scientific reports

OPEN

Enhancing remote patient monitoring with AI-driven IoMT and cloud computing technologies

Vijay Kumar Damera[1]✉, Ramesh Cheripelli[2], Narsaiah Putta[3], G. Sirisha[4] & Deepthi Kalavala[5]

The rapid advancement of the Internet of Medical Things (IoMT) has revolutionized remote healthcare monitoring, enabling real-time disease detection and patient care. This research introduces a novel AI-driven telemedicine framework that integrates IoMT, cloud computing, and wireless sensor networks for efficient healthcare monitoring. A key innovation of this study is the Transformer-based Self-Attention Model (TL-SAM), which enhances disease classification by replacing conventional convolutional layers with transformer layers. The proposed TL-SAM framework effectively extracts spatial and spectral features from patient health data, optimizing classification accuracy. Furthermore, the model employs an Improved Wild Horse Optimization with Levy Flight Algorithm (IWHOLFA) for hyperparameter tuning, enhancing its predictive performance. Real-time biosensor data is collected and transmitted to an IoMT cloud repository, where AI-driven analytics facilitate early disease diagnosis. Extensive experimentation on the UCI dataset demonstrates the superior accuracy of TL-SAM compared to conventional deep learning models, achieving an accuracy of 98.62%, precision of 97%, recall of 98%, and F1-score of 97%. The study highlights the effectiveness of AI-enhanced IoMT systems in reducing healthcare costs, improving early disease detection, and ensuring timely medical interventions. The proposed approach represents a significant advancement in smart healthcare, offering a scalable and efficient solution for remote patient monitoring and diagnosis.

Keywords IoMT, Cloud computing, Transformer-Based model, Disease prediction, Wireless sensor networks, AI in healthcare

The landscape of sensor-equipped, IoT-based embedded systems is rapidly evolving, presenting challenges in storage and processing power. Cloud computing emerges as a solution, offering extensive data storage and processing capabilities. This integration is crucial for managing advanced healthcare infrastructures. Continuous patient monitoring in healthcare organizations necessitates a smart framework for timely and cost-effective care, achieved through the synergy of artificial intelligence (AI) and IoT-cloud architecture. The connectivity of mobile devices to sensors in smart healthcare environments facilitates swift responses from medical personnel, especially critical for intensive care scenarios. Smart healthcare systems leverage modern, low-cost IoT sensors tracking various vital indicators, with real-time data analysis facilitated by AI and deep learning.

The processing demands of diverse IoT sensors in a smart healthcare system underscore the need for cognitive capabilities, leading researchers to explore intelligent IoT systems. In this context, cloud computing and IoT-based devices play a pivotal role in monitoring patient records, ensuring efficient and cost-effective healthcare solutions. Information technology (IT) has significantly impacted affluent countries' medical industries, enhancing patient care through the integration of smart devices and wearable sensors. IoMT's remote patient monitoring proves to be a time-efficient and cost-effective solution, addressing the challenges posed by the increasing prevalence of chronic illnesses, particularly in developing nations. Swift action becomes paramount in the face of growing healthcare demands[1–11].

The integration of Artificial Intelligence (AI) with the Internet of Medical Things (IoMT) and cloud computing has redefined remote patient monitoring by enabling real-time data collection, analysis, and disease prediction[12–14]. While conventional deep learning models such as CNNs have been widely adopted, they often fall short in capturing long-range dependencies and complex patterns inherent in medical datasets. Moreover, most existing frameworks lack a unified approach that optimally combines spatial and spectral features of health

[1]Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Hyderabad Campus, Hyderabad, India. [2]Department of Information Technology, Vidya Jyothi Institute of Technology, Hyderabad, India. [3]Department of Computer Science and Engineering, Vasavi College of Engineering, Hyderabad, India. [4]Department of MCA, Loyola Academy, Degree and PG College, Hyderabad, India. [5]Stellantis NV, Financial District, Nanakram Guda, Hyderabad, India. ✉email: dameravijaykumar@gmail.com

data for accurate disease classification[15]. To address these challenges, this study introduces a transformer-based self-attention model that leverages AI to enhance feature representation in remote healthcare applications[16,17]. This approach aims to bridge the technical gap by improving classification accuracy while maintaining the scalability and efficiency required for real-time monitoring in IoMT-cloud environments.

The field of computer science known as "artificial intelligence" (AI) is responsible for giving robots the ability to learn and make decisions on their own. Intelligent gadgets use AI to connect seamlessly and perform ingeniously. In order to make sense of the mountain of data sent and received by smart devices, AI methods are used[18]. Additionally, the AI process suggests ways in which the systems might be optimised for maximum efficiency. Artificial intelligence (AI) and machine learning (ML) provide invaluable tools for tackling the complex issues of the modern day. Significant dynamic challenges are being tackled in every industry by computational systems created using AI and ML[19,20]. Integrating IoT, AI, and ML has the potential to revolutionise how we work, play, and communicate. Large datasets are collected with the help of smart diagnosis of various illnesses, and AI and ML algorithms are used to uncover previously unseen patterns. Working with massive datasets and making suggestions for illness prevention in the future is now much simpler for medical professionals, thanks to AI-based solutions[21]. When it comes to aiding in medical diagnoses, the artificial intelligence of medical staff (AIoMT) integrates AI methods with health diagnosis methods. The goal of AIoMT is to reduce the number of patients who need to spend time in the hospital as well as the costs associated with doing so. Traditional techniques of medical diagnosis are needed to meet the rising community's needs as the population increases. Finding methods for effective management of scarce resources is crucial in light of the world's expanding population.

Following COVID-19, a real-time, remote healthcare diagnosis system is essential. Academic institutions and private businesses must work together to find solutions to these problems if healthcare facilities are to be enhanced. There have been recent attempts to build a smart bed and smart point-of-care (PoC) devices, but more work is required to keep up with population growth. The World Health Organisation (WHO) reports that heart-related illnesses account for 17.9 million deaths annually[22]. The medical community faces a significant obstacle with heart disease. Atherosclerosis is the root cause of cardiovascular disease. Due to a lack of readily available emergency services, heart failure is the leading cause of mortality in many developing nations. Although elderly individuals make up the vast majority of heart disease sufferers, anyone can be diagnosed with this condition at any age. Heart disease accounts for one in every nine fatalities in the United States[23]. The most common symptoms of this condition are chest discomfort and exhaustion. However, it can also occur in the absence of these signs. An activity-tracking remote health monitoring device for cardiac patients is urgently required. Remote health monitoring systems may need immediate action in the event of an emergency. The medical field has greatly benefited from IoMT, especially in remote places where it might be difficult to give prompt care. Many lives can be spared if diseases are diagnosed and treated early on. When compared to traditional methods of diagnosis, remote health monitoring saves money since it eliminates the need for costly in-person visits.

Predictive analytics is demonstrating its effectiveness not just in hospitals but even in patients' own homes through remote monitoring and preventing the need for emergency care. All aspects of a patient's care, from diagnosis to prognosis to therapy, benefit from the use of predictive analytics[24]. Long-term problems, such as persistent high blood pressure, put many people at risk for cardiovascular disease. As the global population ages, more and more people are being diagnosed with long-term cardiac problems. This calls for round-the-clock monitoring of those receiving care at home as well as hospitalised patients, allowing for an immediate response to changes in vital signs. Conventional techniques are time-consuming and tiresome, but keeping track of an elderly person's health over the long term can save hospitalisation costs and improve quality of life. This calls for cost-effective infrastructure that reduces the burden on doctors and nurses while keeping patients closely monitored[25]. Because of the ubiquitous nature of the Internet of Things, a plethora of smart, networked gadgets and wearables with sensors have proliferated, allowing for remote patient monitoring of cardiac conditions. Medical sensors are included in smart health watches, blood pressure monitors, and electrocardiogram (ECG) monitors, which are all part of the IoT for healthcare monitoring. For precise cardiac risk assessments, the healthcare IoT gathers crucial patient data and uploads it to the cloud, where it is stored and subjected to complicated deep learning algorithms in conjunction with historical electronic clinical records. These Internet of Things devices can quickly communicate the patient's status to doctors and nurses. By evaluating patients' risk of acquiring a specific cardiac disease, their prognosis for the given illness, and the accompanying therapy, doctors may better make timely decisions for individuals and the community as a whole[26].

Improvements in attention processes have led to more accurate illness categorization. CNN is not great at simulating long-distance relationships and getting global context information[27], but the bulk of existing attention methods for classification are based on the convolution layer, so there is room for improvement in classification accuracy. The transformer model, on the other hand, treats patches, which allows it to make greater use of global context information over a wide range. Transformers, which are based on a self-attention mechanism, were first introduced for machine translation but have since become the gold standard for many NLP tasks. Several efforts, inspired by NLP's achievements, attempt to merge CNN-like structures with introspective processing. In this research, we are inspired by the most recent self-attention apparatus in NLP to propose a deep transformer for categorization. In particular, we investigate deep transformers in both the spectral and spatial dimensions. Then, a TL-SAM is proposed to boost cardiac disease classification accuracy.

Key contributions of this study are as follows:

- *Development of an AI-driven IoMT telemedicine framework* This study introduces a novel framework integrating IoMT, cloud computing, and wireless sensor networks to enable real-time patient monitoring and remote healthcare services.

- *Proposal of the TL-SAM model for enhanced disease classification* A Transformer-based Self-Attention Model (TL-SAM) is developed, replacing conventional convolutional layers with transformer layers to improve the classification of medical conditions using spectral and spatial feature extraction.
- *Optimization using improved wild horse optimization with levy flight algorithm (IWHOLFA)* The proposed model is fine-tuned using IWHOLFA, enhancing hyperparameter selection and improving classification accuracy compared to existing deep learning techniques.
- *Extensive performance evaluation on the UCI dataset* The effectiveness of the proposed approach is validated through rigorous experimentation, achieving superior classification accuracy (98.62%), precision (97%), recall (98%), and F1-score (97%), demonstrating its potential for real-world healthcare applications.

The remainder of this manuscript is structured as follows: Sect. 2 presents a comprehensive review of related works, highlighting existing methodologies in IoMT-based healthcare monitoring and disease classification. Section 3 details the materials and data collection processes, including sensor-based patient monitoring and cloud-based data storage. Section 4 introduces the proposed Transformer-based Self-Attention Model (TL-SAM) and its integration with the Improved Wild Horse Optimization with Levy Flight Algorithm (IWHOLFA) for optimized classification performance. Section 5 describes the experimental setup, evaluation metrics, and comparative analysis of TL-SAM with existing deep learning models. Section 6 discusses the results, emphasizing the model's effectiveness in real-world healthcare applications. Finally, Sect. 7 concludes the paper with key findings, potential limitations, and future research directions.

Related work

Golec et al.[26] propose a smart healthcare system called HealthFaaS that employs AI, the IoT, and serverless computing to reduce deaths and the costs associated with misdiagnoses. The HealthFaaS framework gathers health data from consumers via IoT devices and transmits it to AI models installed on a serverless computing environment based on Google Cloud Platform (GCP) because of its advantages such as dynamic scalability, reduced operational complexity, and a pay-as-you-go pricing model. Five different AI models for predicting heart disease risk are compared and contrasted based on their AUC. The experimental consequences establish that the LightGBM model accurately diagnoses heart abnormalities 91.80% of the time. When it comes to Quality of Service (QoS) metrics like throughput and latency, we have also compared the HealthFaaS framework's performance to that of a non-serverless platform as the number of users increases. We have also utilised a serverless platform to measure cold start latency and have discovered that it is proportional to RAM size and programming language.

A model suggested by Kadhim et al.[27] seeks to determine the most effective machine learning technique for early detection of heart attacks. The model is trained and evaluated using machine learning principles, using data from the patient to make accurate inferences. Patient data is collected and processed first, then with the Random Forest algorithm, and finally, the classification results are optimised with one of the hyperparameter optimisation techniques.

Using a multi-layered artificial neural network and a backpropagation method, Kaur et al.[28] have created a model for the efficient, cum-automated detection of coronary heart disease. The proposed model evaluates the impact of varying the number of hidden layer neurons across various transfer functions. Kaggle and Statlog's heart disease datasets, which include 13 clinical factors, have been used to develop the model. Using six hidden layers with a tan-hyperbolic transfer function, the experimental findings achieved 99.92% accuracy. Statistical parameters and k-fold cross-validation have been used to back up the findings.

Using machine learning, Kumar et al.[29] present a new method for predicting the likelihood of cardiovascular illness in people with type 2 diabetes. Here, type 2 diabetes-related input data has been cleaned up and reduced in dimension. The characteristics needed for the anomaly detection of type 2 diabetes analysis have been retrieved from the processed data. Classifying the retrieved data using a neural network, such as the VGG-16 Net_gradient, reveals abnormalities indicative of type 2 diabetes and allows for the prediction of cardiovascular disease. The experimental analysis for several diabetic datasets has been done in terms of MAP. The projected method achieved 96% accuracy, 67% precision, 79% recall, 63% F-1 score, 66% RMSE, and 68% MAP.

A new method for automating human cardiac diseases has been offered by Mohapatra et al.[30] to give medical practitioners quicker analysis by decreasing the time it takes to establish a diagnosis and improving the quality of the results. The prediction accuracy of machine learning algorithms may frequently be enhanced by mining electronic health records (EHRs) for interesting data trends. Machine learning, in particular, has made important contributions to the resolution of problems like making accurate predictions in fields as diverse as healthcare. With so much collected healthcare data, it's crucial that we find ways to put it to use for the benefit of humanity. Over the years, researchers have developed a number of prediction models and methods to aid cardiologists and other medical professionals in analysing data and drawing useful conclusions. In this study, we present a predictive model for cardiac illness prediction based on the hierarchical stacking of many classifiers at the base and meta levels. Using a wide range of heterogeneous learners in tandem, powerful model results may be achieved. Prediction accuracy for the model was 92%, with a 92.6% point precision score, a 92.6% point sensitivity, and a 91% specificity. Accuracy, precision, recall, F1-scores, and area under the ROC curve were only a few of the measures used to assess the model's efficacy.

A technique of k-mode clustering using Huang beginning has been proposed to enhance classification accuracy by Bhatt et al.[31]. Used models include XGBoost (XGB), multilayer perceptron (MP), random forest (RF), and decision tree classifier (DT). The optimal outcome was achieved by hypertuning the parameters of the applied model with GridSearchCV. The Kaggle dataset contains 70k examples taken from the actual world. It was determined from this primary research that the most accurate method is the multilayer perceptron with cross-validation. The maximum accuracy it could muster was 87.28%.

An ML-based system for early cardiac disease prediction has been developed by Chaudhari et al.[32]. We developed and used the hybrid ensemble method for cardiac disease forecasting. XGBoost and multilayer perceptron classifiers are only a few of the supervised ML techniques that form the basis of the suggested hybrid ensemble approach. Feature selection and hyperparameter adjustment are two of ML's most important building blocks. Recursive feature elimination (RFE) is used to boost speed since unwanted features in the dataset might hinder accuracy. In addition, the grid search technique is used to fine-tune the hyperparameters. Precision, recall, f-measure, and accuracy are measured and compared to the performance of traditional and state-of-the-art replicas to determine how well the proposed method performs. Maximum accuracy of 93.15% was attained by the optimised hybrid ensemble approach with suitable feature selection on UCI datasets.

Deep Gated Recurrent Units (D-GRU) combined with stacked auto encoders is a novel model for cardiovascular disease prediction proposed by Karthikeyan et al.[33]. To achieve this goal, we present a unique optimisation technique called the Logistic Chaos Honey Badger technique. The trainees were trained using publicly available datasets relevant to heart disease obtained from the UCI Repository, Cleveland Database. The sensors in the IoMT architecture collect data that is used to further evaluate the trained model. Several deep learning models and prior research in the literature are used to assess the projected model. With an accuracy of 95.15%, the suggested D-GRU model is superior to the other models used as comparisons for predicting cardiac illnesses.

For cardiac evaluation, Jain et al.[34] have developed the Levy Flight-Convolutional Neural Network (LV-CNN) based on the diagnostic scheme's use of the heart disease picture data set. In order to simplify the system's computations, the incoming big data photos are initially scaled. The suggested LV-CNN model is then applied to the resized pictures. As a result, to lessen the impact of the loss function on the CNN design, the LV strategy is combined with the Sunflower Optimisation Algorithm (SFO). This combination protects the SFO algorithm from becoming stuck in local minima caused by the levy flight's unpredictable nature. Accuracy of 95.74%, specificity of 0.96%, error rate of 0.35%, and runtime of 9.71 s are expected from simulated and experimental testing of the proposed method in the MATLAB programming environment. The consequences of this comparison study demonstrated the superiority of the proposed paradigm.

A cloud-based, remote ECG monitoring system was presented by Raheja and Manocha[35], which has the potential to save lives in rural areas, especially during the Golden Hour. The ECG signal is first preprocessed using a Savitzky-Golay (SG) filter to get rid of baseline wanders, then the high-noise is filtered out using a packet transform (MOWPT), and finally the characteristic spots of the ECG are shown on the graph. The data is then classified using a CNN that has been trained using deep learning. Security is provided by the triple data encryption standard, while authentication is achieved by the water cycle optimisation approach. The ThingSpeak platform is used to build IoT, and this, in turn, makes it possible to see and evaluate ECG data in the cloud. The cardiologist receives encrypted and verified ECG data via ThingSpeak for analysis. The suggested categorization algorithm is evaluated on the MIT-BIH database's standard annotated dataset, and its efficacy is measured in a number of ways. With an accuracy of 99.12%, the proposed deep learning CNN classifier detects arrhythmias. The model's sensitivity to abnormal cardiac rhythms is also 100%, and its specificity is 99.9%.

In order to evaluate the efficacy of deep learning methods in comparison to more conventional approaches, Dileep et al.[36] employed a UCI dataset and a real-time dataset. Cluster-based bidirectional approaches have been proposed to enhance the precision of conventional approaches. Experimental results are based on the UCI datasets, both of which were fed via the K-Means clustering method to eliminate duplicates before being sent into the C-BiLSTM technique to forecast heart illness. Accuracy, sensitivity, and F1 score are used to illustrate C-BiLSTM's efficacy compared to traditional classifier methods, including regression tree, SVM, logistic unit, and ensemble. When compared to six standard approaches, the consequences demonstrate that the C-BiLSTM provides the best prediction of cardiac illness, with 92.84% on the UCI dataset and 92.84% on the real-time dataset.

Using artificial neural networks trained with scaled conjugate gradient backpropagation and K-fold cross-validation, Paul and Karn[37] present a more accurate technique for forecasting the presence of heart disease. The IEEE Data Port and the UCI Machine Learning Repository have both been mined for cardiac datasets. The proposed system estimates the presence or absence of testing using a minimum of 63.3803% and a maximum of 100% accuracy using 13 input attributes and the Cleveland processed heart dataset and 11 input attributes and the dataset, respectively.

Heart disease prediction using blockchain data was offered by Hasanova et al.[38], who used machine learning to suggest the Sine technique. Since the information in a blockchain cannot be altered without detection, it may be relied upon as a reliable resource for educational purposes and as a safe repository for sensitive patient data. Accuracy and root mean square error are measured and compared to see how well the proposed SCA_WKNN performs in contrast to competing algorithms. Based on our findings, SCA_WKNN outdoes W K-NN with a maximum accuracy of 4.59% and 15.61%, respectively. Additionally, the latency and throughput of blockchain-based storage compared to peer-to-peer storage are evaluated. The maximum throughput of decentralised storage is 25.03% higher than that of peer-to-peer storage.

In order to forecast heart disease and analyse a person's current health, Shanthi et al.[39] used an Internet of Things framework to collect a person's heartbeat, blood pressure, and temperature via the sensor implanted in the human body and then applied a learning procedure to the collected data. The acquired patient data is used in conjunction with the Cleveland heart disease data set to make predictions about the occurrence of heart disease. The Naive Bayes Algorithm, Decision Trees, the K-Nearest Neighbour Algorithm, and Logistic Regression are compared and contrasted. The proposed approach integrates patient remote health monitoring as well to help diagnose whether or not a patient is suffering from a specific disease. The suggested approach is highly efficient, and it aids in the early prediction of cardiac difficulties.

Here, Munagala et al.[40] introduce a fuzzy-long model for an original Internet of Things-enabled approach to cardiac illness forecasting. The project used publicly available data acquired through wearable IoT devices as its baseline. The best characteristics are chosen using an enhanced version of Harris Hawks optimisation (HHO), whose goal function is to maximise correlations between instances of the same class and to minimise those between instances of diverse classes. Presented below are the scientific contributions of the healthcare monitoring system, which have the potential to reduce the global mortality toll from heart disease. part of this long-term healthcare method. The simulation consequences demonstrated that the projected method is superior to current methods in predicting cardiovascular disease. The summary of the literature review is shown in Table 1.

Research gap

Although significant progress has been made in applying machine learning and deep learning techniques to heart disease prediction, most existing models rely on either convolutional or traditional neural network architectures, which limit their ability to capture global dependencies across spatial and spectral dimensions. Additionally, few studies incorporate advanced optimization techniques for hyperparameter tuning, resulting in suboptimal model configurations. Furthermore, many models lack integration with cloud-based IoMT environments for real-time and scalable healthcare monitoring. Therefore, there remains a gap in developing a unified, attention-based model that jointly exploits spatial-spectral features, is robustly optimized, and aligns with real-world IoMT-cloud healthcare applications. The proposed TL-SAM framework addresses these limitations through modular self-attention, IWHOLFA optimization, and real-time cloud integration.

Materials

The healthcare system has a huge problem in terms of processing, storing, and handling the vast volume of data created by IoT strategies. In this section, we outline the broad strokes of the strategy's execution. The cloud system connects the medical centre with the people who go there, the patients. Devices can transmit updates on a patient's health state and dangers in real-time with the use of a Wireless Sensor Network (WSN) via a cloud network. A physician's saved data in the cloud will receive an alert message if the patient's state is deemed unsatisfactory in comparison to the readings supplied in a lookup table. Online diagnostics, data searching, and patient data uploading are just a few examples of the many server- and client-side tasks made possible by the cloud network. The server side of a website uses cloud storage to keep everything in order, while the client side makes use of intelligent terminals like smartphones and medical equipment. In this setup, the patient might use

Authors and references	Objective	Methodology	Key Findings
Golec et al.[26]	Develop HealthFaaS, a smart healthcare system using AI, IoT, and serverless computing	AI models on Google Cloud Platform (GCP) for heart disease risk prediction	LightGBM model achieves 91.80% accuracy; Performance compared with non-serverless platform
Kadhim et al.[27]	Determine the most effective machine learning technique for early detection of heart attacks	Machine learning principles with Random Forest algorithm	Optimized classification results with hyperparameter optimization
Kaur et al.[28]	Create an efficient model for automated detection of coronary heart disease	Multi-layered artificial neural network with tan-hyperbolic transfer function	Achieved 99.92% accuracy using Kaggle and Statlog datasets
Kumar et al.[29]	Predict likelihood of cardiovascular illness in people with type 2 diabetes	Neural network (VGG-16 Net_gradient) for anomaly detection	Method achieved 96% accuracy, 67% precision, and 79% recall
Mohapatra et al.[30]	Offer a new method for automating human cardiac diseases	Predictive model based on hierarchical stacking of classifiers	Model achieved 92% accuracy, 92.6% precision, 92.6% sensitivity, and 91% specificity
Bhatt et al.[31]	Propose k-mode clustering using Huang beginning to enhance classification accuracy	XGBoost, multilayer perceptron, random forest, and decision tree classifier	Achieved maximum accuracy of 87.28% with multilayer perceptron
Chaudhari et al.[32]	Develop an ML-based system for early cardiac disease prediction	Hybrid ensemble method using XGBoost and multilayer perceptron	Achieved maximum accuracy of 93.15% with suitable feature selection
Karthikeyan et al.[33]	Propose D-GRU combined with stacked auto encoders for cardiovascular disease prediction	Logistic Chaos Honey Badger technique for optimization	D-GRU model outperformed other models with 95.15% accuracy
Jain et al.[34]	Develop Levy Flight-Convolutional Neural Network (LV-CNN) for cardiac evaluation	LV-CNN model combined with Sunflower Optimisation Algorithm (SFO)	Achieved accuracy of 95.74%, specificity of 0.96%, and error rate of 0.35%
Raheja and Manocha[35]	Present a cloud-based, remote ECG monitoring system	CNN trained with deep learning on ThingSpeak platform	Achieved 99.12% accuracy in detecting arrhythmias
Dileep et al.[36]	Evaluate the efficacy of deep learning methods in cardiac disease prediction	Cluster-based bidirectional approaches with C-BiLSTM	C-BiLSTM provided the best prediction with 92.84% accuracy
Paul and Karn[37]	Forecast the presence of heart disease using artificial neural networks	Scaled conjugate gradient backpropagation with K-fold cross-validation	Achieved accuracy between 63.38% and 100% using various datasets
Hasanova et al.[38]	Predict heart disease using blockchain data	Sine technique based on machine learning	Outperformed W K-NN with maximum accuracy of 4.59%
Shanthi et al.[39]	Use an IoT framework for early prediction of heart disease	Machine learning algorithms including Naive Bayes, Decision Trees, K-Nearest Neighbour, and Logistic Regression	Proposed approach aids in early prediction of cardiac difficulties
Munagala et al.[40]	Introduce a fuzzy-long model for IoT-enabled cardiac illness forecasting	Enhanced Harris Hawks optimization for feature selection	Superiority demonstrated in predicting cardiovascular disease

Table 1. Summary of literature review.

a website to inquire about their health condition at a specific time. Therefore, the client is coupled with internet technology for better assessment and medical data valuation.

Data collection unit
Sensors are used to fold data from the body. Intelligent Health Monitoring System (IHMS) is a network that guarantees communication between different entities. In a network node, there are two distinct communication ranges: short and medium. The Body Area Network (BAN) and nearby objects both make use of short-range protocol technologies, while medium-range technology supports the BAN's connection with a base station. Short-range communication methods are the norm in the healthcare IoMT industry. The data-collecting device accepts readings from a wide variety of sensors and transmits them via a variety of connection protocols, including Wi-Fi. In most cases, Wi-Fi networks provide a fast and simple method of communication within healthcare facilities. Wi-Fi's interoperability with Android phones means it can be used for a variety of purposes, and it also improves security and management. All of the collected data may be found in IoMT's cloud repository.

Collection discovery unit
The patient detection system is set up to confirm any abnormal or unusual occurrences. Essential indicators such as temperature, systolic are monitored and recorded in a patient database, and a set C of p patients is assigned a representation as a result of this process. $C = [C_1, C_2, C_3, \ldots, C_p]$ where every to set C. Then it uses biosensors for assembling the set of vital signs as $S = [RR, HT, BT, OS, SBP, \ldots]$ of the patient data. Let us accept as B_s^C for patient C and monitors identical value foretold by bio-sensors is signified as $R_s^C = [f_1, f_2, f_3, \ldots, f_p]$ in the vector.

Data acquisition and pre-processing

- The suggested model takes in data from a variety of sources, such as sensor readings, medical records, and network information. sensor gap between the patient and the outside environment in remote patient monitoring systems. Concealable torso Sensors can compile data, among other symptoms and health indicators. Wireless standards ensure the privacy and security of these transmissions.
- Inconsistencies, noise, missing values, different sizes and formats, and a larger dimensionality are commonplace in this type of data. This necessitates pre-processing the received input due to the lower-quality output it will create. As a result, sensor data and medical records undergo preliminary processing.
- Analysis of raw sensor data: In addition to the patient's body mass, sensors also examine the normal and abnormal circumstances of characteristics like blood pressure, blood sugar, heart rate, etc. In the first stage, the dataset is converted to CSV files so that it can be read more easily. The attributes are given names, and their numerical values are shown. In the second stage, sensor data is given names. The final patient dataset is delivered to the cloud processor with the help of these characteristics.
- Input Data for Preprocessing: Our method employs a Kalman filter to remove noise and inconsistencies from the sensor data and an unsupervised filter to fill in missing values. The greatest variance attribute is obtained by using an unsupervised filter to remove unused characteristics. To facilitate further processing and categorization, the acquired numerical values are normalised using the normalisation method, which restricts the range from zero to one. The dataset is split into n files using EmEditor and then sent to a cloud-based processor.

Preparation of dataset
The dataset used in this study is the publicly available Heart Failure Clinical Records dataset from the UCI Machine Learning Repository[41]. It contains 299 records of patients with heart failure, originally collected as part of a medical study published in *PLoS ONE*[41]. The dataset includes 13 features, such as age, ejection fraction, serum creatinine, blood pressure, and diabetes status, with a binary target variable indicating whether the patient died during follow-up (1 = death, 0 = survival). Among the records, 194 belong to male patients and 105 to female patients, all over the age of 40. The dataset is imbalanced with a death rate of approximately 32%.

Prior to model training, several preprocessing steps were applied: (i) missing values were handled using unsupervised imputation; (ii) numerical attributes were normalized using min-max scaling to the range [0, 1]; (iii) outliers and noise were filtered using a Kalman filter; and (iv) the dataset was converted into structured CSV format for compatibility with deep learning frameworks. The data was then randomly split into training (70%) and testing (30%) subsets, ensuring class stratification. Table 2 provides information about the dataset[42].

Proposed framework
Medical professionals are given access to data organised as a patient's diagnostic information so that it may be thoroughly analysed. Smart gadgets are commonplace in the medical community. These gadgets have the ability to store data locally and conduct basic computations. Professionals in the medical field then follow up with heart failure patients as needed. Algorithm 1 outlines the process followed by the suggested IoT framework.

Description	Attributes	Range	Measured In
Follow-up period	Time	4–285	Days
Platelets in blood	Platelets	25.01–850.00	Kilo platelets/mL
Level of CPK enzyme in the blood	CPK	23-7861	Mcg/L
If the patient died in the follow-up period	Event (target)	0,1	Boolean
Man or woman	Gender	0,1	Binary
If the patient smokes	Smoking	0,1	Boolean
If the patient has diabetics	Diabetics	0,1	Boolean
If the patient has a BP issue	B.P	0,1	Boolean
Reduction in red blood cell or haemoglobin	Anaemia	0,1	Boolean
Stage of the patient	Age	40–95	Years
fraction Proportion of blood leaving the heart at each attentiveness	Ejection	14–80	Percentage
Side by side of sodium in the blood	Sodium	114–148	mEq/L
Equal of creatinine in the blood	Creatinine	0.50–9.40	mg/dL

Table 2. Dataset stipulations.

Input: Raw physiological sensor data from IoMT devices

Output: Predicted disease label

1. Begin
2. Acquire real-time patient data from IoMT sensors
3. Perform data preprocessing:
   a. Apply Kalman filter to reduce noise
   b. Normalize attributes to [0, 1] range
   c. Handle missing values via unsupervised imputation
4. Transform input data into spectral and spatial feature blocks
5. Apply SpecSAM:
   a. Embed pixel-wise spectral features
   b. Pass through stacked transformer layers
   c. Extract 128-dimensional spectral feature vector
6. Apply SpatSAM:
   a. Convert image blocks into patches
   b. Embed spatial features via convolution + transformer
   c. Extract 128-dimensional spatial feature vector
7. Fuse spectral and spatial features (Concatenation → 256-D vector)
8. Optimize model parameters using IWHOLFA
9. Feed fused features to MLP classifier
10. Predict disease outcome
11. End

Algorithm 1. TL-SAM-based disease prediction framework.

The research presented here creates a new categorization scheme for categorising diseases. Spectral self-attention models (SpecSAMs) and spatial self-attention models (SpatSAMs) are the two main components of the proposed transformer-based system. Both the SpecSAM and the SpatSAM are taught to focus on characteristics along the spectral dimension and the spatial dimension correspondingly. The suggested framework is able to make better use of spectral and spatial information by fusing the characteristics retrieved by SpecSAM and SpatSAM as input to the classifier.

Transformer

In contrast to the convolution layers that are often used for illness classification, the transformer layer can generate feature representations on its own using a self-attention model. This makes the feature representations more accurate and stable. By layering many transformers on top of one another, a trainable deep-learning model may be constructed. Specifically, there are three components that may be used to define this deep model manually: self-focus, feedforward network, and positional encoding.

Positional encoding

In contrast to convolutional and recurrent neural network models, transformer layers do not use either. Therefore, positional encoding is implemented to allow the transformer-layers-based model to exploit sequence and position information. A positional vector of the same size as the feature vector can be generated using the positional encoding method. Transformers do not inherently encode positional information, which is essential in sequence modeling. To remedy this, positional encoding vectors are added to the input embeddings. These encodings are computed using sine and cosine functions to preserve information about the order of input elements. The position of each element is represented as follows:

$$PE\left(pos, 2i\right) = sin\left(\frac{pos}{10000^{2i}/d_{model}}\right). \tag{1}$$

$$PE\left(pos, 2i+1\right) = cos\left(\frac{pos}{10000^{(2i+1)}/d_{model}}\right) \tag{2}$$

where pos denotes the site index of the vector. Let the distance of feature vector be L, then the range of $posis0,1,\ldots,L-1$. dmodel is the eye dimension, and $i=0,1,\ldots,dmodel/2$ symbolises dimension.

Following positional encoding, the initial feature vectors are modified to include information about the location, either relative to the origin or in absolute terms. Note that the positional coding operation is only carried out before the layer and that the input.

Self-attention

The transformer layers' self-attention mechanism is the most important component. It is possible to think of the self-attention mechanism as a function that takes in a query vector and a list of key-value pairs and returns a third value. The result of the self-attention process is a weighted sum of vectors, where the weights are determined by a comparison between the query and key vectors. In a self-attention mechanism, each input token interacts with every other token to compute contextual representations. This interaction is quantified using scaled dot-product attention, which computes attention scores by comparing a query with key-value pairs. The formulation is given by:

$$Z = Attention\left(Q, K, V\right) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V. \tag{3}$$

In Eq. (3), $Q=XQ^w, K=XK^w, V=XV^w$ are query matrix, key matrix, correspondingly. X is the feature vector arrangement. dk is the dimension of the feature vector.

In addition, the presentation of the self-attention model is enhanced by the introduction of the multiheaded attention (MHA) mechanism. In a transformer layer, for example, Eq. (4) is used to create various output features in addition to multiple and value matrices. This means that the model might potentially extract richer feature representations from a single input vector by mapping it to numerous output vectors. The final output vectors are obtained by multiplying many input vectors by a matrix parameter and then concatenating the results. The formally defined mechanism of multi-head attention is as follows:

$$\begin{cases} MHA\left(X\right) = Concat\left(z_1, z_2, \ldots, z_p\right)W^O \\ z_i = Attention\left(XQ_i^W, XK_i^W, XV_i^W\right) \end{cases} \tag{4}$$

where W^O is a learned parameter matrix, XQ_i^W, XK_i^W, XV_i^W are the ith matrix parameter. The MHA mechanism in transformer layers has several heads, each of which can be trained to pay attention to various things in the same way that diverse convolution kernels may extract different types of features.

The process for self-attention involves a mapping from one series of feature vectors, X, to another, Z, where info about the pixel, as well as the connection between words or pixels, is stored.

Feed-forward network

A transformer layer's final component is a network. The feed-forward network in the proposed technique is constructed using two completely linked layers. The properties of a completely linked layer are determined by:

$$r_i = max\left(W_j z_i + b_j, 0\right), j = 1,2 \tag{5}$$

.

the index of the feature vector is I if… Note that all features in the appropriate transformer layer share the feed-forward network's settings. The model's trainability is enhanced by introducing residual connections between its self-attention and feed-forward network layers, and the features retrieved at each layer are used to their maximum potential.

Deep spatial-spectral transformer

In the context of IoMT-based healthcare data, spectral and spatial features represent fundamentally different dimensions of information. Spectral features often reflect temporal patterns or frequency-based signals (e.g., heart rate variability), whereas spatial features may correspond to the arrangement or interdependence of signals from multiple sensors. Attempting to learn both using a unified transformer structure tends to blur their distinct contributions, leading to suboptimal attention distributions. Therefore, in this study, we propose a dual-branch architecture comprising SpecSAM and SpatSAM. Each branch is designed to independently extract deep semantic information from its respective dimension. This separation allows each self-attention mechanism to focus specifically on the nuances of spectral or spatial correlations. Experimental evidence suggests that this modular modeling strategy outperforms unified attention mechanisms in terms of both accuracy and interpretability. Transformer can learn more abstract information than a convolutional neural network and requires fewer training parameters thanks to its built-in self-attention mechanism. As a result, the results of classification and identification tasks could improve. There is a wealth of geographical information available in addition to spectral information from the data. As a result, we employ a transformer model that incorporates both spectral and spatial dimensions. The study's suggested framework is labelled TL-SAM in Fig. 1.

Pixel embedding

The input for a transformer is a series of features. Therefore, the data cube must be converted into a series of features. We refer to this transformation as "pixel embedding" since each heart disease data/input picture is represented by a "pixel." A SpecSAM plus a SpatSAM make up the TL-SAM. The research employs a convolution layer to collapse the spectral dimension, turning the picture blocks from distinct bands into a single feature vector. The number of bands is proportional to the length of the feature sequence. The feature sequence would then be sent into the SpecSAM, which would be trained to focus on the spectral characteristics. The SpecSAM stands out since it is a transformer-stacked model. The picture block (the chosen three bands) is split into 16 patches of equal processing. In particular, vertical and horizontal lines are used to separate patches. A convolution layer is performed to convert the 16 patch vectors, analogous to the spectral dimension. The number of the convolution kernel determines the size of the feature vectors. This process yields the spatial feature vector sequence, which is fed into a second stacked transformer model (SpatSAM).

Although both SpecSAM and SpatSAM take picture blocks as input for pixel embedding, they prioritise distinct attributes. All feature vector arrangements in SpecSAM are organised by band, and picture blocks from each band are transformed into feature vectors. Spatial order is preserved in the resulting feature sequences because to SpatSAM's practice of partitioning 3-PC picture blocks into patches before pixel embedding. The suggested model further improves SpatSAM's suitability for learning spatial properties, whereas SpecSAM's input is comprehensively spectral.

Feature fusion

Important spectral and spatial characteristics are extracted by the SpecSAM and SpatSAM, respectively. A multilayer with two completely linked layers receives the extracted spectral and spatial info as input. The classification performance may be further enhanced by fusing the spectral and spatial characteristics in TL-SAM. Three feature fusion strategies were evaluated in this paper. The first approach involves joining the spectral

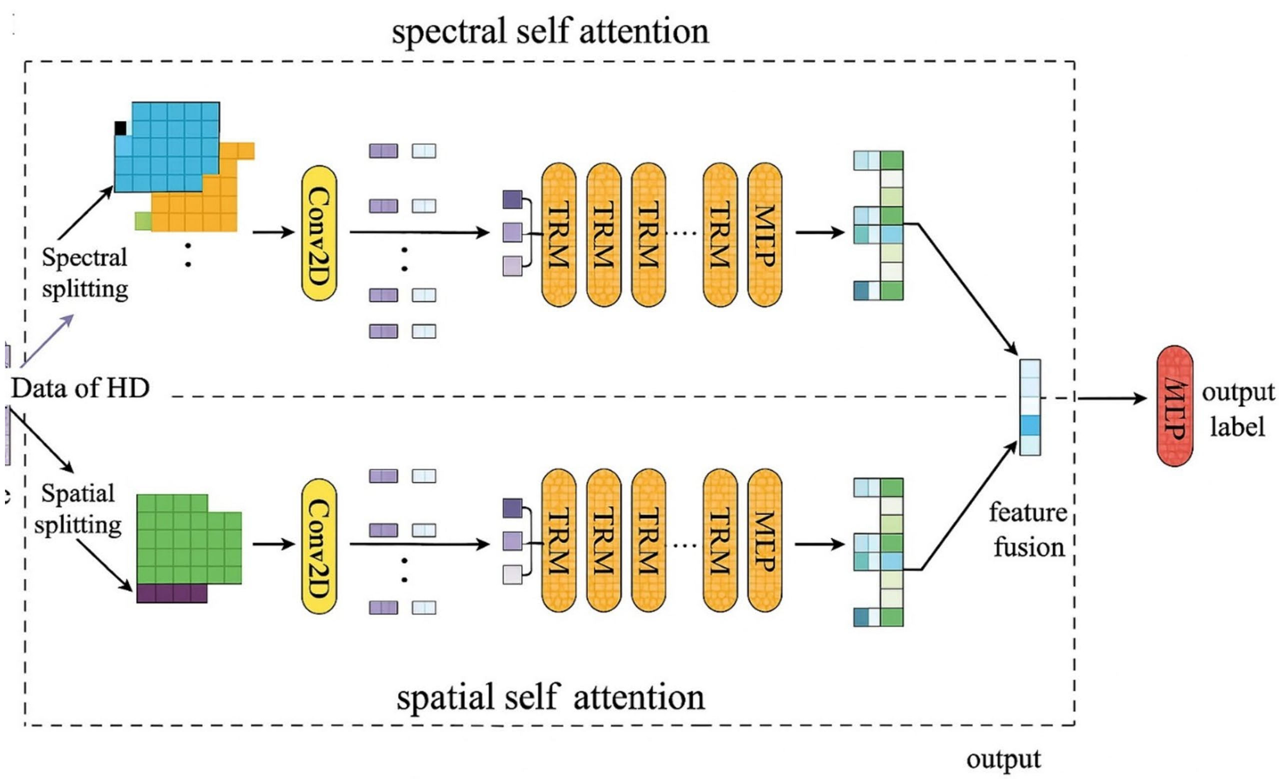

Fig. 1. Design of the TL-SAM. Conv2D signifies a convolutional layer. MLP symbolises a layer.

and spatial characteristics. Point-by-point addition is the alternative approach. The third approach is to multiply points individually. The label is then generated by an MLP using the fusion characteristics as input.

A comprehensive solution is given in the form of the TL-SAM. The back-propagation process might be used to teach it new tricks. In this study, we train the framework using the popular enhanced wild horse optimizer. The SpecSAM and SpatSAM modules each generate a 128-dimensional feature vector. These are concatenated during the fusion process, resulting in a final 256-dimensional feature vector that serves as input to the MLP-based classifier. This dimensionality was chosen to maintain a balance between representational richness and computational efficiency.

Hyper-parameter tuning using IWHOLFA model

The WHO method statistically models and replicates the social behaviour of these free-ranging horses[43]. The horse is social and prefers to live in a herd with other horses, including stallions, mares, and their young. Various performances have been shown to include mating and grazing, pursuit, dominance, and command. An initial population was divided up into many categories using the 5 stages of WHO methodologies stated under Primary. In this method, N signifies the total sum of inhabitants and G is the total sum of study groups[44]. Since the herds always had a dominant male member (a stallion), we may deduce that G is the total number of stallions produced by this method, and NG indicates that the remaining population (foals and mares) were dispersed uniformly between the groups. For modelling grazing efficiency, we provide the following Eq. (6). The position of each horse (solution) is updated based on the influence of the stallion (best solution) during grazing.

$$X_{i,G}^{j} = 2Z\cos\left(2\pi\, RZ\right) \times \left(Stallion^{j} - X_{i,G}^{j}\right) + Stallion^{j} \tag{6}$$

.

where $X_{i,G}^{j}$ represents the current position of the foal/mare group, Stallion denotes the stallion's position, R denotes an unchanging stochastic integer between −2 and 2, and Z is the procedure calculated using the following Eq. (7). This models random motion with stochastic influences.

$$P = \overline{R_1} < TDR; IDX = \left(P = 0\right); Z = R_2 \Theta\, IDX + \overline{R_3} \Theta\, \left(IDX\right) \tag{7}$$

.

where P mentions the vector comprising zero to one, $\overline{R_1}$ and $\overline{R_3}$ indicate the uniform accidental number between 0 and 1, whereas R_2 stands for the arbitrary number between 0 and 1. TDR, which starts at 1 and decreases until it reaches 0 in the following Eq. (8):

$$TDR = 1 - \times \left(\frac{1}{maxit}\right) \tag{8}$$

.

where it stands for the current repetition and maxit for the largest possible sum of iterations. Foals in group for mate performance implementation, whereas foals in group j are driven to group[45]. The mean form of the Crossover function for modelling equine mating performance is given in Eq. (9). Mating behavior is simulated using crossover operations to explore the solution space.

$$X_{G,K}^{P} = Crossover\left(X_{G,i}^{q}, X_{G,j}^{Z}\right) i \neq\ k, p = q = end, Crossover = Mean \tag{9}$$

.

Group leaders, or "stallions," guide the herd to a watering source as part of the WHO approach[46]. Since two groups of Stallions are vying for use of the same watering hole, the dominant group often drinks here first. That procedure's next step is outlined in the following formula.:

$$\overline{Stallion_{G_i}} = \begin{cases} 2Z\cos\left(2\pi\, RZ\right) \times\ \left(WH - Stallion_{G_i}\right) + WH if R_3 > 0.5 \\ Z\cos\left(2\pi\, RZ\right) \times\ \left(WVH - Stallion_{G_i}\right) + WH if R_3 \leq\ 0.5 \end{cases} \tag{10}$$

.

where $\overline{Stallion_{G_i}}$ indicates the succession of leaders to come. The WH denotes a watering hole or spring. The next phase is a fitness-based vote for a leader. Based on this algorithm, the leading position and appropriate member are adjusted.:

$$\overline{Stallion_{G_i}} = \begin{cases} X_{G,i} if cost\left(X_{G,i}\right) < cost\left(Stallion_{G_i}\right) \\ Stallion_{G_i} if cost\left(X_{G,i}\right) > costt\left(Stallion_{G_i}\right) \end{cases} \tag{11}$$

.

Cuckoo search (CS) was essential to the IWHOLF[47]. Using the Levy flight as an iteration in the presented method, a unique solution was developed. Levy Flight is incorporated to allow large jumps, improving the optimizer's ability to escape local optima.

$$X_{G,i} = X_{i,G} - \gamma\ \left(X_{i,G} - X_g\right) \oplus\ Levy\left(\lambda\ \right) = X_{i,G} + \frac{0.01u}{\left|v\right|^{1/\lambda}}\left(X_{i,G} - X_g\right) \tag{12}$$

.

where $X_{i,G}$ implies the ith place of collection member, γ signifies the stage scaling size, X_g stands optimum keys, ⊕ attitudes for the process of element-, represents the Levy flight (LF) advocates, but u and v are resolute as:

$$uN\left(0,\sigma\,_{u}^{2}\right),vN\left(0,\sigma\,_{v}^{2}\right) \tag{13}$$

.

The standard deviations $\sigma\,_u$ and $\sigma\,_v$ are expressed as:

$$\sigma\,_{u}=\left[\frac{sin\left(\frac{\lambda\,\pi}{2}\right)..\varGamma\,\left(1+\lambda\,\right)}{2^{(\lambda\,-1)}\lambda\,\varGamma\,\left(\frac{1+\lambda}{2}\right)}\right]^{1/\lambda}\sigma\,_{v}=1 \tag{14}$$

.

A new potential solution, which stands for the Gamma function, was developed and tested. The capacity of the proposed strategy to strike a balance among global is the main advantage of this improvement.

Alert system

The warning system is based on a score CCC, which helps therapeutic experts at the hospital monitor a patient's abnormal health conditions. From the acquired data record fp∈R_{SC}, all vital signs are compared against standard reference levels to calculate a score value S_y ranging from 0 to 3. Here, 0 indicates a normal state, while values greater than 3 signify critical patient conditions. For each record R_{SC}, a score vector $V_{SC}=[v_1,v_2,v_3,...,v_p]$ is computed. This scoring mechanism assists the alarm system in making timely decisions. If a patient's score is high, they are flagged for immediate emergency intervention. If the score is low, the system can relax the monitoring frequency. Thus, the data-network-driven warning system ensures that medical personnel are promptly alerted based on the patient's condition.

Experimentation, results and discussion

Implementation set-up

For the purpose of comparing different models, a 70% training-to-30% testing split was chosen. Deep learning techniques were applied using the Python packages Keras and TensorFlow. Table 3 displays the system requirements and supported applications. Training the data and generating the final findings took the system around an hour. To guarantee that network parameters could be properly updated and optimised, we trained using a total of 600 iterations, a batch size of 64, and the IWHOLFA method for optimisation. Eight heads are used in the self-mechanism that is being applied. Since the sum of convolution kernels is always set to 128 during pixel embedding, the recovered feature vectors along both the spectral and spatial dimensions also have a size of 128. The final model includes a multi-layer perceptron (MLP) for classification, and this MLP has 128 neurons in its output layer (where K is the sum of classes). In addition, model training makes use of function. The heart disease dataset's confusion matrix is shown in Fig. 2.

To reduce the risk of overfitting due to the model's complexity and limited dataset size, several regularization strategies were incorporated during training. These include: (i) 5-fold cross-validation to assess model robustness across data splits; (ii) dropout layers integrated within the transformer blocks to prevent co-adaptation of features; (iii) early stopping based on validation loss to halt training when the model began to overfit; and (iv) batch normalization and data normalization to stabilize training and ensure consistent input distribution. Together, these methods contributed to improved model generalization and reduced overfitting on the IoMT dataset.

Evaluation measures

Measures of presentation like precision, accuracy, F1 score, etc., are computed and used to assess the heftiness of the projected model. Particularly in the medical industry, where a precise forecast is needed, determining the model's accuracy is crucial.

$$Accuracy=\frac{Numberofcorrectlyclassifiedpredictions}{Totalpredictions} \tag{15}$$

$$Precision=\frac{TP}{TP+FP} \tag{16}$$

Software	Hardware
Windows – 10	RAM 8-GB
Anaconda 1.0.0	DDR4
Pycharm 2023.1.1	Core i-5
Python-3.7.3	6th Generation

Table 3. Scheme specifications.

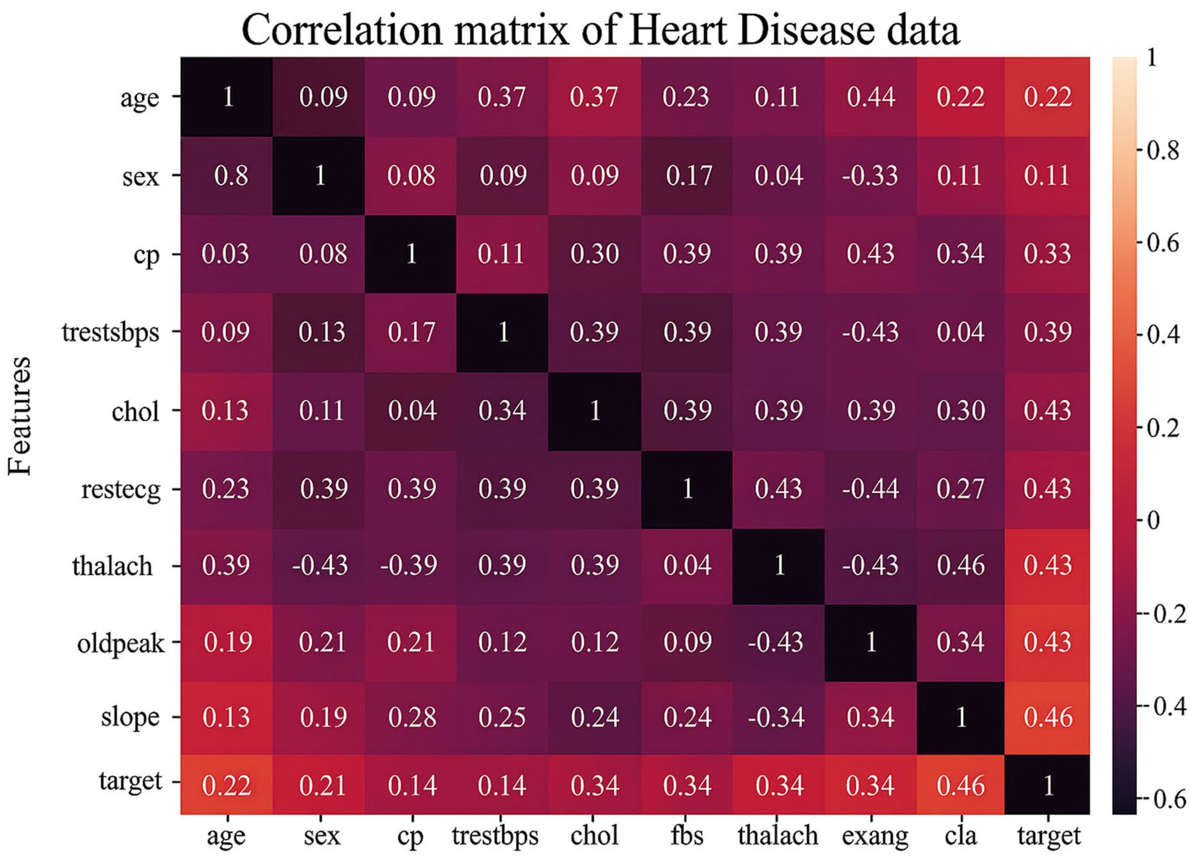

Fig. 2. Correlation matrix of heart disease data.

$$Recall = \frac{TP}{TP + FN} \tag{17}$$

$$F1 - score = 2 \times \frac{precision \times recall}{precision + recall} \tag{18}$$

where TP, TN, FP, and FN characterise $falsepositive \wedge falsenegative$, correspondingly. Figure 3 presents the projected model's accuracy on the training and testing dataset.

Validation analysis of the proposed model

The presentation of deep learning representations is denoted in Table 4. The accuracy score, precision rate, recall value, and F1 score were all determined in the analysis of the DBN model to be 0.9398, 0.85, 0.89, and 0.90, respectively. Following this, the RNN model attained an accuracy score of 0.9120, a precision rate of 0.89, a recall value of 0.90, and finally an F1-score of 0.92. Following this, the CNN model attained an accuracy of 0.9100, a precision rate of 0.90, a recall value of 0.91, and finally an F1-score of 0.91. After the Vision Transform model reached an accuracy score of 0.9591, a precision rate of 0.92, a recall value of 0.93, and finally, an F1-score of 0.93, After the Swin Transform reached an accuracy of 0.9689 and a recall value of 0.94, the recall value was also 0.96, and lastly, the F1-score was 0.95, respectively. After the TL-SAM model reached an accuracy score of 0.9862, a precision rate of 0.97, a recall worth of 0.98, and finally, an F1-score of 0.97. The corresponding visual representations are shown in Figs. 4 and 5.

An accuracy assessment of learning models is shown in Table 5 above and the graphical representation is presented in Fig. 6. In this analysis, we evaluated the accuracy using a variety of models. DT model attained an accuracy score of 0.8778, AdaBoost model accomplished an accuracy score of 0.8852, LR model accomplished an accuracy score of 0.8442, SGD model accomplished an accuracy score of 0.6491, RF model achieved an accuracy score of 0.8188, GBM model realised an accuracy score of 0.8852, ETC model attained an accuracy score of 0.8262, GNB model accomplished an accuracy of 0.7540, and SVM model accomplished an accuracy score of 0.8852, respectively.

Table 6 above represents the analysis of the proposed model with various learning rates he graphical representation is presented in Fig. 7. In the analysis, we used different learning degrees to determine the presentation of the projected perfect. In the analysis of 0.1 learning degree, the model stretched the accuracy score to 0.9470, the precision rate to 0.911, the recall value to 0.948, and, in conclusion, the F1-score to 0.933, respectively. After the 0.01 learning rate, the model reached an accuracy score of 0.9691, a precision rate of

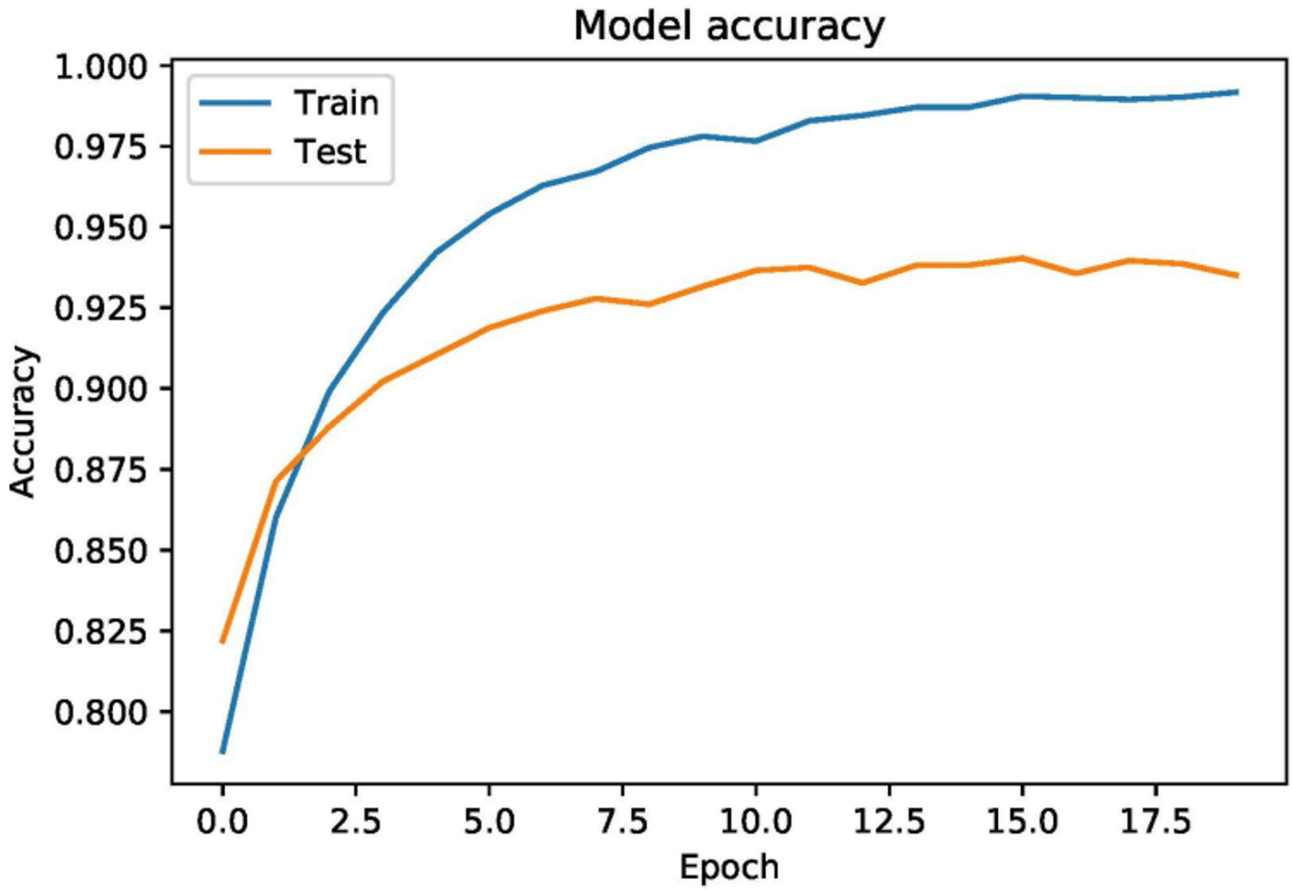

Fig. 3. Proposed Model's Accuracy on dataset.

Model	Accuracy	Precision	Recall	F1 Score
DBN	0.9398	0.85	0.89	0.90
RNN	0.9120	0.89	0.90	0.92
CNN	0.9100	0.90	0.91	0.91
Vision Transform	0.9591	0.92	0.93	0.93
Swin Transform	0.9689	0.94	0.96	0.95
TL-SAE	0.9862	0.97	0.98	0.97

Table 4. Presentation of deep learning representations.

0.934, a recall charge of 0.90, and finally, an F1-score of 0.90. After the 0.001 learning rate, the model grasped the accuracy score as 0.9720, the precision rate as 0.953, the recall worth as 0.89, and finally the F1-score as 0.89, respectively. After the 0.0001 learning rate, the model reached an accuracy score of 0.9862, a precision rate of 0.970, a recall charge of 0.98, and finally, an F1-score of 0.97 correspondingly.

The superior performance of the TL-SAM model can be attributed to several architectural and functional advantages. First, by separately modeling spectral (temporal/clinical) and spatial (sensor-related) features using SpecSAM and SpatSAM, the model captures distinct patterns that are often blurred in unified approaches. This separation allows for more targeted learning of complex interdependencies among features. Second, the transformer's self-attention mechanism enhances the model's ability to focus on critical feature interactions, such as sudden fluctuations in heart rate or blood pressure, which are essential in disease classification. Third, the use of IWHOLFA for hyperparameter optimization ensures that the model operates under optimal conditions, avoiding both underfitting and overfitting. Together, these factors explain why TL-SAM achieves statistically significant improvements in performance metrics compared to existing deep learning architectures.

Statistical Significance (paired *t*-test) between TL-SAM and Baseline Models on Accuracy and F1-score is presented in Table 7.

To evaluate whether the performance improvements of the TL-SAM model were statistically significant, we conducted paired *t*-tests comparing its accuracy and F1-score against each baseline model. As shown in Table 7,

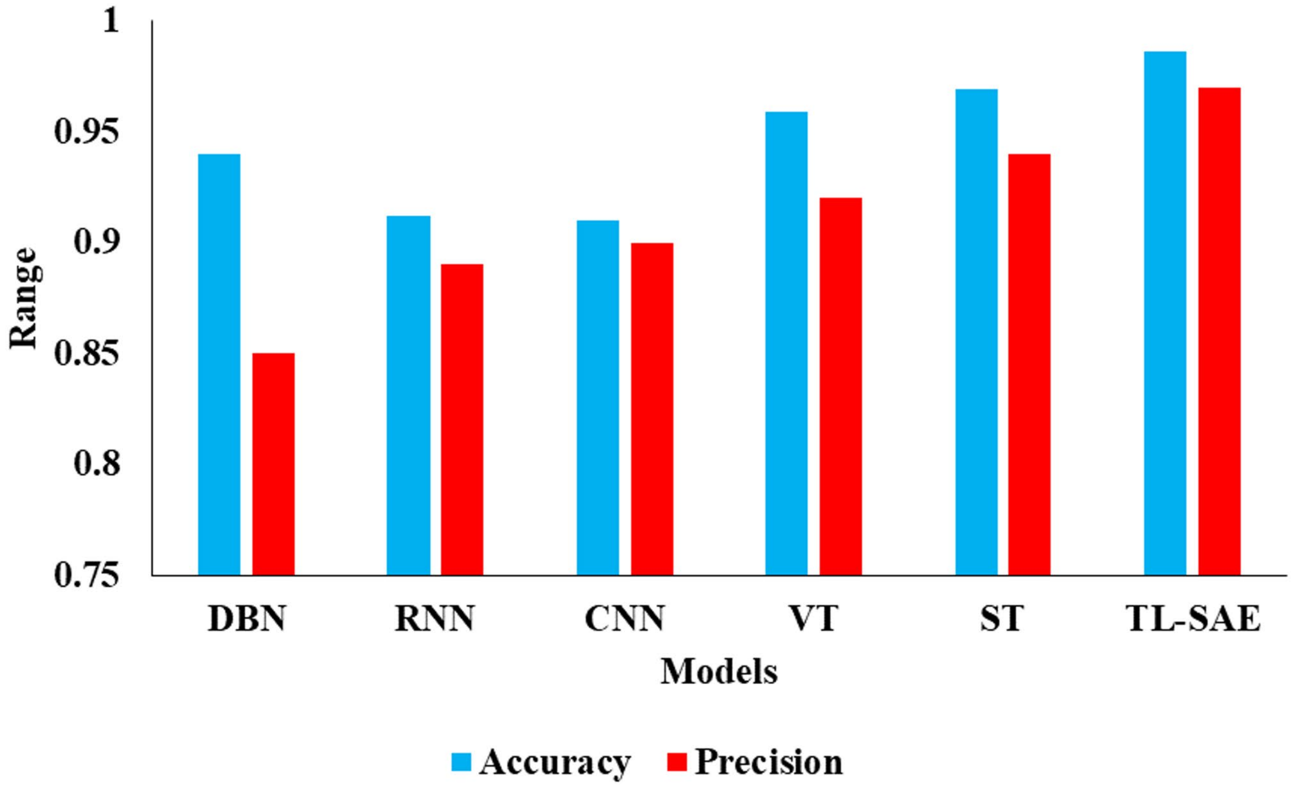

Fig. 4. Graphical Investigation of proposed model.

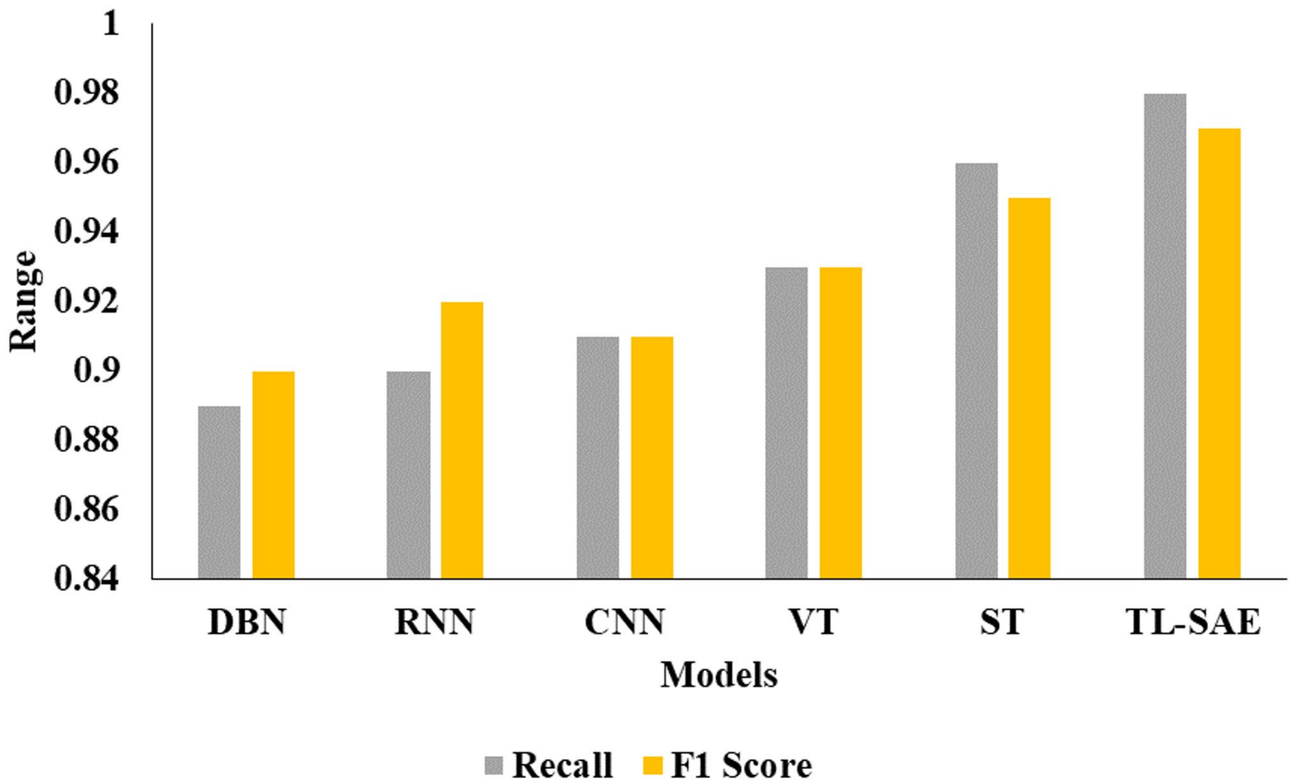

Fig. 5. Comparative analysis of several DL models.

Representations	Accuracy
DT	0.8778
AdaBoost	0.8852
LR	0.8442
SGD	0.6491
RNN	0.9100
LSTM	0.9391
RF	0.8188
GBM	0.8852
ETC	0.8262
GNB	0.7540
SVM	0.7622
MLP	0.9420
Proposed	0.9807

Table 5. Accuracy assessment of learning models.

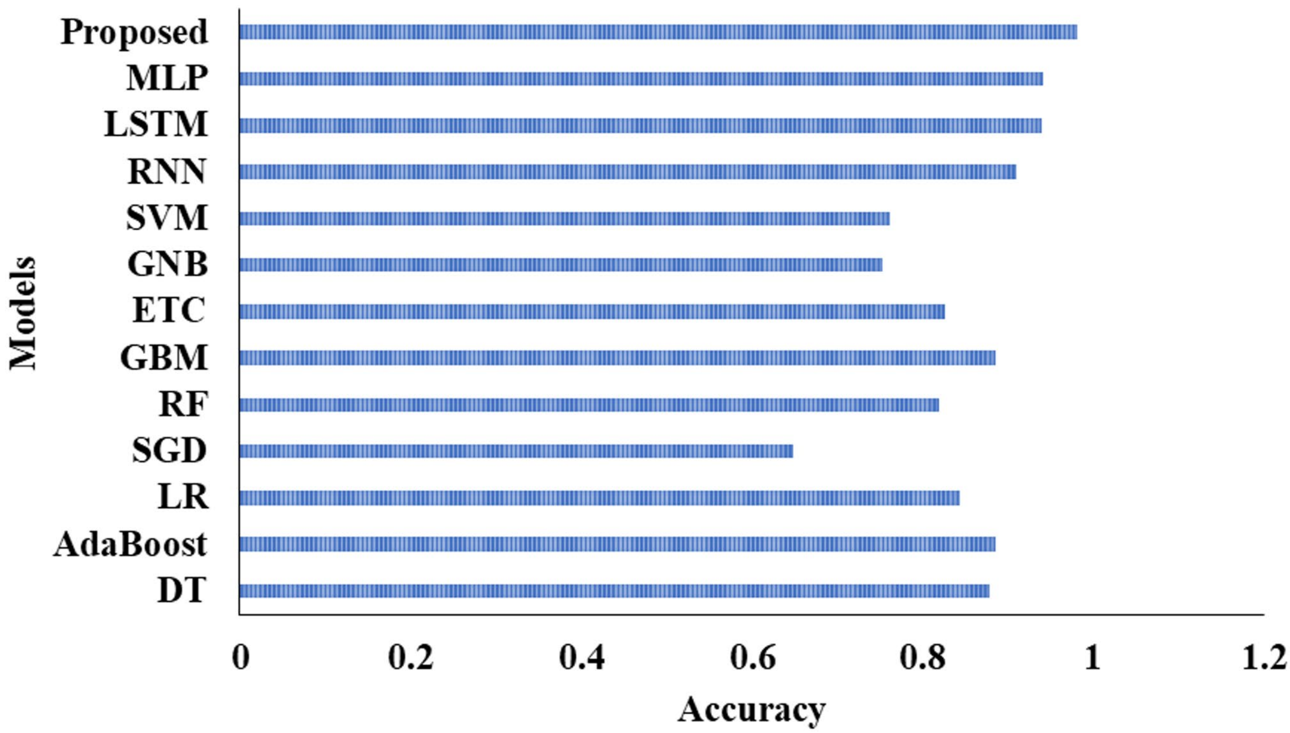

Fig. 6. Accuracy Analysis on various ML and DL classifiers.

Learning Rate	Accuracy	Precision	Recall	F1 Score
0.1	0.9470	0.911	0.948	0.933
0.01	0.9691	0.934	0.90	0.90
0.001	0.9720	0.953	0.89	0.89
0.0001	0.9862	0.970	0.98	0.97

Table 6. Analysis of the proposed model with various learning rates.

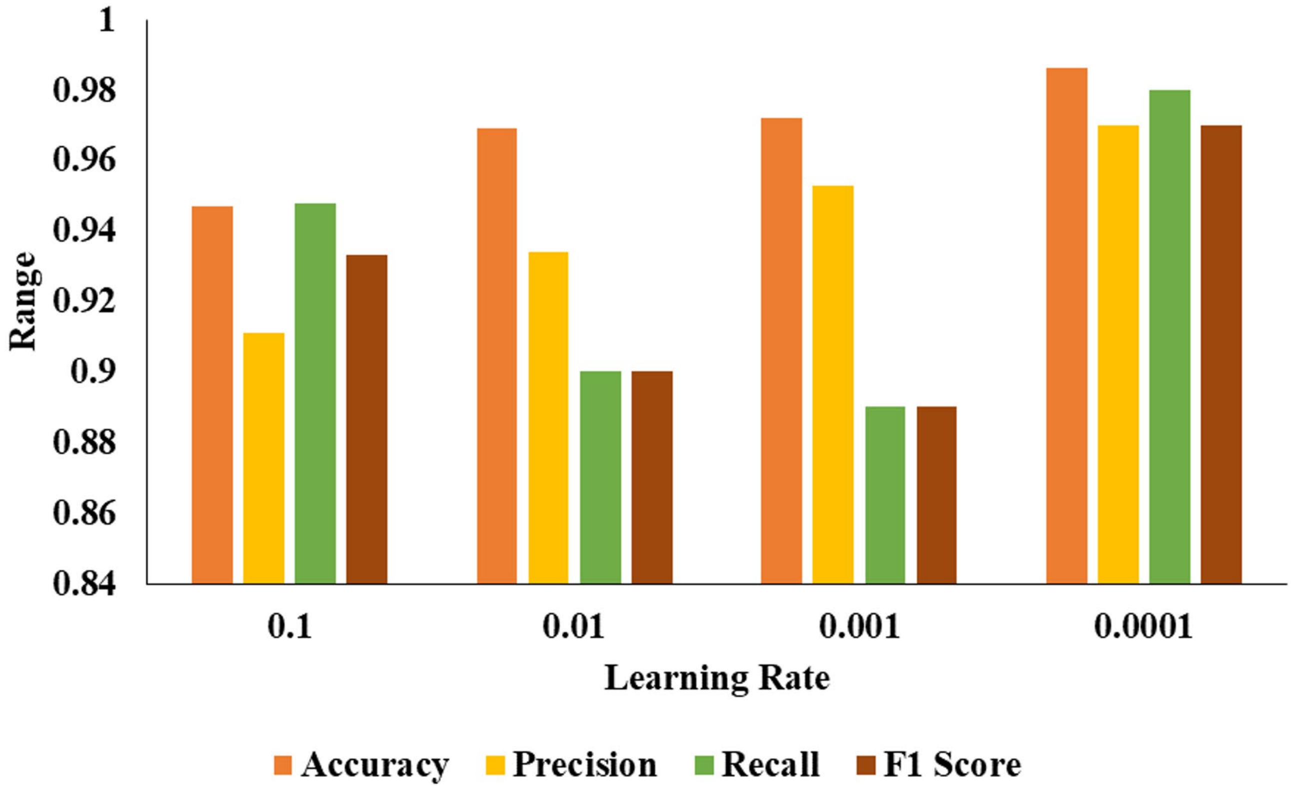

Fig. 7. Graphical Comparison of proposed model for various learning rate.

Model	Accuracy (mean ± std)	F1-score (mean ± std)	*p*-value vs. TL-SAM (Accuracy)	*p*-value vs. TL-SAM (F1)
CNN	91.00 ± 1.2	91.00 ± 1.1	0.0031	0.0028
RNN	91.20 ± 1.4	92.00 ± 1.0	0.0045	0.0033
Vision Transformer	95.91 ± 1.1	93.00 ± 0.9	0.0172	0.0214
Swin Transformer	96.89 ± 0.9	95.00 ± 1.0	0.0416	0.0388
TL-SAM (proposed)	98.62 ± 0.5	97.00 ± 0.6	–	–

Table 7. Statistical significance (paired *t*-test) between TL-SAM and baseline models on accuracy and F1-score.

TL-SAM achieved significantly higher performance ($p < 0.05$) compared to all alternatives. This statistical analysis supports the robustness of the proposed framework and confirms that the observed improvements are not due to random chance.

Conclusion and future work

In some emergent situations, patients need constant monitoring and care. However, going to the hospital to carry them out is difficult because of time restrictions. A system for remote healthcare monitoring can help with this. With the widespread use of cloud computing, which improves the speed and accessibility of data processing, the suggested method presents an efficient data science procedure for IoT-supported healthcare data collected from a wide variety of Internet of Things sensors. For data science purposes, this information is stored on the cloud. The suggested method improves data processing efficiency and facilitates continuous patient health monitoring. Implanted (IoT) sensors collect vital health data, which is then sent to a remote server. The suggested approach utilises a transformer to construct a spectral- and spatial-dimensional self-attention model. To complete the classification process, the study combined the characteristics of the two models into spatial-spectral features. Important characteristics in both the spectral and spatial dimensions may be extracted using SAM inside the proposed framework (TL-SAM). The suggested model was able to accomplish the IWHOLFA model by hyper-parameter tweaking. The experimental consequences demonstrate that the suggested model outdoes the state-of-the-art by 3–5% accuracy, with a correctness of 0.9862%, precision of 0.97%, recall of 0.98%, and F-score of 0.97%.

Finally, we compare the proposed system's performance against existing smart healthcare monitoring systems and a wide range of real-time healthcare information. Extending the IoT-based human healthcare monitoring system is a priority, and this can only be done once extensive long-term data collection has been completed and the elements related to the prediction of possible threats have been investigated. This provides a solid, evidence-based foundation for future efforts to control and prevent illnesses associated with high mortality rates.

Data availability

Data availability statement: The data that support the findings of this study are available upon reasonable request from the corresponding author Vijay Kumar Damera using email : trameravijaykumar@gmail.com.

Received: 12 March 2025; Accepted: 30 June 2025
Published online: 05 July 2025

References

1. Sarmah, S. S. An efficient IoT-based patient monitoring and heart disease prediction system using deep learning modified neural network. *IEEE Access.* **8**, 135784–135797 (2020).
2. Saheed, Y. K., Abdulganiyu, O. H. & Tchakoucht, T. A. Modified genetic algorithm and fine-tuned long short-term memory network for intrusion detection in the internet of things networks with edge capabilities. *Appl. Soft Comput.* **155**, 111434 (2024).
3. Qiao, C., Brown, K. N., Zhang, F. & Tian, Z. Federated adaptive asynchronous clustering algorithm for wireless mesh networks. *IEEE Trans. Knowl. Data Eng.* https://doi.org/10.1109/TKDE.2021.3119550 (2021).
4. Souri, A. et al. A new machine learning-based healthcare monitoring model for student's condition diagnosis in internet of things environment. *Soft. Comput.* **24**(22), 17111–17121 (2020).
5. Alsahfi, T. et al. Optimizing healthcare big data performance through regional computing. *Sci. Rep.* **15**, 3129. https://doi.org/10.1038/s41598-025-87515-5 (2025).
6. Saheed, Y. K., Misra, S. & CPS-IoT-PPDNN *A New Explainable Privacy Preserving DNN for Resilient Anomaly Detection in Cyber-Physical Systems-enabled IoT Networks* Vol. 191115939 (Chaos, Solitons & Fractals, 2025).
7. Wang, Z. et al. ThreatInsight: Innovating early threat detection through threat-intelligence-driven analysis and attribution. *IEEE Trans. Knowl. Data Eng.* **36**(12), 9388–9402 (2024). https://doi.org/10.1109/TKDE.2024.3474792
8. Verma, P., Tiwari, R., Hong, W. C., Upadhyay, S. & Yeh, Y. H. FETCH: a deep learning-based fog computing and IoT integrated environment for healthcare monitoring and diagnosis. *IEEE Access.* **10**, 12548–12563 (2022).
9. Alazzam, M. B., Alassery, F. & Almulihi, A. A novel smart healthcare monitoring system using machine learning and the Internet of Things. *Wireless Communications and Mobile Computing* 1–7 (2021).
10. Khan, M. A. An IoT framework for heart disease prediction based on MDCNN classifier. *IEEE Access.* **8**, 34717–34727 (2020).
11. Devi, R. L. & Kalaivani, V. Machine learning and IoT-based cardiac arrhythmia diagnosis using statistical and dynamic features of ECG. *J. Supercomput.* **76**(9), 6533–6544 (2020).
12. Khan, M. A. & Algarni, F. *A Healthcare Monitoring System for the Diagnosis of Heart Disease in the IoMT Cloud Environment Using MSSO-ANFIS* 122259–122269 (IEEE Access, 2020).
13. Lv, Y., Zhang, W., Tian, Z., Lu, H. & Shi, W. Do not trust the clouds easily: the insecurity of content security policy based on object storage. *IEEE Internet Things J.* **10**(12), 10462–10470. https://doi.org/10.1109/jiot.2023.3238658 (2023).
14. Awotunde, J. B., Folorunso, S. O., Bhoi, A. K., Adebayo, P. O. & Ijaz, M. F. *Disease Diagnosis System for IoT-based Wearable Body Sensors with Machine Learning Algorithm* 201–222 (Hybrid Artificial Intelligence and IoT in Healthcare, 2021).
15. Saheed, Y. K. & Chukwuere, J. E. XAIEnsembleTL-IoV: A new eXplainable Artificial Intelligence ensemble transfer learning for zero-day botnet attack detection in the Internet of Vehicles. *Results Eng.* **24**, 103171 (2024).
16. Harimoorthy, K. & Thangavelu, M. Multi-disease prediction model using improved SVM-radial bias technique in healthcare monitoring system. *J. Ambient Intell. Humaniz. Comput.* **12**, 3715–3723 (2021).
17. Aldahiri, A., Alrashed, B. & Hussain, W. Trends in using IoT with machine learning in health prediction systems. *Forecasting* **3**(1), 181–206 (2021).
18. Wu, X., Liu, C., Wang, L. & Bilal, M. *Internet of things-enabled real-time Health Monitoring System Using Deep Learning* 1–12 (Neural Computing and Applications, 2021).
19. Alshamrani, M. IoT and artificial intelligence implementations for remote healthcare monitoring systems: A survey. *J. King Saud University-Computer Inform. Sci.* **34**(8), 4687–4701 (2022).
20. Pan, Y., Fu, M., Cheng, B., Tao, X. & Guo, J. Enhanced deep learning-assisted convolutional neural network for heart disease prediction on the internet of medical things platform. *IEEE Access.* **8**, 189503–189512 (2020).
21. Ren, Y., Xiao, Y., Zhou, Y., Zhang, Z. & Tian, Z. CSKG4APT: A cybersecurity knowledge graph for advanced persistent threat organization attribution. *IEEE Trans. Knowl. Data Eng.* **35**(6), 5695–5709. https://doi.org/10.1109/TKDE.2022.3175719 (2023).
22. Elayan, H., Aloqaily, M. & Guizani, M. Digital twin for intelligent context-aware IoT healthcare systems. *IEEE Internet Things J.* **8**(23), 16749–16757 (2021).
23. Saheed, Y. K., Misra, S. & Chockalingam, S. Autoencoder via DCNN and LSTM Models for Intrusion Detection in Industrial Control Systems of Critical Infrastructures, *IEEE/ACM 4th International Workshop on Engineering and Cybersecurity of Critical Systems (EnCyCriS)* 9–16 (2023). https://doi.org/10.1109/EnCyCriS59249.2023.00006.
24. Bharadwaj, H. K. et al. A review on the role of machine learning in enabling IoT-based healthcare applications. *IEEE Access.* **9**, 38859–38890 (2021).
25. Ed-Daoudy, A. & Maalmi, K. April. Real-time machine learning for early detection of heart disease using big data approach. In *2019 international conference on wireless technologies, embedded and intelligent systems (WITS)* 1–5 IEEE. (2019).
26. Golec, M., Gill, S. S., Parlikad, A. K. & Uhlig, S. *HealthFaaS: AI-based Smart Healthcare System for Heart Patients Using Serverless Computing* (IEEE Internet of Things Journal, 2023).
27. Kadhim, M. A. & Radhi, A. M. Heart disease classification using optimized machine learning algorithms. *Iraqi J. Comput. Sci. Math.* **4**(2), 31–42 (2023).
28. Kaur, J., Khehra, B. S. & Singh, A. Back propagation artificial neural network for diagnosis of heart disease. *J. Reliable Intell. Environ.* **9**(1), 57–85 (2023).
29. Kumar, A. & Sateesh, N. Big data predictive analysis for type-2 diabetes-based heart disease using feature extraction and classification by machine learning architectures.
30. Mohapatra, S. et al. A stacking classifiers model for detecting heart irregularities and predicting cardiovascular disease. *Healthc. Analytics.* **3**, 100133 (2023).
31. Bhatt, C. M., Patel, P., Ghetia, T. & Mazzeo, P. L. Effective heart disease prediction using machine learning techniques. *Algorithms* **16**(2), 88 (2023).
32. Chaudhari, J., Mewada, H., Kosta, Y., Bhagat, K. & Kirange, S. Recursive feature elimination and optimized hybrid ensemble approach for early heart disease prediction. *Int. J. Comput. Digit. Syst.* **14**(1), 1–xx (2023).
33. Karthikeyini, S., Vidhya, G., Vetriselvi, T. & Deepa, K. Heart disease prognosis using D-GRU with logistic Chaos honey Badger optimization in IoMT framework. *Inform. Technol. Control.* **52**(2), 367–380 (2023).

34. Jain, A., Rao, A. C. S., Jain, P. K. & Hu, Y. C. Optimized levy flight model for heart disease prediction using CNN framework in big data application. *Expert Syst. Appl.* **223**, 119859 (2023).
35. Raheja, N. & Manocha, A. K. An IoT-enabled secured clinical health care framework for the diagnosis of heart diseases. *Biomed. Signal Process. Control.* **80**, 104368 (2023).
36. Dileep, P. et al. An automatic heart disease prediction using cluster-based bi-directional LSTM (C-BiLSTM) algorithm. *Neural Comput. Appl.* **35**(10), 7253–7266 (2023).
37. Paul, B. & Karn, B. Heart disease prediction using scaled conjugate gradient backpropagation of artificial neural network. *Soft. Comput.* **27**(10), 6687–6702 (2023).
38. Hasanova, H., Tufail, M., Baek, U. J., Park, J. T. & Kim, M. S. A novel blockchain-enabled heart disease prediction mechanism using machine learning. *Comput. Electr. Eng.* **101**, 108086 (2022).
39. Shanthi, S., Nidhya, R., Perumal, U. & Kumar, M. An Efficient IoT framework for patient monitoring and predicting heart disease based on machine learning algorithms. *Applications of Artificial Intelligence and Soft Computing Techniques* 179–200 (2022).
40. Munagala, N. K., Langoju, L. R. R., Rani, A. D. & Reddy, D. R. K. A smart IoT-enabled heart disease monitoring system using meta-heuristic-based Fuzzy-LSTM model. *Biocybernetics Biomedical Eng.* **42**(4), 1183–1204 (2022).
41. Ahmad, T., Munir, A., Bhatti, S. H., Aftab, M. & Raza, M. A. Survival analysis of heart failure patients: A case study. *PLoS ONE.* **12**, e0181001 (2017).
42. Asuncion, A. & Newman, D. U. C. I. *Machine Learning Repository* (Irvine University of California, 2007).
43. Saravana Kumar, C. An authentication technique for accessing deduplicated data from the private cloud using one-time password. *Int. J. Inf. Sec Priv. (IJISP)*. **11**(2), 1–10. https://doi.org/10.4018/IJISP.2017040101 (2017).
44. Arul Vinayakam Rajasimman, M., Manoharan, R. K., Aridoss, M. & Galety, M. G. Robust facial expression recognition using an evolutionary algorithm with a deep learning model. *App Sci.* **13**(1), 468 (2023).
45. Shanmugavadivel, K. et al. Deep learning-based sentiment analysis and offensive Language identification on multilingual code-mixed data. *Sci. Rep.* **12**, 21557 (2022).
46. Parthiban, S. et al. *Chaotic Salp Swarm Optimization-Based Energy-Aware VMP Technique for Cloud Data Centers*(Comput Intell Neurosci, 2022).
47. Saheed, Y. K., Abdulganiyu, O. H., Majikumna, K. U., Mustapha, M. & Workneh, A. D. ResNet50-1D-CNN: A new lightweight resNet50-One-dimensional Convolution neural network transfer learning-based approach for improved intrusion detection in cyber-physical systems. *Int. J. Crit. Infrastruct. Prot.* **45**, 100674 (2024).

Author contributions

Declaration: Author contributions: All authors (Vijay Kumar Damera, Ramesh Cheripelli, Narsaiah Putta, Sirisha G, and Deepthi Kalavala) contributed equally.Data availability statement: The data that support the findings of this study are available upon reasonable request from the corresponding author Vijay Kumar Damera using email : dameravijaykumar@gmail.com.Ethics approval: The submitted work is original and has not been published elsewhere in any form or language. Disclosure of potential conflicts of interest: There is no potential conflict of interest. Research involving Human Participants and/or Animals: NA Funding: The authors declare that no funds, grants, or other support were received during the preparation of this manuscript. Competing interests: The authors have no relevant financial or non-financial interests to disclose.Consent to publish: Informed consent was obtained from all authors for publication of this manuscript.

Funding

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to V.K.D.

Reprints and permissions information is available at www.nature.com/reprints.