Ahmed M. Anter
Mohamed Elhoseny
Anuradha D. Thakare *Editors*

Nature-Inspired Methods for Smart Healthcare Systems and Medical Data

Springer

Nature-Inspired Methods for Smart Healthcare Systems and Medical Data

Ahmed M. Anter • Mohamed Elhoseny
Anuradha D. Thakare
Editors

Nature-Inspired Methods for Smart Healthcare Systems and Medical Data

Springer

Editors
Ahmed M. Anter
Egypt-Japan University of Science and Technology (E-JUST)
Alexandria, Egypt

Mohamed Elhoseny
University of Sharjah
Sharjah, United Arab Emirates

Anuradha D. Thakare
Computer Engineering
Pimpri Chinchwad College of Engineering
Pune, India

ISBN 978-3-031-45951-1 ISBN 978-3-031-45952-8 (eBook)
https://doi.org/10.1007/978-3-031-45952-8

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Paper in this product is recyclable.

Preface

This book begins with the basics of smart healthcare systems and medical data and introduces the methodologies, processes, results, and challenges associated with the same. In addition, nature-inspired methods are introduced to serve the healthcare systems and big medical data. The Internet of Things (IoT) industry systems-based nature-inspired and mining algorithms are considered an important field of AI that is used to construct a healthcare application for lowering costs, increasing efficiency, accurate analysis of data, and better care for patients. In addition, nature-inspired methods are helpful tool in the development of a smart and intelligent healthcare system because of their flexibility and robustness characteristics. The fundamental problem is that although mobile devices or sensors can collect more detailed data, analyzing the situation from the data requires a strong forecasting and recognizing tool. Since metaheuristics can be used to improve the performance of algorithms for healthcare or to solve challenges in data mining or machine learning. They will therefore be an essential component of contemporary research on the healthcare system. Since complicated IoT healthcare systems are implemented in real-time and can be modeled with an IoT-based approach to produce good solutions, nature-inspired algorithms serve as a source of inspiration for these systems. Metaheuristics are strong technology for tackling several optimization problems in various fields, especially healthcare systems. The primary advantage of metaheuristic algorithms is their ability to find a better solution to a healthcare problem and their ability to consume as little time as possible. In addition, metaheuristics are more flexible compared to several other optimization techniques. These algorithms are not related to a specific optimization problem but could be applied to any optimization problem by making small adaptations to become suitable to tackle it. This book will focus on involvement of IoT-driven intelligent computing methods, state of the arts, novel findings and recent advances in smart healthcare and medical big data due to new technologies, and faster communication between users and devices. Also, the successful outcome of this book will enable a decision-maker or practitioner to pick a suitable optimization approach when making decisions to schedule patients under crowding environments with minimized human errors. Moreover, this book will target the healthcare domain, undergraduate students, graduate students, researchers, AI engineers, who pursue a strong understanding of IoT and healthcare applications. The book is intended for professionals, lecturers, short courses, researchers, trainers, industry professionals and researchers, e-Health management, IoT community, artificial intelligence community, and biomedicine community.

Alexandria, Egypt — Ahmed M. Anter
Sharjah, United Arab Emirates — Mohamed Elhoseny
Pune, India — Anuradha D. Thakare

Objective of the Book

The main aim of this book is to provide a detailed understanding of IoT-based smart healthcare systems and medical big data with involvement of distinct intelligent nature-inspired optimization methods in the field of computer science. This book brings together high-quality research on developing theories, frameworks, architectures, and algorithms for solving the complex smart healthcare challenges for real industry applications. Also, it investigates the most recent theoretical and practical applications of metaheuristics and optimization in a variety of smart healthcare applications. Furthermore, this book discusses the capability of optimization techniques to obtain the optimal parameters in machine learning and deep learning technologies. It endeavors to endow with significant frameworks, theory, design methods, and the latest empirical research findings in the area of smart healthcare systems and medical big data to foster healthcare sector that can be put to good use.

Organization of the Book

This book is organized into 12 chapters with the following brief description:

1. A Review of Methods Employed for Forensic Human Identification

This chapter explores the application of biometric identification techniques in forensic human identification. It provides an overview of the historical background and development of EEG-based identification, highlighting its significance in forensic investigations. Additionally, the study investigates the integration of different machine learning and deep learning approaches in forensic human identification. Moreover, this chapter sheds light on the importance of biometric identification, provide valuable insights for forensic experts, researchers, and practitioners in the field.

2. AI-Based Medicine Intake Tracker

This study developed an automated reminder system based on android application. The interaction between patients and doctors is emphasized. When it's time to take their medication, patients can set a reminder. The reminder can be customized with a variety of schedules for drugs, such as the date, time, and pharmaceutical summary. According to the patients' preferences, messages within the system will be used to notify them. Patient can choose to look for a doctor who can help them. The system notifies the patient each and every hour of the day as to when and how much to take. Also, the cloud storage is used to store all required test results and medicines for later use. Also patients are informed of the drug's expiration date, and a record of the drug's past usage can be kept for future use. A decent user interface and simple navigation are given top priority in this system. With the aid of the potent CNN-RNN-CTC algorithm, image processing will be precise and effective.

3. Analysis of Genetic Mutations Using Nature-Inspired Optimization Methods and Classification Approach

The purpose of this study is to use an optimization approach for classification on a dataset to increase the precision of classification models. An analysis of Genetic Mutations in cancer patients using nature-inspired optimization methods like Particle Swarm Optimization (PSO), Bee colon Optimization, and Genetic Algorithms is discussed. The mutations are classified using machine learning approach. It is observed that an accuracy of Random Forest classifier with PSO supersites other methods which is 71%. This study is confined to selection of standard dataset whereas the mutations may vary with respect to type of cancer and health history of patient, which in turn will impact the accuracy. Authors are working on experimenting the real-time datasets as a future research.

4. Applications of Blockchain: A Healthcare Use Case

This chapter presents the blockchain technology in the field of the medical healthcare. Blockchain is a new, revolutionary, and decentralized technology. This chapter emphasizes how blockchain for healthcare data management systems may foster innovation and lead to substantial breakthroughs. The ability to modify the current intelligent healthcare systems from highly centralized to distributed, secure, decentralized systems that can aid in improving healthcare and other applicable services is provided by blockchain technology when blended with other modern technologies. It helps ensure patient privacy while providing transparent data to all stakeholders. Malicious attackers can no longer alter vital medical records and are protected from data theft and surveillance. The link blockchain makes in the healthcare industry efficiently coordinates various industries under healthcare and provides transparency. In addition, the characteristics, features, and advantages of the Blockchain's are discussed for the healthcare industry.

5. Comprehensive Methodology of Contact Tracing Techniques to Reduce Pandemic Infectious Diseases Spread

In recent years, infectious diseases such as COVID-19 have posed a serious threat to individuals all over the world. This work investigates a number of applications that contributed to the development of contact tracing method. Additionally, there is a great deal of detail about the challenges that contact tracing software must overcome, such as privacy concerns, the inability to identify contacts, and delays in full identification. In this chapter, using the AI methods, scientists can build models that can explain how epidemics propagate and the effect contact tracing has on containing their spread. Finally, the available research, enumerated applications, domains of use, and open directions are reviewed.

6. High-Impact Applications of IoT System-Based Metaheuristics

The Internet of things (IoT) has become increasingly popular in recent years, with applications in various fields, such as healthcare, finance, agriculture, and smart cities. However, IoT systems face challenges, such as dynamic features, device mobility, and wireless communications. Metaheuristic algorithms, such as swarm intelligence (SI), have shown promise in addressing these challenges and improving IoT-based systems' performance. Remote monitoring of COVID-19 patients using wearable IoT devices is an example of an IoT-based system that employs nature-inspired algorithms to analyze health-related data.

7. IoT-Based eHealth Solutions for Aging with Special Emphasis on Aging-Related Inflammatory Diseases: Prospects and Challenges

This chapter discusses the current status of elderly care using the Internet of Things (IoT) which are currently available, and future possibilities of research and development with possible concerns related to security and patient confidentiality. IoT-based smart healthcare devices bring innovative changes for various age-related conditions of elderly persons. During the worldwide pandemic of COVID-19, IoT significantly plays a vital role in assisting the elderly person to monitor their behavior and movement using smart wearable with multiple sensors installed across their homes and also for continuous monitoring of their health conditions to provide emergency services.

8. Leveraging Meta-Heuristics in Improving Healthcare Delivery: A Comprehensive Overview

Leveraging meta-heuristics in healthcare delivery has the potential to revolutionize the way healthcare is delivered. By optimizing resource allocation, improving decision-making, processing times, and patient outcomes, meta-heuristics can improve the overall quality of healthcare delivery. However, there are also limitations to consider, such as the lack of transparency, bias, limited data, and high complexity. This chapter provides a comprehensive overview of the role of meta-heuristics in the healthcare system, starting with an introduction to the fundamental concepts and various applications of meta-heuristics in healthcare. The chapter delves into various meta-heuristics techniques and their application in solving healthcare problems such as resource allocation, treatment planning, and diagnosis. The strengths and limitations of these techniques are also discussed in detail, along with case studies that demonstrate their successful implementation in healthcare. Moreover, the chapter identifies potential areas for future application of meta-heuristics in healthcare, including personalized medicine, clinical decision support systems, and healthcare supply chain management.

9. Metaheuristics Algorithms for Complex Disease Prediction

Machine learning and metaheuristics algorithms are used to improve the performance of sensor technologies, devices, WSN, and IoT by picking relevant characteristics from raw data to improve device performance or lengthen sensor network lifetime. This chapter provides statistical techniques to select relevant features and provides particle swarm optimization (PSO) to divide the population into those with and those without breast cancer. PSO was used to identify the optimal pathophysiological parameter weights for a diagnosis system and then implemented it on a field programmable gate array (FPGA). Recent research has used metaheuristics to overcome healthcare optimization problems, such as the segmentation of magnetic resonance imaging, computed tomography pictures, and images from other sources. Metaheuristics for the study of large amounts of healthcare data is one of the future trends in research, along with the improvement of data storage systems, pre-processing of data extracted, and data analysis.

10. Printed rGO-Based Temperature Sensor for Wireless Body Area Network Applications

A wireless body area network (WBAN) is a type of wireless network that connects small, low-power, and lightweight devices worn on the body to monitor various physiological parameters, such as heart rate, blood pressure, temperature, and other health-related information. This chapter discusses temperature sensors with a special consideration for temperature sensors with rGO as the sensing material. The fabrication procedure for the sensor is also mentioned. The resulting sensor acts as a negative temperature coefficient resistor. The voltage drop across the resistor is mapped to the temperature in Celsius. The resultant values are optimized using the generic meta-heuristic algorithm to obtain more reliable results. The chapter also discusses the various sensors available in the literature and the materials that are available in the market which will help to build a temperature sensor. Different types of temperature sensors are also discussed.

11. Recent Advanced in Healthcare Data Privacy Techniques

This chapter reviews recent advances in healthcare data privacy techniques, the privacy breaches from the unique identification of an individual, or single/multiple attributes disclosure to the behavioral advertising privacy breaches. The strengths and limitations are discussed of each approach and provide examples of how they have been applied in healthcare contexts. Also, the ethical and legal considerations surrounding healthcare data privacy and the challenges of implementing these techniques in practice are examined.

12. The Ability of the CFD Approach to Investigate the Fluid and Wall Hemodynamics of Cerebral Stenosis and Aneurysm

This chapter conducts a computational investigation on cerebral arteries of patients and examines the hemodynamics parameters affected by the presence of a vascular stenosis or aneurysm. A systematic image-based computational fluid dynamics (CFD) method was introduced to simulate the blood flow in both benchmark and MRI-based models, to understand the hemodynamics of the complex vascular system. The use of CFD simulations involves both idealized and MRI-based models of the cerebral arteries. A patient-specific cerebral artery geometry was reconstructed to create a realistic model. To study the hemodynamics parameters, CFD simulations were employed to analyze the recirculation of the flow (flow vortex) and the wall pressure/shear stress. The boundary conditions for the simulation were obtained from ultrasonography measurements taken from the patients. The results of the stenosis analysis indicate that a specific area of vascular contraction critically influences blood flow, leading to a rise in wall shear stress in the stenosis region. This causes a rise in the flow velocity and the formation of vortices after the narrowing zone, potentially leading to blockage of the blood vessel and stroke. Additionally, the computational fluid dynamics results reveal the presence of the flow recirculation in the aneurysm zone. This significantly affects the flow and wall characteristics of the dilation.

Acknowledgments

First and foremost, all thanks and praises to the Almighty ***Allah*** who gives us the power to complete this work that cannot be accomplished without ***His*** endless support and blessing.

We would like to express our utmost levels of thanks to our parents for all these years of support and encouragement during my study and all my life, without them we are nothing.

Lastly, we would like to thank all our colleagues for their good attitude, care, and support, and the kind atmosphere they provided to us.

Contents

A Review of Methods Employed for Forensic Human Identification

Youssef Mohamed, Noran Mohamed, and Ahmed M. Anter

1 Introduction

The term "Forensic Human Identification"describes the scientific and investigative procedures used to identify the identities of people who have died or who have been involved in criminal activity. Applying scientific methods and principles to establish a person's identity based on the facts at hand is known as forensic human identification. Comparing distinctive traits like fingerprints, DNA profiles, dental records, facial features, skeletal remnants, and other identifiable markers can be one way to do this. The objective is to give precise and trustworthy identification to support forensic situations such as missing person cases, disaster victim identification, legal investigations, and missing person cases. There are some common methods and disciplines that play a significant role in forensic identification such as Fingerprint analysis which is the process of identifying someone by examining and contrasting the distinctive patterns, ridge features, and minute details contained in their fingerprints, DNA analysis that involves comparing samples of DNA taken from crime scenes, victims, or suspects to DNA profiles that are already known. DNA analysis can establish direct connections between people or identify family ties, and Forensic odontology involves the examination of dental remains and dental records to establish identification. To identify people or aid in the identification process, dentists compare dental data, including X-rays and dental charts, with dental remains. Other techniques for forensic human identification are Forensic Anthropology which

Y. Mohamed · N. Mohamed
Egypt-Japan University of Science and Technology (E-JUST), Alexandria, Egypt

A. M. Anter (✉)
Egypt-Japan University of Science and Technology (E-JUST), Alexandria, Egypt

Faculty of Computers and Artificial Intelligence, Beni-Suef University, Beni Suef, Egypt
e-mail: Ahmed_Anter@fcis.bsu.edu.eg

A. M. Anter et al. (eds.), *Nature-Inspired Methods for Smart Healthcare Systems and Medical Data*, https://doi.org/10.1007/978-3-031-45952-8_1

means examining the remains of skeletons to determine characteristics such as age, sex, ancestry, stature, and other unique features that can assist in establishing identity, Forensic Pathology which means examining deceased individuals to determine the cause and manner of death, and Forensic Anthropometry which involves the estimation of a person's physical traits and aid in identification by the measurement and comparison of various body parameters like height, limb lengths, or skull dimensions [1].

Establishing identity is significant for forensic investigations as the accurate identification of individuals contributes to public safety and security. It helps in the prevention and detection of crimes, tracking individuals with criminal histories, and monitoring potential threats to society. Moreover, establishing identity is also important in medical settings, where accurate identification allows for appropriate medical treatment, access to medical history, and proper management of healthcare resources. In forensic science, identity verification ensures the integrity of forensic evidence, chain of custody, and expert witness testimony.

2 Biometric Identification

Biometrics are unique physical characteristics that can be used for automated recognition. The term "biometric" comes from the Greek words" bios" which means life and "metrics" which means measure [2]. Today, a wide range of applications require reliable verification schemes to confirm the identity of an individual. Recognizing humans based on their body characteristics became more interesting in a lot of technology applications. Biometrics cannot be borrowed, stolen, or forgotten, and forging one is practically impossible. So it became a very important tool in the security systems. Humans have been identified based on their voice, appearance, or gait for thousands of years; these distinctive characteristics, or biometric traits, include features such as face, iris, palm print, and voice. Biometrics is now a mature technology that is widely used in a variety of applications ranging from border crossings to identity management.

There exist several biometric identification techniques that can be applied to humans to verify and authenticate people, as follows:

2.1 Fingerprint Recognition

A fingerprint is a pattern of furrows and ridges which are located on the tip of each finger. Fingerprints were used for personal identification for many centuries and the matching accuracy was very high [3]. Fingerprint-based recognition has been the most popular and successful method. Various historical accounts suggest that fingerprints were used in business transactions as early as 500 B.C. in Babylon [4].

High-resolution fingerprint images can capture sweat pores and other details in addition to minutiae points; these extended features are gaining attention as forensics experts appear to use them, particularly for latent and poor-quality fingerprint images. Nearly all forensics and law enforcement agencies worldwide use Automatic Fingerprint Identification Systems. It is the collection of these minutiae points in a fingerprint that is primarily used for matching two fingerprints. Fingerprints are composed of a regular texture pattern made up of ridges and valleys. These ridges are characterized by several landmark points, known as minutiae, which are mostly in the form of ridge endings and ridge bifurcations [3]. Figure 1 represents an example of a fingerprint of a human being.

2.2 Face Recognition

Face recognition refers to the process of identifying or verifying the identity of individuals based on their facial features. It is a biometric technology that analyzes and compares various facial characteristics to match an input face with a preexisting database of faces. Face recognition systems utilize computer algorithms to extract distinctive facial features, such as the arrangement of eyes, nose, mouth, and other facial landmarks, to create a unique representation of an individual's face called a face template [3]. Face recognition systems typically make use of the spatial relationship between the locations of facial features like eyes, noses, lips, chins, and the overall appearance of a face. Face is a natural human trait for automated biometric recognition. The forensic and civilian applications of face recognition technologies pose many technical challenges both for static mug-shot photos and moving faces [5]. Figure 2 represents an example of face recognition technology.

Fig. 1 Fingerprint identification

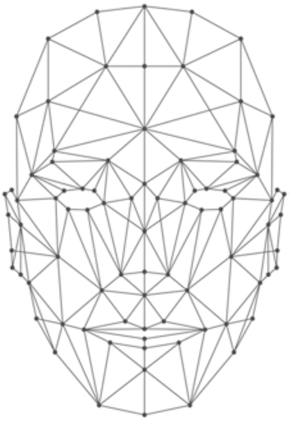


Fig. 2 Face recognition

2.3 Voice Recognition

Speech or voice-based recognition systems recognize a person based on their spoken words. The generation of a human voice involves a combination of physiological and behavioral features. The physiological component of voice generation depends on the shape and size of vocal tracts, lips, nasal cavities, and mouth. The behavioral component of voice includes movement of the lips, jaws, tongue, velum, and larynx. To build a model (typically the Hidden Markov Model) for speaker recognition, the spectral content of the voice is analyzed to extract its intensity, duration, quality, and pitch information. Speaker recognition is highly suitable for applications like telebanking, but it is quite sensitive to background noise and playback spoofing [5].

2.4 Iris Recognition

Iris identification is a biometric technology that allows people to be recognized based on the distinctive patterns in their irises, which are the colorful circular portions around the eyes' pupils. It is a very reliable and safe technique for biometric identification. The intricate and individual patterns on each person's iris include radial lines, freckles, furrows, and crypts. The iris is ideally suited for use as a biometric identification method since these patterns arise during embryonic development and are stable throughout the entire life of an individual [6]. Iris identification systems take a high-resolution image of the iris using specialized cameras or iris scanners. To extract the distinctive iris patterns, computer techniques are used to process the acquired image. These patterns are transformed into the iris code, a binary representation that captures the unique characteristics of the iris. An individual's iris image is taken throughout the enrolling process, and their iris code is generated. Then, to identify or verify people in the future, this iris code is saved in a database.

When a person offers their iris for identification during the recognition process, a new image of their iris is taken. The iris code is produced from the acquired image and contrasted with the database's entries for iris codes. Mathematical procedures like bitwise comparison and Hamming distance are used to determine how similar the iris codes are to one another. A match is discovered and the person's identification is verified if the resemblance reaches a predetermined level. The person's identity is not recognized if the similarity is below the threshold. High precision, dependability, and resistance to fraud or imitation are all characteristics of iris recognition. The iris provides a reliable biometric modality for a variety of applications, including access control, border security, time and attendance management, and identity verification. This is due to the rich and stable patterns found within the iris as well as the sophisticated algorithms employed in iris recognition systems.

2.5 Retina Recognition

Retinal recognition is a biometric system that uses a person's specific retinal blood vessel patterns to identify them. The retina, the eye's deepest layer, is home to a network of blood vessels that feeds the retinal tissue with nutrients and oxygen. To do retina recognition, a specialized camera that makes use of near-infrared light needs to take a picture of the retina. The retina is illuminated by near-infrared light, which enhances the clarity and visibility of the blood vessels. To extract and identify distinct blood vessel patterns, complex algorithms are used to process and evaluate the collected image [2]. The retina's blood vessel patterns are constant and particular to each person. Retina recognition is a trustworthy biometric method since they are created during early development and remain primarily constant throughout a person's lifespan.

An individual's retinal image is taken during enrolment, and the distinctive blood vessel patterns are retrieved and encoded into a template. After that, this template is saved in a safe database for further identification or confirmation. When a person offers their eye for identification throughout the recognition process, their retinal image is once more taken. The collected blood vessel patterns are compared to the database's template data. The similarity of the retinal patterns is assessed using a variety of mathematical methodologies, such as correlation coefficients or pattern-matching methods. A match is discovered and the person's identification is verified if the resemblance reaches a predetermined level. The person's identity is not recognized if the similarity is below the threshold. Images of the retina are taken with sophisticated cameras that record the complex web of blood vessels within the retina. Analyzing the branching structures and distribution of blood vessels at various scales within the retina is a component of fractal dimension analysis. The scaling behavior of the blood vessel patterns is used to calculate the fractal dimension. A more complex and self-similar pattern will have a higher fractal dimension value than one with a lower value. Researchers have looked into how fractal dimension can be used in retina recognition systems as a feature descriptor. An individual's retina's distinctive features can be recorded and compared for identification purposes by assessing the complexity of the blood vessel patterns using fractal dimensions [7].

2.6 Hand Geometry

A sort of biometric identification technology called hand geometry biometric uses measurements and analysis of the physical properties of a person's hand to identify them. It entails taking and examining numerous hand measurements and characteristics, including the dimensions, proportions, and contours of the hand and fingers. Typically, hand geometry systems take an image or 3D model of the hand using a specialized scanner or sensor. The scanner may use optical or thermal imaging methods to precisely capture the features of the hand. For identification or verification, the processed data is subsequently compared to pre-registered templates in a database [2].

Hand Geometry has applications in biometric identification. Hand Geometry can be used as a biometric modality, specifically analyzing the measurements and features of the hand for identification purposes. The methodology used in capturing hand geometry, which involves using specialized sensors or scanners to obtain images or 3D representations of the hand. The authors outline the measurements and features extracted from the hand, such as finger lengths, widths, and distances between key hand landmarks. These measurements are used to create unique templates that represent individuals' hand geometry [7]. One challenge of representing hand geometry measurements is the dimensionality of the data so an entropy-based discretization technique was used to convert continuous hand geometry measurements into discrete values. This discretization process aims to reduce the dimensionality of the data while preserving discriminatory information. The methodology employed involves three main steps: feature extraction, entropy-based discretization, and classification. In the feature extraction step, various hand geometry measurements, such as finger lengths, widths, and palm dimensions, are obtained from the hand images. These measurements serve as the initial continuous feature set [8].

According to the limitations of traditional hand geometry systems that require the hand to be aligned and positioned consistently during enrollment and recognition, a peg-free approach that can handle variations in hand position and rotation was proposed. Hierarchical geometry extraction and form matching are the two primary steps in the process. The hand image is processed to create a hierarchical representation of the hand geometry in the hierarchical geometry extraction step. To extract important elements like the fingers, palm, and hand contour, the hand image is split. The overall structure of the hand is then captured by representing these traits at various hierarchical levels. The hierarchical hand geometry representations are then compared using the shape-matching technique. The authors explain how to match hand shapes using shape context descriptors and the iterative closest point (ICP) technique. The ICP method iteratively aligns and matches the hierarchical representations for recognition while the shape context descriptors record the spatial relationships between the locations on the hand geometry [9].

2.7 Palm Print Recognition

The goal of palm print recognition, a biometric identification method, is to identify people using the distinctive patterns and traits found in their palm prints. It entails photographing and examining an individual's palm's surface characteristics, such as its lines, ridges, and texture patterns, to identify or verify them. Systems for reading palm prints make use of a variety of imaging tools, like specialist cameras or scanners, to record high-resolution photographs of the palm surface. The elaborate patterns made by the palm's distinctive configuration of ridges, lines, and texture are depicted in these pictures [10].

Palm print recognition is a significant biometric modality in applications such as access control, identity verification, and forensic investigations because of the uniqueness and stability of palm prints as a biometric characteristic. It has various techniques, including line-based features, texture analysis, minutiae extraction, and region-based features. Another technique for palm print recognition involves combining multiple algorithms or features to enhance recognition accuracy [11]. Another method is the acquisition of palm print images using specialized imaging devices. The acquired images are preprocessed to enhance image quality and normalize illumination and size variations. The key contribution lies in the utilization of Eigen palms for feature extraction. Eigen palms is a technique similar to Eigen faces but applied to palm print images. Eigen palms represent the global variations in palm print images by analyzing the principal components of a set of palm print images. The Eigen palms are obtained through a dimensionality reduction process, resulting in a lower-dimensional representation of the palm print patterns. The Eigen palms features are extracted from the preprocessed palm print images and utilized for recognition. The matching process involves comparing the Eigen palms features of the query palm print with the stored template palm prints in a database. Similar measures such as Euclidean distance or correlation coefficients are employed for matching and identification [12].

2.8 Vein Recognition

Vein recognition, also known as vascular pattern recognition, is a biometric technology that involves identifying individuals based on the unique patterns of veins within their bodies, typically in the palm, back of the hand, or finger. It utilizes the distinct vein patterns that form a complex network of blood vessels to establish a person's identity. Vein recognition systems capture images of the veins beneath the skin using near-infrared light or other specialized imaging techniques. Unlike other biometric modalities such as fingerprints or iris, vein patterns are not visible to the naked eye, making them difficult to forge or tamper with [13].

There are some advantages of using palm vein recognition such as its high accuracy and stability and resistance to forget. Palm vein recognition has different techniques, including near-infrared imaging, multispectral imaging, and thermal imaging [13]. Finger vein recognition is another biometric modality that utilizes the unique patterns of veins within the fingers for identification or verification purposes. It has several advantages such as, including its accuracy, non-intrusiveness, and resistance to spoofing or forgery. Near-infrared imaging, multispectral imaging, and other specialized techniques are used in finger vein recognition [14].

2.9 Ear Recognition

A biometric technique called ear recognition includes recognizing people by their ears' distinctive traits and characteristics. It establishes a person's identification by using the ear's shape, structure, and pattern. Specialized cameras or sensors are used by ear recognition systems to take pictures of the ear. The external ear structure, comprising the helix, lobule, antihelix, and tragus, is the subject of the photographs, highlighting the variances and shapes particular to everyone [15]. Ear recognition is a biometric modality that utilizes the unique features and characteristics of the ear for identification purposes because of its advantages such as its non-intrusive nature, stability, and potential for reliable identification. Some factors affect ear recognition performance, including pose variations, occlusions, image quality, and data variability [16]. There are various techniques for ear recognition such as Appearance-based Approaches that utilize texture or appearance features of the ear. Methods such as Local Binary Patterns (LBP), Scale Invariant Feature Transform (SIFT), or Histogram of Oriented Gradients (HOG) can be employed to capture texture information from the ear image. 2D Image-based Techniques: In this approach, 2D images of the ear are used for recognition. Various image processing techniques, such as edge detection, feature extraction, and template matching, are applied to extract distinctive features from the ear image for recognition. 3D Image-based Techniques utilize 3D images or models of the ear for recognition. 3D scanning techniques, such as laser scanning or structured light scanning, are used to capture the shape and depth information of the ear. This enables more accurate and robust recognition by considering the ear's three-dimensional structure [16].

2.10 DNA Matching

A forensic method called DNA matching, commonly referred to as DNA profiling or DNA fingerprinting is used to identify people based on their distinctive genetic information. It is predicated on the idea that DNA sequences differ amongst people, except for identical twins who have similar DNA. There are various techniques used in DNA matching such as The Analysis of short sequences of repeating nucleotides (STR loci) found in certain DNA regions is known as a short tandem repeat (STR) analysis. Each individual's DNA profile is different due to the variation in the number of repetitions at each locus. Due to its discriminatory solid power, STR analysis is frequently employed in forensic DNA matching, the PCR method which is used to amplify particular DNA sections, making it simpler to study. Even if the starting DNA sample is little or damaged, it enables the reproduction of specific DNA segments. In DNA matching, PCR is frequently employed to amplify STR loci for further analysis. Y-Chromosome Analysis focuses on the unique regions of the Y chromosome that are passed down from fathers to sons. It can be used for paternal lineage identification, determining male-related biological relationships, or in cases where only male DNA is present in the sample. DNA sequencing techniques involve determining the precise order of nucleotides in a DNA molecule. This technique provides a comprehensive analysis of the DNA sequence and can be used to identify specific variations or mutations in the DNA [17].

2.11 ECG Identification

The technique of identifying people based on their distinctive electrocardiogram (ECG) patterns is known as "ECG identification." The electrical activity of the heart is represented by ECG signals, which offer important insights into cardiac rhythm and function. A biometric method called ECG identification makes use of the particular features of an individual's ECG patterns to identify them. One approach for ECG-based personal authentication is the deep learning approach using LSTM-based deep recurrent neural networks. The aim is to leverage the temporal dependencies and patterns present in ECG signals to develop an accurate and reliable identification system. The authors utilized a deep recurrent neural network architecture with LSTM units to model the temporal dependencies in the ECG signals. The LSTM network is designed to capture and learn patterns from the sequential nature of ECG data. Multiple LSTM layers are stacked to enable deeper representation learning. The results of the experiments demonstrate the effectiveness of the LSTM-based deep recurrent neural network approach for ECG identification and personal authentication. The proposed method achieved high accuracy and outperforms traditional methods in ECG identification tasks. The findings highlight the advantages of utilizing deep learning techniques to capture the temporal dependencies and patterns present in ECG signals, leading to improved identification accuracy and robustness [18].

2.12 *EEG Recognition*

2.12.1 EEG

EEG stands for Electroencephalography which means the non-invasive measurement of the brain's electric fields. On the scalp, electrodes are used to capture voltage potentials caused by current passing through and around neurons. EEG has been around for about a century, and because of this, it has a wide range of uses. On the one hand, brain-triggered neuro-rehabilitation therapies have lately converged with the foundations of EEG in clinical diagnostics. On the other hand, EEG has not only been a mainstay in the field of experimental psychology for supplying brain correlates of constructs but it has also been employed as a real neuroimaging tool with more recent developments in translational as well as computational neuroscience. This "old dog" can still deliver new tricks and innovations thanks to its adaptability and accessibility, as well as developments in signal processing [19, 49, 50].

2.12.2 EEG History

The field of encephalography has advanced significantly during the past 100 years. Richard Caton, an English physician, made the discovery of electrical currents in the brain in 1875. Caton examined the EEG from the exposed monkey and rabbit brains. German doctor Hans Berger amplified the electrical activity of the brain observed on a human scalp in 1924 using his standard radio equipment. He declared that weak electric currents produced in the brain may be visually represented on a strip of paper and recorded without opening the skull. The brain's functioning state, such as sleep, anesthesia, a shortage of oxygen, and some neural illnesses, including epilepsy, affected the activity he saw. Many of the current uses of electroencephalography were established by Berger [20].

2.12.3 EEG Identification

EEG identification, often referred to as EEG-based identification or EEG biometrics, is the process of using electroencephalography (EEG) data to identify people based on the distinctive patterns of their brainwaves. It is a biometric identification technique that depends on the unique features of a person's EEG waves. With EEG identification, electrodes are positioned on the scalp to capture the electrical activity of the brain. The cerebral activity and brainwave patterns connected to particular mental states, cognitive activities, or physiological responses are captured by the recorded EEG signals [21].

The process of EEG identification involves five steps. First is the EEG data acquisition where electrodes are positioned on the scalp, and specialized equipment is used to record EEG signals. The electrodes track the brain's electrical activity and send data for additional processing. Second, in the preprocessing step to eliminate noise, artifacts, or undesired signals, preprocessing processes are applied to the collected EEG data. Filtering, artifact-removal methods, and signal normalization might be used in this. Then the feature extraction step where the preprocessed EEG data is used to extract pertinent features that indicate the distinctive qualities of each person's brainwave patterns. These characteristics may include frequency bands, statistical measures, spectral power, statistical measures, or other pertinent measures generated from the EEG signals. The features are then compressed into a template or representation that captures the individual's particular brainwave patterns. This model is used as a guide for further identification or verification in the feature encoding and template creation step. The collected EEG signals of a person are matched against a database of saved templates during the identification stage. The degree of similarity between the recorded EEG signals and the template data is measured using matching or similarity algorithms. The individual's identity is verified if a match is discovered that is greater than a predetermined threshold in the last step which is the matching and recognition step [22].

The benefits of EEG identification include its non-intrusiveness, difficulty in forging, and potential resistance to spoofing or imitation. Due to the variability of EEG signals caused by variables including mental states, electrode location, and inter-individual variances, it also presents difficulties. EEG identification is used in a variety of fields, such as brain-computer interfaces, biometric authentication, and access control systems. It can be utilized in instances when a safe and trustworthy identification mechanism is needed, especially when other biometric modalities might not be appropriate or practical.

2.12.4 EEG Identification Techniques

There are techniques that are commonly used in EEG identification. Time Domain Analysis is a technique used in EEG identification and this method entails time domain analysis of the unprocessed EEG signals. For identification, features from the EEG signals can be retrieved, including amplitude, peak detection, zero-crossing rate, and statistical metrics (mean, variance, etc.). An example of the Time Domain Analysis technique is the use of Convolutional neural networks (CNNs) for P300 detection in brain-computer interfaces. In EEG signals connected to cognitive processes, particularly those connected to attention and decision-making, the P300 is a notable event-related potential. In brain-computer interface (BCI) applications, detecting the P300 component is essential because it can serve as a foundation for communication or control. The authors suggest using CNN-based techniques to detect P300 in BCI systems. CNNs are useful for studying EEG signals because they excel at extracting pertinent characteristics from sequential data and capturing temporal dependencies. CNNs' primary benefit is their capacity to automatically identify hierarchical features from unprocessed input using subsequent convolutional and pooling layers. The CNN architecture created by the authors is intended exclusively for EEG analysis. Convolutional layers in the architecture act on the unprocessed EEG signals to identify regional patterns and temporal correlations. A sizable dataset of EEG recordings encompassing both non-P300 signals and P300 event-related potentials are used to train the CNN [23].

Another technique is using the Frequency Domain Analysis which involves transforming the EEG signals from the time domain to the frequency domain using methods like the Fast Fourier Transform (FFT) or Wavelet Transform. Power spectral density, frequency bands (e.g., alpha, beta, theta), or spectral entropy can be extracted as features for identification. The frequency domain analysis of EEG signals can be performed using techniques such as the Fourier transform and power spectral density estimation. The Fourier transform was applied to convert the EEG signals from the time domain to the frequency domain. The Fourier transform decomposes a signal into its constituent frequency components. By applying the Fourier transform to the EEG signals, information about the signal's frequency content was obtained. The resulting frequency spectrum provided details about the amplitudes and phases of different frequency components present in the EEG signals. A method for estimating the power distribution across various frequency components in a signal is called power spectral density (PSD) estimation. The PSD estimation was performed to quantify the power or energy contained within specific frequency bands in the EEG signals. This allowed analysis of the relative contributions of different frequency ranges (e.g., delta, theta, alpha, beta) to the overall power of the EEG signals. By applying the Fourier transform and power spectral density estimation, we gained insights into the frequency characteristics of the EEG signals [24].

3 Machine learning Approaches in Personal Identification

Machine learning approaches in the context of EEG identification involve training algorithms to learn patterns and relationships within EEG data to classify and identify individuals based on their unique biometric patterns [25].

3.1 Support Vector Machines (SVM)

SVM is a well-liked classification method that is utilized in many fields [51]. Finding the best hyperplane in a high-dimensional feature space that maximally divides distinct classes is the goal of SVM. To categorize people based on their EEG patterns in the context of EEG identification, SVM can be trained utilizing extracted features from EEG data.

3.2 Random Forests

An ensemble learning technique called random forests mixes many decision trees [52]. Each tree is trained using a subset of features and data, and the results from all the trees are combined to get the final classification. Random forests can handle high-dimensional data, and they are resistant to overfitting. They can be trained using extracted features to categorize people and then used for EEG identification.

3.3 K-Nearest Neighbors (K-NN)

A sample is given a class label by the k nearest neighbors in the feature space using the non-parametric k-NN algorithm [53]. When used for EEG identification, k-NN can be trained using features taken from the signals. When given a fresh EEG sample, it recognizes the k most comparable samples in the feature space to establish the identity of the subject. The K-nearest neighbors (KNN) algorithm can be improved in performance and efficacy using a variety of approaches and techniques: Distance Metrics: Depending on the features of the data, various distance metrics, such as the Manhattan distance, Minkowski distance, or Mahalanobis distance, may be used in place of the Euclidean distance. Selecting the right distance measure can help KNN perform more accurately by improving sample discrimination. Distance Weighting: By giving neighbors varying weights based on their distance from each other, the problem of unbalanced influence can be resolved. Greater weights can be assigned to nearby neighbors, highlighting their importance in the classification process. A distance-weighted KNN can improve the algorithm's capacity for discrimination. Feature Selection: aims to identify the most relevant features in the dataset for classification. By selecting a subset of informative features, the dimensionality of the data can be reduced, which can enhance the performance of KNN. Feature Scaling: Scaling the features to a common range can prevent features with larger scales from dominating the distance calculation. Common techniques for feature scaling include min-max scaling (normalization) and standardization (z-score normalization). Voting Schemes: KNN involves classifying a sample based on the majority vote of its nearest neighbors. Various voting schemes can be employed, such as simple majority voting or weighted voting, where closer neighbors have a higher influence on the classification decision. Table 1 represents the reference index, year, dataset, methods, finding, advantages, and disadvantages of 10 studies about KNN.

4 Deep Learning Approaches in Personal Identification

Convolutional neural networks (CNNs) and recurrent neural networks (RNNs), for example, are deep learning models that have attracted a lot of interest in EEG detection [54, 55]. These models are able to recognize complicated patterns and connections and automatically learn hierarchical representations from unprocessed EEG signals. Deep learning models have demonstrated promising results in identifying persons based on EEG patterns, but they frequently need enormous amounts of labeled data and intensive computational resources.

Table 1 The reference index, year, dataset, methods, finding, advantages, and disadvantages of 10 studies about KNN

Reference Index	Year	Dataset	Methods	Finding	Advantages	Disadvantages
[26]	2015	Financial Tweets Dataset	The mutual k nearest neighbor (MkNN or MNN) approach, Error-based Weighting	Efficiently predicted the yield of particular stocks based on those tweets using KNN.	The authors evaluated the KNN approach in the presence of label noise.	The high fraction of the overall error is due to various types of noise
[27]	2020	Used car which is collected from Kaggle	Euclidean Distance	Successfully predicted the price of used cars.	Using Cross-Validation to inspect the model overfitting	The accuracy was not too good.
[28]	2019	Stock Data which is obtained through the API.	KNN, Euclidean Distance.	making the LQ45 stock index prediction application to predict stock prices.	A simple application to use to help investors make decisions regarding their shares.	The accuracy generated by the application was not too good.
[29]	2020	Wisconsin Diagnostic Breast Cancer (WDBC) Dataset from the UCI repository	KNN classification	Successful cancer detection in the early stages	Effective for small to medium-sized datasets	Outliers can significantly impact results
[30]	2016	Breast Cancer data set	KNN classification	Efficiently detected the breast cancer	Efficient detection with high accuracy	Sensitive to irrelevant features and outliers
[31]	2013	Finance data set	KNN, Euclidean Distance.	Financial distress prediction	Simple and easy to implement.	Sensitive to missing values
[32]	2013	Stock data is extracted from the Jordanian Stock exchange	KNN classification	Predicted stock prices for a sample of six major companies	The results were rational and reasonable	Large datasets can impact performance

(continued)

Table 1 (continued)

Reference Index	Year	Dataset	Methods	Finding	Advantages	Disadvantages
[33]	2007	Irregularity data set	KNN classification	Successfully detected irregularities in images and in video	Simple to implement	Can not handle large data sets
[34]	2020	Covid-19 dataset from Kaggle	KNN classification	Prediction of Covid-19 possibilities using KNN	Effective for small and medium datasets	High storage requirements for large datasets
[35]	2017	Student marks which is a data set collected from the Ministry of Education	KNN classification	Efficient student performance prediction	Simple and easy to implement	The accuracy was not too high

4.1 Long Short-Term Memory (LSTM)

Long Short-Term Memory, also known as LSTM, is an RNN architecture style created to model sequential data and identify long-term dependencies. Since their introduction by Hochreiter and Schmidhuber in 1997, LSTMs have found widespread use in a variety of tasks, including speech recognition, time series analysis, and natural language processing. The main benefit of LSTMs over conventional RNNs is their capacity to ameliorate the vanishing gradient problem, which is characterized by the exponential decay of gradients propagated across the network over time, making it challenging to detect long-term dependencies. By including a memory cell, LSTMs deal with this problem by enabling the network to remember or forget information over time selectively.

The following essential elements make up the LSTM architecture: Cell State: The LSTM‘s cell state serves as its memory. It encompasses the whole sequence and permits information to move freely throughout the network without being considerably altered. Using gates that are governed by sigmoid activation functions, the cell state can choose to selectively remember or forget information. The input gate chooses how much of the fresh input should be absorbed into the state of the cell. It generates a value between 0 and 1, reflecting the significance of the new input, using a sigmoid activation function. How much of the prior cell state should be remembered or retained is decided by the forget gate. It determines whether data from the cell state should be eliminated using a sigmoid activation function. The output gate regulates how much of the cell state should be made available to the LSTM‘s output. The amount of information that is transferred to the subsequent layer or task is controlled using a sigmoid activation function. The input, the previous hidden state, and the current cell state are used to compute the hidden state, which is in charge of transferring information across time steps. It can be thought of as the LSTM's "memory," storing important information from the past and influencing predictions for the future [36].

An LSTM processes each element of the input sequence individually during the forward pass, updating its cell state and hidden state at each time step. The input gate, forget gate and output gate are used to update the cell state, and the updated cell state and output gate are used to determine the hidden state. Backpropagation through time (BPTT), which computes gradients and uses them to update the model parameters, can be used to train LSTMs. The gradients run through all the time steps, enabling the LSTM to take into account long-term dependencies and learn from previous observations to produce predictions. LSTMs are very good at modeling and forecasting data sequences because they incorporate memory cells and solve the vanishing gradient problem.

4.2 Gated Recurrent Unit (GRU)

Gated Recurrent Units (GRUs), a sort of recurrent neural network (RNN) architecture akin to LSTMs (Long Short-Term Memory), are what they are called. Cho et al. presented GRU as an alternative to LSTM in 2014 to streamline the architecture while preserving long-term dependencies in sequential data. To selectively update and maintain knowledge over time, GRU units include gating mechanisms that regulate the flow of information inside the network. Similar to LSTM, GRU is made to deal with the vanishing gradient issue and capture long-term dependencies, both of which are crucial in applications involving sequential data, such as time series analysis, speech recognition, and natural language processing. The network can efficiently update and forget data thanks to GRU units. GRU units can capture important long-term dependencies while avoiding pointless computations by controlling the information flow through the gating mechanisms. GRU has become more widely used in a variety of applications, particularly those that emphasize computing effectiveness or work with smaller datasets. It has proven successful at tasks like sentiment analysis, speech recognition, machine translation, and language modeling [37].

GRU has a less complex design than LSTM and fewer gates. The procedure is simplified and fewer parameters are needed since it creates a single vector by combining the memory cell and hidden state. Typically, GRU units include the following elements: The update gate chooses how much of the prior hidden state should be kept and how much of the fresh input should be absorbed into the present hidden state. Based on the current input, it determines the significance of the previous concealed state and the incoming input. Reset Gate: The reset gate regulates how much the previous hidden state affects how the current hidden state is calculated. It chooses whether or not the network should keep the prior concealed state. Candidate Activation: Using the current input and the prior hidden state, the candidate activation computes a new candidate hidden state. Using the update and reset gates, it is utilized to update the concealed state. GRU offers a less complicated LSTM substitute while retaining the capacity to model sequential data and record long-term dependencies. Both architectures have their advantages and disadvantages, and the decision between LSTM and GRU depends on the job and dataset [38].

There are some approaches and techniques that can be used to enhance the performance of Gated Recurrent Units such as Stacked GRUs, stacking multiple layers of GRUs to increase the capacity of the model and capture more complex dependencies in the data, that can be effective for tasks, Bidirectional GRUs which process the sequence of the input in both directions forward and backward to allow the model access past and future information like in speech recognition and sentiment analysis, Regularization techniques like dropout or recurrent dropout can also be applied to the GRU layers to prevent overfitting and improve generalization, Batch Normalization can be applied to the hidden states of the GRU to accelerate training and improve the stability of the model, Learning Rate Scheduling is another technique that can be used to improve the training efficiency and convergence, Attention Mechanisms can enhance the ability of the model to focus on relevant parts of the input sequence as these mechanisms allow the model to assign weights to different time steps dynamically, Pretrained Embeddings that can provide the GRU model with useful initial representations to help it capture semantic relationships and improve its performance, and Hyperparameter Tuning is another different approach that can optimize the performance of the GRU model. All these techniques can be applied according to the characteristics and requirements of the task. Table 2 represents the reference index, year, dataset, methods, finding, advantages, and disadvantages of 10 studies.

5 Evaluation Methods

The Confusion Matrix is used to evaluate the performance of the model. Table 3 represents the confusion matrix. Moreover, the following measures are extracted from the confusion matrix to represent the stability and robustness of the models used in forensic human identification such as (Accuracy, Recall, Precision, F1-Score, Sensitivity, and Error rate) [56–58].

Performance is later calculated using the following formulas:

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

$$Recall\left(Sensitivity\right) = \frac{TP}{TP + FN}$$

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

Table 2 The reference index, year, dataset, methods, finding, advantages, and disadvantages of 10 studies about GRU

Reference index	Year	Dataset	Methods	Finding	Advantages	Disadvantages
[39]	2014	4 datasets each contain at least 7 hours of polyphonic music	GRU, LSTM, Traditional tanh units	GRU & LSTM are better than traditional tanh units.	Captures sequential information in the data.	May struggle with capturing complex patterns.
[40]	2020	Flight-delays dataset is obtained from Kaggle	GRU, vanilla LSTM, Bi-directional LSTM	Successfully predicted delays encountered in the aviation industry.	Predicted the amount of delay with quietly small error.	Computationally expensive for large datasets.
[41]	2015	Hutter dataset from the human knowledge compression contest	RNN, LSTM, GRU, GF-RNN, Tanh Units	Demonstrated that GF-RNN was a fast and better performance	Helpful with the complicated sequences models	Requires longer training time compared to simpler models
[42]	2021	The datasets are obtained from open-access websites	RNN, Bi-LSTM, GRU, LSTM	GRU presented the most accurate prediction for LTC	Efficient in training and inference	Domain – Specific required
[43]	2020	The dataset is obtained from an online website	GRU, LSTM, Machine Learning, Deep Learning	Efficiently Airborne particle pollution prediction	Accurately forecasting with a reasonable error	High memory requirements for large-scale models.
[44]	2020	A massive amount of water level dataset is collected from the Yangtze River using IOT	CNN, CNN-GRU, LSTM, ARIMA, WANN	Successfully predicted the water level of inland rivers using CNN-GRU Model	The CNN-GRU model has higher accuracy than other classical models	The CNN-GRU has prediction errors
[45]	2016	The used dataset is collected from the PeMS dataset	LSTM, GRU, RNN, ARIMA	GRU & LSTM accurately predicted real-time traffic flow better than the ARIMA model.	Has better performance and accuracy.	High memory requirements for large-scale models.

(continued)

Table 2 (continued)

Reference index	Year	Dataset	Methods	Finding	Advantages	Disadvantages
[46]	2022	The used dataset represents the average power consumption in péronne city	GRU, LSTM, Drop-GRU	Efficiently predicted energy consumption	Overcomes the limitations of the traditional models	There are few prediction errors
[47]	2019	The used dataset is from the official website of the China Meteorological Administration	GRU, LSTM, RNN	Successfully predicted air pollutant concentration using the GRU method	The prediction accuracy is high	The GRU model in this paper is not competent for predicting larger datasets
[48]	2021	The used dataset is from the NetEase Finance and Economics website	GRU, LSTM, PCA-LSTM, PCA-GRU, LASSO-LSTM, LASSO-GRU	Efficiently predicted stock prices using the GRU model and LSTM model	Achieved high accuracy in stock price forecasting.	Higher computational requirements

Table 3 The confusion matrix

	Actually positive (1)	Actually negative (0)
Predicted Positive (1)	True Positive (TP)	False Negative (FN)
Predicted Negative (0)	False Negative (FN)	True Negative (TN)

$$Error\ Rate = \frac{FP + FN}{TP + FP + TN + FN}$$

$$F1 - Score = \frac{2 * TP}{2 * TP + FP + FN}$$

6 Conclusion and Future Work

In conclusion, this research paper highlights the significance of biometric identification techniques in forensic human identification. The study specifically focuses on numerous methods of recognition, especially EEG-based identification. These biometric modalities have proven to be effective in establishing the identity of individuals in various forensic contexts. The integration of machine learning approaches, such as KNN, SVM, random forests, and deep learning. It also covered different architecture styles of RNN such as LSTM and GRU. It has enhanced the accuracy and efficiency of identification systems. The paper provides insights into the historical development of EEG-based identification, emphasizing its relevance in forensic investigations. The use of EEG signals as a biometric identifier holds immense potential, and its application in forensic human identification continues to evolve.

While significant progress has been made in biometric identification for forensic purposes, there are several avenues for future research and development. Further research can focus on improving the accuracy and reliability of EEG-based identification systems, Investigating the fusion of multiple biometric modalities, addressing ethical considerations which become crucial as biometric identification becomes more prevalent, developing real-time identification systems that can process biometric data efficiently with rapid results, and creating standardized benchmark datasets and performance metrics specific to forensic human identification.

By addressing these areas of future work, researchers can contribute to the continuous improvement and advancement of biometric identification techniques for forensic applications, ultimately enhancing the accuracy, reliability, and practicality of identification systems in real-world scenarios.

References

1. Thompson, T., & Black, S. (Eds.). (2006). *Forensic human identification: An introduction*. CRC press. https://ds.amu.edu.et/xmlui/bitstream/handle/123456789/7281/Forensic_Human_Identification.pdf?sequence=1&isAllowed=y
2. Delac, K., & Grgic, M. (2004, June). A survey of biometric recognition methods. In *Proceedings. Elmar-2004. 46th international symposium on electronics in marine* (pp. 184–193). IEEE. https://ieeexplore.ieee.org/abstract/document/1356372/
3. Maio, D., Maltoni, D., Cappelli, R., Wayman, J. L., & Jain, A. K. (2002, August). FVC2002: Second fingerprint verification competition. In *2002 international conference on pattern recognition* (Vol. 3, pp. 811–814). IEEE. https://ieeexplore.ieee.org/abstract/document/1048144
4. Pugliese, J. (2010). *Biometrics: Bodies, technologies, biopolitics* (Vol. 12). Routledge.
5. Mordini, E., & Tzovaras, D. (2012). *Second generation biometrics: The ethical, legal and social context*. Springer Science & Business Media. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=6b6866fbb4354e30ab34db9d6a8a07da4bf25777
6. Daugman, J. (2009). How iris recognition works. In *The essential guide to image processing* (pp. 715–739). Academic Press. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=df0e47d526820bb99c1435541a208a66d4d8bc61
7. Sukumaran, S., & Punithavalli, M. (2009). Retina recognition is based on the fractal dimension. *IJCSNS International Journal of Computer Science and Network Security, 9*(10), 66–67. https://www.researchgate.net/profile/S-Sukumaran/publication/266863563_Retina_Recognition_Based_on_Fractal_Dimension/links/602365fe92851c4ed55ebd47/Retina-Recognition-Based-on-Fractal-Dimension.pdf
8. Kumar, A., & Zhang, D. (2007). Hand-geometry recognition using entropy-based discretization. *IEEE Transactions on Information Forensics and Security, 2*(2), 181–187. https://core.ac.uk/download/pdf/205608391.pdf

9. Wong, A. L., & Shi, P. (2002, December). Peg-free hand geometry recognition using hierarchical geometry and shape matching. In *MVA* (pp. 281–284). https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=25743bbc24b208830189853da5e2b20eedacfeed
10. Kong, A., Zhang, D., & Kamel, M. (2009). A survey of palmprint recognition. *Pattern Recognition, 42*(7), 1408–1418. https://www.academia.edu/download/46520160/j.patcog.2009.01.01820160615-13065-1i89pni.pdf
11. Zhang, D., Zuo, W., & Yue, F. (2012). A comparative study of palmprint recognition algorithms. *ACM Computing Surveys (CSUR), 44*(1), 1–37. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=748675f7ab301351b8d80b0c9816d5763ddd7bf8
12. Lu, G., Zhang, D., & Wang, K. (2003). Palmprint recognition using eigenpalms features. *Pattern Recognition Letters, 24*(9–10), 1463–1467. http://ai.pku.edu.cn/application/files/6815/1124/8094/Palmprint_recognition_using_eigenpalms_features.pdf
13. Wu, W., Elliott, S. J., Lin, S., Sun, S., & Tang, Y. (2020). Review of palm vein recognition. *IET Biometrics, 9*(1), 1–10. https://ietresearch.onlinelibrary.wiley.com/doi/pdfdirect/10.1049/iet-bmt.2019.0034
14. Shaheed, K., Liu, H., Yang, G., Qureshi, I., Gou, J., & Yin, Y. (2018). A systematic review of finger vein recognition techniques. *Information, 9*(9), 213. https://www.mdpi.com/2078-2489/9/9/213/pdf
15. Emeršič, Ž., Štruc, V., & Peer, P. (2017). Ear recognition: More than a survey. *Neurocomputing, 255*, 26–39. https://arxiv.org/pdf/1611.06203
16. Ganapathi, I. I., Ali, S. S., Prakash, S., Vu, N. S., & Werghi, N. (2023). A survey of 3d ear recognition techniques. *ACM Computing Surveys, 55*(10), 1–36. https://www.researchgate.net/profile/Iyyakutti-Ganapathi/publication/363232795_A_Survey_of_3D_Ear_Recognition_Techniques/links/633171616063772afd92c041/A-Survey-of-3D-Ear-Recognition-Techniques.pdf
17. Jain, A. K., & Kumar, A. (2010). Biometrics of next generation: An overview. *Second Generation Biometrics, 12*(1), 2–3. https://www.intechopen.com/chapters/16506
18. Kim, B. H., & Pyun, J. Y. (2020). ECG identification for personal authentication using LSTM-based deep recurrent neural networks. *Sensors, 20*(11), 3069. https://www.mdpi.com/1424-8220/20/11/3069/pdf
19. Biasiucci, A., Franceschiello, B., & Murray, M. M. (2019). Electroencephalography. *Current Biology, 29*(3), R80–R85. https://www.cell.com/current-biology/pdf/S0960-9822(18)31551-3.pdf
20. Teplan, M. (2002). Fundamentals of EEG measurement. *Measurement Science Review, 2*(2), 1–11. http://www.edumed.org.br/cursos/neurociencia/MethodsEEGMeasurement.pdf
21. Marcel, S., & Millán, J. D. R. (2007). Person authentication using brainwaves (EEG) and maximum a posteriori model adaptation. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 29*(4), 743–752. https://infoscience.epfl.ch/record/83215/files/marcel-idiap-rr-05-81.pdf
22. Poulos, M., Rangoussi, M., Chrissikopoulos, V., & Evangelou, A. (1999, September). Person identification based on parametric processing of the EEG. In *ICECS'99. Proceedings of ICECS'99. 6th IEEE international conference on electronics, circuits, and systems (Cat. No. 99EX357)* (Vol. 1, pp. 283–286). IEEE. https://www.academia.edu/download/48955447/Person_identification_based_on_parametri20160919-31290-1j0bord.pdf
23. Cecotti, H., & Graser, A. (2010). Convolutional neural networks for P300 detection with application to brain-computer interfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 33*(3), 433–445. https://pure.ulster.ac.uk/ws/files/11410302/cecotti_pami.pdf
24. Al-Fahoum, A. S., & Al-Fraihat, A. A. (2014). Methods of EEG signal features extraction using linear analysis in frequency and time-frequency domains. *International Scholarly Research Notices*, 2014. https://downloads.hindawi.com/archive/2014/730218.pdf
25. Zangeneh Soroush, M., Maghooli, K., Kamaledin Setarehdan, S., & Motie Nasrabadi, A. (2018). Emotion classification through nonlinear EEG analysis using machine learning

methods. *International Clinical Neuroscience Journal, 5*(4), 135–149. http://eprints.stmdigipress.com/id/eprint/108/1/22921-Article%20Text-100399-1-10-20190524.pdf
26. Buza, K., Nanopoulos, A., & Nagy, G. (2015). Nearest neighbor regression in the presence of bad hubs. *Knowledge-Based Systems, 86*, 250–260. http://real.mtak.hu/26128/1/hubness_aware_regression.pdf
27. Samruddhi, K., & Kumar, R. A. (2020). Used car price prediction using k-nearest neighbor-based model. *International Journal of Innovative Research in Applied Sciences and Engineering (IJIRASE), 4*, 629–632. https://ijirase.com/assets/paper/issue_1/volume_4/V4-Issue-2-629-632.pdf
28. Tanuwijaya, J., & Hansun, S. (2019). LQ45 stock index prediction using k-nearest neighbors' regression. *International Journal of Recent Technology and Engineering, 8*(3), 2388–2391. https://www.researchgate.net/profile/Julius-Tanuwijaya/publication/336715759_LQ45_Stock_Index_Prediction_using_k-Nearest_Neighbors_Regression/links/5dae7e5ca6fdccc99d929d4a/LQ45-Stock-Index-Prediction-using-k-Nearest-Neighbors-Regression.pdf
29. MurtiRawat, R., Panchal, S., Singh, V. K., & Panchal, Y. (2020, July). Breast Cancer detection using K-nearest neighbors, logistic regression, and ensemble learning. In *2020 international conference on electronics and sustainable communication systems (ICESC)* (pp. 534–540). IEEE. https://www.researchgate.net/profile/Ram-Rawat/publication/349928499_Breast_Cancer_Detection_Using_K-Nearest_Neighbors_Logistic_Regression_and_Ensemble_Learning/links/60f029650859317dbde38576/Breast-Cancer-Detection-Using-K-Nearest-Neighbors-Logistic-Regression-and-Ensemble-Learning.pdf
30. Alarabeyyat, A., & Alhanahnah, M. (2016, August). Breast cancer detection using k-nearest neighbor machine learning algorithm. In *2016 9th international conference on developments in eSystems engineering (DeSE)* (pp. 35–39). IEEE.
31. Imandoust, S. B., & Bolandraftar, M. (2013). Application of k-nearest neighbor (knn) approach for predicting economic events: Theoretical background. *International Journal of Engineering Research and Applications, 3*(5), 605–610. https://www.academia.edu/download/88091993/DI35605610.pdf
32. Anter, A. M., Mohamed, A. W., Zhang, M., & Zhang, Z. (2023). A robust intelligence regression model for monitoring Parkinson's disease based on speech signals. *Future Generation Computer Systems, 147*, 316–327. https://www.academia.edu/download/41209628/00b49537c5ed39c749000000.pdf20160115-19908-ad2dgh.pdf
33. Boiman, O., & Irani, M. (2007). Detecting irregularities in images and in video. *International Journal of Computer Vision, 74*, 17–31.
34. Anter, A. M., Elnashar, H. S., & Zhang, Z. (2022). QMVO-SCDL: A new regression model for fMRI pain decoding using quantum-behaved sparse dictionary learning. *Knowledge-Based Systems, 252*, 109323.
35. Amra, I. A. A., & Maghari, A. Y. (2017, May). Students' performance prediction using KNN and Naïve Bayesian. In *2017 8th international conference on information technology (ICIT)* (pp. 909–913). IEEE. https://www.researchgate.net/profile/Ashraf-Maghari/publication/320672596_Students_performance_prediction_using_KNN_and_Naive_Bayesian/links/5c1b8b18a6fdccfc705b14c2/Students-performance-prediction-using-KNN-and-Naive-Bayesian.pdf
36. Staudemeyer, R. C., & Morris, E. R. (2019). *Understanding LSTM – a tutorial into long short-term memory recurrent neural networks.* arXiv preprint arXiv:1909.09586. https://arxiv.org/pdf/1909.09586
37. Dey, R., & Salem, F. M. (2017, August). Gate-variants of gated recurrent unit (GRU) neural networks. In *2017 IEEE 60th international Midwest symposium on circuits and systems* (MWSCAS) (pp. 1597–1600). IEEE. https://arxiv.org/pdf/1701.05923

38. Heck, J. C., & Salem, F. M. (2017, August). Simplified minimal gated unit variations for recurrent neural networks. In *2017 IEEE 60th international Midwest symposium on circuits and systems* (MWSCAS) (pp. 1593–1596). IEEE. https://arxiv.org/pdf/1701.03452
39. Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*. https://arxiv.org/pdf/1412.3555
40. Ballakur, A. A., & Arya, A. (2020, October). Empirical evaluation of gated recurrent neural network architectures in aviation delay prediction. In *2020 5th International Conference on Computing, Communication and Security (ICCCS)* (pp. 1–7). IEEE. https://www.researchgate.net/profile/Arti-Arya-2/publication/347540112_Empirical_Evaluation_of_Gated_Recurrent_Neural_Network_Architectures_in_Aviation_Delay_Prediction/links/636396322f4bca7fd02b1eb7/Empirical-Evaluation-of-Gated-Recurrent-Neural-Network-Architectures-in-Aviation-Delay-Prediction.pdf
41. Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2015, June). Gated feedback recurrent neural networks. In *International conference on machine learning* (pp. 2067–2075). PMLR. http://proceedings.mlr.press/v37/chung15.pdf
42. Hamayel, M. J., & Owda, A. Y. (2021). A novel cryptocurrency price prediction model using GRU, LSTM and bi-LSTM machine learning algorithms. *AI, 2*(4), 477–496. https://www.mdpi.com/2673-2688/2/4/30/pdf
43. Becerra-Rico, J., Aceves-Fernández, M. A., Esquivel-Escalante, K., & Pedraza-Ortega, J. C. (2020). Airborne particle pollution predictive model using gated recurrent unit (GRU) deep neural networks. *Earth Science Informatics, 13*, 821–834. https://idp.springer.com/authorize/casa?redirect_uri=https://link.springer.com/article/10.1007/s12145-020-00462-9&casa_token=4lniCxVM4LIAAAAA:mtakuR83Qkhp6ru8si9f_-_qTwN2OzsiKmd1aHiISlEsr5cawBjti6lpbERzJzwSyd_McuvA484-dlMslQ
44. Pan, M., Zhou, H., Cao, J., Liu, Y., Hao, J., Li, S., & Chen, C. H. (2020). Water level prediction model based on GRU and CNN. *IEEE Access, 8*, 60090–60100. https://ieeexplore.ieee.org/iel7/6287639/8948470/09044367.pdf
45. Fu, R., Zhang, Z., & Li, L. (2016, November). Using LSTM and GRU neural network methods for traffic flow prediction. In *2016 31st Youth academic annual conference of Chinese association of automation (YAC)* (pp. 324–328). IEEE. https://www.researchgate.net/profile/Li-Li-86/publication/312402649_Using_LSTM_and_GRU_neural_network_methods_for_traffic_flow_prediction/links/5c20d38d299bf12be3971696/Using-LSTM-and-GRU-neural-network-methods-for-traffic-flow-prediction.pdf
46. Mahjoub, S., Chrifi-Alaoui, L., Marhic, B., & Delahoche, L. (2022). Predicting energy consumption using LSTM, multi-layer GRU and drop-GRU neural networks. *Sensors, 22*(11), 4062. https://www.mdpi.com/1424-8220/22/11/4062
47. Zhou, X., Xu, J., Zeng, P., & Meng, X. (2019, February). Air pollutant concentration prediction based on GRU method. *Journal of Physics: Conference Series, 1168*(3), 032058. IOP Publishing. https://iopscience.iop.org/article/10.1088/1742-6596/1168/3/032058/pdf
48. Gao, Y., Wang, R., & Zhou, E. (2021). Stock prediction based on optimized LSTM and GRU models. *Scientific Programming, 2021*, 1–8. https://www.hindawi.com/journals/sp/2021/4055281/
49. Thakare, A., Anter, A. M., & Abraham, A. (2023). Seizure disorders recognition model from EEG signals using new probabilistic particle swarm optimizer and sequential differential evolution. *Multidimensional Systems and Signal Processing, 34*, 1–25.
50. Anter, A. M., Abd Elaziz, M., & Zhang, Z. (2022). Real-time epileptic seizure recognition using Bayesian genetic whale optimizer and adaptive machine learning. *Future Generation Computer Systems, 127*, 426–434.
51. Gudadhe, S., Thakare, A., & Anter, A. M. (2023). A novel machine learning-based feature extraction method for classifying intracranial hemorrhage computed tomography images. *Healthcare Analytics, 3*, 100196.

52. Anter, A. M., Oliva, D., Thakare, A., & Zhang, Z. (2021). AFCM-LSMA: New intelligent model based on Lévy slime mould algorithm and adaptive fuzzy C-means for identification of COVID-19 infection from chest X-ray images. *Advanced Engineering Informatics, 49*, 101317.
53. Anter, A. M., Moemen, Y. S., Darwish, A., & Hassanien, A. E. (2020). Multi-target QSAR modelling of chemo-genomic data analysis based on extreme learning machine. *Knowledge-Based Systems, 188*, 104977.
54. Al-Shourbaji, I., Kachare, P. H., Abualigah, L., Abdelhag, M. E., Elnaim, B., Anter, A. M., & Gandomi, A. H. (2022). A deep batch normalized convolution approach for improving COVID-19 detection from chest X-ray images. *Pathogens, 12*(1), 17.
55. Anter, A. M., & Abualigah, L. (2023). Deep federated machine learning-based optimization methods for liver tumor diagnosis: A review. *Archives of Computational Methods in Engineering, 30*(5), 3359–3378.
56. Anter, A. M., & Zhang, Z. (2023). RLWOA-SOFL: A new learning model-based reinforcement swarm intelligence and self-organizing deep fuzzy rules for fMRI pain decoding. *IEEE Transactions on Affective Computing*. https://doi.org/10.1109/TAFFC.2023.3285997
57. Anter, A. M., & Ali, M. (2020). Feature selection strategy based on hybrid crow search optimization algorithm integrated with chaos theory and fuzzy c-means algorithm for medical diagnosis problems. *Soft Computing, 24*(3), 1565–1584.
58. Anter, A. M., Wei, Y., Su, J., Yuan, Y., Lei, B., Duan, G., et al. (2019). A robust swarm intelligence-based feature selection model for neuro-fuzzy recognition of mild cognitive impairment from resting-state fMRI. *Information Sciences, 503*, 670–687.

AI Based Medicine Intake Tracker

Gulbakshee Dharmale, Dipti Patil, Swati Shekapure, and Aditi Chougule

1 Introduction

In this Android-based application an automated reminder system is built where patients can set a reminder for their medications. It emphasizes the contact between doctors and patients. Patients will be notified through a message within the system. They have the option of looking for a doctor for assistance. Patients will be provided with doctor contact information based on their availability. Also, patients will be notified about the expiry date of the medicine and the former history of the medicines can be stored for further reference [1, 2]. The proposed system prioritizes good user interface and easy navigation. We attempted to create a reminder system that is cost-effective, time-saving, and promotes medication adherence. Children, teenagers as well as all fall into the group of patients because we all have a rigorous schedule. If the patient's at home, someone may look after him or her, but if the patient is not at home, is out of the city or state, it is difficult for family members to contact and remind them of their dose schedules on a regular basis [3–5]. We rely entirely on devices, particularly smartphones, in our rapidly growing and technologically reliant lives. Everyone nowadays owns a smartphone. As a result, we will be able to make greater use of technology and make it more beneficial to us. It also plays a vital role in our daily lives, assisting us in remaining fit in a variety of ways. So, we're developing an android application whose goal is to use an alarm. This system is for folks who forget to take their medications on time. Users may establish

G. Dharmale (✉) · A. Chougule
Pimpri Chinchwad College of Engineering, Pune, India
e-mail: gulbakshi.dharmale@pccoepune.org

D. Patil
MKSSS's Cummins college of engineering for Women, Pune, India

S. Shekapure
Marathwada Mitra Mandal's College of Engineering, Pune, India

A. M. Anter et al. (eds.), *Nature-Inspired Methods for Smart Healthcare Systems and Medical Data*, https://doi.org/10.1007/978-3-031-45952-8_2

an alarm using the fields of date, time, and drug description, allowing them to set alarms for numerous medicines at different time intervals.

The problems we are working on:

1. Patients unable to take medicines on time due to their busy schedule.
2. Storing the prescriptions from doctors and health records like x-rays, reports of various tests.
3. Easy retrieval of previous health records.
4. Doctor assistance for patients with modest concerns (such as dosage/medication changes).

2 Related Work

Many different platforms and concepts have been used to construct medication systems. There are numerous worries regarding the functionality of healthcare-related apps, whose use is growing. My Therapy is an app for recording medications. This software not only monitors your medication consumption, but also your emotions and general health [6–8]. You may use the app to keep track of your symptoms and get specific treatment advice. My Therapy records your health data and creates a visual record that you can share with your doctor to help them figure out which aspects of your health need to be addressed [9–11, 17]. The Groove Health app is another. This app's built-in artificial intelligence engine helps you comprehend the medication you're taking by responding to any pharmaceutical or health-related questions a patient may have [18]. The app allows users to learn more about their medication, set personalized reminders to help them stay on track, and share their success with friends, family, or careers. Moreover, Round Health maintains a patient's prescription history, allowing them to keep track of how many pills they have taken and how many they have missed. The user will receive notifications from the Pill Reminder app until they mark the medication as taken [12, 13, 19]. Reminder systems now in use have various drawbacks. Here are a few illustrations: They don't offer any tools for tracking or checking a drug's expiration date. There is no option for automatically adding the medicine name after scanning the prescription. There are currently no facilities for storing patient medical histories and previous medications. Also, none of the systems hold health-related papers, such as maternity documents and various scan reports. Many of the systems need the purchasing of specialized hardware. Some systems include a built-in alert tone that users cannot change.

After comparing REMICARE system with few similar applications and studying the functionality of every module in those applications. It is observed that every application is User/Patient centric. There is no involvement of a doctor in the application. Whereas this application is Patient-Doctor centric. It is also observed that there is no facility of storing original documents on the cloud. And the facility of scanning prescription is not seen in any other similar application. The analysis of REMICARE with two other similar applications is described in the Table 1 below.

Table 1 Analysis of REMICARE

Functionality	REMICARE	My therapy	Care aid
Scanning prescription	Yes	No	No
Real time doctor support	Yes	No	No
Healthcare document Storage	Yes	No	No

3 REMICARE System

This REMICARE system offers a digital solution in the form of an Android application that can store personal health records and remind patients to take their medications on time. In this system, users are able to scan and upload the prescription image, which is then stored in the cloud. After scanning the prescription, the user will be able to set a reminder for the medication's name. The user can also add additional medications if he requires a reminder for his other medications. Additionally, he can manually enter the expiration date, and the application will send an alert message to the user prior to the expiration date. In addition, this application will provide a doctor's assistance for patients with trivial questions, such as dosage changes, the proper time to take a particular medication, and which medications can cause side effects when taken with specific foods, etc. Patients can also input their height and weight to calculate their BMI. Medical records storage is also essential for every individual. A woman can, for instance, store her maternity records. Additionally, it will be beneficial for patients with chronic diseases. During each visit, physicians can consult previous records. The REMICARE System Architecture as depicted in Fig. 1.

The REMICARE system, which runs on Android, uses automatic alarm buzzing and messaging to remind users to take their meds on schedule.

- OCR Technique is used to scan text from prescriptions.
- Patients can maintain their medical records on the cloud and access them when necessary.
- Patients and Doctors can communicate via the Chat-with-Doctor feature of this application to address any small problems.

3.1 Patient Login Module

If the patient needs something or has some inquiries, this app might be helpful. Even after the patient is released from the hospital, this application can remind him to take his medication on time. In order to communicate with their doctors, the user has the option. Doctors are able to modify anything. For instance, if the dosage of a certain medication needs to be altered. Also, it is possible to upload and amend the patient's nutrition record. With the app, the user can contact the doctor with queries

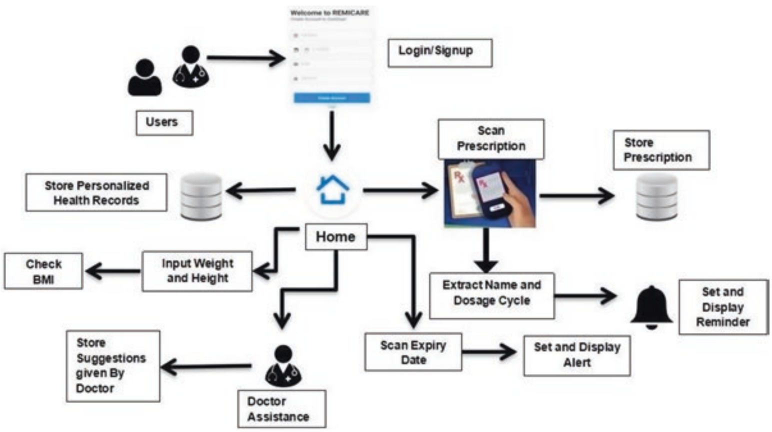


Fig. 1 REMICARE system architecture

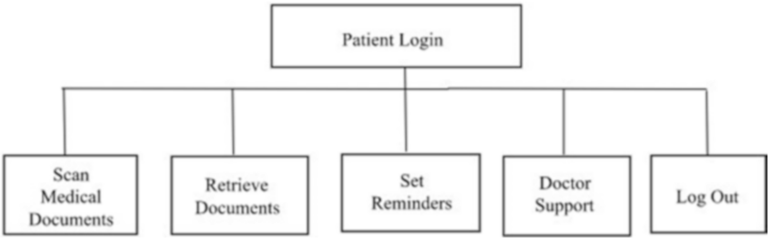


Fig. 2 Patient login module

and offer health and fitness information. If a patient is or was suffering from a sickness or illness, the physician can access the patient's cloud-stored medical history and treat him or her accordingly.

Figure 2 illustrates the patient login module. After logging onto the REMICARE system, the user/patient will be able to view the many options, including scanning medications and necessary papers, create a reminder, receive physician help, retrieve previously stored documents, and log out. After scanning a medicine script, image processing will extract the drug's name, and the remaining information will be set according to the medication schedule. Following this, the patient will receive timely medication reminders. Chat with Doctor Feature is helpful for the treatment of numerous serious illnesses, such as Covid-19. Doctors are able to converse with patients, suggest improvements, and modify doses of drugs based on their needs.

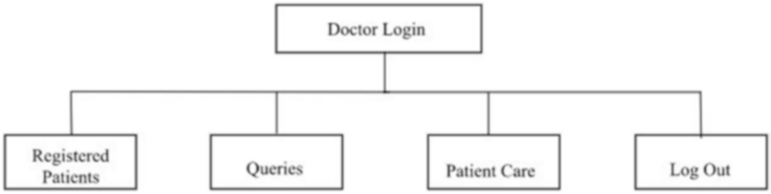


Fig. 3 Doctor login module

3.2 Doctor Login Module

Doctors can login into REMICARE system through doctor's login module. Doctor will keep track of patient's medication history. Doctors will provide suggestion to patient and solve their minor health issues using Chat with Doctor Feature. Notification of patient's medication will send to doctor.

Figure 3 illustrated the login module for physicians. After logging into the system, the doctor is able to view the list of registered patients. Physicians can respond to the questions listed in the questions section. In the patient care section, a physician can also provide recommendations to a specific patient and modify the drug schedule.

4 Result and Analysis

In this system, there are two logins, one for the patient and another for the doctor. For patient login, he/she needs to give some basic details like name, etc. and register first. We have provided different attributes in this application. Whenever a user will log into his account he will encounter the different features of the system. If the user wants to set-up a reminder for his medication, then he will have to enter details like the name of the medicine, time slot, days etc. User can select the notification tune as per his choice. Another feature of this application is that it will alert users about the expiry date of the particular medicine. When a user buys a new medicine, he can enter the medicine name and its expiry date. System can alert the user based on the date he has entered. Keeping medical records is essential for every person. There is one more feature in which we can scan and store all our medical prescriptions with the help of cloud storage. One could save their pregnancy records. During each visit, physicians can consult previous records. Additionally, it will be beneficial for patients with chronic diseases. Additionally, interaction between patient and doctor is also possible. The user can communicate with their doctors directly. Doctors can update everything for e.g.: if there is a need to change the dosage of a particular medicine. It is also accessible to update and upload a patient's diet chart. If a patient has questions, he can contact the doctor and provide wellness updates through the REMICARE system.

This system puts forward an automated solution to remind patients about their medications with the help of an android application. The real objective for creating this REMICARE system is to remind users to take their meds on schedule. User has to enter the time slot and the medicine name and it will set a reminder and alert the user on time. Additionally, it is possible for the user to add extra meds as well as make changes in the previous entries. Along with this, it will alert users about the expiry date of the medicine. Whenever a patient restocks his medicines, he/she can enter the name and expiry date of that particular medicine. Application will also store personal health records. Users have to scan and upload their old prescriptions and will be stored on cloud. Doctors can refer to these records in their further appointments with that particular patient. This will provide access to our medical history to us as well as our doctor.

The REMICARE system will also offer medical advice to patients with minor concerns, such as how often to take a medication, when it is best to take it, which medications may have unwanted side effects if taken with a specific food item, etc. The use of this app is not just limited to chronic diseases but we can also use this for the people suffering from covid-19 or omicron. Doctor – Patient Interaction proves to be very useful if the patient is suffering from Covid-19. All the health data, prescriptions and other reports or medical records can be stored and along with this patient can update the doctor about his/her improvements in health after getting discharged. Doctors can also update about the change in dose or diet if necessary. If a user wants to calculate his BMI, it is possible with the help of this application.

4.1 Login / Sign-Up Module

In this module, the user has to enter his details like email id and set a password. Login id will be registered and account will be created as shown in Fig. 4. For the sign-up page, the user has to enter the details like full name, date of birth, email id and password. Later on, after logging into the system, it will redirect to the home page. This page ensures the user registration and authentication with the firebase database. Users can register him/her on the application and later on they can log in to the application as given in Fig. 5.

4.2 Health Document Storage Module

This module ensures the important health documents of the user are stored safely to the cloud with easy access. These files uploaded to the cloud through the application can be available any time for the user to view and download. For storing documents, we use the Flutter Firebase Storage plugin. This plugin is used to use the Firebase Cloud Storage API. Users can access your files from both Firebase and Google Cloud thanks to Cloud storage for Firebase, which saves them in a Google Cloud

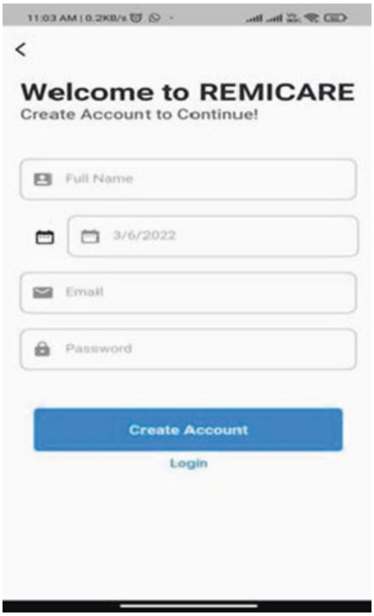


Fig. 4 Sign up page

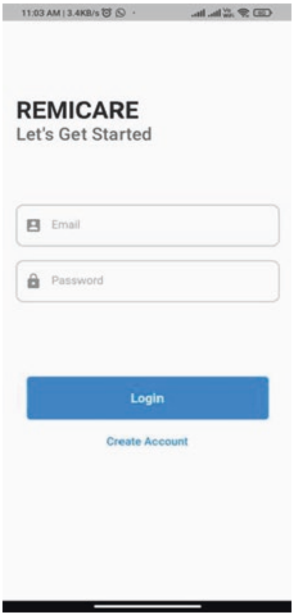


Fig. 5 Login page

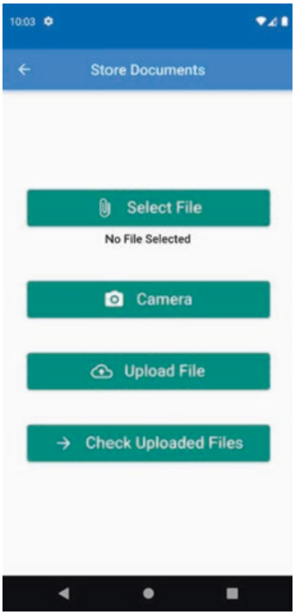


Fig. 6 Module for uploading documents module

Storage bucket. Users have the option to upload and download files from mobile clients thanks to the Firebase SDKs for Cloud Storage.

In this module, the user can scan the prescription, documents and upload it securely on the cloud as shown in Figure 6. Nowadays we tend to misplace the hard copy of the documents or healthcare records. In case of some chronic diseases, it's important to store all the records from the beginning so that the doctor can refer to it every time. Users can also scan the digital prescription printed by the doctor and this application will extract the medicine name, time and dosage from the prescription to set the reminder. It is also possible to store the maternity records so the mother can refer to them a second time if necessary.

Patients can also retrieve the documents uploaded as shown in Fig. 7. This is very beneficial for the user of the system since there is no need to store a bunch of old hardcopies. Users will have to just upload the prescription or reports and use these documents whenever needed. This is beneficial if the user forgets or misplaces the old records or if the user is in some other city.

4.3 Medication Reminder Module

The module is implemented for the patients to be able to set reminders for the tablets, syrups or injections they need to take. The user is supposed to enter the medicine name and the dosage of the medicine. Also, users can take a picture of the

Fig. 7 Retrieve documents

prescription and the application will automatically extract the prescription details. This module uses Google ML Kit API for text recognition. This API works with the OCR algorithm for text recognition. Text is divided into blocks, lines, and elements by the Text Recognizer. In general, a block is a continuous group of text lines, like a paragraph or column [14, 15]. A Line is a continuous line of words arranged along a single axis. A continuous group of alphabetic symbols arranged along the same axis is known as an element. For every detected blocks, lines, and items, the API returns the bounding boxes, corner points, recognized languages, and recognized text [16]. After extracting all the details from the prescription, we can set the alarm by using the Flutter Alarm Clock plugin.

In this module, users can add the reminders shown in Fig. 8. User has to select the medicine type then enter the name of the tablet and dosage. Then the user can add the time, time interval and date so that the remainder will be set accordingly. Another facility for scanning the printed prescription and extracting the data is also provided. User will scan the prescription then the data will be extracted. Patients who frequently fail to take their medications on time are the target audience for this particular module.

Nowadays many doctors are switching to a digital prescription. This system can scan the text from the printed prescription as given in Fig. 9. Data like medicine name, dosage, time etc. will be extracted from the prescription and reminder will be added automatically.

Fig. 8 Add reminder

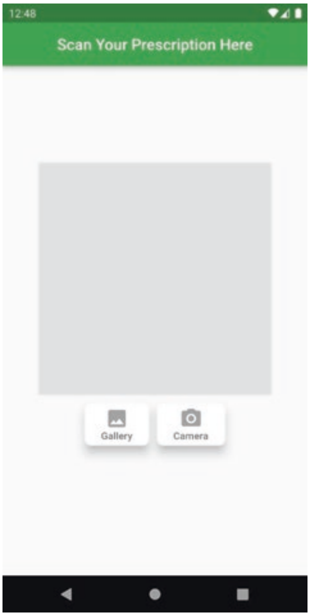


Fig. 9 Scan prescriptions

Fig. 10 Chat with Doctor

4.4 Doctor Support Module

This module focuses on the interaction between patients and doctors. The queries a patient has are solved by the doctor the patient trusts and has visited. The user will be able to take follow ups with the doctor through the application itself. Users can send images of their prescription and documents by using the Flutter Image Picker plugin. Users can send previously saved photographs from their Android smartphone or snap real-time images with the aid of this plugin.

Users will be able to take follow-ups with the doctor through this application. Instead of waiting in the queue, the user can solve his trivial doubts directly. If a doctor wants to update the diet or dosage, then he can do it on the app itself. It can be helpful for some communicable diseases like Covid-19. Figure 10 depicts Chat with Doctor Feature of REMICARE system.

5 Conclusion

In this project we have put forward a digital solution to remind patients about their medications by developing an android application. This application has a unique feature which will remind us to take our medicines on time. Once the user enters the name and time slot of the medicine as prescribed by the doctor, it will set a reminder and alarm the user at that particular time. Users can also select any alarm tone from the given options. Users can also add extra meds in case he needs a reminder for his other medications. If the user did not take his/her meds then it will send the notification to their emergency contacts provided by the user from the application itself. This system provides the feature in which it will alert us about the expiry date of the medicine. Whenever a patient restocks his medicines, he/she can enter the name and expiry date of that particular medicine. Once we enter all the details, the application will automatically remind us about the expiry date when it's near so there is no need to check the expiry date more often. Furthermore, this application also stores personal healthcare records. It is possible for users to scan and upload the image of the prescription which will be stored on the cloud. This will provide access to our medical history to us as well as our doctor if access is given by the patient. Doctors can refer to these records in their further appointments with that particular patient and assess the patient accordingly. In addition to this our application will also provide doctor's assistance for patients having trivial doubts such as change in dose or medicine, proper time to take the particular medicine, which medicine can cause some side effects when taken with a particular food item etc. The use of this app is not just limited to chronic diseases but we can also use this for the people suffering from Covid-19 or Omicron. Doctor – Patient Interaction which is possible during normal diseases impacts greatly in patients' recovery can also be very useful for a Covid-19 patient who is quarantined during treatment. All the health data, prescriptions and other reports or medical records can be stored and along with this patient can update the doctor about his/her improvements in health after getting discharged. Doctors can also update about the change in dose or diet if necessary.

6 Future Scope

- The developed product can be used in various healthcare domains such as:
  Integration of IOT: Integrating IOT Devices with the current application will widen its usefulness. Various IOT Devices like smart watches and fitness bands can be integrated with the app and it can maintain the records such as stress level, SpO2 level, BPM, calories burnt, exercise time and all these records can be saved on the application.
- Diet Recommendation: Provision of doctors providing the patient a diet through the application is something which can be an addition to this solution. Doctors can update the diet chart depending on the patient's requirement.
- Video call: This feature can prove helpful for doctor patient interaction. Patients can arrange a video call instead of going to the clinic for follow-up.

References

1. Vyavahare, S., Sagade, M., Hajari, K., & Surwase, S. (2020). Handwritten cursive English text recognition using deep CNN-RNN based CT. *International Journal of Future Generation Communication and Networking, 13*(2s), 564–569.
2. Anter, A. M., Elaziz, M. A., & Zhang, Z. (2022). Real-time epileptic seizure recognition using Bayesian genetic whale optimizer and adaptive machine learning. *Future Generation Computer Systems, 127*, 426–434.
3. Gupta, A., Srivastava, M., & Mahanta, C. (2011). Offline handwritten character recognition using neural networks. In *ICCAIE 2011–2011 IEEE Conference on Computer Applications and Industrial Electronics*. IEEE. https://doi.org/10.1109/ICCAIE.2011.6162113
4. Olaleye, S. B. (2021). Security of sensitive data on android smartphones using cloud storage with reference to gravitational search algorithm. *International Journal of Computer Science and Mobile Computing, 10*(3), 72–82.
5. Anter, A. M., Mohamed, A. W., Zhang, M., & Zhang, Z. (2023). A robust intelligence regression model for monitoring Parkinson's disease based on speech signals. *Future Generation Computer Systems, 147*, 316–327.
6. Olaleye, S. B., & Kant, S. (2018). Secure use of cloud storage of data on smartphones using atomic AES on ARM architectures. *International Journal of Applied Engineering Research, 13*(5), 2569–2581.
7. Mehala, M., & Viji Gripsy, J. (2020). Voice based medicine reminder alert application for elder people. *International Journal of Recent Technology and Engineering (IJRTE), 8*(6), 2284–2289. ISSN: 2277–3878.
8. Fang, K. Y., Maeder, A. J., & Bjering, H. (2016). Current trends in electronic medication reminders for self care. *Studies in Health Technology and Informatics, 231*, 31–41.
9. Poonguzhali, P. K., Prajyot, D., Chaithanya, M. K., & Patil, M. U. (2016). Secure storage of data on android based devices. *International Journal of Engineering and Technology, 8*, 177–182.
10. Santo, K., Chow, C. K., Thiagalingam, A., et al. (2017). MEDication reminder APPs to improve medication adherence in coronary heart disease (MedApp-CHD) study: A randomized controlled trial protocol. *BMJ Open, 7*, e017540. https://doi.org/10.1136/bmjopen-2017-017540
11. Hayakawa, M., Uchimura, Y., Omae, K., Waki, K., Fujita, H., & Ohe, K. (2013). A smartphone-based medication self-management system with real-time medication monitoring. *Applied Clinical Informatics, 4*, 37–52.
12. Zao, J. K. (SMIEEE), Wang, M. Y., Tsai, P., & Liu, J. W. (FIEEE) (2010, July). Smart phone based medicine in-take scheduler, reminder and monitor. *The 12th IEEE international conference on e-health networking, applications and services*. IEEE. 978-1-4244-6376-31101$26.00 ©2010 IEEE
13. Ameta, D., Mudaliar, K., & Patel, P. (2015). Medication reminder and healthcare – An android application. *International Journal of Managing Public Sector Information and Communication Technologies, 6*, 39–48. https://doi.org/10.5121/ijmpict.2015.6204.20
14. Shi, B., Bai, X., & Yao, C. (2015, July). An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 39*, 99.
15. Bagyalakshmi, N., Sai Adhitya, B. S., & Youvashree, K. M. (2019, March). A review on medicine reminder and adherence system. *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering an ISO 3297: 2007 Certified Organization, 8*(1), 59.
16. Purohit, A. (2016). A literature survey on handwritten character recognition. *(IJCSIT) International Journal of Computer Science and Information Technologies, 7*(1), 1–5.

17. Anter, A. M., & Zhang, Z. (2023). RLWOA-SOFL: A new learning model-based reinforcement swarm intelligence and self-organizing deep fuzzy rules for fMRI pain decoding. *IEEE Transactions on Affective Computing*. https://doi.org/10.1109/TAFFC.2023.3285997
18. Anter, A. M., & Abualigah, L. (2023). Deep federated machine learning-based optimization methods for liver tumor diagnosis: A review. *Archives of Computational Methods in Engineering, 30*(5), 3359–3378.
19. Thakare, A., Anter, A. M., & Abraham, A. (2023). Seizure disorders recognition model from EEG signals using new probabilistic particle swarm optimizer and sequential differential evolution. *Multidimensional Systems and Signal Processing, 34*, 1–25.

Analysis of Genetic Mutations Using Nature-Inspired Optimization Methods and Classification Approach

Anuradha Thakare, Pradnya Narkhede, and Sahil S. Adrakatti

1 Introduction

Cancer is a complex disease caused by genetic mutations that can occur in different parts of the body, leading to abnormal growth of cells and tumor formation. Identifying and classifying genetic mutations is essential for cancer diagnosis and treatment, as it can provide valuable information on the specific type of cancer and its potential response to treatment. However, detecting and analyzing genetic mutations can be difficult due to the vast number of possible mutations and their complex interactions. Early cancer detection can improve the chances of successful treatment and increase the chances of survival. The classification of genetic mutations can provide insights into the specific type of cancer and its potential response to treatment, allowing for personalized and targeted therapies. Failure to detect and classify genetic mutations can lead to misdiagnosis, inappropriate treatment, and poor clinical outcomes [1–3].

Nature-inspired optimization methods, such as Genetic Algorithms (GA) [4] and Particle Swarm Optimization (PSO) [5] are computational algorithms inspired by natural phenomena such as swarms, genetic evolution, and neural networks. These methods have shown promise in solving complex optimization problems, including feature selection and classification in cancer diagnosis. By leveraging the power of these optimization methods, it is possible to improve the accuracy and efficiency of classification models for genetic mutations in cancer patients. Additionally, these methods can be used to identify new biomarkers and potential targets for cancer treatment, leading to better patient outcomes.

A. Thakare (✉) · P. Narkhede · S. S. Adrakatti
Pimpri Chinchwad College of Engineering, Pune, Maharashtra, India
e-mail: anuradha.thakare@pccoepune.org; pradnya.narkhede@pccoepune.org

A. M. Anter et al. (eds.), *Nature-Inspired Methods for Smart Healthcare Systems and Medical Data*, https://doi.org/10.1007/978-3-031-45952-8_3

2 Related Research

The [6] article overviews the genetic alterations contributing to cancer development. It discusses the types of somatic mutations, copy number alterations, and structural variations that can lead to oncogenes' activation or tumor suppressor genes'inactivation. The authors also highlight the importance of identifying genetic alterations in cancer diagnosis, prognosis, and treatment and the challenges and opportunities for precision medicine.

The [7] Paper discusses an overview of cancer genomics, from discovering oncogenes and tumor suppressor genes to developing personalized medicine. The authors discuss the advances in genomic technologies, such as next-generation sequencing that have enabled the identification of genetic alterations in tumors. They also highlight the challenges and opportunities for using genomic information to guide cancer diagnosis, prognosis, and treatment.

In [8], this review article provides an overview of the genomic alterations contributing to cancer development, including somatic mutations, copy number alterations, and structural variations. The authors discuss emerging technologies for detecting and analyzing tumor genetic alterations, such as single-cell sequencing and liquid biopsy. They also highlight the challenges and opportunities for using genomic information to guide cancer diagnosis, prognosis, and treatment.

In [9], the review article provides an overview of the genetic mutation that contribute to breast cancer development, including structural variations, copy number alterations, and somatic mutations. The authors discuss the clinical implications of genetic testing for breast cancer diagnosis and treatment, including targeted therapies and immunotherapies.

In [10], the review article provides an overview of the molecular profiling of cancer, including identifying genetic alterations that drive cancer biology and using genomic information for personalized medicine. The authors discuss the advances in genomic technologies, such as whole-genome sequencing and transcriptomic that have enabled the identification of genetic alterations in tumors. They also highlight the challenges and opportunities for using genomic information to guide cancer diagnosis, prognosis, and treatment.

The [11] review article discusses the applications of machine learning in cancer prediction and prognosis. It provides an overview of various machine learning techniques, including neural networks, random forests, decision trees, and support vector machines, their applications in cancer classification using genetic mutations.

In [12] research article proposes a deep learning approach for classifying cancer types based on copy number alterations. The authors developed a convolutional neural network model and applied it to genomic data from 13 cancer types. They demonstrated that their approach achieved high accuracy in cancer classification and outperformed other machine learning methods.

In [13] This review article discusses the applications of machine learning in predicting the pathogenicity of genetic variants associated with cancer. The authors provide an overview of various machine learning techniques, including Deep learning, random forests, decision trees, and support vector machines, and their applications in cancer genetics.

In [14] This research article proposes machine learning models for predicting oncogenic mutations in cancer patients. The authors developed logistic regression models and applied them to genomic data from cancer patients. They demonstrated that their approach achieved high accuracy in predicting oncogenic mutations and outperformed other machine learning methods.

In [15] This review article discusses the applications of machine learning in identifying driver mutations in cancer genomics. The authors provide an overview of various machine learning techniques, including neural networks, random forests, and support vector machines, and their applications in identifying driver mutations.

In [16] This research article proposes a deep learning approach for classifying genetic variants in cancer genes. The authors developed a deep neural network model and applied it to genomic data from cancer patients. They demonstrated that their approach achieved high accuracy in classifying genetic variants and outperformed other machine learning methods.

A genetic algorithm-based feature selection method for cancer classification utilizing microarray gene expression data was suggested in the study by Shahla Nosrati et al. The authors utilized a support vector machine (SVM) classifier to categorise cancers and a genetic algorithm to choose the most pertinent genes from the microarray data. The proposed strategy beat existing feature selection approaches in terms of accuracy and effectiveness when tested on six different cancer datasets, according to the results [17].

For detecting cancer driver genes, Xiao-Li Li et al. proposed a hybrid optimization technique. To find driver genes linked to cancer, the authors combined the genetic algorithm (GA) with particle swarm optimization (PSO). The proposed method beat previous optimization algorithms in terms of accuracy and stability when evaluated on four cancer datasets, according to the results [18].

A hybrid artificial bee colony optimization algorithm for cancer classification utilizing gene expression data was suggested in the publication by M. A. Arvind et al. The most informative genes from the microarray data were chosen by the authors using the artificial bee colony (ABC) algorithm, and the cancer types were categorized using a support vector machine (SVM) classifier. Three different cancer datasets were used to test the suggested strategy, and the results revealed that it performed better than other feature selection approaches in terms of accuracy and stability [19].

Ant colony optimization (ACO) was suggested by R. Balamurugan et al. as a method for improving gene expression data for the categorization of cancer. The authors utilized a decision tree classifier to categories cancers and an ACO algorithm to choose the most pertinent genes from the microarray data. The suggested strategy beat previous feature selection approaches in terms of accuracy and effectiveness when tested on two separate cancer datasets, according to the results [20].

M. Karthikeyan et al. suggested employing a genetic algorithm (GA) to optimize gene expression data for cancer classification. The most important genes from the microarray data were chosen by the authors using a GA method, and the cancer types were determined using an SVM classifier. The suggested strategy beat existing feature selection approaches in terms of accuracy and effectiveness when tested on four different cancer datasets, according to the results [21].

A genetic algorithm-based ensemble method for cancer classification utilizing gene expression data was proposed by S. Sathishkumar et al. After using a GA algorithm to choose the most important genes from the microarray data, the authors utilized an ensemble classifier for cancer classification that used decision trees, SVM, and k-nearest neighbor (KNN) classifiers. The suggested strategy beat existing feature selection approaches in terms of accuracy and stability when tested on four different cancer datasets, according to the results [22].

3 Proposed Classification Model (Diagram and Description)

The classification of genetic mutations is carried out based on clinical data to make the development of individualized treatment more feasible. Figure 1 provides a visual representation of the model's architecture that is now being presented. The main modules of the proposed system are

1. Input Data: Predefined Doc and Unclassified Doc
2. Exploratory Data Analysis
3. Preprocessing:
   (a) Text Cleaning
   (b) Feature Extraction
   (c) Standardization of Data
   (d) Dimensionality Reduction

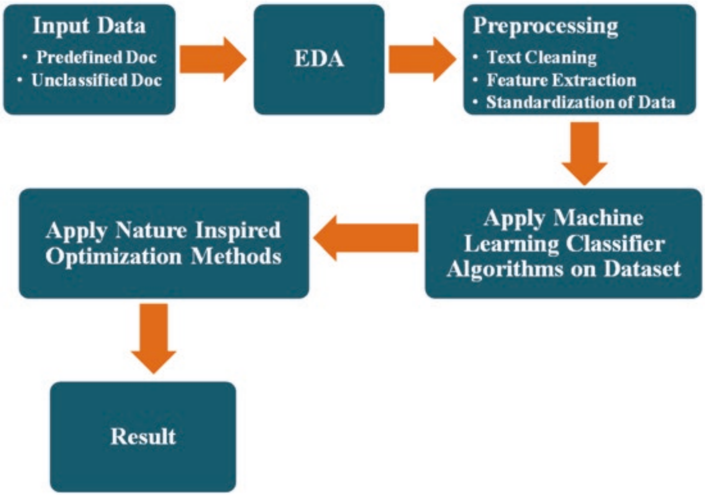


Fig. 1 Architectural diagram for the proposed method

4. Apply Machine Learning Classification Algorithms
   (a) Logistic Regression
   (b) Decision Tree
   (c) Support Vector Machine
   (d) Random Forest
   (e) K-nearest neighbor
   (f) Naive Bayes
   (g) Genetic Naive Bayes
5. Apply Nature Inspired Optimization Methods
   (a) Genetic Algorithm
   (b) PSO Optimization Algorithm
   (c) Bee colony Optimization Algorithm

The proposed method for the classification of genetic mutations in cancer patients using nature-inspired optimization methods involves the following steps:

1. Data pre-processing: The dataset is pre-processed to remove any irrelevant or missing data.
2. Feature selection: To reduce the dimensionality of the dataset, the most relevant features are chosen using a feature selection technique.
3. Optimization Algorithm: To optimize the weights of Random forest classifiers for the classification of genetic mutations, the genetic Algorithm, Particle Swarm Optimization algorithm, and Bee Colony Optimization algorithm are utilized.
4. Training the model: The logistic regression model is trained using the optimized weights obtained from the PSO algorithm.
5. Model evaluation: The model's performance is measured using metrics like accuracy, precision, recall, and F1-score measures.
6. Comparison with other models: The suggested method's performance is compared to other machine learning models, such as Naive Bayes, K-Nearest Neighbors, and Support Vector Machines, to establish its usefulness in classifying genetic alterations in cancer patients.
7. Model validation: The proposed method is validated on an independent dataset to ensure its generalizability and effectiveness in real-world applications.

4 Algorithms for the Proposed Approach

4.1 ML Algorithms

4.1.1 Random Forest

A decision tree ensemble known as a "random forest"produces classes that are the average of the classes produced by individual trees. Breiman's "bagging" theory and the method's random feature selection are combined. It introduces "Bagging"

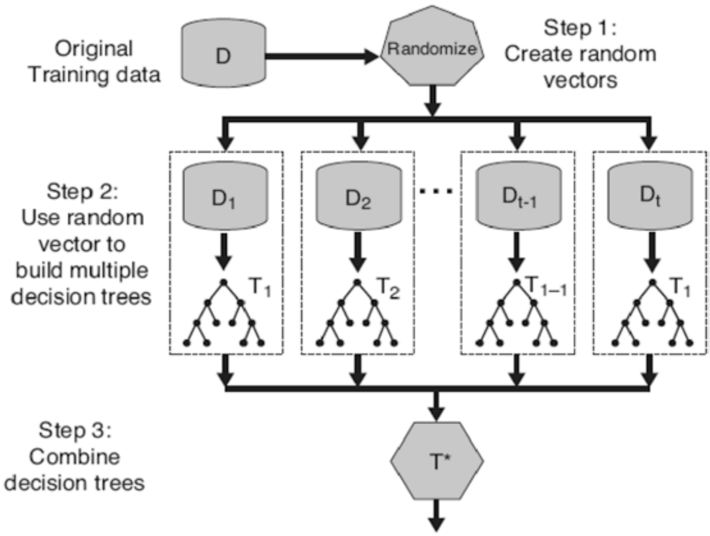


Fig. 2 Random forests

and "Random input vectors" as two sources of randomization. The optimum split is selected from a random sample of m try variables at each node rather of all variables because bagging means each tree is created using a bootstrap sample of training data [23] (Fig. 2).

Each tree is planted & grown as follows:

1. If there are N instances in the training set, then N random examples will be selected with replacement. The tree will be developed using this data set as training data.
2. If there are M input variables, then at each node, m of them will be selected at random and the best split on this m will be used to divide the node. The value of m does not change as the forest expands.
3. Every tree is developed to its full potential. Nothing is pruned.

4.1.2 Support Vector Machine

Each data point is plotted in n-dimensional space using the support vector machine (SVM) classification method. Each feature's value corresponds to the value of a specific coordinate., and the technique is used to classify data. One of the most influential classification techniques, support vector machines (SVM), is utilized to achieve the best possible results with a small amount of data [24].

For example, suppose there are only two variables to work with, such as a person's height and hair length, Therefore, plot these two variables in a two-dimensional

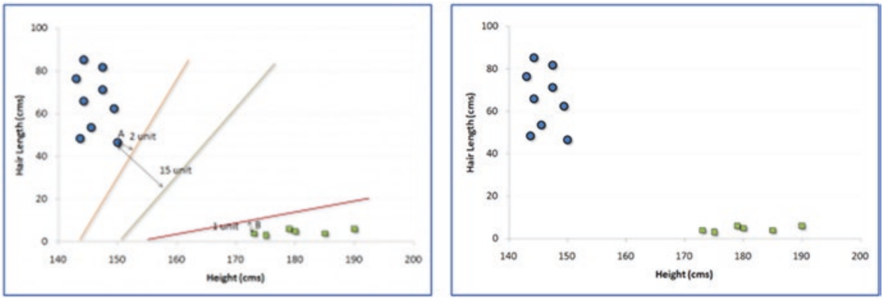


Fig. 3 SVM

graph with each point having two coordinates. The term “support vectors” is used to describe these coordinates.

Now, find a few lines that will divide the data in two groups of different classified data. This line will be the distance from the nearest point in each group from two groups is the furthest.

In the above example shown in Fig. 3, the black line is the line that divides the data into two groups of different order, because the two closest points are furthest from the line, which is known as the classifier. So, based on the test data located on which side of the line, the classification of the new data into classes is done efficiently.

4.1.3 KNN

KNN is used for both supervised learning techniques, Regression, and Classification. It frequently appears in categorization issues in companies. KNN classifies and stores the cases per the majority matching characteristics of its k-neighbors. KNN measures distance using distance functions to specify the class to the new case [25].

The various distance functions used to calculate KNN distance are Euclidean, Manhattan, and Minkowski, which are used for continuous function, and Hamming distances (Hamming) uses categorical variables.

Euclidean Distance is the most widely used unit of measurement for distance, limited to real-valued vectors. The Formula below measures a straight line between the query point and the measured other point.

$$d(x,y) = \sqrt{\sum_{i=1}^{n} (y_i - x_i)^2}$$

If K is equal to 1, the instance is then merely put into the class of its closest neighbor. Choosing K can occasionally be difficult when using KNN modeling (Fig. 4).

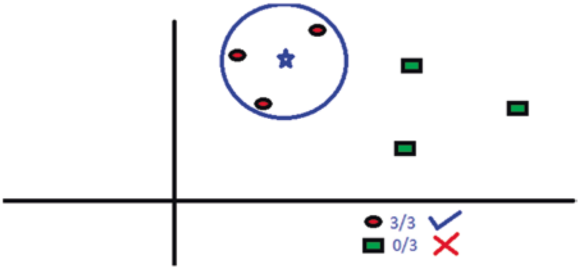


Fig 4 KNN

Things to Think About Before Choosing KNN:

- The computational cost of KNN is high.
- Variables should be standardized to prevent bias caused by greater range variables.

Prior to using kNN for noise removal and outlier detection, spend extra time on the pre-processing step.

4.1.4 Naïve-Bayes

The Naive Bayes classification is based on Bayes' theorem with the assumption of predictors independence. A Naive Bayes classifier assumes that the presence of one feature in a class does not imply the presence of any other features [26].

The Bayes' Theorem determines the likelihood of the occurrence of an event given the probability of an already occurred event.

Bayes' theorem equation is stated as following:

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

where A and B are considered as events with condition P(B)! = 0.

Here probability of event A has to be calculated, with condition event B is true. Event B is known as evidence. The priory of A is denoted by P(A). Posteriori probability of B is denoted as P(A|B).

Bayes Theorem is basis for Naïve Bayes Classifier

$$P(y|X) = \frac{P(X|y)P(y)}{P(X)}$$

where, y is class variable and X is a dependent feature vector (of size n) where: $X = (x_1, x_2, x_3, \ldots \ldots, x_n)$

4.1.5 Logistic Regression

Logistic regression is a binary classification algorithm. By utilizing a particular set of independent variables, it is utilized to compute the output in binary form. Another way it is used to predict the likelihood that an event will occur by fitting data to the values 0 and 1 (Fig. 5).

4.2 Genetic Naive-Bayes

Input: Genetic Mutations and Text dataset
Output: Classified Dataset

1. Input: Genetic Mutations (G) and Clinical text data (T)
2. Combine the datasets (D) = G + T
3. Cleaning of dataset
4. Extract features (H)
5. t = 0;
6. generate random population (P(t)) from extracted features(H);
7. calculate fitness function (F) score for each sample (P(t))
8. while not termination do
9. Pp(t) = P(t). select Parents ();
10. Pc(t) = reproduction(Pp);
11. mutate(Pc(t));
12. evaluate(Pc(t));
13. P(t + 1) = build next generation from (Pc(t), P(t));
14. t = t + 1;

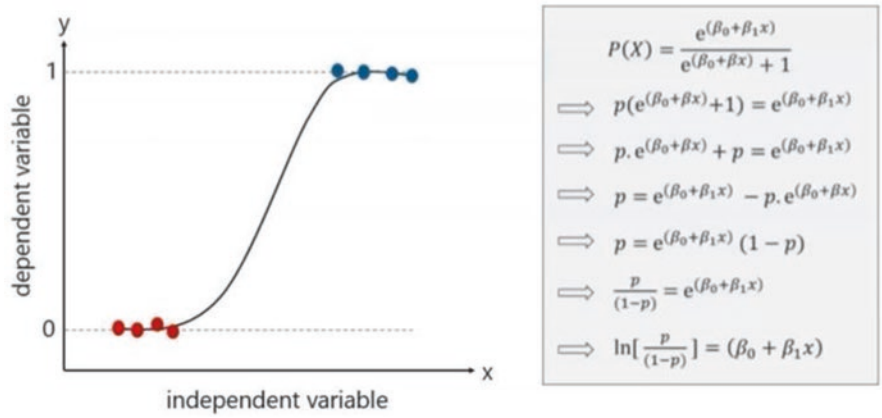


Fig 5 Logistic regression

15. Apply Naive bayes classifier to (P(t + 1))
16. end
17. Output: Classified Genetic Mutations

4.3 Natured-Inspired Algorithms

Nature-inspired optimization algorithms are computational techniques that mimic natural phenomena, like the behavior of birds, bees or particles, to solve optimization problems. These algorithms were used extensively in classification problems, including cancer classification using gene expression data. Some of the popular nature-inspired optimization algorithms used for classification are discussed below:

4.3.1 Particle Swarm Optimization

Particle Swarm Optimization (PSO) is a nature-based optimization algorithm used for solving optimization problems. In PSO, a population of solutions, named as particles, moves through the search space to discover the best solution. Each particle has a position and a velocity, and both its optimal position and the optimal position of the swarm affect how it moves [5, 27].

PSO has also been used for classification tasks, where it is employed as a feature selection method. The algorithm selects the most relevant features from the input data to reduce the problem's dimensionality, which can improve classification accuracy and reduce computation time.

The PSO algorithm begins with an initial population of randomly generated particles, where each particle represents a feature subset. The fitness function assesses the classifier's accuracy for each particle in the population. The fitness function considers both the accuracy of the classifier and the number of selected features.

Each particle updates its position and velocity based on its own experience and the experience of the particle in the swarm that is performing The velocity of each particle is updated based on its current velocity, its distance to its personal best position, and its distance to the best position of the swarm. Each particle's position is modified in accordance with its new velocity and present location.

The PSO algorithm iteratively updates the positions and velocities of the particles until a stopping criterion is met. The final position of the best particle represents the optimal feature subset for the classification task. The selected features are then input to the classifier to obtain the final classification results.

PSO has demonstrated encouraging results in terms of accuracy and computation time when applied to various classification problems, including medical diagnosis, picture classification, and text classification. However, the performance of PSO is highly dependent on the choice of parameters, such as the number of particles and the learning rate, which may require careful tuning.

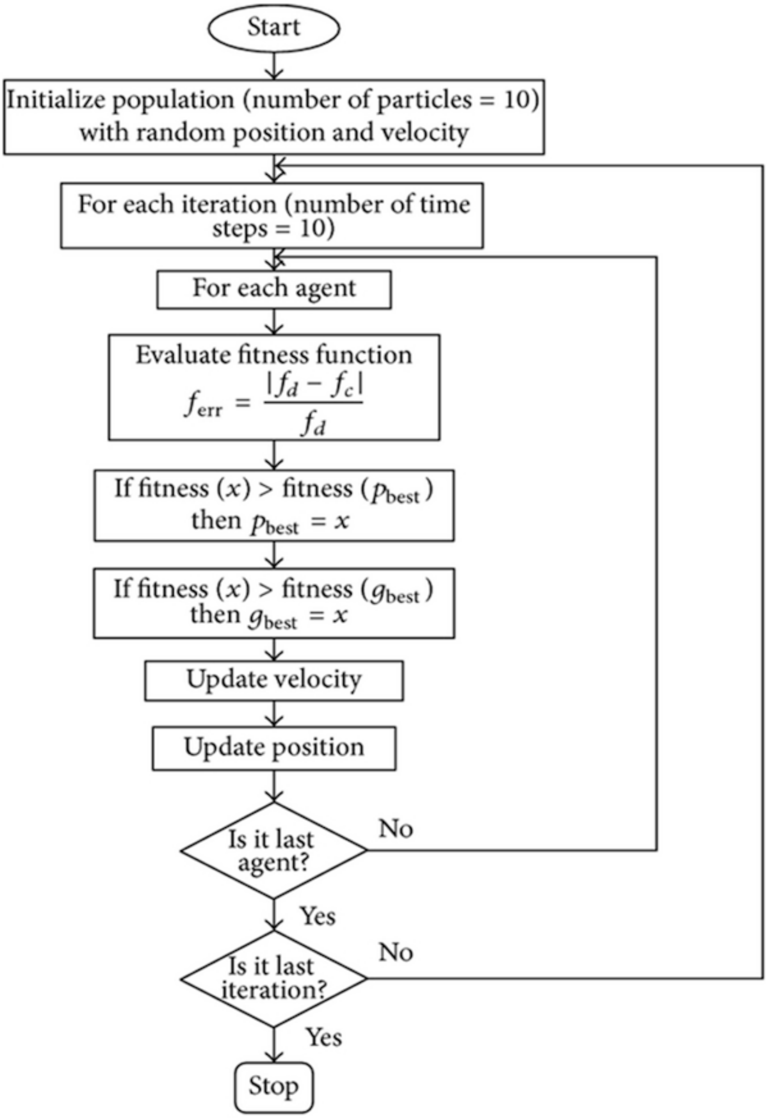


Fig. 6 Flowchart of PSO algorithm

The figure below shows the basic flowchart of the PSO algorithm (Fig. 6).

Some of the control parameters that affect the basic PSO are problem size, particle count, acceleration coefficients, inertia weight, neighborhood size, iterations, and random values that scale the contribution of cognitive and social components. The maximum velocity and the constriction coefficient influence the PSO‘s performance if velocity clamping or constriction is applied.

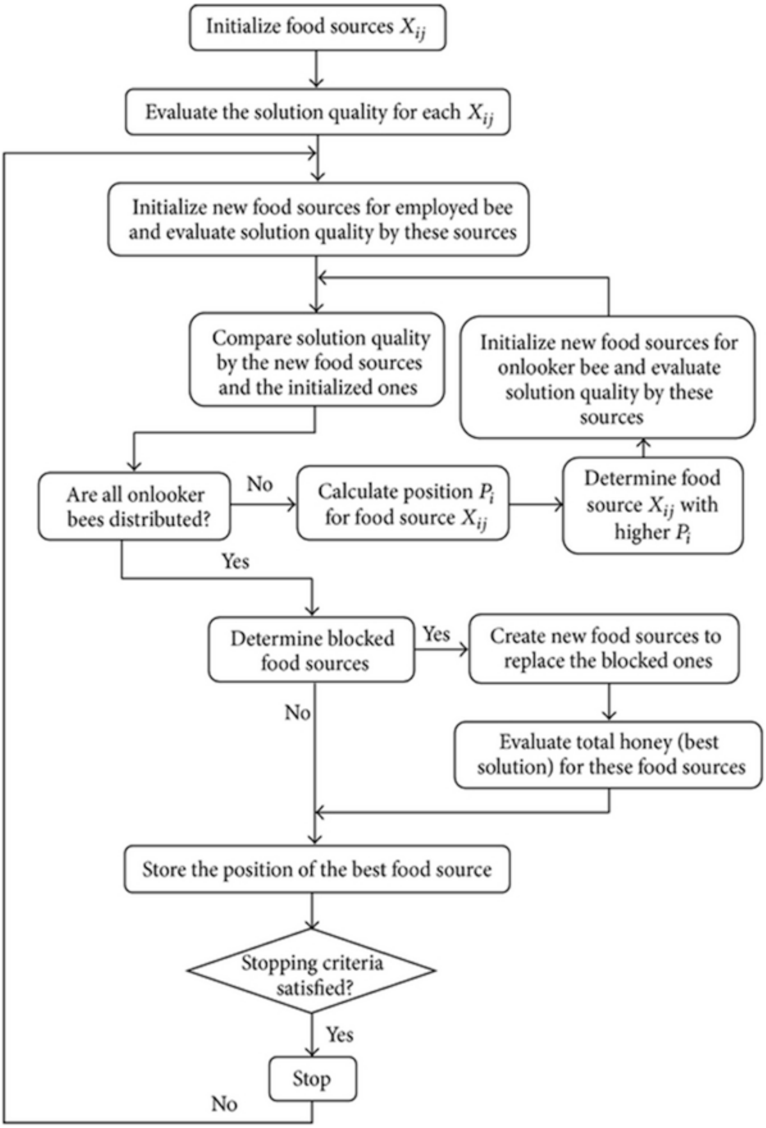


Fig. 7 Artificial Bee colony optimization

4.3.2 Bee Colony Optimization

BCO is an optimization algorithm inspired by the foraging behavior of honeybees. The algorithm uses two types of bees: employed bees and onlooker bees. The employed bees search for food sources in the search space and share the information with the onlooker bees. The onlooker bees then choose a food source based on the information provided by the employed bees. The Fig. 7 below shows the basic flowchart of the BCO algorithm [28].

4.3.3 Genetic Algorithm

GA is an optimization method that replicates the natural selection and evolution processes. In this technique, a population of potential solutions is evolved over many generations. Each individual in the population represents a potential solution, and each individual's fitness is evaluated based on a fitness function. Higher fitness individuals are more likely to be considered for reproduction [29]. The Genetic algorithm's fundamental flowchart is depicted in the Fig. 8.

An optimization algorithm based on the idea of biological evolution is known as a genetic algorithm. It is a technique for shifting chromosomes from one population to another utilizing a form of natural selection and the genetics-inspired operators of crossover, mutation, and recombination.

Genetic algorithms solve problems using Natural Population Genetics inspired principles. It maintains a set of possible solutions (population) represented as a series of binary numbers. New series are produced in each generation by 1. Decoding each series and assessing its ability to solve the problem. Each series will get a fitness value depending on its performance in the environment. 2. Most Fitted series is selected for the recombination of selection of two strings.

Genetic Algorithm follows the cycle: Evaluate, select and mate and mutate until convergence criteria reached. Criteria are: 1. Let the Genetic algorithm run for certain no. of cycles. 2. allow Genetic algorithms to run until a reasonable solution is found.

Procedure of Genetic Algorithm:

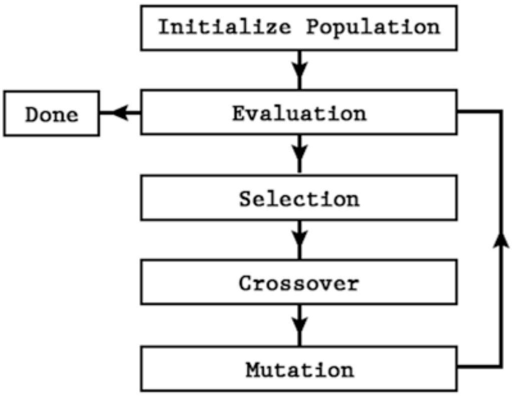


Fig. 8 Generic algorithm flow chart

5 Results and Discussion

5.1 Dataset Description

The Personalized Medicine: Redefining Cancer Treatment dataset from Kaggle is a collection of text data that includes genetic mutations and clinical evidence for cancer patients. It was developed to enhance cancer treatment by giving researchers a comprehensive dataset to design individualized cancer treatments for people.

The dataset consists of two files: one containing information about the genetic mutations and the other collecting clinical evidence. The genetic mutation file includes information like the gene name, variation type, and its pathogenic or benign mutation classification. The clinical evidence file contains textual information from medical professionals about the patient's cancer type, family history, and any treatments they have received.

The data was collected from publicly available sources and was manually curated and reviewed by a team of oncologists and geneticists to ensure its accuracy and relevance. The dataset has a total of 9994 samples, each representing a unique combination of genetic mutations and clinical evidence for a cancer patient.

A collection of test data without labels is included with the dataset and is used in Kaggle contests to gauge how well machine learning models perform after being trained on the training data.

The Training and Test data sets are provided in two different files. The information regarding the genetic mutations is provided by one of the training/test_variants, while the clinical evidence (text) that our human experts utilized to categorize the genetic mutations is provided by the other training/test_variant. Through the ID field, both are related. Some of the test data is produced by the machine to avoid hand labeling.

Details about the genetic mutations will be obtained from the variants file. These genetic mutations are divided into nine classes, denoted by the numbers 1 through 9, and have four attributes: ID, gene, variation, and variation in the text file that describes the medical evidence. It has two attributes: 1. ID 2. clinical evidence. Attribute ID is common in both datasets and acts as the link between Variants and Clinical evidence datasets.

- ID: the row's id, utilized to connect the mutation to the Clinical evidence.
- Gene: Location of Genetic mutation.
- Variation: change for mutation by the amino acid.
- Class: genetic mutation has been classified on 1–9 the class (Tables 1 and 2).

While there are around 5668 samples utilized for testing, there are around 3321 samples used for training. Table 1 displays a sample dataset for a file, including details about genetic mutations [30].

Table 1 Training text data

	ID	Text
0	0	Cyclin-dependent kinases (CDKs) regulate a var...
1	1	Abstract Background Non-small cell lung canc...
2	2	Abstract Background Non-small cell lung canc...
3	3	Recent evidence has demonstrated that acquired...
4	4	Oncogenic mutations in the monomeric Casitas B...

Table 2 Training data for genetic mutation

	ID	Gene	Variation	Class
0	0	FAM58A	Truncating Mutations	1
1	1	CBL	W802*	2
2	2	CBL	Q249E	2
3	3	CBL	N454D	3
4	4	CBL	L399V	4

5.2 Exploratory Analysis

Exploratory Data Analysis (EDA) approach is used for data analysis. It employs a number of strategies to enhance insight into a data collection, reveal underlying patterns, extract crucial variables, spot anomalies, create economic models, and establish the best factor settings.

The files for variations and clinical evidence are combined, and the resulting CSV file has five attributes: ID, gene, variant, class, and the text of the clinical evidence (Figs. 9 and 10) (Table 3).

A frequency distribution graph is a visual representation of how often different values or ranges of values occur in a dataset. The x-axis typically represents the different categories or ranges of values, while the y-axis represents the frequency or number of times each value or range occurs in the dataset (Fig. 11).

5.3 Performance Measures

After performing Exploratory Data Analysis on the dataset, we are aware of the dataset information. Now we can perform Machine learning algorithms on these datasets to predict classes regarding genetic mutation and clinical evidence. Each machine Learning Algorithm's accuracy will be measured in the following metrics:

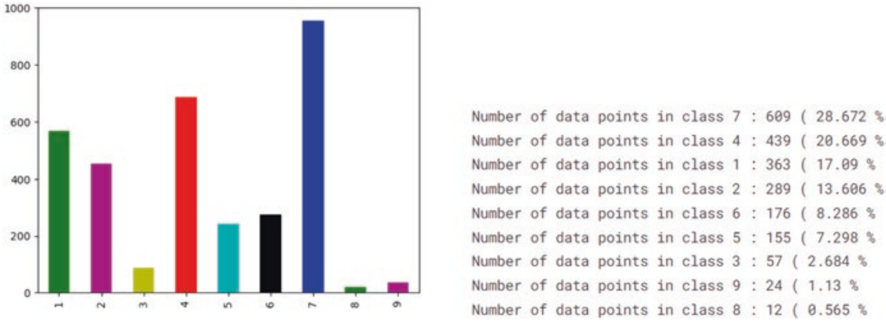


Fig. 9 Distribution of training dataset among nine classes

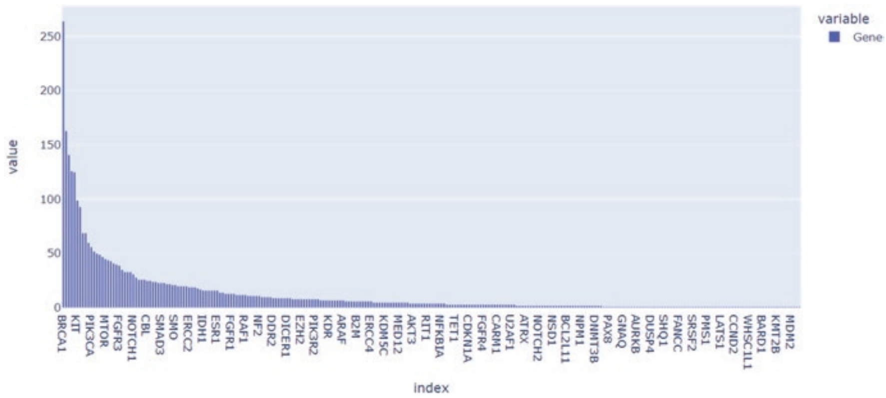


Fig. 10 Frequency distribution of all Gene

Table 3 Combination text and genetic mutation training dataset

	ID	Gene	Variation	Class	Text
0	0	FAM58A	Truncating Mutations	1	Cyclin-dependent kinases (CDKs) regulate a var...
1	1	CBL	W802*	2	Abstract Background Non-small cell lung canc...
2	2	CBL	Q249E	2	Abstract Background Non-small cell lung canc...
3	3	CBL	N454D	3	Recent evidence has demonstrated that acquired...
4	4	CBL	L399V	4	Oncogenic mutations in the monomeric Casitas B...

1. Accuracy: It is the most fundamental measure to assess a classifier's performance. Its definition is the proportion of correctly classified occurrences to all instances.
2. Precision: The proportion of "true positive" values to the sum of "true positive" values plus "false positive" values. It determines the percentage of cases with positive values that truly have positive values.

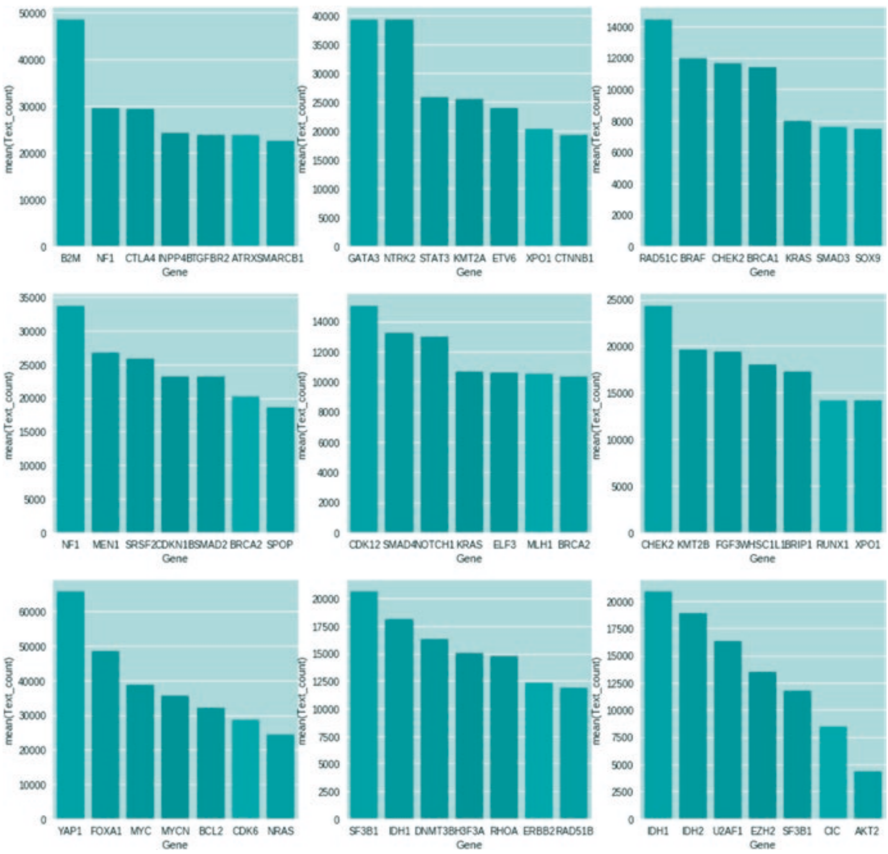


Fig. 11 Gene frequency per class

3. Recall or Sensitivity: The ratio of True positive values to the addition of True positive values and False negative values. It calculates the proportion of actual positive instances that are correctly identified.
4. F1 Score: The average recall and accuracy with time. It strikes a good chord between memory recall and factual accuracy.
5. Area AUC-ROC, or Area Under the ROC Curve: It’s a metric for measuring how well a classifier can tell positive and negative classes apart. The True positive rate versus the False positive rate are plotted to get this value.
6. Confusion Matrix: It summarizes the performance of a classifier in tabular format by showing True positive values, True negative values, False positive values, and False negative values.
7. Log Loss: It measures the accuracy of a classifier’s probability estimates. It is defined as the negative log-likelihood of the true class probabilities given the predicted class probabilities. It is a commonly used metric in binary classification problems.

5.4 Classification of Genetic Mutations Using ML Algorithms

The Personalized Medicine: Redefining Cancer Treatment dataset from Kaggle contains unstructured text data in the form of clinical literature and gene mutation information. Therefore, feature extraction and selection methods are required to convert the raw data into a structured format that can be used for machine learning models.

Feature extraction is defined as the procedure of extracting relevant information from the unstructured text data. In this dataset, the feature extraction methods used:

1. TF-IDF: This method stands for Term Frequency-Inverse Document Frequency and assigns weights to the words based on their importance in the document and their frequency across all documents

Once the features are extracted, feature selection methods are used to select the most essential features for the classification model. The feature selection methods used in this dataset include

1. Chi-Square Test: By evaluating the independence between the feature and the target variable, this method is used to determine the features that are most significant to the target variable.
2. Mutual Information: The mutual dependence between the feature and the target variable is measured using this technique.
3. Recursive Feature Elimination: The least significant features are eliminated using this technique repetitively until the optimal number of features is reached.

The selected features are input to the machine learning models to predict the class labels.

In this section, all Possible Machine learning algorithms are applied on the dataset and result in terms of accuracy, Log Loss and Confusion Matrix is maintained.

5.4.1 Random Forest

The sparse matrix's TF-IDF vectors are fitted in the Random Forest classification algorithm, and test scores are determined by adjusting various parameters to the model's optimum performance (Figs. 12, 13 and 14) (Table 4).

5.4.2 Support Vector Machine

SVMs are effective machine learning classification algorithms. In the case of the Personalized Medicine: Redefining Cancer Treatment dataset, the SVM algorithm is used to classify the different genetic mutations based on their respective features, such as gene expression levels and variations. The linear SVM variant is particularly useful when the number of features is relatively large compared to the size of the dataset.

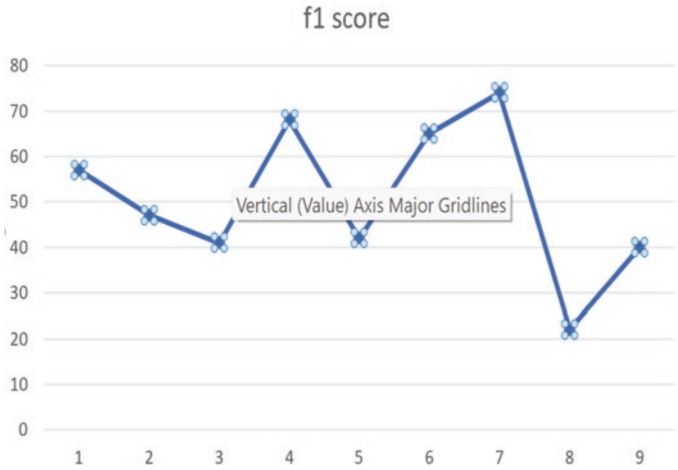


Fig. 12 The graph represents the F1-measure of all nine classifications

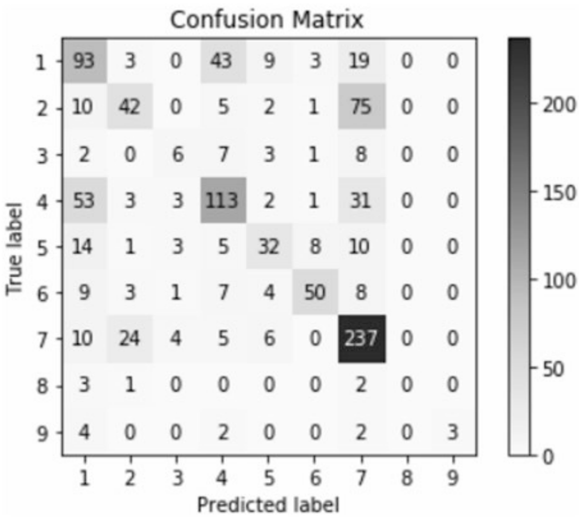


Fig. 13 Confusion matrix using random forest

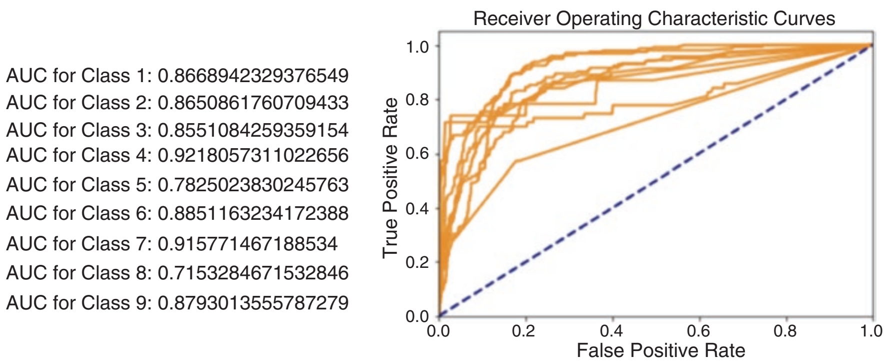


Fig. 14 ROC curve for the classification of mutations

Table 4 Accuracy and log loss for random forest

Sr. No	Feature extraction	Classification	Accuracy	Log loss
1	TF-IDF	Random Forest	64.9%	2.06

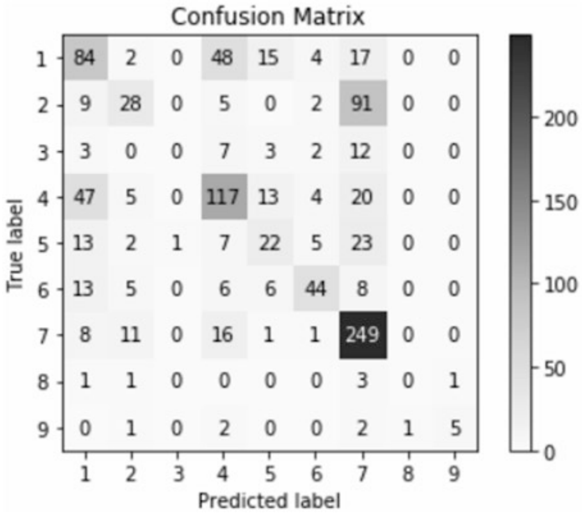


Fig. 15 Confusion matrix for SVM

Table 5 Accuracy and log loss for Support vector machine

Sr. No.	Feature extraction	Classification	Accuracy	Log loss
1	TF-IDF	Support vector machine	62.8%	1.20

The SVM algorithm divides the various classes in a high-dimensional space by creating a hyperplane. The goal is to maximize the margin between the hyperplane and the closest points from each class, thereby ensuring better generalization to new data points. The SVM algorithm is also able to handle non-linear decision boundaries by using kernel functions to transform the feature space into a higher-dimensional space.

Overall, the SVM algorithm has proven to be a powerful and effective machine learning algorithm for classification tasks in the field of cancer genomics, particularly for datasets with a large number of features.

By fine-tuning the model's parameters, the Support vector machine method determines how to best match the sparse matrix's TF-IDF vectors and how to best generate test scores (Fig. 15) (Table 5).

5.4.3 KNN

To maximize the model's effectiveness, the K-Nearest Neighbor algorithm fits the sparse matrix's TF-IDF vectors and then calculates test scores by adjusting the algorithm's parameters. If k = 5 is selected, the results are as shown in below figure (Figs. 16 and 17) (Table 6).

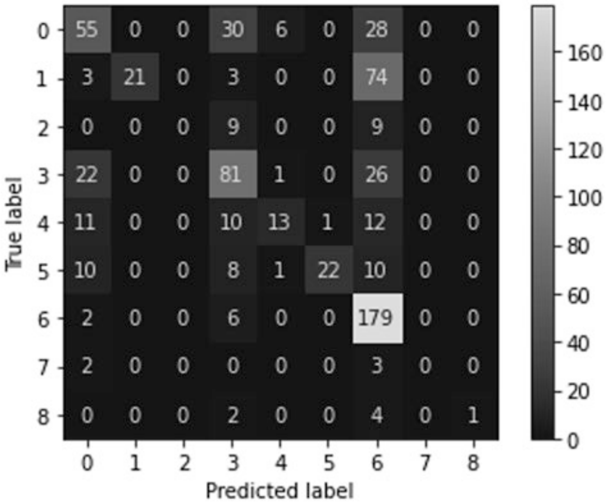


Fig. 16 Confusion matrix for KNN algorithm

Classification report:

	precision	recall	f1-score	support
1	0.39	0.52	0.45	86
2	0.46	0.66	0.54	165
3	0.56	0.47	0.51	36
4	0.53	0.47	0.50	38
5	0.67	0.54	0.60	24
6	0.73	0.62	0.67	29
7	0.48	0.28	0.36	18
8	0.62	0.29	0.40	14
9	0.58	0.26	0.36	23
accuracy			0.55	433
macro avg	0.56	0.46	0.49	433
weighted avg	0.54	0.55	0.53	433

Fig. 17 Classification report for KNN

Table 6 Accuracy and log loss of K nearest neighbor

Sr. No	Feature Extraction	Classification	Accuracy	Log Loss
1	TF-IDF	k-nearest neighbor	67%	1.08

```
Confusion Matrix:
 col_0    1  2   4  5   6    7
Class
1        52  0  26  0   2   90
2         1  3   0  0   0  142
3         0  0   6  2   0   20
4        15  0  96  3   0   94
5        33  0   4  8   4   23
6        10  0   2  1  27   33
7         1  0   1  0   0  285
8         0  0   0  0   0    3
9         0  0   0  0   0   10
```

Fig. 18 Confusion matrix for Naive Bayes

```
Classification report:
              precision    recall  f1-score   support

           1       0.46      0.31      0.37       170
           2       1.00      0.02      0.04       146
           3       0.00      0.00      0.00        28
           4       0.71      0.46      0.56       208
           5       0.57      0.11      0.19        72
           6       0.82      0.37      0.51        73
           7       0.41      0.99      0.58       287
           8       0.00      0.00      0.00         3
           9       0.00      0.00      0.00        10

    accuracy                           0.47       997
   macro avg       0.44      0.25      0.25       997
weighted avg       0.59      0.47      0.40       997
```

Fig. 19 Classification report for Naive Bayes algorithm

5.4.4 Naïve Bayes

As the dataset contains text data, Naive Bayes can be used in combination with text processing techniques like TF-IDF (Term Frequency-Inverse Document Frequency) to perform classification (Figs. 18 and 19).

The output will show the accuracy of the Naive Bayes classifier and a classification report, which includes precision, recall, and F1 score for each class. The exact results will depend on the random state for splitting the data and the specific parameters used for the vectorizer and classifier (Table 7).

Table 7 Accuracy and log loss for Naive Bayes

Sr. No	Feature extraction	Classification	Accuracy	Log loss
1	TF-IDF	Naive Bayes	47.24%	2.73

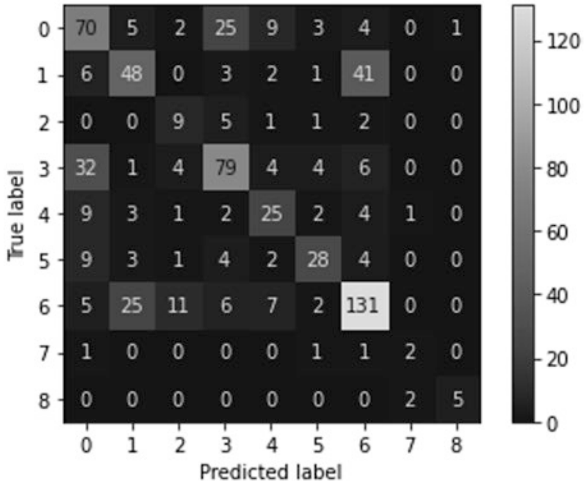


Fig. 20 Confusion matrix for Naive Bayes

Table 8 Accuracy and log loss for logistic regression

Sr. No	Feature extraction	Classification	Accuracy	Log loss
1	TF-IDF	Logistic regression	64%	1.06

5.4.5 Logistic Regression

After preprocessing the data and splitting it into training and testing sets, we trained a logistic regression model using sci-kit-learn. The model achieved an accuracy score of approximately 0.64 on the testing set (Fig. 20) (Table 8).

5.5 Classification of Genetic Mutations Using Genetic Algorithms

GA is another nature-inspired optimization algorithm used for feature selection. Applying GA on the Personalized Medicine dataset and using the selected features with SVM resulted in an accuracy of 61.8% (Figs. 21 and 22).

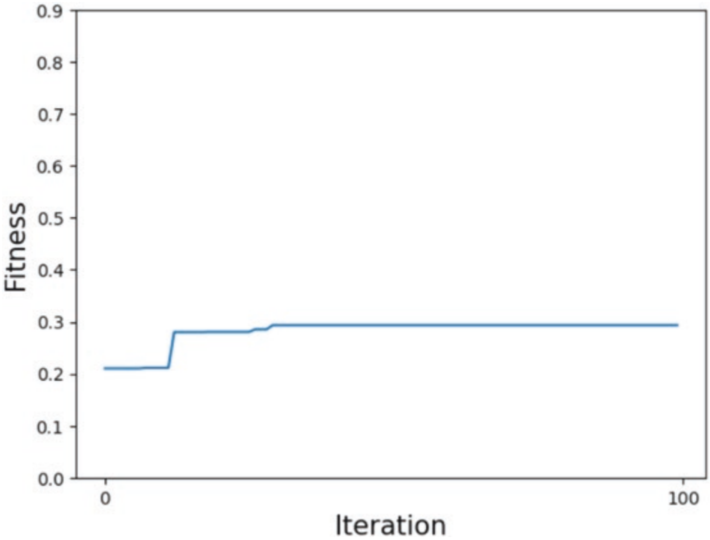


Fig. 21 Fitness vs iteration using genetic classification

Confusion Matrix:

col_0	1	2	3	4	5	6	7
Class							
1	86	4	1	24	1	3	2
2	26	29	0	7	0	0	29
3	6	8	1	3	0	0	4
4	82	1	5	31	0	3	3
5	29	2	0	3	5	2	3
6	25	10	0	5	0	1	0
7	78	31	0	14	0	3	83
8	2	0	0	1	0	0	1
9	7	0	0	0	0	0	1

Fig. 22 Confusion matrix for genetic algorithm

5.6 Classification of Genetic Mutations Using PSO Algorithms

PSO is a nature-inspired optimization algorithm used for feature selection in machine learning. Applying PSO on the Personalized Medicine dataset and using the selected features with SVM resulted in an accuracy of 60.6%.

```
SVC(random_state=42)
Accuracy: 60.60
Log loss: 1.20
```

```
RandomForestClassifier(random_state=42)
Accuracy: 71.0
Log loss: 1.00
```

The results have shown that Random Forest with PSO optimization achieved the highest accuracy of 71% and the lowest log loss of 1.00 (Fig. 23).

5.7 Classification of Genetic Mutations Using BCO Algorithms

ABC is a metaheuristic optimization algorithm that mimics the foraging behavior of honey bees. Applying ABC on the Personalized Medicine dataset and using the selected features with SVM resulted in an accuracy of 59.9%.

Confusion Matrix:

col_0	1	2	3	4	5	6	7
Class							
1	86	4	1	24	1	3	2
2	26	29	0	7	0	0	29
3	6	8	1	3	0	0	4
4	82	1	5	31	0	3	3
5	29	2	0	3	5	2	3
6	25	10	0	5	0	1	0
7	78	31	0	14	0	3	83
8	2	0	0	1	0	0	1
9	7	0	0	0	0	0	1

Fig. 23 Confusion matrix for PSO

6 Conclusion

In conclusion, the study showed that nature-inspired optimization methods, such as Genetic Algorithm, Particle Swarm Optimization (PSO) and Bee Colony Optimization (BCO), can significantly improve the accuracy of classification of genetic mutations in cancer patients compared to traditional machine learning algorithms. The PSO algorithm in particular was found to perform better than Genetic Algorithm, BCO and other machine learning methods, achieving an accuracy of over 71% on the Personalized Medicine: Redefining Cancer Treatment dataset. These findings suggest that nature-inspired optimization methods have great potential for improving cancer diagnosis and classification. Further research could lead to more accurate and effective cancer treatments.

References

1. Anter, A. M., & Hassenian, A. E. (2018). Computational intelligence optimization approach based on particle swarm optimizer and neutrosophic set for abdominal CT liver tumor segmentation. *Journal of Computational Science, 25*, 376–387.
2. ElSoud, M. A., & Anter, A. M. (2016). Computational intelligence optimization algorithm based on meta-heuristic social-spider: Case study on CT liver tumor diagnosis. *International Journal of Advanced Computer Science and Applications, 7*(4), 466.
3. Anter, A. M., Hassanien, A. E., ElSoud, M. A., & Kim, T. H. (2015, July). Feature selection approach based on social spider algorithm: Case study on abdominal CT liver tumor. In *2015 seventh international conference on advanced communication and networking (ACN)* (pp. 89–94). IEEE.
4. Anter, A. M., Moemen, Y. S., Darwish, A., & Hassanien, A. E. (2020). Multi-target QSAR modelling of chemo-genomic data analysis based on extreme learning machine. *Knowledge-Based Systems, 188*, 104977.
5. Thakare, A., Anter, A. M., & Abraham, A. (2023). Seizure disorders recognition model from EEG signals using new probabilistic particle swarm optimizer and sequential differential evolution. *Multidimensional Systems and Signal Processing, 34*, 1–25.
6. Brunner, S. F., Roberts, N. D., Wylie, L. A., Moore, L., Aitken, S. J., Davies, S. E., & Campbell, P. J. (2019). Somatic mutations and clonal dynamics in healthy and cirrhotic human liver. *Nature, 574*(7779), 538–542.
7. Middleton, G., Fletcher, P., Popat, S., Savage, J., Summers, Y., Greystoke, A., et al. (2020). The National Lung Matrix Trial of personalized therapy in lung cancer. *Nature, 583*(7818), 807–812.
8. Torkenczy, K., Langer, E., Fields, A., Turnidge, M., Nishida, A., Boniface, C., & Adey, A. (2020). *Integrated single-cell analysis reveals treatment-induced epigenetic homogenization*. Available at SSRN 3687026.
9. Andrade-Vieira, N. D. et al. (2020). The emerging landscape of genetic alterations in breast cancer. *Breast Cancer Research*.
10. Sicklick, J. K., Kato, S., Okamura, R., Schwaederle, M., Hahn, M. E., Williams, C. B., et al. (2019). Molecular profiling of cancer patients enables personalized combination therapy: The I-PREDICT study. *Nature Medicine, 25*(5), 744–750.
11. Fang, B., et al. (2021). Machine learning approaches in cancer prognosis and prediction. *Cancer Treatment Reviews*.

12. Costa, N. et al. (2021). Classification of cancer types based on copy number alterations using deep learning approach. *PLOS ONE*.
13. Mobiny, A. et al. (2021). Machine learning approaches for predicting the pathogenicity of genetic variants associated with cancer. *Briefings in Bioinformatics*.
14. Giacomelli, E. et al. (2020). Machine learning models for predicting oncogenic mutations in cancer patients. *BMC Cancer*.
15. Liang, Y., et al. (2020). Identifying driver mutations using machine learning approaches in cancer genomics. *Briefings in Bioinformatics*.
16. Habibi, J., et al. (2020). Deep learning approach for classifying genetic variants in cancer genes. *BMC Bioinformatics*.
17. Nosrati, S. et al. (2020). Genetic algorithm-based feature selection approach for cancer classification using microarray gene expression data. *Artificial Intelligence in Medicine*.
18. Li, X.-L., et al. (2020). A hybrid optimization algorithm for detecting driver genes in cancer. *BMC Bioinformatics*.
19. Arvind, M. A., et al. (2021). A hybrid artificial bee colony algorithm for cancer classification using gene expression data. *Journal of Ambient Intelligence and Humanized Computing*.
20. Balamurugan, R., et al. (2020). Optimization of gene expression data using ant colony optimization for classification of cancer. *Biocybernetics and Biomedical Engineering*.
21. Karthikeyan, M., et al. (2020). Optimization of gene expression data using genetic algorithm for classification of cancer. *International Journal of Intelligent Systems and Applications*.
22. Sathishkumar, S., et al. (2020). A genetic algorithm-based ensemble method for cancer classification using gene expression data. *Journal of Ambient Intelligence and Humanized Computing*.
23. Anter, A. M., Gupta, D., & Castillo, O. (2020). A novel parameter estimation in dynamic model via fuzzy swarm intelligence and chaos theory for faults in wastewater treatment plant. *Soft Computing, 24*(1), 111–129.
24. Gudadhe, S., Thakare, A., & Anter, A. M. (2023). A novel machine learning-based feature extraction method for classifying intracranial hemorrhage computed tomography images. *Healthcare Analytics, 3*, 100196.
25. Anter, A. M., & Ali, M. (2020). Feature selection strategy based on hybrid crow search optimization algorithm integrated with chaos theory and fuzzy c-means algorithm for medical diagnosis problems. *Soft Computing, 24*(3), 1565–1584.
26. Anter, A. M., & Zhang, Z. (2019, October). Adaptive neuro-fuzzy inference system-based chaotic swarm intelligence hybrid model for recognition of mild cognitive impairment from resting-state fMRI. In *International workshop on predictive intelligence in medicine* (pp. 23–33). Springer International Publishing.
27. Anter, A. M., & Abualigah, L. (2023). Deep federated machine learning-based optimization methods for liver tumor diagnosis: A review. *Archives of Computational Methods in Engineering, 30*(5), 3359–3378.
28. Davidovic, T. (2016). Bee colony optimization part I: The algorithm overview. *Yugoslav Journal of Operations Research, 25*(1).
29. Anter, A. M., Abd Elaziz, M., & Zhang, Z. (2022). Real-time epileptic seizure recognition using Bayesian genetic whale optimizer and adaptive machine learning. *Future Generation Computer Systems, 127*, 426–434.
30. Gupta, M., Wu, H., Arora, S., Gupta, A., Chaudhary, G., & Hua, Q. (2021). Gene mutation classification through text evidence facilitating cancer tumour detection. *Journal of Healthcare Engineering, 2021*. https://doi.org/10.1155/2021/8689873

Applications of Blockchain: A Healthcare Use Case

Priya Shelke, Nilesh P. Sable, Suruchi Dedgaonkar, and Riddhi Mirajkar

1 Introduction

Blockchain is a new emerging technology for the world, which brings the idea of web 3.0. Fundamental features of Blockchain, such as hashing and data encryption, make it much more difficult to bridge the blockchain security environment and corrupt the data. Every individual has a copy of the entire Blockchain in the network, so it builds trust in new technology. The medical system in today's world also requires significant changes to meet the daily needs and inflow of data. Proper monitoring of health care and individuals' health is necessary to take all the required actions to maintain the health index high. Digitization of the entire medical system is needed today, but many problems will arise, such as data storage capacity, infrastructure, privacy and security issues. Blockchain can solve this problem as its security feature will protect all sensitive data. By delivering exceptional data efficiency and upholding trust, Blockchain is a promising technology that can help to improve healthcare data management processes. It provides various notable and built-in characteristics, including decentralized storage, transparency, immutability, authentication, flexibility in data access, interconnection, and security, enabling more extensive use of blockchain technology for managing healthcare data.

Blockchain leverages intelligent contracts, which establish terms and conditions upon which all the healthcare partners engaged in the network agreed, eliminating the need for an intermediary. It cuts back on wasteful administrative expenses. The

P. Shelke · S. Dedgaonkar · R. Mirajkar
Department of Information Technology, Vishwakarma Institute of Information Technology, Pune, Maharashtra, India
e-mail: Priya.shelke@viit.ac.in; suruchi.dedgaonkar@viit.ac.in; riddhi.mirajkar@viit.ac.in

N. P. Sable (✉)
Department of Computer Science & Engineering (Artificial Intelligence), Vishwakarma Institute of Information Technology, Pune, Maharashtra, India

A. M. Anter et al. (eds.), *Nature-Inspired Methods for Smart Healthcare Systems and Medical Data*, https://doi.org/10.1007/978-3-031-45952-8_4

three primary ideas behind Blockchain are peer-to-peer connections, public key cryptography, and trust procedures. Depending on how managing permission is handled, blockchains are categorized as either public, private, or consortium blockchains. Anyone with an Internet connection can participate in the public blockchains‘consensus process. Public blockchains use proof-of-work or proof-of-stake procedures to incorporate incentives and encrypted digit verification [1, 2]. The whole public blockchain network is visible, yet each participant's identity is kept a secret. A single company controls the web in a private blockchain. As a result, for this particular blockchain type to attain an agreement, a reliable agent is needed. Both formal and informal blockchain networks have advantages combined in the consortium blockchain. Only platforms that want to communicate better with one another should use it. Healthcare firms can adopt any form of blockchain network since each has advantages and disadvantages that can be tailored to individual requirements or even use case scenarios.

2 Traditional Healthcare System and Its Limitations

A healthcare system refers to a network of individuals, organizations, and assets that provides healthcare services to a population. It encompasses all parties involved in the delivery of healthcare, such as insurance companies, governmental bodies, non-governmental organizations, hospitals, clinics, doctors, nurses, and other healthcare workers [3–5].

By preventing and treating illness and disease, a healthcare system's principal objective is to preserve and improve population health. Countries' social, economic, and political settings, as well as their cultural values and healthcare demands, all influence how different their healthcare systems are.

Some healthcare systems depend on private insurance or out-of-pocket payments, while some are publicly supported and offer universal access to healthcare services. The effectiveness and quality of healthcare systems can also differ; while some deliver high-quality treatment for a reasonable price, others suffer with lengthy wait times, a lack of resources, and exorbitant healthcare expenditures.

To guarantee that they are able to satisfy the demands of the community they serve, effective healthcare systems need to be carefully planned, coordinated, and funded. Investment in healthcare infrastructure, professional education and training, the creation of evidence-based standards and procedures, and efficient resource management are all part of this.

Big data has profoundly altered how businesses in every sector handle, examine, and use data. Healthcare is one of the most promising industries where big data may be used to effect change. Big healthcare data has great promise for enhancing patient outcomes, forecasting epidemic breakouts, gaining insightful knowledge, avoiding avoidable diseases, lowering healthcare delivery costs, and generally enhancing quality of life. Yet, choosing the appropriate applications for data while upholding security and patient privacy rights is a challenging issue. No matter how important big data is to the development of medical knowledge and how crucial it is to the success of all healthcare organizations, its utilization is contingent upon security and privacy concerns being resolved [6, 7].

2.1 Common Attacks on Current Healthcare Systems

In recent years, cyberattacks have become more often directed at healthcare institutions. Attacks like this put patient data at risk and have the potential to seriously disrupt healthcare services. Cyberattacks on healthcare systems involve unauthorized access to or destruction of sensitive patient data, disruption of critical healthcare services, and theft of intellectual property or financial information. These attacks can have serious consequences for patient safety, as well as the financial and reputational health of healthcare organizations (Fig. 1).

There are various types of cyberattacks that can be carried out on healthcare systems, including:

Ransomware attacks: These involve the use of malicious software to encrypt sensitive data and hold it hostage until a ransom is paid.

Malware attacks: These involve the use of malicious software to gain unauthorized access to systems, steal data, or cause damage to networks.

Phishing attacks: These involve the use of fraudulent emails or websites to trick individuals into revealing sensitive information or downloading malicious software.

Distributed denial-of-service (DDoS) attacks: These involve overwhelming a system with traffic to cause it to crash or become unavailable.

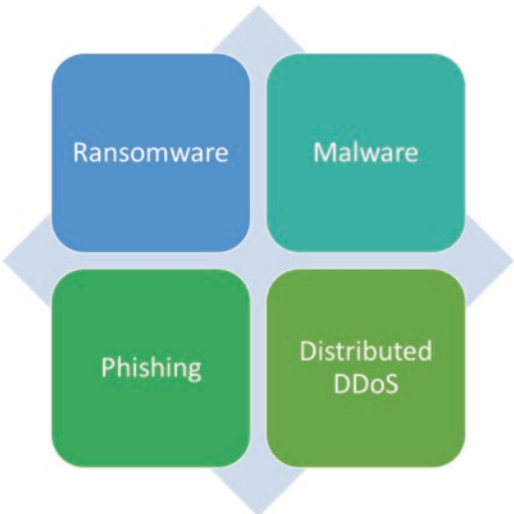


Fig. 1 Types of cyber-attacks on healthcare systems

2.2 Examples of Attacks in the Healthcare Systems

Healthcare systems are particularly vulnerable to cyberattacks due to the sensitive and valuable nature of the data they store, as well as the reliance on technology to deliver critical patient care. As such, it is crucial that healthcare organizations take appropriate cybersecurity measures to protect their systems, including regular security assessments, employee training on best practices, and implementing robust security protocols and technologies. The following recent cases demonstrate the working of the cyberattacks and its effect on the healthcare industry (Fig. 2).

1. An instance of a cyberattack on the healthcare industry happened in November 2020, when the federal cybersecurity agency of the United States issued a warning about a "imminent and heightened cybercrime threat" to healthcare providers. The warning was issued in response to many assaults on the nation's healthcare infrastructure. One of the biggest networks of reproductive clinics in the country, U.S. Fertility, said in a statement that it had been the target of a ransomware attack in September 2020. On September 14, hackers released the ransomware, encrypting systems and blocking access to them. Hackers stated that they had access to "a small number of files" in their statement. These files included personally identifiable information such as names, email addresses, addresses, social security numbers, and other information. Also, it's likely that the hackers had access to patient health records. Hacking organisations continue to target healthcare systems, from huge hospitals to large reproductive clinics. As recently as September and October of 2020, several medical systems were successfully targeted, and in certain cases, everyday operations were stymied. The cybercriminal organisation behind TrickBot, which is likely also the developer of the BazarLoader virus, has continued to produce new tools and features, making victims easier, faster, and more lucrative. As part of their destructive cyber efforts, these threat actors are increasingly using loaders like TrickBot and BazarLoader (or BazarBackdoor). Cybercriminals spread TrickBot and BazarLoader by phishing operations that either include attachments containing the malware or links to malicious websites that host the malware. By dispersing the payload, loaders begin the infection chain. The backdoor is then deployed and executed from the C2 server and installed on the victim's computer.
2. Another noteworthy instance happened in May 2017, when the worldwide WannaCry ransomware assault affected thousands of computers at hospitals and healthcare facilities in over 150 countries and this attack was the most disastrous

Fig. 2 Phases of a ransomware attack

to date. A form of harmful malware called WannaCry Ransomware prevents users from accessing files or systems, encrypting data or whole devices and holding them hostage until the victim pays a ransom in return for a decryption key that enables the user to access the data or encrypted systems.Several healthcare services were adversely affected by this attack, including procedures that were postponed and appointments that were cancelled. These cyberattacks serve as a reminder of the value of cybersecurity safeguards for healthcare systems as well as the necessity for ongoing readiness and attention in the face of new threats. The most severely impacted country by cyberattack was India. Businesses including FedEx, Nissan, Russian Railways, Interior Ministry, megaforTelefonica in Spain, at least 16 NHS organisations in the UK, and German and German train industries were all negatively impacted. Malware was detected on certain computers. Plenty of colleges and students' computers were attacked by assault in china. Only two files, one with instructions on what to do next and the other a decryption application, are left on the victim's computer system after WannaCry locks up all the data. Hackers seek bitcoin as payment. If not, inform the user that the file will be erased. By accessing the file, reading its contents, writing the encrypted contents in place, and then shutting the file, ransomware overwrites the contents of the original file.

3. A major cyberattack on India's biggest public hospital has raised concerns about the country's healthcare system's digitalization as well as the security of the targeted institution. The 3000 bed All-India Institute of Medical Sciences (AIIMS) in New Delhi, one of India's top hospitals, had its computer systems taken offline by a ransomware attack in November 2022.

   An organization cannot access its computer systems unless it pays the hackers a significant sum of money, which is how ransomware works. A wide range of services, including patient registration, online appointments, diagnostic report production, invoicing, as well as administrative systems like salary disbursement and medicine procurement, were rendered inoperable by the AIIMS assault. According to Indian media, the hackers requested a bitcoin ransom. Hospital representatives have declined to comment. Yet the tragedy has caused a lot of introspection and unavoidable concerns about whether a nation that leads the world in information technology has adequate security measures in place for its anticipated digital revolution in healthcare.

   These services had to be handled manually for over a week, which resulted in lengthy lines and longer wait times for customers. The personal information of more than 30 million patients and healthcare professionals, including a number of cabinet ministers and top bureaucrats, may have been compromised, even though online services have now resumed with the data recovered from a backup server. The Indian government has refuted claims in the media that a ransom was asked, and Delhi police are currently looking into the matter.

2.3 Need of Secure System in Healthcare

The need for a secure system in healthcare cannot be overstated as it has a direct impact on patient safety, privacy, and trust. Here are some reasons why a secure system is crucial in healthcare:

- *Patient Safety:* Healthcare data is sensitive and confidential, and if it falls into the wrong hands, it can have serious consequences for patients. A secure system can help protect patient information from unauthorized access, theft, or misuse, thereby safeguarding patient safety.
- *Data Privacy:* Patients have a right to privacy and must be able to trust that their health information will be kept confidential. A secure system can ensure that only authorized individuals have access to patient data and can help prevent breaches that can compromise patient privacy.
- *Compliance:* Healthcare organizations must comply with various regulations such as HIPAA and GDPR, which require the protection of patient data. A secure system can help organizations meet these compliance requirements and avoid penalties for non-compliance.
- *Trust:* A secure system can help build trust between patients and healthcare providers. Patients are more likely to share their health information and seek medical help if they trust that their data is safe and secure.
- *Efficiency:* A secure system can also improve the efficiency of healthcare operations by ensuring that patient data is accurate and up-to-date. This can help healthcare providers make informed decisions, reduce errors, and improve patient outcomes.

Overall, a secure system is critical to ensuring patient safety, protecting data privacy, and building trust in the healthcare system. Healthcare organizations must invest in secure systems and protocols to safeguard patient information and maintain compliance with regulations.

3 Blockchain Technology

Blockchain is a decentralized and distributed digital ledger technology that records transactions and stores data in a secure, transparent, and immutable way. It was invented in 2008 by an unknown person or group using the pseudonym Satoshi Nakamoto as the underlying technology of Bitcoin, the first and most famous cryptocurrency.

Blockchain works by creating a continuously growing chain of blocks that are linked together using cryptographic algorithms, forming a distributed ledger that is maintained by a network of nodes. Each block contains a hash, which is a unique digital fingerprint, of the previous block, along with a timestamp and a record of multiple transactions.

There are three main types of blockchain:

3.1 Public Blockchain

Anyone can participate in the network, read or write transactions, and validate them. Transactions are validated by a decentralized consensus mechanism, such as Proof of Work (PoW) or Proof of Stake (PoS). Examples of public blockchains include Bitcoin, Ethereum, and Litecoin.

3.2 Private Blockchain

Only authorized nodes can participate in the network, read or write transactions, and validate them. Transactions are validated by a centralized consensus mechanism, such as a consortium of trusted validators. Private blockchains are often used for enterprise applications, such as supply chain management, where confidentiality and scalability are important.

3.3 Hybrid Blockchain

It combines elements of both public and private blockchains, allowing authorized nodes to participate in a private network while also allowing public access to selected parts of the blockchain. Hybrid blockchains can provide the benefits of both public and private blockchains, such as transparency, security, and scalability, while also enabling privacy and confidentiality for certain use cases.

4 Contribution of Blockchain Technology in Healthcare

Blockchain technology has the potential to revolutionize the healthcare industry by enhancing security, interoperability, and efficiency. Here are some ways in which blockchain can contribute to the healthcare domain:

4.1 Secure Storage and Sharing of Medical Records

Medical records contain sensitive information that must be kept secure and confidential. Blockchain technology provides a secure and immutable storage solution that can prevent unauthorized access and tampering. Patients can have complete control over their data, and doctors can access only the information they need to provide treatment.

4.2 Improved Interoperability

Healthcare systems often struggle with interoperability, making it difficult for different providers to access and share patient data. Blockchain technology can enable a shared and decentralized database, allowing multiple stakeholders to access and update patient data in real-time. This can enhance collaboration, coordination, and patient outcomes.

4.3 Streamlined Clinical Trials

Clinical trials are critical for the development of new drugs and treatments, but the process can be slow, expensive, and prone to errors. Blockchain technology can create a secure and transparent system for tracking the entire trial process, from patient recruitment to data analysis. This can improve data accuracy, reduce costs, and speed up the time to market for new treatments.

4.4 Supply Chain Management

The pharmaceutical industry is highly regulated and complex, with numerous intermediaries involved in the supply chain. Blockchain technology can provide end-to-end transparency and traceability, enabling manufacturers, distributors, and pharmacists to track and verify the authenticity of drugs, reduce the risk of counterfeiting, and improve drug safety.

4.5 Decentralized Healthcare

Blockchain technology can enable decentralized healthcare systems that are not controlled by a central authority. Patients can have direct access to medical services and providers, and smart contracts can automate payments and ensure fair compensation for providers. This can increase access to healthcare for underserved populations and reduce administrative overhead.

Overall, blockchain technology has the potential to transform the healthcare industry by enhancing security, interoperability, and efficiency. However, it is important to note that widespread adoption of blockchain in healthcare will require collaboration, standardization, and regulatory support.

5 Use Cases of Blockchain in Healthcare

Important use cases for use of blockchain in healthcare are given below:

5.1 *Medical Record Storage (Patient's Medical History Maintenance)*

The World Health Organization views medical knowledge as the most innovative and transportable resource. Medical data is tied to registration, diagnosis, and hospitalization; as a result, privacy and security are more crucial. Additionally, getting increasingly complex and stereochemical is the medical data [8].

By implementing some smart contracts, the medical records are maintained within particular blockchain network nodes.

Registrar Contract (RC)

The identity of the Ethereum address is mapped to member identifying strings using this contract. Both patients and medical professionals are made up of all registered members. The proposed method grants multiple levels of access to each identity [9].

Like,

1. Patients can review their medical data.

Medical Professionals Can Create/Modify Authorized Medical Data

Patient-Provider Relationship Contract (PPR)

A medical certificate is part of any PPR contract. The PPR smart contract tracks medical records, diagnostic information, and access rights to various summary contracts. According to the access rights in the summary contracts, the medical personnel will only be permitted to read or modify after receiving the permits.

Summary Contract

The list of PPRs specified in this contract can be used to find patient medical record history [8]. A categorization framework for patient-focused medical records is used to build the suggested system. Each entry is regarded as a PPR smart contract.

5.2 *The Link Between the Healthcare Industry*

Healthcare is a vast network that includes collaboration from many industries: hospitals, pharmacies, biotech, insurance companies, Health tech, etc. [10] Two industry groupings and six industries can be further distinguished within the healthcare sector:

Medical supplies and services

1. Healthcare equipment and supply companies create everything from the most straightforward products, like bandages and aids, to the most complex ones, like MRI scanners.
2. Companies that own and manage healthcare facilities like hospitals, rehabilitation centres, nursing homes, and rescue organizations are included in the healthcare goods and services providers category.
3. Research and Development in Healthcare Technology include businesses engaged in data analysis to improve existing healthcare practices and research and development (R&D), Biotechnology, pharmaceuticals, and life sciences

1. Companies involved in the production or research and development of goods typically generated from living things are considered to be involved in biotechnology.
2. Organizations that develop and produce medications and immunizations that are often based on chemicals are categorized as pharmaceuticals.
3. Companies that research living things and provide analytical tools, clinical testing and assessment, and full contract research services are included in the life sciences tools and services category.

All industries are interconnected through demand and supply, which makes all of them interdependent. Keeping data consistent and maintaining it is a crucial part of healthcare. For example, tracking equipment from manufacturing to the end of its useful life is necessary. If any dysfunction occurs with the help of the data, we can go down the line and make corrections in the particular chain node in the network. Suppose we create a block of data for each intermediate node, such as manufacturer, dealer, and hospital and link them to each other as a blockchain in such a way. In that case, we can track the entire history of the equipment, which will help to maintain proper coordination between all the industries. Whether it is food provided to the patients, clothes, injections, or medicines, all these patient-related requirements can also be tracked using a unique id. Payment's history and pending amounts by various industries could be maintained, allowing transparency in the trading and supply chain. The significant advantage of using Blockchain to connect interdependent sectors is that it will help to analyse the quality of services and goods. Any defect or unexpected irregularity could be tracked easily and quickly; for example, side effects of particular medicines, we can use the barcode to check the entire history related to treatment, starting from the supply of raw material for manufacturing, transportation, storage, distribution and use can be tracked and investigation of the problem takes less time. Resolution of any dysfunction at any stage of this industry link is possible.

5.3 Logistic Supply Chain

Numerous studies have focused on intra-site patients' movements to enhance their circulation and ensure the security of their stay in care facilities or medical-technical services (Shen et al. 2007). When moving patients between hospitals, the partner network (hospital, laboratory, blood transfusion centre, etc.) must coordinate logistics efficiently because they are expensive, and there is a danger of data loss or medical issues [11]. The study of patient flows also looked at administrative practice's issue with handling and monitoring patients' medical information throughout their treatment. The time it takes nurses to run medical supplies may be cut down if nursing unit storage facilities are decentralised, however, managing these storage sites will require more logistical staff. This outcome emphasises the requirement for managers to incorporate the entire replenishment system and approach inventory management from a comprehensive standpoint. Our endeavour tries to include all the literature on the uses, integration, and implementation of blockchain technology in SCs and logistics, including journal articles, press articles, review papers, and quick surveys [11].

Blockchain is a distributed ledger technology (DLT) used in shared, synchronised environments where users and traceable validation of all transactions. It is immutable, tamper-proof, and uses synchronisation. Every sector relies heavily on supply chain management, but as technologies and methodologies advance, SCM should also be flexible and transparent enough to meet future expectations. Due to its benefits to SCs, such as cost reduction, disintermediation, efficient operations, reduced waste, transparency, authenticity, trust, and security, Blockchain can revolutionise SCM.

Blockchain in supply chain management has several advantages, including calibrating data dispersed across several SCs. The data that is kept is secure. The blockchain network is crucial for information capturing. Blockchain allows for the tracking of an item's progress, which increases process transparency. Task automation, including data processing, is made possible by the network. It delivers total transparency and bases permission levels on a hierarchy.

Increases reaction time Blockchain creates a dynamic, real-time SC that uses its resources better. Blockchain helps to expedite the SC procedure from start to finish. Early detection of faults and issues stops the process from becoming resilient. The management of SC is made possible by using smart contracts, which allow for the customisation and creation of unique agreements that can be specified for each function and coordinated with one another. It aids in the process of designing for operational company needs. The intelligent contract improves visibility and eliminates the requirement for a middleman.

The Blockchain creates a continuous chain of transactions and boosts transaction speed. By offering them quick and straightforward transactions, you may raise the level of trust among process stakeholders. All transactions on the blockchain network have a consensus process, assuring their security. Additionally, Blockchain aids in the development of cooperative intrusion detection systems that allow product codes to communicate and share information along the way [9].

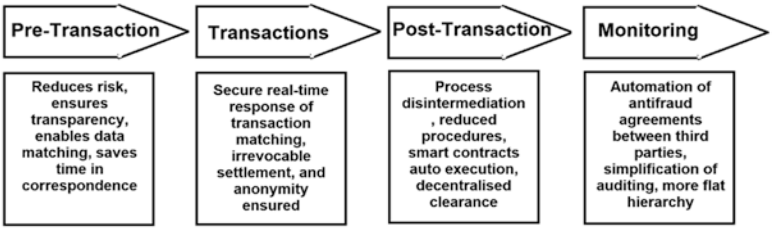


Fig. 3 Supply chain flowchart using Blockchain

As a result, implementing Blockchain in SCs improves stakeholder relations, cuts costs, and increases productivity. Blockchain technology has the potential to alter SCM practises, including reengineering, security, resilience, provenance, process management, and product management. By reducing the impact of disruptions, employing proactive and preventive risk management strategies, and providing multilayer protection for SC networks, Blockchain fosters supply chain resilience. Blockchain offers data and resource provenance, integrity assurance, authentication, secrecy, privacy, and access control in its supply chain management services.

In the healthcare sector, supply chain management mainly involves medications, surgical equipment, infrastructure, etc., all of which require RFID security. Blockchain uses an ultra-lightweight mutual authentication RFID protocol to increase the RFIDs' transparency, data protection, dependability, and cost control. Walmart and IBM used blockchain technology and RFID technologies to track consumer goods (Fig. 3).

5.4 Tracking Diseases & Outbreaks

Blockchain's outstanding capabilities can aid real-time disease reporting and the study of disease patterns that can help identify the disease's origin and transmission factors. Fast outbreaks, high infection rates, and possibly severe sickness are characteristics of infectious illnesses like the flu and SARS. The threats to public health posed by such infectious diseases are becoming more serious. A more recent example, COVID-19, poses a severe risk to people's lives and the ongoing functioning of human society as it quickly spreads over the planet. This global pandemic demonstrates how difficult it is to contain and treat pandemics, given the simple large-scale mobility of people made possible by adequate transportation. Controlling and suppressing infectious diseases requires accurately identifying and tracing the source of illnesses early on. Three crucial elements of disease-direct reporting systems have been the focus of the global COVID-19 pandemic.

1. Traceability
2. Reliability
3. Effectiveness.

Traditional disease reporting systems have a central node of control and a hierarchical structure that begins at the local level (cities and counties), travels up to the regional level (regions), and ultimately reaches the central data repository. These systems are frequently vulnerable to data loss, arbitrary or illegal modifications, and poor node-to-node tracing. Blockchain, a recent disruptive technology, offers a potential answer.

It is required to have data uploaded from numerous nodes, with legitimate, transparent, and tamper-proof data, as well as quick querying and tracking capability at the node level, to identify the source of an infectious disease. Traditional disease reporting systems prioritizing hierarchical upward reporting and management by the superior level find it difficult to implement these aspects. The existing system's hierarchical structure cannot meet the high demands for the rapid and accurate source tracing of illnesses.

Blockchain technology for information storage for infectious diseases will enable information recording and storage throughout an epidemic. In such a disease-direct reporting approach, participants may submit disease-related data or information independently. This procedure is represented in the diagram below. Hospital 1 manages the hospital's information collection on infectious diseases, such as confirmed cases, infected patients, etc. The timestamp of the upload is immediately registered with the time stamp in the Blockchain. The uploaded data would be kept on the Blockchain along with a timestamp and a predetermined hash function. The blockchain agreement permits this, allowing the other blockchain nodes to examine the disease data supplied by Hospital 1 on the Blockchain. The hash function allows any data length to be translated into a fixed-length function value. A hash function can hash the data in uploading disease-related information to the Blockchain and produce a fixed-length hash value. Every piece of data has an irreversible, one-of-a-kind hash value.

The generated hash value will change dramatically when the original data marginally changes or differs. So, to check if data on a blockchain has been changed, all that is needed to do is compute the hash function on the original disease information, compare the result to the hash saved on the Blockchain, and presume the disease data is still the same if the two values are the same (no tampering or forging). To assure data integrity, when the following hospital (Hospital 2) needs to submit a new disease case, the hospital can compare the information that came before to its hash. If Hospital 2 had fresh illness information to add to the Blockchain, the newly loaded data would still be recorded as a hash. The integrity of tracing is guaranteed by this tamper-proof storage approach's tamper-proof disease information storage.

6 Blockchain in Health Care – Case Study or Products

There are several blockchain-based products and initiatives in the healthcare industry. Here are a few examples:

1. Medicalchain: Medicalchain is a platform that uses blockchain technology to store and share electronic health records (EHRs) securely. The platform allows patients to control who has access to their records and enables healthcare providers to share patient data across different systems securely.
2. ProCredEx: ProCredEx is a blockchain-based platform that verifies healthcare provider credentials. The platform uses blockchain technology to create an immutable record of provider credentials, which can be shared with healthcare organizations to streamline the credentialing process.
3. SimplyVital Health: SimplyVital Health is a blockchain-based platform that enables healthcare providers to share patient data securely. The platform uses blockchain technology to create a shared network of patient data that can be accessed by authorized providers, which can help improve patient outcomes and reduce costs.
4. Guardtime: Guardtime is a blockchain-based platform that provides secure supply chain management for healthcare products. The platform uses blockchain technology to create an immutable record of product transactions, which can help reduce fraud and ensure the authenticity of healthcare products.
5. Nebula Genomics: Nebula Genomics is a blockchain-based platform that enables patients to share their genomic data securely. The platform uses blockchain technology to create a decentralized network of genomic data that can be accessed by authorized individuals and organizations, which can help accelerate medical research.

These are just a few examples of the many blockchain-based products and initiatives in the healthcare industry. As blockchain technology continues to evolve, we can expect to see more innovative solutions that address the unique challenges of the healthcare industry (Fig. 4).

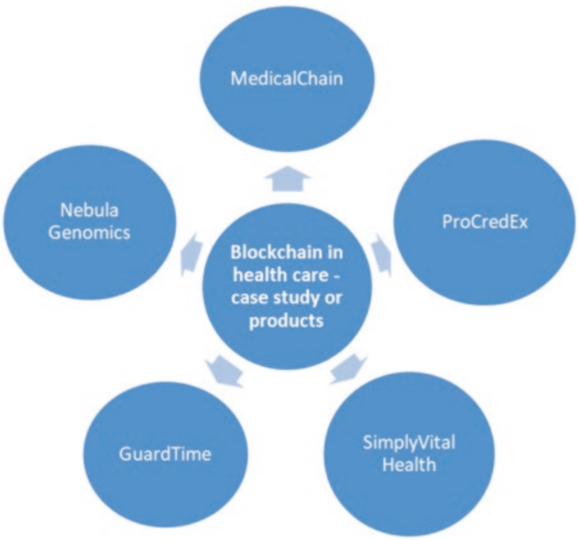


Fig. 4 Blockchain based products in healthcare

6.1 Medical Chain

MedicalChain is a blockchain-based platform that aims to provide a secure and efficient way of managing and sharing medical data. It uses blockchain technology to create a tamper-proof record of medical data that is accessible to authorized parties while protecting patient privacy. Here are some of the key features and benefits of MedicalChain:

Patient-controlled medical records: Patients have control over their medical records and can choose who has access to them. This ensures patient privacy and helps to prevent data breaches.

Secure data storage: Medical data is stored on a decentralized blockchain network, making it difficult for hackers to access or manipulate the data.

Improved data sharing: MedicalChain allows healthcare providers to securely share patient data with other healthcare providers, regardless of their location or the electronic health record (EHR) system they use.

Faster and more accurate diagnosis: MedicalChain allows doctors to access a patient's medical history quickly and accurately, improving the speed and accuracy of diagnosis.

Easy payment and reimbursement: MedicalChain allows patients to pay for medical services using cryptocurrency, making it easier and faster to make payments and receive reimbursements.

Overall, MedicalChain aims to improve the efficiency and security of healthcare data management, while also putting patients in control of their own medical records.

6.2 ProCredEx

ProCredEx is a blockchain-based platform that enables healthcare organizations to securely and efficiently exchange credentials and other sensitive data. The platform uses distributed ledger technology to provide a secure and tamper-proof record of healthcare provider credentials, such as medical licenses, certifications, and training, which can be easily accessed and verified by authorized parties.

ProCredEx aims to solve the longstanding challenge of verifying healthcare provider credentials, which is a critical component of ensuring patient safety and quality care. The platform offers a streamlined solution that replaces the traditional manual process of credential verification, which can be time-consuming, inefficient, and prone to errors.

ProCredEx also offers enhanced security and privacy features, such as data encryption and user authentication, to protect sensitive information and ensure that only authorized parties can access it. By leveraging blockchain technology, ProCredEx provides a trusted and transparent platform that can help improve the overall efficiency and effectiveness of healthcare credentialing processes.

How ProCredEx is Leveraging Blockchain to Speed up Credentialing

ProCredEx is leveraging blockchain technology to speed up the credentialing process by creating a decentralized, secure, and tamper-proof ledger that can be accessed by authorized parties in real-time. Here are some ways in which ProCredEx is using blockchain to speed up credentialing:

Trust and transparency: ProCredEx uses a distributed ledger to provide a secure and tamper-proof record of healthcare provider credentials. This means that healthcare organizations can trust the validity of the data, and that they can easily verify the authenticity of credentials in real-time.

Automated verification: ProCredEx uses smart contracts to automate the verification of credentials. Smart contracts are self-executing contracts that use code to automatically validate and enforce the terms of an agreement. This means that healthcare organizations can quickly and easily verify the validity of credentials without the need for manual intervention.

Data sharing: ProCredEx allows healthcare organizations to securely share credential data with authorized parties in real-time. This means that healthcare organizations can quickly and easily exchange credential data, which can help to speed up the credentialing process and improve patient care.

Reduced administrative burden: ProCredEx streamlines the credentialing process by reducing the administrative burden on healthcare organizations. By automating the verification process and providing a secure and efficient platform for data exchange, ProCredEx can help healthcare organizations to reduce the time and resources required for credentialing.

Overall, ProCredEx is using blockchain technology to create a more efficient, secure, and transparent platform for healthcare credentialing. By leveraging the benefits of blockchain, ProCredEx aims to speed up the credentialing process, reduce administrative burden, and improve patient care.

Four Ways ProCredEx Transforms Data Into Insights

Distributed Ledger Technology: ProCredEx uses distributed ledger technology (DLT) to create data provenance, make it immutable, and provide permanent traceability.

As a result, ProCredEx members can share valid credentials data in a secure and trusted environment. ProCredEx studied numerous blockchain and DLT options before settling on r3's Corda protocol, which provides the security and industrial scalability required to support healthcare credentialing transactions.

Validation Engine: Your organization's credentialing practises are governed by a plethora of laws, rules, guidelines, by-laws, standards, and preferences.

ProCredEx uses its validation engine to capture and apply these rules, bringing accuracy and efficiency to the analysis, curation, and presentation of credentials data.

ProCredEx has developed a distributed ledger of healthcare credentials data that improves the productivity of complicated datasets by making them irreversible and forever traceable. It enables data to be selected to match the needs of certain organisations and shared with approved partners.

The platform employs proprietary validation engines and limits membership to vetted and recognised institutions, allowing health systems to quickly obtain verified credentials while also promoting patient safety and care quality.

6.3 SimplyVital Health

SimplyVital Health is a healthcare technology company that is focused on using blockchain technology to improve patient care and reduce healthcare costs. The company was founded in 2015 and is based in Connecticut, USA.

One of the key products of SimplyVital Health is its blockchain-enabled platform called Health Nexus. The platform is designed to securely store and share patient data between healthcare providers, with the aim of improving care coordination and reducing the administrative burden of healthcare organizations.

The Health Nexus platform uses blockchain technology to create a decentralized, tamper-proof ledger of patient data. This ensures that the data is secure and immutable, and that it can be easily accessed by authorized parties in real-time. The platform also uses smart contracts to automate the exchange of data and to ensure that data is shared only with authorized parties.

Another key feature of Health Nexus is its ability to facilitate value-based care models. The platform allows healthcare organizations to track patient outcomes and to incentivize providers based on the quality of care they deliver. This approach aims to improve patient outcomes while also reducing costs by incentivizing providers to deliver high-quality care.

SimplyVital Health also offers a range of other products and services, including a blockchain-based healthcare billing system called VSee, and a care coordination platform called NexusPay. These products and services are designed to streamline administrative processes, improve care coordination, and reduce costs across the healthcare industry.

Overall, SimplyVital Health is a healthcare technology company that is leveraging blockchain technology to improve patient care and reduce healthcare costs. Its Health Nexus platform is designed to securely store and share patient data, while also facilitating value-based care models that incentivize high-quality care.

SimplyVital Health is developing a SaaS based platform to provide a way to send data to the Blockchain with retrieve and display functions for patients and providers. The company claims that the platform and anlaytics suite can amplify data exchanges and data repositories for better care coordination. The platform will be able to query the Blockchain, providing access to healthcare data. The company's goal is to grow the healthcare network that leverages blockchain technology, our focus is to bring providers, patients, and healthcare technology companies together on the uniform backend that is the blockchain.

SimplyVital Health How it Works

SimplyVital Health's flagship product, Health Nexus, is a blockchain-enabled platform designed to securely store and share patient data between healthcare providers. Here's a step-by-step breakdown of how it works:

Data Entry: Healthcare providers can enter patient data, including medical history, test results, and other health information, into the Health Nexus platform.

Data Storage: Health Nexus uses blockchain technology to securely store this data in a decentralized, tamper-proof ledger. This ensures that the data is immutable and can only be accessed by authorized parties.

Data Sharing: Providers can share patient data with other authorized providers on the network, allowing for seamless care coordination across different healthcare organizations. Health Nexus uses smart contracts to automate the exchange of data and to ensure that data is shared only with authorized parties.

Analytics: Health Nexus also includes a range of analytics tools that allow providers to track patient outcomes and identify areas for improvement. This data can be used to inform treatment decisions, optimize care pathways, and improve patient outcomes.

Incentivization: Health Nexus supports value-based care models that incentivize providers based on the quality of care they deliver. Providers are rewarded based on patient outcomes, which encourages them to deliver high-quality care while also reducing costs.

Overall, Health Nexus is designed to create a more secure, efficient, and collaborative healthcare ecosystem. By leveraging blockchain technology, the platform enables healthcare providers to securely store and share patient data, while also incentivizing high-quality care and improving patient outcomes.

6.4 Guardtime

Guardtime is a technology company that specializes in developing digital verification and data integrity solutions using blockchain technology. The company was founded in Estonia in 2007 and has since expanded globally with offices in several countries.

Guardtime's core technology is called Keyless Signature Infrastructure (KSI), which uses blockchain technology to create a verifiable digital signature that can be used to prove the integrity of data. KSI creates a digital fingerprint of data that is stored in a distributed network of nodes, making it virtually impossible to tamper with the data without detection.

Guardtime's solutions are used in a variety of industries, including healthcare, financial services, and government. For example, Guardtime's KSI technology is used to ensure the integrity of patient health records in Estonia, and it has also been used to secure the data and infrastructure of the Estonian government.

Overall, Guardtime's technology aims to provide a high level of security and trust in digital data, helping organizations to protect their valuable information from tampering and fraud.

GuardTime is a blockchain-based digital timestamping technology that provides proof of integrity and time-stamping for data. It uses a mathematically verifiable method to create a digital signature that proves the existence, integrity, and authenticity of a piece of data at a specific point in time.

Here's how GuardTime works:

Data is hashed: The data to be time-stamped is first hashed, which means it is transformed into a unique and fixed-length string of characters. This hash serves as a unique identifier for the data.

Hash chain is created: A hash chain is created by linking the hash of the data with the hash of the previous piece of data. This creates a chain of hashes that are mathematically linked to each other.

Blockchain is created: The hash chain is added to a blockchain, which is a distributed ledger that is maintained by a network of computers. The blockchain serves as a tamper-proof record of the hash chain and ensures that the data cannot be altered without detection.

Time-stamp is created: The current time is added to the last hash in the chain, creating a time-stamp that proves the existence and authenticity of the data at that specific point in time.

Verification: The authenticity and integrity of the data can be verified at any time by generating a new hash of the data and comparing it to the hash in the blockchain. If the hashes match, it means that the data has not been altered since it was time-stamped.

GuardTime's technology provides an immutable record of data that is resistant to tampering and manipulation, making it ideal for use cases such as financial transactions, healthcare data, legal documents, and other sensitive data.

6.5 Nebula Genomics

Nebula Genomics is using distributed ledger technology to eliminate unnecessary spending and middlemen in the genetic studying process. Pharmaceutical and biotech companies spend billions of dollars each year acquiring genetic data from third parties. Nebula Genomics is helping to build a giant genetic database and incentivize users to safely sell their encrypted genetic data.

Nebula Genomics is a genomics and biotechnology company that is focused on leveraging blockchain technology to enable more secure and efficient sharing of genomic data. The company was founded in 2017 by Harvard genetics professor George Church, and its mission is to empower individuals with greater control over their genomic data and to accelerate scientific discoveries in genomics and personalized medicine.

Nebula Genomics uses a decentralized blockchain network to securely store and manage genomic data. The platform allows individuals to securely and privately store their genomic data, and to control who has access to that data. By using blockchain technology, Nebula Genomics ensures that the data is tamper-proof, transparent, and easily accessible by authorized parties.

One of the key features of Nebula Genomics is its ability to facilitate the sharing of genomic data between individuals and researchers. The platform uses a token-based system to incentivize individuals to share their genomic data with researchers in exchange for compensation in the form of cryptocurrency. This approach aims to create a more open and collaborative ecosystem for genomics research, while also providing greater transparency and control for individuals over their own data.

Nebula Genomics also offers a range of other services, including whole genome sequencing, data analysis, and personalized health recommendations. The platform uses advanced machine learning algorithms to analyse genomic data and provide personalized health insights and recommendations to individuals.

Overall, Nebula Genomics is a cutting-edge biotechnology company that is leveraging blockchain technology to create a more secure and efficient platform for genomic data sharing. Its mission is to empower individuals with greater control over their genomic data, and to accelerate scientific discoveries in genomics and personalized medicine.

7 Conclusion and Future Trends

7.1 Conclusion

The ability to modify the current intelligent healthcare systems from highly centralised to distributed, secure, decentralised systems that can aid in improving healthcare and other applicable services is provided by blockchain technology when blended with other modern technologies. First, it helps ensure patient privacy while providing transparent data to all stakeholders. Malicious attackers can no longer alter vital medical records and are protected from data theft and surveillance.

By using blockchain in medical record maintenance, we have overcome the drawbacks and provided them with more secure storage of data and efficient tracking of the patient's record. The link blockchain makes in the healthcare industry efficiently coordinates various industries under healthcare and provides transparency.

The blockchain helps the supply chain management to track its commodity, maintain the stock in hospitals, and manage the transparency between the vendor and the management. Moreover, blockchain affects explicitly the supply chain to monitor fraud and maintain an anti-fraud system. Blockchain helps with disease tracking as it can maintain the dataset which would be used to identify the growth rate of a particular disease in a particular region and could be transparent to all members.

7.2 Future Trends

Researchers must create additional proofs-of-concept and prototypes because Blockchain technology is currently in its youth in healthcare. It will better understand the technology as it relates to healthcare.

According to current trends in blockchain research in healthcare, other uses of the technology, such as managing supply chains or managing prescription medicine orders, are much less common than data exchange, access control, and health records. Therefore, Blockchain still has a lot of untapped potentials. Most research uses blockchain technology in healthcare and proposes a new framework, design, or paradigm. Additionally, technical information, such as the blockchain platform, consensus method, type of Blockchain, or use of smart contracts, is frequently omitted. In particular, smart contracts might be employed more because they enable processes to be managed on a blockchain. Blockchain technology in medical care is still in its development. Thus, there is still room for new applications and study. In conclusion, Blockchain should continue to be utilised when it makes sense and is required.

References

1. Mamun, Q. (2022). Blockchain technology in the future of healthcare. *Smart Health, 23*, 100223. Available at: https://doi.org/10.1016/j.smhl.2021.100223
2. Sable, N. P., Powar, S. R., Fernandes, Q., Gade, N. A., & Shingade, A. B. (2022). Pragmatic approach for online document verification using block-chain technology. *ITM Web of Conferences, 44*, 03001. https://doi.org/10.1051/itmconf/20224403001
3. Anter, A. M., & Abualigah, L. (2023). Deep federated machine learning-based optimization methods for liver tumor diagnosis: A review. *Archives of Computational Methods in Engineering, 30*(5), 3359–3378.
4. Gudadhe, S., Thakare, A., & Anter, A. M. (2023). A novel machine learning-based feature extraction method for classifying intracranial hemorrhage computed tomography images. *Healthcare Analytics, 3*, 100196.
5. Azar, A. T., Anter, A. M., & Fouad, K. M. (2020). Intelligent system for feature selection based on rough set and chaotic binary grey wolf optimisation. *International Journal of Computer Applications in Technology, 63*(1–2), 4–24.
6. Fan, K., Wang, S., Ren, Y., et al. (2018). Med block: Efficient and secure medical data sharing via Blockchain. *Journal of Medical Systems, 42*, 136. https://doi.org/10.1007/s10916-018-0993-7
7. Du, M., et al. (2021). An optimised consortium blockchain for medical information sharing. *IEEE Transactions on Engineering Management, 68*(6), 1677–1689. Available at: https://doi.org/10.1109/tem.2020.2966832

8. Daraghmi, E.-Y., Daraghmi, Y.-A., & Yuan, S.-M. (2019). MedChain: A design of a blockchain-based system for medical records access and permissions management. *IEEE Access, 7*, 164595–164613. Available at: https://doi.org/10.1109/access.2019.2952942
9. Yaqoob, I., Salah, K., Jayaraman, R., et al. (2022). Blockchain for healthcare data management: Opportunities, challenges, and future recommendations. *Neural Computing and Applications, 34*, 11475–11490.
10. Ratta, P., et al. (2021). Application of blockchain and internet of things in healthcare and medical sector: Applications, challenges, and future perspectives. *Journal of Food Quality, 2021*, 1–20. Available at: https://doi.org/10.1155/2021/7608296
11. Saha, A., et al. (2019). Review on Blockchain technology-based medical healthcare system with privacy issues. *Security and Privacy, 2*(5). Available at: https://doi.org/10.1002/spy2.83

Comprehensive Methodology of Contact Tracing Techniques to Reduce Pandemic Infectious Diseases Spread

Mohammed Abdalla and Ahmed M. Anter

1 Introduction

In December 2019, the Chinese city of Wuhan faced an extraordinary illness outbreak. Coronavirus sickness became known to the rest of the globe in 2019. (COVID-19). The rapid global spread of this infectious disease affected millions of people globally. In response to this unusual virus threat, the World Health Organization (WHO) issued a worldwide emergency alert. According to the WHO, over 500,000 people had died, and over 10 million more were still fighting for their lives. This virus is very contagious and easily spreads from person to person. As a result, authorities around the world enacted lockdown measures for months to combat the spread of the coronavirus. As a result, these measurements have had a significant impact on social and economic activities worldwide [1, 33, 42].

As a result, governments began to seek solutions to mitigate the socioeconomic disasters, and lockdowns were gradually eased. The hunt for reliable technologies to track the mobility of patients who may be infected with COVID-19 and have had contact with a virus-infected individual continues. Because people who are in close proximity to someone who is sick are more likely to develop the illness and possibly pass it to others, properly monitoring these relationships can help stop the spread of the virus. This monitoring approach is referred to as contact tracing.

The contact tracing procedure is a monitoring and mitigation tool used to detect infectious diseases and stop their spread as quickly as possible. The fundamental

M. Abdalla
Faculty of Computers and Artificial Intelligence, Beni-Suef University, Beni Suef, Egypt
e-mail: mohammed.a.youssif@fcis.bsu.edu.eg

A. M. Anter (✉)
Faculty of Computers and Artificial Intelligence, Beni-Suef University, Beni Suef, Egypt

Egypt-Japan University of Science and Technology (E-JUST), Alexandria, Egypt
e-mail: Ahmed_Anter@fcis.bsu.edu.eg

A. M. Anter et al. (eds.), *Nature-Inspired Methods for Smart Healthcare Systems and Medical Data*, https://doi.org/10.1007/978-3-031-45952-8_5

purpose of contact tracing is to prevent the spread of infectious illnesses by first checking people who are already affected, then locating and analyzing each person's contacts. Because contacts are defined as everyone who had direct physical touch with the infected, all subsequent-generation suspected cases will be managed. A number of applications have been suggested and developed to aid with this process. The bulk of these apps employ smartphone technologies to track all of your friends' travels and notify those who are at high or low risk of infection. However, a number of difficulties limit the utility of contact tracking procedures. These challenges include (1) privacy concerns, (2) the difficulty to fully identify contacts, and (3) identification delays.

We study the abilities, methodologies, problems, applications, and use cases of the contact tracing process in this paper. Based on the results of the contact tracing process, the study assesses a number of models that have been utilized and served in this area. This paper's primary contribution is the presentation of a comprehensive view of the contact tracing process as a whole, which will aid researchers in understanding its functions, identifying challenges and unresolved issues, and inspiring them to develop new models, applications, and methods in this area. The study's materials were carefully chosen to highlight the importance of contact tracking in limiting the spread of epidemics. The key search criteria have been clearly established in order to identify publications focusing on contact tracing definition, contact tracing applications in epidemics, contact tracing use in statistical models that forecast epidemic transmission, and contact tracing technologies. Covide-19, epidemic models, contact tracing, and infectious diseases are examples of these terms. The research publications were published between 2003 and 2020. Credible publishers such as IEEE, Springer, BMC, and online medical research journals have also been targeted (JMIR).

1.1 Contact Tracing Process Definition

This section investigates various definitions of the contact tracing process and highlights the various functionalities that must be present in this process.

Definition 1: Authors describe COVID-19 contact tracing as a procedure that supervises and regulates the prevention of the disease's future spread by carefully identifying and analyzing people who had recent direct physical contact with COVID-19 patients [56].

Definition 2: The contact tracing method, according to the authors, involves procedures that ensure and validate monitoring and infection-control measures [13].

Definition 3: According to the authors, the process of tracking and documenting the progression of an infectious disease begins with an investigation into cases that have been diagnosed and whose infections have been confirmed, and then moves on to all other people who have recently had direct contact with the patients [17], [4].

Definition 4: Contact tracing, according to the authors, is a critical step in combating the development of infectious diseases such as the COVID-19 epidemic [32].

Definition 5: Contact tracing, according to the authors, is the practice of keeping track of persons who have recently experienced a bodily reaction to infected individuals. The authors also considered the contact tracing technique as a mitigation strategy to reduce the spread of infectious diseases by quickly discovering, assessing, and preparing for the next wave of cases [61].

Definition 6: The authors define the contact tracing procedure as a strong disease control strategy that minimizes and prevents illness transmission from one person to another. In this regard, this technique is effective since it monitors reported and suspected infected cases before isolating them. In addition, whether infected or suspicious, all individuals who had direct physical contact with these cases must be tracked [26].

Figure 1 depicts the contact tracing approach. Every person who contacts with a sick patient is tracked by this technique. The process then splits these contacts into two categories: infected contacts and non-infected contacts. The infected patients are then classified into two groups: (1) high-risk contacts, or those who are at a high risk of dying and require particular treatment, and (2) low-risk contacts.

Roadmap

The remainder of the paper is structured as follows. Common definitions of the contact tracing method are presented in Sect. 2. The importance of contact tracing and its main applications are shown in Sect. 3. The novel avenues of contact tracing and the main challenges these directions face are discussed in Sect. 5. The common models that have been used to simulate epidemic outbreaks are described in Sect. 6, along with the role that contact tracing plays in halting the spread of epidemics.

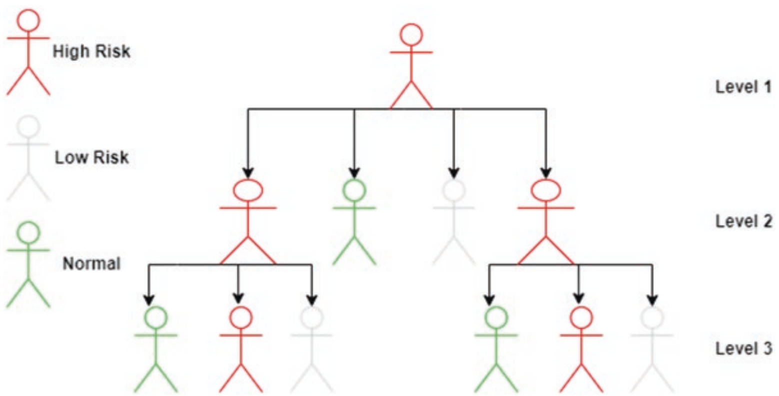


Fig. 1 Contact tracing process

2 Relevance and Concept

This section discusses novel applications for contact tracking as well as its importance in limiting epidemics.

The authors investigate a range of contact tracing strategies, such as pairwise-approximation and entirely random simulation methods, with the purpose of addressing and investigating infected and suspicious situations. As a result, the first step in using these methodologies is to examine the contacts that are associated with infected and suspected cases. This study emphasizes the need to comprehend each individual's disease-transmission routes in detail and the proper modelling of contact tracing methods based on that information. By warning and handling sick and suspected cases by either rapid treatment or isolation, the final findings show that there is a substantial association between the accuracy of contact tracing modelling and the disease reproduction ratio [2, 18, 29].

The authors focus on and discuss the roles and best practices for community surveillance in order to limit and stop the development of the COVID-19 disease. The authors also look into the use of quarantine and contact tracing for COVID-19. COVID-19 illness typically transmits from one person to another through coughing, inhaling, or nasal or oral spray from COVID-19 infected individuals. The writers talked on the importance and criticality of contact for conducting the quarantine competently and promptly. The authors additionally separated persons who have contact with a COVID-19-infected person into two categories: (1) high-risk contacts and (2) low-risk contacts. High-risk contacts are those who have shared close quarters with or had direct physical contact with a COVID-19-infected person. Less severe cases involve those who commute with the patient by train, bus, or other method of transportation or who share a space with Covid-19-infected individuals. The practice of quarantine, which restricts the movement of individuals who have previously interacted with COVID-19 patients, is also critical, as the authors note, in order to stop and restrict the spread of the illness. The World Health Organization (WHO) has established that a 28-day isolation period is necessary for COVID-19. The next section by the authors dealt with issues that frequently restrict contact tracking. These challenges include the following: (1) the challenge of tracking contacts who have travelled by bus, train, or other means of transportation; (2) the challenge of tracking and remembering all contacts who have interacted with the COVID-19 patient in the previous 14 days; (3) the challenge of tracking and locating all people who directly interact with the COVID-19 patient, particularly if the patient attended large gatherings like weddings, religious ceremonies, or markets; and (4) the shared fear between the hysterical and the afflicted [56].

After researching the Ebola virus, the authors of [16] reached to the conclusion that bodily fluids from victims, such as bodily fluids, are one way the infection spreads. The authors of "cite"ref21 also introduced a study that investigates the Ebola virus contact tracing procedure and intends to follow those who were exposed to infected patients for 21 days. To achieve this, the researchers in citeref19 construct and suggest an Ebola transmission model based on contact tracing. This model looks at the ideal window of time for touch-tracing operations. The authors also categorize hospitalization delays, contact identification delays, and contact initiation delays into three groups. The data suggest that early contact identification and hospitalization could result in a 50% reduction in the propagation of the epidemic as compared to delayed recognition. The authors arrived at the conclusion that it is critical to start the contact tracing process as soon as possible in this regard.

The authors of [12, 36] provide insight into contact tracing practices used on social networking sites (SNS). The authors further urge the use of SNS for disease tracking and management by specialized agencies and government-run health sectors. The authors emphasize on utilizing SNS because it is thought to be a key source of information. Additionally, the health sectors will have fantastic opportunity to conduct contact tracing in an appropriate manner as these websites gain popularity and develop for a number of useful mobile devices, such as smartphones and tablets.

The network of people who were in close proximity to one another during the outbreak is carefully analyzed as the main element of the contact tracing method. To find out more about these physical interactions, which will be explored, several projects have been established. For instance, in [60], writers examine extensive Facebook data on individual interactions to better understand and demonstrate the dynamics of sickness in the community. Additionally, [45] authors capture in-person encounters using wireless sensors. According to the writers in [34], the majority of individuals may not have access to or need more resources than those used in these attempts. Because they think a mobile phone is a useful tool for obtaining communication traces for contacts with proximity interactions, the authors develop a model for contact tracing based on one. The scientists claim that efficient contact tracing can greatly lessen the consequences of the Epidemic outbreak. Additionally, because there are a lot less contacts that need to be found, contact tracing is more successful early in the epidemic than it is later, after the Epidemic evolution. In their conclusion, the authors claimed that Epidemic a cell phone to track connections could be a useful strategy for halting the spread of the Epidemic virus.

In [41], the staged-progression model and the differential infection model (DI), which both take contact monitoring and stochastic screening into account, are the two models that the authors aim to utilize to stop the spread of HIV. (SP). Overall, the SP model emphasizes the time variations for the same infected person more than the DI model, which emphasizes individual differences throughout disease transmission. The authors describe contact tracking and stochastic screening as interventions options for stopping the spread of illness. The authors of this study develop a number of formulas to estimate reproduction rates and pinpoint equilibrium points during the epidemic‘s emergence. The outcomes of the stochastic screening and contact tracing are then contrasted for the two models. The findings demonstrate how effective the contact tracing methods used in the DI model are in identifying and containing disease super spreaders, which can help to lessen the overall scope of the Epidemic. This information agrees with that found in the studies cited in [6, 27, 40, 46]. The authors stress that stochastic screening in the SP model is more effective and has a significant impact compared to contact tracing, which is ineffectual at stopping and delaying the spread of the disease. The authors' ultimate finding is that the fundamental reasons why diseases spread can have a big impact on how well interventions work.

The authors of [3, 64] argue that because the Ebola virus has an 11-day incubation period, many people may come into close contact with an infected person. As a result, the disease will spread quickly. As a result, the authors of [26] present a model that takes contact tracing into account for predicting and tracking Ebola infection rates. The intended model takes into account a number of control scenarios in order to precisely analyze the characteristics of the Ebola disease and study how these traits affect the efficiency of the contact tracing technique in containing the Ebola epidemic. These characteristics include the following: (1) Ebola sickness characteristics; (2) methods of surveillance for confirmed and suspected cases; and (3) reporting deadlines for infected cases; (4) epidemic incubation period; and (5) deadlines for managing reported cases. Additionally, to obtain a reproduction number (R_e) less than one, the authors' approach highlights the crucial cases from contacts that were located. The equation reads as follows: $R_e = k\frac{1-q}{q} + k_m$, where k is wnumber of untraced contacts discovered in the first phase, km is the calculated number of infected contacts based on the traced contacts discovered in the first phase's identification, and $\frac{1-q}{q}$ is the likelihood that the stated contact will not be located. To conclude, R_e production number is the ratio of observed reported instances to identified contacts (observed cases *divided by* the total number of contacts traced).

3 Applications and Domains of Usage

This section discusses novel applications for contact tracking as well as its importance in limiting epidemics.

To build a viable contact tracing program that can detect connections of infected people, the system should be loaded onto mobile devices. Because the majority of people own cellphones, a mobile app tailored to each individual will allow law enforcement to rapidly identify contacts.

Figure 2 depicts a functional framework that makes use of an app loaded on users' mobile devices. If two mobile applications are close to one another, they can quickly connect. This could expose the relationships between each people. The application sends these data to the server for storage. People who have been diagnosed with the Covid-19 infection will be warned about their contacts and advised to keep them to themselves while seeking medical assistance. The system also provides information about the spread of Covid-19 to decision-makers so that they can determine whether to adopt quarantine, isolation, or lock-down in specific regions identified by the application.

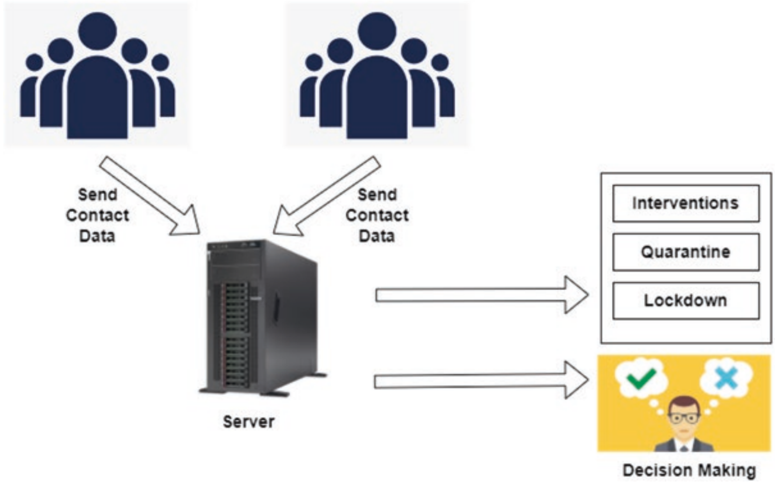


Fig. 2 A depiction of contact tracing framework based on real-time monitoring of individuals

The Singaporean government unveiled Trace-Together on March 20, 2020, an application that uses Bluetooth to track and trace contacts. When two application users are nearby and one of them specifies that he is a COVID-19 patient, the Trace Together program begins to notify the Ministry of Health in order to identify the network of contacts logged to be nearby. Furthermore, the program talks with such contacts and reports them, along with the necessary follow-up steps [10].

The authors create a cutting-edge peer-to-peer smartphone application that successfully completes the contact tracking procedure by utilizing the capabilities of smartphone application technology. The authors address the importance of contact tracing, which is a critical step in avoiding the spread of the COVID-19 pandemic. The primary goal of this research is to create a smartphone application that does contact tracing while protecting user privacy by preventing the acquisition of any personal information or location data. In fact, the method used in this study to perform contact tracing while maintaining privacy is based primarily on a cutting-edge data structure known as a transmission graph, which is composed of nodes and edges and represents the transmission vectors between contact points on a node level and contacts on an edge. Each contact point, which documents the physical interaction between two or more people at a specific time and location, is clearly represented by a node. The transmission graph not only promotes social interaction but also protects personal information [37].

In [17, 23], the authors discuss numerous ways for conducting the contact tracking process. First-order tracing, single-step contact tracing, iterative contact tracing, and retroactive contact tracing are some of these methods. People who had direct physical contact with the sick are identified through the first order tracing process; they must be isolated and need particular medical attention. Furthermore, after the first order has been successfully fulfilled, contact tracing is performed throughout the second order tracing process. The single-step contact tracing method discovers all of the verified ill person's contacts, classifies each of these contacts as infectious, and then identifies each of these contacts' contacts, repeating the process for each of these contacts. The Iterative Contact Tracing technique attempts to discover the virus before it manifests itself by performing diagnostic tests on each individual iteratively. One of these tests is the symptom screening. The final type is a retrospective contact tracing method, which follows the same processes as single-step or iterative methods but also looks backward into the sick individual's recent history to identify everyone with whom the patient comes into touch. The goal of the retrospective contact tracing technique is to locate the person who first contracted the illness.

COVID-Safe, a piece of software, has been made available by the Australian government. This application tries to communicate with the predicted persons who are exposed to the illness in order to limit the spread of the COVID-19 outbreak. The application informs and notifies these individuals to take the necessary emergency safeguards [48, 70].

According to studies [13, 37, 43], the following factors must be considered when evaluating the effectiveness of contact tracing techniques: Contact identification is followed by contact notification, contact identification, automatic contact tracking of movements and direct physical communications, the use of mobile application technology, and narrowcast messaging.

Between 2014 and 2016, the African countries of Liberia, Guinea, and Sierra Leone faced an Ebola virus disease (EVD) outbreak. The research addressed in [51] reports over 11,000 fatalities and 28,000 reported infected patients. Authors in [58] created the Ebola Contact Tracing application (ECT app), a smartphone software, to track and identify Ebola-infected persons in Sierra Leone. This programme is linked to an alert system, which allows it to notify the response centre about symptomatic persons and the districts in which they live.

The Liberian government uses contact tracking to stop the spread of Ebola in this study, published in [54], (2014–2015). They measure the method's efficacy using official county reports on the number of cases. The authors created a model that distinguished between positive predictive value (PPV), which represented persons who had been found, and potential cases, which represented people who had previously experienced symptoms, recovered from them, or died. It bases its findings on 25,830 contact tracing data of probable cases collected between June 4, 2014 and July 13, 2015. These instances account for 26.7% of all EVD cases. Contact tracing demonstrated its ability to locate 3.6% of new cases over its operational period.

The writers of [22] look into the SARS outbreak in Hong Kong in 2003. The authors utilize contact tracing to index cases based on surveys who have visited hospitals. These surveys are conducted by nurses in order to locate the contacts of index cases and counsel them to separate themselves and seek medical attention. The surveys were also utilized to determine the routes of transmission to index cases. The authorities can categories index cases based on a variety of characteristics, including demographic (housing and employment areas), age (friends and schools), and workplaces, by reviewing survey data. These figures can be used by authorities to pinpoint epidemic hotspots, isolate them, and limit disease propagation.

The authors present a case study analysis to explain how the SARS Coronavirus 2 (SARS-CoV-2) propagated in Wuhan, China. The research examines a dataset of patients from the Shenzhen area to determine the effectiveness of using contact monitoring to limit the development of the epidemic. The research focuses on the use of contact tracing to identify infectious sources and prevent future spread. The Shenzhen Centre for Disease Control and Prevention (CDC) undertakes contact tracing to identify affected people' relationships. The findings suggest that these cases were strikingly comparable to those discovered through routine surveillance and hospital records [55].

A range of non-pharmaceutical treatments (NPI) are being employed to reduce and suppress the Covid-19 epidemic, according to the authors' research [50]. There are two types of NPIs: suppression treatments and mitigation treatments. By protecting patients from deadly diseases, the mitigation methods aim to reduce the strain on the hospital system. Preventative measures include social seclusion, isolation, and quarantine. The suppression seeks to reduce the number of new instances while maintaining the status quo for an extended period of time. Suppression techniques include school and university closures, as well as extended periods of social isolation. The authors give research on the effects of individual NPI as well as the combined effects of many interventions on the spread of Covid-19. The writers' research is being carried out using information obtained from the United States, the United Kingdom, and China. The findings demonstrate that integrating various interventions can significantly reduce the likelihood of the pandemic spreading. Because of the immediate and long-term effects these interventions have on countries' economies, countries must carefully select the appropriate NPIs for them.

A. *Sexual Disease Surveillance*

A dangerous viral that weakens the immune system is the Human Immunodeficiency viral (HIV). The CD4 cells are damaged and killed by it, leaving the patient's body susceptible to developing various infections and cancers. The HIV infection can be communicated through abrasive bodily fluids like blood, semen, breast milk, and rectal fluids, according to [68]. Early HIV detection boosts a person's chance of receiving treatment.

Human Immunodeficiency Virus (HIV) is a risky virus that impairs immune function. It causes harm to and kills CD4 cells, making the patient more susceptible to different infections and malignancies. According to [68], abrasive bodily fluids like blood, semen, breast milk, and rectal fluids can transmit the HIV virus. A person's likelihood of receiving treatment increases with early HIV detection.

The authors of [75] offered yet another method for using contact tracking to prevent the spread of STDs. In Taizhou Prefecture, Zhejiang Province, China, they employ contact tracking to detect HIV-positive people. Between 2008 and 2010, information was gathered for the investigation. Information was gathered from volunteers who had their HIV status checked. HIV-positive individuals are given surveys in order to create a network of their contacts. The investigation aids in the discovery of 463 new HIV infections for the authorities. As they develop their network through surveys, 398 incidents serve as the initial index cases for contact tracking. 1403 sexual encounters are found by watching the 398 index instances. Authorities can advise suspects to seek medical assistance when HIV infection is discovered early, which reduces the incidence of AIDS-related late diagnosis.

B. *Networked Social Contacts*

In [62], the authors demonstrate how to create a useful model that can simulate how an epidemic spread over huge social contact networks (SCNs). To execute the simulation methods on large SCNs, extensive computer resources are required. The authors suggested using parallel models to simulate huge SCNs and get findings in a reasonable amount of time. The authors use high-speed computer clusters to parallelize the model's execution. The SCN must be partitioned in order to perform many methods concurrently. The distribution skew generated by certain nodes' connections to densely populated hubs and other nodes' links to less populated networks is the fundamental issue with SCNs. The authors propose that the hubs (malls and metros) be divided into preset divisions so that the algorithm can be applied to each division at the same time. They also propose segregating the agents (visitors) to duplicate the network's behaviors. The results of the authors' studies on various datasets show how effective their strategy was.

The authors of [69] propose creating a model that can simulate epidemic spread. They are concerned with the propagation of an epidemic within interpersonal social networks. They consider the variety of social networks and do not offer each social network member an equal chance of contracting the sickness when the pandemic first began. The writers consider the crowding or protective effect in their approach. They create the Infected-Susceptible-Infected-Recovered (ISIR) model to simulate epidemic spread. According to the findings, only networks with a high number of connections are more likely to become contaminated. The high-degree set and critical set vaccination strategies are the most effective at containing the outbreak, according to the authors' research of four immunization strategies.

In order to work successfully, contact tracing applications must primarily complete four processes. These phases are shown in Fig. 3. The difficulties that these novel directions for applying contact tracing must overcome are discussed in the following section.

Table 1 provides a summary of the key publications' objectives and descriptions in this approach.

4 Challenges and Open Issues

The key difficulties in using contact tracing for practical applications are highlighted in this part.

In [32], the authors point out that privacy issues limit the efficacy of contact tracing applications. They consider the ramifications of these worries and look into

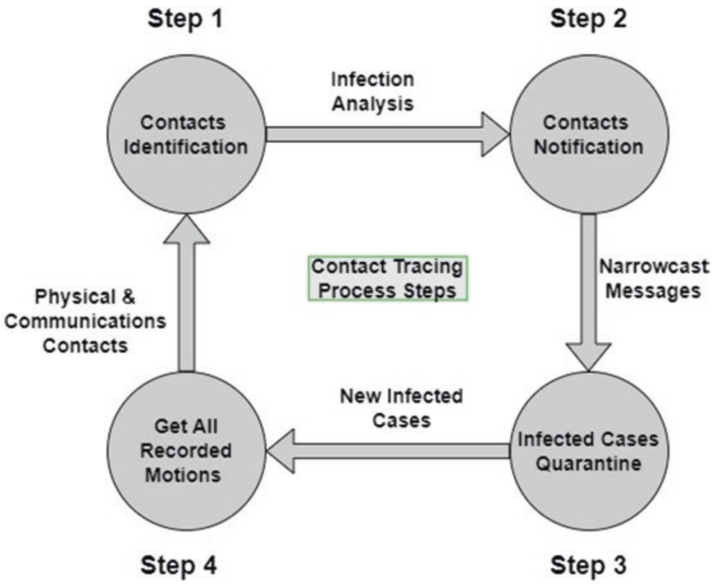


Fig. 3 Stages of contact tracing

possible solutions. The writers' privacy worries can be summed up as follows: protecting authorities' privacy is equally important as protecting contacts' private from prying eyes.

The authors raised the following contact tracing issues: (1) the difficulty in fully identifying contacts, (2) the lack of paper-based reporting systems, (3) incomplete contact lists, (4) ineffective data collection, (5) transcription errors made early in the contact tracing process, specifically during contact identification, and (6) delays in the identification and isolation of suspected individuals from these contacts [6, 9, 52].

The authors highlighted some of the challenges that can arise during the contact tracing process, such as the need for a substantial quantity of labor to precisely and successfully trace the links. As examples, (1) the Canadian Health Sector allows volunteers to perform contact tracing and other relevant tasks, and (2) during the peak of the COVID-19 disease outbreak, the governments of South Korea employed a large number of medical professionals and other workers to trace contacts and control the epidemic, as noted in [5, 14, 47].

The authors of [53] categorize each contact tracing technology into two classes: static and dynamic contact tracing strategies.

Figure 4 displays a classification of contact tracing tools based on contact tracing techniques. The writers include surveys and questionnaires in their definition of static individual contact tracing.

Although these methods have been successful in treating STDs, they have several drawbacks that can be summed up as follows: The spread of the epidemic cannot be determined online, and a labor force is needed. (3) Time-consuming, (4)

Table 1 Relevance and concept (summary)

Paper	Rational	Information sources	Study selection	Keywords	Year
[18]	Pairwise approximation and fully random simulations methods	Sexual relationships in Manitoba, Canada between 82 connected individuals together with the presence of chlamydia and gonorrhea infection	Relevance and Concept	tracing efficiency; transmission networks; pairwise approximations; stochastic simulation	2003
[41]	models for reducing the spread of HIV, these models consider the stochastic screening and contact tracing strategies	Synthetic dataset	Relevance and Concept	AIDS; Epidemic modeling; Screening; Reproductive number	2003
[45]	Employed wireless sensors to capture face-to- face interactions.	dataset covers CPIs of 94% of the entire school population, representing 655 students, 73 teachers, 55 staff, and 5 other persons, and it contains 2,148,991 unique close proximity records	Relevance and Concept	disease dynamics; human interactions	2010
[60]	large-scale Facebook data to clarify the disease dynamics in the community.	Synthetic dataset	Relevance and Concept	dynamics of infectious diseases spread; im munization interventions	2010
[12, 36]	Investigated social net- working sites to control the evolution of diseases.	None	Relevance and Concept	Disease control and surveillance; Social media applications	2013
[26]	Proposed a model for monitoring the reproduction rates for Ebola	Data of West Africa (Sierra Leone and Guinea) Ebola outbreak	Relevance and Concept	Contact tracing; Ebola; Epidemiology; Mathematical modeling.	2015
[61]	Transmission model investigates the perfect timing for contact tracing actions to be done.	Synthetic dataset	Relevance and Concept	Ebola; activity-driven network; compart- mental model	2018
[56]	Highlighted community surveillance roles and 109 guidelines to reduce the spread of COVID-19 disease.	None	Relevance and Concept	Quarantine	2020

Fig. 4 Devices classification for contact tracing

Requires individually identifying each relationship. Online questionnaires and surveys have been suggested as the online versions of these approaches to provide an online evaluation of the spread of an epidemic [15, 20, 24, 25]. Despite their advantages, the authors contend that online techniques can give decision-makers erroneous information and still do not give them real-time information [28]. When analyzing dynamic contact tracking approaches, the writers focus on GPS, wireless sensors, and mobile apps [7, 11, 19, 21, 66, 73, 74]. They highlight a number of drawbacks to these methods. Due to its limited lifespan, the mobile device's battery presents a significant challenge when it comes to mobile applications. Because of privacy issues, participants dislike wearing wearable sensors, and using them to track a large number of individuals is expensive. Radio Frequency Identification Devices (RFIDs) have a limited range for sending and receiving data. Even GPS technology has its limitations because it is useless in confined locations like classrooms and other institutions [53].

Tables 2 and 3 summarizes the main papers with their purposes and description in this direction.

5 Practical and Simulation Models

This section goes over the methods researchers can use to build models that can explain how epidemics spread as well as the effectiveness of contact tracing in preventing epidemics from spreading.

To forecast the spread of epidemics, models are created. Deterministic models and stochastic models are the two types [49].

A collection of differential equations are used to represent the deterministic model, which, provided that the parameters and circumstances are constant, always yields the same result. Deterministic models have the primary benefit of being able to demonstrate how the model's behavior can be influenced by its initial conditions.

Table 2 Applications (summary)

Paper	Rational	Information sources	Study selection	Keywords	Year
[75]	Utilise contact tracking software to find HIV-positive people in Taizhou Prefecture, Zhejiang Province, China.	463 HIV-infected individuals from 2008–2010 in Taizhou Prefecture, Zhejiang Province in China.	Applications	Sexual behavior, Sexual networks, HIV infection	2012
[54]	prospective cases, which represented instances that had previously experienced symptoms, had recovered, or had passed away; and positive predictive value (PPV), which represented contacts that had been traced.	data collected from six counties during June 2014-July 2015, 25,830 records for contacts who had monitoring initiated or were last exposed between June 4, 2014 and July 13, 2015.	Applications	Ebola epidemic	2014
[68]	To stop the spread of HIV in North Carolina (NC), con- tact tracing and phylogenet- ics are combined.	HIV cases reported in North Carolina from 2013–2014 and their reported contacts.	Applications	HIV-1, molecular epidemi- ology, phylo- genetics	2018
[62, 63]	A powerful model predicts how an epidemic would spread in huge social contact networks (SCNs).	3 Social network datasets were employed	Applications	epidemic spread	2019
[58].	developed the ECT app, a smartphone application to track Ebola patients in Sierra Leone.	Port Loko District, Sierra Leone, data capture using a smartphone application; April–August 2015 ECT app which was used to conduct contact tracing activities	Applications	Ebola; Sierra Leone	2019
[37]	A peer-to-peer mobile application efficiently handles contact tracing.	No dataset	Applications	COVID-19; epidemic; pandemic; peer-to-peer tracing; privacy;	2020
[23]	provided numerous methods for carrying out the process of contact tracing	synthetic	Applications	tracing strategies and pro- cedures	2020
[50]	Non-pharmaceutical approaches (NPI) to control and slow the Covid-19 epidemic's progress.	A synthetic dataset to represent the contacts of individ- uals, made within the house- hold, at school	Applications	Ebola epidemic COVID-19 mortality; Interventions	2020

Table 3 Challenges (summary)

Paper	Rational	Information sources	Study selection	Keywords	Year
[52, 59]	Classified the challenges in the contact tracing process; such as identification delays and data col- lection inefficiencies	Ebola Guinea and Sierra-Leone Reports	Challenges	Ebola, outbreak, epidemic	2015
[53]	Identified dynamic contact tracing methods challenges such as; mobile applications, wearable sensors, and Radio Frequency Identification evices	No datasets have been utilized	Challenges	Disease transmission, epidemic modeling	2018
[32]	Highlighted constraints limit the functionality of contact tracing applications such as privacy concerns	Based on synthetic dataset over Trace Together Software	Challenges	Privacy concerns; limitations	2020
[5]	Identified some difficulties with the contact tracing method, including the necessity for a lot of labor to efficiently and accurately track the contacts.	No datasets have been utilized	Challenges	COVID-19, HIV, AIDS	2020

It is possible to build a random model using differential equations. Every time the model is run, at least one probabilistic parameter causes the output to shift [49].

In [71], In order to simulate the COVID-19 outbreak in Wuhan, the authors present conceptual models; these models take into account how people will respond and what the governments will need to do. These models' main goals are to identify, comprehend, and forecast the outbreak's future patterns in terms of ratio. The COVID-19 paths can be captured by these models, and they can also record all expected future patterns for the outbreak. The authors correctly construct the model using essential COVID-19 information. This data also includes the time frame from the commencement of the sickness through hospital discharge or death.

Figure 5 demonstrates the critical phases that each individual goes through from the time they become infected until they stop being contagious.

In [57], The authors' proposed epidemiological model corresponds to reality. The purpose of this approach is to identify high-risk individuals, define how diseases develop, and exert control over them. The leverage of this model, according to the scientists, lies in its capacity to deliver prediction findings fast and how beneficial this will be for emergency response systems.

The Susceptible-Infectious-Recovered-Dead (SIDR) paradigm, which the authors propose employing in [65], is meant to trace the progression of the COVID-19 outbreak. The COVID-19 epidemic outbreak is being tracked using the

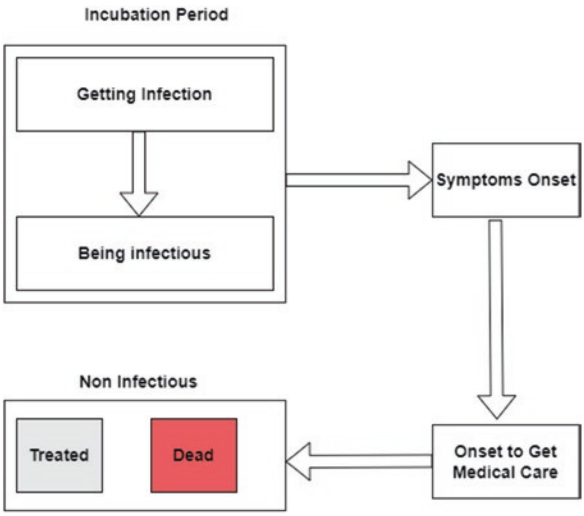


Fig. 5 Progress of Covid-19 infection

following data: (1) infection case estimates, (2) mortality estimates, (3) recovery rate estimates, and (4) fundamental reproduction numbers estimates. These kinds of projections are made every day. This model seeks to forecast the COVID-19 pandemic's evolution as an epidemic 3 weeks in advance.

In [23], the effectiveness of contact tracking in halting the spread of infectious diseases is examined by the writers. The majority of infections can be treated using the contact tracing strategy that the authors explore. This model is used to evaluate a variety of instances. The results demonstrate that (1) the diversity of detection time leads to an effective contact tracing, (2) iterative tracing can boost the effectiveness, and (3) there is generally no set formula to forecast the fraction that will be tracked.

The authors of [38] are interested in developing a simulator that can anticipate the Covid-19 pandemic using stochastic networks. To create the model, they use data from the Republic of Kazakhstan, which acts as a case study. The simulator employs the Susceptible-Exposed-Infectious-Recovered (SEIR) model to simulate how individuals transition between infection states. The authors run the simulator on data acquired from Lombardy (Italy) to test the model's efficacy. The simulator projected the duration of the outbreak and the number of afflicted people when different factors were modified.

The authors of [31] discuss the implications of employing contact trace to slow the transmission of the epidemic using a stochastic model based on the susceptible-infectious-removed. They use a graph to show how the virus has spread across social networks. This makes it easier to find new suspects during contact tracing. The model demonstrates the efficiency of contact tracking in slowing the disease's exponential spread, particularly in medium states. The model also shows how well random screening of people could prevent the epidemic from spreading.

The authors of [67] present a statistical model to assess the severity of Covid-19 using data obtained in Wuhan, Hubei. The information was provided by Wuhan's national and provincial health officials. They contribute data from 37 additional countries to their investigation into Covid-19 mortality in these countries. Key epidemic parameters, such as the period between the onset of the disease and death or hospital discharge, and the case fatality rate, have been approximated using the data. The algorithm is first used to data acquired in China, and then it is applied to data gathered in other countries. Results indicate that the model's predictions were very close to actual data gathered from Chinese authorities and additional countries.

The authors of [49] focus on constructing a stochastic susceptible-infected-recovered (SIR) model that incorporates contact tracing to analyze the importance of contact tracing in epidemic control. Using a tree structure, the techniques replicated the spread of the disease. The authors develop equations to explain how an epidemic spreads and how intervention instruments influence the spread rate. These formulas are then applied to graphs with no tree contacts.

The authors of [72] examine various simulation models that have been presented to mimic the spread of MERS and SARS. The authors describe and thoroughly discuss a number of essential factors that may influence how well the models function, including as the basic reproduction number, incubation time, latent period, and infectious period. They also expand on their research to investigate the effectiveness of intervention strategies such as isolation, contact tracing, and quarantine in slowing or stopping the development of the epidemic.

The three main case types produced by the contact tracing method are (1) infection cases, (2) fatality cases, and (3) recovery cases. Any prediction models that accounts for the infection's reproduction ratio can use these results as an input. Figure 6 summarizes the inputs and outputs used by prediction algorithms that seek to estimate the infection production ratio. Table 4 provides a summary of the key publications' objectives and descriptions in this approach.

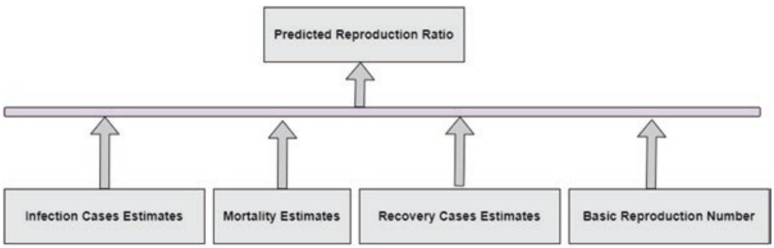


Fig. 6 Prediction models (reproduction ratio)

Table 4 Models (summary)

Study	Rational	Information sources	Study selection	Keywords	Year
[31]	To discuss the impact of implementing contact tracing to stop the epidemic from spreading, stochastic model based on the susceptible-infectious is used.	Synthetic dataset	Models	Epidemic control; epidemic mitigation	2002
[57]	Epidemiological model to explain and manage disease evolution and identify people at high risk of exposure.	Synthetic dataset	Models	Epidemiologica models	l2016
[71]	Simulate the outbreak of COVID-19 in Wuhan based on governmental acts and personal behaviors required to prevent outbreak ratio.	Real dataset contained the daily number and cumulative number of cases and deaths of COVID-19 in Wuhan, China.	Models	COVID-19; Lockdown; epidemic	2019
[41]	Simulate the spread of SARS and the Middle East Respiratory Syndrome epidemics. (MERS).	None	Models	Coronavirus Epidemics, transmission modelling, SARS, MERS	2019
[65]	Model designed to track the spread of the COVID-19 outbreak.	The fresh daily confirmed cases reported for the Hubei province from January 11 through February 10, according to a true dataset, have been used.	Models	Epidemiologica models	l2020
[39]	Represent how people Change from one infectious state to another.	The COVID-19 data collected from Lombardy region, Italy has been used to calibrate the proposed model.	Models	Stochastic epidemic, epidemiology, epidemic suppression, SEIR model.	2020
[67]	A statistical model to assess the impact of Covid-19 based on information gathered in Wuhan, Hubei.	Dataset that contained individual-case data for patients who died from COVID-19 in Hubei, mainland China and for cases outside China from government and reports for 37 countries, as well as Hong Kong, until Feb 25, 2020.	Contact Tracing Models	Covid-19 outbreak	2020

5.1 Case Study: Real-Time Framework to Process Contacts Traces and Identify Suspected Contacts

This section primarily introduces TraceAll, a technique developed by authors in [30] that tracks all contacts exposed to infected patients and creates a list of potentially affected individuals. The importance of TraceAll resides in its ability to trace suspected contacts in real time. The following are the main contributions of this technique: (1) efficiently answering contact tracing queries, (2) quick retrieval of evidence relating to suspected instances that may be infected, and (3) evaluation of technique on real datasets based on multiple metrics.

A. *TraceAll Parameters*

This section presents the parameters that the technique takes as input to trace contacts.

Parameter 1: *Exposure Time* $T_{Exposure}$ is the duration of contact between a healthy individual and an infected individual.
Parameter 2: *Social Distance D*, is the shortest distance between a healthy individual and an infected individual.
Parameter 3: *Contact Tracing Query* (*CTQ*),is the inquiry that tracks the individuals who have been connected to the infected person either directly or indirectly within a certain $T_{Exposure}$ and *D*.
Parameter 4: *Meeting Point*, is the location where an infected person and a healthy person simultaneously come into contact, T_{Diff} (*infected, normal*) = 0, in this study T_{Diff} is measured in minutes.
Parameter 5: *Direct Tracing*, is the designation of people as *high-risk* for a specific tracking time who have a direct physical link to an infected person.
Parameter 6: *In-Direct Tracing*, is the iterative finding of contacts for those who had a direct physical touch with an infected person within a particular tracing time; these contacts are referred to as *low-risk* contacts.

B. *Methodology*

Given are the following variables: the tracing period ($Tracing_period$), the exposure time ($T_Exposure$), the social-spatial distance (*D*) for exposure, and the trajectory data set ($tau_Oothers$) for objects moved in the space. Each trajectory contained in $tau_Oothers$ consists of a list of places visited, each represented by a pair of timestamps and longitude values (*L* and *T*);τ={$(l_1, T_1), (L_2, T_2). (L_N , T_N)$}.

Given are the following variables: the tracing period ($Tracing_period$), the exposure duration ($T_Exposure$), the social- spatial distance (*D*) for exposure, and the trajectory data set ($tau_Oothers$) for objects moved in the area. Every trajectory in $Tau_Oothers$ is a list of timestamps *T* paired with traveled places *L*, where each location is identified by its longitude and latitude values;τ={$(l_1, T_1), (L_2, T_2). (L_N , T_N)$}.

Let the query utilize the infected person, Q_user, who has the trajectory tau_Q. The tau_Q represents the path that the infected person is traveling at a specific time. A set of contacts *C* (*C* subset $tau_Oothers$) exposed to the Q_user during the $Tracing_period$

and whose exposure time was higher than or equal to $T_Exposure$ within the spatial range of D must be determined. An additional marker that shows that this query is being monitored is the $Tracing_M\ ode$ Boolean indicator, which has the values (1) direct and (2) indirect.

Since only contacts connected to the Q_user are returned when using direct tracing, this type of query is also referred to as a snapshot query (first-level query). In the case of indirect tracing, contacts connected to Q_user are returned during the tracing period; initially, the contacts connected to Q_user are retrieved as with direct tracing; then, recursive calls have occurred to the new retrieved list; each person in this list is considered a new Q_user; the break condition is the tracing period's end date; the recursion depth is represented by the $Tracing_period$ ((multi-level query).

Tracing Query Cardinality
C individuals who are suspected of being infected; these contacts may be high-risk or low-risk.

Query Criteria
contacts that link the Q_user to other contacts during the $Tracing_period$ and the time that has passed between the Q_user and other contacts Social distance between Q_user and other acquaintances as a result of $ge\ T_Exposure$ & $Tracing_M\ ode$ & $le\ D$ = (direct or indirect).

Demonstration Example
The tracing method is direct, and the tau_Q record is represented as follows: (20-August - ($Location_1$, 02: 30: 00), ($Location_2$, 02: 35: 00), ($Location_3$, 02: 40: 00), …($Location_N$, 02: 55: 00)). The goal is to gather all contact histories that connect to Q_user between August 6 and August 20, with a 2-meter spatial separation between tau_Q and other contacts' trajectories and a common time of at least 10 minutes between Q_user and other contacts. It is important to note that if the tracing mode is indirect, the list extracted for the Q_user must be fetched the first time, and each item in the list is treated as a new Q_user. The list belonging to each fetched Q_user must then be iterated once more, and so on, until the end of the tracing period.

The Table 5 highlights the contact tracing query's input settings.

C. *Architecture*

The four main steps of the proposed solution *TraceAll* are briefly discussed as follows:

– **Step 1:** restricts the Query User's trajectory. The purpose of this stage is to set space-time bounds for the query user's trajectory. It provides the beginning

Table 5 Parameters table

Parameter name	Value
Tracingperiod	14 days
TExposure	10 min
D	2 meters
TracingMode	Direct

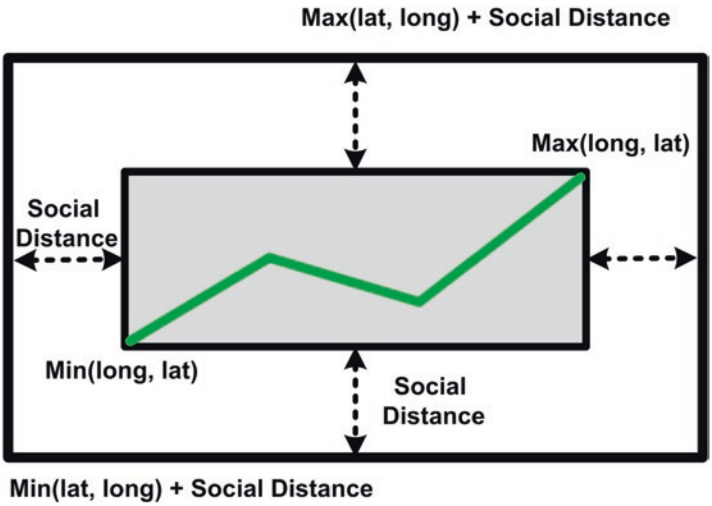


Fig. 7 Overlapped region

and ending times of the trajectory for the inquiry item. Based on the minimum (minimum longitude, minimum latitude) and maximum (maximum longitude, maximum latitude) points in the trajectory, the method constructs a space zone that surrounds the query user's trajectory. Figure 7's inner rectangle depicts this stage.

- **Step 2:** Locate the Overlapping Region. The purpose of this stage is to create a zone that captures the trajectories of all objects that travelled near the query item and may have been affected by it. The region from Step 1 is now finished by adding the D to all of its edges. More specifically, this freshly generated region contains the region that limits the query object's trajectory. This stage is depicted in Fig. 7 by the outer rectangle.
- **Step 3:** Extraction of Overlapping Trajectories. The purpose of this step is to extract any trajectories that cross the overlapped area created in step 2 and that are between the start and finish times of the query object's trajectory. The owners of these trajectories are regarded to be questionable goods that demand additional examination.
- **Step 4:** Extract infected trajectory. The purpose of this stage is to extract all user trajectories in which their infection has been verified based on the illness infection circumstances. For these trajectories to be evaluated, two key characteristics must be met: (1) The distance between the trajectory and the query trajectory must be less than or equal to the social distance identified as a risk factor for infection, and (2) the exposure time must be equal to the exposure time identified as a risk factor for infection.

Two main algorithms (Algorithm 1 and Algorithm 2) describe the full functionality of the proposed technique.

Algorithm 1: TraceAll: Contact Tracer Technique

```
procedure CONTACT_TRACER
INPUT: Tracing_period, T_Exposure, D.
  /* Overlapped Traces list */
  Let OverlappedTracesList OL ← ϕ
  /* Infected list */
  Let SuspectedList IL ← ϕ
  /* Overlapped Region */
  Let OverlappedRegion Region_Overlapped ← ϕ
  /* Region Bounding Query Trajectory */
  Let RegionBounding_τQ ← ϕ
  /* Query TrajectoriesList */
  Let QL_τ ← ϕ
  Tracing_StartDate = CurrentDate
  Tracing_EndDate = CurrentDate − Tracing_period
  for Tracing_StartDate to Tracing_EndDate do
    /* τ_Q Query object's trajectories */
    Get QL_τ in iterated day
    for each τ_Q ∈ QL_τ do
      /* τ_Others Other object's trajectories */
      Get τ_Others in iterated day
      /*Region created from max&min points τ_Q*/
      Regionτ_Q = Bound τ_Q
      /*Add social distance to overlapped region*/
      Region_Overlapped = D_Eucl + Regionτ_Q
      /*Get traces in overlapped region by R-Tree*/
      OL ← R-Tree (Region_Overlapped, τ_Others)
      for each o ∈ OL do
        isExposed=IsExposed(τ_Q,o)
        if isExposed = True then
          add o to IL
        end if
      end for
    end for
  end for
  if TracinMode = InDirect then
    for 0 to Tracing_period do
      /*Increment IL by fetched contacts*/
      Call Mobility_Tracer
    end for
  end if
end procedure
OUTPUT: Return objects belongs to IL
```

Algorithm 1 shows the pseudo-code for the contact tracing method that has been suggested. The algorithm needs three inputs: (a) the tracing period ($Tracing_{p}eriod$) parameter, which represents the duration of the disease's incubation; (b) the time exposure threshold ($T_{E}xposure$); and (c) the social distance (D) parameter, which indicates whether the infection can spread from an infected person to a healthy person if the measured distance is less than this threshold. The algorithm produces a list of

Algorithm 2: IsExposed

1: **procedure** ISEXPOSED
2: **INPUT:** τ_1, τ_2, $T_{Exposure}$, D.
3: let *Exposed* ← *false*
4: **for** *each* $Point_i \in \tau_1$ **do**
5: **for** *each* $Point_j \in \tau_1$ **do**
6: */*Verify* τ_1 *and* τ_2 *intersects*/*
7: **if** $Time_{Diff}(Point_i, Point_j) = 0$
8: $\&distance(Point_i, Point_j) \leq D$ **then**
9: $StartTimeStamp = TimeStamp\ of\ Point_i$
10: $EndTimeStamp = StartTimeStamp + T_{Exposure}$
11: */*Points from startTimestamp to EndTimeStamp* τ_1**/*
12: $Sub\tau_1$ = SubTrajectory(τ_1)
13: */*Points from startTimeStamp to EndTimeStamp* τ_2**/*
14: $Sub\tau_2$ = SubTrajectory(τ_2)
15: */* Average Distance Calculations */*
16: **if** $distance(Sub\tau_1, Sub\tau_2) \leq D$ **then**
17: *Exposed = true*
18: *Exit loops; Return Exposed value*
19: **end if**
20: **end if**
21: **end for**
22: **end for**
23: **end procedure**
24: **OUTPUT:** Return *Exposed*

people who might be affected based on their paths. When the infected patient (query user) gives a flag indicating that he became infected, the algorithm initially takes the date of the query into account as the start date of the tracing. The tracing end date is then obtained by deducting the tracing time from the start date (Lines 13–14). From the tracing start date to the tracing end date, the algorithm iterates at line 15, getting all of the patient's trajectories for this day for each repetition. (line 17). In line 18, the algorithm iterates through the daily trajectory iterations of the inquiry user. The algorithm next built boundaries for the trajectory of the iterated query object, including both spatial and temporal boundaries. (line 22). The algorithm creates a zone called the overlapping region in line 24 to record the trajectories of the closest neighbors that have moved in close proximity to the trajectory of the query item. With the identified *D*, bounds were added from all sides to the trajectory of the query object. The algorithm then retrieves all trajectories that were a part of the query object's trajectory from its start time to its finish time when it was going through the overlapping area created in line 24 using the three-dimensional R-Tree index in line 26. The algorithm considers these trajectories to be suspects that need to be investigated from space- and time-perspective in order to confirm the infection based on the circumstances for disease transmission. The algorithm iterates across these overlapping trajectories, checking each time if the iterated trajectory satisfies the requirements for infection exposure by comparing the query object's route to the iterated trajectory. If so, the algorithm repeats the exposure tests performed by algorithm 2, lines 27 to 32, or it adds the iterated trajectory to the list of trajectories that have been infected. It's crucial to keep in mind that the algorithm configures the tracing settings prior to running. If the tracing mode is set to *InDirect*, the algorithm makes recursive calls to get the contacts of contacts in each iteration using direct tracing (lines 35–40). The system then produces a list of infected users together with their trajectory.

Remember that the proposed method considers any positions with a time difference of zero and a spatial distance that is less than or equal to the determined social distance to be *hotspots* which are places where there is a high risk of hazard.

Algorithm. Algorithm 2 outlines the formula used to determine the difference in exposure between an infected patient and a healthy individual. The result reveals whether the disease was transferred from the infected patient to the healthy individual or not. Four parameters are given to the algorithm: $T_E xposure$ stands for the identified exposure time period and has the same value set in the algorithm 1 $T_E xposure$ represents the trajectory of the infected patient, Tau_2 stands for the suspected user's trajectory, D stands for the identified distance, and $T_E xposure$ has the same setting in algorithm 1. The method generates a boolean flag that indicates whether or not the illness spreads to this healthy person. The algorithm initially iterates over the locations on both trajectories to ensure that the objects of these trajectories cross paths at a specific location. In line 7, the algorithm checks to see if the time difference between the iterated points is zero. Two trajectories are either at the same point or at separate points at the same time when there is no temporal gap between them. The algorithm runs one additional check before starting the tracing. It calculates the separation between the meeting places $Point_i$ and $Point_j$ from tau_1 and tau_2, respectively. The method does not complete the tracing and instead continues the procedure for all additional meeting places if the distance exceeds $D_E ucl$. If not, the algorithm terminates the tracing. After that, the algorithm logs the time of $Point_i$ in line 9 and stores the value in a StartTime parameter. The method then added a new parameter called EndTime in line 10 to maintain the point's time at a particular moment. EndTime was calculated by adding $T_E xposure$ to the start time definition in line 9. The outcome is then stored in a new variable named $Sub_t au_1$ in line 12 after the algorithm obtains the sub-trajectory of tau_1 from a position at the StartTime to a point at the EndTime. Similar to this, the algorithm gets the sub-trajectory from tau_2 starting at the point at the StartTime and ending at the point at the EndTime, and line 14 stores the result in a new variable named $Subtau_2$. In line 16, the algorithm determines the typical separation between the generated sub-trajectories. If the average distance is less than or equal to *Dsocial*, the algorithm returns that the tau_2 belongs to an infected person and ends the loops; otherwise, the algorithm keeps iterating. It is important to remember that when meeting spots are located, the Hausdorff distance method is utilized to calculate the distance between the newly constructed $Sub_t au_1$ and $Sub_t au_2$ ([8, 35, 44]).

The Hausdorff algorithm, in actuality, iterates over the $Subtau_1$-dimensional points, computing the distance between $Subtau_1$ and $Subtau_2$ points in each iteration, obtaining the minimum distance for each point comparison, and then producing the maximum of the minimum distances generated in each iteration. In a manner similar to this, Hausdorff iterates over the $Subtau_2$-points, computing the distance

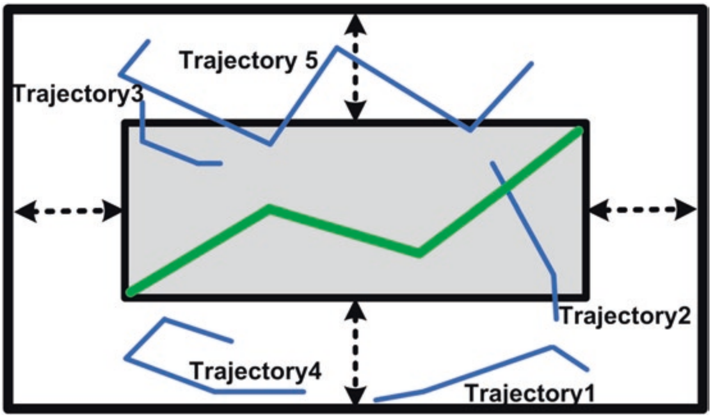


Fig. 8 Overlapped trajectories

between the $Subtau_2$-points and $Subtau_1$-points in each iteration, obtaining the least distance for each point comparison, and then constructing the maximum of the minimum distances generated in each iteration. The distance is calculated as the sum of $Subtau_1$ and $Subtau_2$ since Hausdorff produces the highest values for $Subtau_1$ and $Subtau_2$. To address the specified scenarios, Hausdorff is altered in this study. If any two points' distances exceed the D, the adaptation is performed; Hausdorff then ceases looking for additional points and informs the Algorithm 2, further decreasing processing time, that the computed distance between the compared sub-trajectories exceeds the D. Figure 7 demonstrates the creation of the overlapping region and how the suggested method bound the trajectory of the query object. It begins by looking for the query object's trajectory's minimum and maximum points. Along the trajectory of the query object, the maximum point indicates the maximum longitude and latitude, and the minimum point indicates the minimum longitude and latitude. Then, the dimensions of the rectangle containing the trajectory of the query object are determined. The method starts by generating a larger rectangle to be the overlapped region, all in accordance with the rectangle boundaries created around the trajectory of the query object, and then creating the overlapped region by adding the identified social distance D to all sides of the created rectangle. Figure 8 gives a general description of how the proposed method detects and treats as questionable cases the nearest neighbor's trajectories that move around the query object's trajectory. The system first takes advantage of the overlapped region produced by Fig. 7 and uses a three-dimensional R-Tree to carry out a range search query in order to obtain all trajectories that are within the overlapped region and fall between the start and end times of the query object's trajectory. In this context, the term" three dimensions" refers particularly to two space-related dimensions and one time-related one. A specific type of tree index structure called R-Tree is employed to efficiently manage spatial data objects. The rectangle is represented by the letter R in the term *R-Tree*. The primary concept of the R-Tree structure is to group nearby spatial objects and store them in MBR (minimum bound rectangle) in the next level of the

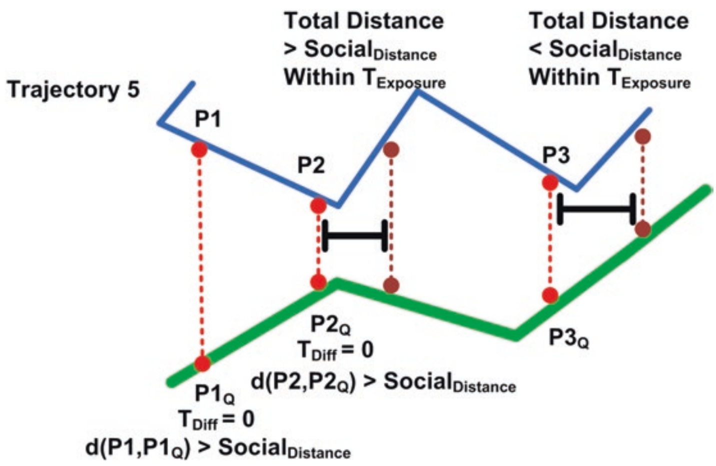


Fig. 9 Computing exposure

tree. R-Tree was chosen for this project because it performs closest neighbor and window queries much more quickly; like inside, covers, and contains.

Figure 9 gives a thorough explanation of the three qualities of meeting venues, social distance, and exposure time as they relate to the tracing process. Besides, Fig. 9 explains how these traits relate to one another. Having obtained the intersecting paths from Fig. 8, Each overlapped trajectory is first examined. Figure 9 outlines three instances that can be examined to determine whether or not this trajectory belongs to a patient who is infected. According to the suggested method, there is no time difference between point P1 and its corresponding point $P\,1_Q$ in the trajectory of the query object in the first scenario. The technique then scans other parts of the trajectory since the distance between these two spots exceeds the calculated social distance D

In the second case, the method establishes that there is no time difference between point P2 and the corresponding position $P\,2_Q$ in the trajectory of the query object, and it further establishes that the distance between the two points is smaller than the established social distance. The technique starts by analyzing the sub-trajectories from the two trajectories (the trajectory of the query object and trajectory 5). These sub-trajectories were created by choosing a segment from each trajectory, going ahead from the point where there is no time difference to the point where the identified exposure time is added. The technique then determines the average distance between these sub-trajectories; in this case, it finds that the average distance is greater than the estimated social distance, thus it continues to scan other parts of the trajectories. Case 3 is similar to Case 2 except that the algorithm recognizes that this trajectory is from an infected user and that the average distance between the produced sub-trajectories is closer to the estimated social distance.

Indeed, our proposed solution is reliable because it reacts to contact tracking requests speedily and successfully in real time. Additionally, the flexibility of retrieving contacts and the flexibility of configuring the disease's incubation period enable the end-user to manage the contact tracing process in real time with ease.

6 Conclusion

The contact tracing process is a surveillance and mitigation approach that looks to discover infectious diseases in order to promptly contain their epidemic. In this work, we investigated a number of applications that contributed to the development of the contact tracing method. Additionally, we went into great detail about the challenges that contact tracing software must overcome, such as privacy concerns, the inability to identify contacts and delays in full identification. In this paper, we've demonstrated the methods, scientists can build models that can explain how epidemics propagate and the effect contact tracing has on containing their spread. Finally, we reviewed the available research and enumerated the applications, domains of use, and open directions.

References

1. Al-Shourbaji, I., Kachare, P. H., Abualigah, L., Abdelhag, M. E., Elnaim, B., Anter, A. M., & Gandomi, A. H. (2022). A deep batch normalized convolution approach for improving COVID-19 detection from chest X-Ray images. *Pathogens, 12*(1), 17.
2. Anter, A. M., Mohamed, A. W., Zhang, M., & Zhang, Z. (2023). A robust intelligence regression model for monitoring Parkinson's disease based on speech signals. *Future Generation Computer Systems, 147*, 316–327.
3. Fraser, C., Riley, S., Anderson, R. M., & Ferguson, N. M. (2004). Factors that make an infectious disease outbreak controllable. *Proceedings of the National Academy of Sciences of the United States of America, 101*(16), 6146–6151.
4. Armbruster, B., & Brandeau, M. (2007). Contact tracing to control infectious disease: When enough is enough. *Health Care Management Science, 10*, 341–355.
5. Nosyk, B., Armstrong, W. S., & Del Rio, C. (2020). Contact tracing for Covid-19: An opportunity to reduce health disparities and end the human immunodeficiency virus/aids epidemic in the United States. *Clinical Infectious Diseases, 71*(16), 2259–2261.
6. Wykoff, R. F., Heath, C. W., Hollis, S. L., Leonard, S. T., Quiller, C. B., Jones, J. L., Artzrouni, M., & Parker, R. L. (1998). Contact tracing to identify human immunodeficiency virus infection in a rural community. *JAMA the Journal of the American Medical Association, 259*(24), 3563–3566.
7. Regueiro, M. A., Viqueira, J. R. R., Taboada, J. A., & Cotos, J. M. (2015). Virtual integration of sensor observation data. *Computers & Geosciences, 81*, 12–19.
8. Atallah, M. J. (1983). A linear time algorithm for the Hausdorff distance between convex polygons. *Information Processing Letters, 17*(4), 207–209.
9. Mcnamara, L. A., Schafer, I. J., Nolen, L. D., Gorina, Y., Redd, J. T., Lo, T., Ervin, E., Henao, O., Dahl, B. A., & Morgan, O. (2016). Ebola Surveillance – Guinea, Liberia and Sierra Leone. *MMWR Supplement, 65*, 35–43.
10. Singapore Government Blog. *Help speed up contact tracing with tracetogether*. https://www.gov.sg/article/help-speed-up-contact-tracing-with-tracetogether. Accessed Mar 2020.

11. Ballari, D., Bruin, S., & Bregt, A. K. (2012). Value of information and mobility constraints for sampling with mobile sensors. *Computers & Geosciences, 49*, 102–111.
12. Mandeville, K. T., Harris, M., Thomas, H. L., Chow, Y., & Seng, C. (2014). Using social networking sites for communicable disease control: Innovative contact tracing or breach of confidentiality? *Public Health Ethics, 7*(1), 47–50.
13. Vazquez-Prokopec, G. M., Montgomery, B. L., Horne, P., Clennon, J. A., & Ritchie, S. A. (2017). Combining contact tracing with targeted indoor residual spraying significantly reduces dengue transmission. *Science Advances, 3*(2), e1602024.
14. Epidemiology COVID-19 National Emergency Response Center, Korea Centers for Disease Control Case Management Team, and Prevention. (2020). Coronavirus disease-19: Summary of 2,370 contact investigations of the first 30 cases in the Republic of Korea. *Osong Public Health Res Perspect, 11*(2), 81–84.
15. Dalton, C., Durrheim, D., Fejsa, J., Francis, L., Carlson, S., d'Espaignet, E. T., & Tuyl, F. (2009). Flutracking: A weekly Australian Community online survey of influenza-like illness in 2006, 2007 and 2008. *Communicable Diseases Intelligence Quarterly Report, 33*, 316–322.
16. Dixon, M. G., & Schafer, I. J. (2014). Ebola viral disease outbreak – West Africa. *MMWR Morbidity and Mortality Weekly Report, 63*(25), 548.
17. Eames, K. T. D. (2006). Contact tracing strategies in heterogeneous populations. *Epidemiology & Infection, 135*(3), 443–454.
18. Eames, K. T. D., & Keeling, M. J. (2003). Contact tracing and disease control. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 270*(1533), 2565–2571.
19. Heavner, M. J., Fatland, D. R., Hood, E., & Connor, C. (2011). Seamonster: A demonstration sensor web operating in virtual globes. *Computers & Geosciences, 37*, 93–99.
20. Dalton, C., Carlson, S., Butler, M., Cassano, D., Clarke, S., Fejsa, J., & Durrheim, D. (2017). Insights from flu-tracking: Thirteen tips to growing a web-based participatory surveillance system. *JMIR Public Health Surveill, 3*, e7333.
21. Salathe, M., Kazandjieva, M., Lee, J. W., Levis, P., Feldman, M. W., & Jones, J. H. (2010). A high-resolution human contact network for infectious disease transmission. *Proceedings of the National Academy of Sciences of the United States, 107*, 22020–22025.
22. Donnelly, C. A., Ghani, A. C., Leung, G. M., Hedley, A. J., Fraser, C., Riley, S., Abu-Raddad, L. J., Ho, L.-M., Thach, T.-Q., Chau, P., Chan, K.-P., Lam, T.-H., Tse, L.-Y., Tsang, T., Liu, S.-H., Kong, J. H. B., Lau, E. M. C., Ferguson, N. M., & Anderson, R. M. (2003). Epidemiological determinants of spread of causal agent of severe acute respiratory syndrome in Hong Kong. *The Lancet, 361*(9371), 1761–1766.
23. Klinkenberg, D., Fraser, C., & Heesterbeek, H. (2020). The effectiveness of contact tracing in emerging epidemics. *Plos ONE, 1*(1), e12.
24. Merino, N., Sanchez, R. L., Munoz, A., Prada, G., Garcia, C. F., & Frank Polk, B. (1990). *HIV-1, sexual practices, and contact with foreigners in homosexual men in Colombia, South America* (pp. 330–334). Johns Hopkins University.
25. Merino, N., Sanchez, R. L., Munoz, A., Prada, G., Garcia, C. F., & Polk, B. F. (2003). Prevalence of unsafe sexual behavior among HIV-infected individuals: The swiss HIV cohort study. *Journal Acquired Immune Deficiency Syndromes, 33*, 494–499.
26. Browne, C., Gulbudak, H., & Webb, G. (2015). Modeling contact tracing in outbreaks with application to Ebola. *Journal of Theoretical Biology, 384*, 33–49.
27. Rutherford, G. W., & Woo, J. M. (1988). Contact tracing to control the spread of HIV-reply. *JAMA The Journal of the American Medical Association, 260*(22), 3275–3275.
28. Klous, G., Smit, L. A. M., Borlee, F., Coutinho, R. A., Kretzschmar, M. E. E., Heederik, D. J. J., & Huss, A. (2017). Mobility assessment of a rural population in The Netherlands using GPS measurements. *International Journal of Health Geographics, 16*, 1–13.
29. Thakare, A., Anter, A. M., & Abraham, A. (2023). Seizure disorders recognition model from EEG signals using new probabilistic particle swarm optimizer and sequential differential evolution. *Multidimensional Systems and Signal Processing, 34*, 1–25.

30. Alarabi, L., Basalamah, S., Hendawi, A., & Abdalla, M. (2021). Traceall: A real-time processing for contact tracing using indoor trajectories. *Information, 12*, 202–222.
31. Huerta, R., & Tsimring, L. S. (2002). Contact tracing and epidemics control in social networks. *Physical Review E, 66*, 056115.
32. Cho, H., Ippolito, D., & Yu, Y. W. (2020). Contact tracing mobile apps for Covid-19: Privacy considerations and related trade-offs. *Arxiv*.
33. Anter, A. M., Oliva, D., Thakare, A., & Zhang, Z. (2021). AFCM-LSMA: New intelligent model based on Lévy slime mould algorithm and adaptive fuzzy C-means for identification of COVID-19 infection from chest X-Ray images. *Advanced Engineering Informatics, 49*, 101317.
34. Cebrian, M., Farrahi, K., & Emonet, R. (2014). Epidemic contact tracing via communication traces. *Plos ONE, 9*(5), e95133.
35. Kim, I.-S., & Mclean, W. (2013). Computing the hausdorff distance between two sets of parametric curves. *Communications of the Korean Mathematical Society, 28*(4), 833–850.
36. Stein, M. L., Rump, B. O., Kretzschmar, M., & Van Steenbergen, J. E. (2013). Social networking sites as a tool for contact tracing: Urge for ethical framework for normative guidance. *Public Health Ethics, 7*(1), 57–60.
37. Elaraby, M. E., & Anter, A. M. (2022). Web-based modernized architecture over cloud computing for facial extraction and recognition. In *Handbook of research on applied intelligence for health and clinical informatics* (pp. 394–414). IGI Global.
38. Kuzdeuov, A., Baimukashev, D., Karabay, A., Ibragimov, B., Mirzakhmetov, A., Nurpeiissov, M., Lewis, M., & Varol, H. A. (2020). A network-based stochastic epidemic simulator: Controlling Covid-19 with region-specific policies. *Medrxiv*.
39. Kuzdeuov, A., Baimukashev, D., Karabay, A., Ibragimov, B., Mirzakhmetov, A., Nurpeiissov, M., Lewis, M., & Varol, H. A. (2020). A network-based stochastic epidemic simulator: Controlling Covid-19 with region-specific policies. *IEEE Journal of Biomedical and Health Informatics, 28*(4), 2743–2754.
40. Landis, S. E., Schoenbach, V. J., Weber, D. J., Mittal, M., Krishan, B., Lewis, K., & Koch, G. G. (1992). Results of a randomized trial of partner notification in cases of HIV infection in North Carolina. *New England Journal of Medicine, 326*, 101–106.
41. Hyman, J. M., Li, J., Ann, E., & Stanley. (2003). Modeling the impact of random screening and contact tracing in reducing the spread of HIV. *Mathematical Biosciences, 181*(1), 17–54.
42. Basha, S. H., Anter, A. M., Hassanien, A. E., & Abdalla, A. (2023). Hybrid intelligent model for classifying chest X-Ray images of COVID-19 patients using genetic algorithm and neutrosophic logic. *Soft Computing, 27*(6), 3427–3442.
43. Maghdid, H. S., & Ghafoor, K. Z. (2020). A smartphone enabled approach to manage Covid-19 lockdown and economic crisis. *Social and Information Networks, 1*(5), 271.
44. Mcmullen, P. (1984). The Hausdorff Distance Between Compact Convex Sets. *Mathematika, 31*(1), 76–82.
45. Salathe, M., Kazandjieva, M., Lee, J. W., Levis, P., & Feldman, M. W. (2010). A high resolution human contact network for infectious disease transmission. *Proceedings of the National Academy of Sciences, 107*(51), 22020–22025.
46. Norwood, C. (1995). Mandated life versus mandatory death: New York's disgraceful partner notification record. *Journal of Community Health, 20*(2), 161–170.
47. Government of Canada. Covid 19: How you can make a difference. https://www.canada.ca/en/public-health/services/diseases/2019-novel-coronavirusinfection/make-a-difference.html. Accessed Apr 2020.
48. Australian Government Department Of Health: Covidsafe App. *Covidsafe App*. https://www.health.gov.au/resources/apps-and-tools/covidsafe-app. Accessed Apr 2020.
49. Okolie, A., & Muller, J. (2020). Exact and approximate formulas for contact tracing on random trees. *Mathematical Biosciences, 321*, 108320.

50. Okolie, A., & Muller, J. (2020). Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce Covid-19 mortality and healthcare demand. *Imperial College London, 10*(77482), 491–497.
51. World Health Organization. (2014). *Contact tracing during an outbreak of Ebola virus disease*. WHO.
52. Ilesanmi, O. S. (2015). Learning from the challenges of Ebola virus disease contact tracers in Sierra Leone. *The Pan African Medical Journal, 22*, 21.
53. Chen, H., Yang, B., Pei, H., & Liu, J. (2018). Next generation technology for epidemic prevention and control: Data-driven contact tracking. *IEEE Access, 7*, 2633–2642.
54. Swanson, K. C., Altare, C., Wesseh, C. S., Nyenswah, T., Ahmed, T., Eyal, N., Hamblionid, E. L., Lessler, J., Petersid, D. H., & Altmann, M. (2018). Contact tracing performance during the Ebola epidemic in Liberia, 2014–2015. *Neglected Tropical Diseases, 12*(9), e0006762.
55. Ma, T., Bi, J. L. Q., Yongsheng, W., Mei, S., Ye, C., Zou, X., Zhang, Z., Liu, X., Wei, L., Truelove, S. A., Zhang, T., Gao, W., Cheng, C., Tang, X., Xiaoliang, W., Yu, W., Sun, B., Suli Huang, Y., Sun, J. Z., & Feng, T. (2020). Epidemiology and transmission of Covid-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: A retrospective cohort study. *The Lancet Infectious Diseases, 20*(8), 911–919.
56. Gowda, G., Holla, R., Ramraj, B., & Gudegowda, K. S. (2020). Contact tracing and quarantine for Covid 19: Challenges in community surveillance. *Indian Journal of Community Health, 32*, 306–308.
57. Rodrigues, H. S. (2016). Application of SIR epidemiological model: New trends. *International Journal of Applied Mathematics and Informatics, 10*, 92.
58. Danquah, L. O., Hasham, N., Macfarlane, M., Conteh, F. E., Momoh, F., Tedesco, A. A., Jambai, A., Ross, D. A., & Weiss, H. A. (2019). Use of a mobile application for ebola contact tracing and monitoring in northern sierra leone: A proof-of-concept study. *BMC Infectious Diseases, 19*(1), 1–12.
59. Sacks, J. A., Zehe, E., Redick, C., Bah, A., Cowger, K., Camara, M., Diallo, A., Gigo, A. N. I., Dhillon, R. S., & Liu, A. (2015). Introduction of mobile health tools to support Ebola surveillance and contact tracing in Guinea. *Global Health: Science and Practice, 3*(4), 646–659.
60. Salathe, M., & Jones, J. H. (2010). Dynamics and control of diseases in networks with community structure. *PLoS Computational Biology, 6*(4), e1000736.
61. Shahtori, N. M., Ferdousi, T., Scoglio, C., & Sahneh, F. D. (2018). Quantifying the impact of early-stage contact tracing on controlling Ebola diffusion. *Mathematical Biosciences and Engineering, 15*, 1165–1180.
62. Yulin, W., Cai, W., Li, Z., Tan, W. J., & Hou, X. (2019). Efficient parallel simulation over large-scale social contact networks. *Transactions on Modeling and Computer Simulation*.
63. Yulin, W., Cai, W., Li, Z., Tan, W. J., & Hou, X. (2019). Efficient parallel simulation over large-scale social contact networks. *ACM Transactions on Modeling and Computer Simulation, 12*(1).
64. WHO Ebola Response Team. (2014). Ebola virus disease in West Africa the first 9 months of the epidemic and forward projections. *New England Journal of Medicine, 371*(16), 1481–1495.
65. Anastassopoulou, C., Russo, L., Tsakris, A., & Siettos, C. (2020). Data-based analysis, modelling and forecasting of The Covid-19 outbreak. *Plos ONE, 15*(3), e0230405.
66. Martinez, K., Hart, J. K., Basford, P. J., Bragg, G. M., Ward, T., & Young, D. S. (2017). A geophone wireless sensor network for investigating glacier stick-slip motion. *Computers & Geosciences, 105*, 103–112.
67. Verity, R., Okell, L. C., Dorigatti, I., Winskill, P., Whittaker, C., Imai, N., Cuomo-Dannenburg, G., Thompson, H., Walker, P. G. T., Han, F., Dighe, A., Griffin, J. T., Baguelin, M., Bhatia, S., Boonyasiri, A., Cori, A., Cucunuba, Z., Fitzjohn, R., Gaythorpe, K., Green, W., Hamlet, A., Hinsley, W., Laydon, D., Nedjati-Gilani, G., Riley, S., Van Elsland, S., Volz, E., Wang, H., Wang, Y., Xi, X., Donnelly, C. A., Ghani, A. C., & Ferguson, N. M. (2020). Estimates of the severity of Coronavirus disease 2019: A model-based analysis. *The Lancet Infectious Diseases, 20*(6), 669–677.

68. Dennis, A. M., Pasquale, D. K., Billock, R., Beagle, S., Mobley, V., Cope, A., Kuruc, J., Sebastian, J., Walworth, C., & Leone, P. A. (2018). Integration of contact tracing and phylogenetics in an investigation of acute HIV infection. *Sexually Transmitted Diseases, 45*(4), 222.
69. Zhang, Z., Wang, H., Wang, C., & Fang, H. (2020). Modeling epidemics spreading on social contact networks. *IEEE Transactions on Emerging Topics in Computing, 3*(3), 410–419.
70. Woodley, M.. *Racgp releases covidsafe factsheet*. Royal Australian College of General Practitioners. https://www1.racgp.org.au/newsgp/professional/racgp-releasescovidsafe-factsheet?feed=racgpnewsgparticle. Accessed May 2020.
71. Lin, Q., Zhao, S., Gao, D., Lou, Y., Yang, S., Musa, S. S., Wang, M. H., Cai, Y., Wang, W., Yang, L., & He, D. (2020). A conceptual model for the coronavirus disease 2019 (Covid-19) outbreak in Wuhan, China with individual reaction and governmental action. *International Journal of Infectious Disease, 93*, 211–216.
72. Kwok, K. O., Tang, A., Wei, V. I., Park, W. H., Yeoh, E. K., & Riley, S. (2019). Epidemic models of contact tracing: Systematic review of transmission studies of severe acute respiratory syndrome and middle east respiratory syndrome. *Computational and Structural Biotechnology Journal, 17*, 186–194.
73. Yoneki, E. (2011). Fluphone study: Virtual disease spread using haggle. In *ACM mobile communication workshop challenged network*. ACM.
74. Yoneki, E., & Crowcroft, J. (2014). Epimap: Towards quantifying contact networks and modelling the spread of infections in developing countries. *Ad Hoc Networks, 13*, 83–93.
75. Lin, H., He, N., Ding, Y., Qiu, D., Zhu, W., Liu, X., Zhang, T., & Detels, R. (2012). Tracing sexual contacts of HIV-infected individuals in a rural prefecture, Eastern China. *BMC Public Health, 12*(1), 1–11.

High-Impact Applications of IoT System-Based Metaheuristics

Shaweta Sharma, Aftab Alam, Akhil Sharma, Prateek Singh, Shivang Dhoundiyal, and Aditya Sharma

1 Introduction

The Internet of Things (IoT), a pervasive technology in recent years, has enabled small devices to make decisions, sense their surroundings, and communicate with one another [1]. The IoT network's data volume, rate of change, and variety are all increasing. System mobility, device mobility, and wireless communications all add to the inherent complexity of IoT systems. Despite these challenges, metaheuristic (MH) algorithms serve as the foundation for sophisticated IoT systems used in real-time processes [2]. Because it provides solutions to difficult practical applications such as healthcare, the field of swarm intelligence (SI) on the Internet of Things has grown to become an important area of study. Smart agriculture systems demonstrate the promise of collaborative swarm intelligence-based Internet of Things (IoT) solutions [3]. The COVID-19 pandemic has had a profound impact on the world, and IoT technology has played a significant role in managing its effects. With the help of IoT devices, healthcare providers and governments have been able to improve patient outcomes and reduce the spread of the virus. For example, remote monitoring devices such as wearables and sensors have enabled healthcare providers to monitor the vital signs of COVID-19 patients remotely, reducing the need for in-person visits and minimizing exposure to the virus. Contact tracing apps and Bluetooth beacons have also been used to track the movements of who may have been exposed to the virus, helping to contain its spread [4–6].

Furthermore, Internet of Things systems based on swarm intelligence have been integrated into the construction of critical infrastructure, such as smart cities. This

S. Sharma (✉) · A. Alam · P. Singh · S. Dhoundiyal · A. Sharma
Department of Pharmacy, SMAS, Galgotias University, Greater Noida, Uttar Pradesh, India
e-mail: shaweta.sharma@galgotiasuniversity.edu.in

A. Sharma
RJ College of Pharmacy, Aligarh, Uttar Pradesh, India

A. M. Anter et al. (eds.), *Nature-Inspired Methods for Smart Healthcare Systems and Medical Data*, https://doi.org/10.1007/978-3-031-45952-8_6

chapter provides an overview of the research being conducted on the application of IoT system-based metaheuristics. It discusses the challenges of IoT systems and the potential of metaheuristics algorithms to address those challenges.

1.1 Internet of Things in Various Domains

IoT system-based metaheuristics have several applications in various domains. Here are some of the applications: IoT-based swarm intelligence is critical in supervising healthcare sectors in smart cities. It can quickly access patient data, gather data using sensors, diagnose ailments, and provide other health services. Swarm intelligence can be used for real-time monitoring of patient health status using IoT-based wearable devices [7]. IoT applications based on swarm intelligence have been designed for financial risk management. Particle-swarm optimization-based back-propagation neural networks are used in banking institutions for IoT deployment to assess credit risk and forecast bankruptcy [8]. Smart agriculture is constrained and topologically difficult. Swarm intelligence can be used to plan agricultural paths accurately, automate the gathering, interpretation, and application of data. Several techniques, such as Bellman-held-kar, Ant colony optimization, K-means clustering, and Christofides based on the nearest neighbour, are used for precision agriculture [9]. Maintaining reliability, efficiency, scalability, and security is a difficult issue for smart cities. In these circumstances, SI is a potential method that emphasizes the self-organizing and collective behaviour of dispersed systems. Swarm intelligence can be used for modelling intelligent infrastructures utilising cutting-edge computers, networks, sensors, digital communication, and embedded intelligence [10]. Swarm intelligence can be used for scheduling heterogeneous jobs in cloud computing platforms. Six state-of-the-art metaheuristic methods, such as particle swarm optimization, ant colony optimization, crow search algorithm, genetic algorithm, artificial bee colony algorithm, and penguin swarm optimization algorithm, have been quantitatively analysed in terms of scheduling parameters like resource utilization cost and make span [11–13].

1.2 Algorithms in IoT in Metaheuristcs

Algorithms are heavily used in IoT systems because they allow connected devices to perform real-time data processing and analysis. Furthermore, metaheuristics are a type of algorithm that may be especially useful in Internet of Things contexts. Metaheuristics, as optimisation algorithms, are useful for solving difficult problems, such as those with multiple goals or constraints. Algorithms like these assist Internet-of-Things systems in maximising the value of their resources such as technology, power, and data. Ant colony optimisation (ACO), a metaheuristic method, have applications in IoT [14]. The ACO algorithm is a metaheuristic approach to problem solving that draws inspiration from ant colony behaviour. In ACO, an artificial ant colony calculates the shortest path between two network nodes. This algorithm can be used in IoT systems to improve the efficiency and lower the power consumption of data packet routing between devices. Another type of metaheuristic algorithm that can be applied to IoT systems is the Genetic Algorithm (GA) [15]. GA is a natural selection-inspired algorithm for optimising data structures. GA breeds the best of the best solutions from a population of candidates to perpetuity over many generations. GA can be used in IoT systems to distribute resources more efficiently such as processing power and storage space among connected devices. To summarise, metaheuristic algorithms can be an effective method for optimising Internet of Things (IoT) systems, particularly when dealing with complex problems involving multiple competing goals and constraints. In IoT systems, these algorithms enable more effective resource allocation, faster data processing, and lower energy consumption.

1.3 Swarm Intelligence

Swarm intelligence is a collective response displayed by self-organized entities in decentralised systems. Swarm intelligence algorithms enables organised and effective collaboration between devices that can improve the performance of IoT systems. Metaheuristics are optimisation methods that can be used to address multiple goals and constraints at the same time. The integration of swarm intelligence and metaheuristics allows for strong, dynamic, and constantly improving Internet of Things (IoT) systems. Figure 1 depicts the integration of swarm intelligence in IoT

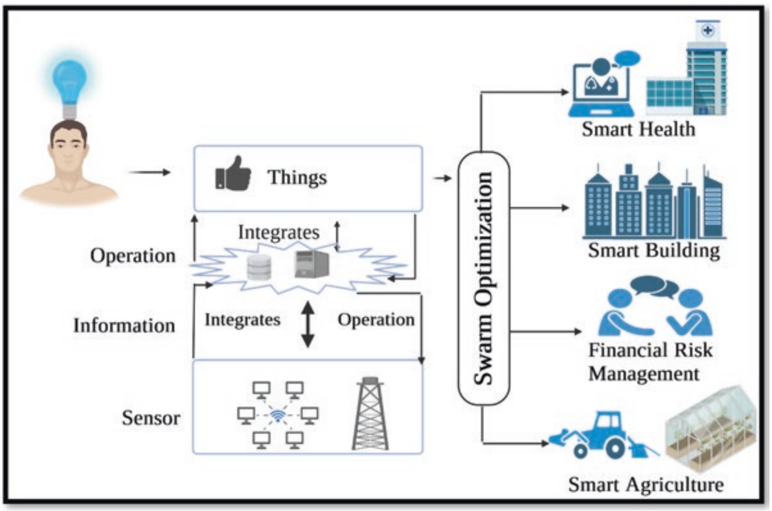


Fig. 1 Integration of Swarm intelligence algorithms in IoT systems

systems. Particle Swarm Optimization (PSO) is one example of a metaheuristic algorithm that takes its cues from animal groups such as bird colonies or fish schools [16]. PSO employs a population of particles to solve problems. The PSO algorithm, which employs swarm intelligence, can quickly and efficiently explore a large solution space, allowing IoT systems to optimise performance. Another example of a swarm intelligence-based metaheuristic algorithm is the Ant Colony Optimization (ACO) algorithm, which was inspired by ant colony behaviour. ACO employs a colony of artificial ants to determine the shortest path between two points in a network. The ACO algorithm uses swarm intelligence to find the most efficient path through a complex network, allowing IoT systems to optimise routing and communication [17]. Finally, metaheuristic algorithms based on swarm intelligence can be a powerful tool for optimising IoT systems. By enabling devices to collaborate in a coordinated and efficient manner, these algorithms can help IoT systems operate more efficiently, with lower energy consumption and faster data processing.

1.4 Applications of IoT System-Based Metaheuristics

IoT-based metaheuristics have the potential to transform various fields by enabling more efficient and effective decision-making processes. One of the most significant applications of IoT-based metaheuristics is in smart manufacturing. By leveraging real-time data from IoT sensors, metaheuristics can optimize various aspects of the manufacturing process, such as scheduling, production planning, and quality control [18]. For example, by analysing data on the factory floor, metaheuristics can identify bottlenecks in the production process and optimize the scheduling of machines and workers, reducing downtime and improving efficiency. Another important application of IoT-based metaheuristics is in energy management. By analysing real-time data on energy consumption from IoT sensors, metaheuristics can optimize the energy usage of buildings, factories, and other facilities. For example, by identifying energy-intensive processes and optimizing the scheduling of those processes, metaheuristics can reduce energy consumption, lower costs, and improve sustainability. IoT-based metaheuristics can also be applied to environmental monitoring. By leveraging data from IoT sensors, metaheuristics can analyse environmental conditions, such as air quality, water quality, and weather patterns, and provide real-time insights to decision-makers [19]. For example, by analysing data from air quality sensors, metaheuristics can identify areas with high levels of pollution and optimize the deployment of resources, such as air purifiers, to improve air quality. Overall, the application of IoT-based metaheuristics has the potential to significantly impact various fields by enabling more efficient and effective decision-making processes. By leveraging real-time data from IoT sensors, metaheuristics can optimize processes, reduce costs, improve sustainability, and provide real-time insights to decision-makers. Taxonomical classification of metaheuristics algorithms has been illustrated in Fig. 2.

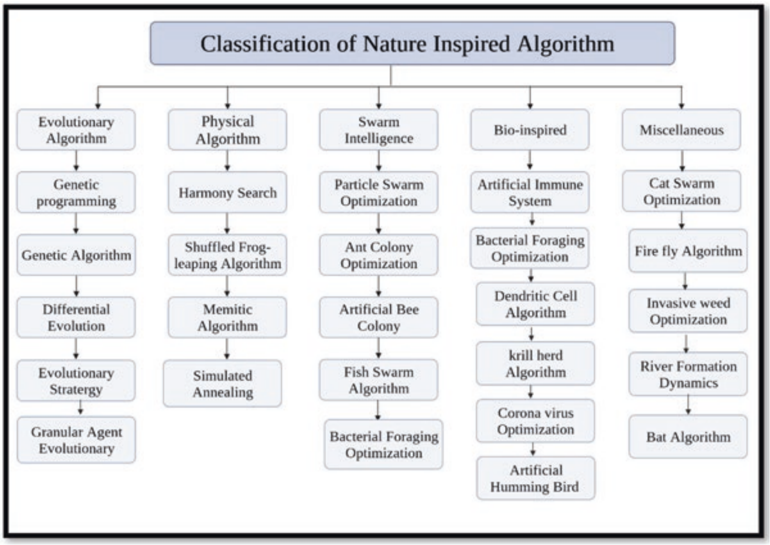


Fig. 2 Taxonomy of Meta-Heuristics Algorithms

2 Metaheuristic Methods

Metaheuristic methods are a type of optimisation algorithm used to solve complex problems that have multiple objectives and constraints. Unlike traditional optimisation algorithms, which use a single approach to solve a problem, metaheuristic methods combine multiple approaches to find the best solution. There are numerous metaheuristic methods, each with its own set of advantages and disadvantages. Among the most popular metaheuristic methods are discussed below. Table 1 classifies algorithms according to their applications.

2.1 Genetic Algorithms

Metaheuristic optimisation techniques, such as genetic algorithms (GAs), can be applied to the complex optimisation problems of IoT. GAs generates a population of potential solutions and then improve them repeatedly through selection, crossover, and mutation, just as the evolutionary processes that inspired the GA. Gas can optimise many aspects of an Internet of Things system, including power consumption, network activity, data transfer, and sensor positioning [20].

Table 1 Application-based categorization of algorithms

Year	Algorithm	Category	Application areas
1983	Simulated Annealing (SA)	Physics based	Multivariate or combinatorial optimization
1989	Memetic Algorithm (MA)	Physics based	Optimization problem for chemical process
1992	Genetic algorithms (GA)	Evolutionary based	High quality solution to optimization and search problems
2001	Bacterial foraging optimization (BFO)	Physics based	Travelling salesman problem, academic optimization, pipe network
2002	Bacterial foraging optimization (BFO)	Bio inspired	Machine learning, pattern recognition, neural network problems
2004	Particle swarm optimization (PSO)	Bird-based	Data clustering, data mining, prediction, scheduling
2005	Ant Colony Optimization (ACO)	Insect-based	Heart disease prediction, clustering, classification, scheduling
2006	Evolutionary Strategy (ES)	Evolutionary based	Real number vector problem
2006	Shuffled frog-leaping algorithm (SFL)	Physics based	Combinatorial optimization
2007	Dendritic Cell algorithm	Bio inspired	Robotic classifier, dendritic cell population tuning
2008	Differential Evolution	Evolutionary based	Numerical optimization problem
2009	Artificial Bee Colony (ABC)	Insect-based	Medical classification, image optimization, routing
2009	Fish swarm algorithm (FSA)	Amphibious-based	Robot control optimization, routing, neural network training
2011	Grey wolf optimizer (GWO)	Wild-based	Satellite image segmentation, clustering, multi-objective problem
2012	Artificial Immune system (AIS)	Bio inspired	Damaged detection in structural health monitoring
2012	Differential Evolution (DE)	Evolutionary based	Model Order Reduction for Single Input, Single Output Systems
2017	Genetic Programming (GE)	Evolutionary based	Heuristic search, hill climbing, multi-expression programming
2018	Beer froth artificial bee colony (ABC)	Insect-based	Job scheduling, flow scheduling, batch scheduling
2019	Spider monkey optimization	Bio inspired	Enhancing the exploitation and exploration of SMO algorithm
2020	Coronavirus optimization	Bio inspired	Identifying how COVID-19 affects people's health
2021	Horse herd optimization algorithm	Bio inspired	Intimating horse behaviour at different ages
2022	Artificial hummingbird algorithm	Artificial hummingbird algorithm	Defining intelligent foraging techniques

2.2 Particle Swarm Optimization

Particle swarm optimisation (PSO) is another metaheuristic optimisation method that can be used in IoT systems and is particularly effective at solving complex optimisation problems. Particle swarm optimisation mimics the cooperative behaviour of animal groups such as flocks of birds or schools of fish by instructing each "particle" to seek the best solution. PSO can be used in an IoT system to optimise parameters such as energy consumption, data transmission, and sensor placement. For example, in a smart city IoT system, PSO can be used to optimise traffic signal timing to reduce traffic congestion and improve traffic flow [21].

2.3 Simulated Annealing

Simulated annealing (SA) is another metaheuristic optimisation technique that can be used in IoT systems to resolve complex optimisation problems. SA was inspired by the metallurgical annealing process, which involves gently cooling a material to eliminate flaws and improve its properties. In an IoT system, SA can be used to optimise parameters such as energy consumption, network traffic, data transmission, and sensor placement. For example, in a wireless sensor network, SA can be used to optimise data packet routing to reduce energy consumption while ensuring that data reaches its destination [22].

2.4 Ant Colony Optimization

Metaheuristic optimisation techniques such as Ant Colony Optimization (ACO) can be used to solve difficult optimisation problems in IoT environments. ACO was inspired by how ant colonies behave when attempting to locate food sources near their colony. ACO can be used to optimise parameters such as energy consumption, network traffic, data transmission, and sensor placement in an IoT system. In a smart agriculture IoT system, for example, ACO can be used to optimise sensor placement to monitor soil moisture levels and adjust irrigation accordingly [23].

2.5 Artificial Bee Colony Algorithm

IoT systems may use the Artificial Bee Colony (ABC) Algorithm, a type of metaheuristic optimisation, to solve difficult optimisation problems. The ABC method was inspired by bee foraging habits to locate the best food source. In a smart building IoT system, for example, the ABC algorithm can be used to optimise the placement of temperature sensors to maintain the desired temperature while reducing energy consumption. The ABC algorithm has been used successfully in IoT systems such as energy management and wireless sensor network optimisation. However, as with other metaheuristic algorithms, there are challenges in using ABC in IoT systems, such as the need for efficient communication protocols, data security, and privacy concerns [24].

2.6 Crow-Search Algorithm

The Crow Search Algorithm (CSA) is one such metaheuristic optimisation technique that can be used in IoT systems to solve difficult optimisation problems. Crows' cooperative hunting behaviour was the inspiration for CSA. Renewable energy sources such as solar panels and wind turbines can be strategically placed in an IoT smart grid system with the help of CSA to improve energy production and reduce costs. The CSA algorithm has been used successfully in IoT systems such as wireless sensor network optimisation and smart grid management. However, as with other metaheuristic algorithms, there are challenges in using CSA in IoT systems, such as the need for efficient communication protocols, data security, and privacy concerns [25].

2.7 Upgraded Grey-Wolf Optimizer Method

The Grey Wolf Optimizer (GWO) is a modification of the Grey Wolf Optimizer (UGWO), which is a metaheuristic optimisation technique inspired by grey wolf hunting behaviour. UGWO can be used in IoT systems to solve complex optimisation problems such as energy management, wireless sensor network optimisation, and smart city planning. The UGWO algorithm improves the GWO algorithm's exploration and exploitation capabilities by introducing an adaptive search mechanism. The adaptive search mechanism enables UGWO to dynamically adjust the search process to better fit the problem at hand [26].

2.8 Merkle-Hellman Knapsack Cryptosystem

Merkle-Hellman Knapsack Cryptosystem (MHKC) is a public-key encryption algorithm that, unlike metaheuristic methods, is based on a mathematical problem rather than an optimisation problem. However, some metaheuristic methods can be used in conjunction with the MHKC algorithm to increase its security. The Genetic Algorithm is one such method that can be used to optimise the selection of the algorithm's super-increasing knapsacks [27]. It is possible to find a set of knapsacks that satisfy the super-increasing property by using a GA, but it is also more difficult to solve using a subset sum problem. This can improve the MHKC algorithm's resistance to attacks. Simulated Annealing is another metaheuristic method that can be used in conjunction with MHKC [28]. SA can be used to find an optimal subset of the knapsack that produces the same sum as the encrypted message. It is possible to improve the security of the MHKC algorithm by using SA to find a subset that is more difficult to solve using a subset sum problem.

Overall, metaheuristic methods are an effective tool for solving complex problems in a wide range of fields, including engineering, finance, and computer science. These algorithms can find the best solution to a problem more efficiently and effectively than traditional optimisation methods because they use a combination of different approaches.

3 Conclusion

In conclusion, IoT system-based metaheuristics provide a powerful approach for optimising complex problems across a wide range of applications. These include energy management, smart city planning, wireless sensor network optimisation, and many others. Each metaheuristic algorithm has advantages and disadvantages, and the best algorithm to use depends on the problem at hand. It is possible to significantly improve system efficiency, cost-effectiveness, and overall performance by leveraging the capabilities of IoT systems and metaheuristic algorithms. As IoT technology advances, the use of metaheuristic algorithms will become increasingly important for addressing the complex challenges of tomorrow's smart systems.

References

1. Porkodi, R., & Bhuvaneswari, V. (2014, March 6). The internet of things (IoT) applications and communication enabling technology standards: An overview. In *2014 international conference on intelligent computing applications* (pp. 324–329). IEEE.
2. Anter, A. M., Mohamed, A. W., Zhang, M., & Zhang, Z. (2023). A robust intelligence regression model for monitoring Parkinson's disease based on speech signals. *Future Generation Computer Systems, 147*, 316–327.
3. Anter, A. M., Abd Elaziz, M., & Zhang, Z. (2022). Real-time epileptic seizure recognition using Bayesian genetic whale optimizer and adaptive machine learning. *Future Generation Computer Systems, 127*, 426–434.
4. Basha, S. H., Anter, A. M., Hassanien, A. E., & Abdalla, A. (2023). Hybrid intelligent model for classifying chest X-ray images of COVID-19 patients using genetic algorithm and neutrosophic logic. *Soft Computing, 27*(6), 3427–3442.
5. Anter, A. M., Oliva, D., Thakare, A., & Zhang, Z. (2021). AFCM-LSMA: New intelligent model based on Lévy slime mould algorithm and adaptive fuzzy C-means for identification of COVID-19 infection from chest X-ray images. *Advanced Engineering Informatics, 49*, 101317.

6. Al-Shourbaji, I., Kachare, P. H., Abualigah, L., Abdelhag, M. E., Elnaim, B., Anter, A. M., & Gandomi, A. H. (2022). A deep batch normalized convolution approach for improving COVID-19 detection from chest X-ray images. *Pathogens, 12*(1), 17.
7. Chandy, A. (2019, September). A review on IoT based medical imaging technology for healthcare applications. *Journal of Innovative Image Processing (JIIP), 1*(01), 51–60.
8. Wang, R., Yu, C., & Wang, J. (2019, August 1). Construction of supply chain financial risk management mode based on internet of things. *IEEE Access, 7*, 110323–110332.
9. Ray, P. P. (2017, January 1). Internet of things for smart agriculture: Technologies, practices and future direction. *Journal of Ambient Intelligence and Smart Environments, 9*(4), 395–420.
10. Mehmood, Y., Ahmad, F., Yaqoob, I., Adnane, A., Imran, M., & Guizani, S. (2017, September 8). Internet-of-things-based smart cities: Recent advances and challenges. *IEEE Communications Magazine, 55*(9), 16–24.
11. Anter, A. M., & Zhang, Z. (2023). RLWOA-SOFL: A new learning model-based reinforcement swarm intelligence and self-organizing deep fuzzy rules for fMRI pain decoding. *IEEE Transactions on Affective Computing*. https://doi.org/10.1109/TAFFC.2023.3285997
12. Anter, A. M., & Zhang, Z. (2020). E-health Parkinson disease diagnosis in smart home based on hybrid intelligence optimization model. In *Proceedings of the International Conference on Advanced Intelligent Systems and Informatics 2019* (pp. 156–165). Springer International Publishing.
13. Anter, A. M., & Zhang, Z. (2019). Adaptive neuro-fuzzy inference system-based chaotic swarm intelligence hybrid model for recognition of mild cognitive impairment from resting-state fMRI. In *International workshop on predictive intelligence in medicine* (pp. 23–33). Springer International Publishing.
14. Dorigo, M., & Stützle, T. (2019). *Ant colony optimization: Overview and recent advances*. Springer International Publishing.
15. Anter, A. M., Elaziz, M. A., & Zhang, Z. (2022). Real-time epileptic seizure recognition using Bayesian genetic whale optimizer and adaptive machine learning. *Future Generation Computer Systems, 127*, 426–434.
16. Okwu, M. O., Tartibu, L. K., Okwu, M. O., & Tartibu, L. K. (2021). Particle swarm optimisation. In *Metaheuristic optimization: nature-inspired algorithms swarm and computational intelligence, theory and applications* (pp. 5–13). Springer Nature.
17. Shin, C., & Lee, M. (2020, September 10). Swarm-intelligence-centric routing algorithm for wireless sensor networks. *Sensors, 20*(18), 5164.
18. Cohen, Y., Naseraldin, H., Chaudhuri, A., & Pilati, F. (2019, December). Assembly systems in Industry 4.0 era: a road map to understand Assembly 4.0. *The International Journal of Advanced Manufacturing Technology, 105*, 4037–4054.
19. Karimi, Y., Haghi Kashani, M., Akbari, M., & Mahdipour, E. (2021, November 10). Leveraging big data in smart cities: A systematic review. *Concurrency and Computation: Practice and Experience, 33*(21), e6379.
20. Kramer, O., & Kramer, O. (2017). *Genetic algorithms*. Springer International Publishing.
21. Wang, D., Tan, D., & Liu, L. (2018, January). Particle swarm optimization algorithm: An overview. *Soft Computing, 22*, 387–408.
22. Delahaye, D., Chaimatanan, S., & Mongeau, M. (2019). Simulated annealing: From basics to applications. In *Handbook of metaheuristics* (pp. 1–35). Springer.
23. Dorigo, M., & Socha, K. (2018, May 15). An introduction to ant colony optimization. In *Handbook of approximation algorithms and metaheuristics* (2nd ed., pp. 395–408). Chapman and Hall/CRC.
24. Li, X., & Yang, G. (2016, April 1). Artificial bee colony algorithm with memory. *Applied Soft Computing, 41*, 362–372.
25. Meraihi, Y., Gabis, A. B., Ramdane-Cherif, A., & Acheli, D. (2021, April). A comprehensive survey of crow search algorithm and its applications. *Artificial Intelligence Review, 54*(4), 2669–2716.

26. Emary, E., Zawbaa, H. M., & Hassanien, A. E. (2016, January 8). Binary grey wolf optimization approaches for feature selection. *Neurocomputing, 172*, 371–381.
27. Kantour, N., & Bouroubi, S. (2020, February). Cryptanalysis of merkle-hellman cipher using parallel genetic algorithm. *Mobile Networks and Applications, 25*(1), 211–222.
28. Abdel-Basset, M., Mohamed, R., & ELkomy, O. M. (2022, April 1). Knapsack cipher-based metaheuristic optimization algorithms for cryptanalysis in blockchain-enabled internet of things systems. *Ad Hoc Networks, 128*, 102798.

IoT-Based eHealth Solutions for Aging with Special Emphasis on Aging-Related Inflammatory Diseases: Prospects and Challenges

Pritha Chakraborty, Shankar Dey, Ritwik Patra, Nabarun Chandra Das, and Suprabhat Mukherjee

Abbreviations

AI	Artificial intelligence
BAN	Body area networks
BPM	Beats per minute
DAMP	damage-associated molecular patterns
ECG	Echocardiogram
EEG	Electroencephalogram
EHR	Electronic health records
EMG	Electromyography
EMR	Electronic medical reports
EPC	Electronic Product Code
GPRS	General packet radio services
GPS	Global Positioning System
IaaS	infrastructure as a service
ICU	Intensive care unit
IoT	Internet of Things
IT	Information technology
LED	Light-emitting diode
MCI	mild cognitive impairment
MEMS	Micro-electro-mechanical systems
MIT	Massachusetts Institute of Technology

P. Chakraborty · R. Patra · N. C. Das · S. Mukherjee (✉)
Integrative Biochemistry & Immunology Laboratory, Department of Animal Science, Kazi Nazrul University, Asansol, West Bengal, India
e-mail: suprabhat.mukherjee@knu.ac.in

S. Dey
Department of Obstetrics and Gynaecology, ESI Hospital, Asansol, West Bengal, India

A. M. Anter et al. (eds.), *Nature-Inspired Methods for Smart Healthcare Systems and Medical Data*, https://doi.org/10.1007/978-3-031-45952-8_7

NLP	Natural language processing
PaaS	Platform as a Service
RFID	Radio Frequency Identification
SaaS	Software as a Service
SDN	Software Defined Network
SJSA	St. John Sepsis Agent
SMS	Short messaging service
WSN	Wireless sensor networking.

1 Introduction

Over a long period, there has been an ongoing debate related to the Internet of Things (IoT) integrating into the healthcare system transforming into e-health. IoT can be defined as the network of physical objects combined with technology that enables interaction with the environment, offers autonomous communication and sensing which has enabled millions of sensors and actuators to connect with the internet via access networks resulting in technologies including wireless sensor networks, infrared, real-time and semantic web services, Radio Frequency Identification (RFID), Global Positioning System (GPS), wearable and wireless sensor networks around smart cities in the field of healthcare, home monitoring, transportation system, and traffic management [1]. IoT is based on 6Cs: convergence, communication, connectivity, computing, collections, and content which helps the individuals in establishing a connection with other individuals using networking services [2].

The Internet has been born in the year 1989, and the term "Internet of Things" was first coined by Kevin Ashton when he described the IoT as technology using the RFID tags to improve the quality of the healthcare system [3]. Development of St. John Sepsis Agent (SJSA) is one of the classic examples which has reduced the effects of sepsis and increases the chances of survival using machine learning algorithms [4]. The basic perception of the IoT has been changing with the development of some ground-breaking technologies such as insulin delivery devices, glucose monitoring devices, smart wearable, home monitoring devices; inhalers are named few proving its high potential and benefits.

With the world population growing, life expectancy of geriatric population has been increased globally and suffering from age-related diseases such as type 2 diabetes, mild cognitive impairment (MCI), neurological disorders, atherosclerosis, cardiovascular diseases to establish an "Inflamm-ageing" relationship between aging and low-grade chronic inflammation and passing away silently without getting hospice care [5, 45]. According to current reports of 2019, there are about 702.9 million people above the age of 60 and 53.9 million above 80 which will probably increase up to 120% soon and among them, around 70% of elderly are dependent on caregivers but for the informal caregivers, it has been a huge challenge [6]. This group requires medical health care to sustain their independent lifestyle. Increased medical expenses and deteriorating medical conditions urges the advancement of technology. IoT has promising objectives which provide real-time health monitoring, online consultations, reducing the burden on the caregivers, supporting the elders with their medical decisions given by their medical professionals continuously [7].

In this chapter, we will be discussing IoT-related e-health solutions that could provide possible solutions in the geriatric healthcare system suffering from age-related inflammatory diseases based on current research papers and review papers. The key component of an IoT-based geriatric healthcare system is based on a cloud computing system that collects the data and processes it for further analysis to enhance their quality of living. We will also discuss managing and combating well-known aging-related diseased conditions including the global pandemic of COVID-19 by using IoT-based techniques. But with the advancement of technologies, there will be limitations posing as a challenge to the developers which will require the need for further attention and research.

2 Learning the IoT-Based Health Care System

IoT has revolutionized the medical field and the healthcare system as they help the doctors to reach out to their patients through the internet by using health monitoring devices that have bought the patient and the doctor closer by analyzing the patients individually and efficiently [8].

2.1 Development of IoT in Healthcare

Ashton and Brock together founded Auto-ID which is an identification technology that can reduce the possibility of error, automation which will result in increased efficiency at the Massachusetts Institute of Technology (MIT) [9]. Auto-Id leads to the development of the Electronic Product Code (EPC) network using microchips and RIFD in the year 2003, which can track objects from their movement that has given IoT commercial means on the global platform and has started a new era of information technology (IT) [10]. The IoT-based healthcare system has been able to perform diagnosis, monitoring, and surgeries remotely by using the resources and putting them up into a networking system connecting all the activities efficiently and extend the services from hospitals to homes with the help of wireless technologies and a centralized data system [11, 12]. IoT in the healthcare system is based on three parameters i.e., the Master which includes the doctors, nurses, and the patients using their specific devices with end-to-end encryption, the Server form the central part and is capable of generating prescriptions, managing the database, analyzing the data, and things will include the physical objects such as the patient resources and the wireless connectivity hence every parameter has been verified by different pioneered exoskeleton applications [10].

Healthcare data includes electronic health records (EHR) or Electronic medical report (EMR) containing the patient's medical history, prescriptions, diagnostic reports, and pathological data which needs to be managed over IoT networks so that it can be accessed instantly and securely. It fails over traditional hardware and software because of its huge volume and diversity [13, 46].

Cloud computing provides the solution to the above traditional challenge by providing configured computing resources which can have flexibility and availability at low cost in healthcare along with the ability to collect the data and store them in huge capacity directly to the cloud with the use of sensors attached to the medical equipment which can be accessed by the medical professional securely [14]. The cloud platform is used in three different ways i.e., the infrastructure as a service (IaaS) which includes the hardware, physical devices, and storage over the internet; Software as a Service (SaaS) which includes the software and other related applications and Platform as a Service (PaaS) which uses programming languages and other tools to deploy the applications [15].

From cloud computing, healthcare experts have decided to update their existing technology and create long-term solutions for patients suffering from several diseases. The data from the EHRs are accessed to integrate into natural language processing (NLP) to form the SmartHealth-stimulated hybrid cloud [16, 47]. Nowadays, 5G mobile community cloud networks have arrived that have enabled intelligent data analytics powered by data mining and machine learning. It also has connected the patients and the doctors remotely. The therapeutic mobile network community has helped the patients to prevent anxiety, depression, and psychological disorders, to connect with support groups and share their experience throughout the world, and become aware of various rare diseases. Public health monitoring globally and locally has provided a possible prediction of the risk of infection such in cases of Covid-19 [17].

2.2 Application of IoT in Healthcare

IoT has been applied to various fields due to the development and advancement of micro-electro-mechanical systems (MEMS), digital electronics, and wireless communications have provided us with miniature devices having nodes to interconnect and forms wireless sensor networking systems (WSN) [18, 48]. The Healthcare industry follows the principle "the right care for the right person at the right time" and thus focuses on detection, diagnosis, prevention, and homecare to provide integrated well-being to the patients [19]. Over 100 years, patients are under surveillance as a part of critical care by monitoring vital signs such as temperature, pulse, blood pressure, respiratory rate, and oxygen saturation which can be achieved using non-invasive monitoring systems via sensors that are attached to the patients for close monitoring and is collecting data about their physiological state and are getting analyzed and stored in the cloud system [20, 21]. Critical care monitoring includes hardware components such as the LM35 series consisting of the coordinated and integrated circuit of temperature sensors that has output values in Celsius directly which is more reliable than thermostats since it does not get affected due to oxidation [22]. For monitoring the heart status, ECG electrodes are attached to the chest to obtain the PR and QT intervals using AD8232 sensor whereas heart rate sensors give beats per minute (BPM) when the device is put on the finger and the LED flashes with the beat in unison followed by blood pressure monitoring which is also measured using sensory device automatically in which the cuff is worn over the arm and the systolic, diastolic and pulse readings are obtained [22, 23].

Remote monitoring or Telehealth is also the part of eHealth in which the activities of the individuals can be tracked on-site and they do not have to visit the hospital resulting in reduced time and expenditure. Healthcare professionals can perform efficient analysis using the computer powered by wireless technology thus improving the patient's quality of living at ease [24].

Mobile health or mHealth is another part of the eHealth solutions developing in the healthcare system which includes mobile devices such as wireless devices, mobile phones, portable monitoring devices, and other personal digital assistants that can collect real-time data from the patients and upload it to the cloud or the server. mHealth can be used to generate awareness using the short messaging services (SMS), voice messages, educate individuals regarding other outbreaks or disease management using the mobile telecommunication over networks (3G or 4G), general packet radio services (GPRS), Bluetooth and GPS are used for remote monitoring, real-time data collection, outbreak tracking and telemedicine [25].

2.3 Technical Framework of IoT in Healthcare

IoT is reshaping the healthcare system tremendously since previously most of the physical objects used to work on traditional networking system but now, with the development of sensors in devices, critical healthcare monitoring, and remote monitoring has been increasing day by day. IoT objects are tagged with RFID for their identification and can measure the acceleration of the body, detection of nearby objects, and body positioning. There are majorly three domains of IoT technology i.e., application, security, and efficiency [26]. Devices based on IoT have 32-bit microcontroller developed by Intel, ARM, and Renesas units which are small in size and have a high capacity of processing and are capable of data acquisition, communication, up-gradation of firmware and security protocols [27]. The structure of IoT is based on the following parts:

- Collection of Data: A device compiled with sensors is generally attached to the person to collect the data and monitor the health state of the patient. Values from the sensors are digitalized and by using wireless technologies the data is communicated further.
- Connecting Gateway: It performs a crucial role in connecting the local processing units with the backend facilitator remotely using IPv4 or IPv6 and acts as a middle entity for the remote IP network and the local sensor networking system [28].

- Backend Service Provider: Storage services are required and are a crucial part of the IoT system which is required to perform analytics and crucial decision making. The backend service provider also manages the security, remote consultancy, and application development [29].
- Application access: The final component of the IoT framework is the access system which can be installed onto any smart device or computer.

There are three layers proposed for the complete functionality of the system:

- Sensor layer: This layer is designed for the proper functioning of the data accumulation by monitoring the patient's health condition using sensor-embedded devices which perform communication, use bar codes and RFIDs for tagging, and IoT gateway for processing of the data further to the backend facilitator via IP format.
- Network access layer: This layer provides connectivity between the IoT gateway and the backend facilitator along with an interface for the sensor-embedded devices to connect with the backend facilitator using internet services and 3G/4G technologies provides the connectivity
- Service access layer: This layer is designed for medical professionals to access health information from the cloud storage uploaded by smart health applications installed on computers and smart devices using specialized database and application protocols hence providing a front-end interface.

The technical framework of IoT architecture in the healthcare system is schematized in Fig. 1. It includes different layers of technology communicating with devices and cloud storage to enhance productivity and efficiency.

3 Iot-Based Health Care Network for Elderly Disease Prediction

In developing countries, the demographic status has been a challenge to the healthcare system since the chronic illness, aging-related inflammatory conditions; neurological disorders have been a rising issue. eHealth provides a solution to this concern but on the contrary, elder persons are more resistant to adopting new technologies and suffer from anxiety due to decreasing cognitive resources and complex technology [30, 31]. For all living organisms, the immune system is the important defence system for survival. For the first line of defence, acute inflammation takes place within the individual against any foreign antigen encounter resulting in various immune cell components participating to resolve the acute inflammation and restore homeostasis. But, if the acute stage of inflammation continues to persist it results in chronic inflammation manifesting accumulation of immune cells involving in various signalling pathways. Aging is one of the natural processes causing deterioration of the immunological defence system including dysregulation of proinflammatory cytokines and chemokines and exposing the individual to chronic inflammatory

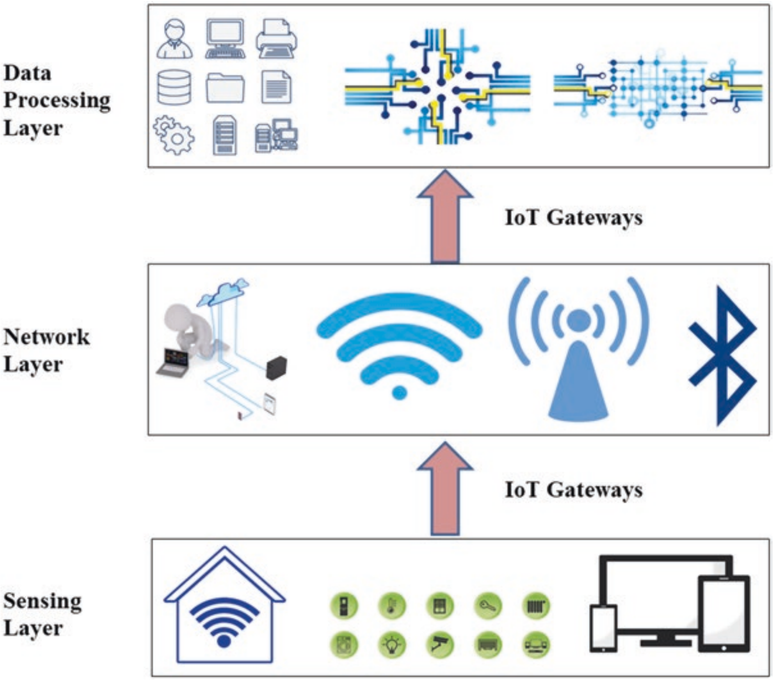


Fig. 1 IoT architecture consisting of three layers setting up the backbone of today's efficient working of healthcare system

conditions [32]. The immune system of the aged person declines due to two phenomenons named immunosenescence in which the immune functions start to decline which deteriorates signalling pathways, pathogen recognition, and clearance. Another phenomenon is called inflammation which is caused due to aging resulting in systemic inflammation due to overactive and ineffective surveillance system of the immune cells [33]. Elderly individuals are prone to metabolic disorders such as obesity, insulin resistance, type 2 diabetes, gout, and fatty liver diseases which are caused due to inflammation within the sites of muscles, pancreas, liver, and adipose tissue interrupting normal tissue functioning [34]. Cardiovascular diseases including atherosclerosis, coronary heart disease, cardiomyopathy, peripheral vascular diseases, dyslipidemia is also prominent aging-related disorder causing high morbidity among the elders in which thrombotic complications are promoted due to TLR$ signalling, damage-associated molecular patterns (DAMPs), oxidative stress and autophagy resulting in cardiac dysfunction [35]. The chronic inflammatory condition can affect the neurological response of the elders causing neurodegenerative disorders such as dementia, Alzheimer's, Parkinson's diseases due to the microglial cells, TLRs, and other cytokines affecting the brain cells resulting in cognitive decline and progression of neurological disorders [36]. Cancer also acts as a contributing factor to elderly morbidity and mortality in which NF-κB plays an important role in the progression of cancer [37]. Lastly, a new threat to the elder population has emerged in the year 2019 by the name of Covid-19 which is affecting the population irrespective of age and gender but the elders possess a greater risk of getting affected due to the presence of above mentioned co-morbid conditions [33].

3.1 Inflammatory Disease Prediction Due to Aging by IoT System

The first requirement of elderly care is to monitor their physical state of health and the current advancement of IT and the development of IoT devices has made it easier for caregivers and medical professionals at a low cost. Hospitals are now modified with newly developed intensive care units (ICU) for assessment of the elders. Healthcare professionals are provided with wireless stethoscopes to monitor the heart rate and breathing sounds, biochemical and physical data collection has been improved which helps in detection as well as diagnosis. All the data are stored digitally in a cloud system so that it can be accessed by the authorized personnel as per the requirement to work on the treatment protocols. Wireless protocol of telemedicine has been followed since 2001 by Gandsas and Montgomery, to obtain physiological data from the patients [38]. RFID micro tags such as VeriChip are attached to the patient's body whereas the readers are installed into the robots which can conduct drug distribution to the elders correctly [39].

3.2 Possible Solutions for the Elders Using IoT System

Chronic disorders of the elders influence the lives of the patients along with their families and society. Remote monitoring has been an utmost possible solution for the elders which can provide them with independent living. Thus, IoT can provide wearable technology to provide state-of-the-art care to patients suffering from various chronic disorders as explained in Fig. 2.

- IoT has been in use for ambient assisted living, providing therapy and entertainment to the elders, indulging them in social activities along with increasing their communication, monitoring their health status togetherly improve their quality of living.
- IoT has been in use for human activity recognition in which their routine lifestyle can be monitored continuously and are being able to detect their abnormal conditions
- It has a wider application in the field of mental health. With the assisted living and development of deep machine learning and artificial neural networking system, it is possible to detect and treat the mental illness of the elders. E.g.,

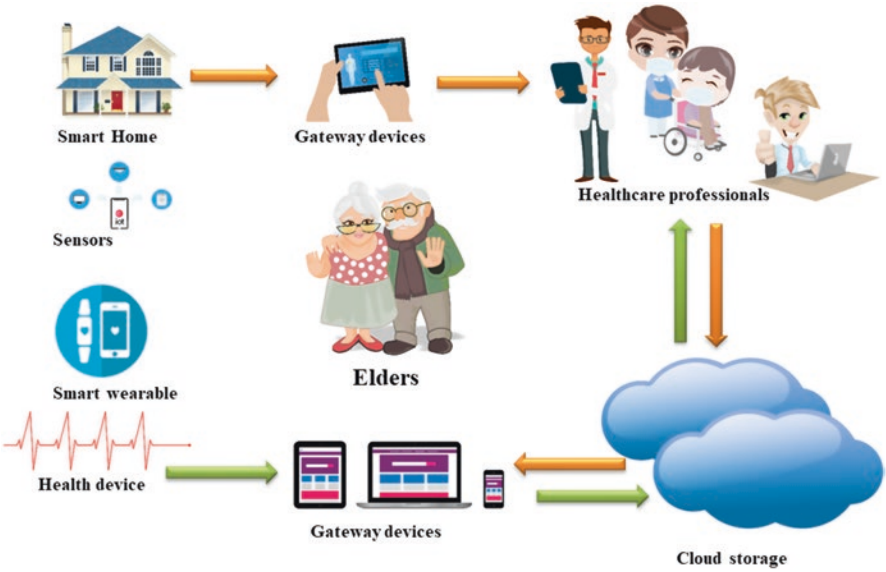


Fig. 2 IoT solutions for health management of elders using smart devices, IoT gateways and cloud computing

dementia can be managed using a smart robot which can be programmed using different algorithms, to reduce the stress level and loneliness of the patients and enhance their activity level [7].

- With age, a patient also suffers from low bone density causing difficulty in free movement resulting in various movement disorders. IoT has been applied in the early detection of movement disorders by using wearable sensors such as patients with Parkinson's disease who suffers from gait problems that can be managed using wearable technology [40].
- Due to aging, elders tend to fall thus IoT has come up with early fall detection and fall risk management to provide emergency response.
- IoTs are developed for rehabilitation as well as a preventive measure for several diseases including Covid-19. In this critical situation, IoT has been used for telehealth consultation, disease monitoring, surveillance, feedback on different parameters such as blood pressure, $SPO_{2,}$ body temperature [25].
- Elders living at their home or in hospice care can use mobile health technology as an aid to receive faster healthcare services in no time.
- IoT has been useful for not only the elders but for the caregivers and the medical professional too. They can access patient's medical records in real-time and give their feedback using wireless technology, the internet, and wireless communication. They can detect emergencies and provide necessary action for the prevention and protection of the elders [41].

There are many applications of IoT currently in use for the elders mentioned in Table 1.

Table 1 Different types of IoT applications currently in use for healthcare management of the elders

Device	Function
Ambient assisted living with robots such as EVA	It has been able to conduct therapies with elders suffering from dementia, Alzheimer's.
Chest sensors such as Cardio Mem, Dexcom	Body glucose, temperature, arrhythmia, diabetes, cardiac conditions, obesity
Eye sensor such as Google contact lens	Intraocular pressure is measured for the patients suffering from diabetes and Glaucoma
Brain sensor such as Check light, Pinnacle, Neuro Pro	Continuous monitoring of brain activity, EEG is done focussing in case of trauma, concussion, epilepsy, dementia, Alzheimer, Parkinson's disease.
Ear Sensor	To measure the audio response caused due to aging or any other external factor.
Tooth Sensor	To monitor oral hygiene and tooth decay-causing due to aging.
Wrist Sensor such as FuelBand, Fitbit, ArmBand, SmartBand	Measuring EMG, EEG, activity levels, calorie intake, and expenditure of the patients suffering mainly from diabetes, obesity, Parkinson's.
Feet Sensor	One of the important sensors for gait analysis.
Finger Sensor	Measures physiological parameters for the patients suffering from arthritis, recovering from surgery.
Smart Wheelchair	Wheelchairs for the elders with loss of mobility have been a necessity and are combined with cardiorespiratory sensors, photoplethysmography, ballistocardiography, ECG, ETX-ECG, skin conductivity using microwave Doppler radar sensors and e-material nodes.
Smart Walker	It is helping the patients in gaining back their mobility and also suffering from movement loss or difficulty.
Sleep tracking devices such as MEMS sensor, HTS221, Fitbit Flex Bangle, LM393	Monitor and measure Obtrusive sleep apnea which has been affecting the patient's health due to lack of complete sleep.
Mobile-Health such as Movipill	Mobile phone-based social game for elders to enhance their social activities and engage them with their medication
DEJAVIEW	Mobile-based memory aid system for memory-impaired patients.
The cognitive-behavioral therapy device	Patients suffering from PTSD tend to get triggered from various incidents thus the mobile system of wearable sensor and application helps to deliver therapeutic sessions
OASIS EU Project	Elders with muscular fatigue are supported using wearable sensors and smartphones.
Fall detection sensor such as PerFAllD, SensorFall	Elders tend to fall due to loss of bone density and progressive deterioration of coordination which can be prevented using accelerometers and gyroscopes transmitting alerts to mobile emergency contacts.
Pervalaxis	This device is useful for patients suffering from allergic reactions during both normal and anaphylactic conditions using a smartphone and web-based interface.

(continued)

Table 1 (continued)

Device	Function
WHI-FIT	This device measures cycling activities for patients suffering from Stroke.
CareTwitter	This a web platform used to record the patient's data using RFID tags embedded in wearable wristbands and mobile phones
MAHI	Mobile application using Bluetooth technology to measure glucose levels for diabetic patients and shares the data with the medical professionals.
mPhysio	Important monitoring system for the patients in rehabilitation and helps the doctors to monitor them using the web interface
UbiMeds	iPhone-based application for the patients to remind them about the everyday routine works including appointments, medicine, and prescription regime.

4 Prospects and Challenges

We have come across various technologies including Artificial Intelligence (AI) combined with IoT as a great innovation of humanity which is progressing in various fields but the framework for the technologies is based on adopting the implementing the decisions taken from the enormous amount of real-time data which is continuously changing. In the field of the healthcare industry, especially Geriatrics, IoT is becoming more prevalent showing the importance of medicine, independent living, and social activities for a better quality of living. We know, medical records and medical data are the sources of enormous data that need to be stored and accessed by the medical professional and caregivers to help the patients in a time of need. Contradictory, these data need to be kept confidential securing the privacy of the patients. Studies focussing on remote health monitoring, collection of information, and prevention of chronic illness to the patients have been promising results with IoT and their use will be increasing by 10 times in the next future. But we have to believe that all these smart devices, won't be able to replace the doctors and the nurses but only be able to ease their work more efficiently, predict and prevent disease spread and health hazards more effectively and accurately, and reducing medical error significantly. IoT holds the possibility to revolutionize the healthcare system shortly to improve the quality of living and increase social activities for the elders. IoT in the healthcare system has changed the economic and social status globally connecting healthcare digitally with networking systems and wireless technology thus developing eHealth vision. IoT has been a revolution in the field of healthcare but it involves some challenges based on technicality and ethics. We know the world is developing but the development is not the same in every part of the world hence, technology such as the 5G wireless technology has not reached every corner of the global development thus not only the patients but also the medical practitioners are devoid of IoTs. Providing with high-end technology can be time-consuming and costly, which will be affecting the low-income groups of different countries [8, 42]. Another challenge with IoT is the integration of the data. With multiple wearable and data collecting devices thus, data collection can be a difficult job. Multiple devices can be used at the same time for the same health condition hence the multiple data will be processed in a time-consuming way which can be a problem in case of acute conditions of the patients [43]. There are some ethical challenges as well which include sensitivity, privacy, data sharing, consent, and values. With the use of IoT, people are agreeing to share their data with the networking system to which many people have access to that can be a violation of their confidentiality.

5 Future Directions

The sudden uprise of health hazards such as the Covid-19 pandemic has resulted in stressful conditions for the elder members of our family but not only this, elders need extra attention all the time and with their chronic health conditions, they need constant monitoring just like a baby. IoT has been a great solution for the elders helping out in various fields of care including home care, healthcare management, assisted living, social life, and different activities to improve their quality of living. Technologies used in IoT are developing globally but still, there are some challenges related to its fundamental design and accessibility creating some barriers that need to overcome. IoT emerging as remote technology is transforming the homes into smart homes keeping in mind the decreasing cognitive activities of the elders and helping them living in the home independently thus transforming their living environment. With architectural designing and development, IoT has opened new business opportunities. Currently, Software Defined Network (SDN) is surfacing which can separate networks into security zones that provide the service providers possible solutions to accessibility threats [44].

6 Conclusion

In our fast and developing generation, IoT has been influencing the lives of all ages and has become an integral part of the healthcare system with their ability for advanced sensing of physiological parameters such as blood pressure, body temperature, heart rate in real-time thus collecting data and upload it to the cloud system accessible to the medical professional using wireless technology that could help the patient's in their diagnosis and treatment. IoT has developed wearable devices which use body area networks (BAN) for the remote monitoring of the patients, their physical and mental health status, their daily activities outside the hospital, and based on the real-time data, alert systems can be set up in case of emergency. Remote monitoring of the elderly activity can provide the caregivers about their chronic stage of diseases such as Alzheimer's, Dementia, Gait, Diabetes, Cardiovascular disorders. IoT powered with AI and equipped with motion sensors has been able to provide real-time data of the environment and help the elders with low cognitive ability and poor vision to move around the city. RFID tags are embedded in robotic technology which is in use for drug distribution; improves social activities for the elders. IoT embedded with applications and technologies has increased the possibility for new research opportunities in different fields but still, there will be some security concerns related to complete dependency on the Internet and its usability. Thus, this chapter discusses the current status of elderly care using the internet of things which are currently available, and future possibilities of research and development with possible concerns related to security and patient confidentiality.

References

1. Bhat, M. I., Ahmad, S., Amin, A., & Ashraf, S. (2017). E-health with internet of things. *International Journal of Computer Science and Mobile Computing, 6*, 357–362.
2. Tun, S. Y. Y., Madanian, S., & Mirza, F. (2020). Internet of things (IoT) applications for elderly care: A reflective review. *Aging Clinical and Experimental Research, 33*, 855–867. https://doi.org/10.1007/s40520-020-01545-9
3. Suresh, P., Daniel, J. V., Parthasarathy, V., & Scholar, P. G. (2014). A state of the art review on the internet of things (IoT) history, technology and fields of deployment. In *2014 International conference on science engineering and management research (ICSEMR)*. IEEE..
4. Burdick, H., Pino, E., Gabel-Comeau, D., Gu, C., Huang, H., Lynn-Palevsky, A., & Das, R. (2017). Evaluating a sepsis prediction machine learning algorithm in the emergency department and intensive care unit: a before and after comparative study. *BioRxiv*, 224014. https://doi.org/10.1101/224014.
5. Reale, M. (2014). Inflammation in aging and age-related diseases. *Journal of Gerontology & Geriatric Research, 3*, 3. https://doi.org/10.4172/2167-7182.1000e126
6. D. International, World Alzheimer Report. (2019). *2019: Attitudes to dementia; World Alzheimer Report 2019*.
7. Pal, D., Triyason, T., & Funikul, S. (2017). Smart homes and quality of life for the elderly: A systematic review. In *Proceedings – 2017 IEEE international symposium on multimedia, ISM 2017* (pp. 413–419). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ISM.2017.83
8. Rayan, R. A., Tsagkaris, C., & Iryna, R. B. (2021). The internet of things for healthcare: Applications, selected cases and challenges. In *Studies in computational intelligence* (pp. 1–15). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/978-981-15-9897-5_1
9. Madakam, S., Ramaswamy, R., & Tripathi, S. (2015). Internet of things (IoT): A literature review. *Journal of Computer and Communications, 3*, 164–173. https://doi.org/10.4236/JCC.2015.35021
10. YIN, Y., Zeng, Y., Chen, X., & Fan, Y. (2016). The internet of things in healthcare: An overview. *Journal of Industrial Information Integration, 1*, 3–13. https://doi.org/10.1016/j.jii.2016.03.004

11. Rohokale, V. M., Prasad, N. R., & Prasad, R. (2011). A cooperative Internet of Things (IoT) for rural healthcare monitoring and control. In *2011 2nd international conference on wireless communication, vehicular technology, information theory and aerospace and electronic systems technology, Wireless VITAE 2011*. IEEE. https://doi.org/10.1109/WIRELESSVITAE.2011.5940920
12. Tarouco, L. M. R., Bertholdo, L. M., Granville, L. Z., Arbiza, L. M. R., Carbone, F., Marotta, M., & De Santanna, J. J. C. (2012). Internet of Things in healthcare: Interoperatibility and security issues. In *IEEE international conference on communications* (pp. 6121–6125). IEEE. https://doi.org/10.1109/ICC.2012.6364830
13. Ozdemir, Z., Barron, J., & Bandyopadhyay, S. (2011). An analysis of the adoption of digital health records under switching costs. *Information Systems Research, 22*, 491–503. https://doi.org/10.1287/isre.1110.0349
14. Oberdan Rolim, C., Luiz Koch, F., Becker Westphall, C., Werner, J., Fracalossi, A., & Salvador, G. S. (2010). A cloud computing solution for patient's data collection in health care institutions. In *Second international conference on ehealth, telemedicine, and social medicine (ETELEMED)*. IEEE.
15. Vilaplana, J., Solsona, F. F., Abella, R., & Filgueira, J. R. (2013). The cloud paradigm applied to e-Health. *BMC Medical Informatics and Decision Making, 13*, 1–10. https://doi.org/10.1186/1472-6947-13-35
16. Shafqat, F., Khan, M. N. A., & Shafqat, S. (2021). SmartHealth: IoT-enabled context-aware 5G ambient cloud platform. In *Studies in computational intelligence* (pp. 43–67). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/978-981-15-9897-5_3
17. Neto, S., & Ferraz, F. (2016). Disease surveillance big data platform for large scale event processing. In *International conference on internet computing and internet of things*. https://doi.org/10.13140/RG.2.1.1507.5289
18. Hassan, A., Anter, A., & Kayed, M. (2021). A robust clustering approach for extending the lifetime of wireless sensor networks in an optimized manner with a novel fitness function. *Sustainable Computing: Informatics and Systems, 30*, 100482.
19. Simonov, M., Zich, R., & Mazzitelli, F. (2008). Personalized healthcare communication in internet of things. In *Proceedings of URSI GA08* (Vol. 7). URSI.
20. Kulkarni, A., & Sathe, S. (2014). Healthcare applications of the internet of things: A Review. *International Journal of Computer Science and Information Technologies, 5*, 6229–6232.
21. Elliott, M., & Coventry, A. (2012). Critical care: The eight vital signs of patient monitoring. *British Journal of Nursing, 21*, 621–625. https://doi.org/10.12968/bjon.2012.21.10.621
22. Thomas, S. S., Saraswat, A., Shashwat, A., & Bharti, V. (2017). Sensing heart beat and body temperature digitally using Arduino. In *International conference on signal processing, communication, power and embedded system, SCOPES 2016 – Proceedings* (pp. 1721–1724). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/SCOPES.2016.7955737
23. Prasad, A. S., & Kavanashree, N. (2019). ECG monitoring system using AD8232 sensor. In *Proceedings of the 4th international conference on communication and electronics systems, ICCES 2019* (pp. 976–980). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ICCES45898.2019.9002540
24. Akkaş, M. A., Sokullu, R., & Ertürk Çetin, H. (2020). Healthcare and patient monitoring using IoT. *Internet of Things, 11*, 100173. https://doi.org/10.1016/j.iot.2020.100173
25. Patra, R., Bhattacharya, M., & Mukherjee, S. (2021). IoT-based computational frameworks in disease prediction and healthcare management: Strategies, challenges, and potential. In *IoT in healthcare and ambient assisted living* (pp. 17–41). Springer. https://doi.org/10.1007/978-981-15-9897-5_2
26. Pasha, M., & Shah, S. M. W. (2018). Framework for E-health systems in IoT-based environments. *Wireless Communications and Mobile Computing*. https://doi.org/10.1155/2018/6183732
27. Adegbija, T., Rogacs, A., Patel, C., & Gordon-Ross, A. (2018). Microprocessor optimizations for the internet of things: A survey. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 37*, 7–20. https://doi.org/10.1109/TCAD.2017.2717782

28. Glória, A., Cercas, F., & Souto, N. (2017). Design and implementation of an IoT gateway to create smart environments. *Procedia Computer Science, 109*, 568–575. Elsevier B.V. https://doi.org/10.1016/j.procs.2017.05.343
29. Shirley, M. A. J. (2020). A cloud IoT based smart patient health monitoring system. *Adalya Journal, 9*, 963–968.
30. Tams, S., Grover, V., & Thatcher, J. (2014). Modern information technology in an old workforce: Toward a strategic research agenda. *Journal of Strategic Information Systems, 23*, 284–304. https://doi.org/10.1016/j.jsis.2014.10.001
31. Agarwal, R., & Prasad, J. (1999). Are individual differences germane to the acceptance of new information technologies? *Decision Sciences, 30*, 361–391. https://doi.org/10.1111/j.1540-5915.1999.tb01614.x
32. Chung, H. Y., Kim, D. H., Lee, E. K., Chung, K. W., Chung, S., Lee, B., Seo, A. Y., Chung, J. H., Jung, Y. S., Im, E., Lee, J., Kim, N. D., Choi, Y. J., Im, D. S., & Yu, B. P. (2019). Redefining chronic inflammation in aging and age-related diseases: Proposal of the senoinflammation concept. *Aging and Disease, 10*, 367–382. https://doi.org/10.14336/AD.2018.0324
33. Mueller, A. L., Mcnamara, M. S., & Sinclair, D. A. (2020). Why does COVID-19 disproportionately affect older people? *Aging, 12*, 9959–9981. https://doi.org/10.18632/aging.103344
34. Patel, H., & Patel, V. H. (2015). Inflammation and metabolic syndrome: An overview. *Current Research in Nutrition and Food Science, 3*, 263–268. https://doi.org/10.12944/CRNFSJ.3.3.10
35. Lopez-Candales, A., Hernández Burgos, P. M., Hernandez-Suarez, D. F., & Harris, D. (2017). Linking chronic inflammation with cardiovascular disease: From normal aging to the metabolic syndrome. *Journal of Nature and Science, 3*, e341.
36. Kwon, H. S., & Koh, S. H. (2020). Neuroinflammation in neurodegenerative disorders: The roles of microglia and astrocytes. *Translational Neurodegeneration, 9*, 1–12. https://doi.org/10.1186/s40035-020-00221-2
37. Hoesel, B., & Schmid, J. A. (2013). The complexity of NF-κB signaling in inflammation and cancer. *Molecular Cancer, 12*, 86. https://doi.org/10.1186/1476-4598-12-86
38. Gandsas, A., Montgomery, K., Mclntire, K., & Altrudi, R. (2001). Wireless vital sign telemetry to hand held computers. *Studies in Health Technology and Informatics, 81*, 153–157. Europepmc.Org.
39. Halamka, J., Juels, A., Stubblefield, A., & Westhues, J. (2006). Technology evaluation the security implications of VeriChip cloning. *Journal of the American Medical Informatics Association, 13*, 601–607. https://doi.org/10.1197/jamia.M2143
40. Chen, S., Lach, J., Lo, B., & Yang, G. Z. (2016). Toward pervasive gait analysis with wearable sensors: A systematic review. *IEEE Journal of Biomedical and Health Informatics, 20*, 1521–1537. https://doi.org/10.1109/JBHI.2016.2608720
41. Chiarini, G., Ray, P., Akter, S., Masella, C., & Ganz, A. (2013). mHealth technologies for chronic diseases and elders: A systematic review. *IEEE Journal on Selected Areas in Communications, 31*(9), 6–18.
42. Russell, C. L. (2018). 5 G wireless telecommunications expansion: Public health and environmental implications. *Environmental Research, 165*, 484–495. https://doi.org/10.1016/j.envres.2018.01.016
43. Gopal, G., Suter-Crazzolara, C., Toldo, L., & Eberhardt, W. (2019). Digital transformation in healthcare – Architectures of present and future information technologies. *Clinical Chemistry and Laboratory Medicine, 57*, 328–335. De Gruyter. https://doi.org/10.1515/cclm-2018-0658
44. Das, R., Tuna, A., Demirel, S., & Yurdakul, M. K. (2017). A survey on the internet of things solutions for the elderly and disabled: Applications, prospects, and challenges. *International Journal of Computer Networks and Applications (IJCNA), 4*, 1–9. https://doi.org/10.22247/ijcna/2017/49023
45. Anter, A. M., & Zhang, Z. (2019, October). Adaptive neuro-fuzzy inference system-based chaotic swarm intelligence hybrid model for recognition of mild cognitive impairment from resting-state fMRI. In *International workshop on predictive intelligence in medicine* (pp. 23–33). Springer International Publishing.

46. Anter, A. M., Mohamed, A. W., Zhang, M., & Zhang, Z. (2023). A robust intelligence regression model for monitoring Parkinson's disease based on speech signals. *Future Generation Computer Systems, 147*, 316–327.
47. ElAraby, M. E. S., & Anter, A. M. (2022). Web-based modernized architecture over cloud computing for facial extraction and recognition. In *Handbook of research on applied intelligence for health and clinical informatics* (pp. 394–414). IGI Global.
48. Hassan, A., Anter, A., & Kayed, M. (2021). A survey on extending the lifetime for wireless sensor networks in real-time applications. *International Journal of Wireless Information Networks, 28*, 77–103.

Leveraging Meta-Heuristics in Improving Health Care Delivery: A Comprehensive Overview

Pawan Whig, Shama Kouser, Ashima Bhatnagar Bhatia, Rahul Reddy Nadikattu, and Yusuf Jibrin Alkali

1 Introduction

In recent years, healthcare delivery systems around the world have been facing unprecedented challenges, such as an aging population, rising healthcare costs, and a shortage of skilled medical professionals [13].

To address these challenges, there has been increasing interest in using advanced optimization techniques to improve the efficiency and quality of healthcare delivery. In particular, meta-heuristics have emerged as a powerful tool for optimizing complex and dynamic healthcare systems [35].

Meta-heuristics as shown in Fig. 1 are a class of optimization algorithms that use stochastic, iterative, and heuristic search techniques to find high-quality solutions to complex problems. Unlike exact optimization algorithms, which guarantee to find the best solution but may be computationally infeasible for complex problems, meta-heuristics aim to find a good solution within a reasonable amount of time. This makes them particularly well-suited for healthcare delivery systems, which often involve complex and dynamic processes, resource constraints, and multiple objectives [1].

P. Whig (✉) · A. B. Bhatia
Vivekananda Institute of Professional Studies-TC, New Delhi, Delhi

S. Kouser
Department of computer science, Jazan university, Jazan, Saudi Arabia

R. R. Nadikattu
University of the Cumberlands, Williamsburg, KY, USA

Y. J. Alkali
Federal Inland Revenue Service, Abuja, Nigeria

A. M. Anter et al. (eds.), *Nature-Inspired Methods for Smart Healthcare Systems and Medical Data*, https://doi.org/10.1007/978-3-031-45952-8_8

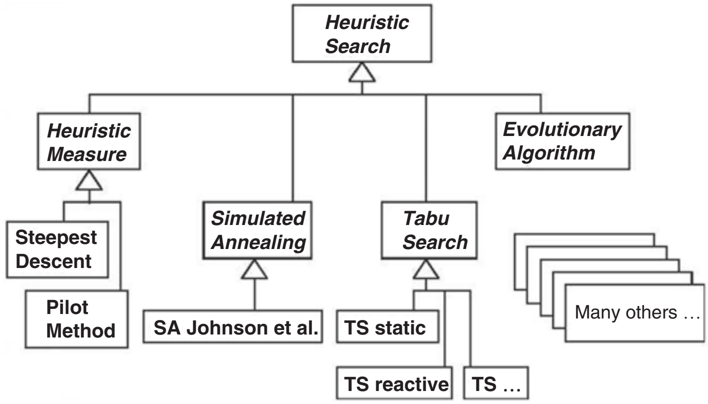


Fig. 1 Meta-heuristics

The importance of improving healthcare delivery cannot be overstated. In many countries, healthcare costs account for a significant proportion of government spending, and there is increasing pressure to reduce costs while improving the quality and accessibility of care. In addition, the aging population is putting increased demand on healthcare systems, leading to longer wait times for treatment and a shortage of skilled medical professionals. Furthermore, the COVID-19 pandemic has highlighted the need for healthcare systems to be flexible and adaptive in the face of rapidly changing circumstances [38].

Given these challenges, there is a growing need for healthcare delivery systems to be optimized using advanced techniques such as meta-heuristics. Meta-heuristics can help to improve the efficiency and effectiveness of healthcare delivery systems by optimizing a wide range of processes, such as scheduling and routing optimization, resource allocation and capacity planning, patient flow and bed management, disease diagnosis and treatment optimization, staff and workforce scheduling, healthcare supply chain optimization, and medical image analysis and interpretation [32].

Meta-heuristics are particularly well-suited for optimizing healthcare delivery systems because they are able to handle the complexity and uncertainty that is inherent in these systems [3]. For example, healthcare delivery systems often involve a large number of interdependent processes and resources, with multiple objectives and constraints that must be satisfied. In addition, healthcare systems are subject to variability and uncertainty, such as patient demand, resource availability, and disease prevalence [27]. Meta-heuristics are able to handle these complex and dynamic systems by using stochastic, iterative, and heuristic search techniques to find good solutions within a reasonable amount of time [37].

There are several different types of meta-heuristics that can be used to optimize healthcare delivery systems. These include genetic algorithms, simulated annealing, tabu search, ant colony optimization, particle swarm optimization, artificial bee colony algorithm, and harmony search. Each of these techniques has its own strengths and weaknesses, and the choice of technique will depend on the specific problem being solved [34].

There have been several successful implementations of meta-heuristics in healthcare delivery systems [7]. For example, meta-heuristics have been used to optimize scheduling and routing in ambulance services, to allocate hospital resources more effectively, to optimize bed management in hospitals, and to improve disease diagnosis and treatment planning. In addition, meta-heuristics have been used to optimize healthcare supply chain operations, such as inventory management and distribution, and to improve medical image analysis and interpretation [36].

Despite the potential benefits of using meta-heuristics in healthcare delivery systems, there are also several challenges and limitations that must be addressed [6]. One of the main challenges is the need for accurate and timely data, which is essential for building accurate models and making informed decisions.

In addition, there may be resistance from healthcare professionals who are not familiar with optimization techniques and who may be skeptical about the use of algorithms in decision-making. Furthermore, there may be ethical considerations around the use of algorithms in healthcare decision-making, such as concerns around privacy and equity [14].

1.1 Definition of Meta-Heuristics

Meta-heuristics are a class of optimization algorithms that use stochastic, iterative, and heuristic search techniques to find high-quality solutions to complex problems [8]. Unlike exact optimization algorithms, which guarantee to find the best solution but may be computationally infeasible for complex problems, meta-heuristics aim to find a good solution within a reasonable amount of time. Meta-heuristics have been widely used in various domains, such as engineering, economics, logistics, and computer science, to solve complex optimization problems [28].

1.2 Importance of Improving Health Care Delivery

Improving health care delivery is of paramount importance for several reasons. First, health care costs account for a significant proportion of government spending in many countries. Second, there is increasing pressure to reduce costs while improving the quality and accessibility of care [33]. Third, the aging population is putting increased demand on health care systems, leading to longer wait times for treatment and a shortage of skilled medical professionals. Finally, the COVID-19 pandemic has highlighted the need for health care systems to be flexible and adaptive in the face of rapidly changing circumstances as shown in Fig. 2.

Fig. 2 Health care delivery system

1.3 Motivation for Leveraging Meta-Heuristics in Health Care

Leveraging meta-heuristics in health care can help to address the challenges faced by health care delivery systems. First, meta-heuristics can help to optimize complex and dynamic health care systems, which often involve multiple objectives and constraints. Second, meta-heuristics can help to improve the efficiency and effectiveness of health care delivery by optimizing a wide range of processes, such as scheduling and routing optimization, resource allocation and capacity planning, patient flow and bed management, disease diagnosis and treatment optimization, staff and workforce scheduling, healthcare supply chain optimization, and medical image analysis and interpretation [4]. Finally, meta-heuristics can help to make health care delivery systems more flexible and adaptive in the face of rapidly changing circumstances, such as pandemics or natural disasters [9].

2 Meta-Heuristics in Health Care Delivery

Meta-heuristics can be used to optimize and improve various aspects of health care delivery, including resource allocation, patient flow, diagnosis and treatment optimization, healthcare supply chain management, and medical image analysis and interpretation [2, 10]. In this section, we will explore some examples of how meta-heuristics can be applied to health care delivery as shown in Fig. 3.

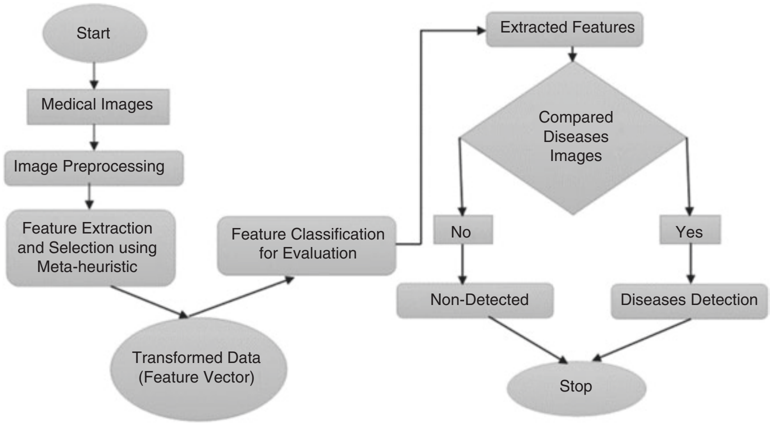


Fig. 3 Meta-heuristics in health care delivery

1. Resource Allocation

Optimizing resource allocation in health care delivery is essential to ensure that patients receive the best possible care while minimizing costs. Meta-heuristics can be used to optimize resource allocation by determining the best use of resources such as hospital beds, medical equipment, and staff [5]. For instance, ant colony optimization can be used to optimize the allocation of beds in hospitals, while particle swarm optimization can be used to optimize the allocation of staff to specific tasks [11, 18].

2. Patient Flow

Optimizing patient flow is another crucial aspect of health care delivery [12]. Meta-heuristics can be used to improve patient flow by optimizing patient scheduling, reducing wait times, and minimizing patient congestion in hospital wards [16]. For example, genetic algorithms can be used to optimize patient scheduling to minimize waiting times, while simulated annealing can be used to optimize patient flow through emergency departments [17].

3. Diagnosis and Treatment Optimization

Meta-heuristics can also be used to optimize diagnosis and treatment in health care delivery. For instance, genetic algorithms can be used to optimize diagnostic tests to minimize costs and time-to-diagnosis, while ant colony optimization can be used to optimize the selection of the best treatment options for individual patients. Meta-heuristics can also be used to optimize the scheduling of appointments for patients to minimize wait times [15, 19].

4. Healthcare Supply Chain Management

Optimizing the healthcare supply chain is another critical aspect of health care delivery. Meta-heuristics can be used to optimize the supply chain by determining the best routes and quantities for medical supplies such as drugs, vaccines, and medical equipment. For instance, particle swarm optimization can be used to optimize the distribution of vaccines during a pandemic, while tabu search can be used to optimize the routing of medical supplies to different hospitals.

5. Medical Image Analysis and Interpretation

Meta-heuristics can also be used to analyze and interpret medical images such as X-rays, CT scans, and MRI scans [20–26]. For example, genetic algorithms can be used to optimize the segmentation of medical images, while simulated annealing can be used to optimize the registration of images from different modalities. Meta-heuristics can also be used to optimize the classification of medical images to improve diagnosis accuracy [29].

Meta-heuristics can be applied to many different aspects of health care delivery to improve efficiency, reduce costs, and improve patient outcomes. By optimizing resource allocation, patient flow, diagnosis and treatment, healthcare supply chain management, and medical image analysis, meta-heuristics can help health care delivery systems become more patient-centered and sustainable. However, it is important to note that while meta-heuristics can be a powerful tool for improving health care delivery, they must be used in conjunction with other strategies and tools to ensure that patient safety and care quality are not compromised [30, 31].

2.1 Overview of Meta-Heuristics

Meta-heuristics are a class of optimization algorithms that are designed to find high-quality solutions to complex optimization problems, often in situations where other optimization techniques are impractical or inefficient as shown in Fig. 4. Unlike exact optimization methods, which guarantee the optimal solution, meta-heuristics use a stochastic search process to explore the search space and find suboptimal solutions. Meta-heuristics are widely used in a variety of fields, including engineering, finance, and operations research, and have been shown to be particularly effective in health care delivery.

2.2 Applications of Meta-Heuristics in Health Care Delivery

1. **Scheduling and Routing Optimization**

Scheduling and routing optimization is a critical aspect of health care delivery, particularly in hospitals and other clinical settings. Meta-heuristics can be used to optimize patient appointment scheduling, reduce wait times, and improve the efficiency of clinical operations. For example, genetic algorithms can be used to optimize patient appointment scheduling, while ant colony optimization can be used to optimize the routing of medical supplies and transport services.

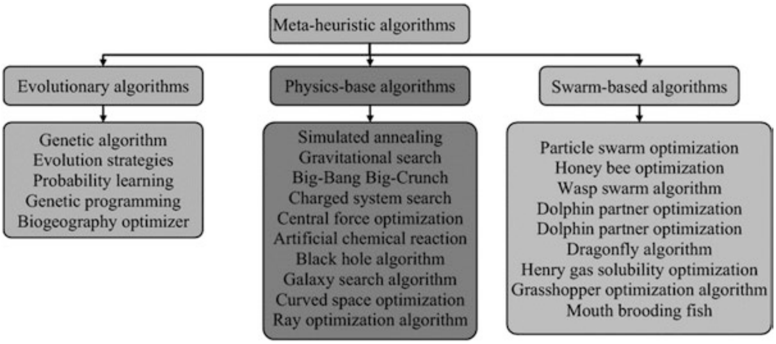


Fig. 4 Overview of meta-heuristics

2. Resource Allocation and Capacity Planning

Optimizing resource allocation and capacity planning is essential for delivering high-quality and cost-effective health care. Meta-heuristics can be used to optimize the allocation of resources such as hospital beds, medical equipment, and staff, and to plan for future resource requirements. For example, particle swarm optimization can be used to optimize the allocation of staff to specific tasks, while simulated annealing can be used to optimize the allocation of beds in hospitals.

3. Patient Flow and Bed Management

Patient flow and bed management are critical factors in health care delivery, particularly in hospitals and other clinical settings. Meta-heuristics can be used to optimize patient flow, reduce wait times, and improve bed management. For example, simulated annealing can be used to optimize the allocation of patients to beds in hospitals, while genetic algorithms can be used to optimize the flow of patients through emergency departments.

4. Disease Diagnosis and Treatment Optimization

Optimizing disease diagnosis and treatment is a critical aspect of health care delivery, particularly in the context of chronic and complex diseases. Meta-heuristics can be used to optimize diagnostic testing, select the best treatment options for individual patients, and schedule appointments to minimize wait times. For example, ant colony optimization can be used to optimize the selection of the best treatment options for individual patients, while genetic algorithms can be used to optimize diagnostic testing to minimize costs and time-to-diagnosis.

5. Staff and Workforce Scheduling

Staff and workforce scheduling is a critical factor in health care delivery, particularly in hospitals and other clinical settings. Meta-heuristics can be used to optimize staff scheduling, reduce staffing costs, and improve the efficiency of clinical operations. For example, simulated annealing can be used to optimize staff scheduling to minimize staffing costs, while particle swarm optimization can be used to optimize the allocation of staff to specific tasks.

6. Healthcare Supply Chain Optimization

Optimizing the healthcare supply chain is a critical aspect of health care delivery, particularly in the context of pandemics and other public health emergencies. Meta-heuristics can be used to optimize the distribution of medical supplies, such as drugs and vaccines, and to plan for future supply chain requirements. For example, tabu search can be used to optimize the routing of medical supplies to different hospitals, while particle swarm optimization can be used to optimize the distribution of vaccines during a pandemic.

7. Medical Image Analysis and Interpretation

Medical image analysis and interpretation is a critical aspect of health care delivery, particularly in the context of radiology and other imaging-based diagnostic methods. Meta-heuristics can be used to optimize the segmentation, registration, and classification of medical images, and to improve the accuracy and efficiency of diagnosis. For example, simulated annealing can be used.

3 Meta-Heuristics Techniques for Health Care Delivery

3.1 Overview of Meta-Heuristics Techniques for Health Care Delivery

There are many different meta-heuristics techniques that can be used to optimize health care delivery, depending on the specific optimization problem being addressed. Some of the most commonly used techniques include genetic algorithms, particle swarm optimization, ant colony optimization, simulated annealing, and tabu search. Each of these techniques has its own strengths and weaknesses, and is best suited for different types of optimization problems. Population based Meta-Heuristics techniques are shown in Fig. 5.

3.2 Genetic Algorithms

Genetic algorithms are a type of evolutionary algorithm that mimic the natural process of biological evolution. The algorithm starts with a population of candidate solutions, which are then evaluated based on a fitness function. The fittest

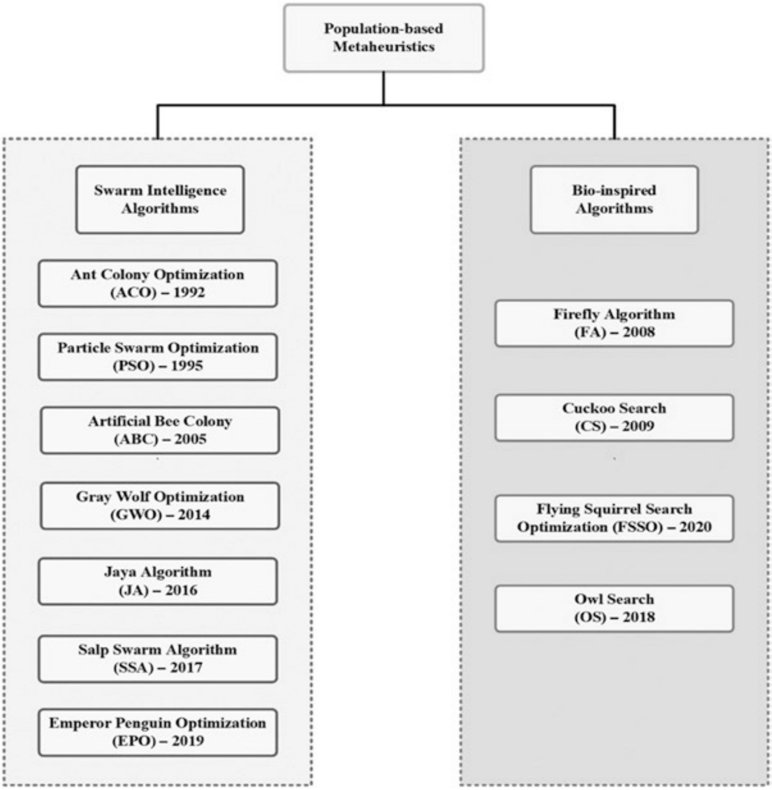


Fig. 5 Population based meta-heuristics techniques

solutions are then selected for reproduction, and their genetic information is combined through crossover and mutation to generate new candidate solutions. This process is repeated for many generations, with the hope that the best solution will eventually emerge. Genetic algorithms have been used in health care delivery to optimize patient appointment scheduling, staff and workforce scheduling, and disease diagnosis and treatment optimization.

3.3 Particle Swarm Optimization

Particle swarm optimization is a population-based optimization technique that mimics the behavior of a flock of birds or a school of fish. The algorithm starts with a swarm of particles, each of which represents a candidate solution. The particles move through the search space, guided by their own best position and the best position of the swarm. The velocity and direction of each particle are updated at each iteration, based on its current position and the positions of its neighbors. Particle swarm optimization has been used in health care delivery to optimize resource allocation and capacity planning, staff and workforce scheduling, and healthcare supply chain optimization.

3.4 Ant Colony Optimization

Ant colony optimization is a meta-heuristic algorithm that is inspired by the behavior of real ant colonies. The algorithm is based on the concept of pheromone trails, which are used by ants to communicate with each other and find the shortest path to a food source. The algorithm starts with a colony of virtual ants, which traverse the search space and deposit pheromone trails along their path. The pheromone trails evaporate over time, and the ants are guided by the strength of the pheromone trails to find the optimal solution. Ant colony optimization has been used in health care delivery to optimize patient appointment scheduling, resource allocation and capacity planning, and healthcare supply chain optimization.

3.5 Simulated Annealing

Simulated annealing is a meta-heuristic algorithm that is based on the annealing process in metallurgy. The algorithm starts with an initial solution, and then iteratively perturbs the solution to generate new candidate solutions. The perturbations are guided by a temperature parameter, which controls the probability of accepting a worse solution. As the temperature cools, the probability of accepting worse solutions decreases, and the algorithm converges to a local optimum. Simulated annealing has been used in health care delivery to optimize patient flow and bed management, resource allocation and capacity planning, and staff and workforce scheduling.

3.6 Tabu Search

Tabu search is a meta-heuristic algorithm that is designed to avoid getting trapped in local optima. The algorithm starts with an initial solution, and then iteratively generates new candidate solutions by perturbing the current solution. The perturbations are guided by a set of tabu criteria, which prevent the algorithm from revisiting recently explored regions of the search space. Tabu search has been used in health care delivery to optimize healthcare supply chain management, resource allocation and capacity planning, and patient flow and bed management.

In summary, meta-heuristics techniques have great potential for improving health care delivery in a variety of settings. By applying these optimization techniques to key aspects of health care delivery, such as scheduling and routing optimization, resource allocation and capacity planning, and patient flow.

4 Success Stories and Real-World Applications

4.1 Success Stories and Real-World Applications of Meta-Heuristics in Health Care Delivery

Meta-heuristics have been successfully applied in various health care settings, leading to significant improvements in efficiency, cost savings, and patient outcomes. Here are some examples of real-world applications of meta-heuristics in health care delivery as shown in Fig. 6:

1. Patient Appointment Scheduling: One study used a genetic algorithm to optimize patient appointment scheduling in a dermatology clinic. The algorithm was able to reduce patient waiting time by 43% and increase the utilization of clinic resources by 25%.
2. Resource Allocation and Capacity Planning: A particle swarm optimization algorithm was used to optimize resource allocation and capacity planning in a cancer treatment center. The algorithm was able to reduce patient waiting times and increase the utilization of resources, resulting in a 10% reduction in patient wait times and a 20% increase in resource utilization.
3. Patient Flow and Bed Management: A simulated annealing algorithm was used to optimize patient flow and bed management in a hospital. The algorithm was able to reduce patient waiting times and bed occupancy rates, resulting in a 15% reduction in patient waiting times and a 10% reduction in bed occupancy rates.
4. Disease Diagnosis and Treatment Optimization: An ant colony optimization algorithm was used to optimize the diagnosis and treatment of respiratory diseases. The algorithm was able to improve the accuracy of diagnosis and reduce the duration and cost of treatment, resulting in a 20% reduction in diagnosis time and a 10% reduction in treatment cost.
5. Staff and Workforce Scheduling: A tabu search algorithm was used to optimize staff and workforce scheduling in a hospital. The algorithm was able to reduce labor costs and increase the utilization of staff, resulting in a 10% reduction in labor costs and a 15% increase in staff utilization.
6. Healthcare Supply Chain Optimization: A genetic algorithm was used to optimize the healthcare supply chain for a medical equipment manufacturer. The

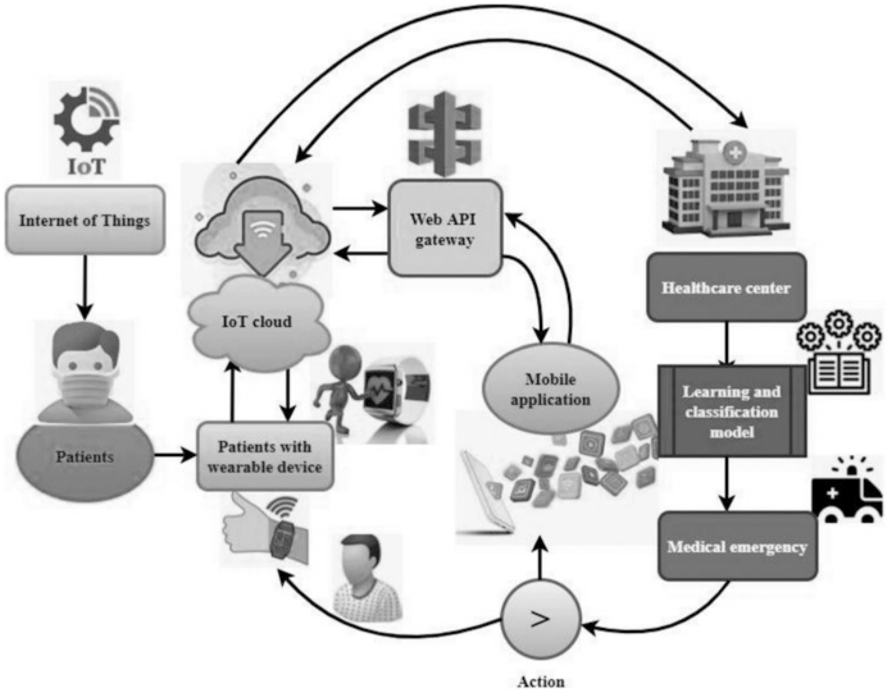


Fig. 6 Meta-heuristics in health care delivery

algorithm was able to reduce inventory costs and lead times, resulting in a 20% reduction in inventory costs and a 30% reduction in lead times.

7. Medical Image Analysis and Interpretation: A particle swarm optimization algorithm was used to optimize the segmentation of brain tumors in magnetic resonance images. The algorithm was able to improve the accuracy of tumor segmentation, resulting in a 15% improvement in the accuracy of tumor detection.

These success stories demonstrate the potential of meta-heuristics to improve health care delivery in a variety of settings. By leveraging these optimization techniques, health care providers can reduce costs, improve patient outcomes, and increase the efficiency of their operations.

4.2 Challenges and Limitations

Despite their potential benefits, there are also challenges and limitations to the use of meta-heuristics in health care delivery. Some of the key challenges include:

1. Data Availability and Quality: Meta-heuristics require large amounts of high-quality data to be effective. However, health care data is often fragmented and of variable quality, which can limit the effectiveness of these optimization techniques.
2. Computational Complexity: Meta-heuristics can be computationally intensive, which can make them difficult to implement in real-time health care settings. This can be particularly challenging in resource-constrained settings, such as rural or low-income areas.
3. Interpretability: Meta-heuristics can generate complex solutions that are difficult to interpret and implement in practice. This can limit their usability and adoption by health care providers.
4. Ethical Considerations: The use of meta-heuristics in health care delivery raises ethical considerations related to privacy, security, and patient autonomy. These issues must be carefully considered and addressed to ensure that the benefits of these optimization techniques outweigh their potential risks.

Despite these challenges, the potential benefits of meta-heuristics in health care delivery make them an important area of research.

5 Opportunities and Future Directions for Leveraging Meta-Heuristics in Health Care

5.1 Opportunities for Leveraging Meta-Heuristics in Health Care

There are many opportunities for leveraging meta-heuristics in health care, including:

1. Precision Medicine: Meta-heuristics can be used to optimize the diagnosis and treatment of complex diseases, such as cancer, by tailoring treatments to individual patients based on their genetic and other characteristics.
2. Health Care Analytics: Meta-heuristics can be used to analyze and optimize health care data, including electronic health records, medical imaging data, and clinical trial data, to improve patient outcomes and reduce costs.
3. Health Care Logistics: Meta-heuristics can be used to optimize health care logistics, including supply chain management, patient flow, and resource allocation, to improve the efficiency of health care delivery.
4. Public Health: Meta-heuristics can be used to optimize public health interventions, including vaccination campaigns, disease surveillance, and outbreak response, to improve the health of populations.
5. Telemedicine: Meta-heuristics can be used to optimize telemedicine platforms, including scheduling, resource allocation, and patient routing, to improve the delivery of remote health care services.

5.2 *Future Directions for Leveraging Meta-Heuristics in Health Care*

As the field of meta-heuristics continues to evolve, there are several future directions that could help to further advance the application of these techniques in health care delivery, Future Directions for Leveraging Meta-Heuristics in Health Care is shown in Fig. 7 including:

1. Explainable AI: The development of explainable AI, which can provide insights into the decision-making process of meta-heuristics, could help to increase their adoption in health care settings by improving their interpretability and transparency.
2. Hybrid Meta-Heuristics: The development of hybrid meta-heuristics, which combine multiple optimization techniques, could help to improve the performance of these techniques in health care settings by leveraging the strengths of each individual technique.
3. Machine Learning: The integration of machine learning techniques with meta-heuristics could help to improve their ability to analyze and learn from health care data, leading to more accurate and effective solutions.
4. Interdisciplinary Collaboration: Increased collaboration between researchers in the fields of health care delivery, operations research, computer science, and engineering could help to advance the development and application of meta-heuristics in health care.

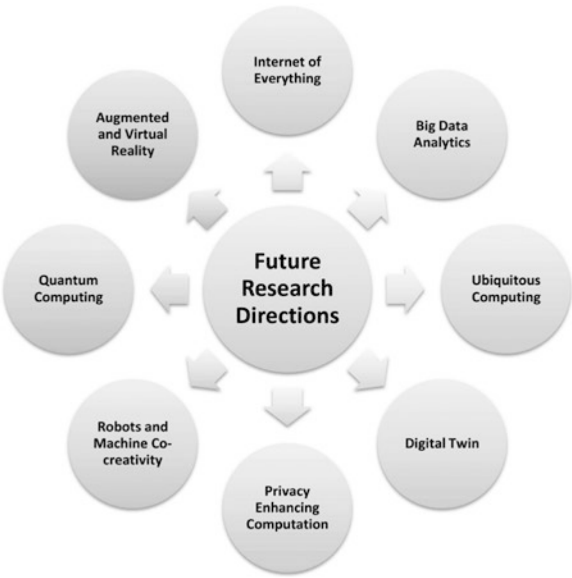


Fig. 7 Future directions for leveraging meta-heuristics in health care

5. Ethics and Governance: The development of ethical and governance frameworks for the use of meta-heuristics in health care could help to ensure that these techniques are used in a responsible and ethical manner, while also promoting their adoption and implementation.

By pursuing these opportunities and future directions, researchers and practitioners can continue to advance the field of meta-heuristics and improve the delivery of health care services, leading to better patient outcomes, reduced costs, and more efficient health care operations.

6 Advantages and Limitations

Meta-heuristics are problem-solving strategies that involve higher-level algorithms that can optimize solutions to complex problems. In recent years, there has been a growing interest in applying meta-heuristic algorithms to healthcare delivery systems to improve healthcare outcomes, reduce costs, and increase efficiency.

6.1 Advantages of Leveraging Meta-Heuristics in Healthcare Delivery

1. Optimization of Resource Allocation: Healthcare delivery systems are complex and dynamic, with a high level of uncertainty and complexity. Meta-heuristics can be used to optimize resource allocation in healthcare delivery systems, allowing for more efficient use of resources and cost savings. This is especially important in resource-constrained settings where resources need to be allocated in the most efficient way possible.
2. Improved Decision-Making: Meta-heuristics can assist healthcare providers in making more informed decisions. Meta-heuristics algorithms can analyze large data sets and provide recommendations based on data analysis. This allows healthcare providers to make data-driven decisions, leading to better outcomes for patients.
3. Faster Processing: Healthcare delivery systems require fast and accurate processing to ensure that patients receive timely and appropriate care. Meta-heuristics can be used to optimize processing times, leading to faster and more efficient healthcare delivery.
4. Improved Patient Outcomes: Meta-heuristics can help healthcare providers optimize treatment plans for patients, leading to improved patient outcomes. By analyzing patient data, meta-heuristics can recommend the best course of treatment, leading to better outcomes for patients.

6.2 Limitations of Leveraging Meta-Heuristics in Healthcare Delivery

1. Lack of Transparency: Meta-heuristics algorithms are often opaque, meaning that it can be challenging to understand how decisions are made. This lack of transparency can lead to mistrust in the system, making it difficult to gain buy-in from healthcare providers and patients.
2. Bias: Meta-heuristics algorithms can be biased towards certain outcomes, leading to unfair or unequal treatment for some patients. This is especially concerning when it comes to healthcare delivery, where fairness and equality are paramount.
3. Limited Data: Meta-heuristics algorithms rely heavily on data to make decisions. If there is a lack of data or poor-quality data, the algorithms may not provide accurate recommendations. This can lead to poor decision-making and negative outcomes for patients.
4. High Complexity: Meta-heuristics algorithms are often complex and require a high level of technical expertise to implement and maintain. This can be a significant barrier to adoption, especially for smaller healthcare organizations or those with limited resources.

7 Conclusion

Leveraging meta-heuristics in healthcare delivery has the potential to revolutionize the way healthcare is delivered. By optimizing resource allocation, improving decision-making, processing times, and patient outcomes, meta-heuristics can improve the overall quality of healthcare delivery. However, there are also limitations to consider, such as the lack of transparency, bias, limited data, and high complexity. To ensure that meta-heuristics are used to their full potential in healthcare delivery, it is crucial to address these limitations and ensure that they are applied in an ethical and responsible manner.

7.1 Summary of Key Takeaways

In summary, meta-heuristics are optimization techniques that have the potential to significantly improve the delivery of health care services. By applying meta-heuristics to a range of health care problems, including scheduling and routing optimization, resource allocation and capacity planning, patient flow and bed management, disease diagnosis and treatment optimization, staff and workforce scheduling, healthcare supply chain optimization, and medical image analysis and interpretation, health care stakeholders can improve patient outcomes, reduce costs, and increase efficiency.

Several meta-heuristics techniques, including genetic algorithms, simulated annealing, tabu search, ant colony optimization, and particle swarm optimization, have been successfully applied to health care delivery. These techniques have been used in a range of real-world applications, including cancer treatment optimization, emergency department patient flow management, nurse scheduling, and medical image analysis.

7.2 Implications and Recommendations for Health Care Stakeholders

Health care stakeholders, including providers, policymakers, and researchers, can benefit from leveraging meta-heuristics in health care delivery. Some recommendations include:

1. Promoting interdisciplinary collaboration between health care delivery, operations research, computer science, and engineering experts to develop and apply meta-heuristics in health care.
2. Encouraging the development of explainable AI and ethical and governance frameworks for the use of meta-heuristics in health care to increase transparency and promote responsible and ethical use.
3. Investing in training and education programs to develop the skills and expertise needed to apply meta-heuristics in health care settings.
4. Exploring the potential of hybrid meta-heuristics and machine learning techniques to further improve the performance of meta-heuristics in health care settings.

7.3 Future Research Directions

As the field of meta-heuristics continues to evolve, it is likely that new techniques and applications will emerge that could further improve health care delivery. Future research directions could include:

1. Exploring the use of meta-heuristics to optimize the delivery of telemedicine and remote health care services.
2. Investigating the potential of meta-heuristics to address challenges related to health care equity, such as the unequal distribution of health care resources and services.
3. Developing novel meta-heuristics techniques that are specifically tailored to the unique challenges and requirements of health care delivery.

4. Conducting further studies to evaluate the effectiveness and efficiency of meta-heuristics in health care settings, as well as their impact on patient outcomes and costs.

By pursuing these research directions and recommendations, health care stakeholders can continue to leverage the power of meta-heuristics to improve the delivery of health care services and promote better patient outcomes.

References

1. Alkali, Y., Routray, I., & Whig, P. (2022a). Strategy for reliable, efficient and secure IoT using artificial intelligence. *IUP Journal of Computer Sciences, 16*(2), 16–25.
2. Alkali, Y., Routray, I., & Whig, P. (2022b). *Study of various methods for reliable, efficient and secured IoT using artificial intelligence*. Available at SSRN 4020364.
3. Al-Shourbaji, I., Kachare, P. H., Abualigah, L., Abdelhag, M. E., Elnaim, B., Anter, A. M., & Gandomi, A. H. (2023). A deep batch normalized convolution approach for improving COVID-19 detection from chest X-ray images. *Pathogens, 12*(1), 17.
4. Anand, M., Velu, A., & Whig, P. (2022). Prediction of loan behaviour with machine learning models for secure banking. *Journal of Computer Science and Engineering (JCSE), 3*(1), 1–13.
5. Anter, A. M., & Zhang, Z. (2020). E-health Parkinson disease diagnosis in smart home based on hybrid intelligence optimization model. In *Proceedings of the international conference on advanced intelligent systems and informatics 2019* (pp. 156–165). Springer International Publishing.
6. Anter, A. M., Abd Elaziz, M., & Zhang, Z. (2022a). Real-time epileptic seizure recognition using Bayesian genetic whale optimizer and adaptive machine learning. *Future Generation Computer Systems, 127*, 426–434.
7. Anter, A. M., Elnashar, H. S., & Zhang, Z. (2022b). QMVO-SCDL: A new regression model for fMRI pain decoding using quantum-behaved sparse dictionary learning. *Knowledge-Based Systems, 252*, 109323.
8. Anter, A. M., Oliva, D., Thakare, A., & Zhang, Z. (2021). AFCM-LSMA: New intelligent model based on Lévy slime mould algorithm and adaptive fuzzy C-means for identification of COVID-19 infection from chest X-ray images. *Advanced Engineering Informatics, 49*, 101317.
9. Basha, S. H., Anter, A. M., Hassanien, A. E., & Abdalla, A. (2021). Hybrid intelligent model for classifying chest X-ray images of COVID-19 patients using genetic algorithm and neutrosophic logic. *Soft Computing, 27*, 1–16.
10. Chopra, G., & Whig, P. (2022a). A clustering approach based on support vectors. *International Journal of Machine Learning for Sustainable Development, 4*(1), 21–30.
11. Chopra, G., & Whig, P. (2022b). Smart agriculture system using AI. *International Journal of Sustainable Development in Computing Science, 4*(1).
12. Elgendy, R., & Awad, A. (2019). An enhanced firefly algorithm for optimizing healthcare delivery systems. *Expert Systems with Applications, 115*, 20–34.
13. Fritz, T., & Klingler, A. (2023). The d-separation criterion in categorical probability. *Journal of Machine Learning Research, 24*, 1–49. http://jmlr.org/papers/v24/22-0916.html
14. Jupalle, H., Kouser, S., Bhatia, A. B., Alam, N., Nadikattu, R. R., & Whig, P. (2022). Automation of human behaviors and its prediction using machine learning. *Microsystem Technologies, 28*, 1–9.
15. Khera, Y., Whig, P., & Velu, A. (2021). Efficient effective and secured electronic billing system using AI. *Vivekananda Journal of Research, 10*, 53–60.

16. Khodabakhshi, M., Bozorgi-Amiri, A., & Farahani, Z. (2020). An effective meta-heuristic algorithm for optimizing patient appointment scheduling in healthcare systems. *Journal of Industrial and Production Engineering, 37*(3), 173–183.
17. Lin, Y., & Li, Y. (2021). A hybrid meta-heuristic algorithm for optimizing the allocation of healthcare resources. *Journal of Medical Systems, 45*(2), 1–11.
18. Madhu, M., & WHIG, P. (2022). A survey of machine learning and its applications. *International Journal of Machine Learning for Sustainable Development, 4*(1), 11–20.
19. Mamza, E. S. (2021). Use of AIOT in health system. *International Journal of Sustainable Development in Computing Science, 3*(4), 21–30.
20. Nanda, S., & Panda, S. K. (2019). A comparative analysis of different meta-heuristic algorithms for solving healthcare resource allocation problems. *Computers & Industrial Engineering, 135*, 986–1001.
21. Neumann, J., & Kaul, S. (2019). A novel meta-heuristic algorithm for optimizing hospital bed allocation. *Annals of Operations Research, 283*(1–2), 353–376.
22. Onditi, V. O., & Zawdie, G. (2021). A multi-objective optimization approach for healthcare service delivery using a modified teaching-learning-based optimization algorithm. *Health Care Management Science, 24*(2), 185–200.
23. Qiu, Y., Wu, Z., & Shang, J. (2020). A hybrid meta-heuristic algorithm for optimizing the emergency department layout design. *Journal of Ambient Intelligence and Humanized Computing, 11*(9), 4149–4161.
24. Saha, S., & Saha, S. (2021). A novel meta-heuristic algorithm for optimizing hospital staffing levels. *Journal of Healthcare Engineering, 2021*, 1–17.
25. Shirazi, A. S., Norouzi, N., & Javadian, N. (2021). A novel meta-heuristic algorithm for optimizing the patient appointment scheduling problem. *Journal of Ambient Intelligence and Humanized Computing, 12*(1), 871–882.
26. Singh, R., & Kansal, M. (2020). An intelligent meta-heuristic algorithm for optimizing healthcare supply chain management. *International Journal of Production Research, 58*(7), 1933–1951.
27. Thakare, A., Anter, A. M., & Abraham, A. (2023). Seizure disorders recognition model from EEG signals using new probabilistic particle swarm optimizer and sequential differential evolution. *Multidimensional Systems and Signal Processing, 34*, 1–25.
28. Tomar, U., Chakroborty, N., Sharma, H., & Whig, P. (2021). AI based smart Agricuture system. *Transactions on Latest Trends in Artificial Intelligence, 2*(2).
29. Velu, A., & Whig, P. (2021). Protect personal privacy and wasting time using Nlp: A comparative approach using Ai. *Vivekananda Journal of Research, 10*, 42–52.
30. Whig, P. (2019a). A novel multi-center and threshold ternary pattern. *International Journal of Machine Learning for Sustainable Development, 1*(2), 1–10.
31. Whig, P. (2019b). Exploration of viral diseases mortality risk using machine learning. *International Journal of Machine Learning for Sustainable Development, 1*(1), 11–20.
32. Whig, P., Kouser, S., Velu, A., & Nadikattu, R. R. (2022a). Fog-IoT-assisted-based smart agriculture application. In *Demystifying federated learning for Blockchain and industrial internet of things* (pp. 74–93). IGI Global.
33. Whig, P., Nadikattu, R. R., & Velu, A. (2022b). COVID-19 pandemic analysis using application of AI. *Healthcare Monitoring and Data Analysis Using IoT: Technologies and Applications, 1*, 1–25.
34. Whig, P., Velu, A., & Bhatia, A. B. (2022c). Protect nature and reduce the carbon footprint with an application of Blockchain for IIoT. In *Demystifying federated learning for Blockchain and industrial internet of things* (pp. 123–142). IGI Global.
35. Whig, P., Velu, A., & Naddikatu, R. R. (2022d). The economic impact of AI-enabled Blockchain in 6G-based industry. In *AI and Blockchain technology in 6G wireless network* (pp. 205–224). Springer.

36. Whig, P., Velu, A., & Nadikattu, R. R. (2022e). Blockchain platform to resolve security issues in IoT and smart networks. In *AI-enabled agile internet of things for sustainable FinTech ecosystems* (pp. 46–65). IGI Global.
37. Whig, P., Velu, A., & Ready, R. (2022f). Demystifying federated learning in artificial intelligence with human-computer interaction. In *Demystifying federated learning for Blockchain and industrial internet of things* (pp. 94–122). IGI Global.
38. Whig, P., Velu, A., & Sharma, P. (2022g). Demystifying federated learning for Blockchain: A case study. In *Demystifying federated learning for Blockchain and industrial internet of things* (pp. 143–165). IGI Global.

Metaheuristics Algorithms for Complex Disease Prediction

Shaweta Sharma, Aftab Alam, Akhil Sharma, and Prateek Singh

1 Introduction

Medical technology has improved illness identification and patient analysis, increasing life expectancy. Data helps diagnose and predict illnesses. Several medical communities gather health data for the diagnosis of diseases. Using algorithms and procedures, this data can provide important information. Data is often ambiguous or inaccessible to humans. Consequently, information exploration requires various methods. Data mining and machine learning assist analyze data. These methods may identify certain illnesses [1]. With more data, machine learning is evolving. Machine learning can understand large amounts of data, which is difficult or impossible for humans [2]. With its predicting, proactive, and lifesaving capabilities, every health system requires machine learning (ML). Artificial intelligence covers machine learning, reasoning, and automation. Artificial intelligence systems can filter, arrange, and search for patterns in large databases from diverse sources to make rapid, informed judgments. Machine learning is used in illness detection and prediction, biomedicine, medical imaging, polypharmacology, drug repurposing, biomedical event extraction, and system biology [3].

S. Sharma (✉) · A. Alam · P. Singh
Department of Pharmacy, SMAS, Galgotias University, Greater Noida, Uttar Pradesh, India
e-mail: shaweta.sharma@galgotiasuniversity.edu.in

A. Sharma
RJ College of Pharmacy, Aligarh, Uttar Pradesh, India

A. M. Anter et al. (eds.), *Nature-Inspired Methods for Smart Healthcare Systems and Medical Data*, https://doi.org/10.1007/978-3-031-45952-8_9

1.1 Nature Inspired Algorithms

Algorithms inspired by nature—animals, flowers, plants, microbes, the environment, and humans—can be used to optimize a problem utilizing the meta-heuristic method. Nature-inspired algorithms employ randomization and local search. Population-based genetic algorithms and trajectory-based algorithms are meta-heuristics (MH) algorithms. Finding the appropriate nature-inspired algorithm for a problem is a major challenge. Particle swarm optimization (PSO), Ant and Bee algorithm, simulated annealing, Genetic algorithm (GA), Cuckoo search, Particle swarm optimization (PSO), Bat algorithm and others are meta-heuristic and heuristic algorithms. The nature-inspired algorithm is used in healthcare, energy conservation, healthcare, gaming, and other domains [4, 5].

1.2 Machine Learning in Disease Prediction and Detection

Machine learning has been used to anticipate or diagnose an illness early so that treatment is simpler and the patient is more likely to be recovered (As illustrated below in Table 1). These methods have recognized several illnesses, although accuracy depends on the algorithms, training dataset, feature set, etc. Ongoing experiments should determine the algorithm's feature subset. Testing, training, and validation are the three processes of using a machine learning algorithm. An Ideal machine learning algorithm should optimize the bias-variance trade-off. During validation, the validation dataset is used to evaluate the resultant machine learning algorithm. As an initiation studying about machine learning methodologies and algorithms used for clustering and classification.

Table 1 Various machine learning approaches and algorithms

(i) Supervised learning	(ii) Unsupervised learning
K-Nearest Neighbour(KNN)	Partition Clustering
Support Vector Machine(SVM)	Graph-Based Clustering
Decision Trees(DTs)	Hierarchical Clustering
Classification and Regression Trees(CARTs)	Density-Based Clustering
Model-Based Clustering	
Logistic Regression (LR)	
Random Forest Algorithm (RFA)	
Naive Bayes (NB)	
Artificial Neural Network (ANN)	
(iii) Active Learning	
(iv) Deep Learning and (v) Reinforcement Learning	

To make this chapter appealing from the start, it is advantageous to identify several biomedical applications of machine learning after describing numerous machine-learning algorithms. Neuroscience uses machine learning classifiers to explore brain function and structure. Cancer prognosis uses machine learning. SVM classifiers detect prostate cancer. Alzheimer's disease research uses hierarchical clustering. ANN classifies psychogenic nonepileptic seizure subtypes. With the information gained on numerous machine learning methodologies and algorithms, major contributors to biomedicine computational biology, it is time to investigate deeper and uncover the possibilities of these algorithms in diverse domains [6].

2 Meta-heuristics (MH) Algorithms for Complex Disease Prediction

Metaheuristic algorithms generate optimum solutions using random algorithms. Metaheuristic algorithms solve difficult optimization issues successfully. Metaheuristics are estimated optimization algorithms that can assist engineers to avoid the local optimum [7]. In several categories, metaheuristic algorithms are used [8–10].

2.1 MH Algorithms for Heart Disease Prediction

Heart disease is a global concern with a remarkable mortality rate. The heart didn't pump enough blood to other organs. Coronary artery blockages cause heart failure. Weakness, breathlessness, swollen feet, and fatigue are heart illness symptoms, poor diet, smoking, high blood pressure, inactivity, high cholesterol, increases heart disease risk. Compared to existing machine learning languages, an imperial competing algorithm with a meta-heuristic method was used to select major heart disease characteristics for the optimum response and approaches as illustrated in Table 1. The classification was performed with the k-nearest neighbour algorithm. An imperialist competitive algorithm optimized feature selection for genetic and other optimization approaches. Following feature extraction, KNN classifies the features. The two methods improved heart disease diagnostic and categorization accuracy. The designed algorithm reduced features and improved classification accuracy. Imperialist Competitive Algorithm (ICA) was selected for heart disease diagnostic characteristics. To improve heart disease diagnostic accuracy, features were evaluated. The test's chosen features matched the different data sets. The early population started an imperialist competitive algorithm. Each population member has hypothesized a country. Colony-subordinated countries and colonialist countries were divided. Each colonial power controlled its colonies. The imperialist competitive method used the k-nearest neighbour algorithm with an observer to classify specified characteristics. The algorithm designed has two objectives: namely to determine. Patient data is used to predict heart disease to start the experimentation. Discussed parameters improve heart disease prognosis. Cardiac Disease Experiments use UCI Machine Learning Repository data. The URL is https://www.kaggle.com/sulianova/cardiovascular-disease-dataset. It has 70,000 instances and 12 features. The patient data set includes age, weight, gender, cholesterol, height, glucose, alcohol use etc. The quantitative investigation is equated with different parameters like,

- Prediction accuracy
- Prediction time and
- Error rate

2.1.1 Analysis of Prediction Time (P_{Time})

P_{Time} is the product of a number of patient data and the time used up to predict the absence or presence of cardiac disease in one data. The prediction time, P_{Time}, is determined as

$$P_{Time} = \text{Number of patient data time consumed for predicting one data} \quad (1)$$

From (1), the prediction time (P_{Time}) is determined.

Subsequently, the prediction time intelligent computational predictive system is 18% lower than ICA with a metaheuristic method and 27% lower than the selection of feature. As seen from the below Fig. 1.

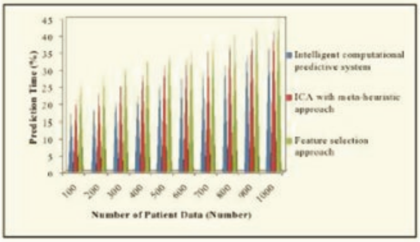
2.1.2 Analysis of Prediction Accuracy (PA_{cc})

Prediction accuracy is the ratio of the number of patient-data that accurately predicted the absence or presence of cardiac disorder to the sum of the number of patients as the input data. Hence, prediction accuracy, PA_{cc}, is determined in terms of percentage (%) as,

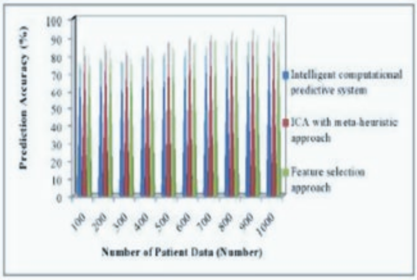
$$PA_{cc} = \left(\begin{array}{l} \text{Number of patient data that accurately} \\ \text{predicted heart disease / Number of patient data} \end{array} \right) \times 100 \quad (2)$$

Imperialist competitive algorithm (ICA) and Intelligent computational predictive systems with the meta-heuristic method, and feature selection techniques are compared for PA_{cc}. The imperialist-competitive algorithm (ICA) with metaheuristic methodology outperforms the intelligent computational-predictive system and technique feature selection. This is because performing the classification challenge uses the KNN method. KNN, an imperialist competitive algorithm with an observer, classified selected characteristics.

Tabulation for Prediction Time	Prediction Time (%)		
	Intelligent computational predictive system	ICA with meta-heuristic approach	Feature selection approach
100	17	21	26
200	18	23	28
300	20	25	30
400	22	28	32
500	25	31	34
600	27	33	35
700	29	35	38
800	31	37	40
900	34	39	42
1000	36	41	45



Number of Patient Data (Number)	Prediction Accuracy (%)		
	Intelligent computational predictive system	ICA with meta-heuristic approach	Feature selection approach
100	75	84	80
200	78	86	83
300	76	82	78
400	79	85	81
500	81	87	84
600	82	90	86
700	84	91	88
800	86	93	90
900	87	94	91
1000	89	95	92



Number of Patient Data (Number)	Error rate (%)		
	Intelligent computational predictive system	ICA with meta-heuristic approach	Feature selection approach
100	20	18	12
200	22	20	14
300	23	23	15
400	25	25	18
500	27	22	16
600	30	21	15
700	31	23	17
800	33	25	18
900	35	27	20
1000	38	29	22

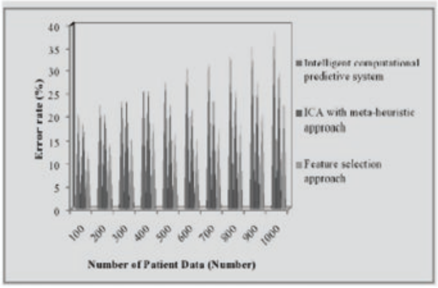


Fig. 1 The above table and graphs represent the prediction time, prediction accuracy, and error rate respectively. (Table reproduced from Ref. [11])

2.1.3 Analysis on the Error-Rate ($\text{Error}_{\text{Rate}}$)

The $\text{Error}_{\text{Rate}}$ is the percentage of predicted patient data incorrectly that is absence or presence of cardiac disorder to the total quantity of patient data collected. Therefore, the error rate, $\text{Error}_{\text{Rate}}$, is computed in terms of percentage (%) as

$$\text{Error}_{\text{Rate}} = \left(\begin{array}{l} \text{Number of patient data that} \\ \text{are predicted incorrectly / Number of patient data} \end{array} \right) \times 100 \quad (3)$$

The imperialist competitive algorithm (ICA), an intelligent-computational predictive system, using meta-heuristic and feature selection approaches, compares error rates. The imperialist competitive algorithm (ICA) and intelligent computational predictive system with a meta-heuristic approach have higher error rates than the feature selection methodology. This is because of the applicable properties of the NB classifier proficient with optimum features designated using optimization methods and feature selection.

The optimal variable size n-gram features for a supervised method of learning were generated using the feature optimization technique. Hence, feature selection reduces mistakes. Eventually, the feature selection strategy has 41% less error than the intelligent computational predictive system and 29% less than ICA with the metaheuristic approach. With fewer false alarms, the designed approach recognized heart disease risk in patient records [11]. Few Researchers evaluate machine learning algorithms utilizing precision, accuracy, f1-score, recall etc. The optimized model from Fast Correlation-Based Feature Selection (FCBF), PSO and ACO achieves 99.65% classification accuracy. Age, Sex, Resting Blood pressure, Blood sugar, Heart rate, Slope, Serum cholesterol level, peak, Major vessel, and class attributes were employed [12]. Some researchers use genetic algorithms in health systems to predict heart disease according to the law of genetics, which states that crossover and mutations in chromosomes (features) generate individuals of the second generation with more diversified characteristics [13]. Another research provided dimensionality reduction and feature extraction and proposed ensemble-based classification to develop a novel heart disease prediction system. Higher order statistics (degree of dispersion and qualitative assessment) and Statistical (central tendency) characteristics are extracted initially during the anticipated phase of feature extraction. Unfortunately, the "curse of dimensionality" appears to be the primary problem in this case, therefore the higher-dimensionality characteristics have to be reduced to lower ones. PCA-based feature reduction was applied. The anticipated ensemble classifier received as input factors these reduced-dimensional characteristics, which include RF, SVM and KNN. Afterward, the categorized results of all three-classifiers were sent to the optimized NN as input, where the training is done by a novel S-CDF optimization algorithm and an enhanced sea lion algorithm by adjusting the ideal weights. The results from optimized NN are more correct. M8's accuracy (=0.957152) is 17.11, 8.5, 17.11, 2.6, 17.11, 2.6, and 13.7% superior to M_1, M_2, M_3, M_4, M_5, M_6, and M_7 respectively. The suggested study has more sensitivity, specificity, and accuracy than previous efforts. Future data dimensionality can be increased to evaluate the heart disease prediction system's performance and optimize prediction algorithms to increase the prediction rate [14]. Normal sinus rhythm and 40 components of ECG signal analysis are used to diagnose six forms of cardiac arrhythmia. The proposed Computer-aided diagnosis (CAD) technology might enable physicians to enhance clinical decision-making accuracy. This study extracts frequency, morphological and nonlinear indices, combines them, identifies the most correlated features using a meta-heuristic multiobjective optimization technique, and classifies them using various algorithms for machine learning. The suggested technique is efficient, automated, low computing complexity, and rapid in diagnosing heart illnesses using ECG signal processing. The experimental findings show an improvement in cardiac anomaly detection precision. Using the FF net classifier, yields the greatest accuracy among the seven classes, validating the reliability of this phrase. The presented method has been shown to be useful in classifying different types of cardiac arrhythmias, with a high degree of accuracy when compared to other approaches using comparable datasets [15].

2.2 MH Algorithms for Breast Cancer Diagnosis

Diagnostic ultrasound imaging methods are useful in identifying breast cancer. CAD is a model that can assist physicians to make precise judgment calls. Using a wavelet neural network (WNN) and the grey wolf optimization method, the authors of this research suggest a computer aided diagnosis (CAD) model for the diagnosis of breast cancer. Images were enhanced with a sigmoid filter for more contrast, and speckle noise was eliminated with IDAD in this setup. The ROI taken from the preprocessed picture was used to obtain a region of interest, after which the texture and morphological characteristics were extracted and merged. Principal component analysis (PCA) was employed to decrease the feature dimensions. The classification challenge was finally performed using grey wolf optimization (GWO)—wavelet neural network (WNN). Classical WNN necessitates more time invested in optimizing and training the parameters. The training and computational periods were shortened since GWO was used to fine-tune the WNN parameters in this suggested strategy. When compared to its contemporaries in terms of classification accuracy, the proposed CAD model system significantly excelled in the simulation results. GWO-WNN robustness, shorter training data, and convergence speed are its main benefits. The GWO algorithm outperforms other bio-inspired algorithms because of its nature-inspired leadership ability mechanism. Most feature selection approaches used meta-heuristic algorithms. The suggested CAD system tuned WNN parameters using the meta-heuristic approach to decrease computing cost and to prevent the local minima problem, improving performance. Cat swarm optimization, ant lion optimization, and other meta-heuristic techniques will be researched to increase classification accuracy [16].

2.3 MH Algorithms for Parkinson's Diagnosis

The novel metaheuristic algorithm was developed to identify Parkinson's disease early, and the findings demonstrate that it predicts the illness with 100% accuracy. Hence, the optimized crow search algorithm (OCSA) facilitates early therapy. The optimized OCSA was designed to detect Parkinson's disease and 20 benchmark datasets using KNN, Decision Tree models, and Random Forest to minimize features, accuracy, and computing time. Using the Random Forest model with the innovative optimized crow search algorithm yields improved results. For 20 benchmark datasets, the proposed algorithm selected less features with a precision of 88.2% and a computational time of 0.20 s compared to the original chaotic crow search algorithm's precision of 84.2% and 0.40 s computational time. The rate of convergence shows the robustness of the suggested algorithm's fitness function. Several study fields can use the proposed algorithm like Handwriting-exam-based Parkinson's diagnostic. Researchers and practitioners can use the algorithm for various PD diagnoses [17].

2.4 MH Algorithms for Prediction of Alzheimer's Disease

Several applications now prioritize image segmentation optimization. The whale optimization algorithm (WOA) solves numerous real-time issues and finds the global optimal solution, is an MH-optimized algorithm. WOAs may choose local optima over global optima as issue complexity rises. This may impair optimal solution finding. This research offers a hybrid algorithm that combines WOA and GWO to segment brain subregions such as the grey matter (GM), white matter (WM), ventricles, corpus callosum (CC), and hippocampus (HC). WOA and GWO comprise this hybrid combination. The hybrid WOA and GWO approach segments brain subregions (SRs) to diagnose Alzheimer's disease (AD) (H-WOA-GWO, which is denoted as HWGO). The segmented region has 92% accuracy after validation with a different measure. After segmentation, the deep learning classifier classified normal and Alzheimer's disease pictures. WOA/GWO accuracy is 90%. So, the recommended strategy is extremely effective for selecting the best solution and is combined with a deep learning algorithm for classification [18].

2.5 MH Algorithms for Prediction of Chronic Kidney Disease and Bone Disorders

Chronic kidney disease, mineral and bone disorders (CKD-MBD) cause vascular and cardiac calcification, which compromises blood pressure compensation. The researcher used BSWEGWO_KELM, a feature selection framework based on an optimized GWO algorithm and kernel extreme learning machine (KELM), to examine 1940 data from 178 Haemodialysis (HD) sufferers. The BSWEGWO_KELM method was evaluated using global optimization and feature selection tests on HD and public datasets. The BSWEGWO_KELM can be capable to screen out critical markers including dialysis vintage, intact parathyroid hormone (iPTH), mean arterial pressure (MAP) and alkaline phosphatase (ALP), according to experimental data. Hence, BSWEGWO_KELM can forecast intradialytic hypotension (IDH) accurately and practically [19].

2.6 MH Algorithms for Immunity Based Ebola Optimization Search Algorithm for Feature Extraction Minimization and Digital Mammography Reduction Using CNN Models

The novel revolutionary biology-dependent metaheuristic algorithm Ebola optimization search algorithm (EOSA) solved deep learning issues rapidly and precisely in recent research. The current EOSA metaheuristics use the spread of the Ebola virus and related sickness to explore by what means exploitation and exploitation stages of optimization may assist solve specific medical optimization challenges.

Researchers created SEIR-HVQD by improving the SIR model of Ebola. The model arises from incorporating into the basic SIR (Susceptible, Infected, Recovered) model the concepts of Exposed (E), Hospitalized (H), Quarantined (Q), Vaccinated (V) and Dead (D) [20].

2.7 MH Algorithm for Classification of White Blood Cells in Healthcare Informatics

In this procedure, blood is examined through blood count using manual and/or automatic procedures. In the search for medical innovations, this necessitated a Deep Learning framework for categorizing white blood cell subtypes in digital images, achieving effectiveness and dependability while attempting to make the methodology more accessible to underprivileged populations. Given this, systems that enable reliable, low-cost medical report procurement are important. For this, it was developed in python utilizing Jupyter notebook and analysis was done using a dataset of 12,500 digital photos of human blood smear fields with nonpathological leukocytes. The approach's accuracy was 85.72%, confirming its excellent dependability. So, the proposal is an accurate, reliable, and economical approach that may be used as a third-practicable blood count procedure among impoverished populations in undeveloped and developing nations [21].

2.8 MH Algorithm for EMG Classification Utilizing PSO Optimized SVM for Neuromuscular Diseases Prognosis

SVM is widely utilized for biological signal classification. This work developed a novel PSO-SVM model to improve EMG signal categorization accuracy by combining PSO and SVM considerably affecting the accuracy of classification. This optimization strategy involves setting kernel parameters in the SVM training method, which affects the accuracy of classification. EMG signals classified the research as neurogenic, normal, or myopathic. The suggested technique divided EMG signals into frequency subbands employing the discrete wavelet transform (DWT) and generated statistical characteristics to reflect wavelet coefficient distribution. The results demonstrate that the SVM approach outperforms standard machine learning methods and that the planned PSO-SVM classification system may increase classification accuracy. The PSO-SVM had 97.41% accuracy on 1200 signals of EMG from 27 subject records, against 96.75%, 95.17%, and 94.08% for the SVM, KNN, and RBF classifiers, respectively. PSO-SVM uses several SVMs as its core to diagnose neuromuscular diseases [22].

2.9 MH Algorithm Using Hybrid Case-Based Reasoning and Particle Swarm Optimization (PSO) Approach to the Detection of Hepatitis Disease

Physicians struggle to diagnose diseases. Misdiagnosis can lead to fatalities. In this context, expert systems and artificial intelligence approaches are used to decrease errors. The diagnoses of serious hepatitis use a combination of Case-Based Reasoning (CBR) and PSO. The CBR approach pre-processes the data and extracts each field's weight that has an influence on diagnosis, followed by PSO clustering. Each record's categorization and patient record are determined by PSO. The CBR-PSO technique outperformed FDT, KNN, SVM, PSO, and NB in hepatitis illness diagnosis with 94.58% accuracy. Comparatively, this technique works better when compared to different algorithms. This approach and fuzzy logic will be used in medical data by future researchers [23].

3 Meta-Heuristic Algorithms in Medical Image Segmentation

Recently, academics have focused on soft computing methods for medical data issue resolution. Similarity measurements are used to segment images. Medical image anomalies volumetric analysis is used for illness detection. Meta-heuristic algorithms support segmentation approaches recently. Highlights of meta-heuristic methods are presented. Meta-heuristic stochastic algorithms such as cellular automata, Memetic algorithms, ant colony optimization, particle swarms, evolutionary computation, Tabu search, and simulated annealing [24] (Fig. 2).

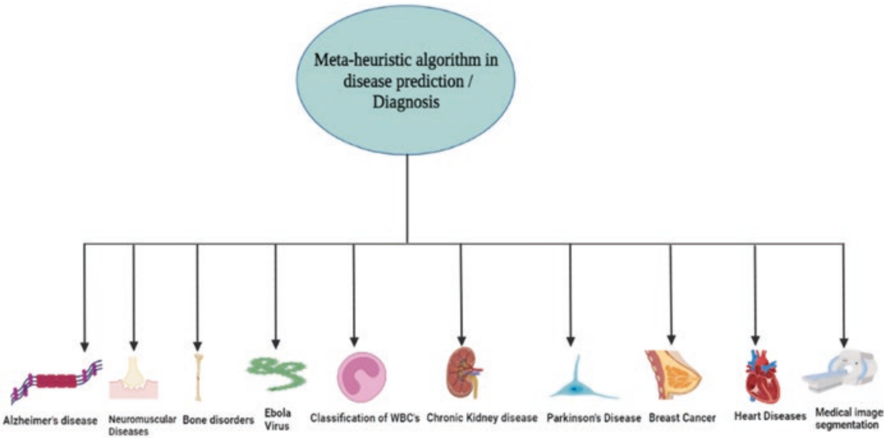


Fig. 2 Meta-heuristic algorithm helps in the prediction/diagnosis of diseases

4 Conclusion

Meta-heuristics algorithms rapidly explore the enormous solution space of candidate characteristics and models to predict diagnose and complicated diseases. Genetic algorithms, particle swarm optimization, and simulated annealing can optimize feature and model parameter selection to improve prediction. Meta-heuristics may become more important in illness prediction as healthcare data grows in volume and complexity. As processing power and machine learning improve, meta-heuristics algorithms will become more complex and accurate, making them a vital tool for healthcare practitioners.

References

1. Kadu, P., & Buchade, A. (2019). Non-communicable disease prediction system using machine learning. *International Journal of Scientific & Technology Research, 8*(9), 1307.
2. Dai, W., Brisimi, T. S., Adams, W. G., Mela, T., Saligrama, V., & Paschalidis, I. C. (2015, March 1). Prediction of hospitalization due to heart diseases by supervised learning methods. *International Journal of Medical Informatics, 84*(3), 189–197.
3. Calamuneri, A., Donato, L., Scimone, C., Costa, A., D'Angelo, R., & Sidoti, A. (2017). On machine learning in biomedicine. *Life Safety and Security, 5*(12), 96–99.
4. Anter, A. M., Bhattacharyya, S., & Zhang, Z. (2020). Multi-stage fuzzy swarm intelligence for automatic hepatic lesion segmentation from CT scans. *Applied Soft Computing, 96*, 106677.
5. Anter, A. M., Huang, G., Li, L., Zhang, L., Liang, Z., & Zhang, Z. (2020). A new type of fuzzy-rule-based system with chaotic swarm intelligence for multiclassification of pain perception from fMRI. *IEEE Transactions on Fuzzy Systems, 28*(6), 1096–1109.
6. Jayatilake, S. M., & Ganegoda, G. U. (2021, January). Involvement of machine learning tools in healthcare decision making. *Journal of Healthcare Engineering, 27*, 2021.
7. Kumar, S. N., Fred, A. L., Miriam, L. J., Padmanabhan, P., Gulyás, B., Kumar, A., & Dayana, N. (2022, January 1). Improved crow search algorithm based on arithmetic crossover—A novel metaheuristic technique for solving engineering optimization problems. In *Multi-objective combinatorial optimization problems and solution methods* (pp. 71–91). Academic Press.
8. Anter, A. M., & Ali, M. (2020). Feature selection strategy based on hybrid crow search optimization algorithm integrated with chaos theory and fuzzy c-means algorithm for medical diagnosis problems. *Soft Computing, 24*(3), 1565–1584.
9. Anter, A. M., & Zhang, Z. (2020). E-health Parkinson disease diagnosis in smart home based on hybrid intelligence optimization model. In *Proceedings of the international conference on advanced intelligent systems and informatics 2019* (pp. 156–165). Springer International Publishing.
10. Anter, A. M., & Zhang, Z. (2019, October). Adaptive neuro-fuzzy inference system-based chaotic swarm intelligence hybrid model for recognition of mild cognitive impairment from resting-state fMRI. In *International workshop on predictive intelligence in medicine* (pp. 23–33). Springer International Publishing.
11. Muthulakshmi, P., & Parveen, M. (2021, December 1). Investigation study on heart disease prediction with patient healthcare data. *Infocomp Journal of Computer Science, 20*(2).
12. Khourdifi, Y., & Bahaj, M. (2019, February). Heart disease prediction and classification using machine learning algorithms optimized by particle swarm optimization and ant colony optimization. *International Journal of Intelligent Engineering and Systems, 12*(1), 242–252.
13. Fofanah, A. J., Bundu, H. R., & Kargbo, J. G. (2022). A generic heart diseases prediction and application of genetic algorithms in healthcare systems: Genetic algorithm and machine learning algorithm approaches. *International Journal of Health Sciences, 6*, 12264–12290.

14. Kumar, P. R., Ravichandran, S., & Narayana, S. (2020, November 26). Ensemble classification technique for heart disease prediction with the meta-heuristic-enabled training system. *Bio-Algorithms and Med-Systems, 17*(2), 119–136.
15. Mazaheri, V., & Khodadadi, H. (2020, December 15). Heart arrhythmia diagnosis based on the combination of morphological, frequency and nonlinear features of ECG signals and meta-heuristic feature selection algorithm. *Expert Systems with Applications, 161*, 113697.
16. Bourouis, S., Band, S. S., Mosavi, A., Agrawal, S., & Hamdi, M. (2022, June 13). Meta-heuristic algorithm-tuned neural network for breast cancer diagnosis using ultrasound images. *Frontiers in Oncology, 12*, 834028.
17. Gupta, D., Sundaram, S., Khanna, A., Hassanien, A. E., & De Albuquerque, V. H. (2018, May 1). Improved diagnosis of Parkinson's disease using optimized crow search algorithm. *Computers & Electrical Engineering, 68*, 412–424.
18. Dhakhinamoorthy, C., Mani, S. K., Mathivanan, S. K., Mohan, S., Jayagopal, P., Mallik, S., & Qin, H. (2023, February 24). Hybrid whale and gray wolf deep learning optimization algorithm for prediction of Alzheimer's disease. *Mathematics, 11*(5), 1136.
19. Yang, X., Zhao, D., Yu, F., Heidari, A. A., Bano, Y., Ibrohimov, A., Liu, Y., Cai, Z., Chen, H., & Chen, X. (2022, June 1). An optimized machine learning framework for predicting intradialytic hypotension using indexes of chronic kidney disease-mineral and bone disorders. *Computers in Biology and Medicine, 145*, 105510.
20. Oyelade, O. N., & Ezugwu, A. E. (2022, October 26). Immunity-based Ebola optimization search algorithm for minimization of feature extraction with reduction in digital mammography using CNN models. *Scientific Reports, 12*(1), 17916.
21. Monteiro, A. C., Iano, Y., França, R. P., & Arthur, R. (2021, July 14). A metaheuristic algorithm for classification of white blood cells in healthcare informatics. In *Metaheuristics in machine learning: Theory and applications* (pp. 219–238). Springer International Publishing.
22. Subasi, A. (2013, June 1). Classification of EMG signals using PSO optimized SVM for diagnosis of neuromuscular disorders. *Computers in Biology and Medicine, 43*(5), 576–586.
23. Neshat, M., Sargolzaei, M., Nadjaran Toosi, A., & Masoumi, A. (2012). Hepatitis disease diagnosis using hybrid case based reasoning and particle swarm optimization. *International Scholarly Research Notices, 2012*, 609718.
24. Dey, N., & Ashour, A. S. (2018). Meta-heuristic algorithms in medical image segmentation: A review. In *Advancements in applied metaheuristic computing* (pp. 185–203). IGI Global.

Printed rGO-Based Temperature Sensor for Wireless Body Area Network Applications

Asha Susan John and Kalpana Murugan

1 Introduction

A wireless body area network (WBAN) is a type of wireless network that connects small, low-power, and lightweight devices worn on the body to monitor various physiological parameters, such as heart rate, blood pressure, temperature, and other health-related information. The devices communicate with each other using wireless communication technologies such as Bluetooth, Zigbee, and Wi-Fi. WBANs are primarily used for healthcare applications and can provide real-time monitoring and analysis of a patient's health status [32, 35]. They are often used for remote patient monitoring, where patients can be monitored from their homes or other remote locations, allowing healthcare providers to make timely decisions and provide appropriate care. WBANs can also be used for sports and fitness monitoring to track an individual's physical activity, heart rate, and other physiological parameters (Fig. 1).

Wireless Body Area Networks (WBANs) offer several advantages for healthcare and wellness monitoring [33]. They are:

1. **Continuous Monitoring:** WBANs provide continuous and real-time monitoring of various physiological parameters, allowing healthcare providers to detect early warning signs of health problems and prevent complications. WBANs can also be used for sports and fitness monitoring, providing individuals with real-time feedback on their physical activity and health status.
2. **Non-Invasive Monitoring:** WBAN devices are typically small, lightweight, and non-invasive, making them easy and comfortable to wear. This can improve

A. S. John · K. Murugan (✉)
Department of Electronics and Communication Engineering, School of Electronics, Electrical and Biomedical Technology, Kalasalingam Academy of Research and Education, Virudhunagar, Tamil Nadu, India

A. M. Anter et al. (eds.), *Nature-Inspired Methods for Smart Healthcare Systems and Medical Data*, https://doi.org/10.1007/978-3-031-45952-8_10

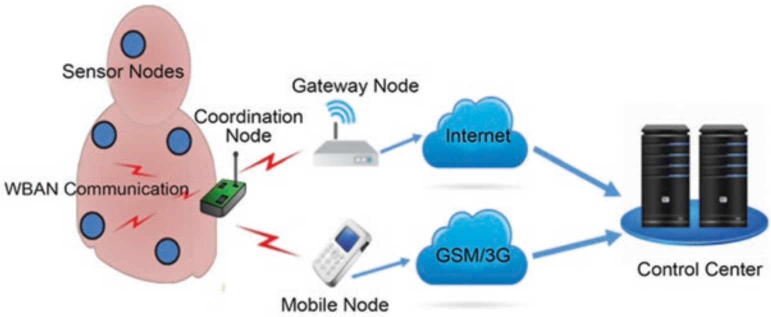


Fig. 1 WBAN Architecture [29]

patient compliance with monitoring protocols, allowing for longer and more accurate monitoring periods.

3. **Mobility:** WBAN devices are designed for mobility, allowing patients to move around freely while still being monitored. This can be particularly useful for patients who require continuous monitoring but want to avoid being confined to a hospital bed.
4. **Personalized Care:** WBANs can provide personalized care by monitoring an individual's specific health parameters and providing tailored feedback and recommendations. This can help individuals make informed decisions about their health and wellness.
5. **Reduced Healthcare Costs:** WBANs can help reduce healthcare costs by enabling remote patient monitoring and reducing the need for hospital visits. This can be particularly useful for individuals with chronic conditions who require frequent monitoring and care.

While Wireless Body Area Networks (WBANs) offer several advantages for healthcare and wellness monitoring, there are also some potential disadvantages to consider:

1. **Security and Privacy**: WBAN networks can collect sensitive personal data, which raises concerns about security and privacy. Data breaches and unauthorized access to sensitive data can compromise patient privacy and lead to identity theft or other forms of harm.
2. **Interference:** WBAN networks can be susceptible to interference from other wireless devices, which can cause disruptions in the data transmission and affect the accuracy of the monitoring.
3. **Compatibility:** The compatibility of different WBAN devices and systems can be a challenge, as different devices may use different communication protocols and standards.
4. **Limited Range:** The range of WBAN networks is limited, which can make it difficult to monitor patients who are far away from the receiver or base station.
5. **Battery Life:** WBAN devices are typically powered by batteries, which can limit the duration of monitoring and require frequent recharging or replacement.

6. **Interoperability** is another challenge, as different WBAN devices may use different communication protocols and standards.

A wide range of sensors is used to monitor various physiological parameters of the body. These sensors are networked using wireless sensing protocols as shown in Fig. 2. The different types of sensors include:

1. **Temperature sensors:** These sensors are used to monitor body temperature and can be used to detect fever, hypothermia, or other changes in body temperature. They are commonly used in medical applications, such as monitoring patients with infectious diseases or post-operative care.
2. **Heart rate sensors:** These sensors measure the electrical signals generated by the heart and can be used to monitor heart rate, heart rhythm, and other cardiac parameters. They are commonly used in sports and fitness applications to track an individual's physical activity and monitor their cardiovascular health.
3. **Blood pressure sensors:** These sensors measure the pressure of blood flowing through the arteries and can be used to monitor blood pressure and other cardiovascular parameters. They are commonly used in medical applications to diagnose and manage hypertension and other cardiovascular conditions.
4. **Oxygen saturation sensors:** These sensors measure the oxygen saturation level in the blood and can be used to monitor respiratory function. They are commonly

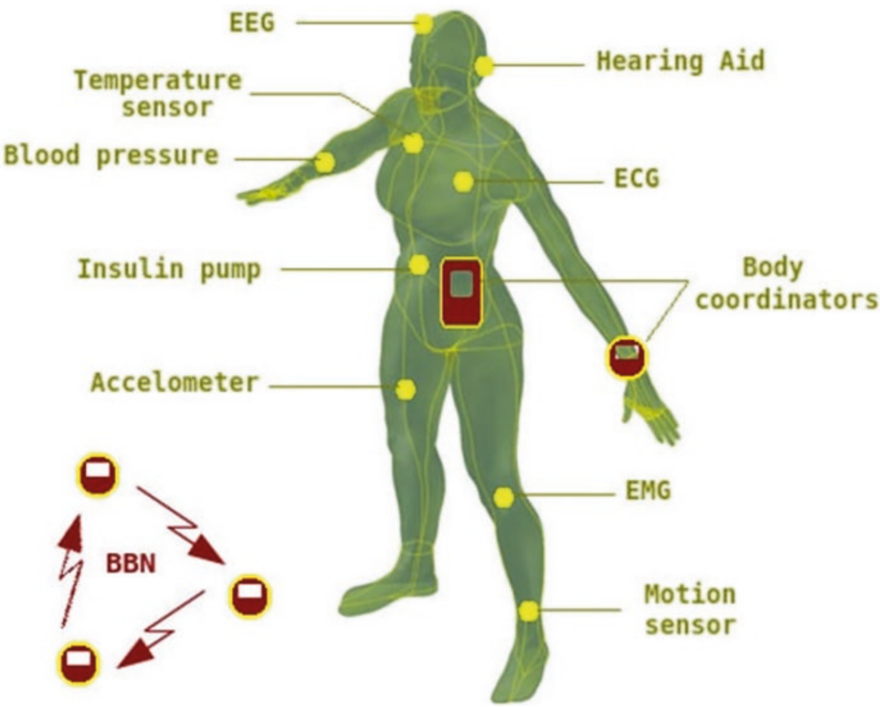


Fig. 2 Patient monitoring sensors in WBAN network [30]

used in medical applications to diagnose and manage respiratory conditions, such as chronic obstructive pulmonary disease (COPD) or sleep apnea.

5. **Glucose sensors:** These sensors measure the glucose level in the blood and can be used to monitor blood sugar levels in patients with diabetes. They are commonly used in medical applications to monitor diabetes management and prevent complications associated with high or low blood sugar levels.
6. **Accelerometers:** These sensors measure the acceleration of the body and can be used to monitor physical activity and detect falls or other accidents. They are commonly used in sports and fitness applications to track an individual's physical activity and monitor their overall health and wellness.

2 Background and Motivation

Temperature sensors are an essential type of sensor used in Wireless Body Area Networks (WBANs) for monitoring body temperature. WBANs are a type of wireless sensor network that enables real-time and continuous monitoring of an individual's physiological parameters, such as heart rate, blood pressure, and temperature, among others. WBANs have numerous applications in medical, sports, and fitness settings, and can help individuals and healthcare providers make informed decisions about their health and wellness [31]. In WBANs, temperature sensors are used to monitor body temperature, which is an important physiological parameter that can indicate various health conditions. Changes in body temperature can be a sign of fever, hypothermia, or other health conditions, and continuous monitoring of body temperature can help in the early detection and management of these conditions.

Temperature sensors used in WBANs are typically small and low power, allowing for continuous monitoring of body temperature without causing discomfort to the patient. They can be attached to the skin using adhesive patches or incorporated into clothing or wearable devices, providing real-time temperature readings that can be transmitted wirelessly to a central monitoring system. The commonly used temperature monitoring devices in WBAN are:

1. **Thermistors:** Thermistors are a type of resistor whose resistance changes with temperature. They are commonly used in medical applications to measure body temperature and can provide accurate temperature readings in a small form factor [26].
2. **Thermocouples:** Thermocouples are made of two dissimilar metals that generate a voltage proportional to the temperature difference between the two junctions. They are commonly used in industrial applications to measure temperature but can also be used in WBANs for body temperature monitoring.
3. **Infrared sensors:** Infrared sensors measure the temperature of an object by detecting the infrared radiation emitted by the object. They are commonly used in non-contact temperature measurements and can be used in WBANs for monitoring skin temperature.

4. **Fibre optic sensors:** fibre optic sensors use optical fibre's to measure temperature changes. They are commonly used in industrial and aerospace applications but can also be used in WBANs for monitoring body temperature.
5. **Integrated circuits (ICs):** Temperature sensors based on integrated circuits can provide accurate temperature readings in a small form factor. They are commonly used in consumer electronics and can be used in WBANs for monitoring body temperature.

These sensors are available in the market in the form of digital thermometers, Infrared thermometers, mercury thermometers, disposable thermometers, chemical thermometers and smart thermometers [38]. Digital thermometers use electronic sensors to measure body temperature. They are available in various forms, including oral, axillary, and rectal thermometers. They typically provide a fast and accurate reading and can display the temperature in Fahrenheit or Celsius. Infrared thermometers use infrared technology to measure body temperature without making contact with the skin. They can be used on the forehead or the ear and provide a quick and non-invasive reading. Mercury thermometers are a traditional type of thermometer that uses mercury to measure temperature. They are usually used for rectal temperature measurements and have largely been replaced by digital thermometers due to safety concerns related to the handling of mercury. Disposable thermometers are single-use thermometers that are often used in clinical settings to reduce the risk of cross-contamination. They are usually made from plastic or paper and are designed for use in the mouth or under the armpit. Chemical thermometers use chemical reactions to measure temperature changes. They are not commonly used for measuring human body temperature but are often used in laboratory settings for other applications. Smart thermometers are digital thermometers that are connected to a smartphone app. They can record and track temperature readings over time, and some models can provide additional health insights and recommendations.

A wearable thermometer is a type of wearable device that measures and monitors a person's body temperature. Unlike traditional thermometers that require physical contact with the body, wearable thermometers can be worn on the body, allowing for continuous temperature monitoring without the need for constant manual measurements. Wearable thermometers can come in various forms, including patches, wristbands, and smartwatches. They use sensors to detect changes in body temperature and transmit the data to a mobile device or other connected devices for further analysis. Wearable thermometers have become increasingly popular in recent years, especially during the COVID-19 pandemic, as they allow for remote temperature monitoring and can help detect potential infections early on. They are also useful for monitoring the temperature of individuals with chronic illnesses or for tracking the temperature of infants and young children. Wearable thermometers are a convenient and effective way to monitor body temperature continuously and remotely, making them a valuable tool for both medical professionals and individuals concerned about their health. Some of the wearable type of sensors used in market include:

1. **Smartwatches:** Many modern smartwatches, such as Apple Watch and Samsung Galaxy Watch, include sensors that can monitor body temperature. These sensors can measure skin temperature on the wrist and can provide continuous temperature tracking.
2. **Thermometer patches:** Thermometer patches are small, flexible adhesive patches that can be applied to the skin to monitor body temperature. They typically use a wireless transmitter to send data to a smartphone or other device for monitoring and analysis.
3. **Wireless temperature monitoring systems:** These are systems that can be worn on the body, such as a bracelet or necklace, and provide continuous temperature monitoring. They typically use wireless technology, such as Bluetooth or Wi-Fi, to transmit data to a smartphone or other device.
4. **Smart clothing:** Smart clothing includes clothing with embedded sensors that can monitor various physiological parameters, including body temperature. These sensors can be integrated into fabrics and can provide continuous temperature tracking.
5. **Earbuds:** Some earbuds, such as the Jabra Elite Sport earbuds, include sensors that can monitor body temperature through the ear canal. These sensors can provide continuous temperature tracking and can be used for sports and fitness tracking.

3 Literature Review

Monitoring body temperature is critical for identifying human health problems and reflecting physiological activities. In conventional temperature sensing, rigid temperature detectors, such as thermometers, are commonly utilised. The main weakness of this approach is its inability to mix with curved surfaces, such as human or animal bodies. Also, thermometers cannot be placed under the arm as required, making them inappropriate for children [2]. Furthermore, when inflexible temperature detectors come into contact with uneven surfaces, they become uncomfortable. As a result, a wearable, flexible temperature sensor capable of making direct contact with the human body is critical. The mechanical robustness, biocompatibility, multifunctionality, comfort, and comfort of flexible sensors enable next-generation wearable technologies. Two examples are RTDs and Thermistors [4]. Significant advances in the development and upgrading of flexible temperature sensors have been made to replace standard sensors on rigid substrates and provide a viable alternative for wearable applications [3]. Temperature sensors of various types include thermocouples and resistive temperature detectors (RTDs) [4].

The resistance of the conductor material in RTDs varies with temperature [1, 5, 6]. Thermistors, on the other hand, utilise variations in electric conductivity of a sensor material to compensate for temperature variation [7]. The main advantages of RTD temperature sensors are their low cost, good linearity, and simple fabrication technique [1, 6]. However, RTD has a slow response time and a low sensitivity. Thermistors, on the other hand, have advantages due to their qualities such as inexpensive cost, wide sensing range, quick response, and extremely accurate temperature measurements [7, 8]. The two thermometer (PTC) types are the negative temperature coefficient (NTC) Thermistor and the positive temperature coefficient. Resistance increases with resistance in the NTC, whereas resistance decreases with temperature in the PTC.The temperature coefficient of resistance (TCR), which represents the sensor‘s sensitivity, has attracted the most attention for all temperature sensors [6, 8]. The sensor becomes more sensitive to small temperature variations as the TCR value increases [9]. Equation 1 [1, 5, 6, 8, 9] drives the TCR.

$$TCR = \alpha = \left(\mathrm{R1} - Ro / Ro\left(\mathrm{T1} - \mathrm{To}\right)\right) \tag{1}$$

where R1 is the thermistor's resistance at absolute temperature T1[C] and R0 is the tested sample's initial resistance at [C]. The two most common processes utilised to make flexible and wearable sensors are photolithography and printing [10]. Photolithography is used to create high-performance gadgets. Unfortunately, due to the need for cleanroom facilities, significant material waste, and high fabrication costs, these procedures are unfortunately prohibitively expensive [10, 11]. As a result, low-cost printing processes that are simple to use and have substantially lower fabrication costs are in demand. Because of inkjet printing, the use of a stencil mask or clean-room lithography is no longer required [12]. It offers precise and continuous deposition of micro- and nanomaterials onto a range of surfaces under ambient conditions [13]. Inkjet also provides simple, flexible print patterns with high resolution, low material usage, and cost savings [10, 13].

Temperature sensors are used in a wide range of applications, from industrial control to biological monitoring. Several types of temperature sensors have been developed and intensively explored in response to the growing demand for high precision and dependability. The following are the most popular temperature sensor technologies, which range in responsiveness and accuracy from excellent to poor:

- Negative Temperature Coefficient (NTC) Thermistors.
- Resistance Temperature Detectors (RTDs)
- Thermocouples.
- Semiconductor-Based Sensors.
- Infrared (IR) temperature sensors
- Fiber optic temperature sensors

A thermistor is a type of thermally sensitive resistor that changes resistance as a function of temperature continuously, slowly, and incrementally. An NTC thermistor has higher resistance at low temperatures. The resistance gradually reduces as the temperature rises, according to the R-T table. Resistance rises per °C allow for exact reflection of minor variations [19]. The output of an NTC thermistor cannot be linearized due to its exponential nature; but, depending on the application, it can be. The effective operating range of regular thermistors or glass-encapsulated thermistors is 150 °C [18]. With a resistance temperature detector, or RTD, the resistance of the RTD element changes with temperature. To make an RTD, a wire or, for greater precision, a film is wrapped around a ceramic or glass core. Although nickel and copper are less expensive, they are not as stable or reproducible as platinum, which generates the most exact RTDs. Platinum RTDs, while more expensive than copper or nickel RTDs, produce a highly precise linear output over a temperature range of −200 to 600 °C [17]. A thermocouple is made up of two wires of different metals that are electrically connected at two points. The changing voltage generated between these two metals reflects proportional temperature changes [14]. Because thermocouples are nonlinear devices, they must be converted using a table when utilised for temperature control and correction. A lookup table is widely used for this. Thermocouples operate throughout the widest temperature range, from −200 to 1750 °C, although their accuracy is limited (0.5–5 °C) [15].

A semiconductor-based temperature sensor is commonly seen in integrated circuits (ICs). To track temperature changes, these sensors use two identical diodes with temperature-sensitive voltage vs. current characteristics. Although having the lowest accuracy of the basic sensor types, they provide a linear response. Furthermore, these temperature sensors respond slowly over the entire temperature range (−70 to 150 °C) [22]. Infrared (IR) temperature sensors enable accurate non-contact temperature measurement in medical applications. The most common applications for this type of temperature sensor are to measure skin, forehead, or ear temperatures. Fiber optic temperature sensors have gained popularity in recent years due to its high precision, resistance to electromagnetic interference, and capacity to detect temperature at multiple locations at the same time. They require an expert for signal processing, though, and are more expensive than conventional sensors [22].

In general, the temperature sensor to be used is determined by the specific application and needs. Despite the ubiquitous use of thermocouples and RTDs, recent advances in fibre optic temperature sensors provide a potential opportunity for highly accurate and scattered temperature readings. The four most common types of temperature sensors in modern electronics are thermocouples, RTDs (resistance temperature detectors), thermistors, and semiconductor-based integrated circuits (IC). A thermocouple is a type of temperature sensor. This sensor is made up of two distinct metal wires that are linked at one end and coupled to a thermocouple thermometer or other thermocouple-capable device at the other [15]. When properly configured, thermocouples may provide temperature measurements over a wide temperature range. A thermocouple is a sensor that measures temperature. It is constructed of two distinct metals that are joined at one end. When the junction of the two metals is heated or cooled, a voltage that is proportional to temperature is created [16]. Thermocouples are a popular temperature sensor due to its simplicity, wide temperature range, and low cost. Unfortunately, they have measurement errors built in, and factors such as temperature gradients and electromagnetic interference can affect their precision. Because of their versatility as temperature sensors, thermocouples are widely used in a wide range of applications, ranging from household appliances and utilities to industrial thermocouples [15]. Because thermocouples come in such a wide range of models and technical specifications, understanding their fundamental construction, operation, and ranges is critical in order to select the optimal kind and material for your application [14].

Thermocouple sensors work on the following principle: A continuous current flows through the thermoelectric circuit when two wires of different metals are joined at both ends and one of the ends is heated. The most common thermocouple types are “Base Metal” thermocouples, which include Types J, K, T, and E. In high-temperature applications, noble metal thermocouples of types R, S, and B are used [16]. The capacity of thermocouple sensors to provide exact measurement values for industrial applications requiring high temperatures is its primary advantage [14]. They are long-lasting, inexpensive, fast to respond, and incredibly effective for many years to come. Straightforward vast range of temperatures selection Rugged It is “self-sufficient.”

The resistance temperature detector is another typical type of temperature sensor that relies on the idea of temperature-dependent resistance (RTD). They provide more accuracy and repeatability than thermocouples, making them excellent for applications requiring highly accurate temperature measurements [18]. Yet, they only work within a small temperature range and require more expensive signal conditioning and processing equipment [17]. Identifying RTDs. The resistance of a sensor known as an RTD (Resistance Temperature Detector) fluctuates with temperature. As the temperature of the sensor rises, so does the resistance. Temperature and resistance have a well-established and constant relationship [18]. Temperature is measured with an RTD, which stands for “Resistance Temperature Detector,” because its resistance varies with temperature. The resistance of the RTD increases linearly with temperature. RTDs are commonly referred to be wire wrapped. They’re manufactured with a glass or ceramic core wrapped in microscopic wire. The wire was made of platinum [18]. The RTD elements are often housed in a protective probe to protect them from the environment and to boost their robustness, which is another noteworthy aspect [17].

Thin film RTDs are less expensive. They are comprised of a basic ceramic that has been coated with a tiny layer of platinum. Let us now look at how an RTD works. As previously stated, an RTD is made up of insulated Platinum wires and a resistance element [17]. RTDs may occasionally incorporate three or even four wires to increase accuracy and minimise connection lead resistance concerns [18]. Platinum is utilised as the resistance element because it is chemically inert, has a wide temperature range, and is exceptionally long-term stable [17]. The RTD functions on a fundamental principle. When a metal’s temperature rises, so does its resistance to the flow of electricity. When an electrical current is conducted through the sensor, the resistance element measures the resistance of the current passing through it [18]. Electrical resistance increases as the temperature of the resistance element rises. The electrical resistance is measured in Ohms. The resistance value can then be converted into temperature using the element’s characteristics. An RTD usually reacts between 0.5 s and 5 s. As a result, they are well suited for a wide range of applications [17]. RTD sensors have numerous benefits. Long-term steadiness, extreme accuracy, repeatability, consistency, and ability to deliver precise measurement even in severe settings. Platinum RTDs can sustain a wider range of temperatures. When compared to thermocouples and thermistors, they are the most dependable and consistent throughout time [16, 17]. RTD disadvantages: A more sophisticated measurement circuit is required due to low sensitivity and a high starting cost. Larger than typical bulbs; low absolute resistance; a current source is required; less rugged in a high vibration environment; a bridge circuit with power supply is required [17]. Thermistors are temperature sensors that operate on the premise that the resistance of semiconductor materials changes with temperature. They have high sensitivity, short reaction times, and low pricing. Self-heating and nonlinearity, however, can reduce their accuracy [18].

A thermistor is a resistor with temperature-dependent resistance or a resistance thermometer. The term is made up of the words "thermal" and "resistor." It is made of metallic oxides that are shaped into a bead, disc, or cylinder shape before being covered in an impermeable substance such as epoxy or glass [19]. The temperature sensor in this case is an inkjet printer thermistor. The resistance of thermos witches, also known as temperature-dependent resistors, fluctuates with temperature. They are sensitive to even minor temperature swings, making them suitable for temperature control. They can measure liquids, gases, or solids depending on the type of thermistor. The body temperatures of infants and children are measured with thermometers [19]. To determine if the sensor is a thermistor or an RTD, measure the resistance between the two dissimilarly coloured wires: An RTD PT100 has a resistance of 100 ohms at 0 °C. The resistance of an RTD PT1000 at 0 °C is 1000 ohms [20]. The thermistor's key advantages are its high sensitivity, small heat capacity, and quick reaction; nevertheless, its main disadvantages are its low interchangeability and non-linear thermoelectric properties, which broaden the measurement range [19].

Thermistors can have errors as low as less than 0.001 °C when utilized over a restricted temperature range of 0–50 °C. Thermistors have more sensitivity than other temperature sensors, which enables them to function well over a narrow temperature range [20].

- They respond quickly
- They are inexpensive and therefore inexpensive to replace.
- Simple to use.
- Compact dimensions that enable them to occupy the tiniest areas.

Semiconductor-based temperature sensors, also known as IC sensors, are made up of two identical diodes on a dual integrated circuit (IC). Diodes and temperature-sensitive voltage are used to measure temperature. Although these sensors produce a relatively linear output, their accuracy falls between 1 °C and 5 °C [21]. An IC temperature sensor is a two-terminal integrated circuit temperature transducer that creates an output current proportional to the absolute temperature. The sensor package is small, has a fast response time, and has a low thermal mass [21]. Temperatures typically vary from −58 to 302 °F (55 to 150 °C). The output of a solid-state sensor can be either analog

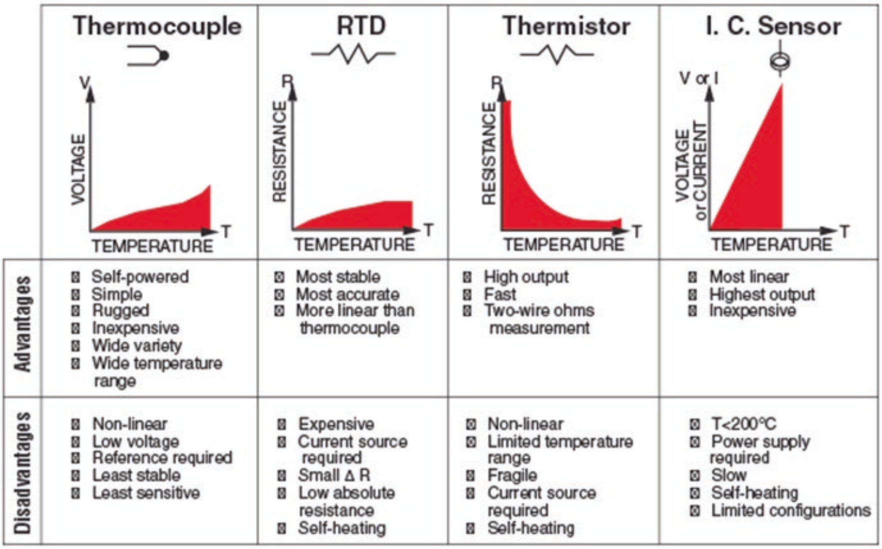


Fig. 3 The comparative chart for various temperature sensors [23]

or digital [20]. The IC transducer is a small device with a fast response time and low thermal mass. Ideally suited for circuit board temperature management, computer CPU temperature control, and telecommunications (cell phones) applications [22]. The advantages and disadvantages of IC sensors, as well as the possibility of analogue or digital outputs and inexpensive pricing, are strengths.

- Direct temperature measurement (1.000 = 100C and 298A = 298 K or 25 °C for some analog devices)
- Linear output, no curve fitting
- Direct voltage, current, or digital output needing no additional circuitry
- Wider interchangeability than most RTDs and thermistors

Limited temperature range: −55 to 150 °C Max

- Large variance in accuracy across models.
- With some immersion systems, small package sizes might be a barrier to low-cost applications (Fig. 3).

4 Materials and Methods

A printed temperature sensor is a type of sensor that is fabricated using printing techniques such as screen printing, inkjet printing, or flexography. These sensors are designed to measure temperature changes in various environments and are typically made from materials such as conductive inks, polymers, or carbon nanotubes.

The design of a printed temperature sensor depends on the specific application for which it will be used. Some sensors are designed for use in wearable devices, while others are intended for use in industrial or automotive applications. The sensor‘s design must take into account factors such as temperature range, sensitivity, accuracy, and response time. The fabrication process for printed temperature sensors involves depositing the sensor‘s material onto a substrate using printing techniques. This process allows for the creation of sensors with complex shapes and patterns, making them suitable for a wide range of applications. Printed temperature sensors have several advantages over traditional temperature sensors, including low cost, flexibility, and ease of integration with other electronic components. They are also more environmentally friendly than traditional sensors, as they use fewer materials and generate less waste during the manufacturing process. The design and fabrication of printed temperature sensors have opened up new possibilities for temperature sensing applications, particularly in the fields of wearables, smart homes, and the Internet of Things (IoT).

4.1 Materials

The choice of material for a temperature sensor that can be used for WBAN application includes the selection of material for temperature sensing as well as the material choice for the substrate on top of which the sensor is fabricated.

4.1.1 Sensing materials

There are several materials that can be used for the fabrication of a printed temperature sensor. These materials are typically chosen based on their thermal and electrical properties, as well as their ability to be printed using various printing techniques. Some common materials used for printed temperature sensors include:

1. **Conductive inks:** These are inks that contain conductive particles such as silver, copper, or carbon. They are commonly used to print the sensing element and the interconnects of the sensor.
2. **Polymers:** Polymers such as polyimide, polycarbonate, or polyethylene terephthalate (PET) are commonly used as substrates for printed temperature sensors. These materials offer good thermal stability and flexibility.
3. **Carbon nanotubes:** Carbon nanotubes are an emerging material for printed temperature sensors. They offer high thermal conductivity, low thermal mass, and excellent mechanical properties [27].
4. **Graphene:** Graphene is another emerging material for printed temperature sensors. It offers high thermal conductivity and excellent electrical properties [24].
5. **Thermocouples:** Thermocouples are often used as temperature sensors in industrial applications. They are made from two dissimilar metals such as cop-

per and constantan, and generate a voltage proportional to the temperature difference between the two junctions.

6. **Metal oxides:** Metal oxide materials such as zinc oxide, tin oxide, and indium oxide are commonly used in printed temperature sensors. These materials have high thermal stability and can be printed using inkjet or screen printing techniques.
7. **Conductive polymers:** Conductive polymers such as polyaniline, polypyrrole, and polythiophene are another class of materials that can be used for printed temperature sensors. These materials offer high electrical conductivity and can be printed using inkjet or spray coating techniques.
8. **Nanoparticles:** Nanoparticles of metals such as gold, silver, or copper can also be used for printed temperature sensors. These materials offer high thermal conductivity and can be printed using inkjet or aerosol-jet printing techniques.
9. **Ceramic materials:** Ceramic materials such as alumina, zirconia, or silicon carbide can be used as substrates for printed temperature sensors. These materials offer excellent thermal stability and can be printed using screen printing or inkjet printing techniques.
10. **Organic materials:** Organic materials such as organic semiconductors, organic dyes, or organic molecules can be used for printed temperature sensors. These materials offer good flexibility and can be printed using inkjet or spray coating techniques.
11. **Metal oxides:** Metal oxide materials such as zinc oxide, tin oxide, and indium oxide are commonly used in printed temperature sensors. These materials have high thermal stability and can be printed using inkjet or screen printing techniques.
12. **Conductive polymers:** Conductive polymers such as polyaniline, polypyrrole, and polythiophene are another class of materials that can be used for printed temperature sensors. These materials offer high electrical conductivity and can be printed using inkjet or spray coating techniques.
13. **Nanoparticles:** Nanoparticles of metals such as gold, silver, or copper can also be used for printed temperature sensors. These materials offer high thermal conductivity and can be printed using inkjet or aerosol-jet printing techniques.
14. **Ceramic materials:** Ceramic materials such as alumina, zirconia, or silicon carbide can be used as substrates for printed temperature sensors. These materials offer excellent thermal stability and can be printed using screen printing or inkjet printing techniques.
15. **Organic materials:** Organic materials such as organic semiconductors, organic dyes, or organic molecules can be used for printed temperature sensors. These materials offer good flexibility and can be printed using inkjet or spray coating techniques [25].

The choice of material for a printed temperature sensor will depend on several factors, including the required temperature range, sensitivity, accuracy, and cost.

4.1.2 Substrate Material

The choice of substrate for a printed temperature sensor will depend on several factors, such as the required temperature range, the desired accuracy and sensitivity, the manufacturing process, and the cost. Some common substrates used for printed temperature sensors include:

1. **Flexible polymeric substrates:** Polymeric substrates, such as polyimide, polyethylene terephthalate (PET), and polydimethylsiloxane (PDMS), are widely used in printed temperature sensors because they offer excellent flexibility, low thermal mass, and good thermal stability. These substrates are suitable for printing using techniques such as screen printing, inkjet printing, and roll-to-roll printing.
2. **Rigid substrates:** Rigid substrates such as glass, silicon, or ceramic materials can be used in printed temperature sensors where high accuracy and stability are required. These substrates are suitable for printing using techniques such as photolithography or inkjet printing.
3. **Paper-based substrates:** Paper-based substrates such as cellulose-based paper or coated paper are low-cost options for printed temperature sensors. These substrates are suitable for printing using techniques such as screen printing or inkjet printing.
4. **Metal foils:** Metal foils such as copper, aluminum, or stainless steel can be used as a substrate for printed temperature sensors. These substrates offer good thermal conductivity and stability but may not be as flexible as polymeric substrates. Metal foils are suitable for printing using techniques such as screen printing or flexography.
5. **Textiles:** Textiles such as cotton, silk, or polyester can be used as a substrate for printed temperature sensors in wearable applications. These substrates offer excellent flexibility, breathability, and comfort. Textiles are suitable for printing using techniques such as inkjet printing or screen printing.
6. **Polyimide (PI):** Polyimide is a popular substrate material for printed temperature sensors because it offers good thermal stability, flexibility, and durability. It can withstand high temperatures and is commonly used in harsh environments.
7. **Polycarbonate (PC):** Polycarbonate is another popular substrate material for printed temperature sensors because it offers good thermal stability and is easy to print on. It is commonly used in low-temperature applications.
8. **Polyethylene terephthalate (PET):** PET is a widely used substrate material for printed temperature sensors because it is flexible, transparent, and offers good thermal stability. It is commonly used in consumer electronics applications.
9. **Ceramic:** Ceramic materials such as alumina, zirconia, or silicon carbide can be used as substrates for printed temperature sensors because they offer excellent thermal stability and can withstand high temperatures. They are commonly used in high-temperature applications.

10. **Metal foils:** Metal foils such as copper or aluminum can be used as substrates for printed temperature sensors because they offer good thermal conductivity and can withstand high temperatures. They are commonly used in high-temperature applications.
11. **Flexible printed circuit boards (FPCBs):** FPCBs are a type of substrate material that can be used for printed temperature sensors. They offer good flexibility, durability, and can be easily integrated into other electronic systems.
12. **Silicone rubber:** Silicone rubber is a flexible and durable substrate material that can be used for printed temperature sensors. It is commonly used in medical and wearable applications.
13. **Metals:** Metals such as stainless steel or titanium can be used as substrate materials for printed temperature sensors. They offer high thermal conductivity and can withstand harsh environments. However, they are typically more expensive than other substrate materials.
14. **Glass:** Glass is a substrate material that offers high thermal stability, chemical resistance, and can be easily printed on. It is commonly used in high-temperature applications and in harsh chemical environments.

The choice of substrate material will depend on the specific requirements of the sensor and the printing technique used to fabricate the sensor. For example, some substrates may be better suited for certain printing techniques than others. Additionally, the substrate material may affect the performance of the sensor, such as its sensitivity and response time, so careful consideration is required during substrate selection. It is important to select a substrate material that can withstand the operating temperature range of the sensor, is compatible with the printing method used, and provides the desired level of durability and flexibility.

4.2 Methods

A temperature sensor can be simulated using COMSOL Multiphysics. It uses finite element analysis (FEA) to simulate the behavior of a temperature sensor under different conditions. This allows designers to optimize the performance of the sensor and identify potential issues before it is fabricated. To design a temperature sensor using COMSOL, the model of the sensor should be first drawn. This model typically includes the geometry of the sensor, the materials used, and the boundary conditions. The heat transfer equations that govern the behavior of the sensor can then be defined and COMSOL's built-in solvers can be used to simulate the behavior of the sensor under different conditions. COMSOL also offers several features that make it ideal for designing temperature sensors. These include:

1. **Multi-physics simulation capabilities:** COMSOL can simulate the behavior of temperature sensors under different physical conditions, such as fluid flow and electromagnetic fields.

2. **User-friendly interface:** COMSOL has a user-friendly interface that makes it easy to set up and run simulations.
3. **High-performance computing:** COMSOL can take advantage of high-performance computing resources to speed up simulations and handle large datasets.
4. **Comprehensive documentation and support:** COMSOL offers comprehensive documentation and support resources to help users design and simulate temperature sensors.

The structure of the temperature sensor using COMSOL Multiphysics can be achieved by using the following steps

1. **Defining the geometry:** The first step is to define the geometry of the sensor. This can be done using the COMSOL CAD tools, or by importing a CAD model of the sensor.
2. **Setting up the model:** The next step is to set up the model by defining the materials and properties of the sensor. This includes specifying the thermal conductivity, specific heat capacity, and density of the materials used in the sensor.
3. **Defining the boundary conditions:** The boundary conditions of the model must be defined, including the heat flux and temperature distribution on the surface of the sensor.
4. **Solving the model:** Once the model is set up, it can be solved using COMSOL's built-in solvers. This will provide a temperature distribution throughout the sensor.
5. **Analysing the results:** The final step is to analyse the results of the simulation. This includes analysing the temperature distribution, heat transfer rate, and any other relevant parameters.
6. Table 1 shows the recent temperature sensors with the sensing material, substrate material, fabrication techniques and encapsulation methods.

5 Fabrication

5.1 Fabrication

This session describes the fabrication of a temperature sensor that can be used for WBAN application. The sensor fabrication is performed in two steps

1. Fabrication of the electrodes
2. Fabrication of the sensing material

The electrode is fabricated on a cladded FR4 substrate using a photolithography process. The process includes:

1. **Cleaning and preparation:** The copper-clad FR4 substrate is cleaned and prepared for the photoresist coating.

Table 1 Materials and methods used for temperature sensors [28]

S. No	Design Type	Sensing Materials	Substrate Materials	Fabrication Technique	Encapsulation
1	RTD	Graphene/PE DOT:PSS	Polyurethane	Inkjet Printing	Electronic Grade Coating
2	RTD	Au/Cr	Polyurethane	Transfer Printing	polyurethane
3	Thermistor	Carbon/PEDO T-PSS	PEN	Inkjet Printing	NA
4	RTD	Ag	Polyimide	Inkjet Printing	NA
5	Thermistor	PEDOT:PSS/ CNTs	PET	Inkjet Printing	Polymer resist
6	Thermistor	Liquid Metal	Filaflex, Recreus	3D Printing	PDMS/Epoxy
7	Thermistor	PEDOT:PSS/ CNTs	PET	Inkjet Printing	NA
8	RTD	Ag	Cellulose	Screen/E-Jet Printing	NA
9	RTD	PEDOT:PSS	PET	Inkjet Printing	Surlyn
10	Thermocoupl e	PEDOT:PSS, AgNPs, Graphene	Textile	Stencil Printing	NA
11	Thermistor	PEDOT:PSS	Textile, Kapton	Printing	NA
12	Thermocoupl e	Cu/CuNi	Kapton	Aerosol Jet	NA
13	Thermistor	GO/PEDOT:P SS	PVC	Shadow Masking	NA
14	Thermistor	rGO-PHB	PDMS	Inkjet printing, Drop casting	PDMS
15	Thermistor	Ag, PEDOT:PSS	PEN	Inkjet Printing	CYTOP

2. **Photoresist coating:** A thin layer of photoresist is applied to the copper-clad FR4 substrate using a spin coater.
3. **Mask alignment and exposure:** The photomask, which contains the desired circuit pattern, is aligned over the photoresist-coated substrate using a UV light aligner. The substrate and the photomask are then exposed to UV light, which causes the photoresist to harden in the areas that are exposed to light.

4. **Developing:** The substrate is then washed with a developer solution that dissolves the unexposed photoresist, leaving only the hardened areas.
5. **Etching:** The exposed areas of the copper are etched away using a chemical solution. This creates the desired circuit pattern on the copper-clad FR4 substrate.
6. **Photoresist removal:** Finally, the remaining photoresist is removed using a solvent, leaving the patterned copper-clad FR4 substrate ready for further processing.

The fabrication of the sensing material which is chosen to be reduced graphene oxide on top of the electrode include:

1. **Preparation of the reduced graphene oxide ink:**

   (a) Synthesize graphene oxide: Graphene oxide (GO) is synthesized from graphite through a modified Hummers method.
   (b) Reduce graphene oxide: Graphene oxide is then reduced to rGO using a reducing agent, such as hydrazine, sodium borohydride, or hydrogen.
   (c) Prepare the ink: The rGO can then be dispersed in a suitable solvent, such as water, ethanol, or isopropyl alcohol, to make the ink. You can add a small amount of surfactant, such as Triton X-100, to improve the stability of the ink.
   (d) Characterize the ink: You should characterize the ink to ensure that it has the desired properties, such as viscosity, surface tension, and conductivity. You can use techniques such as rheology, contact angle measurement, and four-point probe measurement to characterize the ink.

2. **Preparation of the substrate:** The printed substrate was cleaned using de ionized water of resistivity 15MΩm for 5 minutes. The cleaned substrate was dried in sunlight for 10 minutes
3. **Printing the structure:**

   (a) Print the temperature sensor: The rGO ink can be printed onto a suitable substrate such as paper, glass, or plastic using various printing techniques such as inkjet, screen, or flexographic printing. The printed rGO pattern can be designed as a resistor, thermistor, or another temperature-sensing element.
   (b) Sinter the rGO: To improve the electrical conductivity of the printed rGO pattern, it must be annealed sensor at a high temperature, exposing it to UV light, or treating it with a suitable chemical agent.
   (c) Characterize the temperature sensor: The printed and sintered rGO temperature sensor can be characterized by measuring its electrical resistance at different temperatures using a multimeter or other measurement instrument. The sensor‘s response time, sensitivity, and accuracy can also be evaluated to ensure it meets the desired performance specifications.

Reduced graphene oxide (rGO) has excellent temperature-sensing properties due to its high thermal conductivity and sensitivity to temperature changes. When exposed to different temperatures, the electrical conductivity of rGO changes, making it a good candidate for use in temperature sensors. The temperature sensing property of rGO is based on the fact that as the temperature increases, the thermal energy causes the rGO sheets to vibrate more rapidly, which increases the resistance of the rGO material. Conversely, when the temperature decreases, the rGO sheets vibrate less, which decreases the resistance of the rGO material. This change in resistance can be measured and correlated with changes in temperature, allowing rGO to be used as a sensitive temperature sensor. The high thermal conductivity of rGO also helps to improve its temperature sensing accuracy and response time.

6 Generic Metaheuristic Algorithm for Optimization

The sensor thus created was integrated with an IoT device which is used to display the measured temperature value. For the sensor to measure and provide a meaningful temperature value the output obtained from the sensor is mapped to specific values and is optimized using the generic metaheuristic algorithm [34, 36, 37]. The steps involved include:

1. **Initialization:** The initialization step in a generic metaheuristic algorithm for optimizing a temperature sensor involves defining the initial state of the algorithm and the search space. Here are the steps to perform the initialization step:
   (a) Define the search space: The first step in initializing a metaheuristic algorithm is to define the search space, which represents the set of all possible solutions to the problem. In the case of temperature sensor optimization, the search space would consist of all possible configurations of the sensor‘s parameters, such as the placement of the sensor, the type of material used, and the sensitivity of the sensor.
   (b) Generate an initial solution: The next step is to generate an initial solution within the defined search space. This can be done randomly or using a heuristic method. For example, the initial solution could be a randomly selected configuration of the sensor‘s parameters.
   (c) Evaluate the initial solution: Once the initial solution is generated, it needs to be evaluated to determine its quality. In the case of temperature sensor optimization, the quality of the solution would be measured by its ability to accurately measure temperature. This can be done using simulation or experimental testing.
   (d) Set the initial temperature: For metaheuristic algorithms such as simulated annealing, the initial temperature is an important parameter that determines the degree of exploration in the search space. The initial temperature should be set high enough to allow the algorithm to explore the search space, but not so high that it causes the algorithm to become stuck in a local minimum.

(e) Repeat the process: If the initial solution is not satisfactory, the initialization step can be repeated with a new initial solution or search space until a suitable starting point is found.

2. **Evaluation:** The evaluation step involves measuring the quality of a candidate solution generated by the algorithm. Here are the steps to perform the evaluation step:

   (a) Define the evaluation metric: The first step in evaluating a candidate solution is to define an evaluation metric that measures its quality. In the case of temperature sensor optimization, the evaluation metric could be the accuracy or precision of the sensor in measuring temperature.
   (b) Simulate or measure the solution: Once the evaluation metric is defined, the candidate solution needs to be simulated or measured to determine its quality. This could involve running simulations or experiments to measure the accuracy of the sensor in measuring temperature.
   (c) Calculate the fitness value: Based on the results of the simulation or experiment, a fitness value can be calculated for the candidate solution. The fitness value should be based on the evaluation metric defined in step 1 and should reflect the quality of the solution relative to other solutions in the search space.
   (d) Compare the fitness value: Once the fitness value is calculated, it can be compared to the fitness values of other candidate solutions generated by the algorithm. The candidate solution with the highest fitness value is considered the best solution so far.
   (e) Repeat the process: The evaluation step is repeated for each candidate solution generated by the algorithm until the termination criterion is met.

3. **Search:** The search step involves generating new candidate solutions in the search space in an attempt to find the optimal solution. Here are the steps to perform the search step:

   (a) Choose a search operator: The first step in the search step is to choose a search operator that will generate new candidate solutions. There are several search operators that can be used, such as mutation, crossover, or perturbation.
   (b) Apply the search operator: Once the search operator is chosen, it is applied to the current solution to generate a new candidate solution in the search space. This new solution should be feasible, meaning that it should satisfy any constraints imposed by the problem.
   (c) Evaluate the new solution: Once the new solution is generated, it is evaluated using the evaluation metric defined in the evaluation step. The fitness value of the new solution is calculated and compared to the fitness value of the current solution.
   (d) Accept or reject the new solution: Based on the fitness value of the new solution and the current solution, the algorithm decides whether to accept or reject the new solution. If the new solution has a higher fitness value, it is accepted as the new current solution. If the new solution has a lower fitness

value, it may still be accepted with a certain probability determined by the current temperature value in the case of simulated annealing.

(e) Repeat the process: The search step is repeated with the new current solution generated in step 4. The search continues until a stopping criterion is met, such as reaching a maximum number of iterations, or when the improvement in the fitness value is below a certain threshold.

4. **Selection:** The selection step involves selecting candidate solutions for the next iteration of the search process. Here are the steps to perform the selection step:

(a) Sort the candidate solutions: Sort the candidate solutions generated in the search step based on their fitness values with respect to the accuracy or precision of the temperature sensor.

(b) Select the best solutions: Select the best solutions from the sorted candidate solutions for the next iteration of the search process. The number of solutions selected depends on the specific metaheuristic algorithm used.

(c) Perform elitism: In some metaheuristic algorithms, such as genetic algorithms, the best solution found so far is preserved and carried forward to the next iteration. This is important for a temperature sensor optimization problem to ensure that the best sensor calibration settings are not lost.

(d) Generate new solutions: Generate new candidate solutions from the selected best solutions using search operators such as mutation or perturbation. These new solutions should be feasible and satisfy any constraints imposed by the problem.

(e) Repeat the process: Repeat the selection step for each iteration of the search process until a stopping criterion is met, such as reaching a maximum number of iterations or when the improvement in the fitness value is below a certain threshold.

5. **Termination:** The termination step involves deciding when to stop the search process. Here are the steps to perform the termination step:

(a) Define a stopping criterion: The first step in the termination step is to define a stopping criterion, which is a condition that indicates when the search process should stop. The stopping criterion could be a maximum number of iterations, a specific target accuracy or precision, or when the improvement in the fitness value is below a certain threshold.

(b) Check if stopping criterion is met: At the end of each iteration, the stopping criterion is checked to see if it has been met. If the stopping criterion has been met, the search process stops and the best solution found so far is returned as the optimal solution.

(c) Repeat until stopping criterion is met: The search process is repeated until the stopping criterion is met. If the stopping criterion is not met, the algorithm proceeds to the next iteration of the search process.

(d) Return the optimal solution: Once the stopping criterion is met, the best solution found so far is returned as the optimal solution. This solution should provide the best calibration settings for the temperature sensor.

6. **Solution:** The solution step involves returning the optimal solution found by the search process. Here are the steps to perform the solution step:
   (a) Check if stopping criterion is met: At the end of each iteration, the stopping criterion is checked to see if it has been met. If the stopping criterion has been met, the search process stops and the best solution found so far is returned as the optimal solution.
   (b) Return the optimal solution: Once the stopping criterion is met, the best solution found so far is returned as the optimal solution. This solution should provide the best calibration settings for the temperature sensor.
   (c) Validate the optimal solution: Before using the optimal solution in a real-world scenario, it is important to validate the solution. This involves testing the temperature sensor with the optimal calibration settings and comparing the sensor readings with the actual temperature values.
   (d) Implement the optimal solution: Once the optimal solution has been validated, it can be implemented in the temperature sensor. This involves programming the sensor with the optimal calibration settings.

7 Conclusion and Future Works

This chapter discusses on temperature sensors with a special consideration for temperature sensors with rGO as the sensing material. The fabrication procedure for the sensor is also mentioned. The resulting sensor acts as a negative temperature coefficient resistor. The voltage drop across the resistor is mapped to the temperature in celcius. The resultant values are optimized using the generic meta heuristic algorithm to obtain a more reliable result. The chapter also discusses the various sensors available in the literature and the materials that are available in the market which will help to build a temperature sensor. Different types of temperature sensors were also discussed.

References

1. Kuzubasoglu, B. A., & Bahadur, S. K. (2020). Flexible temperature sensors: A review. *Sensors and Actuators A: Physical, 315*, 112282. https://doi.org/10.1016/j.sna.2020.112282
2. Chen, Y., Lu, B., Chen, Y., & Feng, X. (2015). Breathable and stretchable temperature sensors inspired by skin. *Scientific Reports, 5*, 11505. https://doi.org/10.1038/srep11505
3. Nuthalapati, S., Kedambaimoole, V., Shirhatti, V., Nella, N., Goddam, V., Nayak, M. M., & Rajanna, K. (2018). Screen printed rGO-pd nanocomposite films on a flexible substrate as a temperature sensor. In *Proceedings of 3rd International Conference for Convergence in Technology (I2CT)* (pp. 1–4). IEEE. https://doi.org/10.1109/I2CT.2018.8529700

4. Trung, T. Q., & Lee, N.-E. (2016). Flexible and stretchable physical sensor integrated platforms for wearable human-activity monitoring and personal healthcare. *Advanced Materials, 28*(22), 4338–4372. https://doi.org/10.1002/adma.201504244
5. Dankoco, M. D., Tesfay, G. Y., Benevent, E., & Bendahan, M. (2016). Temperature sensor realized by inkjet printing process on a flexible substrate. *Materials Science and Engineering B, 205*, 1–5. https://doi.org/10.1016/j.mseb.2015.11.003
6. Li, Q., Zhang, L., Tao, X., & Ding, X. (2017). Review of flexible temperature sensing networks for wearable physiological monitoring. *Advanced Healthcare Materials, 6*(12), 1601371. https://doi.org/10.1002/adhm.201601371
7. Hsiao, F.-R., & Liao, Y.-C. (2018). Printed micro-sensors for simultaneous temperature and humidity detection. *IEEE Sensors Journal, 18*(16), 6788–6793. https://doi.org/10.1109/JSEN.2018.2850372
8. Wang, Y.-F., Sekine, T., Takeda, Y., Yokosawa, K., Matsui, H., Kumaki, D., Shiba, T., Nishikawa, T., & Tokito, S. (2020). Fully printed PEDOT: PSS-based temperature sensor with high humidity stability for wireless healthcare monitoring. *Scientific Reports, 10*(1), 1–8. https://doi.org/10.1038/s41598-020-59432-2
9. Kuzubasoglu, B. A., Sayar, E., CochranE, C., Koncar, V., & Bahadir, S. K. (2021). Wearable temperature sensor for human body temperature detection. *Journal of Materials Science: Materials in Electronics, 32*(4), 4784–4797. https://doi.org/10.1007/s10854-020-05217-2
10. Baqi, M. A., & Ahmed, T. M. (2019). Temperature error analysis of K-type thermocouples in industrial environments. *IEEE Sensors Journal, 19*(12), 4825–4832.
11. Gao, M., LI, L., & Song, Y. (2017). Inkjet printing wearable electronic devices. *Journal of Materials Chemistry C, 5*(12), 2971–2993. https://doi.org/10.1039/C7TC00038C
12. Soni, M., Bhattacharjee, M., Manjakkal, L., & Dahiya, R. (2019). Printed temperature sensor based on graphene oxide/PEDOT: PSS. In *Proceedings IEEE international conference flexible printable sensors system (FLEPS)* (pp. 1–3). IEEE. https://doi.org/10.1109/fleps.2019.8792268
13. Fernandes, D. F., Majidi, C., & Tavakoli, M. (2019). Digitally printed stretchable electronics: A review. *Journal of Materials Chemistry C, 7*(45), 14035–14068. https://doi.org/10.1039/C9TC04246F
14. Igorevich, F. I. (2020). Thermocouple condition monitoring using thermocouple resistance. experimental study. In *2020 ural symposium on biomedical engineering, radio electronics and information technology (USBEREIT)* (pp. 0349–0352). IEEE. https://doi.org/10.1109/USBEREIT48449.2020.9117727
15. Maseko, M. L., Agee, J. T., & Davidson, I. (2022). Thermocouple signal conditioning using augmented device tables and table look-up neural networks, with validation in J-thermocouples. In *2022 30th southern African universities power engineering conference (SAUPEC)* (pp. 1–4). IEEE. https://doi.org/10.1109/SAUPEC55179.2022.9730718
16. Mofei, W., Zhigang, C., Yuzhao, H., & Deming, L. (2017). Design of intelligent temperature controller for thermocouple automatic calibrating furnace. In *2017 International conference on industrial informatics – Computing technology, intelligent technology, industrial information integration (ICIICII)* (pp. 147–150). IEEE. https://doi.org/10.1109/ICIICII.2017.63
17. Wobschall, D., & Poh, W. S. (2004). A smart RTD temperature sensor with a prototype IEEE 1451.2 internet interface. In *ISA/IEEE Sensors for Industry Conference, 2004. Proceedings the* (pp. 183–186). IEEE. https://doi.org/10.1109/SFICON.2004.1287157
18. Mazzini, G., Capineri, L., Zanobini, A., & Marchesi, R. (2021). Metrological characterization of a new textile sensor for temperature measurements and a comparison with a Pt100 sensor. In *2021 IEEE international workshop on metrology for industry 4.0 & IoT(MetroInd4.0&IoT)* (pp. 469–472). IEEE. https://doi.org/10.1109/MetroInd4.0IoT51437.2021.9488523
19. Khalaf, A. M., Issa, H. H., Ramírez, J. L., & Mohamed, S. A. (2022). All inkjet-printed temperature sensors based on PEDOT:PSS. *IEEE Access, 10*, 61094–61100. https://doi.org/10.1109/ACCESS.2022.3176822

20. Sui, Y., Kreider, L. P., Bogie, K. M., & Zorman, C. A. (2019, Article no. 2500704). Fabrication of a silver-based thermistor on flexible, temperature-sensitive substrates using a low-temperature inkjet printing technique. *IEEE Sensors Letters, 3*(2), 1–4. https://doi.org/10.1109/LSENS.2019.2893741
21. Wenyi, Z., Uchida, H., Katsube, T., Nakatsuboi, T., & Nishioka, Y. (1997). A novel semiconductor NO gas sensor operating at room temperature. In *Proceedings of International Solid-State Sensors and Actuators Conference* (Transducers '97) (Vol. 1, pp. 569–572). IEEE. https://doi.org/10.1109/SENSOR.1997.613714
22. Pandian, A. P., Palanisamy, R., Narayanan, M., Senjyu, T., & (Eds.). (2022). *Proceedings of third international conference on intelligent computing, information and control systems* (Vol. 1415). Springer Nature Singapore. https://doi.org/10.1007/978-981-16-7330-6
23. Wang, C., & Liu, M. (2012). Fiber-optic temperature sensors based on fiber Bragg gratings: A review. *Sensors, 12*(8), 12291–12320.
24. Ali, S., Khan, S., & Bermak, A. (2019). Inkjet-printed human body temperature sensor for wearable electronics. *IEEE Access, 7*, 163981–163987. https://doi.org/10.1109/ACCESS.2019.2949335
25. Sahatiya, P., Puttapati, S. K., Srikanth, V. V. S. S., & Badhulika, S. (2016). Graphene-based wearable temperature sensor and infrared photodetector on a flexible polyimide substrate. *Flexible and Printed Electronics, 1*(2), 025006. https://doi.org/10.1088/2058-8585/1/2/025006
26. Malik, J., Andersson, H., Forsberg, V., Engholm, M., Zhang, R., & Olin, H. (2018). PEDOT: PSS temperature sensor ink-jet printed on a paper substrate. *Journal of Instrumentation, 13*(12), C12010. https://doi.org/10.1088/1748-0221/13/12/C12010
27. Adebayo, A. O., & Adebayo, A. B. (2016). Modeling of thermistor-based temperature sensor and its application to temperature measurement in a refrigeration system. In *IEEE international conference on electro/information technology* (pp. 0234–0239).
28. Ozioko, O., Kumaresan, Y., & Dahiya, R. (2020). Carbon nanotube/PEDOT: PSS composite-based flexible temperature sensor with enhanced response and recovery time. In *Proceedings IEEE international conference flexible printable sensors system (FLEPS)* (pp. 1–4). IEEE. https://doi.org/10.1109/fleps49123.2020.9239431
29. Khan, S., Ali, S., Khan, A., & Bermak, A. (2023). Wearable printed temperature sensors: Short review on latest advances for biomedical applications. *IEEE Reviews in Biomedical Engineering, 16*, 152–170. https://doi.org/10.1109/RBME.2021.3121480
30. Arefin, M. T., Ali, M. H., & Haque, A. K. M. F. (2017). Wireless body area network: An overview and various applications. *Journal of Computer and Communications, 05*(07), 53–64. https://doi.org/10.4236/jcc.2017.57006
31. Meharouech, A., Elias, J., & Mehaoua, A. (2019). Moving towards body-to-body sensor networks for ubiquitous applications: A survey. *Journal of Sensor and Actuator Networks, 8*(2), 27. https://doi.org/10.3390/jsan8020027
32. Anter, A. M., Mohamed, A. W., Zhang, M., & Zhang, Z. (2023). A robust intelligence regression model for monitoring Parkinson's disease based on speech signals. *Future Generation Computer Systems, 147*, 316–327.
33. Murugan, K., Reddy, S. L., Prasad, B. S., & RamPrasad, P. (2022). Integration of pH and temperature sensor for biomedical applications. In *2022 7th international conference on communication and electronics systems (ICCES)* (pp. 354–361). https://doi.org/10.1109/ICCES54183.2022.9835719
34. Anter, A. M., Moemen, Y. S., Darwish, A., & Hassanien, A. E. (2020). Multi-target QSAR modelling of chemo-genomic data analysis based on extreme learning machine. *Knowledge-Based Systems, 188*, 104977.
35. Anter, A. M., Abd Elaziz, M., & Zhang, Z. (2022). Real-time epileptic seizure recognition using Bayesian genetic whale optimizer and adaptive machine learning. *Future Generation Computer Systems, 127*, 426–434.

36. Thakare, A., Anter, A. M., & Abraham, A. (2023). Seizure disorders recognition model from EEG signals using new probabilistic particle swarm optimizer and sequential differential evolution. *Multidimensional Systems and Signal Processing, 34*, 1–25.
37. Anter, A. M., & Ali, M. (2020). Feature selection strategy based on hybrid crow search optimization algorithm integrated with chaos theory and fuzzy c-means algorithm for medical diagnosis problems. *Soft Computing, 24*(3), 1565–1584.
38. Hassan, A., Anter, A., & Kayed, M. (2021). A robust clustering approach for extending the lifetime of wireless sensor networks in an optimized manner with a novel fitness function. *Sustainable Computing: Informatics and Systems, 30*, 100482.

Recent Advanced in Healthcare Data Privacy Techniques

Waleed M. Ead, Hayam Mohamed, Mona Nasr, and Ahmed M. Anter

1 Introduction

Many various strategies were offered several years ago to retain and preserve privacy. There are several approaches for privacy-preserving data mining (PPDM) and publication (PPDP). We shall discuss such strategies and algorithms shortly. So, here's a summary of a few research on data mining and publishing with privacy protection. With developments in data mining techniques, the volume and scope of privacy breaches is expanding from the unique identification of an individual or SA disclosure to behavioral advertising privacy breaches in PPDP. Since Sweeney's proposal of k -anonymity to minimize privacy leakage and exposure during data publishing, several anonymization approaches [1, 2] and privacy models [2, 3] have been created in [1, 4, 5]. The author proposed k-anonymity [2] to assure that each raw data record cannot be distinguished from the other k-1 entries on QIDs. They also discovered a vulnerability in k-anonymity, namely that homogeneity and background knowledge attacks can be employed on k-anonymized datasets. They introduced the l-diversity model with the SA diversity constraint to better privacy protection. In [5] they introduced a t-closeness model with distribution constraints

W. M. Ead · A. M. Anter (✉)
Egypt-Japan University of Science and Technology (E-JUST), Alexandria, Egypt

Faculty of Computers and AI, Beni-Suef University, Beni Suef, Egypt
e-mail: Ahmed_Anter@fcis.bsu.edu.eg

H. Mohamed
Faculty of Computers and AI, Beni-Suef University, Beni Suef, Egypt

M. Nasr
Faculty of Computers and Artificial Intelligence, Helwan University, Helwan, Egypt

A. M. Anter et al. (eds.), *Nature-Inspired Methods for Smart Healthcare Systems and Medical Data*, https://doi.org/10.1007/978-3-031-45952-8_11

to ensure privacy after identifying skewness and similarity assaults on l-diversity. T-closeness cannot sufficiently protect the privacy of atypical values since they are more vulnerable to privacy exposure. They created the Mondrian top-down greedy approximation multidimensional k-anonymization technique using local recoding in [6]. Two clustering-based algorithms were created in [7], and they outperformed Mondrian in terms of information loss. In [8], the connection between QID and SA features is broken down. The authors presented Anatomy, a technique for anonymizing microdata. Anatomy, on the other hand, provides specific QID numbers. As a result of the precise QID values it broadcasts, presence assaults are conceivable. Susan et al. [9] presented an anatomization with slicing method: a novel technique to privacy preservation for numerous sensitive attributes. The sensitive attributes are divided into the sensitive table and the quasi attributes table using the anatomization approach. This project promotes and employs the k-anonymity and l-diversity principles. It protects published data from membership, identity, and attribute linkage attacks, but the slicing technique must be applied individually to QIT and ST, and Anatomy is vulnerable to presence attacks because it releases specific QID values. Onashoga et al. [10] presented a KC-slice-based approach. By combining the properties of the LKC privacy model, it is a dynamic privacy-preserving data publishing strategy for many sensitive attributes. However, the writers did not provide instances of several releases, therefore they only described the technique in its entirety for a single release. They developed the LDP-Miner local differential privacy algorithm in [11] to anonymize set-valued data. These approaches are only appropriate for posting results; they cannot be used to publish secure data collections. They presented an effective method for posting microdata for numerous sensitive attributes in [12]. It employs the (p, k)-Angelization method to anonymize numerous sensitive attributes. It protects the privacy of published data by removing the threat of background knowledge and membership attacks. It is restricted to the circumstance where the dataset contains a single record relating to an individual. They proposed the G-model in [13], which guards against gender-specific SA attacks by keeping separate groups and caches of male and female SAs. The G-model, too, avoids generalization. It is unable to protect against Semantic Correlation attacks on privacy. A bidirectional personalized generalization approach was introduced in [14] to solve the higher privacy and reduced utility loss for multi-record datasets. This model can withstand a bi-directional chain attack by using a hierarchical generalization technique. Despite its high performance, the model could only manage one sensitive attribute, and its generalization results in information loss. They devised the QIAB-IMSB technique in [15] to anonymize the set-valued dataset. Vertical partitioning was employed to split the table. The system defends against a sensitive linking attack by utilizing hierarchical generalization, and classification models were utilized to compare the data's correctness.

1.1 Query Answering

Authors in [16] proposed a utility-based strategy for ensuring privacy. The attributes of sensitive values represented in queries are taken into account in their approach. They allowed data owners to designate attribute weights and used generalisation boundaries to anonymize queries only when the sensitive attribute's values surpassed the cut-off. A value is ascribed to a property based on how helpful it is. Generalisation borders are used to anonymize attributes whose combined weights above the threshold values, while the remaining inquiries can be published immediately without any alterations. They assessed the utility of the data by comparing the distributional changes between the original and anonymized data using Cluster Analysis Measure and Empirical CDF Measures. As a result of their efforts, the utility has expanded and the risk of data loss has decreased.

In [17] due to the high accuracy of the results obtained by the proposed SPLU framework, releasing data with a high data utility is achievable for large sum aggregate searches. SPLU provides significant imprecision for tiny sum aggregate searches to protect privacy. The framework randomises the perturbation of the sensitive values. Based on the amount of data involved in each, this strategy differentiates the reconstruction for utility issues from the reconstruction for privacy concerns. They use relative error to demonstrate the privacy and utility levels, where a higher relative error correlates to more privacy and less utility.

In [18] mention GUPT to increase the accuracy and accessibility of data analysis for a certain privacy framework. According to the term "privacy budget" used to characterise privacy utilising existing variable privacy frameworks, better privacy is attained with a lower data security budget. However, because it cannot simply convert into the value of the programme, this Privacy Measurement Unit is difficult for data analysts who are not privacy experts to evaluate. Furthermore, analysts can efficiently distribute this privacy budget across multiple queries on a data set to avoid faulty analysis and limit the amount of inquiries that can be carried out on the dataset safely. The GUPT goes beyond the constraints of traditional differential anonymity by encouraging enterprises to evaluate their datasets differently. The new procedure could be implemented. Authors in [19] described a strategy for replying to personal questions in an ongoing public dataset with a large number of differentially personal results. They found and used the connected information contained in continuous datasets, and they invented the concept of coupled sensitivity, which addressed the demand for differential privacy while causing less query noise. The results of their trials proved that their system is dependable, efficient, and protects privacy with minimal benefit loss.

Authors in [20] provided a survey on the privacy protection of published electronic health records. Their study looked at more than 45 algorithms for disseminating structured electronic health recorded data under privacy limitations. The algorithms chosen for analysis are effective in maintaining privacy and usability. In addition to data privacy, three ways are being researched to protect data utility: (1) They aim to minimise data loss by utilising an optimisation technique to calculate it. (2) to outline the data analysis technique to be utilised with publicly available data and to assure its accuracy; and (3) to take into account the data owners' consideration requirements. Data that meets these specifications should be submitted.

1.2 Data Classification

Authors in [21] proposed a literature review on how to increase the usefulness of anonymised data from the standpoint of data classification. They use differentiated privacy in the use of decision-making arrangements to anonymize and disseminate data that demonstrates considerable quality improvements [22]. A detailed investigation was carried out; the proposed method revealed considerable privacy and efficiency gains focused on the linkages between the features of different datasets. To protect their privacy, they identify data distributions using entropy and the utility of information. The proposed research can categorise distinct associations based on the data's usage and context. Furthermore, [23] implement an analysis to release data for categorization purposes after the inclusion of a number of noises. As a result, the current systems' anonymous data are less useful. Identity cannot be sufficiently concealed in [24], and private information cannot be delivered truthfully. Jaiswal et al. [25] proposed a k-anonymity technique to safeguard the publishing of health-care data. Full-domain generalisation is the core of their proposed algorithm. Zaman and Obimbo [26] proposed an h-ceiling concept to limit overgeneralization levels in order to address the diminished information utility caused by full domain generalisation. They also suggested adding bogus entries in the appropriate classes to meet the K-anonymity but not the h-ceiling [27]. The sensitive data is drawn at random from the relevant attribute domain, and the false records share the same quasi-identity characteristics as similar class data. Furthermore, all phoney records are listed in a database of false records. The adversaries are provided an anonymous table, which they index and attempt to extract sensitive information from using fake data. A record that contains a false one should be removed (Table 1).

In [28] presented a shared data classification data sharing approach for privacy awareness. Their approach to privacy mechanisms is focused on assembling privacy utilities that eliminate the most ineffective data accuracy and privacy characteristics. A k-anonymization method was presented [29] to cover tuples with very sensitive values in each equal category of severe weighting and packing. They devised a random sampling technique to establish a balance between privacy and utility. Because criminal data contains a substantial amount of sensitive information, it solely considers that. As a result, information efficiency matrices have not been used. The application of generalisation techniques resulted in the documenting of several procedures. Furthermore, [30] proposes hybrid generalisations that integrate generalisations with the data relocation paradigm. During cell migration, some cells undergo modifications in order to supplement tiny, indivisible clusters of tuples. This methodology seeks to reduce the consequences of over allocation and outliers.

Table 1 anonymization models

Anonymization model	Identity disclosure	Attribute disclosure	Anonymous operation	Multiple sensitive attributes	Multiple records
k-Anonymity	√		Generalization with predefined taxonomy tree	No	No
l-Diversity	√	√	Generalization with predefined taxonomy tree and suppression	No	No
(α, k)-Anonymity	√	√	Generalization with predefined taxonomy tree	No	No
t-Closeness		√	Generalization with predefined taxonomy tree	No	No
m-Invariance	√	√	Generalization with predefined taxonomy tree	No	No
IR (k, l)-anonymity	√		Generalization with predefined taxonomy tree	No	Yes
IR (α,β)-anonymity	√		Generalization with predefined taxonomy tree	No	Yes
EIR (k, l)-anonymity	√	√	Set generalization	No	Yes
EIR (α, β)-anonymity	√	√	Set generalization	No	Yes
(K, C) L-Anonymity	√	√	Suppression	No	No
(p, k)-Anonymity + Angel			Angelization techniques	Yes	No
(p, l)- Angelization			Angelization techniques	Yes	Yes
(k, l)-diversity	√		Generalization techniques	No	Yes

2 Privacy Preserving Techniques

In the following section, we will go over the algorithms and strategies to see which one performs best. We shall contrast them based on data utility, information loss, and execution time.

2.1 Privacy Preserving

To protect the privacy of patients and research participants, legal and organizational measures must be implemented in order to treat data in accordance with these criteria. It is necessary to establish effective solutions to strike a balance between health care data sharing and privacy protection. Although different approaches have been developed to solve these challenges, the majority of them focus on a specific aspect of the problem using a single theory. In this study, we presented a paradigm for exchanging data while maintaining privacy and enhancing utility with a view to a broader practical application.

2.1.1 Why Need for Privacy

Electronic data is increasing as the modern world becomes more digitized. It is critical to examine the socioeconomic patterns of individuals in society. When it comes to data disclosure, privacy is a major concern. As an example, medical data comprises sensitive data because it contains information on the ailments of the patients. Before making this data available for data mining, it is critical to privatize it. In medical contexts, it is critical to protect the mining model with effective privacy; otherwise, incorrect predictions will result. Personal information that could be regarded unethical must not be disclosed. When data mining is performed on aggregate results, privacy is defined as the protection of unauthorized publication of information. Privacy must be addressed at all the levels while mining is carried out.

3 Privacy Preserving Technologies

Privacy-preserving technologies enable users to protect the security of their personally identifiable information (PII) submitted to and processed by service providers or apps while keeping the functionality of data-driven systems. Privacy-preserving technologies enable users to protect the security of their personally identifiable information (PII) submitted to and processed by service providers or apps while keeping the functionality of data-driven systems. In this section, we present a few well-known privacy models used to safeguard the confidentiality of healthcare data. As a typical PPDP technology, we give great attention to data anonymization privacy.

3.1 Data Anonymization

Figure 1 depicts a scenario for data posting utilizing data anonymization. A modified original database is constructed before being issued as an anonymized database by applying generalization and suppression to the original database. The anonymized database could be studied instead of the original database.

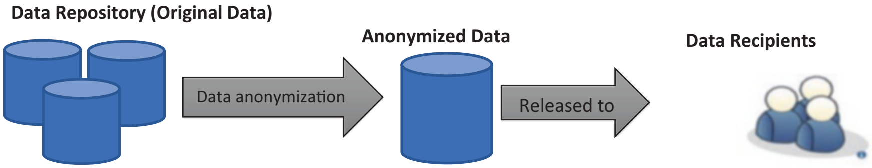


Fig. 1 Data anonymization scenario

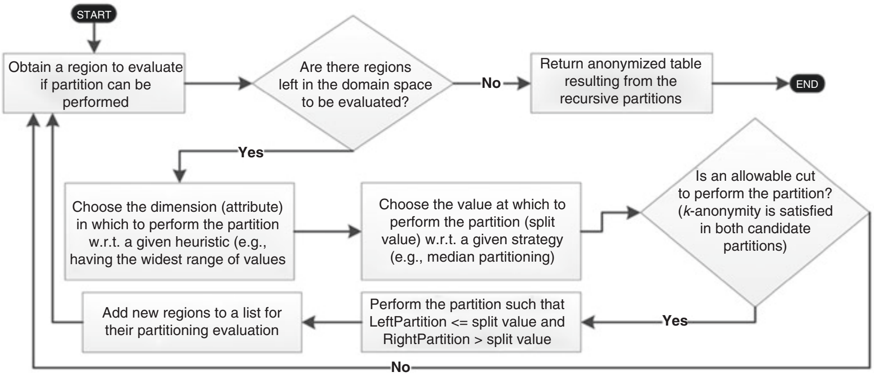


Fig. 2 Core process of the mondrian algorithm

3.1.1 Mondrian Algorithm on IT

The Mondrian method [6] is a popular tool for anonymizing relational data. It divides the data space into a few smaller parts initially using KD-tree. Then, each of these partitions is anonymized using multidimensional generalization.

It can successfully anonymize the QID from the top down. Each characteristic is compelled to follow Mondrian's intuitive arrangement. There are no generalization hierarchies for categorical attributes. As a result of this strategy, less information is lost. However, the semantic outcomes are less acceptable. Mondrian's upper bound problem is that potential data utility declines.

Mondrian anonymization includes partitioning multidimensional data in geometric space to find the most effective partitions for data anonymization. However, one disadvantage of Mondrian anonymization is that it consumes too much memory and processing time when dealing with high-dimensional data. Another concern is that because to the Mondrian-based privacy, the performance of data mining algorithms may become unstable.

The steps of the Mondrian algorithm are:

1. To categorize the raw dataset into k groups, use a kd-tree. Make certain that each K-group contains k records or more.
2. Due to the generalization of each k-group, make sure that each k-group has the same QID (Fig. 2).

3.1.2 M-Generalization Algorithm

For 1: M dataset anonymization, 1: M generalization is used. It is made up of two sub-algorithms: Partition for transactional processing and Mondrian for relational processing. Both are simple based on (k, l). When publishing 1: M data, the (k, l)-diversity technique can secure QID and SA information by enforcing k-anonymity on the SA fingerprint and l-diversity on each equivalence class. This algorithm partitions data twice. SA anonymization should be performed before QID generalization. Use the Mondrian [6] and Partition [40] to generalize QID.

The primary steps of 1:M-generalization are as follows:

1. Convert 1:M microdata to 1:1 microdata as the first step.
2. Analyse the SA fingerprint in step two.
3. Anonymize SA diversity and QID.

Faster R-CNN includes a district application network that detects areas of the image where an object is more likely to be present, as well as a classifier that turns all of those objective applications into a class in the training dataset [12] (Figs. 3, 4, and 5).

3.1.3 Semantic Anonymization Method

The L-diversity methodology and semantic extraction methods are used to build the semantic-based anonymization. The owner of the public data might specify semantic rules on which the semantic extraction is based [8]. Describe the relationship between two delicate values. The term "effective semantic rules" refers to semantic rules that generate data that compromises privacy and leads to similarity attacks. The anonymizer identifies the key data that must be anonymized before selecting the efficient semantic rules that produce these data. The strategy is separated into two main parts: Data extraction using practical semantic ideas determines the rules

Algorithm 1: Partition
Partition (partition, k)
Adding the partition to the global return list if the partition cannot be separated;
else //
The following steps are taken: / choose a node with the maximum information gain split_Node choose node (partition); // distribute records to sub partitions
Sub_ Partitions ← data distribution (partition, split_Node);
// deal with subp_artitions that have fewer than k records balance partitions (sub Partitions);
for sub_ Partition in sub _Partitions do
Partition (sub _Partition, k);
	Algorithm 2: Mondrian for data in 1: M Dataset
	Mondrian (partition, l)
	Add the partition to the global return list if the part's ion cannot be divided;
	else
	/* It is best to choose the property with the widest (normaliZed) range of values. */
	dm ← Select the attribute (partition);
	When dm is a number,
	threshold ← Choose Threshold (dm, partition);
	lHS← {t ∈ partition: T [dm] ≤threshold};
	rHS← {t ∈ partition: T [dm] > threshold}; sub Partition ← {lHS} _ {rHS};
	else
	split_Node ← split (partition, dm);
	sub_ Partitions ← share information (partition, split Node);
	for sub Partition in sub Partitions do
	Mondrian (sub _Partition, l) ;

Fig. 3 Anonymization Algorithms (Partition, Mondrian)

Algorithm 3: 1: M-generalization.
Begin
Enter T, K, and L as input
Output: $T*$
$T1 \leftarrow$ Transform T into 1:1 dataset;
$T2 \leftarrow$ Partition (T, K);

T $*\leftarrow$ Mondrian (IT, l);
return T $*$;
End

Fig. 4 1: M-generalization

Algorithm4: algorithm for semantic extraction
L-diversity Table as an input (T)
Output: effective set of semantic guidelines
Clustering table into ECs;
While T isn't end do
For each equivalence class V from T do
Apply semantic rules on the sensitive attribute of EC;
If there is extracted data impacts on privacy resulted by a semantic rule, then

a) Keep the semantic rule on file as a valid semantic rule.
b) Assign this rule an anonymization action.

End while

Fig. 5 Semantic Anonymization algorithms

of anonymization. Table 2 lists the ECs, the semantic rule that applies to them, and the recommended anonymization step in accordance with the values of sensitive features.

Our Suggested Guidelines
Using the INFORMS dataset:

Rule 1: A respiratory infection disease is present if "the bronchitis, lung cancer, cough, flu, and bronchitis" are present.
Rule 2: Very low salary if "income with values 0 to 2000".

Table 2 privacy models

Privacy models/ Techniques	Evaluation	Privacy disclosure	Data utility	Reference
Slicing	Designed to handle high-dimensional data, slicing can restore the original tuples after random tuple permutation if distinct tuples have the same QIDs and SAs.	Skewed and similarity-based attacks	Missing information	[31]
Anatomization With Slicing	The procedure generates several tables and takes a long time to finish. The answer is really difficult.	Demographic Knowledge attack	Loss of information	[9]
SLOMS	It removed the link between MSA. It issued numerous tables with information loss due to the generalization process.	Knowledge attack regarding demography	Information loss	[32]
Clustering Multi-sensitive Bucketization	It can only analyse numerical data, despite having a low suppression rate and little information loss.	—	Minimal information loss	[33]
MSA (α, l)	Employed suppression and generalizations when discussing anatomy, which reduced its usefulness.	—	Increased information loss	[34]
ANGELMS	Information loss results from the lack of a correlation between QIDs and SAs.	Skewness attacks, Similarity attacks	High information loss.	[35]
LKC- privacy With slicing	The KC slice method was suggested for uploading dynamic MSA data; the approach has only been presented in its full for a single release because the authors did not provide examples of multiple releases.	—	Information loss	[10]
k -anonymity + l –diversity	Proposed privacy models for 1: M datasets. An adversary can reidentify individuals using the datasets if they are aware of the SA fingerprints.	—	Information loss	[36]
(k, l)-diversity	The proposed solution protects the privacy of 1:M datasets. As a result of QI and SA generalisations, there has been tremendous information loss. In the case of MSA, this is not applicable.	—	Information loss	[37]

(continued)

Table 2 (continued)

Privacy models/ Techniques	Evaluation	Privacy disclosure	Data utility	Reference
(p, k)-Anonymity + Angel	The suggested method protects the privacy of MSA's healthcare microdata. Its basis is the (p, k) angelization heuristic method.	Attacks correlated with MSA	Minimal information loss	[12]
(p, l)- Angelization	The 1:M dataset and the MSA dataset are combined in the suggested technique.	Generalization correlation attacks	Low information loss	[32]
G-Model	By keeping separate groups and caches of male and female SAs, the G-model guards against gender-specific SA attacks. The G-model avoids generalisation as well.	Semantic correlation attacks.	Fail to provide optimal privacy protection	[13]
Slicing (Multiple Column Multiple Attributes)	MSA anonymization was suggested, but QI attributes weren't taken into account. In the 1: M occurrence of tuples in microdata, it displays wrong results.	Skewness attacks, Similarity attacks, Sensitivity attacks	Information loss	[38]
(k, k m)-anonymous	The proposed approach is quite unrealistic even though it lessens information loss.	—	Low information loss	[36]
Decomposition Plus	Utility was lost when noise was added to the proposed method. It also failed to prevent attribute and identity leaking.	Skewness attack Similarity attack	High information loss.	[39]

Using the Heat Disease Dataset

Rule 1: Heart_disease_exists if "cp is typical angina or atypical angina."

Rule 2: If the "trestbps from 150 to 200" range, then the resting blood pressure is high.

Rule 3: Blood_sugar_sickness if "fbs > 120."

Rule 4: States_that_heart_disease_exists if "thalach from 170 to 202."

Rule 5: Heart_disease_is_present if "restecg ST>.5".

The red cells in Table 3 relate to the above-mentioned effective semantic rules that generate data that has an influence on privacy and can lead to privacy exposure. To solve this issue, we anonymized the data, which is shown in Table 4. Table 3 displays the ECs, the applicable semantic rule, and creates anonymization action based on the semantic similarity of sensitive attribute values. It is clear that the EC with QI [30, 40], F and the EC with QI [40, 50], M have been blended into a single EC with QI [30, 50], * and all of the sensitive values from the two EC. As a result, prevent an attacker from revealing SA's privacy (Tables 5 and 6).

Table 3 original microdata (the heart disease dataset before anonymization)

Age	Sex	cp	trestbps	fbs	thalach	Restecg
28 Year	S_M	Typical_angina	145.00	108.00	150.00	0
29 Year	S_M	Asymptomatic	130.00	101.00	162.00	1
21 Year	S_M	Atypical_angina	120.00	110.00	148.00	0
41 Year	S_M	Non-anginal_pain	140.00	105.00	153.00	0
50 Year	S_M	Atypical_angina	138.00	110.00	151.00	1
48 Year	S_M	Asymptomatic	120.00	110.00	114.00	1
36 Year	S_F	Typical_angina	160.00	125.00	187.00	1
37 Year	S_F	Atypical_angina	150.00	130.00	172.00	1
30 Year	S_F	Atypical_angina	172.00	130.00	178.00	1

Table 4 Entropy_3_Diversity (the heart disease dataset after 3_Diversity)

Age	Sex	cp	trestbps	fbs	thalach	Restecg
[20–30]	S_M	Typical_angina	145.00	108.00	150.00	0
[20–30]	S_M	Asymptomatic	130.00	101.00	162.00	1
[20–30]	S_M	Atypical_angina	120.00	110.00	148.00	0
[30–40]	S_F	Atypical_angina	172.00	130.00	178.00	1
[30–40]	S_F	Typical_angina	160.00	125.00	187.00	1
[30–40]	S_F	Atypical_angina	150.00	130.00	172.00	1
[40–50]	S_M	Non-anginal_pain	140.00	105.00	153.00	0
[40–50]	S_M	Asymptomatic	120.00	110.00	114.00	1
[40–50]	S_M	Atypical_angina	138.00	110.00	151.00	1

Table 5 the effective semantic rules and anonymization action

Age	Sex	cp	trestbps	fbs	thalach	Restecg	Rules	action
[20–30]	S_M	Typical_angina	145.00	108.00	150.00	0	—	—
[20–30]	S_M	Asymptomatic	130.00	101.00	162.00	1	—	—
[20–30]	S_M	Atypical_angina	120.00	110.00	148.00	0	—	—
[30–40]	S_F	Atypical_angina	172.00	130.00	178.00	1	Heart disease_ blood sugar disease_ high resting blood pressure	Generalization, Adding +10 Suppress Sex attributes
[30–40]	S_F	Typical_angina	160.00	125.00	187.00	1		
[30–40]	S_F	Atypical_angina	150.00	130.00	172.00	1		
[40–50]	M	Non-anginal_pain	140.00	105.00	153.00	0	—	—
[40–50]	M	Asymptomatic	120.00	110.00	114.00	1	—	—
[40–50]	M	Atypical_angina	138.00	110.00	151.00	1	—	—

Table 6 3_Diversity after anonymization

Age	Sex	cp	trestbps	fbs	thalach	Restecg
[20–30]	S_M	Typical_angina	145.00	108.00	150.00	0
[20–30]	S_M	Asymptomatic	130.00	101.00	162.00	1
[20–30]	S_M	Atypical_angina	120.00	110.00	148.00	0
[30–50]	$	Atypical_angina	172.00	130.00	178.00	1
[30–50]	$	Typical_angina	160.00	125.00	187.00	1
[30–50]	$	Atypical_angina	150.00	130.00	172.00	1
[30–50]	$	Non-anginal_pain	140.00	105.00	153.00	0
[30–50]	$	Asymptomatic	120.00	110.00	114.00	1
[30–50]	$	Atypical angina	138.00	110.00	151.00	1

3.1.4 Rating Privacy Preservation Method

In Fig. 4, a Rating system is used to distribute sensitive info. The AT (Attribute Table) and IDT (ID Table) tables for grading releases are based on varying sensitivity coefficients for distinct qualities. This strategy preserves the confidentiality of a large number of microdata correlations while simultaneously preserving the confidentiality of a large number of sensitive variables. The utility of publicly available data while protecting many sensitive elements from prying eyes [41]. The (SCi (sensitivity coefficient)) is calculated using this method (Fig. 6) (Tables 7, 8, and 9).

3.1.5 (p, k) Angelization with Weight Calculation for MSAS Data Publishing

Because each bucket contains entries from the p categories, each bucket partitioning follows the (p, k)-anonymity principle. This approach (Angelization) includes bucket and batch partitioning pairings. Each bucket contains at least k tuples, where k is the smallest group size needed to guard against linking attacks.

Because the degree of sensitivity or weight attributed to each sensitive feature changes, this method applies weighted SA (sensitive attributes) assessments. Each sensitive attribute's weights are calculated to establish its level of sensitivity. Let W = w1, w2,… w1 and w2 are the weights of s1 and s2, respectively. Table 8 shows the weights allocated to each sensitive attribute in the se., and so on. As a result, the two published tables (a sensitive batch table and a generalized table, as shown in Tables 10 and 11) are prepared.

A single table has numerous, closely related characteristics. Tables 11 and 12, for example, show "cp, thalach" and "restecg, trestbps, fbs" respectively. Horizontal partitioning is used to ensure that each bucket has l-diversity (Fig. 7) (Tables 13, 14, 15, and 16).

Algorithm: *AiIDj*-Creation

input: T with SC1, SC2, SCd+1

Results: AiIDj

1: j=0 *AiIDj=Ø*

2: For each property Ai, multiply by $(1 \leq i \leq d + 1)$

3: based on their values, hash the Aitk in Ai (each bin according to value)

/* The group-creation process is in lines 4–9. */

4: so long as at least SCi non-empty hash buckets exist;

/* lines 5 to 9 create a new AiIDj. */

5: j=j+1; *AiIDj = Ø*

6: S is the collection of SCi buckets with the largest size at the moment.

7: for each bucket in S

8: remove a random tuple Aitk from the collection

9: *AiIDj = AiIDj U {Aitk}*

/* The step of residue assignment is lines 10 to 13. */

10: for each non-empty bucket

/* There is just one value for this bucket; see Lemma 3. */

11: *Ait'k* = the bucket's sole remaining asset

12: *S'* = the collection of AiIDj without the Ait'k

13: pick a random AiIDj in S' to receive Ait'k.

Fig. 6 Rating method

Table 7 microdata used for rating method

Age	Sex	cp	trestbps	fbs	thalach	Restecg
28Y	S_M	Typical_angina	145.00	108.00	150.00	0
29 Y	S_M	Asymptomatic	130.00	101.00	162.00	1
21 Y	S_M	Atypical_angina	120.00	110.00	148.00	0
30 Y	S_M	Atypical_angina	172.00	130.00	178.00	1
36 Y	S_M	Non-anginal_pain	160.00	125.00	187.00	0
37 Y	S_F	Atypical_angina	150.00	130.00	172.00	1
41 Y	S_F	Non-anginal_pain	140.00	105.00	153.00	0
48 Y	S_F	Asymptomatic	120.00	110.00	114.00	1
56 Y	S_F	Atypical_angina	138.00	110.00	151.00	1
65 Y	S_F	Asymptomatic	150.00	105.00	148.00	0

Table 8 Rating publishes AT with SC (3; 2; 2; 2; 2; 2; 2) for microdata in Table 5

t	Age (A_1)	Sex (A_2)	cp (A_3)	Trestbps (A_4)	Fbs (A_5)	Thalach (A_6)	Restecg (A_7)
t_1	A_1ID_1	A_2ID_1	A_3ID_2	A_4ID_1	A_5ID_1	A_6ID_1	A_7ID_1
t_2	A_1ID_1	A_2ID_2	A_3ID_2	A_4ID_2	A_5ID_2	A_6ID_2	A_7ID_2
t_3	A_1ID_1	A_2ID_3	A_3ID_1	A_4ID_3	A_5ID_3	A_6ID_3	A_7ID_3
t_4	A_1ID_2	A_2ID_4	A_3ID_3	A_4ID_4	A_5ID_4	A_6ID_4	A_7ID_4
t_5	A_1ID_2	A_2ID_5	A_3ID_5	A_4ID_5	A_5ID_5	A_6ID_5	A_7ID_5
t_6	A_1ID_2	A_2ID_6	A_3ID_4	A_4ID_1	A_5ID_1	A_6ID_1	A_7ID_1
t_7	A_1ID_3	A_2ID_2	A_3ID_4	A_4ID_2	A_5ID_2	A_6ID_2	A_7ID_2
t_8	A_1ID_3	A_2ID_3	A_3ID_3	A_4ID_3	A_5ID_4	A_6ID_3	A_7ID_3
t_9	A_1ID_3	A_2ID_4	A_3ID_5	A_4ID_5	A_5ID_5	A_6ID_4	A_7ID_5
t_10	A_1ID_3	A_2ID_5	A_3ID_1	A_4ID_4	A_5ID_3	A_6ID_5	A_7ID_4

Table 9 For microdata in Table 5, rating publishes IDT with SC (3, 2; 2; 2; 2)

ID	Age (A1)	Sex (A_2)	cp (A_3)	Trestbps (A_4)	Fbs (A_5)	Thalach (A_6)	Restecg (A_7)
ID_1	21–28–29	M-F	Aty-asy	108–130	108–130	150–172	0–1
ID_2	30–36–37	M-F	Ty-asy	130–140	101–105	162–153	0–1
ID_3	41–48–56–65	M-F	Aty-asy	120–120	110–105	148–114	0–1
ID_4	–	M-F	Aty-non	172–150	130–110	178–151	0–1
ID_5	–	M-F	Non-aty	160–138	125–110	187–148	0–1

aty atypical angina, *Ty* typical angina, *asy* asymptomatic, *non* non-anginal pain

Table 10 Calculation of weight

Sensitive-attributes	Identified by	Dependency	Weightage
Chest-pain-type (S1 = cp)	S2–S3–S4	3	3
Resting-blood-pressure (S2 = trestbps)	…	0	0
Resting-electrocardiographic-results (S3 = restecg)	S2–S5	2	2
Maximum-heart-rate-achieved (S4 = thalach)	S1–S2–S5	3	3
Fasting-blood-sugar (S5 = fbs)	S2	1	1

Table 11 QUASI-TABLE (QIT) (publicly release all quasi- identifiers for each person)

Tuple id	Age	Sex
P1	28	M
P2	29	M
P3	21	M
P4	30	F
P5	36	F
P6	37	F
P7	41	M
P8	48	M
P9	46	M

Table 12 Sensitive-attribute-table (ST) (publish all sensitive-attributes for each individual)

cp	trestbps	fbs	thalach	Restecg
Typical_angina	145.00	108.00	150.00	0
Asymptomatic	130.00	101.00	162.00	1
Atypical_angina	120.00	110.00	148.00	0
Atypical_angina	172.00	130.00	178.00	1
Typical_angina	160.00	125.00	187.00	0
Atypical_angina	150.00	130.00	172.00	1
Non-anginal_pain	140.00	105.00	153.00	0
Asymptomatic	145.00	115.00	114.00	1
Atypical_angina	138.00	110.00	151.00	1

Input:

1: dataset

A) Explanatory identifier (E)

B) quasi-identifier characteristics

C) Sensitive characteristics

2: **(External factor)** EX

Output:

A) Sensitive Batch Table

B) Generalized Table

Fig. 7 "(p, K) – Angelization" [14].

Table 13 Sliced-sensitive-attributes (CP, THALACH)

Tuple ID	cp	thalach	Group
P_1	Typical_angina	150.00	1
P_2	Asymptomatic	162.00	
P_3	Atypical_angina	148.00	
P_4	Atypical_angina	178.00	2
P_7	Non-anginal_pain	187.00	
P_9	Atypical_angina	151.00	
P_5	typical_angina	153.00	3
P_8	Asymptomatic	114.00	
P_6	Atypical_angina	172.00	

Table 14 Sliced-sensitive-attributes (RESTEC, TRESTBPS, FBS)

Tuple ID	restecg	trestbps	fbs	Group
P_1	0	145.00	108.00	1
P_2	1	130.00	101.00	
P_5	0	160.00	125.00	
P_4	1	172.00	130.00	2
P_3	0	120.00	110.00	
P_6	1	150.00	130.00	
P_7	0	140.00	105.00	3
P_8	1	145.00	110.00	
P_9	1	138.00	115.00	

Table 15 Generalized table

cp	trestbps	fbs	thalach	Restecg	Batch ID
Typical angina, asymptomatic	145, 130	108, 101	150, 162	0,1	1
Non-anginal pain, Atypical angina	140, 172	105, 130	153, 178	0,1	2
Atypical angina, asymptomatic	120, 145	110, 115	148,115	0,1	3
Typical angina, Atypical angina	160, 138	125, 110	187, 151	0,1	4

Table 16 Sensitive batch table (SBT)

Age	Sex	Batch ID
[20–30]	Person_	**1**
[20–30]	Person_	
[20–30]	Person_	
[30–38]	Person_	**2**
[30–38]	Person_	
[38–45]	Person_	**3**
[38–45]	Person_	
[45–50]	Person_	**4**
[45–50]	Person_	

4 Conclusion and Future Direction

Finally, recent improvements in healthcare data privacy strategies have allowed healthcare organisations to share data while maintaining patient privacy. Homomorphic encryption, differential privacy, and federated learning are just a few of the advanced solutions created to solve healthcare data privacy concerns. As healthcare data becomes more valuable, these strategies are likely to expand and improve, allowing for even deeper insights while yet respecting patient privacy. The chapter finishes with a consideration of future research topics and issues that must be addressed in order to advance healthcare data privacy even further. It addresses difficulties such as enhancing the scalability of advanced privacy strategies, addressing data quality concerns, and developing new privacy techniques for emergent healthcare data types.

References

1. Sweeney, L. (2002). Achieving k-anonymity privacy protection using generalization and suppression. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 10*(05), 571–588.
2. Nergiz, M. E., & Clifton, C. (2007). Thoughts on k-anonymization. *Data & Knowledge Engineering, 63*(3), 622–645.
3. Kifer, D., & Gehrke, J. (2006). Injecting utility into anonymized datasets. In *Proceedings of the 2006 ACM SIGMOD international conference on Management of data*. Association for Computing Machinery.
4. Ciriani, V., De Capitani di Vimercati, S., Foresti, S., & Samarati, P. (2007). Microdata protection. In *Secure data management in decentralized systems* (pp. 291–321). Springer.
5. Li, N., Li, T., & Venkatasubramanian, S. (2006). t-closeness: Privacy beyond k-anonymity and l-diversity. In *2007 IEEE 23rd international conference on data engineering*. IEEE.
6. LeFevre, K., DeWitt, D. J., & Ramakrishnan, R. (2006). Mondrian multidimensional k-anonymity. In *22nd International conference on data engineering (ICDE'06)*. IEEE.
7. Abdelhameed, S. A., Moussa, S. M., & Khalifa, M. E. (2018). Privacy-preserving tabular data publishing: A comprehensive evaluation from web to cloud. *Computers & Security, 72*, 74–95.

8. Xiao, X., & Tao, Y. (2006). Anatomy: Simple and effective privacy preservation. In *Proceedings of the 32nd international conference on Very large data bases*. VLDB Endowment.
9. Susan, V. S., & Christopher, T. (2016). Anatomisation with slicing: A new privacy preservation approach for multiple sensitive attributes. *Springerplus, 5*(1), 1–21.
10. Onashoga, S. A., et al. (2017). KC-Slice: A dynamic privacy-preserving data publishing technique for multisensitive attributes. *Information Security Journal: A Global Perspective, 26*(3), 121–135.
11. Qin, Z., et al. (2016). Heavy hitter estimation over set-valued data with local differential privacy. In *Proceedings of the 2016 ACM SIGSAC conference on computer and communications security*. ACM.
12. Anjum, A., Ahmad, N., Malik, S. U. R., Zubair, S., & Shahzad, B. (2018). An efficient approach for publishing microdata for multiple sensitive attributes. *The Journal of Supercomputing, 74*(10), 5127–5155.
13. Albulayhi, K., Tošić, P. T., & Sheldon, F. T. (2020). G-model: A novel approach to privacy-preserving 1: M microdata publication. In *2020 7th IEEE International Conference on Cyber Security and Cloud Computing (CSCloud)*. IEEE.
14. Li, X., & Zhou, Z. (2020). A generalization model for multi-record privacy preservation. *Journal of Ambient Intelligence and Humanized Computing, 11*(7), 2899–2912.
15. Jayapradha, J., & Prakash, M. (2021). An efficient privacy-preserving data publishing in health care records with multiple sensitive attributes. In *2021 6th International conference on inventive computation technologies (ICICT)*. IEEE.
16. Darwish, N. R., & Hefny, H. A. (2015). Utility-based anonymization using generalization boundaries to protect sensitive attributes. *Journal of Information Security, 6*(03), 179.
17. Fu, A. W.-C., et al. (2014). Small sum privacy and large sum utility in data publishing. *Journal of Biomedical Informatics, 50*, 20–31.
18. Mohan, P., et al. (2012). GUPT: Privacy preserving data analysis made easy. In *Proceedings of the 2012 ACM SIGMOD international conference on management of data*. Association for Computing Machinery.
19. Zhu, T., et al. (2018). Answering differentially private queries for continual datasets release. *Future Generation Computer Systems, 87*, 816–827.
20. Gkoulalas-Divanis, A., Loukides, G., & Sun, J. (2014). Publishing data from electronic health records while preserving privacy: A survey of algorithms. *Journal of Biomedical Informatics, 50*, 4–19.
21. Anter, A. M., & Abualigah, L. (2023). Deep federated machine learning-based optimization methods for liver tumor diagnosis: A review. *Archives of Computational Methods in Engineering, 30*(5), 3359–3378.
22. Aldeen, S., Alsahib, Y. A., & Salleh, M. (2016). A hybrid K-anonymity data relocation technique for privacy preserved data mining in cloud computing. *Journal of Internet Computing and Services, 17*(5), 51–58.
23. Anter, A. M., Oliva, D., Thakare, A., & Zhang, Z. (2021). AFCM-LSMA: New intelligent model based on Lévy slime mould algorithm and adaptive fuzzy C-means for identification of COVID-19 infection from chest X-ray images. *Advanced Engineering Informatics, 49*, 101317.
24. Zaman, A. N. K., Obimbo, C., & Dara, R. A. (2016). A novel differential privacy approach that enhances classification accuracy. In *Proceedings of the ninth international C* conference on computer science & software engineering*. ACM.
25. Jaiswal, J. K., Samikannu, R., & Paramasivam, I. (2016). Anonymization in PPDM based on data distributions and attribute relations. *Indian Journal of Science and Technology, 9*, 37.
26. Zaman, A. N. K., & Obimbo, C. (2014). Privacy preserving data publishing: A classification perspective. *International Journal of Advanced Computer Science and Applications, 5*(9).
27. Majeed, A., Ullah, F., & Lee, S. (2017). Vulnerability-and diversity-aware anonymization of personally identifiable information for improving user privacy and utility of publishing data. *Sensors, 17*(5), 1059.

28. Lee, H., et al. (2017). Utility-preserving anonymization for health data publishing. *BMC Medical Informatics and Decision Making, 17*(1), 1–12.
29. Kohlmayer, F., et al. (2012). Flash: Efficient, stable and optimal k-anonymity. In *2012 International conference on privacy, security, risk and trust and 2012 international confernece on social computing*. IEEE.
30. Nergiz, M. E., & Gök, M. Z. (2014). Hybrid k-anonymity. *Computers & Security, 44*, 51–63.
31. Li, T., et al. (2010). Slicing: A new approach for privacy preserving data publishing. *IEEE Transactions on Knowledge and Data Engineering, 24*(3), 561–574.
32. Han, J., et al. (2013). G-model: A privacy preserving data publishing method for multiple sensitive attributes microdata. *JSW, 8*(12), 3096–3104.
33. Liu, Q., Shen, H., & Sang, Y. (2015). Privacy-preserving data publishing for multiple numerical sensitive attributes. *Tsinghua Science and Technology, 20*(3), 246–254.
34. Abdalaal, A., Nergiz, M. E., & Saygin, Y. (2013). Privacy-preserving, publishing of opinion polls. *Computers & Security, 37*, 143–154.
35. Luo, F., et al. (2013). ANGELMS: A privacy preserving data publishing framework for microdata with multiple sensitive attributes. In *2013 IEEE Third international conference on information science and technology (ICIST)*. IEEE.
36. Poulis, G., et al. (2013). Anonymizing data with relational and transaction attributes. In *Joint European conference on machine learning and knowledge discovery in databases*. Springer.
37. Gong, Q., et al. (2017). Anonymizing 1: M microdata with high utility. *Knowledge-Based Systems, 115*, 15–26.
38. Dhumal, T. S., & Patil, Y. S. (2015). Implementation of slicing for multiple column multiple attributes: Privacy preserving data publishing. *International Journal on Recent and Innovation Trends in Computing and Communication, 3*, 4261–4266.
39. Das, D., & Bhattacharyya, D. K. (2012). Decomposition +: Improving ℓ-diversity for multiple sensitive attributes. In *International Conference on Computer Science and Information Technology*. Springer.
40. Terrovitis, M., Mamoulis, N., & Kalnis, P. (2011). Local and global recoding methods for anonymizing set-valued data. *The VLDB Journal, 20*(1), 83–106.
41. Anil, P. D., & Ravindar, M. (2012). Privacy preservation measure using t-closeness with combined l-diversity and k-anonymity. *International Journal of Advanced Research in Computer Science and Electronics Engineering, 1*(8), 28–33.

The Ability of the CFD Approach to Investigate the Fluid and Wall Hemodynamics of Cerebral Stenosis and Aneurysm

Talaat Abdelhamid and Ahmed G. Rahma

1 Introduction

Investigation of hemodynamics characteristics in the vascular system of patients holds significant value for understanding and diagnosing various diseases [1, 2] such as stenosis [3], aneurysms of different types (saccular [4], fusiform [5], false, and dissection [6]), …, etc. Each year, 15 million people worldwide experience a stroke, with 33.3% of them resulting in death and another 33.3% resulting in permanent disability. Although stroke is not a common occurrence among the young, it can occur due to high blood pressure when it does [7]. Currently, a major concern is gaining insight into cerebrovascular diseases and discovering the relationships they have with parameters of hemodynamics, such as the fluid vortex, and wall pressure/shear stress distribution. Luo et al. [8] conducted a computational fluid dynamics investigation on an MRI-based model of a cerebral artery with stenosis in order to examine its effect on the hemodynamics characteristics. The study began with a benchmark to ensure the boundary conditions and the computational methods, which were adopted by a straight pipe with an ideal symmetric stenosis of $d/D = 0.4$. Where the d/D, is the contraction diameter ratio to the pipe diameter. In the part two of the simulation. The simulations are done on MRI-based model of a cerebral

T. Abdelhamid (✉)
Physics and Mathematical Engineering Department, Faculty of Electronic Engineering, Menoufiya University, Menouf, Egypt

Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China
e-mail: Talaat_abdelhamid@el-eng.menofia.edu.eg

A. G. Rahma
Physics and Mathematical Engineering Department, Faculty of Electronic Engineering, Menoufiya University, Menouf, Egypt

Department of Physics and Engineering, University of Strasbourg, Strasbourg, France

A. M. Anter et al. (eds.), *Nature-Inspired Methods for Smart Healthcare Systems and Medical Data*, https://doi.org/10.1007/978-3-031-45952-8_12

artery with 2 inlets and 15 outlets and measured boundary conditions such as the pressure outlet and the velocity inlet profiles are used. The results indicate that the stenosis is leading to a loss of pressure in the distal area.

Intracranial aneurysms are bulges in the intracranial arteries caused by a weakness in the blood vessel wall [9–11]. They affect a global population of between 0.5% and 6%. Unfortunately, rupture of these intracranial aneurysms (IAs) can occur in approximately 0.25% of cases, leading to death in 60% of cases [12, 13]. The study conducted by Souza et al. [14] aimed to analyze the impact of the hemodynamics and the structure on the behavior of intracranial aneurysms. By utilizing CFD simulations on a numerical model, they analyzed the parameters that have a significant effect on an aneurysm's rupture. The simulations were run at various Reynolds values (1–1000) using the Fluent 2020 R2 commercial package. The results revealed that an increase in *Re* leads to the formation of flow recirculation in the aneurysm. Additionally, the maximum values of WSS, displacement, and strain were observed at the highest Reynolds number.

Additionally, CFD simulations can be performed using two types of models [8, 15], the first is the benchmark model which is a simplified version of the vascular system [16–18]. The second is the patient-specific model, which is based on the real geometry of the patient's vasculature [19–21]. These simulations provide valuable information regarding the flow patterns and the wall stress/pressure distribution at cerebral vessels, which are crucial for understanding and diagnosing cerebrovascular diseases [22–24]. The results obtained from these simulations can be compared with the literature and validated to ensure the reliability of the CFD approach. Perinajová et al. [25] focused on analyzing the turbulence and WSS in the Narrowing of the aorta using computational fluid dynamics (CFD) simulations. To validate their findings, the researchers utilized a flow phantom of a vessel with 180 degree bending and narrowing with pulsating flow and compared the results with those obtained from a 4D-flow MRI technique. The results demonstrated a remarkable link between the experimental outcomes from the experimental approach and the CFD approach.

Xiang et al. [26] studied the effects of the hemodynamics on the rupture of IA. They found that the use of CFD simulations is effective in evaluating the risk of intracranial aneurysms when real patient data are available. The results indicated the importance of hemodynamic analysis in predicting the likelihood of an aneurysm rupture. Chen et al. [27] aimed to simplify the modeling process for studying the blood flow in a stroke patient's cerebral artery. Their research revealed that numerical simulations can be a time-efficient method for pre-and post-surgery planning, making it a valuable tool for medical professionals.

In this chapter, numerical simulations are performed for the entire cerebral arteries of two patients. The computational domain used for the simulation is divided into two segments (the benchmarks and the actual arteries reconstructed from the MRI). The problem is described, and the system of incompressible Navier-Stokes equations is introduced. Furthermore, a comprehensive study of the blood flow and wall hemodynamic parameters, including the WSS, wall pressure, and flow recirculation, is conducted on benchmarks before the solution is refined to incorporate actual patient-specific cerebral arteries.

2 Methodology

2.1 Problem Outline and the Numerical Approach

In this section, we examine two types of cerebral cardiovascular conditions – stenosis and aneurysm. The defined parameters of the simulation at the boundary conditions are derived from the ultrasonography scans. We employ a consistent CFD model to analyze the hemodynamics in both benchmark segment and patient-specific cerebral arteries segment with modified velocity/pressure profiles.

2.1.1 Cerebral Arteries Stenosis

In the first instance, the benchmark study involves simulating blood flow in pipes with a diameter of 3.5 mm and length of 61.6 mm, with a range of vascular Narrowing ratios variation from (0% to 60%) as depicted in Fig. 1. This contraction is assumed to represent an inborn medical condition in the patient. In the second case, a realistic scenario for MRI based model of the cerebral artery is taken into account [27], as shown in Fig. 16. Both studies are conducted using a consistent CFD model to present the blood flow dynamics in the given geometries.

2.1.2 Cerebral Arteries Aneurysm

The CFD simulations provide a comprehensive analysis for the flow characteristics of a benchmark of the intracranial artery (IA), as depicted in Fig. 2. The model is built and analyzed using boundary conditions derived from the literature [14]. Subsequently, the same characteristics are investigated for an MRI based model of

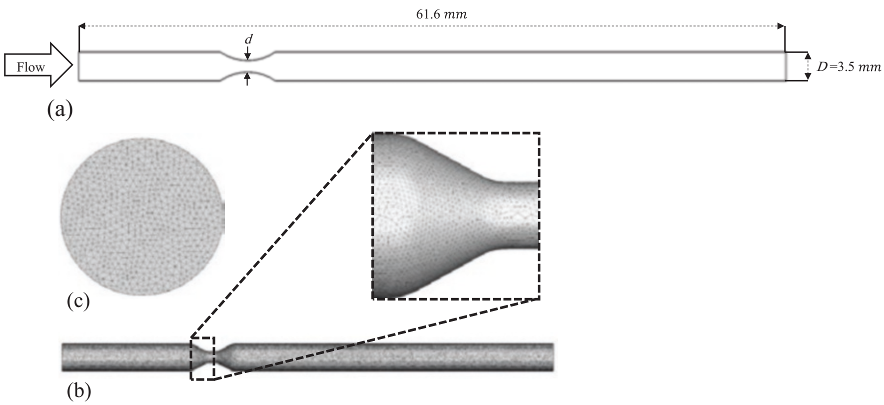


Fig. 1 (**a**) The benchmark domain of computation, (**b**) The model meshing and zoom in of the stenosis zone, and (**c**) zoom-in for the gridding system of the vessel inlet

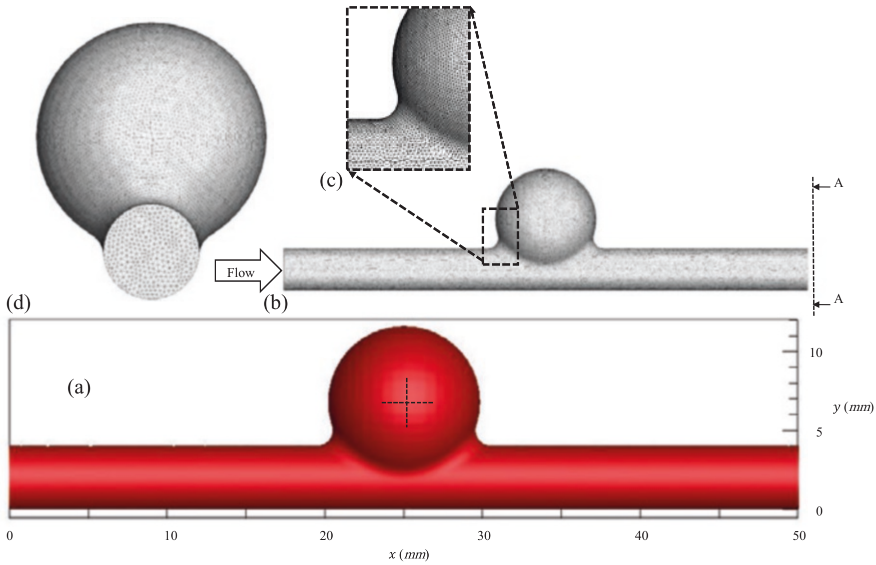


Fig. 2 (**a**) The computational domain (**b**) The meshing for the benchmark (**c**) Zoomed in for the meshing system of the aneurysm zone, and (**d**) Section A-A

a cerebral aneurysm at the anterior communicating artery (AAA). This model is built from MRI data by the commercial software package SimVascular (Fig. 17).

2.2 *Governing Equations*

Steady state, three-dimensional CFD simulations were conducted for both geometries. The flow of blood is governed by the differential equations presented in [28, 29]. The motion of blood was described using the Navier-Stokes equations in a laminar flow state, as follows:

$$\Delta(\rho u) = 0 \tag{1}$$

$$\Delta(\rho u u) = -\frac{\partial p}{\partial x} + \Delta(\mu \nabla u) \tag{2}$$

$$\Delta(\rho u v) = -\frac{\partial p}{\partial y} + \Delta(\mu \nabla v) \tag{3}$$

$$\Delta(\rho u w) = -\frac{\partial p}{\partial z} + \Delta(\mu \nabla w) \tag{4}$$

Table 1 Properties of blood flow used for the benchmarks with stenosis

Properties	Value
Blood density	1063 (Kg.m^{-3})
Velocity in	10.525 (cm/s)
Pressure at the inlet (Gauge)	11,999 Pa (90 mmHg)

Where "*u*", "*v*", and "*w*" are the blood *x*, *y*, *z* velocity components, respectively. The ρ parameter is representing the blood density while pressure defined as *P*. Blood is treated as an incompressible non-Newtonian fluid, with a viscosity model described by the Carreau model, and ρ = 1063 kg/m^3. The walls of the computational domain are assumed to be rigid, with no slipping, as indicated in Tables 1, 2, and 3 for the blood properties at the stenosis, aneurysm and the four model parameter law (Carreau model) [14, 30]. The equations for the Carreau flow model can be found in [29].

$$\mu = H(t)\left(\mu_\infty + (\mu_0 - \mu_\infty)\left[1+\gamma^2\lambda^2\right]^{\frac{n-1}{2}}\right) \tag{5}$$

Where (H(t)) represents the temperature dependence, which is described by the Arrhenius law.

$$H(t) = \exp\left[\alpha\left(\frac{1}{T-T_0}\right)-\left(\frac{1}{T_\alpha - T_0}\right)\right] \tag{6}$$

where α is the ratio of the activation energy to the thermodynamic constant and T_α is a reference temperature for which $H(t) = 1$. The T_0 is the temperature shift. The λ is the time constant.

Simulations were performed in two segments using the Ansys Fluent 19.2 software package [29] to solve the governing equations in a 3D computational domain. The Reynolds number was less than 2200, indicating laminar flow. Second-order upwind schemes were applied for all conservation equations. The pressure-velocity coupling used a coupled scheme and pressure was calculated with a second-order scheme. The pseudo transient for fluid and solid zones were 0.7 and 1.0, respectively, to achieve strong convergence for the complex blood flow.

2.3 Meshing

To ensure the accuracy of the results, a grid independence study was conducted and the resulting mesh can be seen in Figs. 1, 2, 16b, and 18e. The mesh was carefully constructed to be refined near the vessel wall and to maintain a *y* + value of

Table 2 Properties of blood flow used for the benchmarks with the aneurysm

Properties	Value (Unit)
Blood density	1063 ($Kg.m^{-3}$)
Velocity in	88.8 (cm/s)
Pressure at the inlet (Gauge)	0 (*Pa* Gauge)

Table 3 Carreau model four parameter values [30]

μ_∞ $(Pa.\ s)$	μ_0	n	λ (s)
1.0	0.4360	0.360	3.30

approximately 1. Additionally, the standard wall treatment was employed to facilitate a smooth transition. As the CFD study of blood flow is strongly reliant on the wall effect, efforts were made to maintain a y + value close to 1, despite the laminar nature of the flow.

2.3.1 Cerebral Arteries Stenosis

The results of the computational fluid dynamics (CFD) simulations are presented by utilizing the tetrahedral mesh. The mesh size was established based on the recommendations of a previous study [8]. The total elements in the grid system are 897,865 for the benchmark with stenosis and 1,292,839 for the MRI-based model with stenosis. To ensure the accuracy of the results, a comprehensive mesh sensitivity study was imported from a previously study, which improved the confidence in the benchmark and the MRI-based model with stenosis models results [8].

2.3.2 Cerebral Arteries Aneurysm

For the IA benchmark, 522,323 tetrahedral elements are used. For the patient-specific IA, 722,724 tetrahedral elements are used to grid the realistic aneurysm model, for more details regarding the mesh size see [8].

2.4 Validation

2.4.1 The Benchmark with a Stenosis

The accuracy of the CFD results obtained in this study have been validated by comparing them with previously published results [8]. As seen in Fig. 3, the velocity and pressure contours produced by the present CFD model match well with those reported in [8]. Furthermore, a comparison of pressure drop values for the

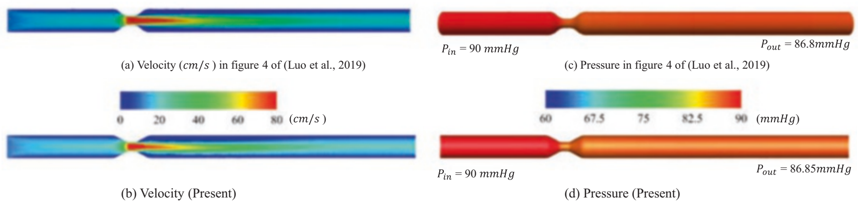


Fig. 3 The velocity and pressure contours for the benchmark with a 40% stenosis (*d/D* = 0.4) with velocity inlet 10.525 cm/s, for (**a**), (**c**) Luo et al. [8], and (**b**), (**d**) current study

benchmark between the present CFD results and those reported in [8] is shown in Table 4. The absolute error in the pressure drop between the current study and Luo, et al. [8] is found to be only 1.56%, which is considered to be acceptable. This validation confirms the validity of the CFD approach to investigate the hemodynamic of the blood.

2.4.2 The Benchmark with an Aneurysm

The CFD simulation outcomes were verified using data from previous studies [14]. The absolute error of the pressure drop across the (IA) benchmark was found to be around 3.261% according to Table 5 [14]. Additionally, Fig. 4 provides a visual representation of the results validation by comparing the wall pressure contours from the current study and the previous study in the literature review [14]. This comparison confirms the capability of the current CFD simulations to accurately evaluate the hemodynamic parameters, particularly for IA cases.

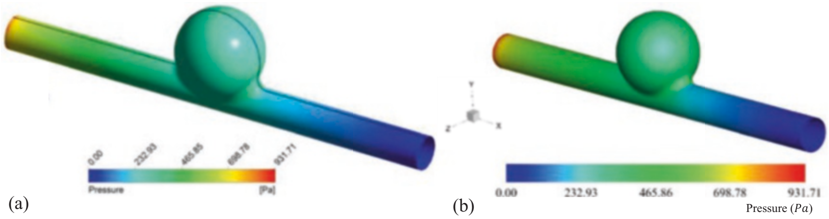
3 Results

3.1 Segment One, Results for the Benchmarks

3.1.1 The Benchmark with a Stenosis

This section presents the outcomes of the CFD simulation for the benchmark scenario involving the idealized straight vessel with a reduced area of flow. The impact of the contraction area ratio (*d/D*) on the artery inlet to outlet pressure ratio is depicted in Fig. 5. The results demonstrate that as the ratio increases, the pressure drop (pressure difference) across the artery contraction decreases, indicating a reduced energy loss for the flow passing through the contraction.

A. **The pressure distribution:**
Figure 6 illustrates the pressure contour across the benchmarks with stenosis. The pressure drop increases with decreasing the *d/D* ratio. This denotes that as the



(a) Wall pressure in figure 4 of (Souza et al. 2022)

(b) Wall pressure (Present)

Fig. 4 The wall pressure distribution across the benchmark of an aneurysms where the inlet velocity = 88.8 cm/s, for (**a**) MacDonald et al. [14], and (**b**) present study

Table 4 The pressure drops for the benchmark with a stenosis *d/D* = 0.4, compared with the literature results

Reference	$P_{in} - P_{out}$
Luo et al. [8]	3.200 (mmHg)
Present study	3.150 (mmHg)
Absolute error	1.563 (%)

Table 5 The pressure drop across the benchmark with an aneurysm, compared with the literature results

Reference	$P_{in} - P_{out}$
Souza et al. (2022) [14]	726.1200 (Pa)
Present study	749.8021 (Pa)
Absolute error	3.26 (%)

cross-section area of the contraction decreases, the required energy to overcome the losses also increases. The results of the pressure difference can be seen in Fig. 5, which shows the inlet and outlet pressure ratio with the contraction area ratio. The development of the flow in the benchmark can be observed in Fig. 6g. The pressure profile along the centerline of the model demonstrates that the flow has reached a fully developed state, and the entrance length is adequate to achieve this. The error was found to be 2.1%, which is considered acceptable within computational standards. Reducing the area of the vessel can result in an increase in the blood velocity and the formation of vortices as in Figs. 7, 8, and 9. These changes in hemodynamics can greatly affect the flow dynamics.

B. **The flow structures**

The contraction of the blood flow leads to the formation of eddies and vortices, which are particularly evident at lower area ratios. This flow behavior is a result of the contraction or expansion of the blood flow and can extract energy from the flow and even impede its passage in some cases. A lower area of contraction leads to

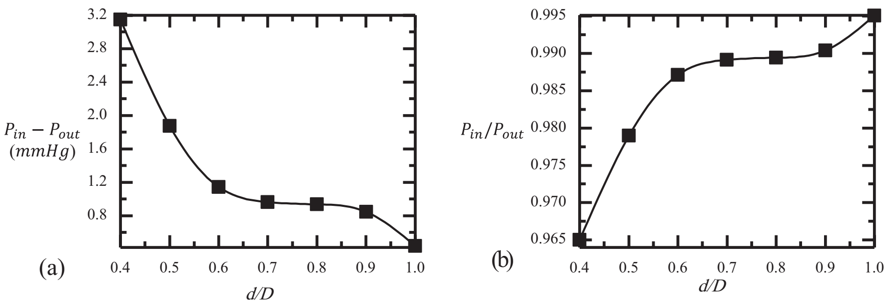


Fig. 5 The artery from inlet to outlet (**a**) The pressure drop, and (**b**) Pressure ratio of the benchmarks with stenosis

blood flow accumulation, which is indicative of a medical issue and may necessitate a therapeutic intervention. These flow characteristics can be easily visualized in the streamline plots presented in Figs. 8 and 9. Additionally, the alteration of the blood flow hemodynamics is further demonstrated by the increase in blood velocity and the creation of vortices, as shown in Figs. 7, 8, and 9. The formation of vortices, due to an increase in fluid flow turbulence, causes asymmetry in the flow and is evident in values of *d/D* such as 0.4, 0.5, and 0.6, as depicted in Figs. 7a–c and 8a–c.

C. **The wall shear stress (WSS):**

The relationship between the stenosis geometry and the wall shear stress (WSS) is illustrated in Fig. 10. The WSS value at the stenosis zone increases as the area of the stenosis (*d/D*) decreases. However, beyond the stenosis zone, the WSS value begins to decline as the *d/D* value decreases. This decline in WSS value can result in the formation of flow recirculation, particularly when *d/D* is less than or equal to 0.4, as demonstrated in Figs. 9 and 10h.

3.1.2 The Benchmark with an Aneurysm

This section presents a thorough examination of the blood flow characteristics, including the wall pressure/shear stress, and recirculation flow within the IA benchmark.

A. **The pressure distributions**

In Figs. 4 and 11, a decrease in wall pressure is observed in the direction of the flow, with an exception for the ideal IA zone. In this zone, the wall pressure remains constant due to the presence of recirculation flow. In the region near the trailing edge of the intracranial aneurysm, the flow direction is observed to reverse due to the interaction between the wall and the flow. This interaction results in an increase in wall pressure values.

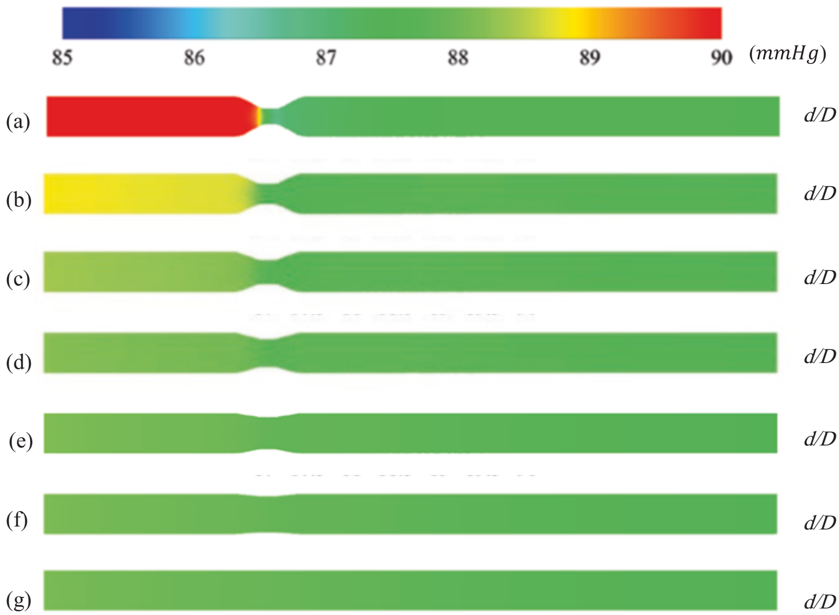


Fig. 6 Pressure distribution contours at center cross section for the benchmark with different *d/D* ratio

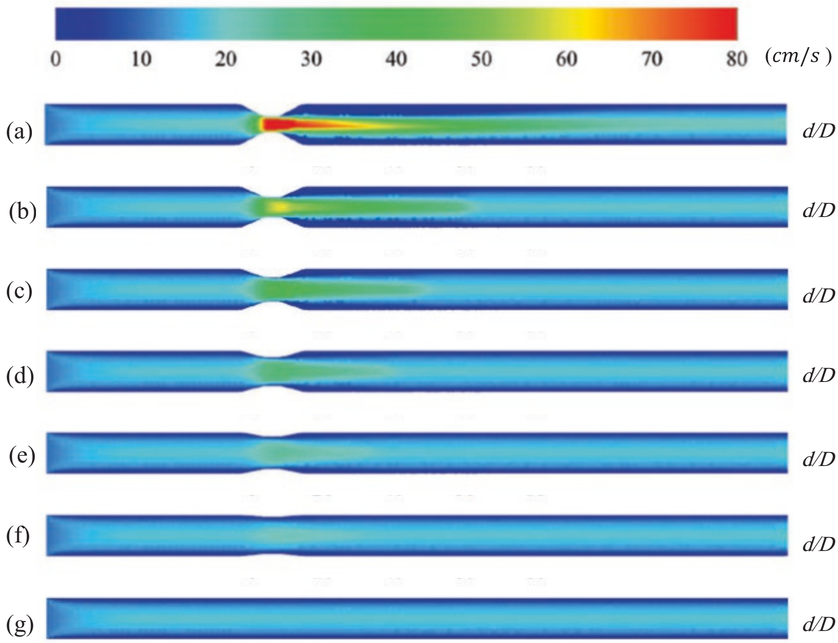


Fig. 7 Velocity distribution contours at center cross section for the benchmark with different *d/D*

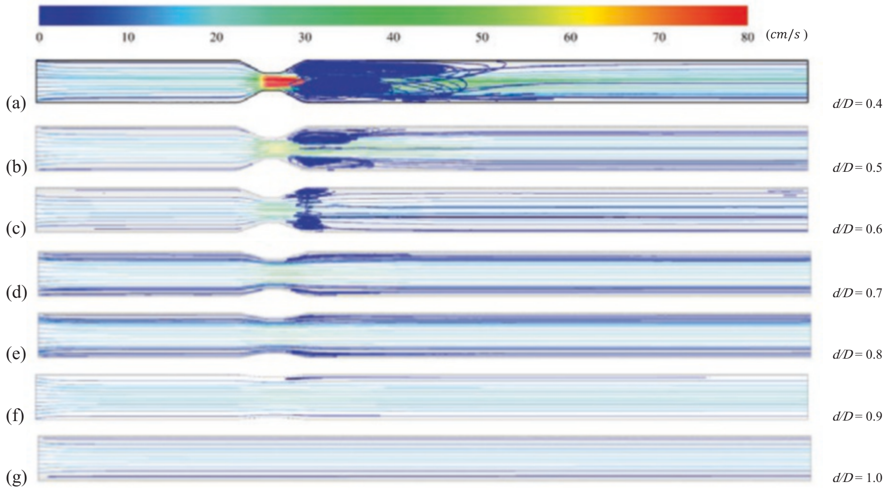


Fig. 8 The velocity streamline contours for the benchmark with different *d/D*

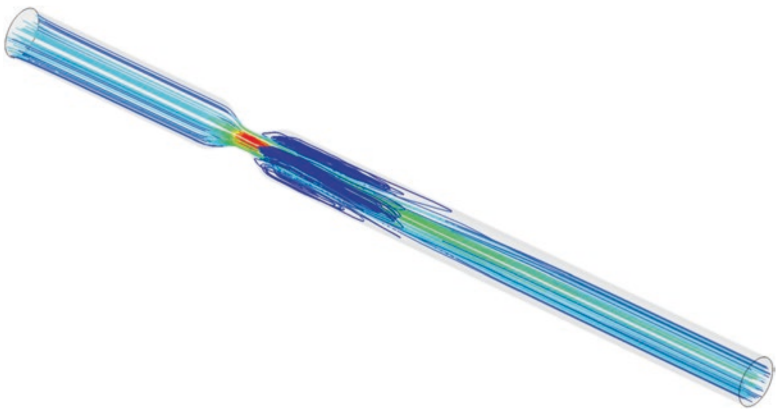


Fig. 9 Three dimensional velocity streamlines contours for the benchmark at *d/D* = 0.4

B. **The flow structure:**

The blood velocity has a profound influence on the occurrence of the flow recirculation. As the velocity distribution increases, the likelihood of flow recirculation also increases [14, 31]. This can be seen in Fig. 12 where two flow regimes are present. The flow regime is seen near the inlet and the outlet of the benchmark, while the formation of vortices is seen close to and inside the dilation. The recirculation of the flow in the aneurysm zone is a result of using a high flow velocity (88.8 cm/s), which leads to separation at the trailing edge of the IA and the formation of the flow recirculation.

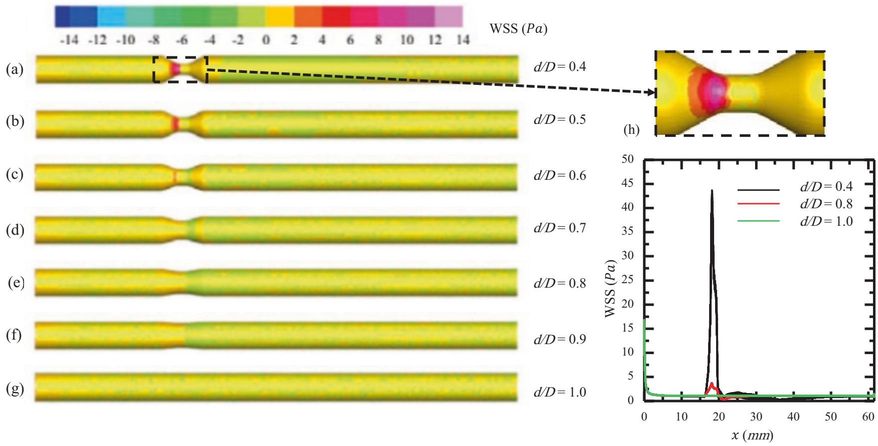


Fig. 10 The WSS distribution contours for the benchmark with variable *d/D*, and (**h**) Zoomed-in view for the stenosis zone at *d/D* = 0.4

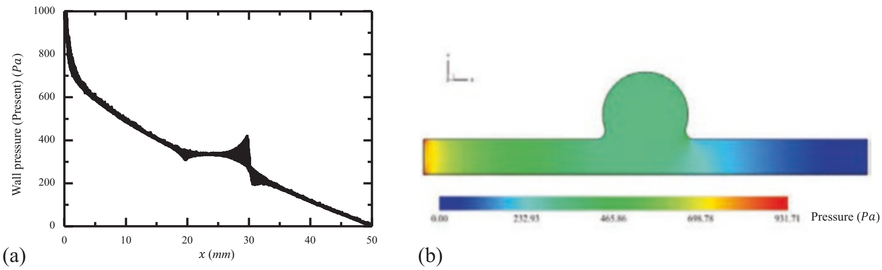


Fig. 11 The magnitude of the wall pressure (**a**) along x axis and (**b**) xy cross-section contour of pressure

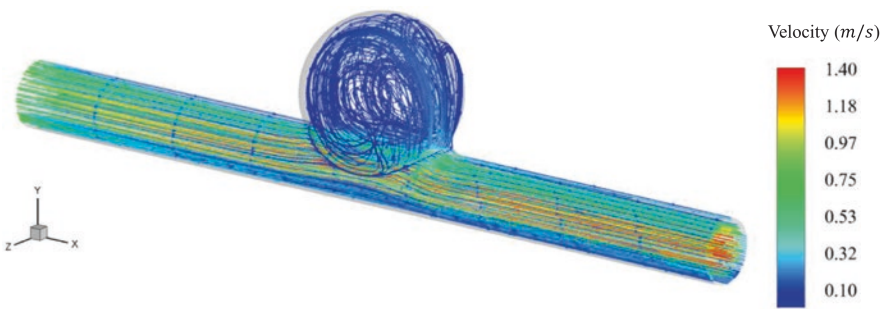


Fig. 12 Three-dimensional velocity streamlines contours for the intracranial aneurysm benchmark

The velocity profile at different locations is shown in Fig. 13. Three different cuts along the x direction are taken to display the velocity contours at three different positions – inlet, outlet, and the aneurysm center. The velocity profiles at the inlet and the outlet exhibit a symmetric distribution around the center of the y axis. The flow at the inlet is fully developed and follows a plug velocity boundary condition. The flow at the outlet is also fully developed. However, the presence of the saccular aneurysm results in a change in the velocity profile's symmetry. The right side of the aneurysm velocity profile exhibits a higher decrease in velocity compared to the center of the cylinder.

C. **The wall shear stress**

In Fig. 14, the configuration of the WSS around the dilation is analyzed. The behavior of the wall shear stress is consistent with the wall pressure configuration displayed in Figs. 4 and 11. The WSS decreases along the flow direction, except around the bulge. The WSS at the start of the aneurysm region stays unchanged because of the existence of flow recirculation. Interaction and separation at the rear edge of the bulge causes an increase in the WSS at its outlet.

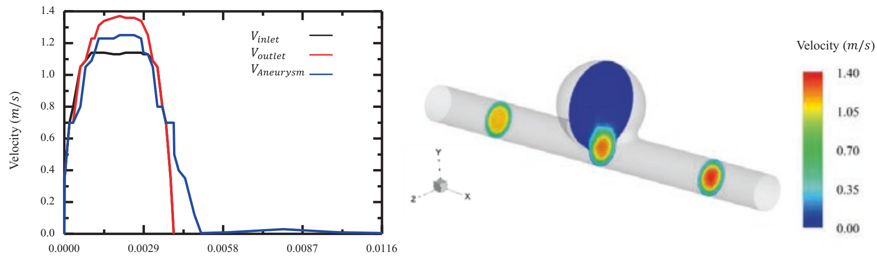


Fig. 13 (**a**) Absolute velocity curves at three different positions (**b**) threes cross-sections in the benchmarks corresponding to the plots at (**a**), respectively

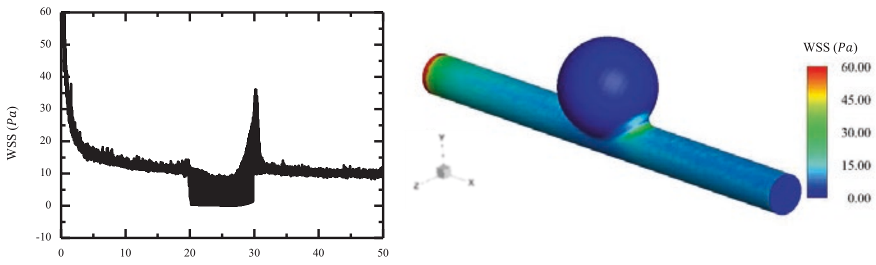


Fig. 14 (**a**) The magnitude of the WSS along the x direction, and (**b**) the magnitude of the WSS contour along the benchmark with intracranial aneurysm, respectively

3.2 Segment Two, Results for MRI-Based Models of the Cerebral Arteritis

In this section, CFD simulations were performed on realistic models of a stenosis and aneurysm in a patient-specific cerebral artery, reconstructed from MRI data. The computational domains were created using the SimVascular software as outlined in the methodology. An actual MRI of a cerebral artery was utilized, with a stenosis presented in the artery left middle part of the artery as shown in Fig. 15. The stenosis exhibits a contraction of approximately 55%, the model involved 2 inlets and 35 outlets (Fig. 16).

Regarding the aneurysm, the dimensions of the *yz*, *xy*, and *xz* planes are 9.91 *mm*, 6.62 *mm*, and 8.082 *mm*, respectively. We observe a significant increase of about 200% in the *yz* plane compared to the main vessel branch, Figs. 17 and 18.

Additionally, the velocity profiles at the artery inlet and the pressure profiles at the artery outlet that were both computed and measured are displayed in Fig. 19. The measured velocity profiles at the inlet and the pressure profiles at the outlet obtained through ultrasonography are shown over 4 cycles in Fig. 19a, c, respectively. The computed velocity profiles at the artery inlet and the pressure profiles at the artery outlet for one cycle, represented by red lines, are shown in Fig. 19b, d, respectively. These were extracted using polynomials equations obtained through curve fitting.

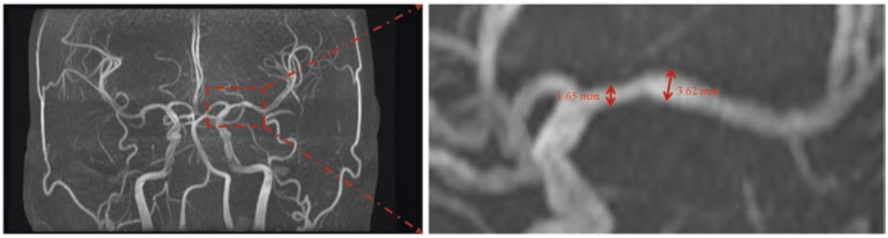


Fig. 15 MRI of the patient cerebral arteries with 55% area reduction stenosis [27]

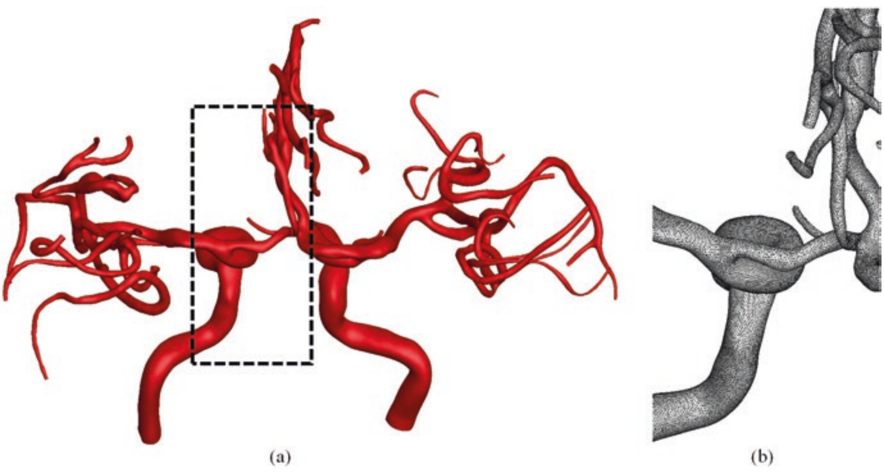


Fig. 16 (**a**) The MRI- based computational domain, and (**b**) a zoomed-in view of meshing system

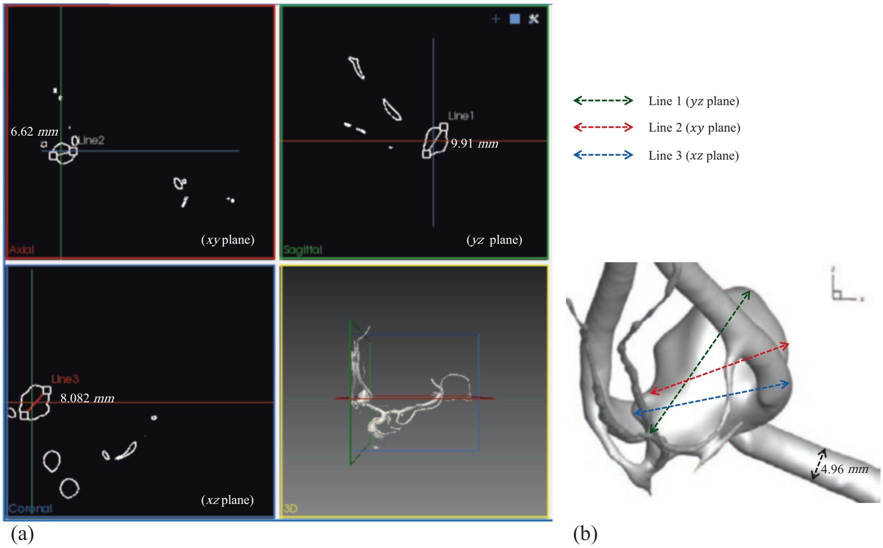


Fig. 17 (**a**) The aneurysm dimensions in the (*yz*, *xy*, and *xz*) directions, and (**b**) the dimension of the vessel before the aneurysm inlet

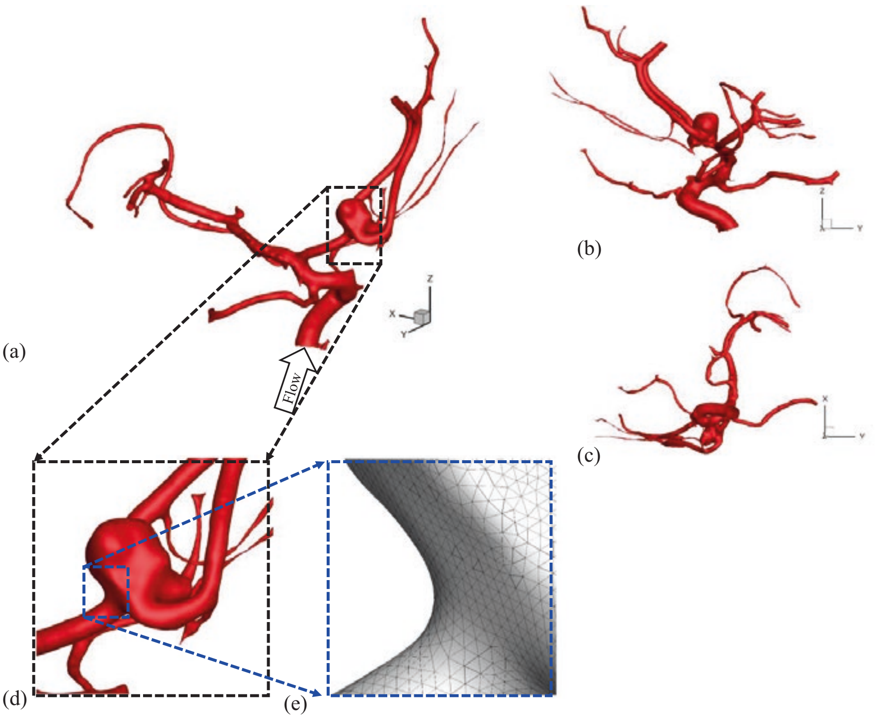


Fig. 18 (**a**) The MRI- based IA computational domain, (**b**, **c**) side-views of the model at *yz*, *xy* planes, respectively, (**d**) Zoomed in view of the IA aneurysm, and (**e**) zoomed-in view of the meshing system

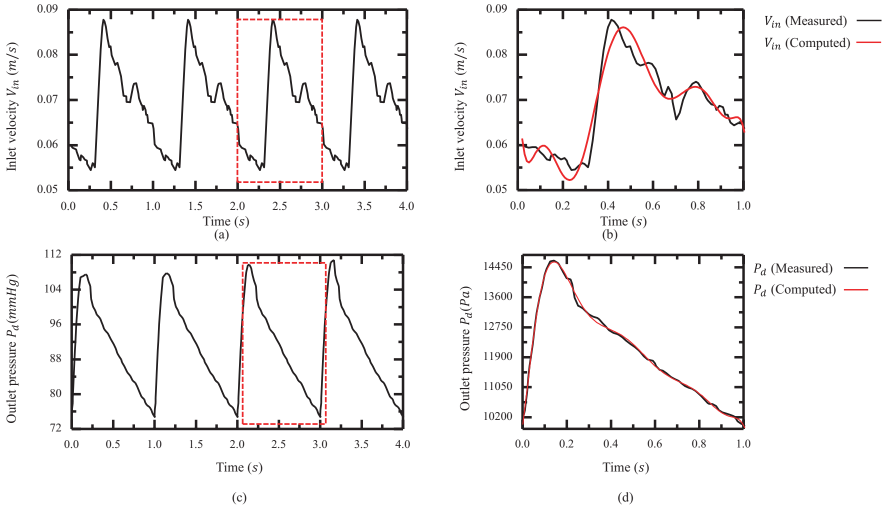


Fig. 19 (**a**–**b**) The cerebral artery inlet velocity profiles, and (**b**–**c**) the cerebral artery outlet pressure (**c**) in *mmHg*, and (**d**) *Pa*

3.2.1 MRI-Based Model With Stenosis

A. **The flow structure:**

1. Figure 20 displays the velocity streamlines and contours distribution. The velocity streamlines are displayed with a range from 0 to 0.7 meters per second. The zoomed-in views support the predictions made based on the results from segment one, as shown in Figs. 8 and 9. The formation of blood eddies and vortices can be seen as a result of the reduced area of the artery, causing a change in the symmetrical blood flow distribution near the wall. In Fig. 21, the velocity at the stenosis area and in the middle of the artery is higher compared to that near the vessel wall. The velocity configuration is not symmetrical, as seen in the zoomed-in view of the stenosis artery, when compared to the section of the normal artery without stenosis.

B. **Wall hemodynamics:**

2. In Fig. 20b, the pressure is seen to vary from 74 to 80 *mmHg*, and as anticipated from the benchmark results, the pressure drop is observed to increase between the inlet and outlet of the constricted area at the stenosis. Figure 22 shows that the WSS in the vicinity of the stenosis is more significant than that at other regions. A decrease in the WSS after the stenosis region is also observed in MRI-based model. High wall shear stress in the stenosis zone could produce an embolism due to the fluid-thrombus interaction, leading to vessel blockage and stroke.

3.2.2 MRI-Based Model with Aneurysm

A. **The flow structure:**

In Fig. 23a, b, the formation of vortex at the anterior communicating artery aneurysm (AAA) greatly impacts the wall pressure/shear stress and flow regimes at the outlet. The interaction of the fluid with the dilation wall results in flow separation.

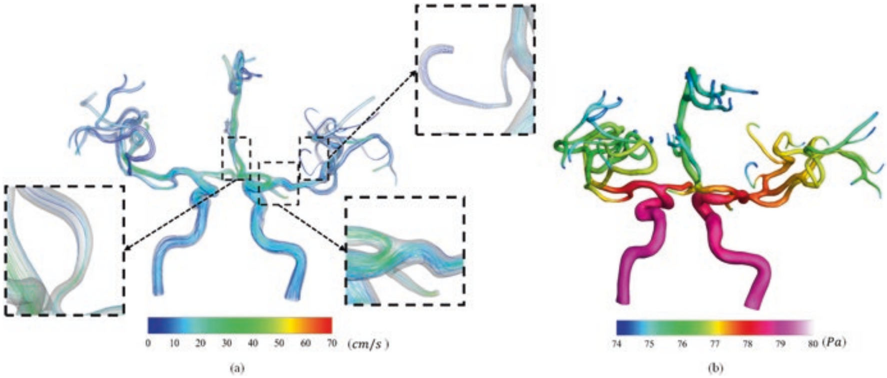


Fig. 20 (**a**) The velocity streamlines contours and (**b**) wall-pressure contour for the MRI-based model

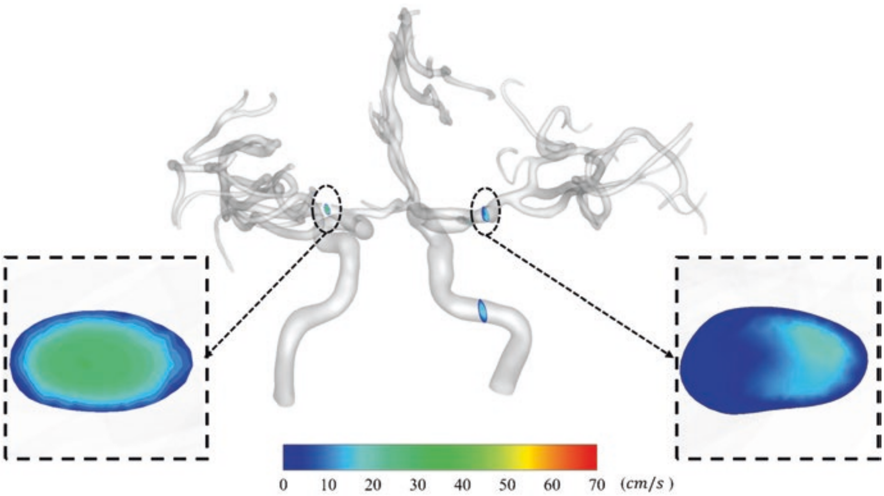


Fig. 21 Two cross-section views of velocity contours for the MRI-based model

The zone of separation experiences the highest values of both wall pressure and wall shear stress.

The velocity contours of the IA zone is displayed in two cross-sections, one in the *yz* plane and one in the *xz* plane, as shown in Fig. 24. In the *yz* plane, the highest velocity value is seen at the inlet of the artery and gradually decreases in the *z* direction, with the exception of the flow recirculation area. The interaction between the fluid and the dilation wall is depicted in Fig. 24c, as the velocity distribution vector shows the reverse flow which is responsible for flow recirculation in the z direction. In the *xz* plane, the highest magnitude of velocity is seen at the core of the flow inlet and decreases gradually in the *z* direction, except for the values close to the wall in

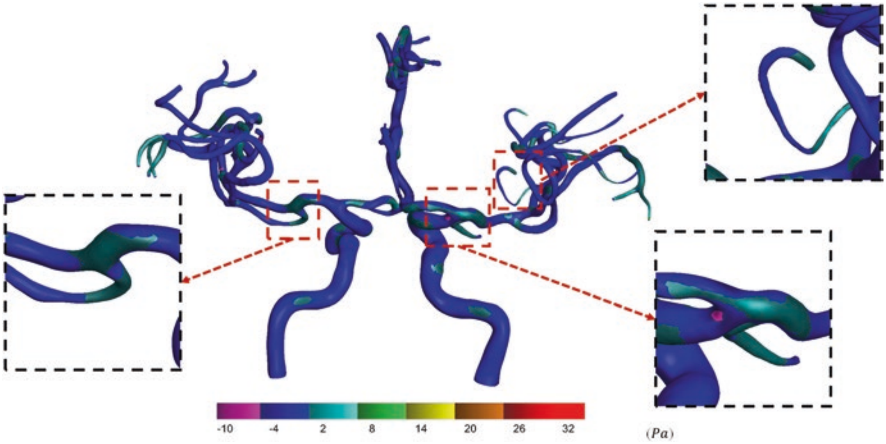


Fig. 22 The wall shear stress contour and zoom-in views of the areas close to the stenosis, and the branching for the MRI-based model

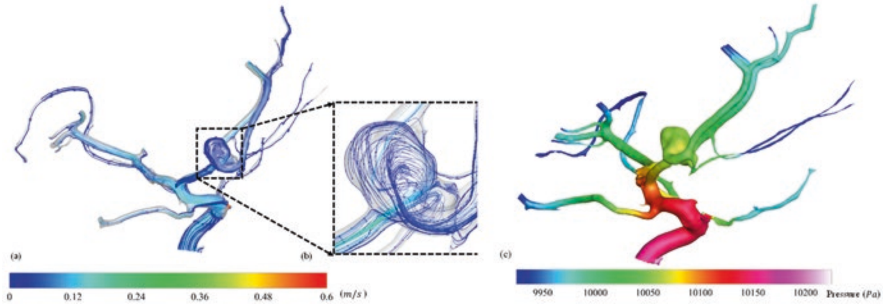


Fig. 23 (**a**) The velocity streamlines for the AAA (**b**) Zoomed-in view of the IA, and (**c**) the wall pressure contour along the MRI-based model of the IA

the *x* direction. The velocity is greater next to the wall in the right part of the *x* direction, due to the presence of the recirculation flow near the flow-wall interaction zone, as demonstrated in Fig. 24b and d.

B. **Wall Hemodynamics:**

The examination of the wall characteristics like the WSS and the wall pressure is crucial in comprehending the behavior of aneurysm. Figures 25 and 23c, the blood flow patterns impact the hemodynamics wall parameters. The wall shear stress and the wall pressure exhibit a gradual decline as the blood flows through, with the exception of the aneurysm area which is depicted in the benchmark at segment one (Figs. 11 and 14). In Fig. 25b, c, the wall shear stress near the inlet of the dilation is higher compared to the outlet zone. The fluid interaction with the edges results in an increase in the WSS at the outlet edges. The changes in WSS help predict the likelihood of aneurysm rupture, as research indicates that ruptures tend to occur in regions with low wall shear stress [32].

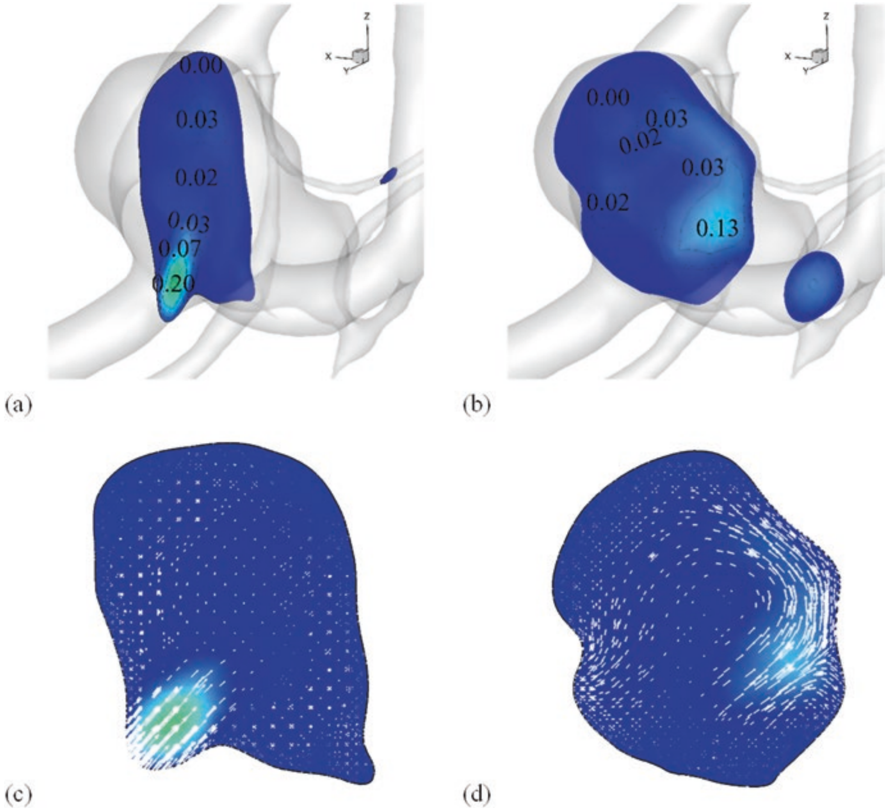


Fig. 24 Cross-sections of velocity distribution at (**a–b**) *yz*-plane and *xz*- plane, respectively. Cross-section of the velocity vector at (**c–d**) *yz* plane and *xz* plane, respectively

4 Conclusion

In this chapter, a systematic image-based computational fluid dynamics (CFD) method was introduced to simulate the blood flow in both benchmark and MRI-based models, to understand the hemodynamics of the complex vascular system. The results showed that the pressure difference across the arterial stenosis was reduced with increasing area ratio (*d/D*), resulting in less flow energy loss. The stenosis lead to increasing the velocity of the blood and produce vortexes, in some cases leading to vessel occlusion. The velocity at the stenosis areas was found to be relatively high and unevenly distributed, with high wall shear stress (WSS) values near the contraction area which could increase the risk of embolism and vessel occlusion. The presence of the saccular aneurysms in the vascular system revealed recirculation flow and uneven variation in WSS that may indicate the location of rupture. The wall pressure and the WSS behavior were found to be qualitatively matching.

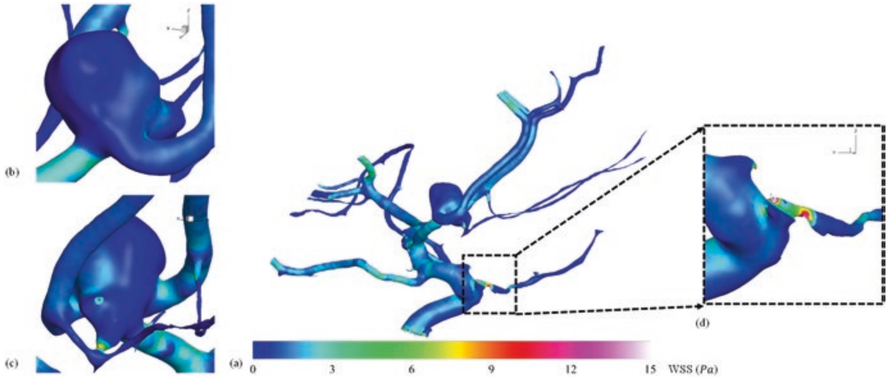


Fig. 25 (**a**) The WSS contour along the *MRI-based model of the intracranial aneurysm.* Zoom in view of (**b**) the inlet of the aneurysm inlet, (**c**) the outlet of the aneurysm, and (**d**) the stenosis zone

Acknowledgment This work was partially supported by the Science and Technology Development Fund (STDF) of Egypt, project No: 39385.

Availability of Data The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflict of Interest The authors declare that they have no conflict of interest.

References

3. Nixon, A. M., Gunel, M., & Sumpio, B. E. (2010). The critical role of hemodynamics in the development of cerebral vascular disease. *Journal of Neurosurgery, 112*, 1240–1253.
4. Cecchi, E., Giglioli, C., Valente, S., Lazzeri, C., Gensini, G. F., Abbate, R., & Mannini, L. (2011). Role of hemodynamic shear stress in cardiovascular disease. *Atherosclerosis, 214*, 249–256.
5. Rahma, A. G., Yousef, K., & Abdelhamid, T. (2022). Blood flow CFD simulation on a cerebral artery of a stroke patient *SN. Applied Sciences, 4*, 261.
6. Rahma, A. G., & Abdelhamid, T. (2023). Hemodynamic and fluid flow analysis of a cerebral aneurysm: A CFD simulation *SN. Applied Sciences, 5*, 62.
7. Elhanafy, A., Guaily, A., & Elsaid, A. (2019). Numerical simulation of Oldroyd-B fluid with application to hemodynamics. *Advances in Mechanical Engineering, 11*, 1–7.
8. Karmonik, C., Yen, C., Grossman, R. G., Klucznik, R., & Benndorf, G. (2009). Intra-aneurysmal flow patterns and wall shear stresses calculated with computational flow dynamics in an anterior communicating artery aneurysm depend on knowledge of patient-specific inflow rates. *Acta Neurochirurgica, 151*, 479–485.
9. Department of Health Statistics and Informatics. (2021). World Health Organization Geneva. Health topics. Stroke, Cerebrovascular accident.
10. Luo, L., Shiu, W. S., Chen, R., & Cai, X. C. (2019). A nonlinear elimination preconditioned inexact Newton method for blood flow problems in human artery with stenosis. *Journal of Computational Physics, 399*, 108926.

11. Thompson, B. G., Brown, R. D., Amin-hanjani, S., Broderick, J. P., Cockroft, K. M., Connolly, E. S., Duckwiler, G. R., Harris, C. C., Howard, V. J., Johnston, S. C. C., Meyers, P. M., Molyneux, A., & Ogilvy, C. S. (2015). AHA/ASA Guideline Guidelines for the Management of Patients with Unruptured Intracranial Aneurysms. *Stroke, 46*(8), 2368–2400.
12. Schievink, W. I. (1997). Intracranial aneurysms. *The New England Journal of Medicine, 336*, 28–40.
13. Johnston, S. C., Higashida, R. T., Barrow, D. L., Caplan, L. R., Dion, J. E., Hademenos, G., Hopkins, L. N., Molyneux, A., Rosenwasser, R. H., Vinuela, F., & Wilson, C. B. (2002). Recommendations for the endovascular treatment of intracranial aneurysms a statement for healthcare professionals from the committee on cerebrovascular imaging of the american heart association council on cardiovascular radiology. *Stroke, 33*, 2536–2544.
14. Amenta, P. S., Yadla, S., Campbell, P. G., Maltenfort, M. G., Dey, S., Ghosh, S., Ali, M. S., Jallo, J. I., Tjoumakaris, S. I., Gonzalez, L. F., Dumont, A. S., Rosenwasser, R. H., & Jabbour, P. M. (2012). Analysis of nonmodifiable risk factors for intracranial aneurysm rupture in a large, retrospective cohort. *Neurosurgery, 70*, 693–701.
15. MacDonald, D. E., Najafi, M., Temor, L., & Steinman, D. A. (2022). Spectral Bandedness in high-Fidelity computational fluid dynamics predicts rupture status in intracranial aneurysms. *Journal of Biomechanical Engineering, 144*, 061004.
16. Souza, M. S., Souza, A., Carvalho, V., Teixeira, S., Fernandes, C. S., Lima, R., & Ribeiro, J. (2022). Fluid flow and structural numerical analysis of a cerebral aneurysm model. *Fluids, 7*, 100.
17. Nordgaard, H., Swillens, A., Nordhaug, D., Kirkeby-Garstad, I., Van Loo, D., Vitale, N., Segers, P., Haaverstad, R., & Lovstakken, L. (2010). Impact of competitive flow on wall shear stress in coronary surgery: Computational fluid dynamics of a LIMA-LAD model. *Cardiovascular Research, 88*, 512–519.
18. Lin, K. Y., Shih, T. C., Chou, S. H., Chen, Z. Y., Hsu, C. H., & Ho, C. Y. (2016). Computational fluid dynamics with application of different theoretical flow models for the evaluation of coronary artery stenosis on CT angiography: Comparison with invasive fractional flow reserve. *Biomedical Physics and Engineering Express, 2*, 1–11.
19. Shanmugavelayudam, S. K., Rubenstein, D. A., & Yin, W. (2010). Effect of geometrical assumptions on numerical modeling of coronary blood flow under normal and disease conditions. *Journal of Biomechanical Engineering, 132*, 61004.
20. Katritsis, D., Kaiktsis, L., Chaniotis, A., Pantos, J., Efstathopoulos, E. P., & Marmarelis, V. (2007). Wall shear stress: Theoretical considerations and methods of measurement. *Progress in Cardiovascular Diseases, 49*, 307–329.
21. Murayama, Y., Fujimura, S., Suzuki, T., & Takao, H. (2019). Computational fluid dynamics as a risk assessment tool for aneurysm rupture. *Neurosurgical Focus, 47*, E12.
22. Morris, P. D., Narracott, A., Von Tengg-Kobligk, H., Soto, D. A. S., Hsiao, S., Lungu, A., Evans, P., Bressloff, N. W., Lawford, P. V., Rodney Hose, D., & Gunn, J. P. (2016). Computational fluid dynamics modelling in cardiovascular medicine. *Heart, 102*, 18–28.
23. Wellnhofer, E., Osman, J., Kertzscher, U., Affeld, K., Fleck, E., & Goubergrits, L. (2010). Flow simulation studies in coronary arteries–Impact of side-branches. *Atherosclerosis, 213*, 475–481.
24. Boutsianis, E., Guala, M., Olgac, U., Wildermuth, S., Hoyer, K., Ventikos, Y., & Poulikakos, D. (2009). CFD and PTV steady flow investigation in an anatomically accurate abdominal aortic aneurysm. *Journal of Biomechanical Engineering, 131*, 1–15.
25. Ford, M. D., Nikolov, H. N., Milner, J. S., Lownie, S. P., DeMont, E. M., Kalata, W., Loth, F., Holdsworth, D. W., & Steinman, D. A. (2008). PIV-measured versus CFD-predicted flow dynamics in anatomically realistic cerebral aneurysm models. *Journal of Biomechanical Engineering, 130*, 1–9.

26. Hoi, Y., Woodward, S. H., Kim, M., Taulbee, D. B., & Meng, H. (2006). Validation of CFD simulations of cerebral aneurysms with implication of geometric variations. *Journal of Biomechanical Engineering, 128*, 844–851.
27. Perinajová, R., Juffermans, J. F., Mercado, J. L., Aben, J. P., Ledoux, L., Westenberg, J. J. M., Lamb, H. J., & Kenjereš, S. (2021). Assessment of turbulent blood flow and wall shear stress in aortic coarctation using image-based simulations. *Biomedical Engineering Online, 20*, 1–21.
28. Xiang, J., Tutino, V. M., Snyder, K. V., & Meng, H. (2014). CFD: Computational fluid dynamics or confounding factor dissemination? The role of hemodynamics in intracranial aneurysm rupture risk assessment. *American Journal of Neuroradiology, 35*, 1849–1857.
29. Chen, R., Wu, B., Cheng, Z., Shiu, W. S., Liu, J., Liu, L., Wang, Y., Wang, X., & Cai, X. C. (2020). A parallel non-nested two-level domain decomposition method for simulating blood flows in cerebral artery of stroke patient. *International Journal for Numerical Methods in Biomedical Engineering, 36*, 1–20.
30. Versteeg, H. K. (1955). *An introduction to computational fluid dynamics: The finite volume method*. Longman Scientific & Technical/Wiley.
31. Fluent, A. (2013). *Ansys fluent theory guide*. ANSYS Inc., USA 15317 724–746
32. Zouggari, L. (2018). The role of biomechanics in the assessment of carotid atherosclerosis severity: A numerical approach OPEN ACCESS *World*. *Journal of Vascular Surgery, 1*, 1007.
33. Usmani, A. Y., & Patel, S. (2018). Hemodynamics of a cerebral aneurysm under rest and exercise conditions. *International Journal of Energy for a Clean Environment, 19*, 119–136.
34. Yoshiki, K., Misaki, K., Nambu, I., Fukui, I., Mohri, M., Uchiyama, N., & Nakada, M. (2017). Intraoperative rupture of unruptured cerebral aneurysm during craniotomy: A case report. *Case Reports in Neurology, 9*, 261–266.

Index

A. M. Anter et al. (eds.), *Nature-Inspired Methods for Smart Healthcare Systems and Medical Data*, https://doi.org/10.1007/978-3-031-45952-8