.
- Visualization Area, comprising of a Grid which shows a maze.

Our project's advanced visualization features and algorithmic capabilities enable users to easily analyze and compare various pathfinding, searching, and sorting algorithms. With the ability to visualize and compare algorithms, users can gain a deeper understanding of their functions and operations, enabling them to make more informed decisions when selecting algorithms for specific tasks. This functionality is especially useful in the field of educational technology in higher education, where algorithmic understanding and proficiency are essential for success.

16.5 CONCLUSION

In conclusion, the development of e-Learning tools has revolutionized the education sector, providing learners with a flexible and accessible mode of learning. Our web application is a testament to the importance of e-Learning, as it simplifies complex algorithms' visualization, making it easier to understand their operation. The application caters to users of all levels and offers numerous features that enhance the user experience, making it a valuable tool for anyone interested in learning about algorithms.

As we look to the future, we see vast potential for expanding the application's scope and capabilities. One area of development we envision is the creation of a mobile application that runs offline, allowing users to access the application's functionalities anytime, anywhere. The mobile application will be designed to include all the features of the web application, enabling users to visualize and compare the performance of algorithms on the go.

Furthermore, the mobile application can include more complex algorithms based on algorithmic strategies such as Divide and Conquer, Greedy Strategy, and others. These algorithms are essential in developing applications that can solve complex problems and including them in our application will provide users with a deeper understanding of algorithmic strategies.

Overall, the development of e-Learning tools such as our web application has created a new paradigm in education, enabling learners to access information and learn at their own pace. As we continue to expand and develop the application, we aim to provide users with a comprehensive and effective tool for learning algorithms, paving the way for a more efficient and enjoyable learning experience.

KEYWORDS

- **algorithm visualization**
- **searching**
- **sorting**
- **pathfinding**
- **digital learning**
- **heuristic algorithms**
- **educational technology**

REFERENCES

1. Prabhakar, G.; Gaur, S.; Deshwal, L.; Jain, P. In *Analysis of Algorithm Visualizer to Enhance Academic Learning*, 2022 2nd International Conference on Innovative Practices in Technology and Management (ICIPTM), 2022; pp 279–282, DOI: 10.1109/ICIPTM54933.2022.9753906.
2. Yadav, N.; Dhameja, K.; Chaubey, P. In *Path Finding Visualizer Application for Shortest Path Algorithm*, 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N), 2021; pp 1669–1672, DOI:10.1109/ICAC3N53548.2021.9725716.
3. Goswami, B.; Dhar, A.; Gupta, A.; Gupta, A. In *Algorithm Visualizer: Its features and working*, 2021 IEEE 8th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON), 2021; pp 1–5, DOI: 10.1109/UPCON 52273.2021.9667586.
4. Nishant, S. Pathfinder Visualizer of Shortest Paths Algorithms. *Int. J. Modern Trends Sci. Technol*. **2021,** *6,* 479–483. DOI: 10.46501/IJMTST061293.
5. Nagaria, B.; Evans, B. C.; Mann, A.; Arzoky, M. In *Using an Instant Visual and Text Based Feedback Tool to Teach Path Finding Algorithms: A Concept*, 2021 Third International Workshop on Software Engineering Education for the Next Generation (SEENG), 2021; pp 11–15, DOI: 10.1109/SEENG53126.2021.00009.
6. Ghandge, A. B.; Udhane, B. P.; Yadav, H. R.; Thakare, P. S.; Kottawar, V. G.; Deshmukh, P. B. In *AlgoAssist: Algorithm Visualizer and Coding Platform for Remote Classroom Learning*, 2021 5th International Conference on Computer, Communication and Signal Processing (ICCCSP), 2021; pp 1–6, DOI: 10.1109/ICCCSP52374.2021.9465503.

7. Jain. In *Realizing Algorithms Using GUI*, 2021 10th International Conference on System Modeling & Advancement in Research Trends (SMART), 2021; pp 451–455, DOI: 10.1109/SMART52563.2021.9676269.
8. Jamil, A.; Jamil, A.; Maslan, Z.; Oliinyk, A.; Azwan, A.; Rahman, A.; Baharudin, Z. A.; Fakulti, T.; Kejuruteraan, E.; Dan, E.; Universiti, T.; Malaysia, M.; Hang, J.; Durian, T. The Development of System for Algorithms Visualization Using Simjava. *J. Eng. Appl. Sci*. **2020,** *15,* 3024–3034.

CHAPTER 17

EDUCATIONAL WEBSITE WITH SEARCH ENGINE OPTIMIZATION

MRUNAL JAGTAP[1], YOGESH DONGARE[2], VIKAS CHAVAN[1], ASHISH BHISE[1], and RADHIKA KOKATE[1]

[1]*Department of Computer Engineering, MMIT, Lohgaon, Pune, India*

[2]*Faculty of Computer Engineering, Savitribai Phule Pune University, India*

ABSTRACT

An educational website aims to improve practicality and provide a comprehensive software package that fulfills users' needs. It allows users to store and access their valuable information easily and efficiently for an extended period. The software and hardware used are readily available and user-friendly. By implementing such an educational website, users can access services more securely and focus on other activities rather than record keeping. This organization will benefit from streamlined and efficient resource utilization and automated record-keeping. Our website also features an AI-based chatbot that employs TensorFlow and NLP models. To ensure security, we have incorporated the B-crypt algorithmic rule. Additionally, we have used SEO techniques to optimize our website's performance, allowing for seamless operation.

17.1 INTRODUCTION

In the current era, websites are the primary means of communication between organizations and their users or visitors. Websites are essential tools

Applied Soft Computing Techniques: Theoretical Principles and Practical Applications. Samarjeet Borah, Ratna Raja Kumar Jambi, Sharifah Sakinah Syed Ahmed, & Mahendra Prabhakar Deore (Eds.)
© 2025 Apple Academic Press, Inc. Co-published with CRC Press (Taylor & Francis)

for sharing information and are critical to an organization's communication success. A website's success is determined by three critical factors: accessibility, SEO, and quality, with usability being the most essential. Usability is a comprehensive concept that goes beyond user-friendliness, according to the ISO 9241-11:2018(en) standard. It refers to the efficacy, efficiency, and satisfaction of a system, product, or service's capacity to be used by defined users to achieve stated goals in each context of usage. The correctness of a system evaluation is determined by the usability evaluation techniques and tools, although the findings produced by various techniques and tools might vary. For reviewing and evaluating websites, there are several offline and online tools and strategies available. University websites are becoming more and more common because of the development of the Internet and domestic campus networks. These websites offer quicker, easier, more accurate, and direct access to resources and information. The IP address or domain name of each university website, the login process, or where to access published material are all frequently forgotten by users. As a result, individuals now get information online mostly through search engines. Search engine optimization (SEO) should be taken into consideration to raise the value of a university website's brand and enhance the number of people who can quickly and easily obtain the necessary information. SEO gives a website the chance to top search engine result pages for particular keywords. The website of the school of continuing education at Ocean University of China will be used as an example in this article to examine website optimization techniques for increasing exposure. In general, this research study attempts to present a thorough examination of a suggested web platform that seeks to improve dialogue, cooperation, and information sharing among college students. Technology is used to advance the planned platform.

17.2 LITERATURE REVIEW

Akoglu shared a fascinating case study on the University of Istanbul's Architectural Department website, where he introduced a cutting-edge usability evaluation tool. The data collection process for the study involved two different methods: laboratory-based tests and remote usability tests conducted via the Internet.[1] The results obtained from the collected data II highlighted the efficiency of the online tool in acquiring data and producing accurate output. The testing tools offer several advantages, including the ability for administrators to access user databases and their responses, which can be opened using Microsoft Access and Microsoft Excel software. This advanced feature allows for seamless data analysis, which is crucial in designing effective user interfaces.

Mustafa and his colleagues.[2] analyzed nine Jordanian websites, utilizing both the Toolbox and Webpage Analyzer tools. These tools were found to be highly effective in evaluating the usability of the websites, a critical aspect of website design. In a similar vein, Carta and his team created a new tool for evaluating website usability, which operates on the principles of customer data and JavaScript events. The tool attributes a proxy-based design that allows for precise performance evaluation and enables the evaluator to classify activities based on averages and optimization levels. This advanced tool promises to be a valuable asset in designing user-friendly websites.

Costa and his team examined the websites of many African nations using the WebQual programme.[3] Their research' findings allowed for a comparison of the web sites' architectural complexity and accessibility level. The analysis found a large number of accessibility problems, which prompted the authors to suggest fixes for enhancing website accessibility. In the meanwhile, Site Analyzer and Qualidator, two automated tools, are used by Sukhpuneet and his colleagues to analyze websites based on numerous criteria at Punjabi institutions in India. The results of the study showed that the chosen websites were among the best, emphasizing the value of good website design in boosting usability.

The chat interface plays a crucial role in collecting user queries and serving as a communication link between the user and the system.[3] It can be integrated with the machine learning layer and packaged as a smart app. To accomplish this, a Python web framework can be utilized to combine web applications with the AI technology of the NLP framework. When a user poses a question, the chatbot will use the knowledge database at its disposal to provide an answer. If the conversation introduces a concept outside the bot's programmed understanding, it will transfer the query to a human operator. This interaction provides an opportunity for the chatbot to learn and enhance its knowledge base, improving its effectiveness over time. Thus, the chatbot's capabilities and importance will continue to expand as it acquires new information and improves its functionality.

17.3 SIGNIFICANCE OF THE STUDY

In the current digital era, the Internet's extensive use and the quickly growing domestic campus networks have resulted in the creation of portals for the majority of schools and universities. People may quickly, easily, accurately, and directly access information and resources through university websites. Many people, however, find it difficult to recall the domain name or IP address of each university website, so they turn to search engine websites to locate the information they want. The preferred way for users to look for information online is to utilize search engines to look for information and relevant websites. Utilizing search engine optimization (SEO) techniques can help university websites build their brand value and make sure that consumers can swiftly and easily find the information they need. We must employ search engine optimization (SEO) techniques to raise the perceived value of university websites and guarantee that visitors can swiftly and easily find the information they want. To increase a website's exposure for particular keywords on search engines and consequently its rating and site traffic, SEO is used. In this essay, we will explore website optimization tactics that can raise website rankings and efficiently provide information to target audiences by using the website of the School of Continuing Education at the Ocean University of China as an example.

17.4 EXPERIMENTAL METHODS AND MATERIALS

17.4.1 TECHNOLOGIES USED

A. React: The JavaScript library React is used to create UI and its components. React gives programmers the ability to effectively manage the application's state and generate reusable user-interface components. It is used to design user interfaces that are dynamic and interactive as well as the front end of websites. The skeleton of a React application is a component. A component is a standalone piece of code that specifies a particular user interface element. Every component can have a changing state that can be updated and altered independently of other components. Many tiny components that make up React apps are often reused and integrated in various ways to produce sophisticated user interfaces.

   Other components of React include:

   - Virtual DOM: React uses a virtual representation of the Document Object Model to manage changes to the user interface. The virtual DOM is a light-weighted copy of the actual DOM that is updated and modified by React, rather than by the browser. This allows React to update the user interface quickly and efficiently, without needing to re-render the entire page.

- Props: By sending data from a parent component to the child component, they allow a child component to render data dynamically. The child component cannot modify props since they are read-only.
- State: State is a way to manage data within a component. The state is used to store data that can change over time, such as user input or application state. When the state changes, React automatically re-renders the component to reflect the updated state.
- Lifecycle methods: At various stages of a component's life cycle, React offers a collection of methods that are invoked. These methods let developers do specific actions, such as altering the state or obtaining data from an API, by hooking into particular moments in the component's life- cycle.
- Overall, these components and features of React provide a powerful toolkit for building dynamic and interactive user interfaces. They allow developers to create reusable and composable components, manage the application state efficiently, and respond to user input and changes in the application state.

B. Node.js: A JavaScript runtime called Node.js is based on the V8 JavaScript engine. It allows programmers to build scalable and rapid network applications on the server side using JavaScript. The backend of the website is built using it, in addition to managing server-side features like user authentication and database queries.

C. MongoDB: 2. The NoSQL document database MongoDB stores data in documents that resemble JSON and have a dynamic structure. Without first defining a schema, it enables developers to save and retrieve data fast and simply. Data from the website, including user accounts and content, is stored there so that it may be promptly and effectively retrieved.

D. Bootstrap: As mentioned earlier, Bootstrap is a front-end UI framework that provides prebuilt HTML, CSS, and JavaScript components for creating responsive and mobile-first websites. It is often used in conjunction with React to create a visually appealing and responsive user interface.

E. jQuery: The traversal and manipulation of HTML documents, event management, and animation are made simpler by the rapid, compact, and feature-rich JavaScript package known as jQuery. To generate interactive and dynamic UI elements, it is frequently used in combination with React.

F. Google Tag Manager: Google Tag Manager is a free tag management solution that allows website owners to manage various tags and track codes on their websites without editing the code. It is used to implement third-party tracking and analytics tools like Google Analytics and Google Remarketing Tag, without needing to edit the website's source code.
G. SaaS: SaaS (Software as a Service) is a software delivery model where the software is hosted on a cloud infrastructure and accessed through a web browser. It is used to host and deliver websites as a service to users, without needing to set up and maintain servers.
H. Google Analytics 4: Google Analytics 4 is a free web analytics service that provides insights into website traffic and user behavior. It is used to measure website's performance and user engagement and to optimize the website for a better user experience.
I. Delivery: delivery is a free, open-source CDN (Content Delivery Network) that delivers JavaScript libraries and other web assets faster by catching them on multiple servers around the world. It is used to improve the website's performance and speed by delivering static assets from the closest server to the user's location.
J. Netlify: The cloud-based platform Netlify offers server-less backend services for contemporary web applications, continuous deployment, and web hosting. It has functions like automated builds, asset optimization, and CDN caching and is used to launch and host the website.
K. Google Font API: Google Font API is a free web service that provides a collection of fonts that can be used on websites. It is used to improve the website's typography and visual appeal by offering a wide range of font styles and weights that are optimized for web use.

Overall, these technologies work together to create a modern, scalable, and efficient web application that provides an optimal user experience. React.js and Node.js are used to create a fast and responsive front-end and back-end, respectively, while MongoDB provides a flexible and easy-to-use database. Other technologies, such as Bootstrap, jQuery, and Google Tag Manager, are used to improve the website's performance, design, and tracking capabilities.

17.4.2 SYSTEM ARCHITECTURE

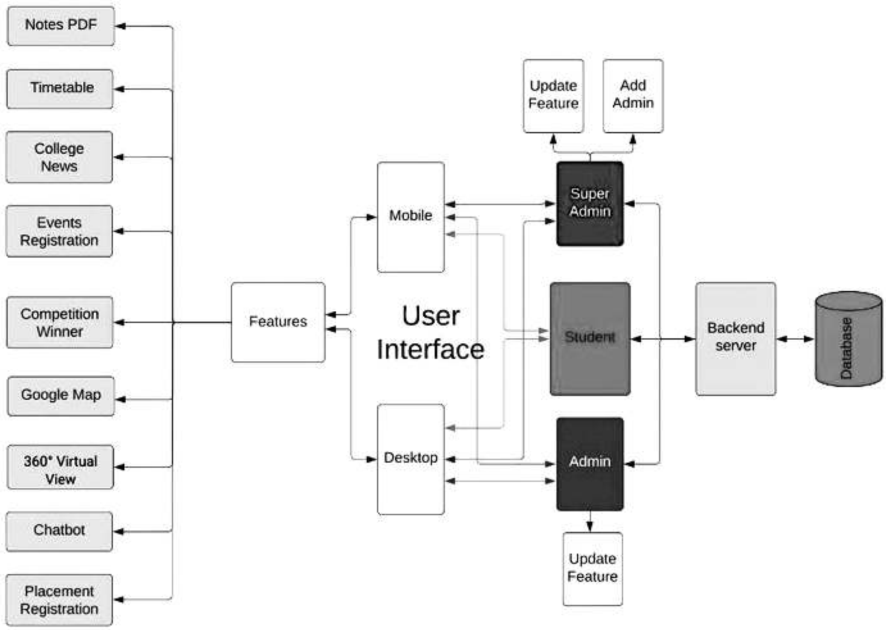


FIGURE 17.1 Architecture of the system.

17.4.3 SIGNIFICANCE OF ARCHITECTURE

17.4.3.1 DESCRIPTION OF THE SYSTEM ARCHITECTURE

Our architecture diagram is mainly divided into two parts, i.e., front end and back end. Back end is divided into a two-part server and database. Our project backed is developed by using JavaScript libraries and framework. We used the MongoDB database to store the data. The reason behind choosing the nonrelational database is we mainly deal with the nonrelational data. We created the server by using the node.js and express.js. The servers require handling the request's response. The main motive to choose the node. js because of the node.js is the trending technology for back end. In the project we have hierarchical access to the users. For the hierarchical access needs of authentication, we used the Json web token for the token for the hierarchical access or authentication. There are three hierarchical levels of access in our project, that is, a super admin, admin, and general users, in our case we assume as student. The Super admin has a high level of access. The super admin can add an admin, give them access, and change the website or database data. It will be capable of adding update features to the website. The super admin can be added and removed the admin. Admin can update the feature and remove the data from the specific features. Like admin can add notes and study material and remove that data from database. Admin is not able to add the admin and remove the admin. The last access is for the general user or student. The end user can only view the web pages through desktop or mobile. The web page has a responsive design for all devices. The end user can view feature of website, i.e., notes pdf, timetable, collage news, event registration, competition winner, google map, 360° virtual view, chat bot, placement registration, etc.

17.5 ANALYSIS OF THE WEBSITE

The website was created without any prior analysis or SEO considerations. However, after purchasing a domain, an SEO score and analysis were conducted which revealed the website's current SEO status. Without specific details about the analysis, it is difficult to provide a comprehensive report on the website's SEO strengths and weaknesses. However, based on typical SEO factors, some potential areas for improvement could include optimizing on-page elements such as titles, meta descriptions, and headers, as well as improving the website's content with high-quality, relevant information. Additionally, building backlinks from authoritative websites and implementing a solid internal linking strategy could help boost the website's overall SEO performance.

In general, the before and after SEO report of the website likely showed a significant difference in terms of its SEO performance. Before implementing SEO, the website may have been difficult to find and may not have been ranking for relevant search terms. After conducting an SEO analysis and implementing best practices, the website likely saw improvements in its search engine visibility, organic traffic, and overall user engagement. However, without specific data and metrics, it is difficult to provide a more detailed analysis of the website's current status and potential areas for improvement.

Based on the analysis of the Educational Website, we have identified several areas for improvement to enhance its SEO performance. We have implemented various techniques, such as adding a robots.txt file, changing the heading, conducting keyword research, improving page speed, optimizing for mobile devices, and implementing a strong internal linking strategy.

In addition, we have also analyzed website analytics regularly to identify potential areas for improvement and make data-driven decisions. These efforts have resulted in improved search engine visibility, organic traffic, and overall user engagement. However, we recommend continuing to monitor and improve the website's SEO performance by regularly analyzing website analytics, updating content with relevant keywords, and building high-quality backlinks from authoritative websites.

17.6 RESULTS/OUTPUT

17.6.1 LANDING PAGE

Figure 17.2 redirects the user to the homepage of the website. It serves as the starting point for users who want to explore different sections or return to the main page from anywhere on the site. It provides easy navigation and acts as a central hub for accessing other sections.

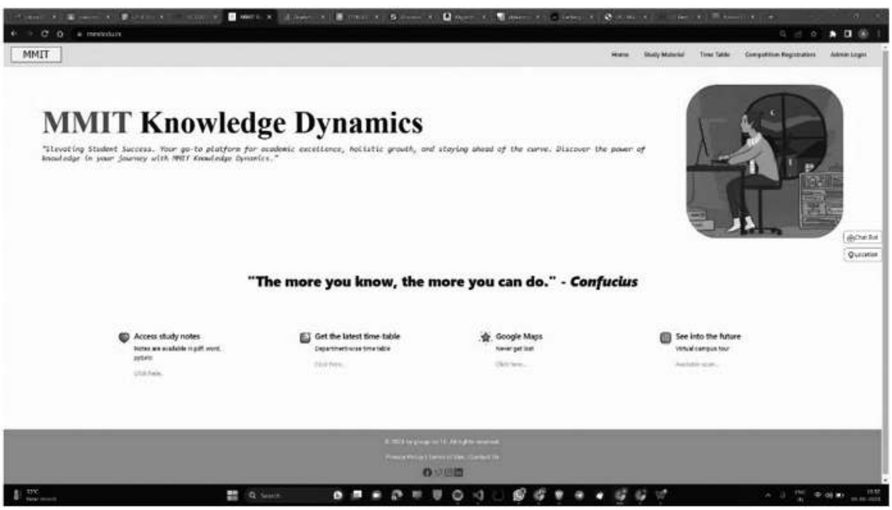


FIGURE 17.2 Landing page.

17.6.2 ANALYSIS OF THE WEBSITE

In Figure 17.3, the first graph provides insights into user engagement and website performance. It indicates that there have been a total of 76 users who

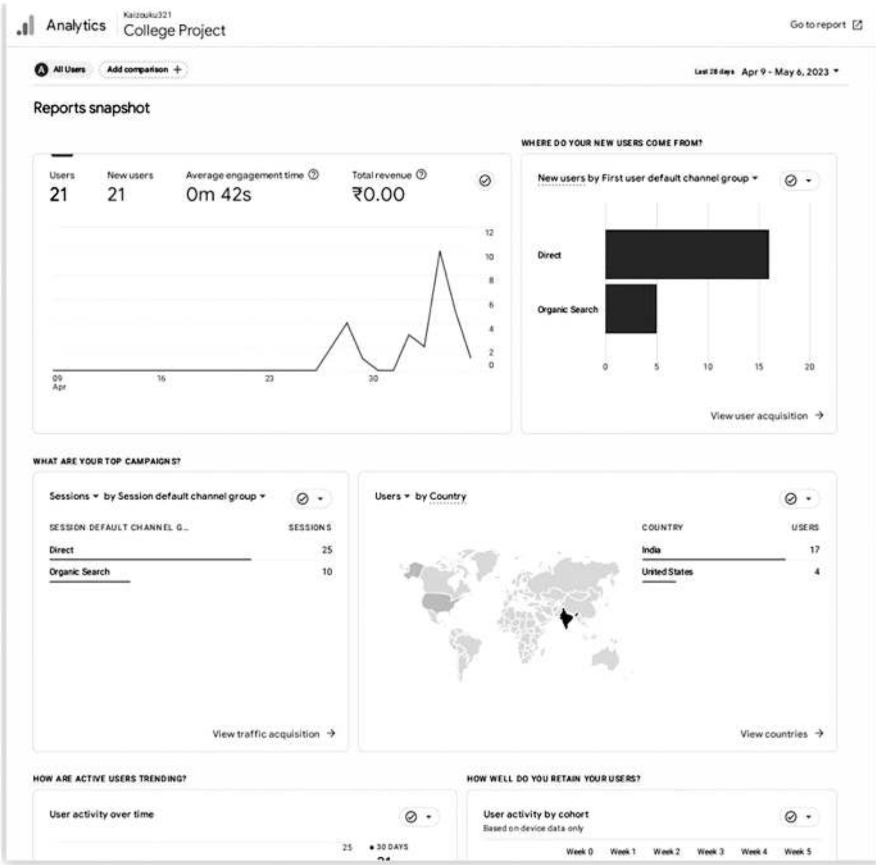


FIGURE 17.3 Webpage analysis.

have visited the website on the domain www.mmitedu.in. The graph displays the average engagement time of these users, ranging from 1 min to 2 min. However, the revenue generated by the website is zero since there are no clickable ads present on the site. This suggests that the website may rely on other means for monetization or is focused on providing information rather than generating revenue. Moving on to the second graph focuses on the users who accessed the website through direct search or organic search via Google. The data reveals that some users directly searched for "mmit.edu" and landed on the website, while others discovered the website through organic search results where the search engine suggested the website as part of their search query. This indicates that the website has been indexed by Google and appears in relevant search results, attracting users who are actively seeking information related to "mmit.edu." The analytics of the website, as depicted in the third graph, offer valuable insights into user engagement and session statistics. The index column presents a total of 160 sessions, which represents the entirety of user interactions on the website. Out of these sessions, 75% are categorized as engaged sessions, suggesting that users actively interacted with the website's content. On average, each session lasted for 1 min and 39 s, indicating a relatively short duration of user engagement. The metric of engaged sessions per user averages 1.17, suggesting that users tend to have multiple engaged sessions during their visits. Additionally, there was an average of 19.63 events per session, signifying user interaction with various features or elements on the website. The engagement rate of 58.67% implies a moderate level of user involvement. However, despite the significant user activity, there were no conversions or revenue generated during the analyzed period, indicating a need for further analysis and optimization. Lastly, the fourth graph provides insights into the geographical distribution of users who visited the website. The data is segmented by country, indicating the number of users from different regions. The analysis reveals that approximately 47 users were from India, while 29 users were from the United States and 1 from Spain. This information highlights the website's reach and global audience, showcasing its popularity among users from these particular countries. Understanding the geographical distribution of users can help tailor content or marketing strategies to cater to specific regions or target audiences. Overall, website analysis using Google Analytics offers valuable information about user engagement, sources of traffic, and user demographics. These insights can assist in improving website performance, enhancing user experience, and optimizing strategies to attract and retain visitors.

17.7 CONCLUSION

In today's digital age, websites have become the primary means of communication between organizations and their users or visitors. They serve as a powerful tool for sharing information and engaging with audiences. However, for a website to be successful, it must meet certain criteria, namely accessibility, search engine optimization (SEO), and usability. Out of these three factors, usability is the most critical success factor. A website's usability determines its effectiveness in engaging and retaining visitors, and thus plays a pivotal role in its overall success. In the case of educational institutes, a web application can aid in accurately, quickly, and reliably performing regular activities such as accessing study material, placement information, and timetables. This, in turn, can enhance the quality of work at educational institutes by improving efficiency and reducing errors. To further enhance the security and accessibility of websites, several innovative technologies such as AI, Android applications, and blockchain can be employed. AI can provide a more personalized and seamless user experience, while creating Android applications can help increase accessibility on mobile devices. Additionally, the blockchain can provide a secure and decentralized platform for website transactions and data storage. Overall, in today's digital landscape, the success of a website depends on several factors, including accessibility, SEO, and usability. By utilizing cutting-edge technologies and tools, organizations can create more secure, accessible, and user-friendly websites that can better engage and serve their audience.

KEYWORDS

- **search engine optimization**
- **accessibility**
- **usability**
- **educational website**
- **website optimization**
- **AI chat-bot**

REFERENCES

1. Quality Analysis of An Educational Website in Terms of Search Engine Optimization (SEO) Analysis Subject: The Website of the Faculty of Electrical Engineering (www. electro. pub. ro) IEEE, 2015.
2. Brajnik, G. In *Automatic Web Usability Evaluation: What Needs to be Done*, Proc. Human Factors and the Web, 6th Conference, June 2000.
3. Mustafa, S. H.; Al-Zoua'bi, L. F.; In *Usability of the Academic Websites of Jordan's Universities An Evaluation Study*, Proceedings of 9th International Arab Conference for Information Technology, December 2008, pp 31–40.
4. Costa, N. F.; Neves, S.; Duarte, C.; Hijon-Neria, R.; Carrico, L. In *Web Accessibility in Africa: A Study of Three African domains*, HUMAN-COMPUTER Interaction-INTERACT2013, Springer Berlin Heidelberg; pp 331–338.

5. Carta, T.; Paterno, F.; De Santana, V. F. In *Web Usability Probe: A Tool for Supporting remote Usability Evaluation of Web Sites*, Human-Computer Interaction-INTERACT, Springer Berlin Heidelberg, 2011; pp 349–357.
6. Tao, Z.; Li, L. Based on Linked Website Search Engine Optimization Strategy. *J. Hubei University Technol.* **2010,** *25* (5), 61–63.
7. Yunfeng, Mo. The Enterprise Website Development Strategy Research for the Search engine Optimization. *Zhejiang University Technol.* **2009,** 35–36.
8. Sukhpuneet, K.; Kulwant, K.; Parminder, K. Analysis of Website Usability Evaluation Methods **2016,** *978* (9), 1043–1046.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

PART IV
IMAGE PROCESSING

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

CHAPTER 18

A COMPARATIVE ANALYSIS ON SUPPORT VECTOR MACHINE, K-NEAREST NEIGHBORS, NAIVE BAYES, AND DECISION TREE CLASSIFIERS APPLIED FOR HUMAN FACE RECOGNITION

MD RAHMATUL KABIR RASEL, HASMOT ALI,
MD. MAHFUJUR RAHMAN, and MOHAMMAD JAHANGIR ALAM

Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh

ABSTRACT

Image processing is one of the most significant areas of computer science. Every day we have to deal with face recognition for various purposes such as security authentication systems, identification, matching images, etc. This research paper proposed a system model for recognizing human face using the combination of Principal Component Analysis (PCA) and Support Vector Machine (SVM). PCA is used to process data and extract useful features from images including reduction of dimensionality, increasing interpretability, and diminishing information loss of the given image. This model provided a comparative analysis based on the results of recognition accuracy achieved by (SVM), k-Nearest Neighbors (k-NN), Naive Bayes (NB), and Decision Tree (DT) classifier. The aim of this paper is to provide the most efficient and accurate method to recognize human face including data preprocessing using a small dataset and lower-end device. Various soft computing-based

Applied Soft Computing Techniques: Theoretical Principles and Practical Applications. Samarjeet Borah, Ratna Raja Kumar Jambi, Sharifah Sakinah Syed Ahmed, & Mahendra Prabhakar Deore (Eds.)
© 2025 Apple Academic Press, Inc. Co-published with CRC Press (Taylor & Francis)

approaches have been considered in this study. The experimental result of this method reveals that the way we processed our data, the SVM technique displays robust performance and increases the efficiency of the human face recognition technique with the utmost level of accuracy than the other three.

18.1 INTRODUCTION

Face recognition has added a new dimension to modern technology because of the increasing demand for machines having computer vision features. Applications such as the field of information indexing, network security, and video technology convenience from human face recognition techniques because "humans" are the focus of attention in a lot of videos.[1] Therefore, the study of face recognition is also developing day by day, and new face recognition techniques, and algorithms are being invented. Basically, there are two different ways to recognize faces. One of them is training a neural network model in which the accuracy of detecting face is highly influenced by the number of input data because a neural network model requires billions of input data to acquire the expected outcome, but the other is the One-shot Learning technique which supposed to learn from the statistics of object category from a very few numbers of input data.[2]

A comparative analysis between different face recognition algorithms will visualize the performance of these algorithms for helping to choose the most accurate one. This paper is going to show a comparative study on SVM, k-Nearest Neighbor (k-NN), NB, and DT classifier on face recognition technique. For face recognition at first, the study will extract features of faces from given dataset images using PCA and then apply our classifier algorithms SVM, k-NN, NB, and DT one by one.

PCA is a familiar data processing technique. For the purpose of working with different datasets, methods are used to severely decrease the dimensionality in an elaborate way, so that the most important information in the data is protected and usable.[2,3] In this paper, PCA is used for adjusting data, creating a covariance matrix, finding the eigenvector and eigenvalue of the covariance matrix, prioritizing the eigenvalue, and creating a feature vector. This provides new data with a lower dimension than the original data which is our ultimate preprocessed target to minimize the computational complexity. Computer vision technology requires powerful and high computational devices. So it is quite a challenging task to implement computer vision technology in a lower-end device.

Four different classifiers observed in this paper consist of SVM, k-NN, NB, and DT is the most applied supervised machine learning algorithms. SVM has been developed recently on the basis of statistical learning theory. It is successfully used in various applications such as time series prediction, face recognition, and also in genetic data processing for medical analysis,[4] on the other hand, k-NN classifier concentrates on mechanisms for measuring similarity (distance), computational problems in recognizing nearest neighbor and procedure for reducing the measurement of data.[5]

Whereas the DT classifier checks a collection of values for a given set of features. The decision made by DT classifier is based on checking each individual feature followed by the branches that represent its values. The values are used to traverse through terminal nodes. NB is a probabilistic-based classifier that uses Bayes' theorem to calculate the probability which is used to make the decision. For the comparative study, the contribution shows the accuracy and cross-validation score of SVM, KNN, NB, and DT classifier.

In Section 18.2 a brief introduction, as well as the analysis of related work of face recognition, is explained, Section 18.3 contains the structure of the method and a discussion regarding the data preprocessing and analysis of two different classifiers. The experimental results and conclusion of this paper are given in Sections 18.4 and 18.5.

18.2 LITERATURE REVIEW

Face recognition is a major research issue with comprehensive area and order. To have a huge number of practical implementations such as entrance control, safety monitoring, surveillance system, and bankcard identification. For efficient communications and exchanges among human, these fundamental human behaviors are essential.[6] The early face recognition system started by detecting face patterns in cluttered prospect, standardizing the face image on the basis of geometrical and illumination changes using appropriate classification algorithms.[7,8] Recognizing human faces using Eigenfaces is one of the familiar ways in recent technology.[9] As face recognition technology is developing day by day, there is a lot of comparative analysis of how face recognition technology is performing. Ashwini et al.[10] discussed three different face recognition techniques using PCA with Back Propagation Neural Network, Gabot Filter, and Wavelet Transform and concluded that Wavelet Transform performs 98% accuracy which is better than the other two. Considering pose illumination and facial expression Shinwari et al., 2019[11] performed a comparative study in terms of acquiring better accuracy and said that the Local Binary Pattern Histogram algorithm provides 99.532% accuracy which is better than LDA and PCA.

As this paper performs a comparative analysis between SVM and KNN for face recognition, so there is a lot of face recognition contribution has been performed using these two classifiers. Using the Nearest Center Classification (NCC), Guo et al., 1970[12] compared a face recognition method with the standard eigenface approach which is an SVMs-based recognition technique. A system working by searching the regions of faces using skin color information was proposed by Yeh et al. in 2009[13] which is based on SVM and generates a possible block or face map of faces for detecting faces. A component-based face detector is introduced by Kukenys and Mccane in 2008[14] which presents their improvement of current results with the previous result to achieve sufficient speed and accuracy using SVM classifiers. On the other hand, Prabin et al.[15] proposed a three-step method including various color space models, eigenvectors, and nearest neighbor classifiers and said that the accuracy of this method is notably better than others. A face recognition method was performed by Dhriti and Kaur in 2012[16] where they used the Gabor filter and PCA for feature extraction and KNN as the classifier and showed that the k-nearest Neighbor technique shows impressive accuracy. A face recognition approach using Bagging KNN was performed by Ebrahimpour and Kouzani in 2007[17] where a collective learning group of methods was employed as the system result and said that gaining higher accuracy is the ultimate goal of their system.

There are also a lot of other technologies used nowadays for the application of face recognition. The performance of the human face recognition technique is described by Arunkumar et al. in 2019[18] which provides a lot of information regarding human face recognition technique and their performance. El-Bakry and Abu Elsoud[19] introduce a human face identification system based on a neural network that uses a cluttered scene having a 20 × 20 pixel image having a human face or not. A new face recognition procedure is done by face feature extraction, drawing the angle of the mouth, drawing (x,y) axes in the face and eyes extraction, and combining the whole feature as one parameter for human face recognition performing three-dimensional recognition procedures by Al-Ghamdi et al. in 2010[20]

In this research work, our goal is to propose a face recognition system, which provides the most efficient method with the best performance compared to the remaining system and minimum computational cost among

SVM, k-NN, NB, and DT. We are also focusing on the computational power cost with respect to the device specification as it is also an important part to consider while building real-life machine learning applications. Although the technology is quite powerful nowadays and researchers are focusing on performance rather than computational cost, it is also important to develop highly powerful AI for a low-end device.

18.3 METHODOLOGY

The ultimate destination of this research is to find the most efficient way of recognizing human face using SVM, k-NN, NB, and DT classifier and perform the comparative analysis between four classifiers. As we are going to follow the One-shot Learning technique, we used a comparatively small dataset to train our face recognition model. Then we perform several processes like splitting train and test data, performing preprocessing of image data like resizing, color to grayscale conversion, pixel dimension finalization for making the most suitable data for best performance, performing PCA for dimension reduction, training four classification algorithms, evaluation, and testing. We also perform cross-validation for model evaluation and compare the performance of the previously trained model and the newly trained model after cross-validation.

The flowchart (see Figure 18.1) illustrates the overall workflow of this paper to recognize human face more efficiently.

18.3.1 DATASET DESCRIPTION

In this paper, we used Olivetti dataset for training and testing data.[21] The dataset contains 40 different faces and ten different images for every face. The overall number of data is 400 images with a dimension of 40×40 and the pixel values are scaled to [0, 1].

The dataset containing images of the human face having many different facial expressions like open eyes or closed eyes, face when smiling, eyes having glasses or without glasses, and the parameter of different lighting varieties and images taken at different times. All the images were captured against a dark background with the faces in an upright, frontal posture. This dataset is focused on classification problems.

Figure 18.2 shows the 40 unique faces contained in the Olivetti dataset.

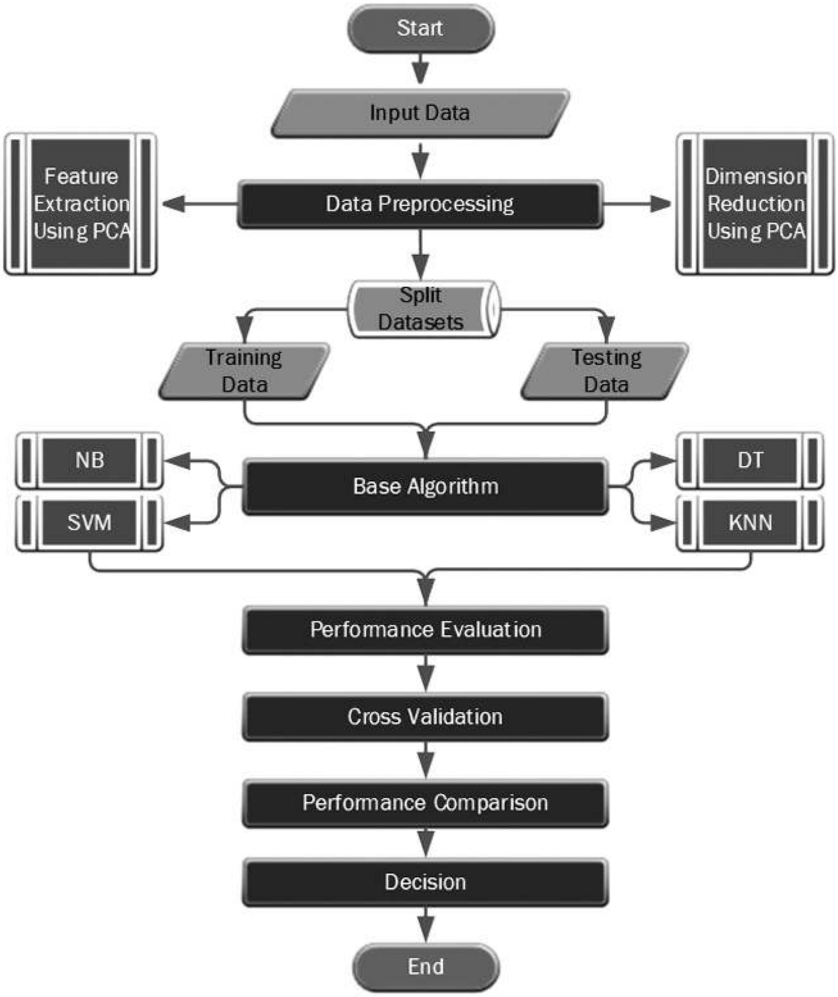


FIGURE 18.1 Process overview.

Figure 18.3 shows the sample of different varieties of face posture in the Olivetti dataset.

The images are scaled to 256 gray levels and stored in format of unsigned 8-bit integers which will be converted into floating point values on the range of [0, 1] in the time of processing. The scaling and conversion of this configuration make it easy to work with different algorithm. This configuration also helps us to train our data using a lower-end device and making the preprocessing and training process easier with minimum computational cost.

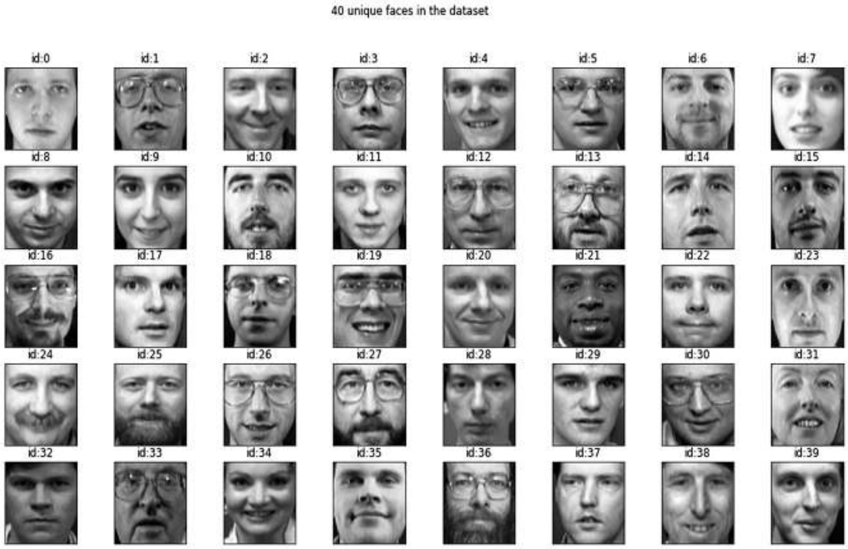


FIGURE 18.2 Forty unique faces.

FIGURE 18.3 Different face postures.

18.3.2 DATASET PREPROCESSING

Before training the model, we have feature extraction process. In this paper, we have used PCA for extracting facial features. It is applied to see likenesses and variances in the dataset. This method transforms those images into an insignificant set of feature images containing face. This is named as "Eigen faces" which is called principal components of the original training images. Then the initial dataset and the new image are predictable on the Eigen faces. For performing PCA, there are several tasks that have to be done like adjusting data, creating a covariance matrix from datasets, finding eigenvalues and eigenvectors of the covariance matrix, taking eigenvalue having the highest priority, and creating a feature vector. Which characterizes a multivariate data form as a smaller set of variables (summary catalogs) in terms of observation of clusters, trends, and outliers.

Finally, we gain new data in which dimension is lower than the original data and it is our main target to minimize our computational complexity. This process helps us representing the data with a different form without any information loss as well as making the data dimension lower for reducing computational cost.

If we consider the way of getting feature from original data, we can write the formula as

$$\text{Final Data} = (\text{Feature Vector})\ T \times (\text{Data Adjust})\ T \tag{18.1}$$

where T indicates transform matrix.

Figure 18.4 illustrates the phase of data before and after performing PCA.

18.3.3 MODEL IMPLEMENTATION

SVM is a very potential classifier that aims to classify data as well as find the best possible boundary. It maintains the largest distance from the points. We actually wanted the points to stay as far as possible from the boundary. Therefore, the paper added two other boundaries that are parallel with the main line and have equal distance from the main boundary line and maximize the distance or margin between the two. For minimizing the error, we consider two ways to measure the model including the number of points which is misclassified and the wide of a margin. We can calculate total error using this formula:

$$\text{Error} = \text{Classification Error} + \text{Margin Error} \tag{18.2}$$

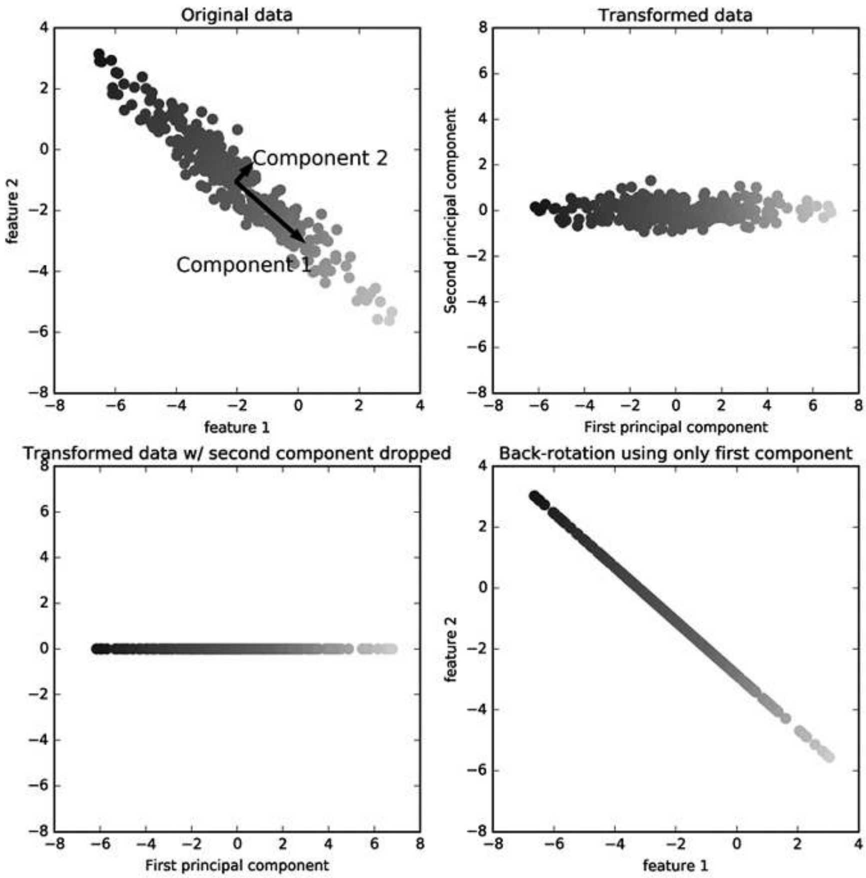


FIGURE 18.4 PCA.

The closer points get punished less and the far points get punished more. This idea will be to minimize error using Gradient Descent. The C parameter sometimes we need to select the large margin model or sometimes-small margin model for minimizing our error. Therefore, the method depends a lot on the data and the problem we approaching which means some flexibility here. Those flexibilities are going to be provided by the C parameter. So the Error Function can be written as the formula

$$\text{Error} = \text{C Classification Error} + \text{Margin Error} \tag{18.3}$$

The large C Classification Error focuses on classifying points and the small C Classification Error focuses on a large margin. SVM faces Overfitting and Underfitting problem, which can be regulated by Regularization. λ is Regularization parameter which is promotional to the decreasing Weight Vector. There is no Regularization for the 0 value of λ.

k-NN is a famous and very easy classification algorithm. Main steps of KNN including measuring the distance from the unknown point to all other points, sorting the data points in ascending order on the basis of distance values, from the sorted data points getting K-number of points. In this K-number of points which class's point has in most of the time, the unknown point has to be pointed in that class.

18.4 RESULT ANALYSIS

In this section, the performance of the mentioned four classifiers is presented with corresponding performance evaluation metrics. At first, we considered Confusion Matrix to describe the complete performance of the model.

TABLE 18.1 Confusion Matrix.

Classifier name	Predicted actual	True	False
SVM	True	38	2
	False	4	36
k-NN	True	29	11
	False	12	28
NB	True	35	5
	False	7	33
DT	True	27	13
	False	15	25

Harmonic Mean between precision and recall is defined as F1 Score. The range for F1 Score is [0, 1]. F1 Score is used to measure the preciseness and robustness of a classifier.

In this paper, as we are performing a comparative investigation among SVM, k-NN, NB, and DT classifier so the performance of these four classifiers has been calculated and analyzed differently.

Then, we again performed cross-validation method to check our acquired performance and to perform the comparison between the classifiers performance of our used model. We used K-Fold cross-validation model with k value of 5 to gain the mean cross-validation score of SVM, K-NN, NB, and DT.

TABLE 18.2 Comparison of Accuracy Score.

Algorithm	Accuracy	False positive	False negative	Precision	F1 score
SVM	91%	5%	11%	95%	91%
k-NN	71%	28%	29%	72%	71%
NB	85%	13%	16%	87%	85%
DT	65%	34%	35%	67%	65%

The performance of cross-validation is measured using the following equation:

$$Performance = \frac{1}{5}\sum_{i=1}^{5} Performance_i \tag{18.4}$$

Figure 18.5 shows the cross-validation process to validate the performance.

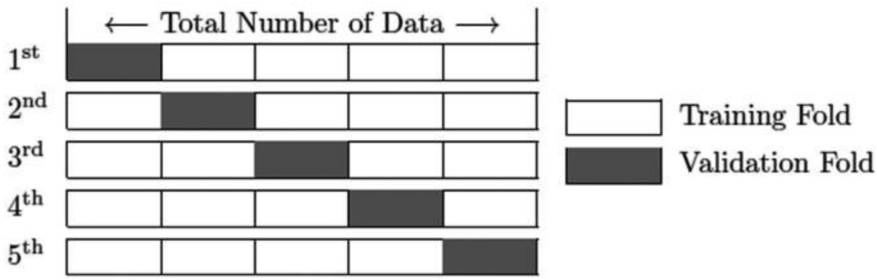


FIGURE 18.5 Cross-validation.

Here is the performance of algorithms after performing cross-validation.

TABLE 18.3 Comparison of Mean Cross-Validation Score.

Algorithm	Accuracy
SVM	88%
k-NN	70%
NB	81%
DT	63%

From the above analysis, we determine that SVM performs better accuracy than k-NN, NB, and DT. As we see in Table 18.3, SVM performs accurate prediction for 91% of cases, where the k-NN, NB and DT perform accurate prediction for 71%, 85%, 65%, respectively. Then we perform cross-validation model to validate the performance and ensure the accuracy of SVM, k-NN, NB, and DT in terms of providing the most suitable classifier to detect human face correctly and efficiently. Here, we have gained the accuracy for SVM classifier is 88% where k-NN classifier is 70%, NB classifier is 81% and DT classifier is 63% with the face recognition technique which defines SVM performs better than other three classifier as usual. So, the performed comparative analysis among SVM, k-NN, NB, and DT described in this paper can help to inform researchers to choose the best algorithm for their work.

18.5 CONCLUSION

In this research paper, we have worked with the applications of image processing called face recognition and also performed a comparative analysis among four classifier algorithms SVM, k-NN, NB, and DT. Firstly, the dataset goes through a preprocessed method, and then the feature is extracted using PCA. Four mentioned classifiers algorithms are used to find the best accuracy in terms of recognizing human face. From the overall investigation, we find that SVM gives better results than others three classifier by analyzing the accuracy score and mean validation score for both of them. K-Fold cross-validation has been used to find out the accuracy of the algorithms. In future we want to work with our own datasets.

As comparative study on classifier algorithms will be very helpful for programmers to decide during choosing classifier algorithms to train the recognizer model. For this paper, face recognition techniques face some problems such that poor image quality, small size images, different face angles, data processing, and storage can cause poor accuracy of classifier algorithms during training the dataset.

KEYWORDS

- **comparative study**
- **face recognition**
- **PCA**
- **SVM**
- **k-NN**

REFERENCES

1. Lin, S. H. An Introduction to Face Recognition Technology. *Inf. Sci. Int. J. Emerg. Transdiscip.* **2000,** *3*. 10.28945/569.
2. Wójcik, W.; Gromaszek, K.; Junisbekov, M. Face Recognition: Issues, Methods and Alternative Applications. In *Face Recognition–Semisupervised Classification, Subspace Projection and Evaluation Methods,* 2016. 10.5772/62950.
3. Jolliffe, I.; Cadima, J. Principal Component Analysis: A Review and Recent Developments. *Philos. Trans. A Math. Phys. Eng. Sci.* **2016,** *374*, 20150202. 10.1098/rsta.2015.0202.
4. Evgeniou, T.; Pontil, M. In *Support Vector Machines: Theory and Applications*, 2001; vol *2049*, pp 249–257. 10.1007/3-540-44673-7_12.
5. Cunningham, P.; Delany, S. J. K -Nearest Neighbour Classifiers. *Mult. Classif. Syst.* **2007,** 1–17, DOI: 10.1016/S0031-3203(00)00099-6.
6. Ahmad, T.; Ali, E. B.; Ahmed, E. H. Face Recognition: A Literature Review. *Int. J. Signal Process.* **2005,** *2*, 88–103.
7. Hazim, N.; Al-Dabbagh, S. S. M.; Esam Matti, W. Face Recognition: A Literature Review. *Int. J. Appl. Inf. Syst.* **2016,** *11*, 21–31. 10.5120/ijais2016451597.
8. Kaur, P.; Krishan, K.; Sharma, S. K.; Kanchan, T. Facial-Recognition Algorithms: A Literature Review. *Med. Sci. Law* **2020,** *60* (2), 131–139. DOI: 10.1177/0025802419893168
9. Turk, M. A.;Pentland, A. P. In *Face Recognition using Eigenfaces*, Proceedings of 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition; Maui, HI, USA, 1991; pp 586–591. DOI: 10.1109/CVPR.1991.139758.
10. Ashwini, B.; Akkawar, A.; Burange, M. S. Review and Comparative Study of Face Recognition using Different Neural Network Algorithms. *Int. J. Eng. Res Gen. Sci.* **2015,** *3* (2).
11. Shinwari, A.; Balooch, A.; Alariki, A.; Abduljalil Abdulhak, S. A Comparative Study of Face Recognition Algorithms under Facial Expression and Illumination, 2019; pp 390–394. 10.23919/ICACT.2019.8702002.
12. Guo, G.; Li, S.; Chan, K. In *Face Recognition by Support Vector Machines*, Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition, 1970.
13. Yeh, J.; Pai, Y. C.; Wang, C. W.; Yang, F. W.; Lin, H. J. Face Detection using SVM-Based Classification. *Far East J. Exp. Theor. Artif. Intell.* **2009,** *3*.
14. Kukenys, I.; Mccane, B. Support Vector Machines for Human Face Detection, 2008.
15. Prabin, J.; Poornima, P.; Kumar, K. A Novel Method for Color Face Recognition using KNN Classifier, 2012. 10.1109/ICCCA.2012.6179151.
16. Dhriti, D.; Kaur, M. K-Nearest Neighbor Classification Approach for Face and Fingerprint at Feature Level Fusion. *Int. J. Comput. Appl.* **2012,** *60*, 13–17. 10.5120/9759-1517.
17. Ebrahimpour, H.; Kouzani, A. Face Recognition Using Bagging Knn, 2007.
18. Arunkumar, S.; Sharma, R.; Kumar, D.; Puranik, V.; Gautham, K. Performance Analysis of Human Face Recognition Techniques, 2019; pp 1–4. 10.1109/IoT-SIU.2019.8777610.
19. El-Bakry, H.; Abu Elsoud, M. Human Face Recognition using Neural Networks, 1999; pp C28/1–C28/8. 10.1109/NRSC.1999.760912.
20. Al-Ghamdi, B.; Allaam, S.; Soomro, S. Recognition of Human Face by Face Recognition System using 3D. *JICT* **2010**.
21. https://www.kaggle.com/imrandude/olivetti

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

CHAPTER 19

STUDY AND ANALYSIS OF COMPUTER VISION FOR CASUALTY DETECTION

MANGESH D. SALUNKE, NILESH N. THORAT, SARTHAK SURVE, AMMAR SAKRIWAL, ATHARV KALEKER, and KARAN DABARE

Department of Computer Engineering, Marathwada Mitra Mandal's Institute of Technology, Lohegaon, Pune, India

ABSTRACT

To draw in new researchers, the study of abnormal and violent activity detection has become an important topic of study in computer vision and image processing. We are all familiar with intelligent security systems based on video monitoring, where a security guard witnesses violence and alerts the appropriate authorities. But as a result, the process becomes reliant. Delay in reporting the casualty may result in significant harm or damage. However, the design model enables the end user to instantly report the most frequent casualties that occur in public, such as fires, automobile accidents, violence, and weapons (guns), by simply visiting the website and uploading the causality image, which will immediately report to the relevant authority. This paper's goal is to visually identify the casualty. The YOLO-V7 algorithm is utilized by the suggested technique. The suggested approach compares the number of false positives and false negatives with the Faster RCNN by using the YOLO-V7 algorithm.

19.1 INTRODUCTION

Casualty is a phenomenon that can occur at any time, at any given moment. If reported late, some casualties might have a significant negative impact on

Applied Soft Computing Techniques: Theoretical Principles and Practical Applications. Samarjeet Borah, Ratna Raja Kumar Jambi, Sharifah Sakinah Syed Ahmed, & Mahendra Prabhakar Deore (Eds.)
© 2025 Apple Academic Press, Inc. Co-published with CRC Press (Taylor & Francis)

the environment. The created machine learning model is capable of spotting violence, fire, and weapons (including guns). The YOLO V7 method is used by the model to find the casualties. For instance, if a fire breaks out in a public area and is not reported right away, the amount of destruction will rise. A weapon (gun) can cause defacement in public if it is held out in plain sight by an unidentified person. One's life can be saved by reporting the car collision right away. Also, it is critical to swiftly stop public violence.

Our project's goal is to create a website powered by Machine Learning and Deep Learning that can identify and notify the authorities of incidents like fires, shootings, accidents, and fights. This website can be quite helpful in delivering prompt aid and improving public safety in the modern world, where similar situations are occurring increasingly frequently. The proposed system can be used to alert casualties caused at various public places as soon as possible to the authorities. Casualties can be caused by human behavior or natural calamities.

Advanced machine learning techniques are used on our website to analyze user-uploaded image data. The technology is capable of precisely identifying patterns and objects connected to the episodes of interest, such as flames, guns, damaged vehicles, and physical altercations. The website instantly alerts local authorities when an incident is found, giving them access to real-time information about the issue and empowering them to take the necessary action.

Our website, in our opinion, has a great deal of potential to increase public safety and save lives. We can quickly recognize and respond to crises by utilizing the power of machine learning, which reduces response times and boosts the overall performance of emergency services. With the help of our project, we seek to make law enforcement, emergency personnel, and the general public's lives safer and more secure.

Overall, the idea significantly improves public safety by speeding up the process of reporting hazardous situations and doing so in a more thorough and understandable way. Also, this will make it simpler and faster for users to record the appropriate casualty.

The main objective of this research is to categorize violence as well as abnormal activity detection from videos utilizing computer vision, as well as to identify the best practices currently in use.

Using the aid of the systematic review approach, we analyze the research that is currently available in this field. The study that is provided is assessed using predetermined standards. Sort the evaluation process's results into categories based on their applicability. Articles on the identification of violence are included in the sample of the systematic examination.

A fire detection engine based on video processing will be presented in this study. The contributions of this work include the detection of fire in real-time video streams, the separation of controlled from uncontrolled fires, and the generation of uncontrolled fire alarms. To detect movement in the frames, the suggested engine first subtracts the backdrop from them. The color detection module that locates the fire region is applied to the resulting frame. Once the fire region has been located, the wavelet coefficients of the area are used to calculate the fire signal's energy. The evaluation distinguishes between fire and non-fire regions based on the energy present in these areas.

19.2 PROPOSED SYSTEM DESIGN

The developed model is a combined approach of Web development, Machine Learning, Deep Learning, Artificial Intelligence, and Computer Vision. The suggested method looks for three crucial qualities to identify fire. These are the energy, color, and mobility of fire. Both regulated and potentially dangerous fires have these characteristics. In order to differentiate between controlled and hazardous fire, the rate of fire region increment is used. The suggested method looks for three crucial qualities to identify fire. They represent the energy, color, plus mobility of fire.

Both regulated and potentially dangerous fires have these characteristics. In order to differentiate between controlled and hazardous fire, the rate of fire zone increase is used. The suggested method looks for three crucial qualities to identify fire. They represent the energy, color, plus mobility of fire. Both regulated and potentially dangerous fires have these characteristics. In order to differentiate among controlled versus hazardous fire, the rate of fire zone increase is used.

Figure 19.1 shows the architectural model of the proposed system that is Computer Vision for Casualty Detection. In that, it shows about the process to follow for the actual detection of causality using computer vision, machine learning algorithms, and prediction techniques. The proposed engine detects moving things when it starts up, the live video feed. Using moving objects is an identified approach for background subtraction. The color of the object that moves is then compared to fire using a color detection technique. To enable additional fire features, the selected pixel values were given this wavelet processing. Object detection establishes an object's presence, while localization establishes the object's position inside the image.

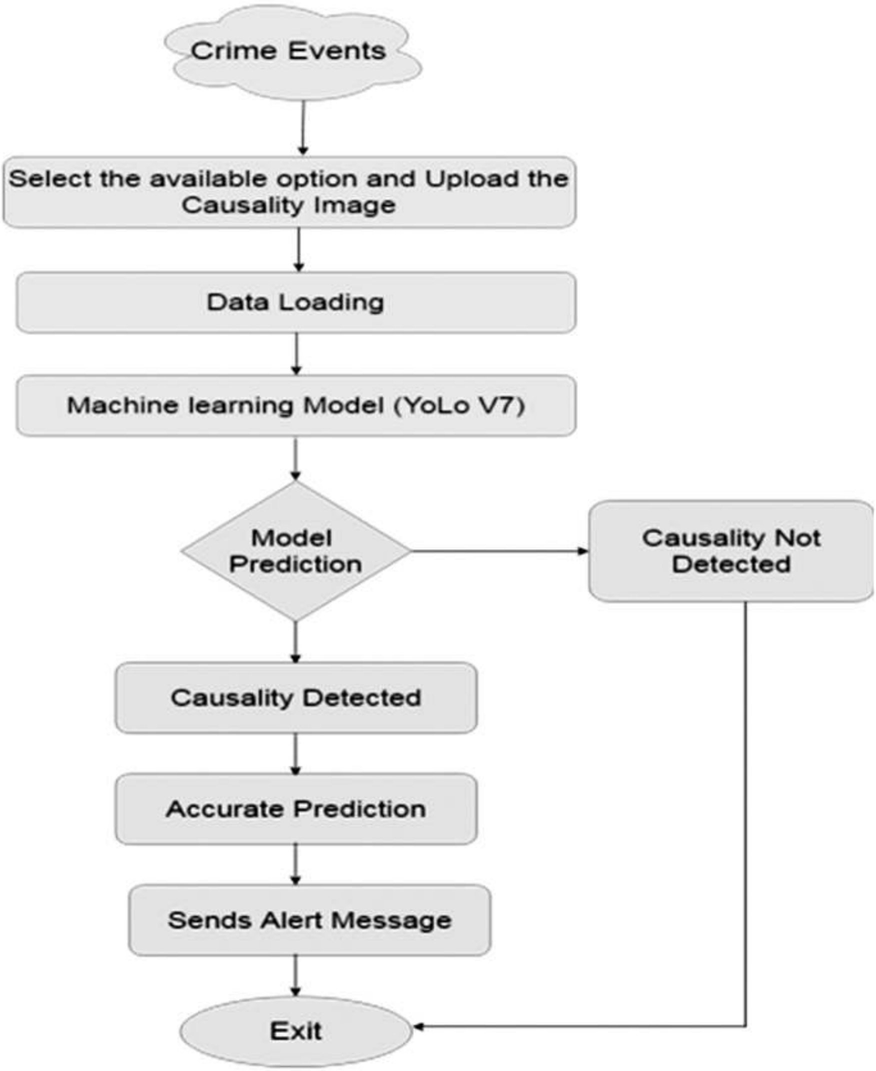


FIGURE 19.1 System design architecture.

19.2.1 USER INTERFACE

User interface is a key means by which the user will interact with the model, this interaction is possible with the help of a website where a detailed description of the system usage is provided.

19.2.2 MACHINE LEARNING MODEL

The developed system uses the YOLO V7 algorithm for detection purposes and is completely responsible for analyzing the detection patterns in the uploaded images and generating the desired output.

19.2.3 IMAGE DETECTION

Once the image is detected for the happening casualty an alert message will be promoted to the managing authority. If the model does not find any type of damage in the uploaded image it will not process toward the next iterations.

19.3 RESULTS AND DISCUSSION

In this section, we will analyze some of the expected outcomes or results of the proposed system, during testing as well as training the proposed system we are able to find the following desirable outcomes from the proposed system architecture.

19.3.1 FIRE DETECTION

To identify flames or smoke in an image or video feed, this module uses computer vision techniques. It can also pick up temperature variations that can point to a fire. The device may set off sirens, alert authorities, and offer information about the location and intensity of the fire when it detects one.

19.3.2 ACCIDENT DETECTION

To recognize vehicles, pedestrians, and other things in an image or video feed, this module utilizes object detection and recognition algorithms. Additionally, it can recognize patterns of motion that can point to an accident, like abrupt stops or impacts. The technology may immediately notify emergency personnel and offer details about the location and severity of the accident when one is identified.

FIGURE 19.2 Fire detection model before.

FIGURE 19.3 Fire detection model after.

FIGURE 19.4 Accident detection before.

FIGURE 19.5 Accident detection after.

19.3.3 VIOLENCE DETECTION

To identify when a physical altercation or fight is occurring, this module employs computer vision and audio processing algorithms. It can pinpoint the number of participants, the severity of the violence, and the scene of the altercation. The system can inform security staff or law police when a conflict is present and provide real-time video feeds to aid in their response.

FIGURE 19.6 Violence detection before.

FIGURE 19.7 Violence detection after.

19.3.4 GUN DETECTION

This module detects guns or other weapons in an image or video feed using object detection and recognition techniques. Also, it can spot movement patterns that can point to a danger. The device can notify security guards or law police when a gun is present and offer details about the threat's location.

FIGURE 19.8 Gun detection before.

19.4 PROPOSED SYSTEM

Our project is a website powered by machine learning that seeks to identify and notify authorities about fire, gun, accident, and conflict situations. The suggested system is made up of a number of parts, such as data preparation and collecting, model selection and training, and website construction.

FIGURE 19.9 Gun detection after.

19.4.1 DATA GATHERING AND PREPARATION

The initial part of the suggested method entails gathering a sizable dataset of labeled photos associated with the relevant incidents. We get pictures from many places, such as social media sites and open databases. In order to preprocess the data, we resize the photographs to a standard size, remove noise, and format them appropriately.

19.4.2 MODEL SELECTION AND TRAINING

To identify occurrences from the photos, the second part of the proposed system comprises choosing and training a deep learning model. We compare the performance of numerous cutting-edge models and choose the best one. Using the preprocessed dataset, we then refine the model to attain high accuracy and reliability in recognizing occurrences.

19.4.3 WEBSITE DEVELOPMENT

The final element of the suggested strategy entails creating the website itself. We create a straightforward user interface that makes it simple for users to upload photographs. The website uses the trained deep learning model to scan uploaded images and identify any instances of fire, gunfire, accidents, or fights. Local authorities receive real-time information about the situation via the website when an event is detected, allowing them to react swiftly and efficiently.

The suggested approach, taken as a whole, offers a considerable improvement in the use of machine learning for public safety. Our technology, we believe, may contribute to a safer and more secure society, safeguarding lives and property and enhancing the overall efficacy of emergency services by enabling rapid identification and reaction to emergencies.

19.5 ALGORITHM

You Only Look Once, or YOLO is a technique which employs neural networks to recognize objects in real-time. The effiectiveness and rapidity of this technique account for its success. It is currently used in many different contexts to distinguish between animals, people, parking meters, including traffic signals. The leading object detection algorithm in the field of computer vision is called YOLO, and it is a state-of-the-art algorithm.

In the past, people have employed methods like R CNN, Fast R CNN, and Faster R CNN, as well as sliding window object detection. But because of its speed and precision, YOLO has since its creation in 2015 become the industry standard for object recognition.

To recognize things in real-time as well, a clever convolutional neural network (CNN) known as YOLO is applied. The technique applies a single neural network to the entire image, segments it into regions, and estimates boundaries, and possibilities for every area. These boundaries are weighted using anticipated likelihood. YOLO is well-liked since it can operate in real-time while achieving outstanding precision. The method "only looks once" at the image in the sense that it just needs to execute one forward propagation run through the neural network to generate predictions, neural network for forecasting.

19.6 CONCLUSION

To sum up, our project is a website powered by Machine Learning, Deep Learning, and Artificial Intelligence that can identify and notify authorities about Fire, Gun, Accident, and Violence occurrences from photographs posted by users. Advanced machine learning methods are used in the project to evaluate photographs and find certain objects and patterns connected to the episodes of interest.

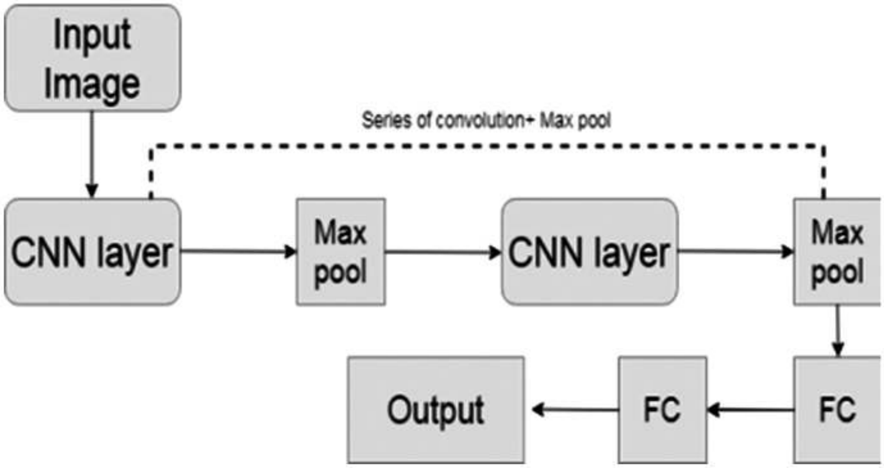


FIGURE 19.10 Proposed algorithm working.

We have shown how machine learning can improve public safety and emergency response through the project's progress. Our solution can increase the overall efficiency of emergency services, cut down on response times, and ultimately save lives by enabling rapid identification and reaction to emergencies.

We think that our initiative might be a useful resource for first responders, law enforcement, and the public. Authorities can respond swiftly and effectively to incidents by using the system's ability to offer real-time information about them.

In the future, we understand how crucial it is to continue working together and innovating in this field. Advanced machine learning algorithms that can increase the precision and dependability of event detection require continual research and development. We also support responsible and open deployment of such technologies while acknowledging the ethical and legal issues involved.

In conclusion, our study marks a significant advancement in the use of machine learning for public safety, and we anticipate that it will help create a society that is safer and more secure.

KEYWORDS

- **computer vision**
- **casualty detection**
- **YOLO-V7**
- **false positives**
- **false negatives**
- **RCNN**

REFERENCES

1. Zaman, T.; Hasan, M. A.; Ahmed, S.; Ashfaq, S. Fire Detection Using Computer Vision. Department of Computer and Information Systems Engineering, 2021.
2. Warsi, A.; Abdullah, M.; Husen, M. N. In *Gun Detection System using YOLOv3*, IEEE 6th International Conference on Smart Instrumentation, Measurement and Applications (ICSIMA 2019), Aug 27–29, 2019; Kuala Lumpur, Malaysia, 2019.
3. Patel, S.; Patel, N.; Deshpande, S.; Siddiqui, A. In *Custom Object Detection System with YOLO Algorithm*, 2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (ISMAC), 2021.
4. Rai, N.; Gulair, D.; Shiningwala, J.; Molawade, M. H. A Review on State-of-the-Art Violence Detection Techniques. P *Int. Adv. Res. J. Sci. Eng. Technol.* **2021,** *8* (7).
5. Nayak, R.; Behera, M.; Patil, U. Video Based Real Time Intrusion Detection System for Smart City Application. *Int. Adv. Res. J. Sci. Eng. Technol.* **2019,** *8* (7).

CHAPTER 20

TRAFFIC SIGN RECOGNITION USING CONVOLUTIONAL NEURAL NETWORK

UMA KARANJE, CHAITANYA BHOSALE, GAYATRI KINGE, AAKANKSHA MAGAONKAR, APARNA GADE, and JANVI KANKHAR

Department of CSE, Marathwada Mitra Mandal Institute of Technology (SPPU), Pune, India

ABSTRACT

In the era of self-driving cars, ADAS (Autonomous Driver Assistance Systems) implement the functionality of Traffic Sign Recognition (TSR). Implementation of the TSR technique as a driver assistance technique can help the vehicle to recognize the image and then categorize images of traffic sign boards on the road. This paper pre-processes and trains various traffic sign board images using Convolutional Neural Networks (CNN). The CNN model learns from the standard German Traffic Sign Recognition Benchmark (GTSRB) dataset which includes 35,000 images that are divided into 43 classes such that each class is having similar images. Max pooling layers are used to extract highly accurate features of input images. In this paper, sequential API is used to arrange the Keras layers in sequential order along with ReLU and SoftMax activation functions. The model gets fit with 43 neurons in the dense layer and the epoch value equals 20. The model once trained and tested gives an accuracy of 94%. The model is then deployed using the Python Flask server.

20.1 INTRODUCTION

Self-driving cars are one of the most talked about topics of this decade. These vehicles are based on intelligence, they are driverless cars that themselves can

Applied Soft Computing Techniques: Theoretical Principles and Practical Applications. Samarjeet Borah, Ratna Raja Kumar Jambi, Sharifah Sakinah Syed Ahmed, & Mahendra Prabhakar Deore (Eds.)
© 2025 Apple Academic Press, Inc. Co-published with CRC Press (Taylor & Francis)

detect and identify the traffic signals and traffic signs and take the required actions. It may happen that due to some conditions, the traffic sign may not be noticed by humans, in such cases the Intelligent Transport System can detect the signs and recognize them. In a country like India where there is a huge crowd on roads, there is a high probability of accidents, hence, to avoid them and to prevent heavy traffic, it is important to follow the traffic signs which are present on either side of the road. There are various network architectures used to build a TSR system, and each of them has its own accuracy, for example, RMR CNN, Mask R CNN, Fast R CNN, etc.

In this paper, the GTSRB standard dataset is used, which consists of 35,000 instances of traffic sign board images divided into 43 different classes. In deep neural networks, convolutional neural network is a type of technique that is majorly used in computer vision. It consists of various layers which all together help to recognize and classify the images in the dataset. It consists of activation functions named ReLU and SoftMax. The accuracy of the model is approximately 0.9403. TSR model is implementable in various autonomous vehicles as well as an application of driver assistance systems.

20.2 LITERATURE SURVEY

The research in the field of TSR is vast. Researchers have developed and implemented the TSR model according to its applications, mainly TSR has two types of functional applications, TSD and TSC. TSC and TSD Researchers of "Indian traffic sign detection and recognition using deep learning" studied and implemented the robust TSR model using advanced and efficient CNN algorithms like RMRCNN, they proposed the accuracy of their model to be 97.08%.[1] LeNet-5 is another CNN-based framework that is used to classify the images which involved the color probability generation using HSV.[6] Researchers from Shiga University, Japan studied the various problems faced by the TSR model while detecting and recognizing the traffic sign images.[5] Students from Mumbai University proposed the model of TSDC (Traffic sign detection and classification), where they used the deep learning framework YOLO.[4] Additionally, deep learning-based techniques for recognition algorithms have a high accuracy recognition rate, but some issues, including high complexity and lengthy processing times, do exist.[2] As a result, it is vital to improve the classification process of deep learning algorithms. The support vector machine (SVM) is used to recognize traffic signs. Adaboost first screened the candidate recognition images before the SVM classed them. Although the detection time was lengthy, the recognition accuracy was good.[2] Numerous research organizations have made available traffic sign data sets. The GTSRB, one of these data sets, is commonly used (Standkamp et al., 2011). This collection was created for the International Joint Conference on Neural Networks (IJCNN) TSR competition in 2011. Featuring a total of 51,839 items, the GTSRB set consists of 43 categories (39,209 for training and 12,630 for testing). Some of the classifications included are restricted speed, mandatory, cautionary, de-restriction, and other enforceable signs. Pixels on the signs, which can be anywhere from 1515 to 222193, have a 10% margin. Labeled photos with actual ROI size and position make verification of results easy in the GTSRB, which is focused on recognition of signs with small backdrops. For detecting signs, the GTSDB was made for an IJCNN 2013 competition and includes 900 photos in three categories (prohibitory, required, hazard) with 600 for training and 300 for testing. Employing machine learning algorithms to differentiate traffic lights in images and video sequences has become a popular method for a wide range of features. However, relying solely on manmade features limits the effectiveness of techniques when distinguishing true signals from fake signals in complex driving scenarios (Yuan et al., 2019). Neural networks (NN)[13] have emerged as a prominent detection approach for both classification and detection as demonstrated by Prem Kumar et al. (2019). Scene images can be accurately segmented using random forests as shown by Ellahyani and El Ansari (2017a). As part of their preprocessing phase, the authors implemented a mean-shift clustering technique. Following that, the relevant sections were identified using a random forest classifier utilizing appropriate colors. Additionally, Ellahyani and El Ansari (2017b) used SVMs to classify color-segmented shapes as triangles, circles, or rectangles.

Table 20.1 gives a summary of the literature survey with their advantages.

20.3 METHODOLOGY

In this paper, the processing and working of CNN are studied. CNN is an interesting computer vision algorithm to recognize as well as categorize photos based on their content. Figure 20.1 shows the system architecture in which feature extraction and classification is done with different convolutional layers.

TABLE 20.1 Literature Survey.

Sl.No.	Year	Title	Advantages	Remark
1	2022	"Indian Traffic sign detection and recognition using deep learning."	Demonstrated strong performance under a variety of circumstances, including changes in scale, light, and orientation.	A broad range of traffic sign categories was successfully used to validate the developed CNN variation.
2	2021	"Robust traffic sign detection and Recognition using YOLOv2."	With various data augmentations, training data showed increased resilience.	Additional difficulties were addressed, such as the paucity of training data and the rise in the number of detections in each environment.
3	2021	"Detecting traffic signs on the road by first classifying traffic sign images using CNN on GTSRB and then on Indian traffic signs using the Indian dataset."	Classification of traffic signs using CNN with different filters helped improve the system by selecting the most accurate model with proper accuracy.	Good accuracy of the model helps the system to improve its work.
4	2020	"Classification of the image for extracting information from the picture by marking the image pixels to different classes."	Studied all facets of machine learning as well as the fundamentals of the CNN architecture.	A comparison of hybrid CNNs was done, and it was discovered that CNN-LSTM produced greater performance.
5	2018	"To evaluate performance on image recognition and detection dataset using CNN models."	Using MNIST datasets, recognize handwritten digits using OCR (Optical Character Recognition), verify signatures, etc.	The accuracy of MNIST is good; however, training with several epochs can increase the accuracy of CFAR-10.

20.3.1 CONVOLUTIONAL OPERATION

This is the very first step in CNN where the image pre-processing takes place. Using the sliding window approach the filters of n × n dimensions are placed over the actual grid of the image. The values of the filter grid get multiplied by the pixel values of the concerned ROI of the input image. This will eventually indicate the presence of edges in the image, such as horizontal and vertical edges.

20.3.2 POOLING LAYER

Pooling is another layer of CNN operation. In this paper, the MaxPooling layer methodology is implemented. MaxPooling helps in extracting the predominant features of an input image by using a suitable stride value. MaxPooling extracts the feature which is having a high average convolutional value. This operation is implemented after the convolutional layer to enhance and sharpen the features of images.

20.3.3 FLATTENING

In this layer, the features extracted from the convolutional layers and the pooling layers are transformed into the one-dimensional matrix.

20.3.4 FULL CONNECTION

The flattened matrix of accurate image features is then supplied as input to the ANN. The connection of CNN operations with this ANN is known as the Full connection step, where the hidden layers in this ANN have neurons and have a connection between each neuron of the first layer with that of the next layer.

20.3.5 DENSE LAYER

Fully connected layers in ANN creates a dense network which is having 43 final neurons (this represents 43 different classes of our traffic sign images). The dense layer prefers using the SoftMax activation function which is used for multiclass classification by converting the integer values to the probability that showcases and predicts the class labels of respective images.

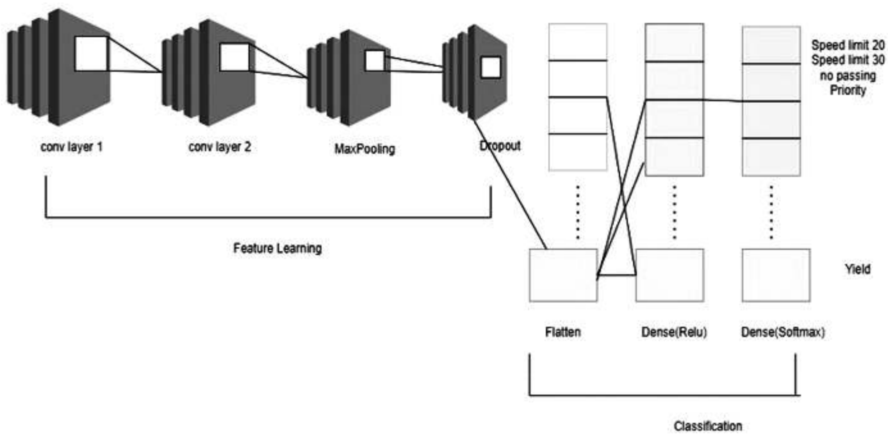


FIGURE 20.1 System architecture.

The photos of traffic signs available in GTSRB are obtained from footage acquired by the vehicle-mounted camera.

The GTSRB comprises 43 kinds of traffic signs, as illustrated in Figure 20.2, and there are different numbers of the various types of traffic signs. Each sort of traffic sign has its own catalog, which is made up of a CSV file with a class description annotation on it and a single image from several tracks (each track includes 30 pictures). GTSRB may also be split into six groups based on the diverse instruction contents, as indicated in Figure 20.3: speed limit, hazard, required, impermissible, de-restricted, and distinctive traffic signs. Different resolutions, illumination settings, atmospheric conditions, diffraction degrees, tilt levels, and other photos are included in the same type of traffic sign, making the dataset more consistent.

Because the photos had varying dimensions and were RGB images, pre-processing was required before uploading the images to the CNN network. The photographs are converted to greyscale versions, and their size is also roughly decreased to 30 × 30 pixels. The photos are all meticulously tagged and stored in their appropriate files because there are 43 output classes in all, each of which represents an Arabic traffic sign. From one class to another, the number of photos per class varies.

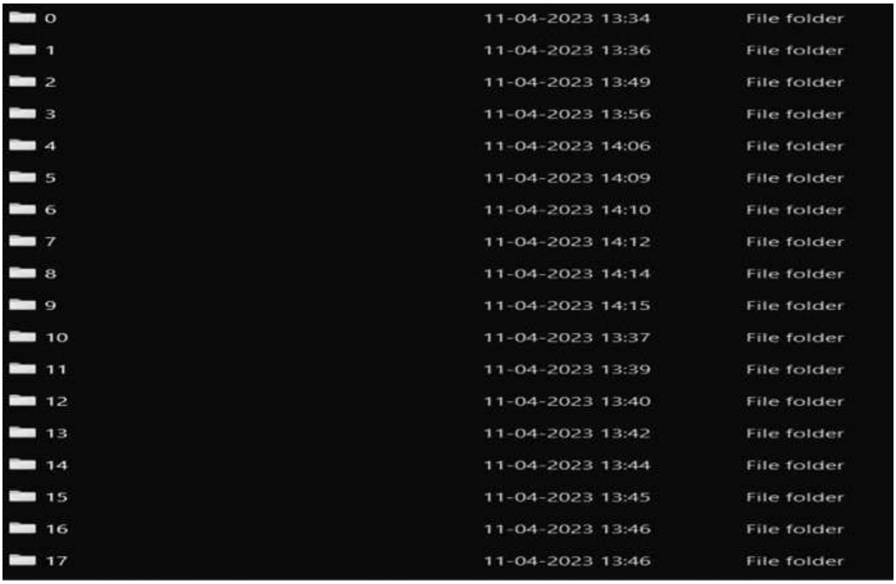


FIGURE 20.2 Classes containing images.

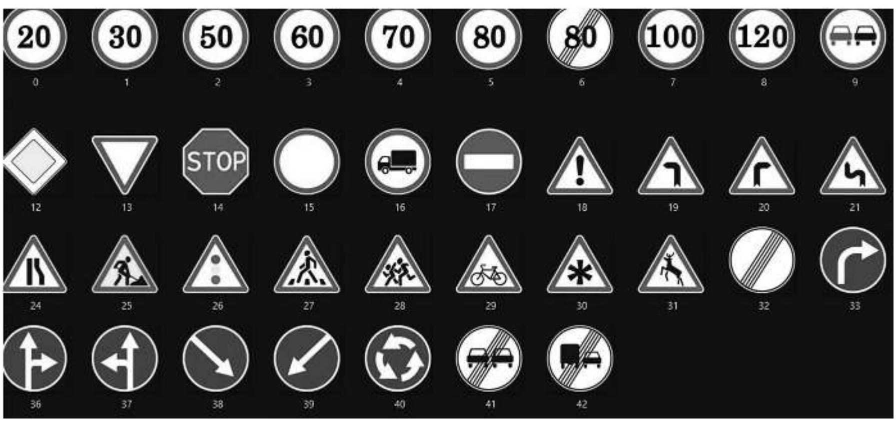


FIGURE 20.3 Metadata.

The initial steps in building the CNN model architecture are as follows:

Step 1. First, we must add the layers in the sequence such that: convolutional layers, a MaxPooling layer, one Dropout layer, and a flattening layer.

Step 2. The count of filters in the convolutional layer is given. It convolutionally transforms the source pictures and creates the feature map.
Step 3. Positive feature values are preserved while negative ones are eliminated by the ReLU. The pooling layer employs MaxPooling to reduce image dimensions while enhancing important features using the rectified feature map as a starting point.
Step 4. During the training process of a neural network, dropout layers are utilized to prevent overfitting by deactivating certain input neurons.
Step 5. By contrast, flattening layers create a one-dimensional array from the input feature map. All outputs from the prior layer are then transmitted to the following completely connected layer by the dense layer.
Step 6. After these layers have been added, the model is created and trained using the dataset's processed image data.
Step 7. Once training is complete, the trained model is used to predict the test data and the resulting output includes the assigned sign name and class ID.

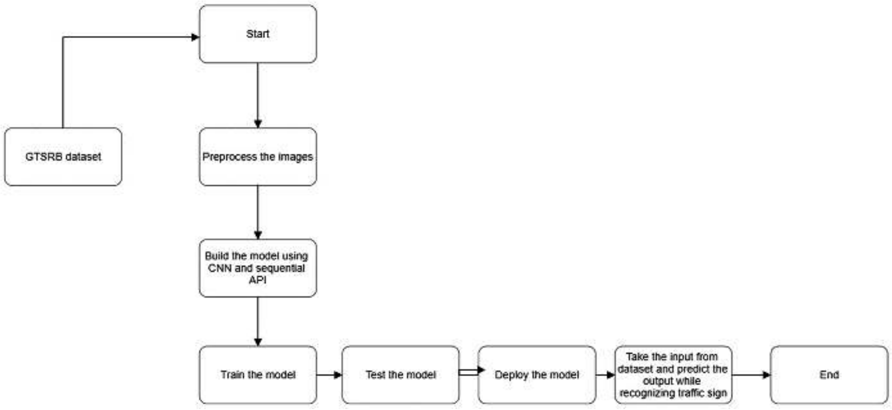


FIGURE 20.4 Flowchart to understand how a sign is recognized.

Figure 20.4 shows the flowchart to understand how a sign is recognized by processing an input image from a dataset. An input can be passed through a filter during convolution, leading to activation. The resulting feature map shows the positions and intensity of identified features in an input, like an image. This linear process, much like a standard neural network, then multiplies the input by a set of weights. The approach in question was designed for analyzing two-dimensional input. To achieve this, the filter or kernel, which is a two-dimensional array of weights, multiplies with an input data array. The outcome is a single value that signifies the application of the filter upon the input array. By repeating this process multiple times and applying the filter to different sections of the input array, we get a two-dimensional output array that represents features in our desired output map hence the name "feature map."

The ReLU transform function is applied to nodes in a neural network and only triggers activation if the input surpasses a set threshold. If the data falls below that designated point, the output will be zero. Additionally, this transformation boasts a linear relationship with its dependent variable. During pooling, we decrease the picture stack size. The process involves passing through an activation layer before flattening the retrieved feature map into a one-dimensional array.

The network's last layer connects all the neurons from the previous levels to the succeeding ones. This can be compared to how high-level thinking explores all possible routes from input to output. The downsized picture is then placed in a single list, obtained by passing it through two layers of convolution, ReLU, and pooling before converting it into a single file or vector. Figure 20.5 shows pre-processing and training results on the input image.

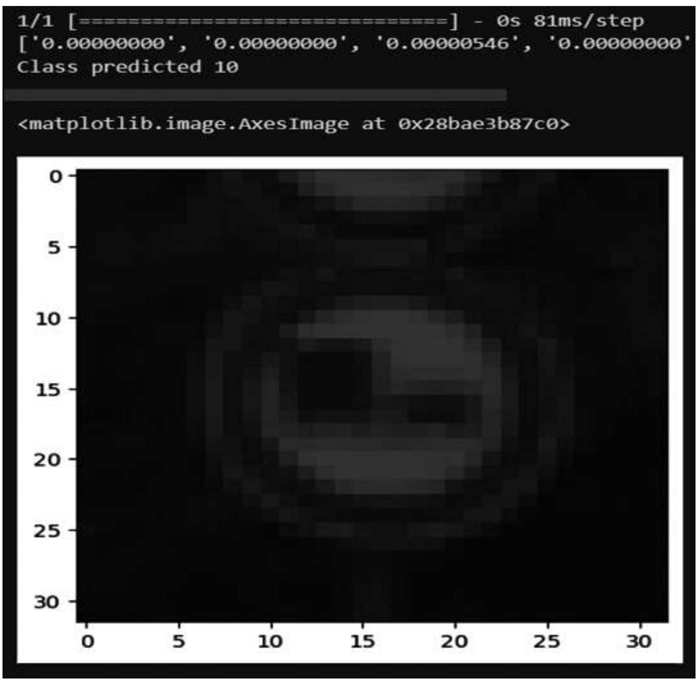


FIGURE 20.5 Pre-processing and training.

20.3.6 RESNET

The two possibly most well-known backbone networks nowadays are ResNet and DenseNet, and they are used in most deep learning-based systems. Despite their outstanding popularity and competitive performance, both have intrinsic flaws. ResNet's representational capacity is constrained by the identity shortcut that stabilizes training, DenseNet, on the other hand, has more capacity because of multi-layered feature combining. Merging densely, on the contrary, introduces a new challenge by needing expensive GPU RAM and additional training time. This contributes to the fact that choosing between ResNet and DenseNet is not an easy decision. To analyze them, this study offers a unified perspective on dense summation, which makes it easier to perceive their fundamental differences. In addition, we propose dense weighted normalized shortcuts. The dense shortcut we propose follows the design philosophy of ResNet and the fundamental design of DenseNet. The outcome of the experiment demonstrates that the DSNet suggested ResNet and DenseNet while using much fewer computational resources on a few benchmark datasets.[11]

The residual block serves as ResNet's primary base component. The complexity of processing increases as we delve further into the network with several levels. In their stead, we attempt to fit a residual mapping. Each layer of these stacks attempts to uncover some underlying mapping of the desired function which is shown in Figure 20.6.

Because the source of these blocks' input is merely an input, we will utilize deep neural layers to test if they fit some residual of our h(x) - x rather than the intended function h(x). Essentially, the input is just accepted and sent through as an identity to the conclusion of this section, where the skip link is taken. If no weight layers were present, the input would just be the identity. The outcome would be the same if we did not use further weightage layers in order to learn some delta for some residual from our x.[12]

Here, because there are so many levels in the network, understanding H (X) gets progressively more difficult as we go deeper. As a result, in this example, we made use of the skip link and discovered f(x) as the input of x. As a result, f(x) is referred to as an output.

ResNet organizes all blocks as a stack, quite closely. The advantage of the highly complex design is, that it enables approximately 150 layers, which are then periodically built on top of one another. Furthermore, we employ stride as two up to spatially down samples and increase the number of filters. The thousandth layer was eventually fully connected to the output classes. Figure 20.7 shows the Code snippet of ResNet50.

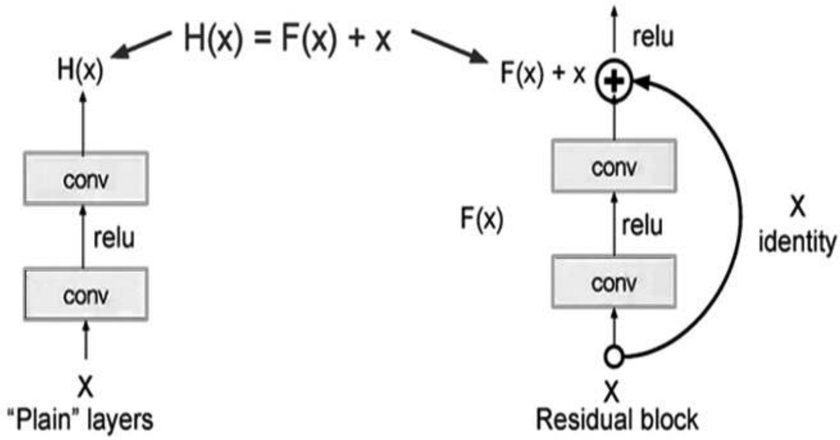


FIGURE 20.6 ResNet system architecture.

```
#ResNet50

base_model=tf.keras.applications.resnet50.ResNet50(weights=r'imagenet', include_top=False,input_shape=(IMG_SIZE,IMG_SIZE,3))

x=(base_model.output)

x = tf.keras.layers.GlobalAveragePooling2D()(x)

predictions = tf.keras.layers.Dense(43, activation="softmax")(x)

ResNet50_model = tf.keras.models.Model(inputs=base_model.input, outputs=predictions)
ResNet50_model.compile(optimizer='adam', loss='CategoricalCrossentropy', metrics=['accuracy']) #compiling the DenseNet model
history=ResNet50_model.fit(x=X_train,y=Y_train,batch_size=32,validation_data=(X_test,Y_test),epochs=5) #Fitting the DenseNet model
```

FIGURE 20.7 Code snippet of ResNet50.

20.3.7 DENSENET

Dense blocks are the basic units of DenseNet. Layers are closely interwoven within these blocks: each layer receives input from the output maps of elements of previous levels. Since each layer receives more supervision from the layer above it because of the excessive reuse of residuals, deep supervision results, which makes the network more powerful.[10]

Each layer in the DenseNet design is strongly connected to the layers above it, increasing the use of the residual mechanism. Because they are all shared through common knowledge, the learned characteristics are not replicated due to the model's compactness, to train densely coupled deep networks where the short connections brought on by implicit deep supervision allow the gradient to flow back more readily.

In a significant part, DenseNet was developed to address the deteriorating accuracy brought on by high-level neural networks' disappearing gradient. Simply said, the increased distance between the input and output layers causes the information to disappear before it reaches its destination. Figure 20.8 shows the Code snippet of DenseNet.

```
#DenseNet model

base_model=tf.keras.applications.densenet.DenseNet121(weights=r'imagenet', include_top=False,input_shape=(IMG_SIZE,IMG_SIZE,3))

x=(base_model.output)

x = tf.keras.layers.GlobalAveragePooling2D()(x)

predictions = tf.keras.layers.Dense(43, activation="softmax")(x)

densenet_model = tf.keras.models.Model(inputs=base_model.input, outputs=predictions)
densenet_model.compile(optimizer='adam', loss='CategoricalCrossentropy', metrics=['accuracy']) #compiling the DenseNet model
history=densenet_model.fit(x=X_train,y=Y_train,batch_size=32,validation_data=(X_test,Y_test),epochs=5) #Fitting the DenseNet model
```

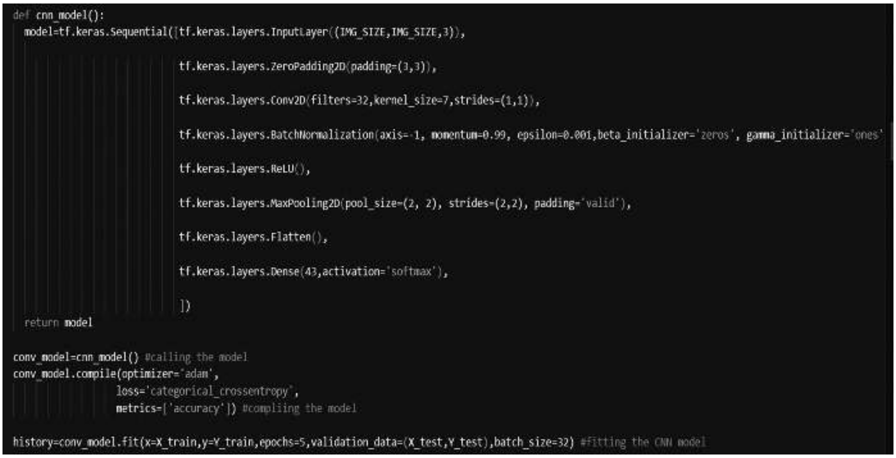
FIGURE 20.8 Code snippet of DenseNet.

20.4 RESULTS AND DISCUSSION

In the base CNN model, the input layer is of shape 200 × 200 followed by the zero-padding, Next, a 2D convolution network layer with momentum 0.99 and epsilon 0.001 was added, followed by a batch normalization layer, pooling, and the ReLU Activation Layer are next, and the SoftMax activation function is followed by a Flatten layer and a Dense layer which is shown in Figure 20.9.

Then, we assess the precision and loss of the trained model, and it can be given as model loss and model accuracy graph as shown in Figures 20.10 and 20.11, respectively.

Then, we use DenseNet where we download the pre-trained weights for the DenseNet121 model, change the input shape for the top layer, add the pooling layer, and then add a dense network of SoftMax activation layer. Then, we use REsNet50 to reduce the gradient problem as it stacks all those convolutional layers that have no significance or less significance. Finally, we train and test the model with a validation accuracy of 0.9048. Figures 20.12 and 20.13 show output for these stages.



```
def cnn_model():
  model=tf.keras.Sequential([tf.keras.layers.InputLayer((IMG_SIZE,IMG_SIZE,3)),

                              tf.keras.layers.ZeroPadding2D(padding=(3,3)),

                              tf.keras.layers.Conv2D(filters=32,kernel_size=7,strides=(1,1)),

                              tf.keras.layers.BatchNormalization(axis=-1, momentum=0.99, epsilon=0.001,beta_initializer='zeros', gamma_initializer='ones'

                              tf.keras.layers.ReLU(),

                              tf.keras.layers.MaxPooling2D(pool_size=(2, 2), strides=(2,2), padding='valid'),

                              tf.keras.layers.Flatten(),

                              tf.keras.layers.Dense(43,activation='softmax'),

                              ])
  return model

conv_model=cnn_model() #calling the model
conv_model.compile(optimizer='adam',
                   loss='categorical_crossentropy',
                   metrics=['accuracy']) #compliing the model

history=conv_model.fit(x=X_train,y=Y_train,epochs=5,validation_data=(X_test,Y_test),batch_size=32) #fitting the CNN model
```

FIGURE 20.9 CNN model.

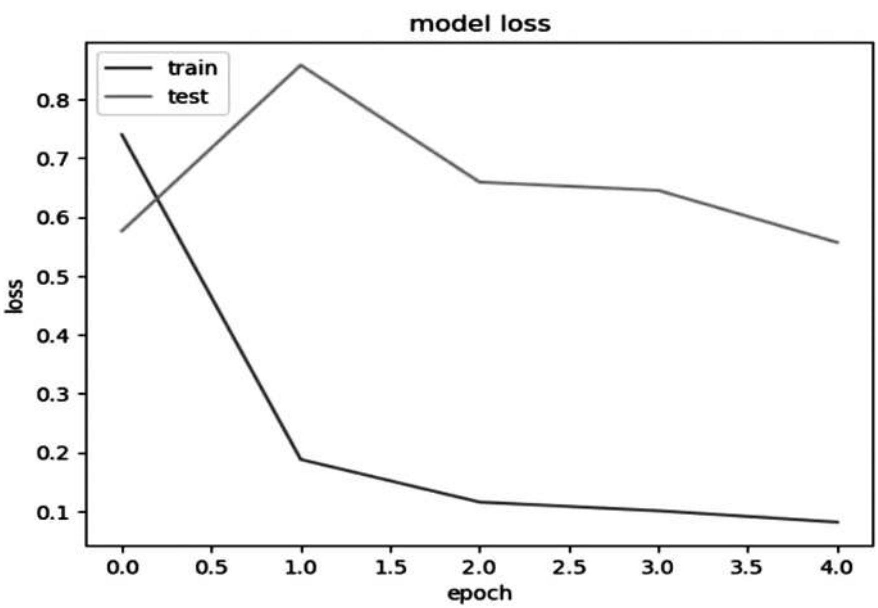


FIGURE 20.10 Model loss.

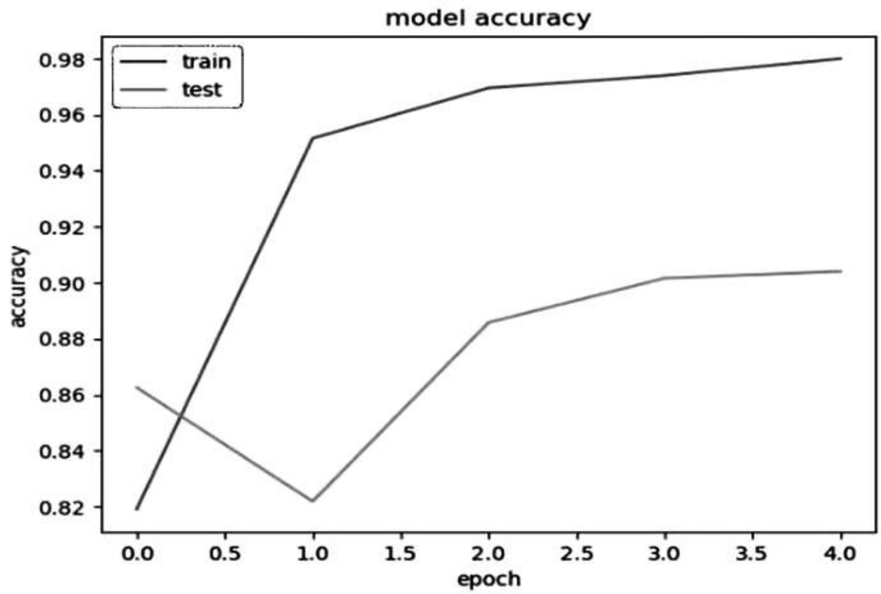


FIGURE 20.11 Model accuracy.

FIGURE 20.12 Home page.

20.5 CONCLUSION

Based on the GTSRB standard dataset, traffic sign classification and recognition tests are carried out in this study. The CNN-based classification model is continuously trained and tested to obtain the favorable prediction and precise identification of traffic signs. Focus is placed on CNN research and the capacity of deep learning models to categorize pictures. The model's implementation result after training and testing indicates that it is around 94.01% accurate and correct. We used the value Epoch = 20 to fit the model. Forty-three classes were created from 35,000 photos. As demonstrated in Figure 20.5, accurate selection and extraction of picture features using the CNN algorithm contributed to improving the deep learning model's accuracy. When a network is flattened, every neuron in the topmost layer is connected to every neuron in the layers below and above it. This represents higher-level thinking where all possible connections between the input and the outcome are considered. Once it has been through two layers of convolution, ReLU, and pooling and has been transformed into a single file or a vector, take the downsized picture and place it in the single list. In Table 20.2, we can give a comparison between various deep learning-based classification techniques.

FIGURE 20.13 Output UI.

TABLE 20.2 Comparison Between Various Deep Learning-Based Classification Techniques.

Neural network architecture	Description	Accuracy
LeNet-5	A classic CNN architecture consisting of convolutional and fully connected layers.	~98.8%
AlexNet	A deep CNN architecture with multiple convolutional and fully connected layers.	~99.3%
VGG-16	A deep CNN architecture with 16 convolutional layers and fully connected layers.	~99.5%
GoogLeNet (Inception)	A deep CNN architecture with parallel and multi-scale convolutional operations.	~99.1%
ResNet-50	A deep residual network with 50 convolutional layers, using skip connections.	~99.6%
DenseNet-121	A densely connected CNN architecture with feature reuse among layers.	~99.7%
MobileNetV2	A lightweight CNN architecture optimized for mobile and embedded devices.	~99.0%
EfficientNet-BO	A scalable CNN architecture that balances model size and computational cost.	~99.2%

20.6 FUTURE SCOPE

The TSR technique using deep learning technology can be implemented and used for further research on real-time detection of traffic signs on roads. The camera can be mounted on the front side shield of the car and capture images of traffic sign boards, these images can be supplied as input to the TSR model which can classify the images and display the output label.

KEYWORDS

- **CNN**
- **max pooling**
- **ReLU**
- **Keras**
- **SoftMax**
- **traffic sign classification**
- **hidden layers**

REFERENCES

1. Megalingam, R. K.; Thanigundala, K.; Musani, S. R.; Nidamanuru, H.; Gadde, L. Indian Traffic Sign Detection and Recognition using Deep Learning. *Int. J. Transp. Sci. Technol.* **2022**. https://doi.org/10.1016/j.ijtst.2022.06.002.
2. Cao, J.; Song, C.; Peng, S.; Xiao, F.; Song, S. College of Automotive Engineering, Jilin University, China, Improved Traffic Sign Detection and Recognition Algorithm for Intelligent Vehicles, Changchun 130022, China.
3. Höferlin, B.; Heidemann, G. Selection of an Optimal Set of Discriminative and Robust Local Features with Application to Traffic Sign Recognition, May 21, 2014.
4. Bichkar, M.; Bobhate, S.; Chaudhari, S. Department of Computer Engineering, Datta Meghe, India, [Online] May 18, 2021. http://ijsrcseit.com.
5. Hasegawa, R.; Iwamoto, Y.; Chen, Y. W. In *Robust Detection and Recognition of Japanese Traffic Sign in the Complex Scenes Based on Deep Learning*, IEEE 8th Global Conference on Consumer Electronics (GCCE), 2019.
6. Prakash, S.; Ayyalu, S.; Vigneshwaran, D.; Jayanthi Sree, S. In *Traffic Sign Recognition using Deep Learning for Autonomous Driverless Vehicles*, Proceedings of the Fifth International Conference on Computing Methodologies and Communication (ICCMC 2021) IEEE Xplore Part Number: CFP21K25-ART, 2021.
7. Ellahyani, A.; El Ansari, M. In *Complementary Features for Traffic Sign Detection and Recognition*, 2016 IEEE/ACS 13th International Conference of Computer Systems and Applications (AICCSA), 2016; pp 1–6.
8. Jin, Y.; Fu, Y.; Wang, W.; Guo, J.; Ren, C.; Xiang, X. Multi-Feature Fusion and Enhancement Single Shot Detector for Traffic Sign Recognition. *IEEE Access* **2020,** *8*, 38931–38940.
9. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K. Q. In *Densely Connected Convolutional Networks*, IEEE Conference on Computer Vision and Pattern Recognition, 2017; pp 4700–4708.
10. Zhang, C.; Benz, P.; Argaw, D. M.; Lee, S.; Kim, J.; Rameau, F.; Bazin, J. C.; Kweon, I. ResNet or DenseNet, Introducing Dense Shortcuts to ResNet, Oct 23, 2020.
11. Patel, K. Architecture Comparison of AlexNet, VGGNet, ResNet, Inception, DenseNet-beb8b116866d, Mar 8, 2020.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

CHAPTER 21

AN OPTIMIZED SOIL FEATURES PREDICTION USING SATELLITE IMAGE DATABASE

MALLEKEDI ANAND[1], ANUJ JAIN[1], and MANOJ KUMAR SHUKLA[2]

[1]*School of Electronics and Electrical Engineering, Lovely Professional University, Punjab, India*

[2]*Department of Robotics and Automation, Symbiosis Institute of Technology, Pune, India*

ABSTRACT

The world population is increasing rapidly, but supplying food for such a population is very difficult. If the farmers follow traditional techniques, crop yield is very low. So, in order to overcome this problem, the farmers have to follow advanced techniques where suitable crop fields are used. For this purpose, satellite images are used. The satellite images are preprocessed, and features are extracted by using optimization techniques. In this work, hence, to find the soil's chemical properties and suitable crops, the coati optimization algorithm (COA) was developed with the required features. The imported satellite images were initially filtered and entered into the classification phase to forecast the present chemical features and suitable crops in specific soils. After detecting the chemical features, soil textures like sandy, silt, and clay were categorized; crops like jute, rice, lentil, and potato were considered. Finally, the proposed model is executed in the MATLAB environment and has gained better results by achieving the lowest error rate of 6% and a high prediction score of 94%. The proposed model has achieved

Applied Soft Computing Techniques: Theoretical Principles and Practical Applications. Samarjeet Borah, Ratna Raja Kumar Jambi, Sharifah Sakinah Syed Ahmed, & Mahendra Prabhakar Deore (Eds.)
© 2025 Apple Academic Press, Inc. Co-published with CRC Press (Taylor & Francis)

exact soil texture analysis for complex satellite images. The gained texture analysis outcome is quite better than other existing models.

21.1 INTRODUCTION

Soils can be classified by considering the real-time data by using satellite images over the last 50 years,[1] and because of this, the crop can be identified easily.[2] In recent times, there has been a growing influence about the environmental hazards as a result of the excessive pesticides and fertilizers in the farmlands, which rate the spatially variable landscapes.[3] The most essential indicator of the fertility and quality of the soil, food security, and agricultural productivity is termed soil organic matter (SOM), which is the necessary component of the worldwide carbon cycle.[4] SOM needs increased chemical application rates of some herbicides because of the absorption of herbicides on the cation exchange complex. It elaborates on the organic matters of diverse nature and characteristics.[5] It is the core variable for the evaluation of agricultural or farmland management. The quick and minimum cost determination of SOM spatial distribution offers the timely management of the agricultural lands.[6] The essential negative correlation between the SOM and spectral reflectance of the soil predicts the SOM spatial distribution. The SOM management also influences the environmental impact of the soil organic carbon (SOC) storage capacity.[7] This minimizes the effects of CO_2 on global warming and climatic changes. Therefore, soil management needed the evaluation of the SOM and its variability in spatial and temporal characteristics.[8] Also, this study is essential for the efficient land utilization and protection of surroundings. An accurate determination of SOM contents will develop sufficient knowledge on maintaining the soil nutrients and crops.[9]

In the current prominent technological improvements, the adoption of sensitive agriculture is a necessary process. For this adoption, the development of numerous technologies for the computation of SOM contents reduces the long-term soil analysis and the cost of mapping of SOM.[10] The study of the characteristics of the soil through the traditional approaches requires suitable temporal and spatial resolution, which is not enough, and a high cost to adapt.[11] So, the study of the soil data using the remote sensing data is introduced for the assessment of the soil features, and it becomes an irreplaceable method for the identification of spatial patterns at a high resolution and minimizes the notable fieldwork.[12]

21.2 LITERATURE REVIEW

Soil chemical property prediction is the required objective for the agricultural land to cultivate the different crops. Chaitanya B. Pande et al.[13] have used the wavelet transforms to forecast the soil chemical properties. Hence, to proceed with this process, the satellite Landsat image database has been utilized.

To find the appropriate techniques to estimate the salinity score, a detailed review has been conducted by D. L. Corwin and E. Scudiero.[14] From the review assessment, it was verified that the knowledge-based spatial model has earned the finest results in analyzing the soil types of features. Also, the mapping procedure has helped to attain the graphical outcomes. However, it is complex in design and takes more time to execute the process.

JiePeng et al.[15] have designed the Cubist regression model to analyze the soil features on the basis of climate changes. Here, dual models have been employed to measure the soil parameters: soil sampling and regression models. In the soil sampling, the satellite images were utilized. Moreover, the study area was adopted from China. However, it needs more memory spaces for execution.

Xiangtial et al.[16] developed the SOC prediction model using SVM, RF, and the back propagated network. This method used the hyperspectral data for the analysis. Here, the prediction is performed through the selection of three optimal bands and indices. It attained the highest prediction accuracy results and also acquired a very low rate of root mean square error (RMSE). However, the increase in decomposition scale can affect the accuracy results.

The different spectral characteristics aiming the soil samples may vary the SOM contents. So, Yilin et al.[17] proposed a new grouping strategy for the prediction of SOM content in the soil. It establishes the RF model for the estimation, and DT performs the grouping strategy for the prediction model. Here, the competitive adaptive reweighted sampling minimizes the input counts and increases the accuracy. However, at global scale data, the prediction accuracy gets decreased due to the soil's varying spectral heterogeneity.

Xin et al.[18] introduced the regional-based SOM prediction model. It utilized the multitemporal satellite imagery data, and the optimal input variables for the system are identified by the time interval among the snow, plowing, and rainfall. The multidate images have diminished the influence of variation in soil moisture content. It improved the accuracy and provided support for digital soil mapping. However, multiple representative sample points are needed in this model due to the considerable spatial variability of soil in order to guarantee a respectable degree of accuracy.

Zipeng et al.[19] introduced the partial least square SVM for the prediction of the SOM content. This model utilizes the newly created three-band index using band optimization for the accurate prediction of organic matter in soil. It increases the estimation and the sensitivity accuracy and maximizes the system performance. The three-band index has a robust capacity for biochemical parameter estimation. However, the increase in derivative order increases the noise level.

Lifei et al.[20] suggested the hyperspectral inversion model for the SOM content determination based on the AdaBoost algorithm. The characteristic band was cantered into the suggested model. By performing the grid search, the developed model attains higher accuracy results with maximum precision at high speed. Also, it leads the way for the improvement of precision agriculture. However, the data quality may affect the model's performance.

On the other hand, the OLI satellite was employed by Abdelgadir Abuelgasim and Rubab Ammad[21] to map the soil attribute from the image data. The designed salinity-index approach has earned the highest mapping exactness score in analyzing the soil attributes. Moreover, its performance is validated with other empirical models. This index model is verified with semiarid and nonagricultural surfaces. It required more time to implement.

Analyzing the crop parameters is an important function in choosing soil types for planting. So, Bikash Ranjan Parida et al.[22] have designed the parameter index to measure the biochemical features in the paddy. Several biological parameters were calculated from the trained satellite image database. Here, the decision models were used to measure the plant growth and soil feature monitoring. In addition, the yield rate has been measured on the basis of biochemical parameters. However, more resources have been required to execute the process. In some cases, the prediction rate became less because of image complexity.

21.3 DATASET

A dataset has been required to analyze the proposed model's working function. After analyzing previous works, it was discovered that the primary issue in estimating soil type and plant biological parameters is the amount of time and resources required. This problem is caused by the complexity of the images used in the process. When images are too complex, more resources are needed, resulting in longer execution times and lower prediction accuracy. To address this issue, a suitable filtering approach is necessary to filter the trained data. Accurate soil feature estimation is crucial for achieving high yields in the agricultural field under the correct climatic conditions. However, the complex nature of imaginary data and inaccurate filtering methods pose a challenge for ML models in classifying these features. To address this issue, a preprocessing function has been implemented in the deep learning framework to filter out noisy content before the feature analysis process. This step is crucial for achieving a higher prediction score. Hence, the soil satellite images were utilized. Moreover, the sample satellite images are shown in Figure 21.1.

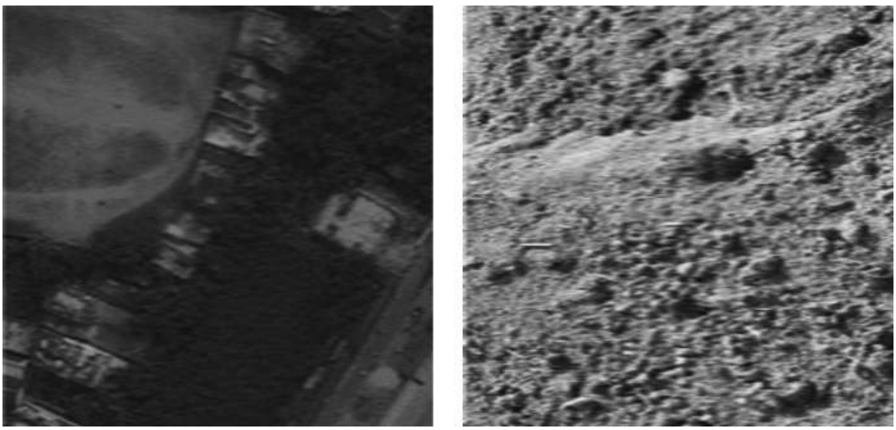


FIGURE 21.1 Satellite images.

21.4 DESCRIPTION OF THE SYSTEMS

The satellite images are used to measure the soil features and plant biological parameters to find the suitable crop for different soils. The proposed research is helpful for agricultural applications to cultivate crops with better yields.

21.4.1 NOVELTY OR THE PROPOSED WORK

Initially, the images were preprocessed and imported to the classification module for predicting the soil features and plant features. Hence, an improved coati convolutional neural-based organic prediction (CCNbOP) system has been framed for the soil feature analysis and plant biological parameter estimation processes. Initially, the satellite images have been gathered and imported into the MATLAB system. The CNN is used for classifying the satellite images through extracted features and provides better classification by combining with the COA algorithm. Figure 21.2 shows the basic structure of the classification method.

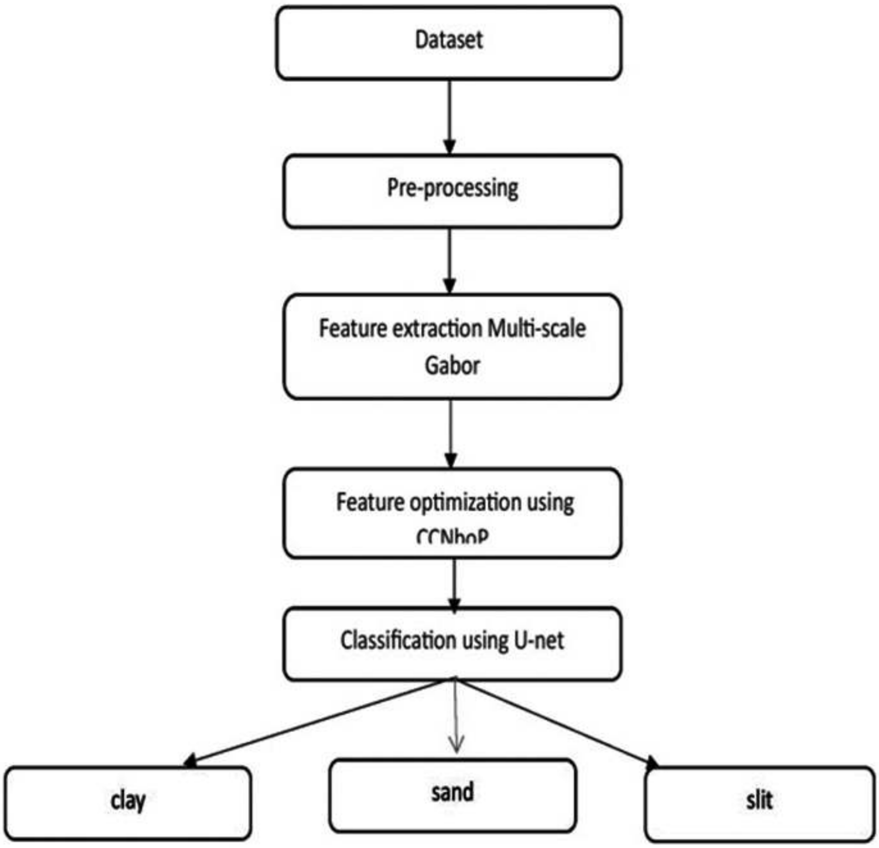


FIGURE 21.2 Methodology.

21.4.1.1 PREPROCESSING

The proposed method is preprocessing the satellite images, which is considered the first step. The unwanted noise components in the images are removed by using preprocessing. At last, the interference of irrelevant regions is reduced by using ROI generation.

21.4.1.2 EXTRACTING THE FEATURES

The selected features represent the uniqueness of the images for defining a class. The extraction of features from many frequencies or scales is aligned at different angles with the help of filtering.

21.4.1.3 FEATURE SELECTION

The most challenging task is to select the correct number of features for better classification. The problem of overextraction and underextraction can be formed by an incorrect choice. PPCA is an effective tool for reducing the dataset dimension and allows the reconstruction of the optimal shape. While retaining most of the variations, the dataset contains a vast number of interrelated variables. For reducing the feature dimension, projection vectors are used, which contribute to the highest covariance. In this work, K-eigenvectors are selected from the feature input by PPCA for choosing the best features from the whole extraction of features.

21.4.1.4 FEATURE OPTIMIZATION

The coati neural-based organic prediction is a metaheuristic algorithm, which will be focusing on attacking and hunting the coatis. The obtained features are optimized by the proposed algorithm. Important metrics are used to identify the performance of the algorithm.

21.5 RESULTS AND DISCUSSION

21.5.1 PERFORMANCE PARAMETERS

In this section, the outcome of the proposed method is evaluated by different parameters such as accuracy, sensitivity, RMSE, MAE, and error rate. The metrics used for this parameter are described as true positive (TP), true negative (TN), false positive (FP), and false negative (FN).

Sensitivity: Sensitivity is described in eq 21.1 as the ratio of the TP number to the addition of both TP and FN.

$$Sensitivity = \frac{No.\,of\ TP}{No.\,of\ TP + No.\,of\ FN} \times 100 \tag{21.1}$$

Accuracy: By using the measure of sensitivity and specificity, accuracy can be evaluated. This measurement can be explained in the following eq 21.2.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \tag{21.2}$$

21.5.2 RMSE AND MAE

RMSE and MAE are calculated to check the proposed BUFPS in predicting the soil parameters and suitable crop yields. If the method has gained the lowest MAE and RMSE, it has earned the finest R^2 and prediction rate.

$$\text{RMSE} = \sqrt{\sqrt{\sqrt{\frac{1}{3}}}} \sum (pj - oj) \tag{21.3}$$

$$\text{MAE} = \frac{|(zi - zj)|}{O} \tag{21.4}$$

The metric RMSE has been calculated to find the root values of the calculated MSE.

21.5.3 PERFORMANCE ANALYSIS

In the following section, various existing methods are compared with the proposed method for validating the performance results in kidney image classification. The method considered half of the data for training and the remaining half for the testing process (i.e., 80:20 sample data). The parameter values such as accuracy and sensitivity are computed in Table 21.1.

TABLE 21.1 The Values of Several Parameters of Proposed and Existing Methods.

	UR	MR	BP	NN	RF	UNET	BUFPS	Proposed
ACC	0.5625	0.5625	0.8125	0.8125	0.8125	0.92041	96.1	97.2
SEN	0	0.333333	0	0	0	0.87765	96	96.5
RMSE	1.414	1.454	1.159	0.814	1.277	1.515	0.492	0.391

21.6 CONCLUSIONS

The obtained results are shown in Table 21.1. By observing the table, it is clearly known that the obtained results are more accurate than the existing ones. In the future, the proposed algorithm will be implemented by using real-time data. This article presents a new coati optimization that predicts soil types in Tarakeswar and suggests suitable crops by validating biological standards with soil types. The model's performance was measured in two phases, before and after applying bat functions. In the preprocessing phase, a satellite image was filtered and modeled in the hidden layer of the UNet. The preprocessed image was then entered into the feature analysis and soil type specification phase. The model achieved an accuracy of 97.2%, with the validation of key metrics such as RMSE, MAE, and R2. In the future, specifying the farming region and classifying the specific crop yield range will offer a useful outcome. Additionally, hyperparameter tuning strategies will provide improved results beyond the current work.

KEYWORDS

- **satellite images**
- **soil textures**
- **crops**
- **coati optimization**
- **classification**

REFERENCES

1. Prabhavathi, V.; Kuppusamy, P. In *A study on Deep Learning based Soil Classification, 2022 IEEE 4th International Conference on Cybernetics, Cognition and Machine Learning Applications (ICCCMLA)*, Goa, India, 2022, pp 428–433, doi: 10.1109/ICCCMLA56841.2022.9989293.
2. Hébert, M. P.; Fugère, V.; Gonzalez, A. The Overlooked Impact of Rising Glyphosate Use on Phosphorus Loading in Agricultural Watersheds. *Front. Ecol. Environ.* **2019,** *17*(1), 48–56.
3. Dinesh, G. K.; Sinduja, M.; Priyanka, B.; Sathya, V.; Karthika, S.; Meena, R. S.; Prasad, S. Enhancing Soil Organic Carbon Sequestration in Agriculture: Plans and Policies. In *Plans and Policies for Soil Organic Carbon Management in Agriculture*; Springer Nature Singapore: Singapore, 2022; pp 95–121.
4. Hoffland, E.; Kuyper, T. W.; Comans, R. N.; Creamer, R E. Eco-Functionality of Organic Matter in Soils. *Plant Soil.* **2020,** *455,* 1–22.

5. Bakhshandeh, E.; Hossieni, M.; Zeraatpisheh, M.; Francaviglia, R. Land Use Change Effects on Soil Quality and Biological Fertility: A Case Study in Northern Iran. *Eur. J. Soil Biol*. **2019,** *95*, 103119.
6. Xu, X.; Du, C.; Ma, F.; Qiu, Z.; Zhou, J. A Framework for High-Resolution Mapping of Soil Organic Matter (SOM) by the Integration of Fourier Mid-Infrared Attenuation Total Reflectance Spectroscopy (FTIR-ATR), Sentinel-2 Images, and DEM Derivatives. *Remote. Sens*. **2023,** *15* (4), 1072.
7. Chen, S.; Lin, B.; Li, Y.; Zhou, S. Spatial and Temporal Changes of Soil Properties and Soil Fertility Evaluation in a Large Grain-Production Area of Subtropical Plain, China. *Geoderma*. **2020,** *357,* 113937.
8. Zhang, H.; Wang, L.; Tian, T.; Yin, J. A Review of Unmanned Aerial Vehicle Low-Altitude Remote Sensing (UAV-LARS) Use in Agricultural Monitoring in China. *Remote. Sens.* **2021,** *13* (6), 1221.
9. Raj, M.; Gupta, S.; Chamola, V.; Elhence, A.; Garg, T.; Atiquzzaman, M.; Niyato, D. A Survey on the Role of Internet of Things for Adopting and Promoting Agriculture 4.0. *J. Netw. Comput. Appl*. **2021,** *187,* 103107.
10. Gao, F.; Anderson, M.; Daughtry, C.; Karnieli, A.; Hively, D.; Kustas, W. A Within-Season Approach for Detecting Early Growth Stages in Corn and Soybean Using High Temporal and Spatial Resolution Imagery. *Remote Sens. Environ.* **2020,** *242,* 111752.
11. Jiang, Y.; Zhang, L.; Yan, M.; Qi, J.; Fu, T.; Fan, S.; Chen, B. High-Resolution Mangrove Forests Classification with Machine Learning Using Worldview and UAV Hyperspectral Data. *Remote. Sens*. **2021,** *13* (8), 1529.
12. Tan, K.; Ma, W.; Wu, F.; Du, Q. Random Forest–Based Estimation of Heavy Metal Concentration in Agricultural Soils with Hyperspectral Sensor Data. *Environ. Monit. Assess*. **2019,** *191,* 1–4.
13. Segarra, J.; Buchaillot, M. L.; Araus, J. L.; Kefauver, S. C. Remote Sensing for Precision Agriculture: Sentinel-2 Improved Features and Applications. *Agronomy* **2020,** *10* (5), 641.
14. Zhang, N.; Yang, G.; Pan, Y.; Yang, X.; Chen, L.; Zhao, C. A Review of Advanced Technologies and Development for Hyperspectral-Based Plant Disease Detection in the Past Three Decades. *Remote. Sens*. **2020,** *12* (19), 3188.
15. Nawar, S.; Mouazen, A. M. On-line vis-NIR Spectroscopy Prediction of Soil Organic Carbon Using Machine Learning. *Soil Tillage Res*. **2019,** *190,* 120–127.
16. Zhao, Z.; Yang, Q.; Sun, D.; Ding, X.; Meng, F. R. Extended model prediction of high-resolution soil organic matter over a large area using limited number of field samples. *Comput. Electron. Agric*. **2020,** *169,* 105172.
17. Song, J.; Gao, J.; Zhang, Y.; Li, F.; Man, W.; Liu, M.; Wang, J.; Li, M.; Zheng, H.; Yang, X.; Li, C. Estimation of Soil Organic Carbon Content in Coastal Wetlands with Measured VIS-NIR Spectroscopy Using Optimized Support Vector Machines and Random Forests. *Remote. Sens.* **2022,** *14* (17), 4372.
18. Chen, D.; Chang, N.; Xiao, J.; Zhou, Q.; Wu, W. Mapping Dynamics of Soil Organic Matter in Croplands with MODIS Data and Machine Learning Algorithms. *Sci. Total Environ*. **2019,** *669,* 844–855.
19. Brook, A.; De Micco, V.; Battipaglia, G.; Erbaggio, A.; Ludeno, G.; Catapano, I.; Bonfante, A. A Smart Multiple Spatial and Temporal Resolution System to Support Precision Agriculture from Satellite Images: Proof of Concept on Aglianico Vineyard. *Remote Sens. Environ.* **2020,** *240,* 111679.

20. Hasan, M. E.; Nath, B.; Sarker, A. R.; Wang, Z.; Zhang, L.; Yang, X.; Nobi, M. N.; Røskaft, E.; Chivers, D. J.; Suza, M. Applying Multi-Temporal Landsat Satellite Data and Markov-Cellular Automata to Predict Forest Cover Change and Forest Degradation of Sundarban Reserve Forest, *Bangladesh*. *Forests*. **2020,** *11* (9), 1016.
21. Jaskulak, M.; Grobelak, A.; Vandenbulcke, F. Modelling Assisted Phytoremediation of Soils Contaminated with Heavy Metals–Main Opportunities, Limitations, Decision Making and Future Prospects. *Chemosphere* **2020,** *2*.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

PART V

NATURAL LANGUAGE PROCESSING

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

CHAPTER 22

A NOVEL TECHNIQUE ENABLING TEXT-TO-SPEECH SIGNAL CONVERTING SYSTEM USING RASPBERRY PI

PRABHU CHAKKARAVARTHY[1], G. PARTHASARATHY[2], P VISU[3], A. MOHAN[1], and SANJU DAS[4]

[1]*Department of Computer Science and Engineering, SIMATS School of Engineering, Saveetha University, Chennai, Tamil Nadu, India*

[2]*Department of Computer Science and Engineering, Reva University Bangalore, Karnataka, India*

[3]*Department of Computer Science and Engineering, Velammal Engineering College, Chennai, Tamil Nadu, India*

[4]*Department of Computer Science, Assam University Silchar, Assam, India*

ABSTRACT

Raspberry Pi is highly known for technological advancements in speeding up. However, the adoption of technology is generally fairly modest. It is common knowledge that most individuals find it challenging to read text from paper and books. We thus offer a mechanism that separates the text images and transmits them to the Raspberry Pi. The Raspberry Pi reads the content from the speaker after processing the text pictures. This makes it possible to convert text to speech.

Applied Soft Computing Techniques: Theoretical Principles and Practical Applications. Samarjeet Borah, Ratna Raja Kumar Jambi, Sharifah Sakinah Syed Ahmed, & Mahendra Prabhakar Deore (Eds.)
© 2025 Apple Academic Press, Inc. Co-published with CRC Press (Taylor & Francis)

22.1 INTRODUCTION

Voice/speech synthesis is the process of creating computer systems that synthesize written text. It is a method for a computer system to convert written text communications into speech utilizing a telephone or a microphone. Even though speech recognition technology is still a new field, most developers are still unfamiliar with it. However, developers want to learn and exploit a number of subtle and potent capabilities offered by computerized speech as understanding and implementing the fundamental operations of speech synthesis and speech recognition just takes a few minutes. Automatic voice synthesis is one of the most rapidly developing fields in speech science and engineering. It is necessary to create the speech recognition capabilities after installing the system. As a new generation of computing technology is supported, it is necessary to create IVR (interactive voice response) systems, which represent the most significant advancement in human–machine interaction, after implementing the functionality of speech recognition. The main idea behind text-to-speech (TTS) technology is to generate synthetic speech from the input text.

The speech synthesis can be done in a variety of ways:

a. Easy voice recording and playback as needed.
b. Breaking down speech into 30–50 phonemes (fundamental linguistic components) and reassembling them to form a fluent speech pattern.
c. Approximately 400 diaphones were employed (phrases were separated at the phoneme center rather than the transition).

The primary feature of modern speech synthesis systems is the artificial simulation or production of human speech. How closely artificial speech may imitate actual human speech is referred to as its "naturalness." The purpose of this research is to improve the artificial intelligence and naturalness of speech synthesis systems.

22.2 LITERATURE SURVEY

In terms of text-to-speech (TTS) and voice conversion (VC), this work proposes novel approaches for high-quality statistical parametric speech synthesis based on modulation spectrum (MS). Although statistical parametric voice synthesis has advantages over concatenative speech synthesis, the synthetic speech quality is still subpar when compared to concatenative speech synthesis or actual speech. One of the main obstacles to effecting quality loss is the oversmoothing effect, which is often present in the generated speech parameter trajectories. It has been proven advantageous to maintain the global variance (GV) of the generated speech parameter trajectories at a level that is comparable to that of natural speech. The oversmoothing effect is known to be strongly associated with the GV. However, there is still a difference in the quality of natural and synthetic speech. Additionally, the correction takes time because each phoneme that is produced must be checked separately. In this paper, we present a method for finding contrastive phonetization hypotheses to automatically detect problems with the conversion of graphemes to phonemes. A signal-dependent phonetization system is constructed using a lattice-based forced alignment technique. An additional sequence-to-sequence neural network model is utilized to demonstrate a context-dependent grapheme-to-phoneme conversion paradigm. A French dataset was used to show that our system has the potential to identify 86.3 percent of the errors made by a commercial grapheme-to-phoneme system. Furthermore, it is maintained that less than 10% of the evaluated entire data contains error markings. As a result, the amount of time needed to manually examine phonemes can be considerably decreased without compromising the accuracy of the phonemic transcription.[3,4] Several works on speech recognition using soft computing and machine learning are found in the literature.[24,25]

To ensure 1:1 mapping for each utterance with its mirror text file and to reduce the signal-to-noise ratio (SNR), vocal leveling, normalization, syllable splitting, and merging the recordings were edited under the free digital audio workstation (DAW) audacity using continuous frequency profiling techniques. The performance of the proposed model was evaluated using specialized transcript files; the files were generated from an audio dataset of 10 different speakers of recorded speech data, including both males and females. According to the experimental results, it is observed that for the test dataset, the proposed model obtains an overall accuracy of 71.7%.

22.3 PROPOSED SYSTEM

The difficulty of the existing system is overcome with the proposed system, as shown in Figure 22.1. The system's brain is identified as "Raspberry Pi". Here, an image of the text is being taken with a camera. The image processing unit receives the captured image from the camera. The filter takes in the output of the image processing and filters out noise signals. The output of the filter is sent to the edge reduction component. The output is delivered to the background separation unit. Finally, the result is forwarded to the OCR (optical character recognition). Input for the Raspberry Pi comes from the OCR output. The Pi recognizes the visual content and produces an audio signal as an output signal.

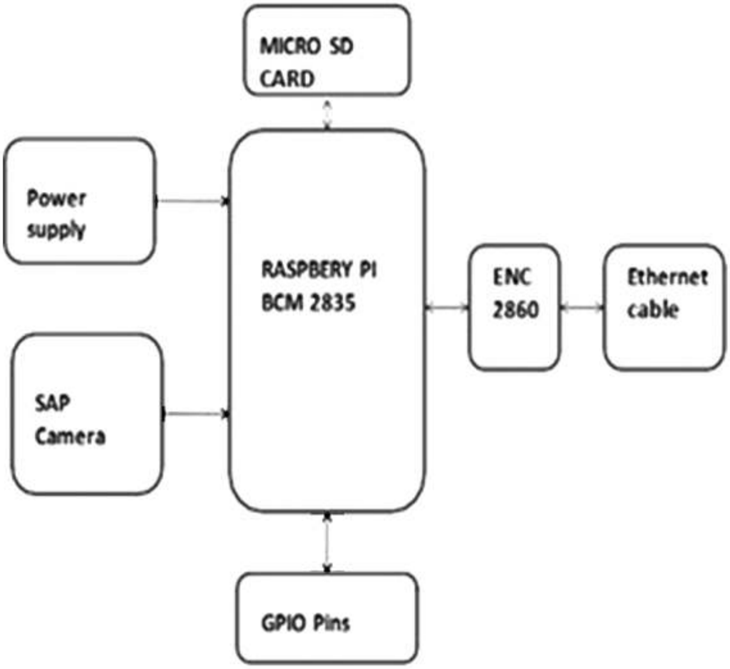


FIGURE 22.1 Block diagram of Raspberry Pi.

The Raspberry Pi 3 is the name of the third Raspberry Pi model. In February 2016, Pi3—which is depicted in Figure 22.2—replaced the Raspberry Pi 2 Model B. The Raspberry Pi 3 has the same form factor as the Pi 2 (and Pi 1 Model B+) and is backward compatible with the Pi 1 and Pi 2. The Pi 3's best feature is that it has the same shape, connections, and mounting holes as the Pi 2. All of them include Open GL ES 2.0, OpenVG hardware acceleration, and 1080p30 H.264 high-profile decode.[6,7,10,13]

22.3.1 GENERAL DESCRIPTION OF A WEB CAMERA

For any type of video device, such as TV boards, analog cameras, and USB cameras, connected to capture cards, it is possible to use active Webcam to capture photos up to 30 frames per second using FireWire (IEEE 1394) camcorders and network cameras. The program may send you the gathered pictures through email, start streaming or recording a video, and sound an alarm as and when the motion is experienced in the monitoring area. The program has options for adding text captions and image logos to pictures, as well as for adding a date and time stamp to each video frame and modifying the video's frame rate, size, and quality.

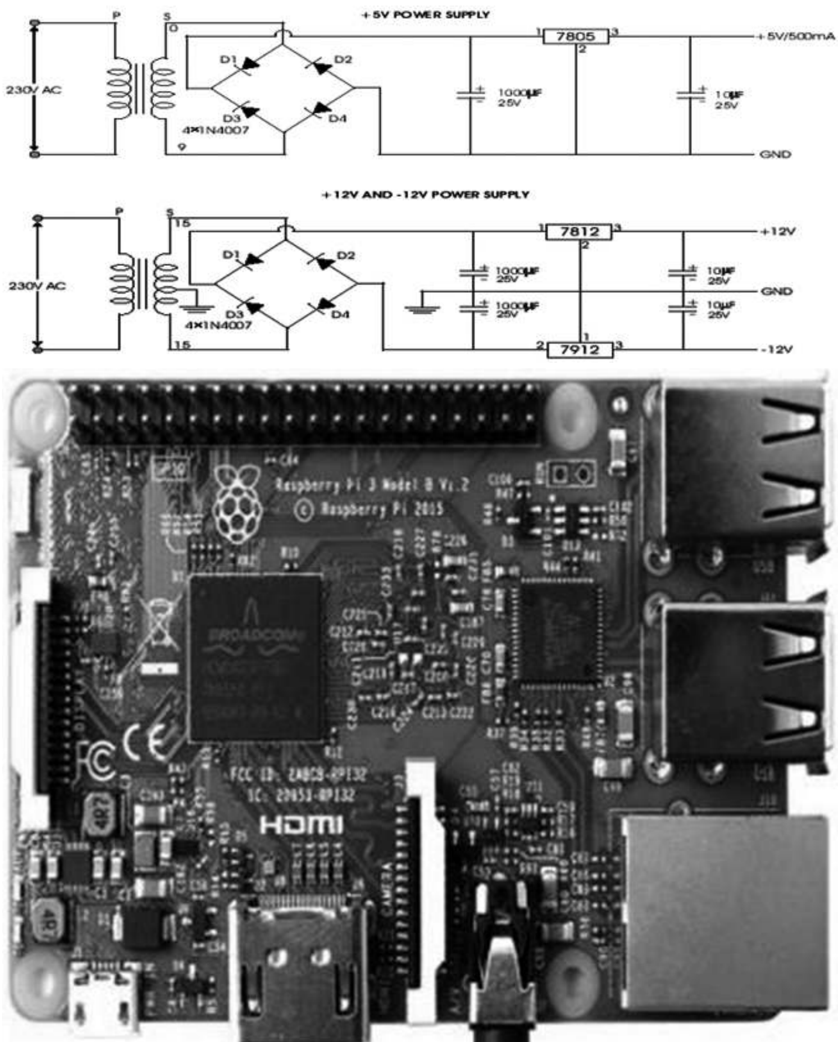


FIGURE 22.2 Circuit diagram for Raspberry connections and Raspberry Pi 3.

22.3.2 DESCRIPTION OF THE PRODUCT

The webcam device is basically a video camera, which helps to transmit and receive real-time images between computers through the internet. These images can be downloaded, viewed, or forwarded to other devices as an attachment using the internet when a video stream is "stolen". It is also possible to store, view, or resend a video stream to another site in contrast to an IP camera, which is connected through Ethernet.

In general, a camera is attached to computer hardware, such as a desktop, laptop, or connected to an external device using a USB cable. Specifications: minimum 30 frames per second frame rate.[9,11]

- Resolution
- Constant autofocus
- Microphones

Applications

- Video conferencing, video calling, and video surveillance.
- Healthcare services (for capturing arterial pulse rate).
- Control devices for input.
- Business (e.g., Webcam social shopper).
- Household safety.

22.3.3 ALGORITHM: RECOGNITION OF OPTICAL CHARACTERS

OCR is a text recognition system that can identify text in digital photographs. Although it can be employed for a number of different tasks, its primary usage is for text recognition in scanned documents.

In a digital image, OCR software finds and recognizes characters like letters, numbers, and symbols. OCR software varies in its ability to convert characters in an image into editable text—some simply export the text, while others offer more advanced features. The size, formatting, and location of the text on a page may all be obtained as output by more advanced OCR software.

OCR (or soft copy) technology can be used to transform a hard copy of a document into an electronic version. Using an OCR program, you may extract the text from a multipage document that has been scanned into a digital picture, such as a TIFF file, and turn it into an editable text file. A document can be scanned and instantaneously converted into a word-processing document using some OCR software.

OCR was created with the intention of recognizing printed text, but it may also be used to recognize digital content. The recognition and validation of handwritten content is another application for it. OCR software is used by postal systems like the USPS to automatically manage mail and deliveries depending on the address. The computer checks the scanned data against a database of legitimate addresses to verify the mailing address. OCR software, which makes use of your device's camera, is included in the Google Translate app. You can use it to translate text from books, journals, signs, and other materials into another language instantaneously. It can be instantaneously translated into another language, as shown in Figure 22.3.

22.3.4 ARCHITECTURE DIAGRAM

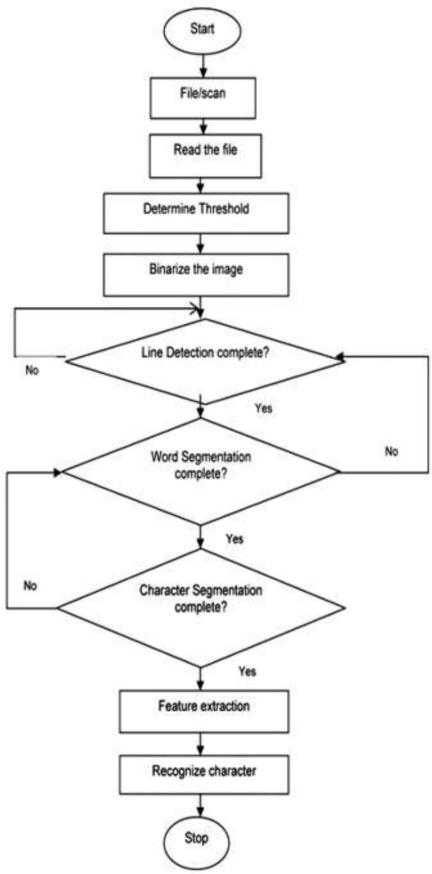


FIGURE 22.3 Flow chart of the proposed system.

22.3.5 PREPROCESSING

Data mining plays an important role in data preprocessing. For data mining and machine learning projects, the adage "garbage in, trash out" is applicable. The methods used to collect data are frequently unregulated, which leads to a variety of issues, including missing data, impossible data combinations, and out-of-range figures. Without thoroughly examining the data for these problems, analysis can produce inaccurate results. Before conducting an analysis, it is necessary to consider the quality and representation of the data. For any machine learning project implementation, such as computational biology, the data preprocessing is one of the most important tasks.

When there is a lot of redundant, irrelevant, noisy, or erroneous data, it is more difficult to uncover knowledge during the training phase. Feature extraction, cleaning, normalization, instance selection, transformation, and many other operations are included in data preprocessing. The final training set is the final outcome of data preprocessing.

22.3.5.1 PREPROCESSING OF IMAGES

The most fundamental image abstraction operation is the preprocessing of an image. By performing operations on images, the intensity of images serves as both the input and the output. An intensity image is typically described as a matrix of image function values (brightness), and images are identical to the sensor's original data. Although geometric transformations of images such as rotation, scaling, and translation are categorized here as preprocessing methods because similar techniques are applied, the idea of preprocessing is to enhance the image data by removing undesirable image features for further processing.

22.3.5.2 PREPROCESSING IMAGE PERSPECTIVES

A single chapter cannot adequately cover the vast subject of image processing. The objective is to raise awareness of the science of local and global feature description, which is frequently ignored in feature description discussions. Applications and image data will be in charge of the picture preprocessing step; therefore, a few general image processing concerns are discussed here while taking feature descriptions into account. The intention is to illustrate rather than to give instructions. As the preprocessing changes the true nature of the original data, some people realize that preprocessing images is a bad idea. However, intelligent picture preprocessing offers several advantages, addressing issues and leading to improvements in both local and global feature detection. We examine how common picture enhancement and correction techniques impact downstream feature analysis in vision. The pipeline may be advantageous or disadvantageous depending on how the approaches are applied.

The quality of the results from feature extraction and image analysis may be significantly improved by image preprocessing. It is quite similar to preprocessing of images, which is a common step in many feature descriptor systems and involves mathematically normalizing data collection. Preprocessing of an image as a sound system with different controls, such as sound with no volume controls, volume control with a simple knob to control tone, volume control with bass and mid, or volume control with a full graphics equalizer, effects processing in an acoustically superior room. This paper presents picture preprocessing by portraying it as a set of improvements and corrections that play a significant role in the computer vision pipeline.

To highlight the similarities and differences between the standard picture preprocessing techniques that can be used before feature description, instead of prescribing or limiting the techniques employed, this article's goal is to illustrate them. Comparisons of point-pair pixel intensities are the focus of local binary features. Image preprocessing may not be necessary to obtain useful results because the comparisons are very insensitive to lighting, brightness, and contrast. The majority of the existing literature's descriptions of local binary pattern approaches are based on a threshold value that can be changed to take light or contrast into account rather than involving considerable image preprocessing. SIFT (operates on small area gradients) and SURF (uses HAAR-integrated pixel values over local regions) are two spectral descriptors that provide a wide variety of picture preprocessing options. Picture pyramids are widely employed in techniques that use image preprocessing to present the data in scale space and use Gaussian filtering to smooth the higher levels of the pyramid. Before creating gradients to enhance the image, simple lighting and filtering can be helpful, for instance, increasing contrast of a picture within a gradient-edged range of intensities. Preprocessing the image data before feature extraction is an opportunity because the literature rarely discusses the advantages or disadvantages of any particular strategies used. A Fourier transform applied to a whole image or block of an image is an example of a global space feature that spans a regular-shaped polygon. However, basis space data such as the Fourier spectrum of the LBP histogram, which can be evaluated over the histogram bin values of a local descriptor, may be used to determine rotational invariance. Another example of a Fourier descriptor is to create polygon factors for the radial line segment lengths to guarantee rotational invariance. The most challenging descriptor family is one based on polygon shape, which necessitates a number of measures of the polygon structure and shapes in the image for image enhancement, structural morphology, and segmentation algorithms. There are phases in the preprocessing of the image, which may all be used in polygon shape description pipelines. Picture preprocessing for polygon feature extraction usually takes more time to set up compared to other methods because segmentation and thresholds need to be adjusted for optimal results. Polygon shape descriptors are not local patterns but rather broader regional structures with features spanning tens or even hundreds of pixels, which makes processing them time-consuming. Preprocessing of images is commonly required to correct problems that would otherwise have a negative impact on feature description.

22.3.5.3 PREPROCESSING OF THE POLYGON SHAPE FAMILY

Polygons are most challenging in terms of photo preprocessing activities because of the vast array of options. The most challenging aspects of preprocessing images may be polygon forms; there are a lot of different preprocessing techniques that may be used, and the choice of one depends heavily on the available data. Polygon shape measurements are rarely applied in specific fields due to their limitations and specified use cases, such as cell biology. The two most common methods for preparing images before measuring polygon shapes are physically fixing the lighting and selecting the background. For instance, the fluorescent dye is utilized to highlight cell features in automated microscopy applications, and the lighting angle and location are then carefully changed under magnification to offer a uniform backdrop under each cell feature to be assessed.

22.3.6 SIMPLER TO SEGMENT

22.3.6.1 ENHANCEMENT OF IMAGES

Image segmentation is the way of dividing digital images into sets of pixels, also known as superpixels. By simplifying or otherwise altering the representation of a picture, segmentation aims to improve comprehension and evaluation of the image. In order to identify object boundaries, lines, and curves in images, image segmentation is widely used. Segmentation can also be represented in another way as it is the process of identifying each pixel in a photograph so that pixels with the same label have unique identification.

In terms of some property or feature of a photograph, such as color, intensity, and texture, the pixels in a specific region are comparable. The images having the same attribute(s) differ dramatically between nearby positions. The contours produced by applying image segmentation to a stack of images, as is customary in medical imaging, can be used to produce 3D reconstructions using interpolation methods like marching cubes.

Segmentation searches for patterns that can be found in any regularity in an image utilized to reduce it. This is how these two principles are related. The method describes the texture and border form of each segment. A probability distribution function is used to mimic each of these components, and the coding length is established as follows:

The encoding of boundaries makes use of the smooth contour that naturally occurring picture sections have. This is done in Huffman coding before the contours of an image are encoded using a difference chain code. Therefore, the coding length decreases with increasing border smoothness.

This is so because the largest entropy belongs to the normal distribution of all the distributions with the same mean and covariance since the reduced length cannot be more than the true length of code achieved by the approach.

With this method, the segmentation is used to determine the amount of bits required to encode a picture. The objective is to identify the image segmentation from all feasible segmentations that yield the most compact coding length. This is obtained by using a straightforward agglomerative clustering method. The degree of segmentation coarseness is identified by applying the lossy compression distortion method, and the ideal value for each image may change. Heuristically, this value can be estimated using texture contrast in an image. For example, when a picture's textures are similar, more sensitivity and thus lower quantization are necessary, as in camouflage photos.

22.3.6.2 THRESHOLDING

The most fundamental technique for segmenting images is the threshold approach. In this method, a grayscale image is converted into a binary image by applying a threshold value. Another approach called balanced histogram threshold is also available.

The secret to this technique (or values when many levels are specified) is choosing the threshold value. In the business world, the most widely used techniques are maximum possible entropy, maximum variance Otsu's approach, and k-means clustering.

Recent years have seen the development of techniques for computed tomography (CT) image thresholding. The key concept is that radiographs use the reconstructed image to set the threshold values and picture, unlike Otsu's method.

New approaches are proposed using the nonlinear thresholds based on multidimensional fuzzy rules. Each pixel is assigned a membership value. Images are segmented using evolutionary algorithms and rules built on different dimensions using fuzzy logic that take into account the application and illumination environment of the picture.

With the use of this technique, which creates a histogram out of each pixel in the picture, it is possible to identify the clusters in the image by looking at the peaks and troughs in the histogram. Color or intensity could be used as the measurement. This method can be improved by recursively dividing the large clusters in the image into smaller clusters using the histogram-seeking technique. Until there are no additional clusters to be generated, this process is repeated with progressively smaller clusters. The difficulty of identifying important peaks and troughs in the image may be a drawback of the histogram-seeking strategy. While maintaining single-pass effectiveness, expanding histogram-based methods is quick to cover many frames. The histogram can be calculated in a number of different ways when several frames are taken into account. Once the findings are combined, once-difficult-to-identify peaks and valleys are more likely to differ now when the same methodology employed with one frame is used with numerous frames. The histogram can also be used to analyze data at the pixel level, allowing one to determine the color that appears most frequently in each pixel. Segmentation depends on moving objects in a static context; this technique creates a unique sort of segmentation that is advantageous in video tracking.

22.3.6.3 DETECTING THE EDGES

A mature field unto itself is edge detection. Region edges and boundaries are where the intensity frequently changes sharply; they are closely connected. Thus, another segmentation technique has been designed with the help of edge detection techniques.

Edge detection usually reveals disconnected edges. The separation of an object from a picture requires closed-area borders. The desired edges are those that divide these items or spatial taxons.

Spatial taxons are information grains made up of a sharp pixel region that are positioned at abstraction levels in a hierarchically layered scene architecture. They share similarities with Gestalt's psychological figure-ground classification, but they also include clusters of objects, the main object, and noticeable object sections. It is extremely useful to employ methods for spotting edges in a strategy.

Segmentation techniques can be used to separate the edges that edge detectors have detected. Li and Lindeberg established an integrated strategy for part-based object recognition that divides edges based on a suitable easy description length of the need into both straight and curving edge segments using a strategy equivalent to split-and-merge that may create breakpoints derived to get more likely sites at which subdivision into distinct components from similar junction cues is assessed.

22.3.6.4 GROWING TECHNIQUES BY REGION

The underlying premise of region-growing algorithms is that adjacent pixels inside a region have comparable values. Comparing one pixel to those in its immediate surroundings is a common technique. A pixel can be attributed to one or more of the cluster's neighbors if an analogy requirement is achieved. Since noise affects results in every situation, the similarity criterion that is selected is crucial.

Implementing 4-connectedness to create a network of pixels and edges considered by the difference in intensity's absolute value is the first step of the statistical region merging (SRM) method. Each pixel initially creates a single-pixel region. After that, SRM prioritizes certain edges and utilizes a statistical predicate to select whether or not to integrate present locations that correspond to pixels toward the edge.

One strategy for expanding areas is the method of seeded regions. This technique accepts the picture and a list of seeds as input. Every segmented object is recognized using the seeds. The zones are repeatedly extended by comparing them to all surrounding, unallocated pixels. The discrepancy between a quantity of intensity at a pixel and the average of the vicinity is a metric for similarity. The region is given to the part of the image having the least amount of change as determined in this way. Until all of the pixels in an area have been assigned, this technique is repeated. The segmentation results depend on the seed selection since seeded region growth requires the insertion of seeds, and picture noise can result in the seeds being improperly positioned.

One variant of this technology is the pixel intensity method, which comprises the intensity, scatter, and mean of the portion. Implementing the values of the prospective pixel, a statistical evaluation is calculated. If the statistical test is modest enough, the region's mean and scatter are modified when the pixel is inserted. If not, a new region is made and the pixel is discarded. Split-and-merge segmentation utilizes an image's quadtree division. This technique begins at the highest point of the picture's tree, which symbolizes the entire picture. If detected to be uneven, divide into four smaller squares, and so forth. On the other hand, four identical kid squares are joined into a number of connecting parts. A segmented component is an item in a tree. The split or merge process is repeated until no more splits or merges are possible. The time complexity of the method's implementation can reach its best algorithm with a special data structure.

An approach applied to the fields of picture processing and vision for machines to remove the foreground of an image for future processing (object recognition, etc.) is background subtraction, also known as foreground detection. The areas of interest in a photograph are typically those in the foreground, such as people, vehicles, text, etc. This technique may be used for object localization following the stage of image preprocessing (which may include picture denoising, postprocessing like morphology, and so on).

In static camera movies, removing the background makes it easier to see moving objects. The method is based on identifying moving objects using the difference between two points, both the present structure and a reference structure, often referred to as a "background image" or "background model." Background removal is typically utilized when a picture is a component of a video feed. Computer vision applications such as surveillance tracking and human posture estimation benefit from background subtraction.

Backdrop subtraction is frequently predicated on a static background hypothesis that may not always apply in real-world situations. Reflections or animated pictures on screens cause background alterations in interior scenarios. Similarly, static backdrop solutions struggle with outside images due to wind, rain, or lighting variations brought on by the weather. Morphological operations are a group of image processing operations that process images based on their forms. A structuring element is applied to an object using morphological techniques and identical-sized output and input pictures. Every pixel in the resulting picture is determined by comparing the matching pixel in the source frame along with adjacent ones in a morphological operation.

You can create a morphological operation that is sensitive to specific shapes by modifying the nearby area's shape and dimension, and the input image is transformed.

The two most fundamental structural processes are dilation and erosion. Dilation increases pixel content at object boundaries in an image, whereas erosion decreases pixel content at object boundaries. The amount of pixels that are added to or removed from the size and shape of the structure's components influences the size and shape of the image's objects used to process the picture. To establish the position of any specific pixel in the ultimate picture, morphological erosion and dilation apply a framework for matching the pixel and the adjacent cells in the source image. The pixel processing guideline specifies the dilation or erosion process. This table lists the guidelines for both erosion and dilation. A collection of nonlinear techniques known as morphological image processing deals with the shape or morphology of visual elements. According to Wikipedia, morphological techniques are ideal for processing binary images since they only depend on the proportional location of pixel values rather than the associated numbers. To enhance light transfer, morphological adjustments can also be made to grayscale photos. The sum of the respective pixel values is either of little or no interest since the functions are unknown.

A computer technology known as object detection looks for instances of semantic entities belonging to a particular group (people, buildings, or cars) in computerized photographs and videos. It relates to machine perception and processing of pictures. Two comprehensive object identification fields are identification of faces and movement recognition. The recognition of objects is useful for many applications involving machine vision, such as picture retrieval and video surveillance. It helps in the detection and identification of faces. It can also be used to follow the movement of objects like a football during a game, a bat for cricket, or someone in a movie. Each group of objects has a set of distinctive properties. This helps to categorize the group because, for instance, all circles are round. In order to identify an object class, these attributes are used. For instance, while searching for circles, items at a particular range between the center are searched. Similarly, when looking for squares, one needs objects with equal side lengths and perpendicular corners. A similar method is used for face recognition, in which characteristics like skin tone and eye distance are identified in addition to the features of the eyes, nose, and mouth. The art of identifying objects from the real world, like faces, bicycles, and buildings, in still images or moving pictures is known as object detection. Object identification algorithms frequently make use of extracted features and learning strategies to discriminate between instances of an item category. In addition to other applications, it is frequently used in advanced driver assistance systems (ADAS), security, and surveillance.

22.4 RESULTS AND DISCUSSION

Sample Input 1

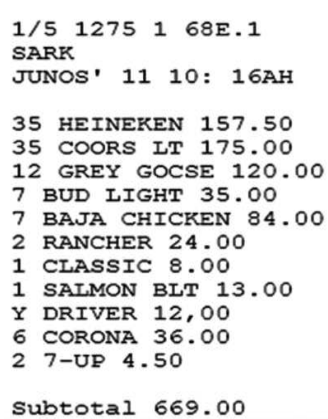

THE SUNCADIA RESORT
TC GOLF HOUSE
1103 CLAIRE 7
1/5 1275 GST 1
SARK
JUN05'11 10:16AM

35	HEINEKEN	157.50
35	COORS LT	175.00
12	GREY GOCSE	120.00
7	BUD LIGHT	35.00
7	BAJA CHICKEN	84.00
2	RANCHER	24.00
1	CLASSIC	8.00
1	SALMON BLT	13.00
1	DRIVER	12.00
6	CORONA	36.00
2	7-UP	4.50
	Subtotal	669.00
	Tax	53.52
3:36	Amnt Due	$722.52

FOR HOTEL GUEST ROOM CHARGE ONLY
Gratuity________________

FIGURE 22.4 Sample input 1.

Output 1


1/5 1275 1 68E.1
SARK
JUNOS' 11 10: 16AH

35 HEINEKEN 157.50
35 COORS LT 175.00
12 GREY GOCSE 120.00
7 BUD LIGHT 35.00
7 BAJA CHICKEN 84.00
2 RANCHER 24.00
1 CLASSIC 8.00
1 SALMON BLT 13.00
Y DRIVER 12,00
6 CORONA 36.00
2 7-UP 4.50

Subtotal 669.00

FIGURE 22.5 Sample output for input 1.

Sample Input 2

FIGURE 22.6 Sample input 2.

Output 2

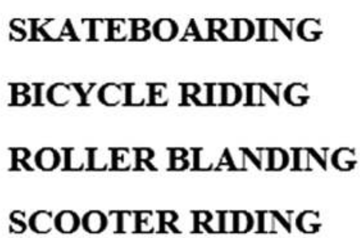


FIGURE 22.7 Sample output 2.

The procedure used for text-to-speech conversion is explained as follows.

OCR Pipeline Extraction

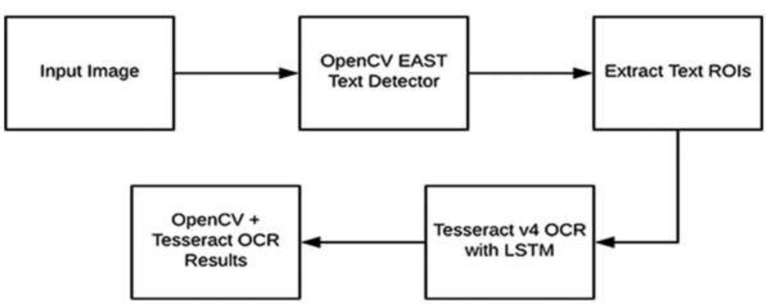


FIGURE 22.8 Processing of input file and output file.

22.4.1 SETTING UP TEST PHOTO IMAGE STORAGE

The following are the most efficient methods for obtaining test images:

a. Find pictures online by searching for terms like "road sign," "restaurant menus," and so forth.
b. Utilizing a snipping tool, save photographs from online books, novels, and articles.
c. Take screenshots of captions or suggestions put on essential domestic items using your camera.

The following are the least efficient methods for obtaining test images:

a. Identify a book and then enter the initial couple of sentences into an electronic document. Next, print it on A4 paper, scan it as a PDF, or create another picture file from it.
b. Learn the principles of pixels so you can place letter blocks on a 128x128 canvas. If you think it is excessive because the pixel-filling process takes a long time, think about enrolling in domains in programing and computations to create a few programs to automate the entire process.

The following photos were saved as test images:

- A highway sign.
- A scanned copy image of an abstract from a published paper.
- Images with numerical values.

Preprocessing a black-and-white photograph with dark writing on a white background:

- Remove the alpha channel (save the image as a jpeg/jpg rather than a png).
- PSM settings for fine-tuning (page segmentation mode).

API Calls and Usage

For a single image, use with-statement.

a. Page segmentation mode selection
b. Installing more languages

22.5 CONCLUSIONS

The effectiveness of the suggested method and other existing algorithms has been tested under our dataset. It is observed that the suggested strategy is better compared to other existing approaches. The suggested model obtained an accuracy of 71.7% (shown in Figures 22.6–22.8). With the help of the proposed approach, people who are visually challenged can now listen to anything they want. Apart from that, they can translate the text into the target language with the aid of translation tools, for which they can then use Google Speech Recognition to convert to voice. As a result, they will be independent. Additionally, compared to other approaches, it is very less expensive.

KEYWORDS

- **interactive voice response**
- **optical character recognition**
- **text-to-speech**
- **voice conversion**
- **global variance (GV)**
- **signal-to-noise ratio (SNR)**
- **digital audio workstation (DAW)**

REFERENCES

1. Lynn, H. Exploring Computing Education in Rural Schools in India [Online]. 2014. https://www.raspberrypi.org/blog/exploring-computing-education-in-rural-schools-in-india/
2. Slaven, C. The Raspberry Pi Phenomenon: Global Education Uses. 2015. http://students.ecohouseinitiative.org/%EF%BF%BCthe- raspberry-pi-phenomenon-global-education-uses.
3. Severance, C.; Fontichiaro, K. *Raspberry Pi*. Cherry Lake Publishing: North Mankato, 2013.
4. Kothari, C. R. *Research Methodology: Methods and Techniques*, 3th ed.; New Age International Pvt Ltd Publishers: India, 2013.
5. Agrawal, N.; Singhal, S. In *Smart Drip Irrigation System Using Raspberry Pi and Arduino*, Proceedings of International Conference on Computing, Communication & Automation (ICCCA), 2015; pp 928–932.
6. Danymol, R.; Ajitha, T.; Gandhiraj, R. In *Real-time Communication System Design using RTL-SDR and Raspberry Pi*, Proceedings of International Conference on Advanced Computing and Communication Systems (ICACCS), IEEE: Coimbatore, India, 2013; pp 1–5.

7. Kyuchukova, D.; Hristov, G.; Zahariev, P.; Borisov, S. In *A Study on the Possibility to use Raspberry Pi as a Console Server for Remote Access to Devices in Virtual Learning Environments*, Proceedings of International Conference on Information Technology Based Higher Education and Training (ITHET), 2015; pp 1–4.
8. Srinivasan, M.; Anand, B.; Antony Venus, A.; Victor, A.; Narayanan, M.; Sree Rakshaa, S.; Vijayaraghavan, V. In *GreenEduComp: Low-cost Green Computing System for Education in Rural India: A Scheme for Journal of Computers 297 Volume 13, Number 3, March 2018 Sustainable Development Through Education*, Proceedings of IEEE Global Humanitarian Technology Conference (GHTC), 2013; pp 102–107.
9. Soetedjo, A.; Mahmudi, A.; Ashari, M.; Nakhoda, Y. In *Raspberry Pi Based Laser Spot Detection*, Proceedings of International Conference on Electrical Engineering and Computer Science (ICEECS), 2014; pp 7–11.
10. Byrne, J.; Fisher, L.; Tangney, B. In *Computer Science Teacher Reactions Towards Raspberry Pi Continuing Professional Development (CPD) Workshops using the Bridge21 Model*, Proceedings of the 10th International Conference on Computer Science & Education (ICCSE), IEEE: Cambridge, 2015; pp 267–272.
11. Fung, P.; White, D.; Jouet, S.; Singer, J.; Pezaros, D. In *The Glasgow Raspberry Pi Cloud: A Scale Model for Cloud Computing Infrastructures*, Proceedings of IEEE 33rd International Conference on Distribute Computing Systems Workshops (ICDCSW), IEEE: Philadelphia, PA, 2013; pp 108–112.
12. Kaup, F.; Gottschling, P.; Hausheer, D. In *PowerPi: Measuring and Modeling the Power Consumption of the Raspberry Pi*, Proceedings of IEEE 39th Conference on Local Computer Networks (LCN), IEEE: Edmonton, AB, 2014; pp 236–243.
13. McLeod, S. Experimental Method. 2012. Retrieved August 30, 2016 [Online]. http://www.simplypsychology.org/experimental-method.html.
14. Tsukada, M.; Santa, J.; Matsuura, S.; Ernst, T.; Fujikawa K. On the Experimental Evaluation of Vehicular Networks: Issues, Requirements and Methodology Applied to a Real Use Case. *EAI Endorsed Transactions on Industrial Networks and Intelligent Systems*. 2014.
15. Wen, K.; Chang, S. In *A Study of Visualization Analysis Method in Crowd Flow Spatial Survey*, Proceedings of the 22nd International Conference on Geoinformatics, 2014; pp 1–7.
16. Meg, S.; Mary, M. Using Cost Analysis in Evaluation. 2015. Retrieved August 30, 2016, from the website: http://ag.arizona.edu/sfcs/cyfernet/cyfar/Costben2.htm
17. Layard, R.; Glaister, S. *Cost-Benefit Analysis,* 2nd ed.; Cambridge University Press: Cambridge, United Kingdom, 1994.
18. Borysowich, C. CBA Sample. 2009. Retrieved September 30,2015, from the website: http://www.scribd.com/doc/20757582/CBA-Sample#scribd
19. Mishan, E. J.; Quah, E. *Cost-Benefit Analysis*. Taylor & Francis: Abingdon, Oxford, United Kingdom, 2007.
20. Markgraf, B.; Media, D. How to Calculate the Total Operating Costs & Breakeven Volume. http://smallbusiness.chron.com/calculate-total-operating-costs-breakeven-volume-63129.html
21. How Much Does it Cost. 2015. Retrieved September 30, 2015, from the website: http://www.powershop.co.nz/how- much-does-it-cost/
22. Allen, D. Cost/benefit Analysis for Implementing ECM, BPM Systems. *Inform. Manag. J.* **2007,** *41* (3), http://www.arma.org/bookstore/files /Allen.pdf

23. Bhattacharya, S.; Borah, S.; Mishra, B. K. Emotion Detection from Multilingual Audio Using Deep Analysis. *Multimed. Tools Appl.* **2022,** *81,* 41309–41338. https://doi.org/10.1007/s11042-022-12411-3
24. Banerjee, N.; Borah, S.; Sethi, N. Intelligent Stuttering Speech Recognition: A Succinct Review. *Multimed. Tools Appl. Springer* **2022,** *81,* 24145–24166. https://doi.org/10.1007/s11042-022-12817-z

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

CHAPTER 23

SELF-LEARNING ARTIFICIAL INTELLIGENCE SMARTBOT WITH NLP AND SPEECH RECOGNITION

P. VISU[1], A. MOHAN[2], A. PRABHU CHAKKARAVARTHY[2],
G. PARTHASARATHY[3], and SANJU DAS[4]

[1]*Department of Computer Science and Engineering, Velammal Engineering College, Chennai, Tamil Nadu, India*

[2]*Department of Computer Science and Engineering, SIMATS School of Engineering, Saveetha University Chennai, Tamil Nadu, India*

[3]*Department of Computer Science and Engineering, Reva University, Bengaluru, Karnataka, India*

[4]*Department of Computer Science, Assam University, Silchar, Assam, India*

ABSTRACT

Chatbots are used to reduce tedious human work and these bots provide quite accurate information than getting information from the physical information provider. However, the problem with the existing chatbots is not mutable and any updates cannot be done without editing the source code or by editing the knowledgebase. This bot purely uses a dynamic knowledge base. This Smartbot was made to rectify the problems faced by the regular bots. The Smartbot that we developed is made up of Python language and uses speech recognition for getting voice input from the user and giving the response to the user in the text format. This Smartbot scrapes the information from the given website and uses it as a knowledge base. Before using the knowledge

Applied Soft Computing Techniques: Theoretical Principles and Practical Applications. Samarjeet Borah, Ratna Raja Kumar Jambi, Sharifah Sakinah Syed Ahmed, & Mahendra Prabhakar Deore (Eds.)
© 2025 Apple Academic Press, Inc. Co-published with CRC Press (Taylor & Francis)

base this Smartbot preprocesses the scraped data. Hence this Smartbot uses Python tools for NLP and speech recognition. The user input is searched against the knowledge base for a similar word. It calculates the similarity using the cosine similarity function. Since this chatbot can learn from the website, it can be self-updated. Hence the uttermost advantage of this Smartbot is this bot can update itself when the website gets updated and the additional speech recognition feature provides easier user access.

23.1 INTRODUCTION

Artificial Intelligence (AI) is referred to as the simulation of human intelligence on a machine that is programed to think like humans. The term may exhibit to the machine with a human mind such as artificial thinking and problem-solving the most important character of AI is its ability to choose the right thing and act which has the best outcome and high-yielding path to achieve a specific goal. AI continuously evolves in many industries. Mathematics is wired using an interdisciplinary approach based in mathematics, computer science, information technology, psychology, and even more. Application of AI is more, and it can be applied in many sectors. AI can be used in the healthcare industry for dosing drugs and different treatments in patients and for surgical procedure in the operation theatre. Some examples of AI are computers playing chess, self-driving vehicles, and chatbots.

AI can be categorized into strong AI and Weak AI. Strong AI is the one whose intelligence is equivalent to human intelligence. Weak AI is the one whose intelligence is limited to a particular domain. Knowledge engineering is the field of AI that deals with enabling a machine or a system to mimic the thought process of a human expert. AI is the logical investigation of AI calculations and factual models that a framework uses to play out a specific undertaking without utilizing manual directions. AI is utilized in a wide variety of calculations, such as email sifting and personal computer (PC) vision. AI is firmly connected with computational measurements which centers around making expectations utilizing PCs. Information mining is the field of study that manages investigating information through solo learning. ML is additionally called as prescient examination. AI has a few learning techniques some of them are: Supervised machine learning, Unsupervised machine learning, semi-supervised machine learning, and reinforcement machine learning. Supervised machine learning works with experience and uses that knowledge to predict the future. Unsupervised machine learning predicts the future output by using training knowledge acquired in the training phase.

Chatbot is an AI software in a device just like Siri, Alexa, or Google Assistant. These Chatbots assist the user in performing a particular task like providing information about a product, providing necessary information about the college, commercial transactions, inquiries such as hotel booking, bill payment, form submission, etc. Today almost every company uses chatbots and integrates them into their official websites for answering the queries raised by the clients. Chatbot can be integrated with Natural language processing (NLP) and speech recognition for getting audio queries from the user. Chatbots can be created using programing languages such as Java, Python, C, C++, etc. and it can also be created by using AI scripting languages such as artificial intelligence markup language (AIML), etc.

Integration of chatbots with the AI algorithms results in high accuracy and it may provide the precise information that the user needs for the user. Some of the major uses of chatbots in industries are to deliver flight information to connect with the user and the company and to support the user by satisfying their queries with appropriate responses. This project describes the implementation of a self-learning chatbots called Smartbot which uses user voice as the input and gives the corresponding answer as the output in the text format. This Smartbot's knowledge base was created by providing a website uniform resource locator (URL). This bot will scrape the data from the URL and store it as a knowledge base.

23.2 LITERATURE SURVEY

Wang et al.[1] Proposed to increase brain network-based engineering with information installing and information mindfulness on a Talkbot (mindful per user) to integrate outer literary information into dialog model work with the discourse and demonstrate the age through which they assessed that their bot could create more cognizant and data exchange.

Daniel et al.[7] Proposed Xatkit—multifaceted low code visit bot improvement Edge work Xatkit handles the issues on the past structure by giving a bunch of areas explicit language to characterize chatbots in a stage's freeway.

AL-Khalifa[8] proposed a framework with a mix of a variety of coding strategies, literal interpretation, and text-to-discourse advancements which makes an instructive apparatus for learning the Arabic language this application can likewise be broadened wide different dialects for better interpretation administration. This framework performs encoding and introducing variety change of each work and characters for interpretation. This framework can have utilized for youthful Arabic students and different people groups who are attempting to realize this language.

Tapsai[13] proposed another model for recovering regular language information from CSV records. It cases normal language handling and procedures such as semantic examples, philosophy, and intuitive discussion framework, for Investigating the culmination and nuances of regular language. These frameworks permit the client to alter the flaws articulations (or) words in the record. The blunders caused in this framework come from those point sources, for example, changing the word significance, utilizing equivocal words, and wrong place of words in the normal language explanation.

chen et al.[16] proposed an original normal organization language for discourse acknowledgment of the succeeding word RNNCM, Su-RNNLM to resolve the issues looked in the existing model as opposed to utilizing an intermittent unit to catch the total future word text, a feed-forward unit in used to display a proper limited number of succeeding words.

Lotifan and Busso[19] proposed a brain network model from engineered discourse to differentiate close to home substance of a discourse signal. This framework needs an example of a reference sign to distinguish the objective sign. This task portrays building three reference sentences hoisting the development in discourse union and alongside it, this paper dissects whether the synthesized signal provides a valid template reference to explain natural speech using feature analyses and perceptual Analysis.

Various processes and strategies employed in a typical speech system are summarized in a novel scheme of total perspective on voice recognition.[21] With the help of these techniques, some major difficulties can be lessened, including speaker and language unpredictability, ambient odors, and vocabulary size.

Li and Rafiei[14] proposed an alternative method for handling the Natural language data due to its rapid growth in size. This proposal explores two relevant areas of overlap to the DB community (1) managing natural language text data in relational databases and (2) developing Natural language interfaces to databases.

Carlos de Olivera et al.[11] proposed a Telegram bot that communicates wither the user efficiently through Telegram messenger. They integrated Telegram bots with the Arduino platform for developing the Internet of Things (IoT)-based chatbots. This bot allows different users to communicate with different hardware for different purposes.

Mathew et al.[16] proposed a chatbot for Disease prediction and treatment using machine learning. This bot predicts the disease which is given by the user and responds to the user with appropriate treatment recommendations. This chatbot is built to reduce the work of doctors and to reduce time costs. More intelligent speech processing-related works can be found in Refs. [17, 18].

Pawar and Mago[9] proposed a system to resolve the conflicts in identifying semantic similarity between two words by incorporating corpora–based statistics into a standardized semantic similarity algorithm. It achieved word similarity in the rate of $r = 0.8753$ and sentence similarity of $r = 0.8793$, and it gives the similarity for the sentences involving compositional knowledge (SICK) dataset in the value of $r = 0.8324$ which specifies the efficacy of this algorithm.

Kades and Woods[6] proposed a voice chatbot based on the study on natural language chatbots such as Siri, Google Assistant, Alexa. These chatbots were developed using AIML. This work discusses the similarities and difference in the previous techniques and examines in particular the Loebner Prize winning chatbots.

Thakur et al.[4] proposed an AI chatbot using Android for answering queries based on college admission. This chatbot was integrated with specific colleges and answered students about the college-related queries.

Indumathi et al.[2] proposed a healthcare care chatbot using AI that provides information about the diseases upon the given symptoms by the user. This chatbot was built using artificial XML (AIML). This proposal is like the idea implemented by Bani and Singh[5] proposed a college inquiry chatbot based on the study of ALICE bot and developed a simple chatterbot for answering the students' quires about their college. This chatterbox takes the user's natural language as input and gives the information in text format. This research is also using the concept of neural network proposed by Wang and Yuho[17] proposed a brain network model for an information base [knowledge base (KB)]-based single connection question addressing (SR-QA). This model has two principal modules which are substance connecting and connection location. Their paper basically concentrated based question portrayal in SR-QA. In the connection location module, they need staggered target consideration (MLTA) to use the staggered portrayal of connection.

23.3 PROPOSED MODEL

This paper emphasizes voice queries given by users. As the proposed model uses both NLP and speech recognition techniques the model helps the user to process the query as per the requirement of user. The chatbot receives the voice query from the user and returns the result from the knowledge base according to the user query. All the existing approaches were domain specific.

The proposed chatbot model explains about enhancing the knowledge base of a Smartbot using Python and incorporating NLP and speech recognition also it overcomes the existing domain-specific problems. The proposed model uses the following processes:

23.3.1 NLP AND SPEECH RECOGNITION

NLP is the branch field of phonetic software engineering, data designing, and man-made brainpower principally centers around the association between two PCs or the PC and human communication. Generally, How the figure is handling and examining the enormous measure of regular language information. Challenges in NLP include discourse acknowledgment, normal language understanding, and regular language age. Discourse acknowledgment is the capacity or the ability of a machine or a program to distinguish the words expressions and sentences communicated in language and convert them into a machine coherent configuration. Discourse acknowledgment works with the assistance of calculations through acoustic displaying and language demonstrating.

23.3.2 TOKENIZATION

Tokenization is the process of replacing sensitive or text data with appropriate unique symbols. The symbols may be numbers or codes etc. This tokenization concept comes under NLP. Here the Natural language is given as training data. The bot trains itself with the training data. In this paper, we use training data as text. The bot split that text into sentences and then into words. Then the bot assigns each word a peculiar number. These words are called tokens. The process of generating tokens is called tokenization. The numbers assigned to tokens are not continuous and they are randomized. By which if a word occurs recurrently, it has the number which was assigned to it during its first occurrence.

23.3.3 LEMMATIZATION

Lemmatization is the process of gathering the inflected forms of a word into a single sentence so that they can be searched as a single item. Lemmatization is just like the lexical analysis. However, here the word is analyzed with respect to their tenses. Lemmatization slightly varies from the stemming process. The major difference between lemmatization and stemming is, stemming creates nonexistent words. Stemming takes the stem part of the word into consideration. However, the lemmatization creates actual words that are in use. For example, consider the words "better" and "good" are in the same lemma so they are the same. However, in stemming those words are different.

23.3.4 SIMILARITY ANALYSIS

Similarity analysis refers to analyzing the similarity between two terms. The terms may be two sentences or two words. So, one of the best ways to deal with finding likeness is the cosine comparability strategy. This technique is more proficient than Euclidean distance. Cosine closeness is a methodology used to track down the likeness of reports. Cosine closeness estimates the cosine point between two vectors projected in a complex space. Assuming two reports are isolated and far separated by the Euclidean distance considering the size, they might in any case have a more modest cosine point between them. So, the more modest the cosine point, the higher the likeness.

23.4 SYSTEM ARCHITECTURE

The user will start the process by giving the URL to the bot. The bot will scrape the data from the website for scraping it using a Python library called Newspaper3k. The Python newspaper has a package called article. This article package needs the URL of the webpage to be scrapped. After scraping, the bot will tokenize the words on the webpage. Then lemmatize the words to find the correct match. Next, it asks the user to provide the input. The bot will get voice input from the user. Bot searches for appropriate similarity against the knowledge base. After processing bot gives the corresponding output. If the user gives "Bye" or "Thanks" as input the bot will terminate the process. The overall working of the proposed system is shown in Figure 23.1.

The system architecture is given as follows:

23.5 RESULTS AND DISCUSSION

An automated approach called web scraping is used to retrieve large amounts of data from websites. Usually, the data found on the website are in unstructured form. Web scraping helps in gathering unstructured data and storing them in structured form as shown in Figure 23.2.

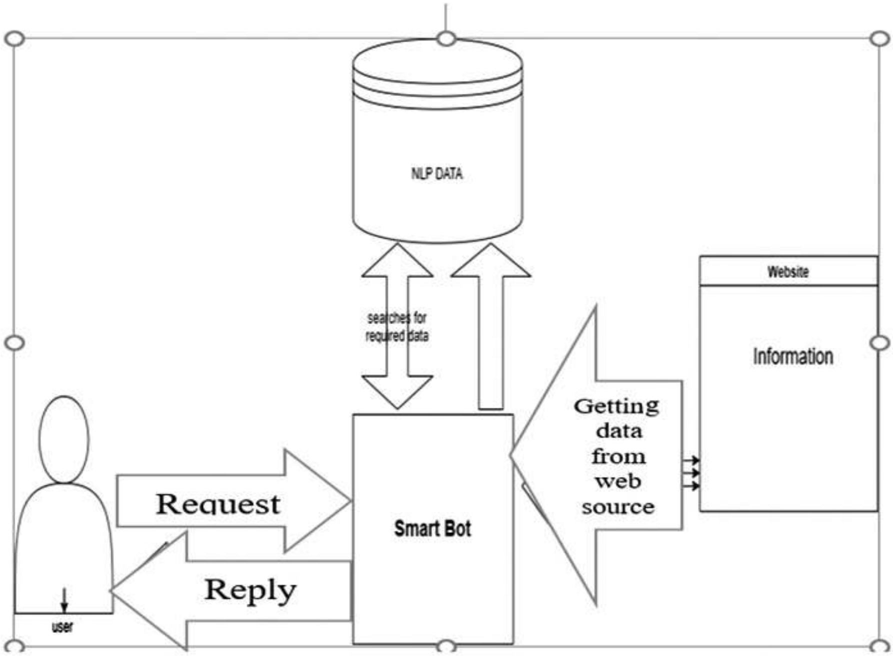


FIGURE 23.1 Architecture of the proposed system.

The stored structured data are being split into words using Python programing. The split(param) technique is used to separate a string into smaller substrings. A string object is part of this method. The parameter is optional and can be divided on a certain string or character. The string can be broken up into words from a given sentence. Like this, a paragraph is also possible to divide into individual characters as in Figure 23.3.

Figure 23.4 shows the result of the proposed bot model based on the voice input given by the user. The bot retrieves the correct match from the knowledge base.

When the user input does not have the correct match in the knowledge base, the bot returns an "apology" message as "I do not understand," otherwise the proposed Smartbot will return all the relevant information as per the user query as shown in Figure 23.5.

```
LEAVE YOUR QUESTIONS RELATED TO ADMISSION PROCEDURE , APPLICATION , ELIGIBILITY , INFRASTRUCTURE , CUT OFF WE WILL
GET BACK TO YOU AS SOON AS POSSIBLE ..
['VEL TECH HIGH TECH DR. RANGARAJAN DR. SAKUNTHALA ENGINEERING COLLEGE ADMISSION 2020\n\nEstablished Year:2002\n
\nVeltech Hightech Dr Rangarajan Dr Sakunthala Engineering College is located at Chennai,2019 admissions
started.Admission Process , Eligibility , Call 9700019482.', 'The institute was established in 2002.', 'VEL TECH
was set up by the VEL TECH group of educational institutions that was set up in the year 1990 by the celebrated
industrialist duo (couple) Dr. R. Rangarajan and Dr. R. Sakunthala Rangarajan.', 'They have been in the public
service for many years and they are experienced too.', 'The major facilities are library with a huge collection of
books, magazines and e-journals, computer centre, seminar halls, convention hall, auditorium, separate hostels for
boys and girls, medical facilities, transportation facilities for day scholars etc.', 'The major recruiters are
Adithya Tradings Pvt Ltd, American Mega Trends, Convergys Information, Esab India Pvt Ltd, India Bulls Ltd, Midas
Communication, Siemens Information Sys Ltd, Vernalis Systems etc.', 'The students of the institute undertake
various cultural and sports activities.', 'Facilities:\n\nClassroom Facilities\n\nThe institute provides separate
hostel facilities for boys and girls.', 'Laboratory Facilities\n\nThe classrooms/ lecture halls in the college are
fully furnished.', 'Extra - Curricular Facilities\n\nRegular personality development workshops.', 'Regular
seminars, workshops and conferences.', 'Library Facilities\n\nCollege has a well stocked library.', 'Sports
Facilities\n\nThe college has a state-of-art computer laboratory equipped with 60 P-IV nodes on LAN system with
free internet access through 512 Kbps broadband facility.', 'NAAC Grade: A\n\nRanking: AAA+\n\nCourses Offered:\n
\nUnder Graduate:\n\nB.E.', 'Civil Engineering\n\nB.E.', 'Computer Science and Engineering\n\nB.E.', 'Electrical
and Electronics Engineering\n\nB.E.', 'Electronics and Communication Engineering\n\nB.E.', 'Mechanical Engineering
\n\nB.Tech.', 'Bio Technology\n\nB.Tech.', 'Chemical Engineering\n\nB.Tech.', 'Information Technology\n\nWhy Join
vel tech dr rangrajan dr sakunthala technical university\n\nRanked 31st in Engineering survey conducted by Higher
Education Review Magazine 2015\n\nScholarships fund worth 7 crores for meritorious students\n\nAICTE approved and
ISO 9001:2008 certified\n\nNBA accredited for MBA, CSE and ECE\n\nMOUs with national and international industries/
institutions for research, faculty and student exchange etc.', 'Accredited by TCS and CTS\n\nMore than 400 research
publications by faculty\n\nConducts conferences, industrial visits, guest lectures, seminars and workshops\n
\nPlacements:\n\nApart from the placement facility, students go on industrial tours, attend personality development
programmes (PDP) and participate in placement related activities like extempore, mock interviews, technical tests,
aptitude tests and other latest trends etc.', 'To create promising career opportunities for Vel Tech students in
reputed corporate nationally and internationally.', 'To provide right job for the right person in right place at
the right time.', 'To ensure our students will start the career and move forward in the right direction for better
quality living.', 'To provide ultimate satisfaction to our valuable students by offering the companies of their
choice according to their eligibility.', 'To provide career guidance through counseling and one to one interactions
with Average and below average students.', 'To enhance the skills of students for Industry ready employer.', 'To
better our previous best achievement year on year.', 'To educate our first year students regarding the interview
process through co ordination in placement drives.', 'To instill professional behavioral skills which includes
respects values, ethics, and beliefs and maintains confidentiality.', 'Finally to align with our Founder cum
President Vel Shri R .Rangarajan's ultimate objective of becoming Veltech as the top University in the Nation in
terms of quality and quantity placement.', 'Admission Procedure:\n\nVTUEEE 2017 is an All India Engineering
```

FIGURE 23.2 Scraping the website data.

```
!"#$%&'()*+,-./:;<=>?@[\]^_`{|}~
{33: None, 34: None, 35: None, 36: None, 37: None, 38: None, 39: None, 40: None, 41: None, 42: None, 43: None, 44:
None, 45: None, 46: None, 47: None, 58: None, 59: None, 60: None, 61: None, 62: None, 63: None, 64: None, 91: None,
92: None, 93: None, 94: None, 95: None, 96: None, 123: None, 124: None, 125: None, 126: None}
['vel', 'tech', 'high', 'tech', 'dr', 'rangarajan', 'dr', 'sakunthala', 'engineering', 'college', 'admission',
'2020', 'established', 'year2002', 'veltech', 'hightech', 'dr', 'rangarajan', 'dr', 'sakunthala', 'engineering',
'college', 'is', 'located', 'at', 'chennai2019', 'admissions', 'startedadmission', 'process', 'eligibility',
'call', '9700019482', 'the', 'institute', 'was', 'established', 'in', '2002', 'vel', 'tech', 'was', 'set', 'up',
'by', 'the', 'vel', 'tech', 'group', 'of', 'educational', 'institutions', 'that', 'was', 'set', 'up', 'in', 'the',
'year', '1990', 'by', 'the', 'celebrated', 'industrialist', 'duo', 'couple', 'dr', 'r', 'rangarajan', 'and', 'dr',
'r', 'sakunthala', 'rangarajan', 'they', 'have', 'been', 'in', 'the', 'public', 'service', 'for', 'many', 'years',
'and', 'they', 'are', 'experienced', 'too', 'the', 'major', 'facilities', 'are', 'library', 'with', 'a', 'huge',
'collection', 'of', 'books', 'magazines', 'and', 'ejournals', 'computer', 'centre', 'seminar', 'halls',
'convention', 'hall', 'auditorium', 'separate', 'hostels', 'for', 'boys', 'and', 'girls', 'medical', 'facilities',
'transportation', 'facilities', 'for', 'day', 'scholars', 'etc', 'the', 'major', 'recruiters', 'are', 'adithya',
'tradings', 'pvt', 'ltd', 'american', 'mega', 'trends', 'convergys', 'information', 'esab', 'india', 'pvt', 'ltd',
'india', 'bulls', 'ltd', 'midas', 'communication', 'siemens', 'information', 'sys', 'ltd', 'vernalis', 'systems',
'etc', 'the', 'students', 'of', 'the', 'institute', 'undertake', 'various', 'cultural', 'and', 'sports',
'activities', 'facilities', 'classroom', 'facilities', 'the', 'institute', 'provides', 'separate', 'hostel',
'facilities', 'for', 'boys', 'and', 'girls', 'laboratory', 'facilities', 'the', 'classrooms', 'lecture', 'halls',
```

FIGURE 23.3 Splitting sentences into words.

23.6 CONCLUSION

This research work of improving the chatbot learning ability (or) capacity of that bot's Knowledge was enabled by using Python and AI. It is well clear that this Smartbot is domain independent and answers the user's query almost precisely and efficiently. Thus, an efficient self-learning domain independent Smartbot was developed using NLP and speech recognition and it was integrated with a website that provides information for experimental purposes. And it shows, the response is almost precise and on point.

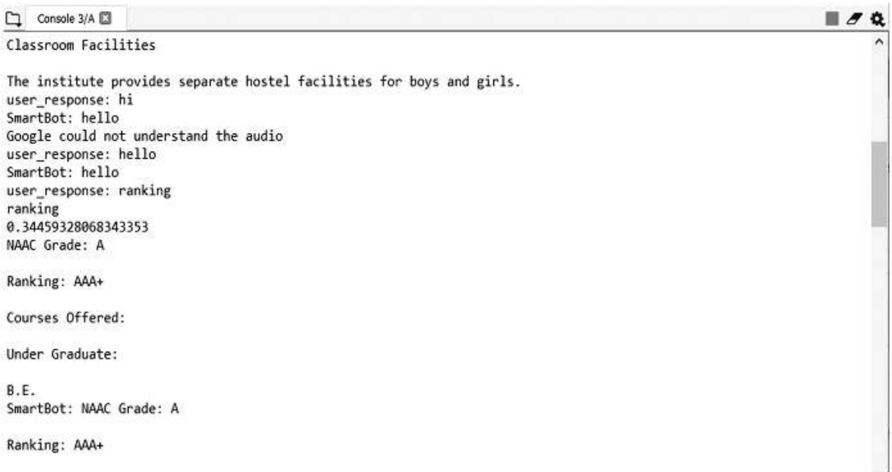


FIGURE 23.4 Greeting the bot and asking query 1.

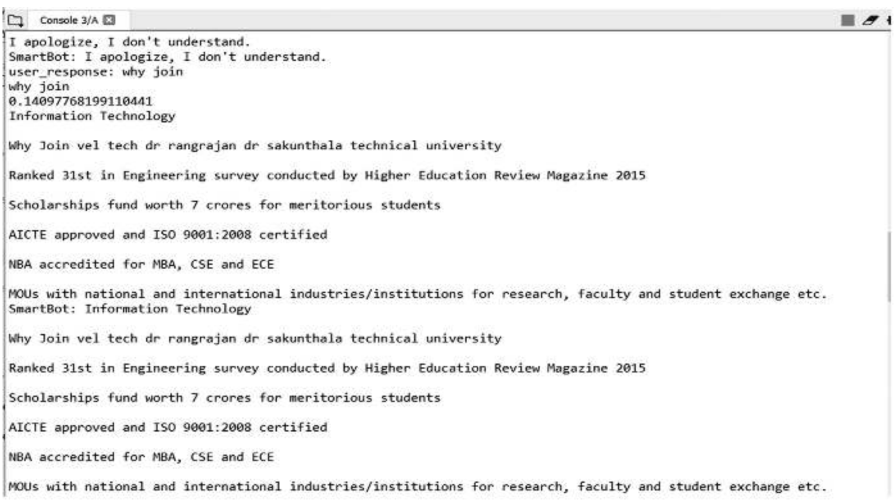


FIGURE 23.5 Query results from the bot.

23.7 FUTURE ENHANCEMENT

Since this bot provides an appropriate response for the user who raises different queries against it. The data scraped from the website can only be possible If the web page is in the right text format. Hence, it is quite difficult to get data from the website which incorporates various levels of data abstractions.

KEYWORDS

- **natural language processing**
- **speech recognition**
- **tokenization**
- **chatbots**
- **artificial intelligence**
- **supervised machine learning**
- **unsupervised ML**

REFERENCES

1. Wang, Y.; Rong, W.; Ouyang, Y.; and Xiong, Z. Augmenting Dialogue Response Generation with Unstructured Textual knowledge. *IEEE Transa. RNN*, **2019**.
2. Divya, S.; Indumathi, V.; Ishwarya, S.; Priya Sankari, M.; Kalpana Devi, S. A Self-diagnosis Medical Chatbot Using AI. *J. Web Dev. Design.* **2019,** *3* (1).
3. Sharma, V.; Goyal, M.; Malik, D. An Intelligent Behavoiur Shown by Chatbot System. **2017,** *3* (4), 2454–4116.
4. Thakur, N.; Hiwrale, A.; Selote, S.; Shinde, A. Arificially Intelligent Chatbot. *URR* **2017,** *4* (6).
5. Bani, B. S.; Singh, A. P. College Enquiry Chatbot Using A.L.I.C.E. *Int. J. N. Technol. Res.* **2017,** *3* (1).
6. Kades, S. A. A.; Woods, J. Survey on Chatbot Design Technique in Speech Conversation Systems. *IJACSA* **2015,** *6* (7).
7. Daniel, G.; Carbot, J.; Derivelle, L.; Derras, M. XALKIT: A Multimodel Low Code Chatbot Development Framework. *IEEE Transac. AI.* **2016**.
8. Al-Khalifa, H. S. A System for Decoding and Coloring Arabic Text for Foreign Language Learners. *IEEE Trans. Nat. Lang. Proc.* **2019,** *7*.

9. Pawar, A.; Mago, V. Challenging the Boundaries of Unsupervised Learning for Similarity. *IEEE Trans. Nat. Lang. Process.* **2017,** *7* (1).
10. Mathew, R. B.; Varghese, S.; Joy, S. E.; Alex, S. S. In *Chatbot for Disease Prediction and Treatment Recommendation using Machine Learning*, IEEE Conference on Trend in Electronics and Informatics, 2019.
11. Carlos de Olivera, J.; Santos, D. H.; Neto, M. P. In *Chatting with Arduino Platform through telegram Bot*, IEEE Conference on Consumer Electronics, 2016.
12. Zhang, W.; Wang, Ren, K.; Song, J. In *Chinese Sentence Based Lexical Similarity Measure for Artificial Intelligence Chatbot*, ECAI, International Conference, 8th ed., 2016.
13. Tapsai, C.; In *Information Processing and Retrival From CSV File by Natural Language*, IEEE 3rd International Conference on Communication and Information System, 2018.
14. Li, Y. Y.; Rafiei, D. In *Natural Language Data Management and Interfaces: Recent Development and Open Challenges*, IEEE Conference Paper, 2017.
15. Xi, Q. G.; Zhou, X.; Wong, J.; Gao, X.; Xichen, C. Matching Real World Facilities to Building Information Modelling Data Using Natural Language Processing. *IEEE Trans. Big Data Analy.* **2019**.
16. Kumar, A.; Mittal, V. Speech Recognition: A Complete Perspective. *Int. J. Recent Technol. Eng.* **2019,** *7* (6C).
17. Bhattacharya, S.; Borah, S.; Mishra, B. K. Emotion Detection from Multilingual Audio using deep Analysis. *Multimed. Tools Appl.* **2022,** *81,* 41309–41338. DOI: https://doi.org/10.1007/s11042-022-12411-3
18. Banerjee, N.; Borah, S.; Sethi, N. Intelligent Stuttering Speech Recognition: A Succinct Review. *Multimed. Tools Appl.* **2022,** *81,* 24145–24166. DOI:https://doi.org/10.1007/s11042-022-12817-z

CHAPTER 24

APPLICATION OF MULTILAYER PERCEPTRON IN SPEECH EMOTION RECOGNITION

NAMAN GUPTA and NIKUNJ AGARWAL

Symbiosis Institute of Computer Studies and Research, Symbiosis International (Deemed University), Atur Centre, Gokhale Cross Road, Model Colony, Pune, Maharashtra, India

ABSTRACT

Emotion is the state of mind based on which we express ourselves. Emotion recognition is a branch of the speech recognition domain which is witnessing a great leap in artificial intelligence (AI). It can be used in various technologies such as voice assistant (Google Assistant, Alexa, Siri, and Cortana), monitoring suspicious behavior and fraud detection. While there are numerous models to recognize emotions implementing machine learning and deep learning, this project aims to use the multilayer perceptron model to identify the speaker's emotions using audio files as input and discover the multilevel representation of the signal. Multiple datasets were explored as potential input sets and a dataset with 7356 files by 24 actors was selected. A sample of 1440 files was taken as input and the final accuracy of the model using the multilayer perceptron model taking 75% training data and 25% test data came out to be ~73%. This study would help us to build a model which would help in better distinguishing the human emotions in future.

Applied Soft Computing Techniques: Theoretical Principles and Practical Applications. Samarjeet Borah, Ratna Raja Kumar Jambi, Sharifah Sakinah Syed Ahmed, & Mahendra Prabhakar Deore (Eds.)
© 2025 Apple Academic Press, Inc. Co-published with CRC Press (Taylor & Francis)

24.1 INTRODUCTION

Artificial intelligence (AI)[1] has reached to a phenomenon level in today's world, where it covers almost every aspect of life. This evolution of AI and Machine Learning[2] has led us to a technology that is one of the most called-for systems at the current time, the speech emotion recognition model.[3] The very basic idea of this model is that it reads audio files as input signals and predicts the human emotions expressed in them.

Out of all the methods by which we communicate, speaking is the fastest and most natural method of communication among human beings. Hence, speech recognition[4] could be a fast and efficient way of interacting with machines as well. The significance of emotion recognition through human speech has inclined recently to maintain the efficiency and convenience of the human–machine interaction.[5] Researchers are continuously working on the detection of emotions from speech signals.[6,7].

The workflow of our model begins with the system reading the input of various audio files using different libraries—soundfile and librosa. Then Mel Frequency Cepstral Coefficient (MFCC), Chroma, and Mel features[8] are extracted from the input sound signal. The modules will be trained, and the trained model will be loaded followed by the process of audio recognition where every audio file will be processed. After saving these files the major process of feature extraction will take place and on its successful completion the emotions will be predicted along with the accuracy and classified report.

24.2 BACKGROUND INFORMATION

24.2.1 SPEECH RECOGNITION

Speech recognition or speech-to-text is the technology which recognizes human speech and converts them into readable text. Through major developments in the sphere of AI and Machine Learning, 'Automatic Speech Recognition'[10] has been getting better and more accurate. This has created a new industry in the form of personal assistant through voice input such as Google Home, Amazon's Alexa, Apple's Siri, and Windows Cortana. Communicating with machines hands-free through voice[11] is more convenient and has a great scope for the future.

24.2.2 EMOTION RECOGNITION

Emotion recognition delves into methodologies of recognizing human emotions. With an emotion recognition system, we can detect the emotions based on either facial expressions or through voice inputs. There are various techniques to capture and process emotions such as machine learning, signal processing, and artificial neural networks[13]. Facial emotion recognition has various applications in domains such as psychology, fraud or lie detection, feedback systems, and monitoring suspicious behavior. Multilayer perceptron (MLP) is a type of artificial neural network (ANN) that falls under the umbrella of soft computing. MLPs are a popular class of feedforward neural networks that consist of multiple layers of interconnected neurons. They are capable of learning complex patterns and making nonlinear decisions.

24.2.3 SPEECH EMOTION RECOGNITION

Speech emotion recognition (SER)[13] is the segment of machine learning which deals in recognizing and inferring human emotions based on audio input signals. Advancements in emotion detection will improve many preexisting systems and provide more accurate results. Determining the emotions on audio signals is a challenging task as different people show emotions differently, also lack of data in various languages makes major obstacles in building a trustworthy system.

24.2.4 APPLICATIONS OF EMOTION RECOGNITION

Emotion recognition can be applied for better analyzing and treating patients with mental illnesses.[14] SER can be applied to improve the functionalities of voice assistants and speech-to-text translation. It can be used for detecting suspicious behavior in crowded or ceremonial places to prevent public unrest. Lie detectors can be equipped with emotion recognition technology to further amplify the system effectiveness.

24.3 LITERATURE REVIEW

McCarthy et al. in the Dartmouth College summer research paper gave proposals for research papers on various applications of AI. Claude E.

Shannon proposed to work on applying the concepts of information theory to brain models and computers by classification of environmental models and using a mathematical model to express it. Marvin L. Minsky came up with the idea of developing a machine that helps in pairing sensory and motor abstractions together to create sensory scenarios that depict the alterations in the environment that may be anticipated if the associated motor act occurred. Nathaniel Rochester gave a proposal to create an artificial language to program a computer to solve problems such as conjecture and self-reference.

Raghu Vamsi et al. proposed an emotion recognition system using deep learning and MLP. They worked on three different datasets with a total of 5732 unique audio files with features such as pitch, rhythm, energy, loudness, etc. After extracting the features, they applied algorithms and techniques such as the MLP classifier and convolutional neural network (CNN) classification. The flask framework was used to deploy accurate techniques of classification into the web app. About 75% of their dataset was assigned as the training data and 25% was the testing data. The MLP classifier gave an overall accuracy of 69.49% with an accuracy of 84.09 in calm emotion. CNN also performed best on emotional calm with an overall accuracy of 72.59%.

Graves et al.[4] worked on creating a speech recognition model using deep recurrent neural networks (RNNs). They focused on end-to-end training where the RNN is trained to directly map an auditory sequence to a phonetic sequence. They parameterized a differentiable distribution over all feasible phonetic output sequences given an audio input sequence using the network outputs. They described two methods for defining output distribution, which will help them to train their neural network. The two methods they described were connectionist temporal classification and RNN transducer. After which they used beam search to decode RNN transducers and produce an n-best list of potential transcriptions and then they regularized the RNN for good performance. They used two regularizes, weight noise, and early stopping. The test set error for RNNs on the TIMIT phoneme recognition benchmark was 17.7%.

Cowie et al.[5] worked on explaining the applicability of emotion detection in computer-human interactions. They proposed two channels one to transmit implicit messages and the other one to transmit explicit message. To understand the other person emotions, they used techniques that involved analysis and single processing. The implicit channel explains to people "how to take" what is communicated through the explicit channel. When a conversation is in full swing as opposed to brief, stereotypical exchanges, this becomes especially important.

Garg et al.[8] worked on MFCC, Mel, and Chroma-based prediction of emotions from audio speech signals. They developed a SER system for tracking the emotions from the audio speech signals. To develop the emotion recognition system, they extracted a lot of audio features and combined them to develop feature vectors. The efficiency of this system is based on the types of classifiers used to detect emotions and what kind of features are extracted. They created vectors using different combinations of features. They also compared the performances using confusion metrics. Their approach seeks for the optimum classification technique and significantly enhances emotion prediction. The Multilayer Layer Perceptron model gave them an accuracy of 47.18%. CNN model gave an accuracy of 60.6%. The long short-term memory network model gave an accuracy of 51.29%.

Tang et al.[9] came up with the idea of an extreme learning machine (ELM) for MLP. Self-taught extraction of features and supervised feature classification are the two primary parts of their architecture, and hidden weights with random initialization are used to connect them. For feature extraction, they used unsupervised multilayer coding along with ELM sparse autoencoder as it achieves more condensed and insightful feature visualization. Before decision-making, they randomly projected the order-wise encoded code to achieve better and faster learning. The hidden layers of their architecture are taught in forward way as compared to that of deep learning (Greedy Layer-wise). The weight of the present layer is fixed after establishing the previous layers, this is done to get better learning efficiency.

Park et al.[10] came up with SpecAugment which is a data augmentation technique for automated speech recognition. A neural network's feature inputs, filter bank coefficients, are directly subjected to SpecAugment. The features are warped, frequency channel blocks are hidden, and time step blocks are hidden as part of the augmentation policy. For peer-to-peer speech recognition tasks, they used SpecAugment on the attend, listen and spell networks. Without using any language model, they got 6.8% WER on LibriSpeech, while along with shallow fusion they achieved 5.8% WER. In contrast, the prior cutting-edge hybrid system had a WER of 7.5%. they got 7.2%/14.6% on the Switchboard component of the Hub5'00 test set for switchboard without using any language model and 6.8%/14.1% along with shallow fusion, compared to 8.3%/17.3% WER for the prior state-of-the-art hybrid system.

Schofar[11] researched the scientific foundations of voice communication between Humans and Machines. To identify areas for potential future research, this paper emphasized on the fundamental scientific and technological problems in voice-based human–machine communication. The major topics covered by him are the implementation of the system's hardware and software, speech understanding and recognition for voice input, voice output using speech synthesis, and usability elements pertaining to how people interact with machines.

Li et al.[14] came up with an emotion recognition system with the use of ANN and voice processing technology to incorporate existing instruments to recognize patients' and medical professionals' emotions in clinics. They created a database of emotional speech in Taiwanese, and each doctor's voice sample was also incorporated with their patient's voice sample, this sample database was used to build the training model. They categorized the 610 voice samples into four emotions neutral, happy, sad, and angry. They preemphasized the voice samples to highlight the parts with high frequency using a high pass filter after which they applied MFCC as eigenvalues for the extraction of speech input signals features. They imported around 40 filters from a Mel-scale triangular bandpass filter bank to ensure a smooth frequency spectrum and reduce the harmonic effect, bringing attention to the fundamental components of the original audio sample. The ANN MATLAB neural pattern recognition classifier was used in this study to classify the vectors. They assigned 70, 15, and 15% of the data to the training, testing, and validation sets, respectively. The accuracy of their emotion recognition model on the testing data was 91.3%.

Hossain and Muhammad[15] proposed an emotion recognition system using deep learning model with audio–visual emotional big data, which presents in what way emotions can be perceived and inferred taking audio and video input signals. They used two different datasets for this, first being a big data dataset consisting of both audio and video signals, and the second being eNTERFACE database. For the audio, they extracted the speech input signals features for obtaining the Mel spectrogram, that is fed to a two-dimensional (2D) CNN and ELM for fusing the scores. Some representational parts from the video input are obtained and served to the 3D CNN and ELMs for the fusion of scores. Both outputs are then passed through a Support Vector Machine Algorithm to obtain the concluding output for classifying emotions based on speech and video input signals.

Using neural networks, Pawan Kumar Mishra and Arti Rawat[16] put forward a SER model which give a description of how to recognize emotions through audio input signals. In their study, they have proposed to filter the sound based on a High Pass Filter which would remove any unwanted noise from the input signal. They used MFCC for extracting the features from the input sound, and then passed it through Neural network for classifying the final emotions as the output.

Xu et al.[17] implemented an Anatomically Constrained Neural Network model using the head fusion method for creating an attribute known as Xfusion, taking the MFCC as input features for recognizing emotion through speech input signals and validating the results studying four emotions—neutral, sad, angry and excited.

Mekruksavanich et al.[18] projected a system for determining negative or positive emotions in Siamese or the Thai language. They used the deep learning approach using two 1D CNN and trained their models on multiple datasets–Toronto Emotional Speech Set (TESS), Ryerson Audio-Visual Database of Emotional Speech (RAVDESS), Crema-d, and Surrey Audio-Visual Expressed Emotion (SAVEE) separately for classifying the input into positive or negative emotion. After successful completion, they tested their model on Thai input datasets.

Xu et al.[19] put forward a SER model using multiscale area attention as well as data augmentation, in which they used the multiscale area attention to the model for designing an attention-based CNN. In this system, they first used the librosa library to draw out the Mel-based frequency and features. They are then sent to multiple hidden layers whose output is fed to four convolution layers which generates an 80-channel representation, which are further forwarded to the attention layer before finally classifying them into emotions.

24.4 PROPOSED METHODOLOGY

MLP is a type of ANN that falls under the umbrella of soft computing. MLPs are a popular class of feedforward neural networks that consist of multiple layers of interconnected neurons. They are capable of learning complex patterns and making nonlinear decisions. The proposed methodology is using MLP in SER and can be discussed as follows:

Data Preprocessing:

- Collect a dataset of speech recordings labeled with the corresponding emotion categories.
- Preprocess the audio data by converting the raw audio signals into a suitable format for analysis, such as MFCC or spectrograms.
- Split the dataset into training, validation, and testing sets.

Feature Extraction:

- Extract relevant features from the preprocessed audio data. Common features used in SER include MFCCs, spectral contrast, chroma features, and pitch.
- Optionally, you can apply dimensionality reduction techniques such as principal component analysis (PCA) or linear discriminant analysis (LDA) to reduce the feature space.

Data Representation:

- Convert the extracted features into a suitable format for feeding into the MLP. This usually involves vectorizing the features and normalizing the data to a common scale (e.g., between 0 and 1).

Model Training:

- Initialize the MLP model with the chosen architecture.
- Train the model using the training set. This involves feeding the input data forward through the network and calculating the loss using a suitable loss function.
- Monitor the model's performance on the validation set and adjust hyperparameters (e.g., learning rate, batch size, number of hidden layers) as necessary to improve performance.
- Repeat the training process until the model converges or reaches a satisfactory level of performance.

Model Evaluation:

- Evaluate the trained MLP model on the testing set to assess its performance and generalization ability.
- Calculate metrics such as accuracy, precision, recall, and F1 score to measure the model's performance across different emotion categories.

24.4.1 WORKFLOW DIAGRAM

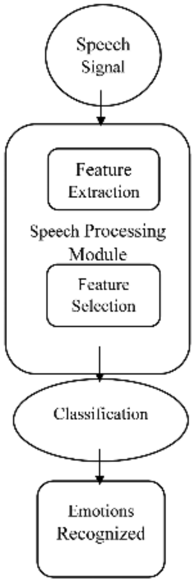


FIGURE 24.1 Proposed methodology.

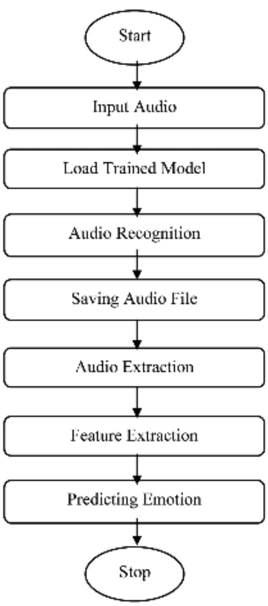


FIGURE 24.2 Workflow diagram.

24.4.2 FEATURE EXTRACTION

As a machine cannot understand voice input, it must be first converted into digital data. In feature extraction, the voice input signals are first transformed into digital data using the librosa package which has many libraries for extracting and processing audio data. The most important feature is the MFCC,[20] which analyses the data based on the short-term spectrum of sound. It breaks down the input into frames and identifies which frequencies are present in each frame, which are then passed for the mapping on the Mel scale. This helps in understanding and processing the audio signal better and increases the robustness in determining the emotions. The second feature is the Mel scale, which provides information about the frequency of the input signal on the Mel scale. The next feature is Chroma which identifies the properties of sound based on the pitch of the provided input file. Tonnetz is used to render Chroma on a 6D basis which represents the perfect fifth, minor third, and major third each as 2D coordinates. The fifth feature is contrast which depicts the highs and lows in the pitch of the sound file.

24.4.3 FEATURE SELECTION

Feature selection[21] is the method by which we reduce the number of input values while developing a predictive model. It is important as it not only reduces the complexity of the system, but also reduces the time and cost in implementation and execution of the model. For our model, we would take only the most required features of the MFCC (~40), 12 features of Chroma, and all 128 features of Mel. A total of about 180 features from each audio input file would be given to the model for training and testing the data.

24.4.4 MULTILAYER PERCEPTRON

MLP[22–24] an architecture of ANN based on supervised learning. In the perceptron learning rule, an additional input signal known as bias is provided which helps to strengthen the signal. It contains three major layers–The input layer for providing the input signals, Hidden layer for applying weights and bias to the input signal and sending them to the activation function, and output layer which gives the output. MLP model is also known as the backpropagation algorithm. In this, if the goal state is not reached, then the weights are recalculated based on the current output and then it is applied for the next iteration.

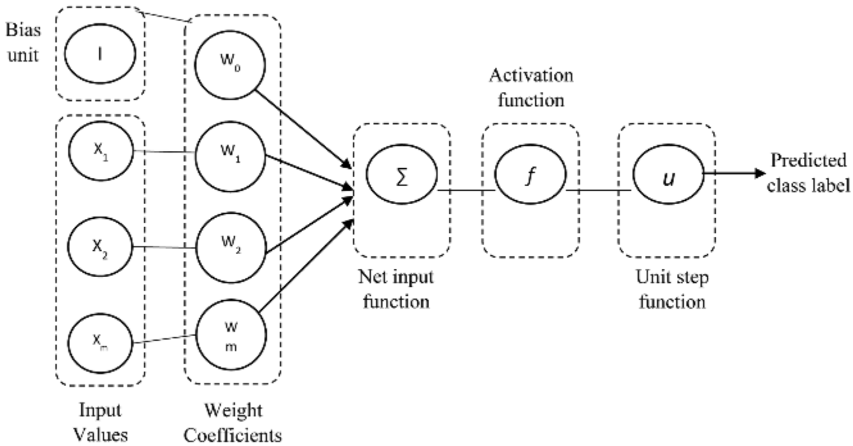


FIGURE 24.3 Perceptron model.

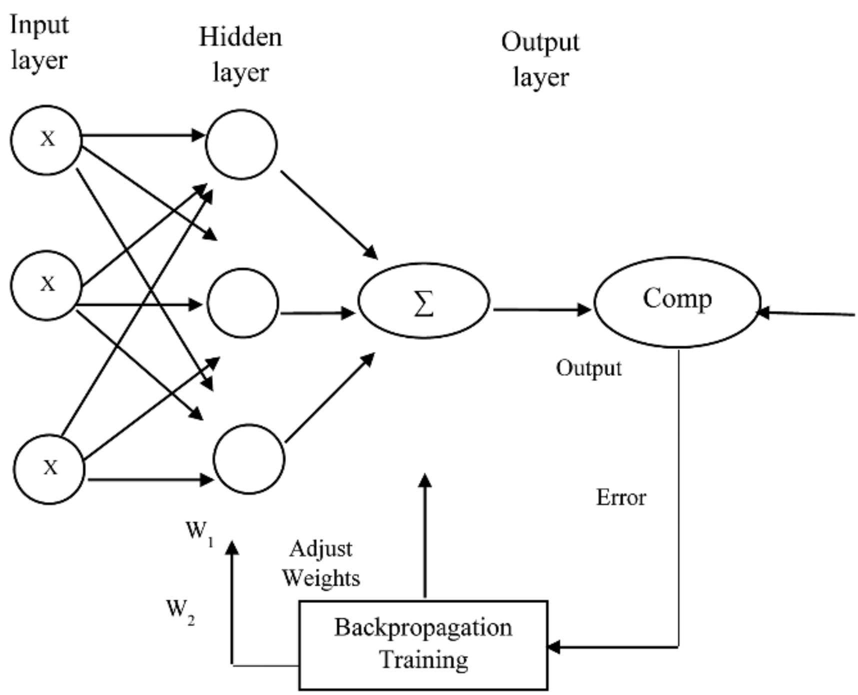


FIGURE 24.4 Backpropagation algorithm.

24.5 RESULTS AND ANALYSIS

The model was applied to the dataset using cross-validation. The dataset was into training data and testing data in the range of 60–85% and 15–40%, respectively. The highest accuracy was reached while keeping the training and testing data at 75 and 25% respectively, which came out to be around 73%. Taking training data more than this would result in overfitting issues and the output would not be reliable.

The following tables show the outputs obtained at different amounts of training data and testing data:

TABLE 24.1 Result at Training Data= 60% and Testing Data = 40%.

	Precision	Recall	f1-score	Support
Angry	0.75	0.72	0.74	72
Happy	0.62	0.67	0.64	83
Neutral	0.5	0.5	0.53	41
Sad	0.65	0.56	0.6	73
Accuracy			0.64	269
Macro average	0.63	0.63	0.63	269
Weighted average	0.64	0.64	0.64	269

TABLE 24.2 Result at Training Data = 65% and Testing Data = 35%.

	Precision	Recall	f1-score	Support
Angry	0.81	0.7	0.75	63
Happy	0.64	0.74	0.69	73
Neutral	0.52	0.43	0.47	37
Sad	0.58	0.62	0.6	63
Accuracy			0.65	236
Macro average	0.64	0.62	0.63	236
Weighted average	0.65	0.65	0.65	236

The following table shows the accuracy of execution for different sets of training and testing datasets.

On changing the considered emotions from angry, happy, neutral, and sad to angry, calm, fearful, and sad, the following results were obtained with 75% training data and 25% testing data:

TABLE 24.3 Result at Training Data = 70% and Testing Data = 30%.

	Precision	**Recall**	**f1-score**	**Support**
Angry	0.74	0.74	0.74	53
Happy	0.69	0.62	0.62	66
Neutral	0.41	0.37	0.37	30
Sad	0.52	0.62	0.57	53
Accuracy			0.61	202
Macro average	0.59	0.59	0.59	202
Weighted average	0.62	0.61	0.61	202

TABLE 24.4 Result at Training Data = 75% and Testing Data = 25%.

	Precision	**Recall**	**f1-score**	**Support**
Angry	0.83	0.62	0.71	32
Happy	0.49	0.63	0.55	27
Neutral	0.3	0.21	0.25	14
Sad	0.56	0.64	0.6	28
Accuracy			0.57	101
Macro average	0.55	0.53	0.53	101
Weighted average	0.59	0.57	0.57	101

TABLE 24.5 Accuracy at Different Proportions of Dataset.

Sr. no	**Training data (%)**	**Testing data (%)**	**Accuracy (%)**
1	60	40	63.94
2	65	35	64.83
3	70	30	61.39
4	75	25	73.81

TABLE 24.6 Result on Changing the Emotions; Overall Accuracy = 80.21%.

	Precision	**Recall**	**f1-score**	**Support**
Angry	0.75	0.78	0.75	40
Calm	0.89	0.96	0.93	52
Fearful	0.79	0.73	0.76	45
Sad	0.78	0.73	0.75	55
Accuracy			0.8	192
Macro average	0.8	0.8	0.8	192
Weighted average	0.8	0.8	0.8	192

On changing the considered emotions and making some tweaks to the model, the accuracy was changed a lot. It can be noted that the model finds it difficult to identify neutral emotions correctly. On changing the considered emotions, accuracy of other emotions was also affected, meaning that the model is dependent on all parameters and the result changes accordingly.

24.6 CONCLUSION

In this paper, we used MLP model on audio input signals using the MFCC approach to recognize human emotions. It is shown that the model's performance changes with respect to changes in the emotions being detected. In future works, more emotions can be added to the model along with increase in the input dataset. By using the basic principles of machine learning, it is still not possible to detect emotions in uncontrolled environment with high accuracy. In reality, emotions are very subjective and differ from person to person. It is highly difficult for a machine to analyze and infer what someone's mood or state of mind is based on audio input alone. In the upcoming years, it will be necessary to have a system that can classify emotions via speech accurately and reliably. However, with the current advancements in the subject of machine learning and AI, it would not be much longer before a system would be developed that could accurately identify emotions just through speech inputs.

KEYWORDS

- **speech signal**
- **speech recognition**
- **artificial intelligence**
- **multilayer perceptron model**
- **emotion**

REFERENCES

1. McCarthy, J.; Minsky, M. L.; Rochester, N.; Shannon, C. E. A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. Published in 1955, Connected to the Birth of Artificial Intelligence.

2. Pitts, W.; McCulloch, W. A Logical Calculus of the Ideas Immanent in Nervous Activity. 1943.
3. Vamsi, U. R.; Chowdhary, B. Y.; Harshitha, M.; Theja, S. R.; Divya Udayan, J. Speech Emotion Recognition (SER) using Multilayer Perceptron and Deep learning techniques. *GITAM Institute Technol.* **2021,** *27* (5), 386–394.
4. Graves A.; Mohamed A.-R.; Hinton G. In *Speech Recognition with Deep Recurrent Neural Networks*, ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing–Proceedings, 6638947, pp. 6645–6649, 2012.
5. Cowie, R.; Douglas-Cowie, E.; Tsapatsoulis, N.; Votsis, G.; Kollias, S.; Fellenz, W.; Taylor, J. G. Emotion Recognition in Human-computer Interaction. *IEEE Signal Process. Mag.* 6892132 **2001,** *18* (1), 32–80.
6. Banerjee, N.; Borah, S.; Sethi, N. Intelligent Stuttering Speech Recognition: A Succinct Review. *Multimed. Tools Appl., Springer* **2022,** *81,* 24145–24166. https://doi.org/10.1007/s11042-022-12817-z
7. Bhattacharya, S.; Borah, S.; Mishra, B. K.; Mondal, A. Emotion Detection from Multilingual Audio Using Deep Analysis, Multimedia Tools and Applications **2022,** *81,* 41309–41338, https://doi.org/10.1007/s11042-022-12411-3
8. Garg, U.; Agarwal, S.; Gupta, S.; Dutt, R.; Singh, D. In *Prediction of Emotions from the Audio Speech Signals using MFCC, MEL and Chroma*, 12th International Conference on Computational Intelligence and Communication Networks, CICN 20209242635, pp. 87–91, 2020.
9. Tang, J.; Deng, C.; Huang, G.-B. Extreme Learning Machine for Multilayer Perceptron. *IEEE Trans. Neural Netw. Learn. Syst.* **2016,** *27* (4), 7103337, 809–821.
10. Park, D. S.; Chan, W.; Zhang, Y.; Chiu, C.-C.; Zoph, B.; Cubuk, E. D.; Le, Q. V. In *SpecAugment: A Simple Data Augmentation Method for Automatic Speech Recognition*, Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, 2019-September, pp. 2613–2617.
11. Schofar, R. W. Scientific Bases of Human–Machine Communication by Voice. *Proc. Natl. Acad. Sci. U. S. A.* **1995,** *92* (22) 9914–9920.
12. Schmidhuber, J. Deep Learning in Neural Networks: An Overview. *Neural Netw.* **2015,** *61,* 85–117.
13. Sapra, A.; Panwar, N.; Panwar, S. Emotion Recognition from Speech. *Int. J. Emerg. Technol. Adv. Eng.* **2013,** *3* (2), 341–345.
14. Li, H.-C.; Pan, T.; Lee, M.-H.; Chiu, H.-W. Make Patient Consultation Warmer: A Clinical Application for Speech Emotion Recognition. *Appl. Sci.* **2021,** *11* (11), 4782.
15. Hossain, M. S.; Muhammad, G. Emotion Recognition Using Deep Learning Approach from Audio-Visual Emotional Big Data. *Inform. Fusion* **2019,** *49,* 69–78.
16. Mishra, P. K. Rawat, A. Emotion Recognition through Speech Using Neural Network. *Int. J. Adv. Res. Comput. Sci. Software Eng. (IJARCSSE)* **2015,** *5* (5), 422–428.
17. Xu, M. et al.; In *Improve Accuracy of Speech Emotion Recognition with Attention Head Fusion*, 2020 10th Annual Computing and Communication Workshop and Conference (CCWC). IEEE, pp. 1058–1064, 2020.
18. Mekruksavanich, S.; Jitpattanakul, A.; Hnoohom, N. In *Negative Emotion Recognition using Deep Learning for Thai Language*, Proceedings of the Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT and NCON), Pattaya, Thailand, pp. 71–74, 2020.

19. Xu, M. et al.; Speech Emotion Recognition with Multiscale Area Attention and Data Augmentation. arXiv:abs/2102.01813, 2021.
20. Prabakaran, D.; Sriuppili, S. Speech Processing: MFCC based Feature Extraction Techniques-An Investigation. *J. Phys.: Conf. Ser.* **2021,** *1717* (1), 012009.
21. Rong, J.; Li, G.; Chen, Y. P. P. Acoustic Feature Selection for Automatic Emotion Recognition From Speech. *Inf. Process. Manag.* **2009,** *45* (3), 315–328.
22. Pal, S. K.; Mitra, S. Multilayer Perceptron, Fuzzy Sets, and Classification. *IEEE Trans. Neural Netw.* **1992,** *3* (5), 683–697.
23. Chaudhuri, B. B.; Bhattacharya, U. Efficient Training and Improved Performance of Multilayer Perceptron in Pattern Classification. *Neurocomputing* **2012,** *34* (1–4), 11–27.
24. Gori, M.; Scarselli, F. Are Multilayer Perceptrons Adequate for Pattern Recognition and Verification? *IEEE Trans. Pattern Anal. Mach. Intell.* **1998,** *20* (11), 1121–1132.

CHAPTER 25

BANGLA AND ODIA MACHINE TRANSLATION USING EM ALGORITHM: EXPERIMENTAL EVALUATION

BISHWA RANJAN DAS[1], HIMA BINDU MARINGANTI[1], and NILADRI SEKHAR DASH[2]

[1]*Department of Computer Application, MSCB University, Baripada, Odisha, India*

[2]*Linguistic Research Unit, Indian Statistical Institute, Kolkata, West Bengal, India*

ABSTRACT

Expectation–maximization (EM) is a critical and very accurate machine translation method used in the Bangla–Odia system. Here, the entire mathematical equation is computed and displayed using a corpus of Bangla–Odia words. To determine the probability value with the highest degree of likelihood, the "argmax" function, which examines the one-to-one, one-to-many, and many-to-many mappings between two or more words in source and target language phrases, is used in conjunction with this EM technique. A mathematical formula can be used to determine the lexical connections between the terms of the parallel sentences to show which word in the target language is congruent with which word in the source language. Because the EM technique loops, it is simple to compute specific probability values in terms of maximum likelihood estimation (MLE) and then iteratively determine the word associations between the source and destination languages. Because

Applied Soft Computing Techniques: Theoretical Principles and Practical Applications. Samarjeet Borah, Ratna Raja Kumar Jambi, Sharifah Sakinah Syed Ahmed, & Mahendra Prabhakar Deore (Eds.)
© 2025 Apple Academic Press, Inc. Co-published with CRC Press (Taylor & Francis)

machine translation is currently one of the main focuses of natural language processing (NLP) research work with some practical obstacles, determining the MLE or maximal a posterior (MAP) of the parameters in the probability model is a very challenging task. The purpose of this study is to describe the different lexical issues that can arise when reading bilingual materials that were translated from Bangla into Odia. The way the EM algorithm and MLE method handle the “word divergence” or “lexical divergence” problem is one of the fundamental aspects of the MT system. The MLE technique is employed in part to enhance the probability values and provide more precision, even if the EM algorithm is used to solve it. Bilingual dictionaries or lexical databases are employed in this situation because they are typically used at the phrase level to address word divergence or lexical divergence difficulties. The biggest challenge is in identifying single-word units from the source text that are transformed into multiword units in the target text. The findings of the trials leverage the bilingual Bangla–Odia corpus, which was produced using the EM algorithm, to pinpoint a productive combination for accurately resolving the lexical divergence problem.

25.1 INTRODUCTION

A method for finding the precise and comparable term between two parallel corpora is word alignment. This exemplifies one of the interactions between parallel phrases that include word translation. Word divergence is the process through which a term is occasionally translated by one or more words. The main goal of word alignment, assuming parallel sentences are provided, is to determine the correspondence between words in source and target sentences that may be one-to-one, one-to-many, or many-to-many. Aligning similar phrases in the source and destination languages is the answer to phrase-based translation. Words from the source phrase that cannot be adequately translated into the target language are simply given the value “null.” Another component of word alignment is moving the translated words from the source sentence to the proper place in the target phrase. The use of multilingual machine translation may need word reordering. Word alignment facilitates this. The biggest challenge is in identifying single-word units from the source text that are transformed into multiword units in the target text. The findings of the trials leverage the bilingual Bangla–Odia corpus, which was produced using the EM algorithm, to pinpoint a productive combination for accurately resolving the lexical divergence problem.

25.2 RELATED WORK

In Ref. [1], the specifics of word alignment, many strategies are used, including a hybrid strategy that groups words locally in Hindi phrases while also including additional techniques such as dictionary searches, comparing transliterations, predicting English terms, and finding the closest aligned neighbors. The probabilities between pairs of sentences that are tiny and large are thoroughly examined. Here, a very basic description of the many topics, issues, problems, and challenges is given. In-depth descriptions of several approaches are also provided. The expectation–maximization (EM) algorithm is used to carefully confront and solve the majority of problems, and statistical techniques are used to accurately and vividly represent the entire notion. Here, most of the challenges and problems are resolved.[4] In this study, the Bangla–Odia lexical divergence problem is resolved using multiple mapping strategies, including one-to-one and many-to-one.[5] Word dictionaries are used to map the sentences' parallel English and Hindi words.[6] Various techniques, including the boundary detection approach, the minimum distance function, and dictionary searches, have been used to perform automatic word alignment. The most crucial aspect of machine translation, which divides a word's entirety into its various meanings, is compound word spitting. The difficulties encountered when translating from one language to another are examined, along with several methodologies and their benefits and drawbacks.[8] In this study, a new probabilistic model for word alignment is presented, where word alignments are connected to alignment categories that are motivated by linguistics. For this goal, a unique semi-supervised learning method is being suggested and implemented for the joint prediction of word alignment and alignment types.[9] A collection of pairs of sentences that are translations of one another are demonstrated, and the idea of word-by-word alignment between such pairs of sentences is defined. A series of five statistical models of the translation process is detailed, along with algorithms for estimating their parameters.[10] The most well-known methods used in statistical machine translation are introduced in detail and with clarity in this book. Many online MT programs, such as Babelfish, Google Translator, Bing Translator, Yahoo, and AltaVista, are available to translate text straight from one language to another.[11] In this study (Moore, 2004), the alignment error rate was reduced by about 30%. Additionally, the probability of alignment to the null word was given extra weight, the probability values for rare words were smoothed, and a straightforward heuristic estimation method was used to initialize, or take the place of, EM training of the model parameters. In addition to the usual lexical and syntactic properties,[12] the semantic relation can be used to enhance the accuracy of the word alignment. Here, a neural network-based algorithm is employed to extract the word similarity from the monolingual data.[13] Jorg T. (2000) created the most significant and practical word aligner, known as UWA, which is used to parallel text PLUG. Both multiword units and single-word units are included in this model, which only refers to link units.[14] This study introduces the "Clue alignment" approach. The word alignment clue links words and phrases together and depends on elements such as frequency, phrase type, part-of-speech (POS) tagging, and actual wordform strings.[15] Word alignment is the relationship between individual words and groups of words in a parallel corpus. The best word alignment can be found using a variety of methods, and all algorithms are evaluated using a word alignment method based on association cues.

25.3 EXPECTATION–MAXIMIZATION (EM) ALGORITHM

When a statistical model depends on unobserved latent variables, an EM method is an iterative procedure or approach to determine the maximum likelihood estimation (MLE) or maximum a posteriori estimation (MAP) of the model's parameters. In the EM iteration, two steps alternately take place: the expectation (E) step, which provides a function for the expected log-likelihood assessed using the most recent estimate for the parameters, and the maximization (M), which computes the parameters that maximize the anticipated log-likelihood determined on the E step. The selection of the best result or value is the main objective of this method.

25.4 PROBLEMS FACED

Different types of problems are faced during word alignments such as, one-to-many, many-to-one, many-to-many, named entity problems, and one having multiple alignments. However, to solve these problems using EM and various mathematical algorithms.

25.5 WORD ALIGNMENT WITH METHODOLOGY

The Bangla–Odia conditional probability model, abbreviated $P(B|O)$, is discussed in this essay. The refers to model parameters for a dataset pair D

of n phrases that are translated from one another and are known as parallel text.; $D = (B_1,O_1), (B_2,O_2), (B_3,O_3), ..., (B_n,O_n)$. In the parallel corpus, every pair is unique from the others. The hidden word pair between the translation pairs for Bangla and Odia is guaranteed to be always detected by the model. The corpus data thoroughly trains the model to anticipate the reality of the misaligned word alignment. There are various methods for defining $P(B|O)$. Let us say that an array of words with index J, $(O_1, O_2, O_3, ..., O_J)$ represents an Odia phrase, while an array of words with index I, $(B_1, B_2, ..., B_I)$ represents a Bangla sentence. Currently, it is thought that every word in Bangla has at least one connection to Odia. Array "*a*" of length I denoted as $[a_1, a_2, a_3, ..., a_i]$, where $a_1, a_2, a_3, ..., a_i$ are one–one alignment variables, is used to express this. *ai* is an alignment variable that ranges from 0 to J. If $a_i = 0$, this indicates that there is a null alignment between the words in Bangla and Odia. Take the sentences Bangla–Odia as an example.

25.5.1 BANGLA SENTENCE

নিজেদের দাবি নিয়ে নির্মাণ কার্য বন্ধ করার জন্য কৃষকদের সংগঠনের মধ্যে আলোড়ন সৃষ্টি হয়েছে।

25.5.2 TRANSLITERATION

"Nijeder dAbi niye nirmAn kAryya bandha karAr janna krishakder sangaThener madhye aloRon srisTi hayechhe".

25.5.3 ODIA SENTENCE

ନିଜର ଅଧିକାରକୁ ନେଇ ନିର୍ମାଣ କାର୍ଯ୍ୟ ବନ୍ଦକରିବାକୁ କୃଷକ ସଂଗଠନଗୁଡ଼ିକରେ ହଟଚମଟ ସୃଷ୍ଟି ହୋଇଯାଇଛି।

25.5.4 TRANSLITERATION

"Nijara adhikaraku nei nirmana karjya bandakaribaku krushaka sangathanagudikare hatachamata shrusti haijaichhi".

The Bangla sentence is 14 words long, while the Odia sentence is 11 words long. Here is a list of the words in both sentences B_1, B_2, B_3, B_I, and O_1, O_2, O_3, and O_J, respectively. The alignment array value '*a*' will be {1, 2, 3, 4, 5, 6, 7, 7, 8, 9, 9, 10, 11}. It is expected that the statistical model uses a typical procedure to translate the Bangla text into Odia. The probability distribution should be used. After finding the value of I, which denotes the length of a Bangla phrase, we obtain $P(I|J)$, in this case $P(14|11)$. Each Odia word position is then aligned to a Bangla word (or null) in accordance with the valid sentence alignment of the standard corpus (ILCI-Indian Language Corpora Initiative, Government of India) $P(a_i = j|J)$. The probability distribution function of the aligned Odia word is then used to translate each Bangla word B_i. As a result, this alignment raises P(Nijara|Nijeder), P(adhikaraku|dabi), i(nei|nie), and so forth.

$$\mathrm{P}\left(B, a \mid O\right) = \mathrm{P}\left(\mathrm{I}|\mathrm{J}\right)\prod_{i=1}^{I}\mathrm{P}\left(a_i|\mathrm{J}\right).\,\mathrm{P}\left(B_i|O_{a_i}\right) \tag{25.1}$$

The aforementioned equation represents the joint likelihood of the Bangla text and its alignment requirement. This is a simple product of all probability values. This equation has three parts: $P(I|J)$, which denotes the length of the Bangla sentence I, the Odia sentence J, and $P(B|O)$, which denotes all word pairings that appear in both Bangla and Odia simultaneously. The range of allowable fractional numbers for word alignment is reflected in the probability values in tabular form.

$$\forall_{O,B} P\left(B\,|O\right) \in \left[0,1\right] \tag{25.2}$$

$$\forall_{O} \sum_{B} P\left(B\,|O\right) = 1 \tag{25.3}$$

25.6 MAXIMUM LIKELIHOOD ESTIMATION

Only focus on $P(B|O)$ and MLE to estimate the approximate value to see the alignment. Before beginning to align the words between Bangla and Odia, the sentence alignment has first been thoroughly discussed. However, unlike what the English-to-French[19] translation indicates, such a circumstance does not occur in Bangla–Odia. For instance, the majority of French words frequently coincide with English words, whereas Bangla–Odia never uses this. To evaluate the plausibility of the input parameters from the perspective of comprehension, an MLE function is described here.

$$\prod_{n=1}^{N} P_\theta\left(B^{(n)}, a^{(n)}|O^{(n)}\right) = \prod_{n=1}^{N} P(I^{(n)}\,|J^{(n)}) \prod_{i=1}^{I^{(n)}} P(a_i^{(n)}\,|J^{(n)} \cdot P(B_i^{(n)}\,|\,O_{a_i}^{(n)}) \tag{25.4}$$

N stands for the number of sentences. "n" is the index number of each and every sentence, "a" is the alignment, "I" is the length of each and every Bangla sentence, and "J" is the length of each and every Odia phrase. Now that data has been collected and parameters have been calculated, the highest value should now be determined using a probability function since our data (value) closely resembles this model.

$$\hat{\theta} = \underset{\theta}{argmax} \prod_{i=1}^{N} P_{\theta}(B^{(n)}, a^{(n)} | O^{(n)}) \tag{25.5}$$

In eq 25.5, where $\hat{\theta}$ For each word in a sentence, it uses the argmax function to find the word alignment with the highest probability value. Machine translation is fundamentally a search problem from an infinite pool of potential texts. After translation in line with the corpus, one sentence is chosen from a variety of potential sentences. A prediction and search difficulty exists here. The model described below is started by an algorithm designed to learn from the suggested aligned data. For each pair of Bangla–Odia words, the data is scanned, the alignments are noted, and the alignments are counted (i.e., aligned data, Bangla–Odia pairs). To acquire the probability values (aligned word pair Bangla–Odia), all counts (means probability values) are normalized by the frequency of the relevant Bangla word appearing in an Odia word. The algorithm described in this article is quite simple.

25.7 ALGORITHM 1

- START
- Initialize all counts to 0
- For each *n* value between 1 and *N*
- For each *i* value between 1 and *I*
    a. For each *j* value between 1 and *J*
        i. Compare *ai* = *j* up to *n*, that is, *i* value
        ii. Count [(*Bi*, *Oj*)] ++
        iii. Count [*Oj*]++
- For each (*Bi*, *Oj*) value in count do
    i. $P(B|O)$ = Count(B,O)/Count(O)
- STOP

The linear alignment variable used in this approach to gather counts is discussed. When the unobserved alignments are left, only the approach that finds all pairs of a word will be relevant, and this algorithm can be expanded. Depending on the length of the sentence, this method counts all word pairings that are quadratic in nature. After a few iterations, two formulas are employed to determine the alignment probability. Although this algorithm is not completely hypothetical, some sentences from the corpus are taken with observed alignments that a small amount of observed data is trained, but this would not produce a good estimate θ. So, the EM algorithm is applied for good estimation of θ. The main problem of MLE for the huge number of sentences, the alignment "*a*" is not observed. So, the EM algorithm is to learn the parameters of a latent variable model as it happened in IBM model 1. EM has two advantages, to replace the previous count (observed count) of alignment links with present counts (expected counts) of alignment links and computed with respect to some previous estimate of θ. Next is to compute the expected counts with respect to each and every previous estimate of θ value and improve it. Two things are necessary, observed count and expected count of alignment links.

There are some extra advantages or inspiration for using EM algorithm from MLE. So need to do slight modification of eq 25.4 to find another equation, as follows:

$$\prod_{n=1}^{N} P_\theta\left(B^{(n)} | O^{(n)}\right) = \prod_{n=1}^{N} P(I^{(n)} | J^{(n)}) \sum_{a^{(n)}} \prod_{i=1}^{I^{(n)}} P(a_i^{(n)} | J^{(n)}).P(B_i^{(n)} | O_{a_i}^{(n)}) \tag{25.6}$$

Here still maximize the likelihood of the observed data, but the observed data is not included the alignments, so just remove the third expression, and expand the fourth expression. In other way, as alignment is not seen, the probabilities values of all possible alignment is added that produced the exact alignment or output because this gives the marginal probability of *B* which is to be maximize. However, there are little bit errors in that the sum over the alignments prevents an analytics solution. However, whatever problems have, that is to be solved algorithmically. However, this algorithm is very similar to MLE to turn out the problem or error as the expected count of the link, is calculated by the Bayes's rule.

$$P(a_i = j | B, O) = \frac{P(B | a_i = j, O) P(a_i = j | O)}{P(B | O)} = \frac{P(B, a_i = j | O)}{P(B | O)} \tag{25.7}$$

The posterior probability above finds the percentage of alignment link between *Bi* and *Oj*. Importance of posterior probability many to know the links percentage. Eq 25.7 says that to find the posterior probability is the sum of the probability values of all of the alignments containing a link between *Bi* and *Oj* divided by the sum of the probability of all the possible alignments. So the resultant value is a probability number. In the MLE case, the link value is incremented by 1, i.e., already linked and counted of the two words each, if the value is zero means there is no link. So, this posterior probability value is always between 0 and 1, that is, formally indicating uncertainty whether a link between B_i and O_j exists or not. If the value is close to 0, then there is no link and if it is close to 1, it means that a link exists. If the value exists in the middle, not certain at all (which means that there exists no link at all). The objective of the EM is to replace counts of observed events with posterior probabilities. These posterior probabilities are to be calculated, and to do so. It has been observed that

$$\frac{P(B,\, a_i = j \mid O)}{P(B \mid O)} = \frac{P(I \mid J)\sum_{a:a_i=j}\prod_{i'}^{I} P\left(a_i|J\right).P(B_{i'} \mid O_{a_i'})}{P\left(I|J\right)\sum_{a}\prod_{i=1}^{I} P(a_i|J)\,P(B_i \mid O_{a_i})} = \frac{\sum_{a:a_i=j}\prod_{i'=0}^{I} P(B_{i'} \mid O_{a_{i'}})}{\sum_{a}\prod_{i=1}^{I} P(B_i \mid O_{a_i})} \tag{25.8}$$

Now $P(I|J)$ is canceled because $P(ai|J)$ is uniform, the expression $\prod_{i=0}^{I} P(a_i \mid J)$ is constant across all alignments of the sentence. Now it is canceled. This expression is written purely in terms of lexical translation probability. Both the numerator and denominator in eq 25.8 are summed over exponentially many terms. Again, writing the eq 25.8 by computing the summation over a.

$$\frac{\sum_{a:a_i=j}\prod_{i'=0}^{I} P(B_{i'} \mid O_{a_{i'}})}{\sum_{a}\prod_{i=1}^{I} P(B_i \mid O_{a_i})} = \frac{P\left(B_i|O_j\right)\sum_{a_1=0}^{J}\cdots\sum_{a_{i-1}=0}^{J}\cdots\sum_{a_I=0}^{J}\prod_{i'=0}^{I} P(B_{i'} \mid O_{a_{i'}})}{\sum_{a_1=0}^{J}\cdots\sum_{a_I=0}^{J}\prod_{i=1}^{I} P(B_i \mid O_{a_i})} \tag{25.9}$$

The eq 25.9 state that the denominator is a sum over product of exactly I terms, to simplify the expression. The summation of a_i values first requires $P(B_1|O_1)$ is multiplied by all possible permutations of probabilities of translations of word $B2$ through BI. The upper part is the sum of all permutations multiplied by $P(Bi|Oj)$.

$$\frac{P\left(B_i|O_j\right)\prod_{i'=0}^{I}\sum_{a_1=0}^{J} \cdots \sum_{a_{i-1}=0}^{J}\cdots\sum_{a_{i+1}=0}^{J}\cdots\sum_{a_I=0}^{J} P(B_{i'} \mid O_{a_{i'}})}{\prod_{i=1}^{I}\sum_{a_1=0}^{J}\cdots\sum_{a_I=0}^{J} P(B_i \mid O_{a_i})} = \frac{P(B_i \mid O_j)}{\sum_{a_i=1}^{J} P(B_i \mid O_{a_i})} \tag{25.10}$$

Here in eq 25.10, calculating the posterior probability $P(a_i = j|B,O)$ in a linear way of all parameter $P\ (B|O)$. To calculate the expected counts of a link is the counts are collected across the entire dataset and then divide, just like MLE that means expected counts of a link divided by MLE. For EM, computing the expected counts called expectation step (E-step) are normalized by the M step called maximization. This problem is not fatal however assigning some initial value to θ which is called θ_0, then compute the expected counts each time θ_0, θ_1, θ_2 and so on. The result is quite improving the estimation of θ is the second key result of the EM algorithm. So, for this, the code of algorithm 1 is extended to EM.

25.8 ALGORITHM 2

1. $k = 0$
2. Initialize θ_0
3. repeat
4. $k = k + 1$
5. Initialize all counts to 0
6. for each n in [1, 2, 3, …, N] do
   a. for each i in [1, 2, 3, …, $I(n)$] do
      i. $z = 0$
   b. for each j in [1, 2, 3, …, $J(n)$] do
      i. $z += P_{\theta_{k-1}}(B_i^{(n)} \mid O_j^{(n)})$
      iii. for each j in [1, 2, 3, …, $J(n)$] do
         1. $c = P_{\theta_{k-1}}(B_i^{(n)} \mid O_j^{(n)} / z$
         2. $\text{count}\left[< B_i^{(n)}, O_j^{(n)} >\right] += c$
         3. count $[e_j] += c$
7. for each <B, O> in count do
   a. $P_{\theta_k}\left(B|O\right) = \text{count}(B,O) / \text{count}(O)$

There are two answers of θ_0. Here two things are very important in EM, first is the likelihood of θ_i will always be greater than of θ_{i-1}. The second one is, the model belongs to likelihood function is convex. These properties say, the limit of the EM algorithm must converge to the exact MLE. Finding the value of θ_0 repeatedly or in a recursive way. The experimental result is something different. Most people run EM only 3–5 times and this shows that in this case,[11] using heuristic initialization of θ_0 can significantly improve its accuracy. When EM is run, a good value of θ is searched and that θ is used to predict the "a"; this "a" is assigned to the highest probability value and indicated as $\hat{a}$

$$\hat{a} = \underset{a}{\operatorname{argmax}} P(a|O,B) \tag{25.11}$$

This is very easy, the alignment factor over the Bangla words.

$$\hat{a}_i = \underset{a_i}{\operatorname{argmax}} P(I|J)\prod_{i=1}^{I} P(a_i|J)\cdot P\left(B_i|O_{a_i}\right) \tag{25.12}$$

Since the only quantity that depends on a_i is $P\left(B_i|O_{a_i}\right)$, this reduces to $\hat{a}_i = \underset{a_i}{\operatorname{argmax}} P(B_i|O_{a_i})\cdot$

This probability value can be refined more by developing some new algorithms. The Expected count with mapping alignment between Bangla–Odia is here.

$$C\left(B_i \leftrightarrow O_j; B^s \leftrightarrow O^s\right) = \frac{P(O_j \mid B_i)}{\sum_x P(x \mid B_i)} * \left(\# B_i \in B^s\right) * \left(\# O_j \in O^s\right) \quad E-\text{step} \tag{15.13}$$

where C = expected count of $B_i \leftrightarrow O_j$ mapping in the context of the parallel corpus $B^s \leftrightarrow O^s$. $\#B_i \in B^s$ = number of times B_i occurs in B^s. $\#O_j \in O^s$ = number of times O_j occurs in O^s. "s" is represented as a parallel sentence pair.

25.9 RESULT AND DISCUSSION

In a bilingual dictionary based on the agriculture domain, a modest number of sentences (about 5000), about 50,000 words, and observable alignments are documented in a well-formatted, scientific style for easy access. The training process involves all observed alignments, and the result is a reliable estimate, as shown in eq 25.5. The corpus size needs to be expanded to obtain accurate estimates. One-to-one, many-to-one, and many-to-one word correspondences are all present. First, linkages (as a one-to-one mapping) have an equal chance of occurring. After one iteration, the model discovers that connections are created between the words from two parallel phrases that are the most similar by calculating the likelihood. Another iteration reveals that the probability value of the current word is more likely to be connected to earlier similar words. By calculating projected probability values recursively, the EM algorithm is an iterative process that maximizes the model or dictionary. It is feasible to decide on a possible alignment based on the likelihood values. Therefore, there are times when the Unigram approach is the most suitable one for determining the probability value of what to maximize and what our expectation value is. The two techniques used to evaluate a sentence's likelihood and word alignment are bigram and trigram. To calculate all probability values, a bigram in the form of a matrix is used. The expected count, updated expected count, and revised alignment probability values are calculated for each word in each parallel phrase. The updated alignment probabilities provide a better degree of approximation value between the words in the parallel text. As opposed to the average entropy with heuristic alignment, which is 1.53, the average entropy with EM alignment is already 1.4. The probability distribution is improving.

$$1.4 < 1.53$$

There are further mathematical procedures that can be used to improve this percentage value.

25.10 CONCLUSION AND FUTURE WORK

If a parallel corpus is closely matched at the sentence level, the computer can translate words from one language to another quickly. Bilingual dictionaries and phrase-based translation are frequently employed to address challenges such as one-to-many and many-to-one alignment. The creation of correspondents (Bangla–Odia) between two languages on a one-to-one, one-to-many, and many-to-one basis is made possible by bilingual dictionaries. With stated level translation, word divergence issues may occasionally be more successfully resolved. The EM method is used to determine which word combination (Bangla–Odia) has the highest probability value. Additionally, it is helpful to translate sentences and words one at a time while meticulously locating the crucial keyword in the target language. One of the most important factors to consider while translating a sizable amount of text is temporal complexity. As a result, caution should be exercised to obtain a better outcome because optimizing this is a challenging process. Likely, space complexity would not be reduced because of the size of our corpus data; rather, space should be increased because memory utilization is a concern.

KEYWORDS

- **probability**
- **alignment**
- **corpus**
- **Odia**
- **Bangla**

REFERENCES

1. Aswani, N.; Gaizauskas, R. Aligning Words in English-Hindi Parallel Corpora. *Assoc. Comput. Linguist.* **2005,** *19,* 115–118.
2. Das, B. R.; Maringanti, H. B.; Dash, N. S. In *Word Alignment in Bilingual Text for Bangla to Odia Machine Translation*, Presented in the International Conference on Languaging and Translating: Within and Beyond on 21-23, Feb 2020, IIT Patna: India.
3. Das, B. R.; Maringanti, H. B.; Dash, N. S. In *Challenges Faced in Machine Learning-Based Bangla–Odia Word Alignment for Machine Translation*, Presented in the 42nd International Conference of Linguistic Society of India (ICOLSI-42) on 10-12 Dec 2020, GLA University: Mathura, UP, India.
4. Das, B. R.; Maringanti, H. B.; Dash, N. S. *Bangla–Odia Word Alignment using EM Algorithm for Machine Translation*. Published in the Journal of Maharaja Sriram Chandra Bhanjadeo (erstwhile North Orissa) University: Baripada, India.
5. Dubey, S.; Diwan, T. D. Supporting Large English-Hindi Parallel Corpus using Word Alignment. *Int. J. Comput. Appl.* **2012,** *49,* 16–19.
6. Jindal, K. In *Automatic Word Aligning Algorithm for Hindi-Punjabi Parallel Text*, International Conference on Information Systems for Indian languages, pp 180–184; 2011.
7. Koehn, P.; Knight, K. Empirical Methods for Compounding Splitting. *EACL '03 Assoc. Comput. Linguist.* **2003,** *1,* 187–193.
8. Mansouri, A. B. Joint Prediction of Word Alignment with Alignment Types. *Trans. Assoc. Comput. Linguist.* **2017,** *5,* 501–514.
9. Peter, F. B.; Pietra, S. A. D.; Pietra, V. J. D.; Mercer, R. L. The Mathematics of Statistical Machine Translation: Parameter Estimation. *Computat. Linguist.* **1993,** *19* (2), 263–311.
10. Koehn, P. Statistical Machine Translation 2010.
11. Moore, R. C. In *Improving IBM Word-alignment Model 1*, Proceedings of the Association for Computational Linguistics, 2004.
12. Songyot, T.; Songyot; Chiang, D. Improving Word Alignment using Word Similarity. In *Empirical Methods in Natural Language Processing*; 2014**;** pp 1840–1845.
13. Tidemann, J. In *Word Alignment Step by Step*, Proceedings of the 12th Nordic Conference on Computational Linguistics, pp 216–227, 1999, University of Trondheim: Norway.

14. Tidemann, J. In *Combining Clues for Word Alignment*, Proceedings of the 10th Conference of the European Chapter of the Association for Computational Linguistics (EACL), pp 339–346, Budapest, Hungary, April 2003.
15. Tidemann, J. In *Word to Word Alignment Strategies*, International Conference on Computational Linguistics, 2004.
16. Bhattacharyya, P. *Machine Translation*, CRC Press, 2017 Print.
17. Jurafsky, D.; Martin, J. H. *Speech and Language Processing*, 4th ed.; Pearson, 2011.
18. https://en.wikipedia.org/wiki/Expectation%E2%80%93maximization_algorithm.
19. https://www.cs.sfu.ca/~anoop/students/anahita_mansouri/anahita-depth-report.pdf.
20. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.421.5497&rep=rep1&type=pdf

INDEX

D

E

F

G

H

I

N

O

P

R