



# OPEN IoT driven smart health monitoring for heart disease prediction using quantum kernel enhanced sardine diffusion and CNN

Dibas Kumar Hembram<sup>1✉</sup>, Satya Narayan Tripathy<sup>1</sup>, CH Narsimha Reddy<sup>2</sup>, Debabrata Dansana<sup>3</sup>, Quadri Noorulhasan Naveed<sup>4</sup>, Shafat Khan<sup>4</sup>, Mohammed Kareemullah<sup>5,6</sup> & Addisu Frinjo Emma<sup>7✉</sup>

Heart disease is one of the major causes of death worldwide, and the traditional diagnostic procedures typically cause delays in treatment, particularly in low-resource regions. In this article, we propose a novel IoT-based Quantum Kernel-Enhanced Sardine Diffusion Attention Network (Qua-KSar-DCK-ArNet) for real-time prediction of heart disease. The system is capable of continuously monitoring heart-related data such as ECG and heart rate via IoT sensors. Quantum Clustering with k-Means is applied to cluster the data, and Z-score Min–Max Normalization is applied for preprocessing. Fast Point Transformer is utilized to identify salient features. The Qua-KSar-DCK-ArNet model, a combination of quantum and classical deep learning methods, classifies the data for predicting the risk of heart disease. The system is fast and accurate, with an accuracy of 99%, significantly improving patient outcomes, especially in resource-scarce regions.

**Keywords** Heart disease prediction, Smart health monitor, Sardine Optimization Algorithm, Quantum-Classical Convolutional Neural Network, Diffusion Kernel Attention Network

## Abbreviations

Qua-KSar-DCK-ArNet	Quantum Kernel-Enhanced Sardine Diffusion Attention Network
CNN	Convolutional Neural Network
IoT	Internet of Things
CHD	Congenital heart disease
LSTM	Long short-term memory
DKAN	Diffusion Kernel Attention Network
SOA	Sardine Optimization Algorithm
FPT	Fast Point Transformer
LSA	Lightweight Self-Attention
MLP	Multilayer Perceptron

Heart disease claims the lives of more than three million people a year and has been the cause of 53% of deaths in the last 5 years<sup>1</sup>. Obesity, sugar, cholesterol, and high blood pressure are some causes. Improved treatment and death prevention are possible with early detection. Artificial intelligence, neural networks, and machine learning are some technologies used in research to forecast cardiac disease<sup>2,3</sup>. However, machine learning cannot see data in three dimensions and needs many data. Physicians who use real-time monitoring systems get better results, and IoT-based prediction allays worries about security and privacy. This service is also compatible with mobile applications for earlier prediction. Many industries, including industry, healthcare, agriculture, and urban areas, are utilising the Internet of Things (IoT)<sup>4,5</sup>. The Internet of Things (IoT) provides practical, reasonably priced,

<sup>1</sup>Berhampur University, Berhampur, Odisha, India. <sup>2</sup>CSE-SOET, MNR University, Fasalwadi, Telangana, India.

<sup>3</sup>Rajendra University, Balangir, Odisha, India. <sup>4</sup>Department of Computer Science, College of Computer Science, King Khalid University, Abha, Saudi Arabia. <sup>5</sup>Department of Mechanical Engineering, Graphic Era (Deemed to be University), Clement Town, Dehradun 248002, India. <sup>6</sup>Centre for Research Impact & Outcome, Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, Punjab 140417, India. <sup>7</sup>College of Engineering and Technology, School of Mechanical and Automotive Engineering, Geddo Zone, South Ethiopia Regional State, Dilla University, P.O. Box 419, Dilla, Ethiopia. ✉email: dkh.rs.cs@buodisha.edu.in; addisuf@du.edu.et

and ecologically friendly solutions, such as wireless sensor networks, smart mobile devices, and identification procedures.

It might lengthen equipment lifespans, optimise resource scheduling, and decrease equipment unavailability<sup>6,7</sup>. Through remote patient monitoring and the ability to store patient condition histories in databases, IoT also makes early disease detection and diagnosis possible. To tackle challenges and trends in the healthcare sector, researchers are exploring the use of IoT in healthcare<sup>8,9</sup>. For example, deep learning methods have been employed to detect heart disease, a complex condition that requires careful management<sup>10</sup>. Heart disease is the leading cause of death globally and is linked to various behavioural risk factors. Congenital heart disease (CHD) prevention and monitoring are being revolutionised by the Internet of Things (IoT). IoT-based solutions link medical equipment to the internet to collect background data on the surroundings and health of a patient<sup>11,12</sup>.

Compared to conventional testing procedures, this data can be utilised to follow changes, detect abnormalities, and provide quick diagnosis, potentially saving lives. Patients with Congenital Heart Disease (CHD) have a better long-term prognosis because of wearable sensors. Continuous monitoring is made possible by IoT-based technologies, which help physicians spot issues early<sup>13,14</sup>. Machine learning algorithms examine patient data to find risk variables. While telemedicine improves accessibility and patient care, remote monitoring lowers the number of doctor and hospital visits. Lengthy short-term memory (LSTM), appropriate for processing and forecasting crucial events in time series with lengthy intervals and delays, is used in the study for CVD research. However, due to varying time intervals, ordinary LSTM cannot successfully learn essential components of a patient's medical state in the medical industry<sup>15</sup>.

## Novelty and contribution

The Novelty and contribution of this paper is given below:

- This manuscript proposes Quantum Kernel-Enhanced Sardine Diffusion Attention Network for Heart Disease Prediction in IoT-Based Smart Health Monitoring (Qua-KSar-DCK-ArNet).
- Implement Quantum Clustering with k-Means to categorize heart-related data into distinct clusters based on disease risk factors, enhancing the organization and analysis of patient information. This method enhances predictive accuracy and enables focused interventions for managing heart disease within healthcare environments.
- To preprocess heart-related data using the Z-score Min–Max normalization technique, ensuring improved data quality and consistency. This approach effectively scales the data, which helps improve model performance and accuracy in predicting heart disease, all while preserving the integrity of the underlying information for analysis.
- To classify processed heart-related data by creating the Quantum Kernel Sardine Diffusion Classical Kernel Attention Network (Qua-KSar-DCK-ArNet). This novel method combines the Quantum–Classical Convolutional Neural Network (Q2CNN) with the Diffusion Kernel Attention Network (DKAN), optimizing parameters using the Sardine Optimization Algorithm (SOA) to enhance heart disease prediction accuracy.
- Integration of cloud storage for secure, scalable, and accessible storage of patient data, enabling healthcare providers to access diagnostic results remotely.
- Facilitation of remote diagnosis by sending detected TB data directly to healthcare professionals, accelerating response time and improving patient care in underserved regions.

The rest of the manuscript is organized as follows: Section “[Literature survey](#)” provides a literature review, Section “[Proposed methodology](#)” details the proposed technique, Section “[Result and discussion](#)” summarizes the findings, and Section “[Discussion](#)” closes with future studies. The Research work organization is given in Fig. 1.

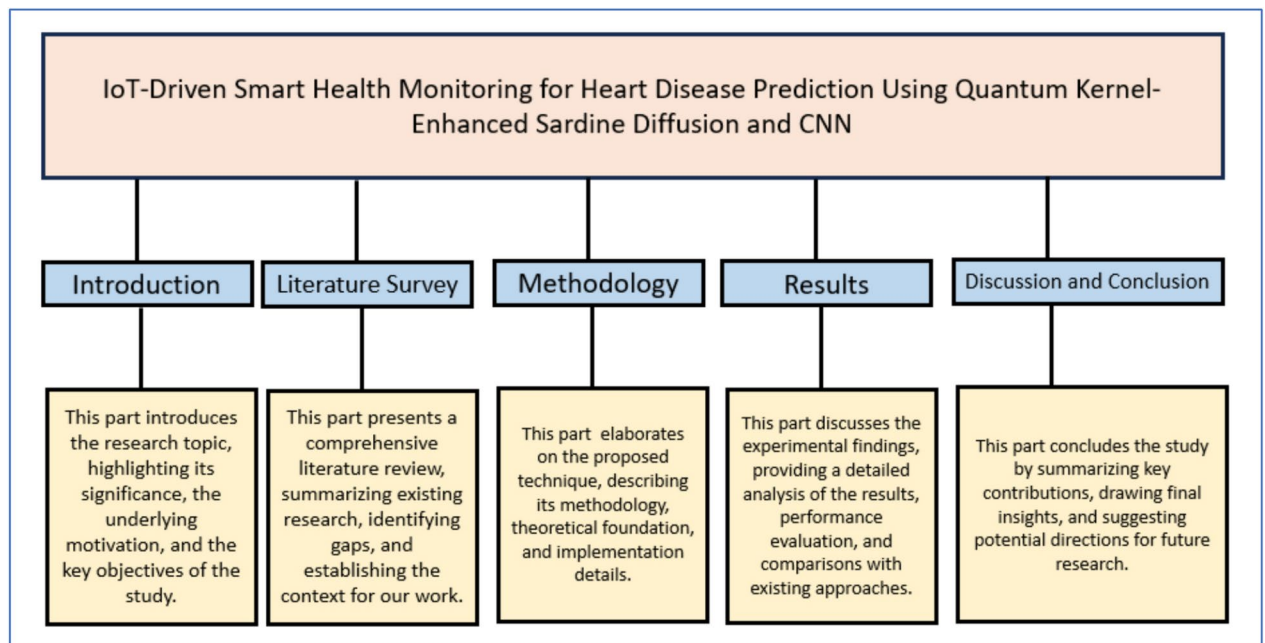
## Literature survey

The papers related to Smart Health Monitor for Early Heart Disease Prediction: An IoT-Based Patient Monitoring System using Deep Learning Methods are given below:

In 2024, Yenukar et al.<sup>16</sup> introduced an ambulatory blood pressure monitoring (IABPM) system for Smart Health monitoring for early heart disease prediction of IoT-based patients. A cutting-edge IoT ambulatory blood pressure monitoring system can offer real-time blood pressure, systolic, diastolic, and pulse rate measurements at predetermined intervals. The system employs machine learning methods such as Naïve Bayes, K-Nearest Neighbors, random forest, decision tree, and Support Vector Machine to predict the early warning score for CHD. The system's accuracy in forecasting blood pressure, a critical component of real-time intensive care for congestive heart failure, has reached 99.44%.

In 2023, Umer et al.<sup>17</sup> introduced an Extra Tree Classifier-Convolutional Neural Network (ET-CNN) for Smart Health monitoring for early heart disease prediction of IoT-based patients. Due to the COVID-19 pandemic's severe effects on healthcare, intelligent health monitoring systems have become more popular. Artificial intelligence (AI) and the Internet of Things (IoT) have transformed data collecting and communication. The activities of cardiac patients are monitored and reported using IoT and AI-based technologies in this study. An ensemble model called ET-CNN achieved an accuracy score of 0.9524.

In 2024, Alzakari et al.<sup>18</sup> introduced a Bidirectional Long-Short-Term Memory (Bi-LSTM) for a Smart Health Monitor to predict early heart disease in IoT-based patients. The development of medical technology has revolutionized remote healthcare monitoring, especially in managing chronic conditions like heart disease and high blood pressure. A novel approach integrates ECD from comprehensive medical records and physical data from patients' routine monitoring with regular medical monitoring.



**Fig. 1.** Research work organization.

In 2023, Almujaally et al.<sup>19</sup> introduced a Multilayer Perceptron Neural Network (MPNN) for Smart Health monitoring for early heart disease prediction of IoT-based patients. The research centres on a health monitoring system designed to track the activities of heart patients and diagnose illnesses through AI and IoT technology. The system enables real-time monitoring and disease classification by gathering data from various sources with intelligent sensors and AI. Experimental results demonstrate the system's effectiveness, highlighting its potential to improve healthcare, especially during the challenges posed by COVID-19.

In 2023, Akhbarifar et al.<sup>20</sup> introduced a Body Area Network and Personal Area Network (BAN-PAN) for Smart Health monitoring for early heart disease prediction of IoT-based patients. The COVID-19 pandemic has led to a surge in the need for remote patient screening and monitoring. However, the presence of sensitive health information raises significant security concerns. This study presents a method for remote health monitoring that safeguards health and medical data in cloud-based IoT environments through lightweight block encryption. The approach ensures the protection of sensitive information while utilizing data mining techniques to predict essential events.

In 2023, Islam et al.<sup>21</sup> introduced a Convolutional Long Short-Term Memory Network (CLSTMN) for Smart Health monitoring for Early heart disease prediction of IoT-based patients. This research proposes an Internet of Things (IoT)-based system designed for remote health monitoring and the early detection of health issues in home clinical settings. The system uses three different types of sensors: the MLX90614 non-contact infrared sensor for body temperature, the AD8232 ECG sensor module for ECG signal data, and the MAX30100 for blood oxygen levels and heart rate. A pre-trained deep learning model classifies potential diseases once the gathered data is sent to a server using the MQTT protocol. The technology immediately puts the user in contact with a physician for a more precise diagnosis.

In 2024, Pachiyannan et al.<sup>22</sup> introduced a Convolutional Neural Network (CNN) for Smart Health monitoring for early heart disease prediction of IoT-based patients. A novel healthcare tool called the Machine Learning-based Congenital Heart Disease Prediction Method (ML-CHDPM) helps quickly diagnose and categorize congenital heart disease (CHD) in expectant mothers. Using cutting-edge machine-learning techniques, the algorithm classifies instances according to clinical and demographic parameters. The model's performance is assessed using six different measures, showing that it can accurately predict and categorize cases of CHD, which is a big step towards early detection and diagnosis. Table 1 shows the Summary of the reviewed approach.

### Issues with the integration of IoT in healthcare and solutions

The integration of IoT in healthcare presents challenges such as data security risks, interoperability issues, high infrastructure costs, and data overload. Ensuring secure data transmission through encryption and block chain technology can mitigate security concerns. Standardized protocols and AI-driven analytics can address interoperability and data management issues, improving efficiency. The adoption of 5G and Mobile Edge Computing (MEC) enhances real-time data processing and connectivity, enabling faster decision-making. Additionally, public-private partnerships can help reduce costs and promote widespread adoption, ensuring seamless and efficient healthcare services.

References	Methods	Advantages	Disadvantages
Yenurkar et al. <sup>16</sup>	IABPM	High precision (99.44%) in blood pressure prediction for immediate critical care	Combining different machine learning methods for the best results may be complicated
Umer et al. <sup>17</sup>	ET-CNN	Vital accuracy (0.9524) results from ensemble learning that combines CNN's and ET's advantages	It can need a lot of processing power, especially when training the model
Alzakari et al. <sup>18</sup>	Bi-LSTM	Outstanding forecasting accuracy (99.4%) due to Bi-LSTM's sophisticated time-series modelling capabilities	Bi-LSTM models may incur significant computational costs, particularly in real-time applications
Almujally et al. <sup>19</sup>	MPNN	Effective in diagnosing and classifying diseases thanks to a flexible and basic neural architecture	MPNNs' efficacy for time-dependent data may be limited by their inability to capture temporal trends
Akhbarifar et al. <sup>20</sup>	BAN-PAN	Ensures privacy in cloud-based systems by utilising lightweight block encryption to secure health data	Lightweight encryption may raise security issues, especially when dealing with susceptible medical data
Islam et al. <sup>21</sup>	CLSTMN	Combines information from several sensors to provide fast diagnosis and thorough health monitoring	Several sensors combined could lead to more complicated systems and measurement error margins
Pachiyannan et al. <sup>22</sup>	CNN	Congenital heart disease (CHD) may be quickly and accurately diagnosed, which can help with early pregnancy detection	Its decision-making process might need to be explained, which is essential for medical diagnosis
Tehseen Mazhar <sup>23</sup>	Fuzzy	Enables early detection of diseases, improves physician decision-making, and enhances treatment accuracy	Requires a well-structured knowledge base, may struggle with uncertain data
Muhammad Amir Khan <sup>24</sup>	(M-ABC) and (KNN)	Increases prediction accuracy, reduces training time, and optimizes feature selection	High dependency on dataset quality, requires computational resources
Tehseen Mazhar <sup>25</sup>	IoMT	Enhances diagnostic precision, enables personalized treatment, reduces surgical errors, and improves patient monitoring	Ethical concerns, data privacy risks, and high implementation costs
Yazeed Yasin Ghadi <sup>26</sup>	MEC	Improves real-time decision-making, reduces latency, and enhances patient care efficiency	Security concerns, decentralization challenges, and infrastructure costs

**Table 1.** Summary of the reviewed approach.

### Types of attacks and security mechanisms

The communication layer in IoT healthcare faces various security threats. MAC layer attacks like collision and DoS are countered with error-correction codes and rate limitation. Network threats like sinkhole, Sybil, and wormhole are mitigated using authentication and redundancy checks. Transport and application layers face SYN flooding, buffer overflow, and logical errors, defended by client puzzles and trusted computing. Physical and privacy attacks like PUEA, jamming, and traffic analysis are secured using spread-spectrum, homomorphic encryption, and ongoing routing.

### Research gap

Heart disease is still a leading killer globally, and traditional diagnostic processes are typically ineffective and time-consuming, especially in low-resource settings. Due to the complexity of this system, health systems fail to diagnose diseases accurately and on time. Patients are at high risk, as most are treated late, and the dangers come from the many heart-related data points published by IoT devices. This research is proposed to solve these issues.

### Proposed methodology

Quantum Kernel-Enhanced Sardine Diffusion Attention Network for Heart Disease Prediction in IoT-Based Smart Health Monitoring (Qua-KSar-DCK-ArNet) is explained in this section.

Figure 2 shows the workflow diagram of Qua-KSar-DCK-ArNet. The system begins by utilizing an IoT-based sensor worn by the patient to continuously collect heart-related data such as ECG, heart rate, and other relevant metrics. The data is transmitted through a mobile phone and securely stored in a cloud-based database. After that, the collected heart-related data is divided into clusters using Quantum Clustering with k-Means to categorize this data according to the disease risk factors. After that, the clustering of the data is followed by the preprocessing using the Z-score Min–Max Normalization technique for better data quality for further analysis. Next is feature extraction, which employs the Fast Point Transformer to identify some of the significant features in data, such as arrhythmia and variability in heart rate. The main subnetwork of the system is the Quantum Kernel Sardine Diffusion Classical Kernel Attention Network (Qua-KSar-DCK-ArNet), which is used to classify processed data and determine the probability of heart disease.

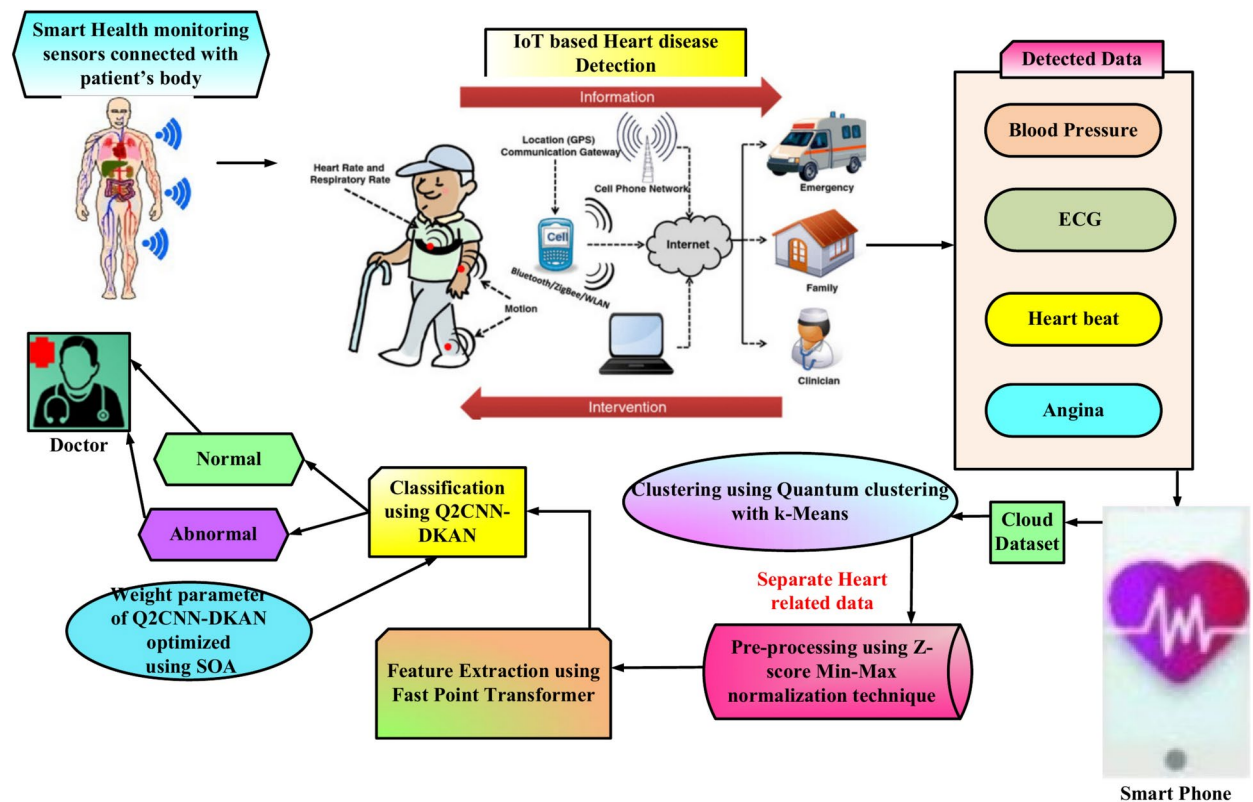


Fig. 2. Workflow diagram of Qua-KSar-DCK-ArNet.

## BEGIN

### 1. Collect Data:

- Gather health data (Blood Pressure, ECG, Heartbeat, Angina) from sensors
- Send data to the cloud via IoT

### 2. Pre-process Data:

- Normalize data using Min-Max scaling
- Remove noise and errors

### 3. Cluster Data:

- Use *k*-Means Quantum Clustering to group similar data

### 4. Extract Features:

- Identify key patterns using Fast Point Transformer

### 5. Classify Data:

- Use Q2CNN-DKAN to determine if the condition is Normal or Abnormal
- Optimize the model using SOA

### 6. Take Action:

- If Normal → No action needed
- If Abnormal → Alert doctor, family, and emergency services

### 7. Store and Display Data:

- Save results in the cloud
- Show results on a smartphone

END

Pseudo code of OpHyp-NASH-Gra-Net



## Data acquisition

The input dataset is from the IoT-based smart health monitoring sensors<sup>19</sup>. Then, this health data is stored in a cloud dataset. This section provides an overview of the dataset sourced from the UCI Machine Learning Repository, focusing on clinical records related to heart failure. The dataset comprises patient records for acute heart failure, featuring 11 clinical characteristics collected during the follow-up period. Of 299 records, 194 pertain to male and 105 to female patients. These IoT-detected health data are given to the Quantum clustering with k-Means<sup>27</sup> to separate heart-related data for detecting heart disease, and its explanations are given below:

## Clustering using quantum clustering with k-Means

Clustering using Quantum clustering with k-Means<sup>27</sup> is used for clustering for heart disease prediction. The procedure entails grouping related records by clustering a dataset of histories into clusters and allocating every entry to a cluster represented by a centroid. Given a dataset  $F = \{\vec{q}_1, \dots, \vec{q}_P\}$  of entries in which each entry is a  $Q$ -dimensional vector, the purpose of grouping is to allocate every record to one of the different clusters  $\{B_1, \dots, B_l\}$  denoted through centroid correspondingly for comparable records to get the same assignment based on a particular distance measure.

### *k*-Means and $\delta$ - *k*-Means

A popular technique for clustering is the *k*-Means algorithm, which uses a certain distance measure to assign each element to its nearest centroid until a predetermined convergence condition is satisfied. It usually uses the Euclidean distance, which is given by Eq. (1):

$$f(\vec{m}, \vec{n}) = \|\vec{m} - \vec{n}\|_2 = \sqrt{\sum_{j=1}^P (m_j - n_j)^2} \quad (1)$$

here are two-dimensional real vectors. The second step entails computing a new cluster centre for every cluster to update the centroids for the subsequent iteration.

The grouping distribution step is the main emphasis of this, and its classical complexity in time is where the quantity of centroids is, is the number of records, and is their dimension. The quantum variant of *k*-Means, known as quantum *k*-Means, incorporates noise in cluster assignment while preserving classical centroid update steps, simulating traditional methods in a quantum environment.

Then, the modernized version of the measured  $\delta$  - *k*-Means. Let  $\vec{b}$  be the centroid closest to the point. Then,  $\delta$  - *k*-Means defines the set of possible labels  $K_\delta(\vec{q})$  for  $\vec{q}$  is given in (2):

$$K_\delta(\vec{q}) = \left\{ m : \left| f^2(\vec{b}, \vec{q}) - f^2(\vec{b}_m, \vec{q}) \right| \leq \delta \right\} \quad (2)$$

When  $\delta = 0$ ,  $\delta$  - *k*-Means is comparable to the norm, *k*-means that since there is no consideration of uncertainty,  $K_\delta(\vec{q})$  contains only the centroid computed. Conversely, though a high level of permits means to contain centroids far from the minimal cluster label, the assignment rule randomly chooses a cluster label from the collection, adding additional noise to the process. While quantum *k*-means guarantees centroid assignments without uncertainty by reducing to standard *k*-means, using more significant  $\varepsilon$  values to achieve proper grouping creates noise.

### Quantum distance estimate

An amplitude-encoded quantum circuit for calculating the quantum Euclidean distance between two quantum states is shown in this study. The circuit makes use of an extra qubit that is entangled with the states. Between the two vectors, the Euclidean distance represents the likelihood of measuring the ancilla in a given state. Calculating the quantum Euclidean distance between several vectors in superposition can be done generally using this process. This work uses similar principles to define the quantum *k*-Means algorithm, allowing for the practical assignment of records to clusters based on proximity.

In summary, the clustering methodology aims to enhance heart disease prediction by leveraging classical and quantum techniques for efficient data clustering, leading to improved outcomes in predictive analytics. Then, this separated heart disease data are given to the pre-processing stage to remove unwanted noise, and its explanations are given below:

## Pre-processing using Z-score Min–Max normalization technique

This section uses the Z-score Min–Max normalization technique<sup>28</sup> to preprocess the input dataset for efficient heart disease prediction. The methodologies of data preparation for health data for modelling include data scaling techniques such as Z-score normalization and Min–Max normalization. Then, Z-score normalization normalizes the data to a mean of one mean and scales it according to the standard deviation. In contrast, Min–Max normalization scales the data to range over certain limits.

These approaches are used on several health indicators, including blood pressure levels, cholesterol and other health factors, including cardiovascular ones, to reduce biases and enhance the efficiency by which models are developed to predict heart diseases. Missing values should be dealt with, and certain characters, such as commas, should be deleted to keep the data clean. The comical ways of dealing with missing data include mean imputation forward and backward filling. The Z-score Min–Max normalization technique preserves the data from clutter and is vital for the correct heart disease evaluation. These approaches are instrumental in the

activation functions, such as sigmoidal functions and guarantee that each feature's contribution to the model is proportional. For efficient heart disease prediction, the Min-Max normalization scales the data using the following formulas (3–4):

$$V^P = \frac{v - \min(v)}{\max(v) - \min(v)} \quad (3)$$

$$i^p = \frac{i - \min(i)}{\max(i) - \min(i)} \quad (4)$$

where  $V^P$ ,  $i^p$  Represents the normalized value of a financial metric, xxx is the original value, and  $\min(i)$ ,  $\min(v)$  and  $\max(v)$ ,  $\max(i)$  these are the minimum and maximum values of the dataset, respectively. With daily/weekly/monthly data on patient's health, most of the vital signs are standardized so that they are on the same scale to enhance comparison. This process also reduces the tendency and bias that might be seen in the model, and the model's efficiency improves. Where  $i^p$  is denoted as the normalized value and  $i$  signifies the current stock price output. Then the input  $[J]_J$  and output  $[O]_0$  pre-processed output is given in Eqs. (5–6):

$$[J]_J = \begin{bmatrix} V^P \\ i^p \end{bmatrix} \quad (5)$$

$$[O]_0 = \begin{bmatrix} V^{ps} \\ i^{ps} \end{bmatrix} \quad (6)$$

Such techniques are used in the everyday record of the historical health data obtained from IoT-based monitoring systems, which ensures that each health parameter can be easily correlated with others to enhance the efficiency of the prediction model. Following this preprocessing stage, the data is then forwarded to the feature extraction stage to identify and extract the most relevant features for heart disease prediction, as explained below:

### Feature extraction based on fast point transformer

After the preprocessing stage, the data moves on to feature extraction to predict outcomes related to heart disease, including the chances of disease presence, its severity, and various patient risk factors. This step uses the network architecture known as the Fast Point Transformer (FPT)<sup>29</sup>, designed to integrate complex spatial patterns specific to HL datasets while preserving local and global context dependencies. The FPT performs feature extraction in three main steps: The three components of centroid-aware voxelization, LSA and centroid-aware de-voxelization as defined below:

#### Step 1 Centroid-aware Voxelization

The FPT is initially divided into units, with each health unit having patient coordinates, health features, and details of centroids. Because the positions are generally relative to the others, a centroid-to-point positional encoding is used to maintain semantic and context features. This encoding captures the relationship between data points and their respective units. Given an input set of health data, where  $M^{in} = \{(M_p, i_p)\}_{p=1}^P$ , where  $M_p$  represents the embeddings of the patient health metrics, the FPT voxelized this data into discrete units of the  $p$ -th patient health metrics and  $i_p$  represents the embeddings of the equivalent to patient health metrics, the FPT voxelizes this data into discrete units  $u = \{(u_i, f_i, a_i)\}_{i=1}^I$ . Each unit  $a_i$  includes its coordinates  $f_i$ , features  $f_i$ , and centroid coordinates  $a_i$ .

To preserve semantic and contextual information during voxelization, a centroid-to-point positional encoding  $f_p \in \mathbb{R}^{D_{enc}}$  is introduced. This encoding retains the relationship between health metrics and their corresponding units. The centroid of a unit is computed as the average of all data points within that unit, defined in Eq. (7):

$$f_p = \mu_{enc} (M_p - a_{i=\eta(p)}) \quad (7)$$

where  $a_{i=\eta(p)}$  is the centroid of the unit containing the health data  $M_p$ , and  $\mu_{enc}$  positional encoding maps the relative position to a higher-dimensional feature space. The centroid  $a_i$  is computed as the average of all patient data points within a unit, as shown in Eq. (8):

$$a_i = \frac{1}{|P(i)|} \sum_{q \in P(i)} M_p, P(i) \quad (8)$$

where  $P(i)$  denotes the set of patient data points within the  $i$ -th unit, and  $\eta(i)$  maps the patient data points  $p$  to its corresponding unit  $i$ .

The unit feature  $f_i \in \mathbb{R}^{D_{in}+D_{enc}}$  is then computed by concatenating the input patient data points  $M_p$  (e.g., blood pressure, cholesterol levels) and the positional encoding  $f_p$  and its equation is given in (9):

$$f_i = \Psi_{p \in P(i)}(i_p \oplus f_p) \quad (9)$$

where  $\oplus$  denotes vector concatenation and  $\Psi$  it is a permutation-invariant operation, such as averaging over the features in the unit.

#### Step 2 Lightweight Self-Attention (LSA)

Post-voxelization, the features are processed through the Lightweight Self-Attention (LSA) block. This mechanism refines the unit features by updating them with contextual information from neighboring units. In the best way possible, the LSA gathers essential information from the surrounding area, enhancing feature representations to predict heart disease rates better. This inevitably guarantees the computational efficiency of the mechanism while at the same time keeping the level of complexity controllable.

#### Step 3 Centroid-aware devoxelization

Finally, translations of the newly updated unit feature representations back to the original data space are obtained. This is done through positional encoding by matching unit-wise features to continuous health data points through centroid-to-point style. The resulting feature is computed using a Multilayer Perceptron (MLP) with the updated unit feature combined with positional encoding. Following the self-attention process, the updated unit features  $u = \{(u_i, f_i, a_i)\}'$  are devoxelized back into the heart disease outcome space to produce the output set  $P^{out} = \{(M_q, l_q)\}$  where  $l_q$  is the output feature associated with health metrics?  $M_q$ .

The devoxelization process uses centroid-to-point positional encoding.  $f_p$  To assign unit-wise features to continuous health data points. Specifically, the output feature  $o_q \in \mathbb{N}^{D_{out}}$  is computed using a Multilayer Perceptron (MLP) applied to the concatenation of the updated unit feature  $f'_{i=\eta(p)}$  and the positional encoding  $f_p$  and its equation is given in (10):

$$o_q = MLP(f'_{i=\eta(p)} \oplus f_p) \quad (10)$$

where  $f'_{i=\eta(p)}$  maps the question  $M_q$  to its corresponding unit  $i$ ,  $MLP$  maps the health data to its corresponding output features.

The method used in the FPT to perform feature extraction is particularly viable for producing feature representation with robust local and global dependencies. These features are subsequently used by the Quantum Kernel-Enhanced Sardine Diffusion Attention Network (Qua-KSar-DCK-ArNet) to predict heart disease outcomes, as detailed below.

### Heart disease detection using Quantum Kernel-Enhanced Sardine Diffusion Attention Network

After completing feature extraction using FPT, the extracted features are processed by the Quantum Kernel-Enhanced Sardine Diffusion Attention Network (Qua-KSar-DCK-ArNet). The hybrid network enhances learning by combining quantum and classical networks with nested convolutional layers for feature abstraction, multiple-dense nested layers for feature patterns, and an attention mechanism for representation. This innovative approach merges the Quantum-Classical Convolutional Neural Network (Q2CNN)<sup>30</sup> with the Diffusion Kernel

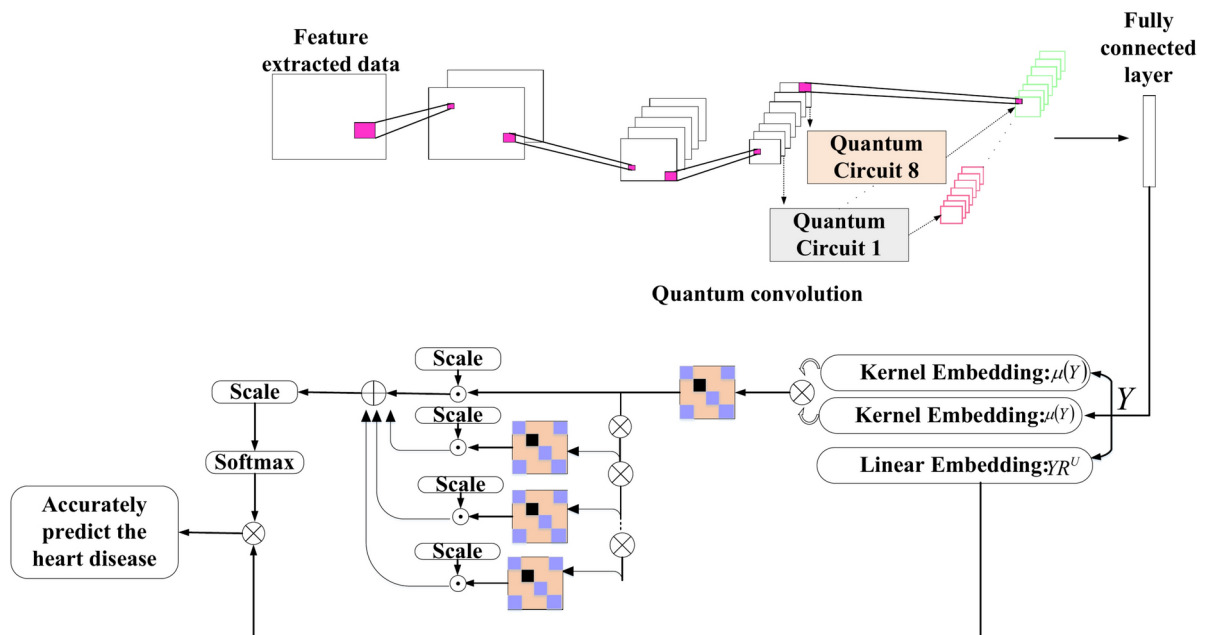


Fig. 3. Architecture of Qua-KSar-DCK-ArNet.



Attention Network (DKAN)<sup>31</sup>, optimizing parameters through the Sardine Optimization Algorithm (SOA)<sup>32</sup> to improve the accuracy of heart disease predictions, with further explanations provided below in Fig. 3.

#### Quantum-classical convolutional neural network

The Quantum-Classical Convolutional Neural Network (Q2CNN) integrates quantum computing principles with classical deep learning methods to enhance computational efficiency and predictive accuracy in detecting heart disease from IoT intelligent environments. In this hybrid model, circuits process extensive health data from multiple IoT devices and find latent correlations that classical networks may not see. The first type of layers, the classical convolutional layers, are oriented towards feature extraction and pattern recognition of the health metrics. The second type of layers, quantum layers, bring entanglement and superposition into play while incorporating them into more expressive data representations. This amalgamation of quantum and classical parts allows the Q2CNN to improve the performance of relevant tasks such as heart disease detection with perfect precision while reducing the computational parameter of the structure.

##### a. Quantum state preparation

To process the classical health data to be fed to quantum processing systems, real-time parameters from IoT devices are collected. PPM also defines conditions such as body functions, including blood pressure, heart rates and activities, environments, and health records. This data is normalized and encoded into a quantum-compatible format. The data vector  $e$  is represented as shown in Eq. (11):

$$e = [e_1, e_2, \dots, e_{q^2}] \quad (11)$$

This vector  $f$  is converted into angle values  $\vartheta$  as (12):

$$\vartheta = \pi f \quad (12)$$

where  $\vartheta = [\vartheta_1, \vartheta_2, \dots, \vartheta_{q^2}]$ , these angles  $\vartheta$  is a function mapping the health metrics to angle values suitable for quantum state preparation. For a quantum system with  $q^2$  input the quantum state  $|\Psi_{data}\rangle$  is created by applying a rotation gate  $s_y$  to the initial quantum state  $|0\rangle^{\otimes p^2}$  as (13),

$$|\Psi_{data}\rangle = \otimes_{l=1}^{p^2} s_y(\vartheta_l) |0\rangle \quad (13)$$

This state  $|\Psi_{data}\rangle$  represents the rotation gate parameterized by  $|\Psi_{data}\rangle$ .

##### b. Quantum-classical convolution operation

In the quantum-classical convolutional layer, a quantum kernel  $u(\varphi)$  is designed using quantum circuits to transform the quantum state  $|\Psi_{data}\rangle$ . The kernel parameters  $(\varphi_1, \varphi_2, \varphi_3, \varphi_4, \varphi_5)$  control the quantum operations applied to  $|\Psi_{data}\rangle$ . The transformation performed by this kernel is described by Eq. (18):

$$|\Psi_{conv}\rangle = R(\varphi) |\Psi_{data}\rangle \quad (14)$$

where  $R(\varphi) = u_1(\varphi)u_2(\varphi)\dots u_l(\varphi)$  represents the sequence of unitary operations applied by the quantum kernel to extract relevant features from the health data.

##### c. Quantum pooling and measurement

To reduce the dimensionality of the quantum convolution results, a quantum pooling gate  $O$  consisting of controlled-NOT gates, is applied. This pooling operation maps the convolution results to specific inputs and is defined in Eq. (15):

$$|\Psi_{pool}\rangle = O |\Psi_{conv}\rangle \quad (15)$$

After pooling, measurements are performed on the quantum state in the Z-basis to obtain expected values, as shown in Eq. (16):

$$T = \langle \Psi_{data} | O^+ Z_1 O V(\varphi) | \Psi_{data} \rangle \quad (16)$$

where  $Z$  is the operator applied to different inputs and  $O$  represents the pooling gate.

The resulting vector  $T$ , with dimensions  $1 \times (q - 1)^2$ , is the fed attention network for further processing. The resulting vector  $T$ , with reduced dimensions, serves as input to the attention network for further processing, enabling more accurate predictions of heart disease outcomes based on data from the IoT bright environment. Then, to improve the accuracy, the Q2CNN is integrated with DKAN, and its explanations are given below.

#### Diffusion kernel attention network (DKAN)

In the context of heart disease prediction from intelligent IoT devices, conventional self-attention mechanisms face challenges in effectively modeling indirect relationships between various health-related variables, such as heart rate, blood pressure, physical activity, and lifestyle factors. Traditional methods typically focus on pairwise relationships without indirectly capturing the more profound and complex interactions that may exist among these variables. This limitation can reduce the accuracy of heart disease prediction models and hinder the identification of hidden patterns influencing patient health. To address this problem, this study proposes the Diffusion Kernel Attention Network (DKAN) as a subroutine of the proposed Q2CNN. The DKAN enhances how the model computes both first-degree and other-level connections between health-related issues to make better decisions and predict better heart disease risk. The DKAN uses the random walk process to capture adjacent and non-adjacent correlations between various measures of health gathered from connected intelligent IoT devices, including pulse rate, movement and eating patterns.

This random walk is characterized by a transition matrix  $R \in \mathbb{R}^{P \times P}$ , representing the chance of passing from one health variable to another within the data set. The transition matrix  $R$  is computed by normalizing the kernel attention  $K$  matrix derived from the self-attention mechanism, ensuring that relationships between variables are encoded effectively, and its equation is given in (17):

$$R_{pq} = \frac{T_{pq}}{\sum_{p=1}^P T_{pq}} \quad (17)$$

where  $R_{pq}$  denotes the probability of transitioning from variable  $p$  to variable  $q$  in one step. By performing  $S$ . The sum ensures proper normalization of the matrix. Thus, the ordering of diffusion kernel attention is combined using the DKAN through a weighted sum. This approach allows the model to sort out indirect relationships between health variables of different levels while considering all the discussed levels of interaction. The weights ensure. This transition matrix forms the basis for modeling the direct and indirect relationship between health metrics.

Since a given clinical quantity may depend on several other clinical quantities, and the clinical quantities may, in turn, depend on other clinical quantities in a two-tiered fashion, the multi-step random walks of the DKAN create awareness of much broader liaisons in the health-related network beyond just duality. An  $A$   $k$ -step random walk expands the spectrum of relations considered, relationships between the variables that are connected not directly. The condition of the heart disease prediction network following  $k$ -step random walks is represented in Eq. (18):

$$T_{fk} = R^K T \quad (18)$$

where  $R^K$  denotes the  $T$ -th power of the transition matrix  $R$ . This formulation effectively captures the state of the network after considering interactions up to  $k$  orders of connectivity. In the context of predicting heart disease, this enables the model to gain a deeper insight into the intricate and nuanced factors that contribute to the risk of heart disease, as outlined in Eq. (19):

$$T_R = \sum_{q=0}^K \alpha_q R^q T \quad (19)$$

here are adaptive weights learned during training to optimize the model's performance. The DKAN combines the different orders of diffusion kernel attentions into a weighted sum to integrate them. The above method helps the model to focus on the relative indirect couplings between health variables at different interaction levels while incorporating all of them. The weights add up some inputs to have a normalized weight for various layers of indirect relationships, so working with the DKAN improves the ability to predict the risks of heart disease. In an attempt to increase accuracy and decrease the error rates, computational complexity and cost, the weight parameters Of Q2CNN-DKAN are fine-tuned with the Sardine Optimization Algorithm (SOA) and its explanations are given below:

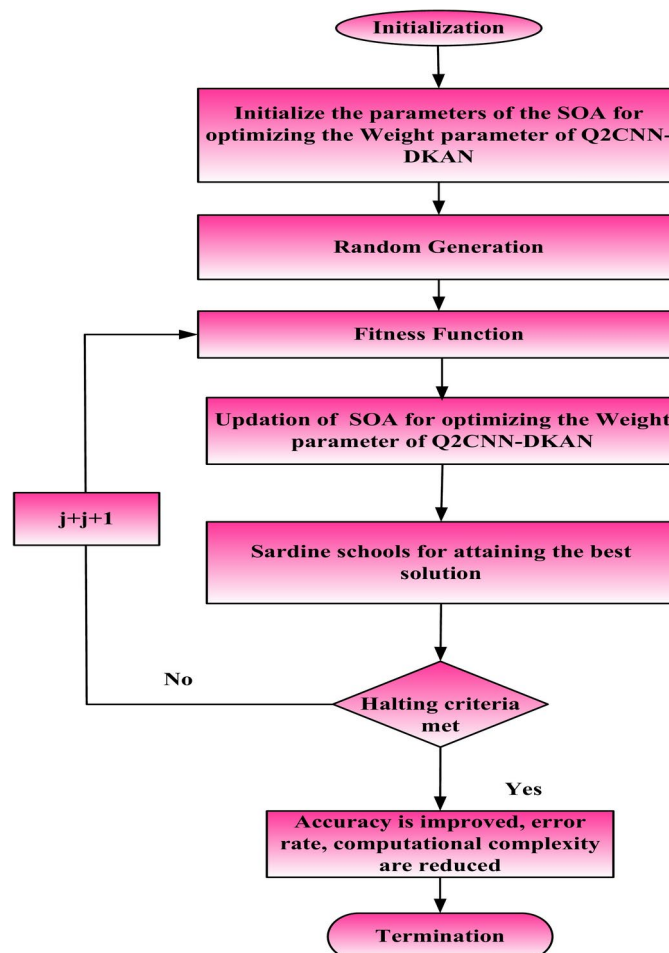
#### Optimization with Sardine Optimization Algorithm (SOA)

To further enhance the accuracy of heart disease prediction from smart IoT devices, the weight of the proposed model, namely, Q2CNN-DKAN, is fine-tuned by employing a metaheuristic named the Sardine Optimization Algorithm (SOA). This optimization aids in enhancing the model that suppresses the risk levels as well as the health condition of the patients by reducing the rates of mispredictions and computations of the model. In particular, the SOA strategy allows the exploration and exploitation of the given space and the search for such a set of model parameters, which will achieve the maximum accuracy in defining the risks of

heart diseases. Figure 4 shows the flow chart of Qua-KSar-DCK-ArNet. The step-by-step procedure of Qua-KSar-DCK-ArNet are given below:

1. Initialize parameters of the SOA for optimizing the weight parameter of Q2CNN-DKAN
2. Randomly generate initial solutions
3. Evaluate the fitness function for each solution
4. Repeat until halting criteria are met:
  - a. Update SOA for optimizing the weight parameter of Q2CNN-DKAN
  - b. Perform Sardine School optimization to find the best solution
  - c. Check halting criteria:
    - If met: Proceed to step 5
    - Else: Increment iteration counter and repeat step 4
5. Accuracy is improved, error rate and computational complexity are reduced
6. Terminate the algorithm

#### Pseudo code of Qua-KSar-DCK-ArNet



**Fig. 4.** Flow chart of Qua-KSar-DCK-ArNet.

#### Step 1 Initialization

The optimization process begins with the random distribution of sardine schools, representing a potential solution. Every solution is connected with a set of parameters, such as the weight of the Q2CNN-DKAN measure or weights of the network components and the attention layers in the Q2CNN-DKAN architecture. These parameters are crucial to span the risk of heart disease indicators, including physiological measures and behavioral aspects.

#### Step 2 Random generation

During this phase, the random generation of the school of sardines is the first step in exploring the SOA space to ensure proper exploration. The purpose here is to identify sets of parameters that would exhibit the activity of health data, including voltage signals that represent HRV, systolic/diastolic BP, etc. In the implementation process, the algorithm will go through most of the possible parameter values to identify settings that enhance the reliability of heart disease predictions.

#### Step 3 Fitness function

Through the measures of the fitness function, the ability of the Q2CNN-DKAN model is evaluated based on the accuracy, error rate and amount of time it takes for the model to make predictions. The purpose is to minimize the loss function, which aids in improving the forecasts of risk factors of heart disease in the model. The fitness function can be expressed as in Eq. (20):

$$\text{Fitness function} = \text{Optimize}(\alpha) \quad (20)$$

#### Step 4 Sardine schools for attaining the best solution

The optimization process involves the movement and transformation of sardine schools. Operators, such as the “oval operator,” are used to iteratively update each sardine’s position in the solution space, guiding the search toward the optimal solution. The sardine schools are updated based on adjacent odd and even dimensions  $(y_p, x_f)$ , adjusting their positions according to the school’s radius and optimization settings. This movement follows these Eqs. (21–23):

$$\begin{cases} \chi = \text{rand}() \times 2\Pi \\ y_p = q \cos(\delta_l + \delta) \\ x_f = q \sin(\delta_l + \delta) \end{cases} \quad (21)$$

$$\left\{ \sigma_l \mid \sigma_l = l \frac{2\Pi}{TSize} \mid 1 \leq l \leq TSize \cap l \in W \right\} \quad (22)$$

$$q = \frac{t \times \text{SearchScope}}{6} \times \left| \frac{\text{UpLimit} - \text{LowLimit}}{2} \right| \times e^{-5 \times \left(\frac{k}{K}\right)^2} \quad t = 1, 2, 3, \dots, 6 \quad (23)$$

An odd dimension is an even dimension, and it is school size. Returns a random number from 0 to 1.  $q$  is the radius of the school,  $K$  is the maximum number of evaluating sardines (i.e., the maximum number of using fitness function), and is the present number of evaluating sardines. With  $k$  increasing, it decreases gradually. Besides, there  $\text{UpLimit}$  and  $\text{LowLimit}$  are the upper and lower limits in each dimension. It is a setting parameter that is equal to 30%. Worth noting that it  $\frac{t \times \text{SearchScope}}{6}$  has six values. A set parameter is set to 30%, and each school randomly selects a value from six values (5%, 10%, 15%, 20%, 25%, 30%). These generated points are uniformly distributed around the center  $B$ ,  $\alpha$  representing the optimization parameters needed to achieve the best solution (improving accuracy, reducing error rates, and minimizing computational complexity).

#### Step 5 Termination

The optimization process continues until the termination criteria are met, when the optimal solution is identified, or a pre-defined number of iterations is reached. The best solution corresponds to the parameters that maximize the model’s accuracy in predicting heart disease risk. By optimizing the regularization coefficients, the model effectively balances minimizing error rates and computational complexity, thus enhancing predictive performance for heart disease risk assessment based on data collected from intelligent IoT devices.

This manuscript presents the Quantum Kernel-Enhanced Sardine Diffusion Attention Network (Qua-KSar-DCK-ArNet) for heart disease prediction using IoT-based sensors. The system collects heart data, clusters it via Quantum Clustering, preprocesses it with Z-score Min–Max normalization, and extracts features using the Fast Point Transformer. It combines quantum and classical techniques for accurate predictions, facilitating timely clinical decision-making.

## Result and discussion

This section presents the results and discussions of Quantum Kernel-Enhanced Sardine Diffusion Attention Network (Qua-KSar-DCK-ArNet) for heart disease prediction using IoT-based sensors. Here, the concept of An IoT-Based Patient Monitoring System is presented.

## Dataset descriptions

The input dataset is taken from the IoT-based smart health monitoring sensors. Then, this health data is stored in a cloud dataset. This section provides an overview of the dataset sourced from the UCI Machine Learning Repository, focusing on clinical records related to heart failure. The dataset comprises patient records for acute heart failure, featuring 11 clinical characteristics collected during the follow-up period. Of 299 records, 194 pertain to male and 105 to female patients. Among them, 70% of data are taken for training and 30% for testing. The training process involves initially dividing the dataset into training, validation, and test sets (70%, 15%, and 15% respectively). Hyperparameters include a learning rate of 0.001, batch size of 32, number of epochs set to 100, quantum clustering k-value of 5, and diffusion optimization iterations set at 50. The Adam optimizer was used to minimize the cross-entropy loss function. Evaluation criteria encompass accuracy, precision, recall, F1-score providing comprehensive metrics for model performance assessment.

## Performance metrics

The performance metrics such as Accuracy, precision, recall, sensitivity, Computational complexity, computational time, and error rate analysis of neural network models such as IoT-based ambulatory blood pressure monitoring (IABPM)<sup>16</sup>, Extra Tree Classifier-Convolutional Neural Network (ET-CNN)<sup>17</sup>, Bidirectional Long Short-Term Memory (Bi-LSTM)<sup>18</sup>, Multilayer Perceptron Neural Network (MPNN)<sup>19</sup>, Body Area Network and Personal Area Network (BAN-PAN)<sup>20</sup>, Convolutional Long Short-Term Memory Networks (CLSTMN)<sup>21</sup> and Convolutional Neural Network (CNN)<sup>22</sup> respectively are analyzed.

Table 2 shows the performance metrics equations are given below:

True Positive  $\lambda$ : Normal correctly calculates as usual.  
 False Positive  $h$ : Normal incorrectly calculates as abnormal.  
 False Negative  $\ell$ : Abnormal incorrectly calculates, abnormal as usual.  
 True Negative  $\xi$ : Abnormal correctly calculates, abnormal.

## Performance analysis

The examinations of the introduced method are clarified here.

Figure 5 shows the examination of Training Accuracy over Epochs with 99.9% accuracy. By evaluating the method's output using the training dataset, training accuracy offers invaluable information about how effectively the model assimilates the info it has encountered. A critical measure of the model's training effectiveness is the loss function. A smaller loss function value indicates better model performance.

Figure 6a shows the Convergence plot over Epochs with existing methods. The convergence plot across epochs shows the performance of the proposed model compared with existing methods. It illustrates the consistency and quickness of the learning process by tracking accuracy or loss over iterations. While a fast convergence indicates that the model learns the best patterns quickly, a steady curve indicates robustness. When our model is compared with current methods, it shows speedier convergence, lower loss, more robust optimization, better generalization, and higher accuracy over alternatives in complex tasks.

Figure 6b compares the ROC curves with existing methods. The proposed model's ROC curves show improved sensitivity and specificity compared to existing processes, resulting in more accurate predictions, fewer misclassifications, and an overall diagnostic or detection accuracy improvement.

Figure 7a shows the Accuracy across epochs (stacked). The suggested Qua-KSar-DCK-ArNet routinely beats alternative techniques such as IABPM, EFCMN, BI-LSTM, MPNN, and BAN-PAN over several epochs (100–500), as shown by the accuracy graph. This suggests that the proposed model performs exceptionally well in

Performance metrics	Equations (24–31)
Accuracy	$\frac{\lambda + \ell}{\lambda + \xi + h + \ell} \quad (24)$
Precision	$\frac{\lambda}{\lambda + h} \quad (25)$
Recall	$\frac{\lambda}{\lambda + \ell} \quad (26)$
Specificity	$\frac{\ell}{\ell + h} \quad (27)$
Sensitivity	$\frac{\lambda}{\lambda + \ell} \quad (28)$
F1-score	$2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (29)$
Disease Prevalence	$\frac{\lambda + \ell}{\lambda + \xi + h + \ell} \quad (30)$
Negative predict value	$\frac{\lambda}{\lambda + h} \quad (31)$

**Table 2.** Performance metrics equations.



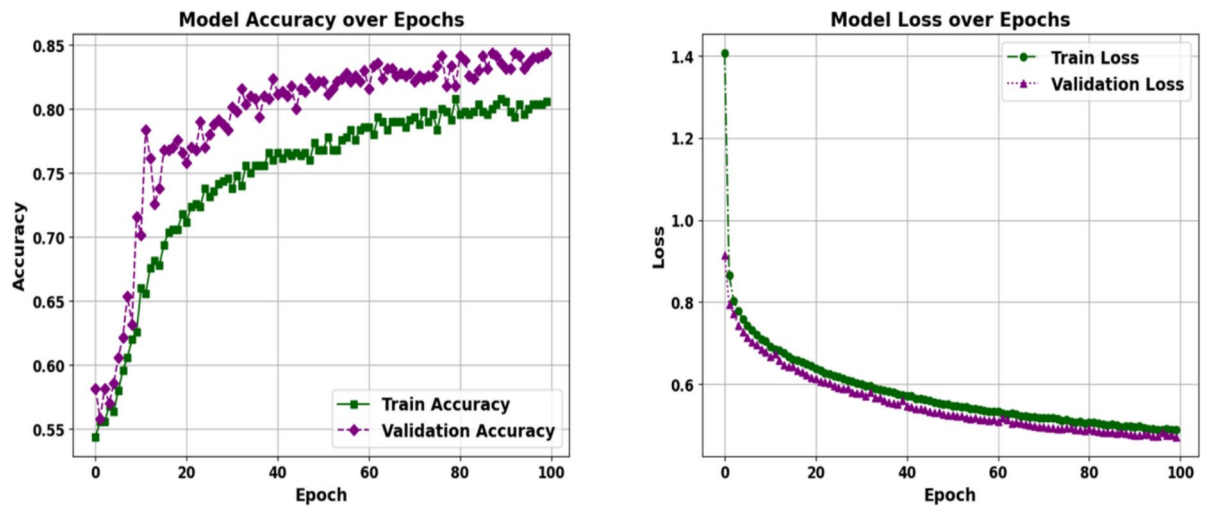


Fig. 5. **a** Training Accuracy, **b** Training loss function over Epochs for Qua-KSar-DCK-ArNet.

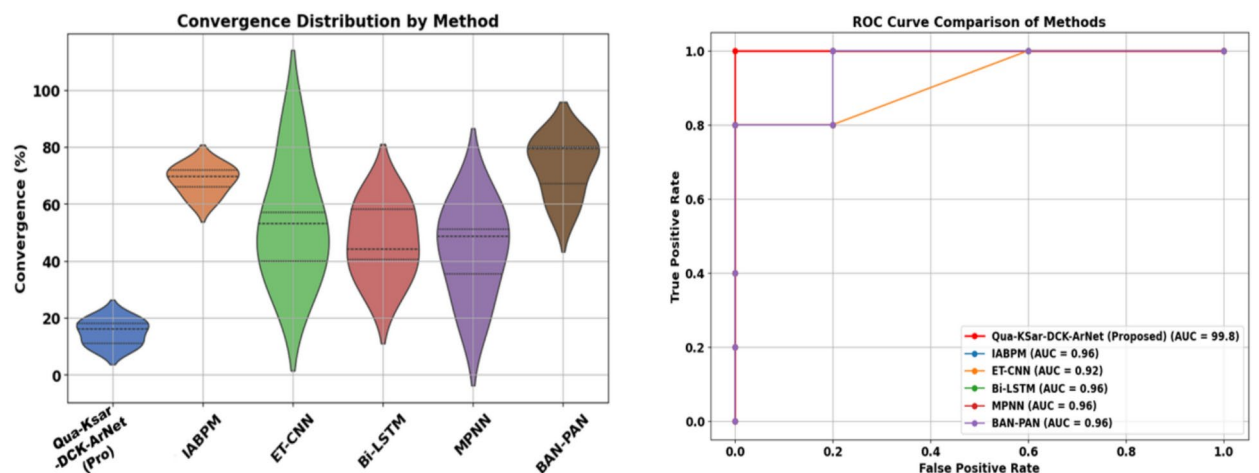


Fig. 6. **a** Convergence plot over Epochs and **b** ROC curves comparison with existing methods.

predicting heart illness with its excellent classification performance. It is essential for efficient health monitoring in the Internet of Things-based intelligent systems.

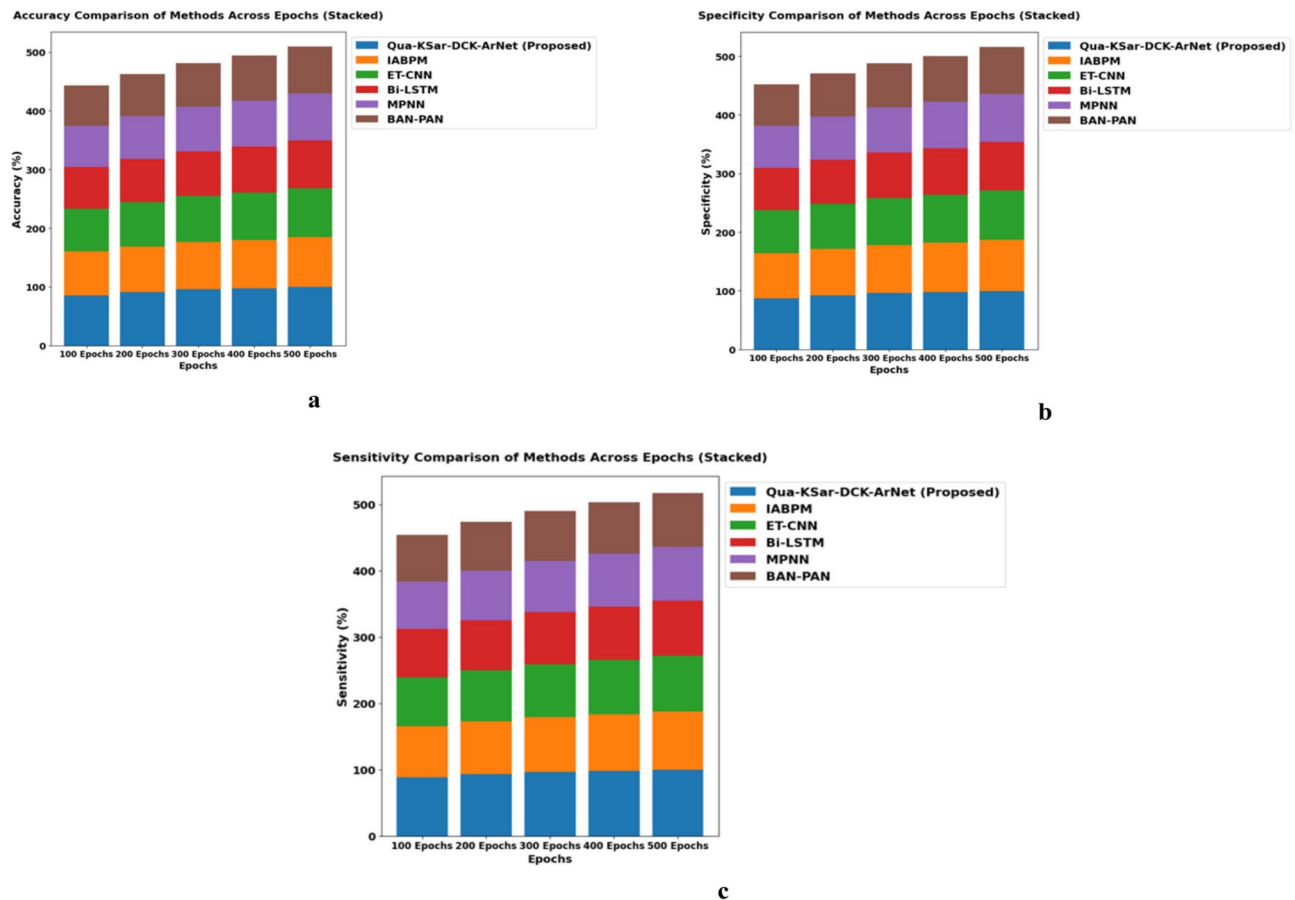
Figure 7a shows the Specificity across epochs (stacked). Compared to IABPM, EFCMN, BI-LSTM, MPNN, and BAN-PAN, Qua-KSar-DCK-ArNet performs better overall epochs and maintains fewer false positives, according to the specificity comparison. In IoT-based health monitoring systems, high specificity guarantees the system correctly detects those without cardiac disease, minimizing needless warnings.

Figure 7a shows the Sensitivity across epochs (stacked). The suggested Qua-KSar-DCK-ArNet leads consistently in recognizing true positives across several epochs in terms of sensitivity. To maintain patient safety in IoT-based health systems, high sensitivity is essential for the prompt and accurate diagnosis of cardiac disease, proving the model's efficacy in decreasing missed diagnoses.

Figure 8a shows the F1 scores across epochs (stacked). Higher F1 scores are routinely achieved by the proposed Qua-KSar-DCK-ArNet, demonstrating its superior ability to handle unbalanced heart disease datasets. This guarantees accurate forecasts in IoT-based intelligent health monitoring systems, where it's essential to reduce false positives and negatives.

Figure 8b shows the recall across epochs (stacked). The recall graph shows the accuracy of various approaches to identifying heart disease patients. With its consistent recall leadership, Qua-KSar-DCK-ArNet efficiently catches true positives, minimizing the number of missed heart disease diagnoses. High recall is crucial in IoT-based health systems to guarantee early detection and prompt medical actions.

Figure 8c shows the error rate across epochs (stacked). The Qua-KSar-DCK-ArNet approach, as compared to alternative methods, consistently maintains a reduced error rate over several epochs, as seen by the error rate comparison. This illustrates its resilience in lowering false positives and negatives, essential for precise and trustworthy heart disease prediction in Internet of Things-based intelligent health monitoring systems.



**Fig. 7.** **a** Accuracy, **b** Specificity and **c** Sensitivity Comparison of existing methods across epochs.

Table 3 and Fig. 9 shows the performance metrics of various models in detecting abnormal and normal cases. The proposed Qua-KSar-DCK-ArNet demonstrates exceptional performance, achieving the highest accuracy (99.98%), precision (99.34%), recall (99.56%), and F1-score (96.34%). This indicates its effectiveness in distinguishing between abnormal and typical cases, significantly outperforming other models in the analysis.

Table 4 and Fig. 10 show the Positive Predictive Value (PPV) percentages for different approaches across multiple data points (DP) displayed in the table. With near-perfect PPV values, the suggested Qua-KSar-DCK-ArNet consistently beats the other methods, particularly with larger datasets (99.79% for 567 records and 99.79% for 5000 records). While techniques like IABPM, ET-CNN, Bi-LSTM, and MPNN exhibit lower and more variable PPV, indicating decreased accuracy in positive categorization, BAN-PAN performs well.

Table 5 and Fig. 11 show that the proposed Qua-KSar-DCK-ArNet model is among the models whose Negative Predictive Value (NPV) percentages are displayed across various data points (DP) and record counts. Among other models, Qua-KSar-DCK-ArNet is particularly good at identifying negative cases in prediction tasks; it consistently outperforms other models and achieves the highest NPV across all data sets.

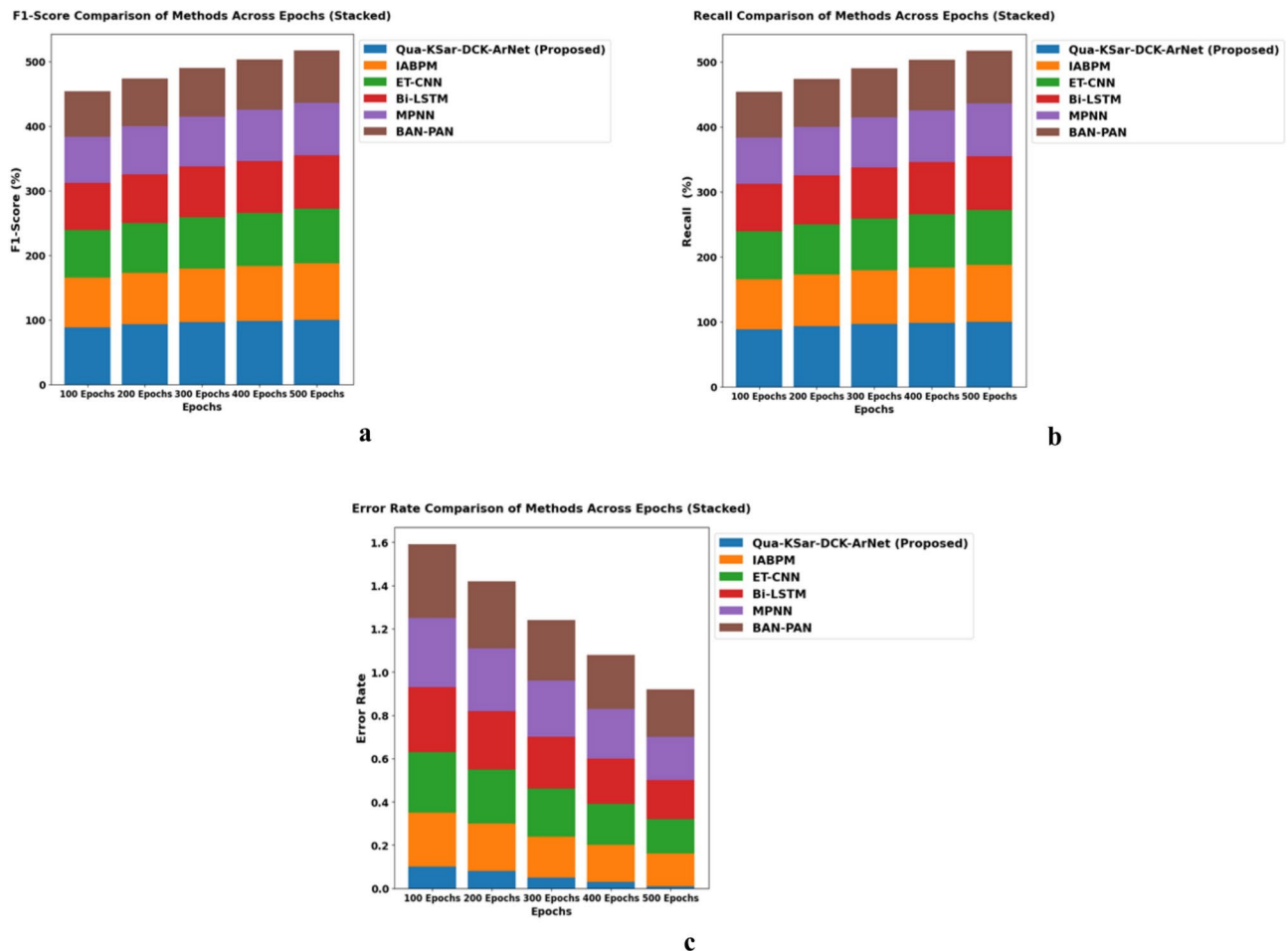
Table 6 and Fig. 12 compare the performance of the proposed Qua-KSar-DCK-ArNet with existing methods across various metrics. Qua-KSar-DCK-ArNet achieves superior results in accuracy (99.97%), recall (98.47%), and precision (99.59%), along with the lowest Mean Squared Error (0.01) and computational costs. These results highlight its effectiveness and efficiency in heart disease detection, outperforming all other methods.

Table 7 and Fig. 13 present a comparison of various optimization techniques used for training neural networks, highlighting their fitness regarding training execution duration and accuracy. These methods include the Sea-Horse Optimizer Algorithm (SHOA), Green Anaconda Optimization Algorithm (GAOA), Greylag Goose Optimization Algorithm (GGOA), Triangulation Topology Aggregation Optimizer Algorithm (TTAOA), Gorilla Troops Optimizer Algorithm (GTOA) and the Quantum Kernel-Enhanced Sardine Diffusion Attention Network (Qua-KSar-DCK-ArNet). Notably, the proposed Qua-KSar-DCK-ArNet attains the highest training accuracy of 99% in the experiment. 89%, while also highly efficient, with an execution time as low as 0.1 s.

### Statistical analysis of the proposed method vs. existing methods

The effectiveness of the projected method is evaluated using five statistical tests: the Shapiro–Wilk test (SW test), Wilcoxon Signed-Rank Test (WSR), Friedman Test (FT), Kruskal–Wallis H-test, and Kolmogorov–Smirnov test (KS test).

#### a. SW Test



**Fig. 8.** **a** F1-Score **b** Recall and **c** Error Rate Comparison of existing methods across epochs.

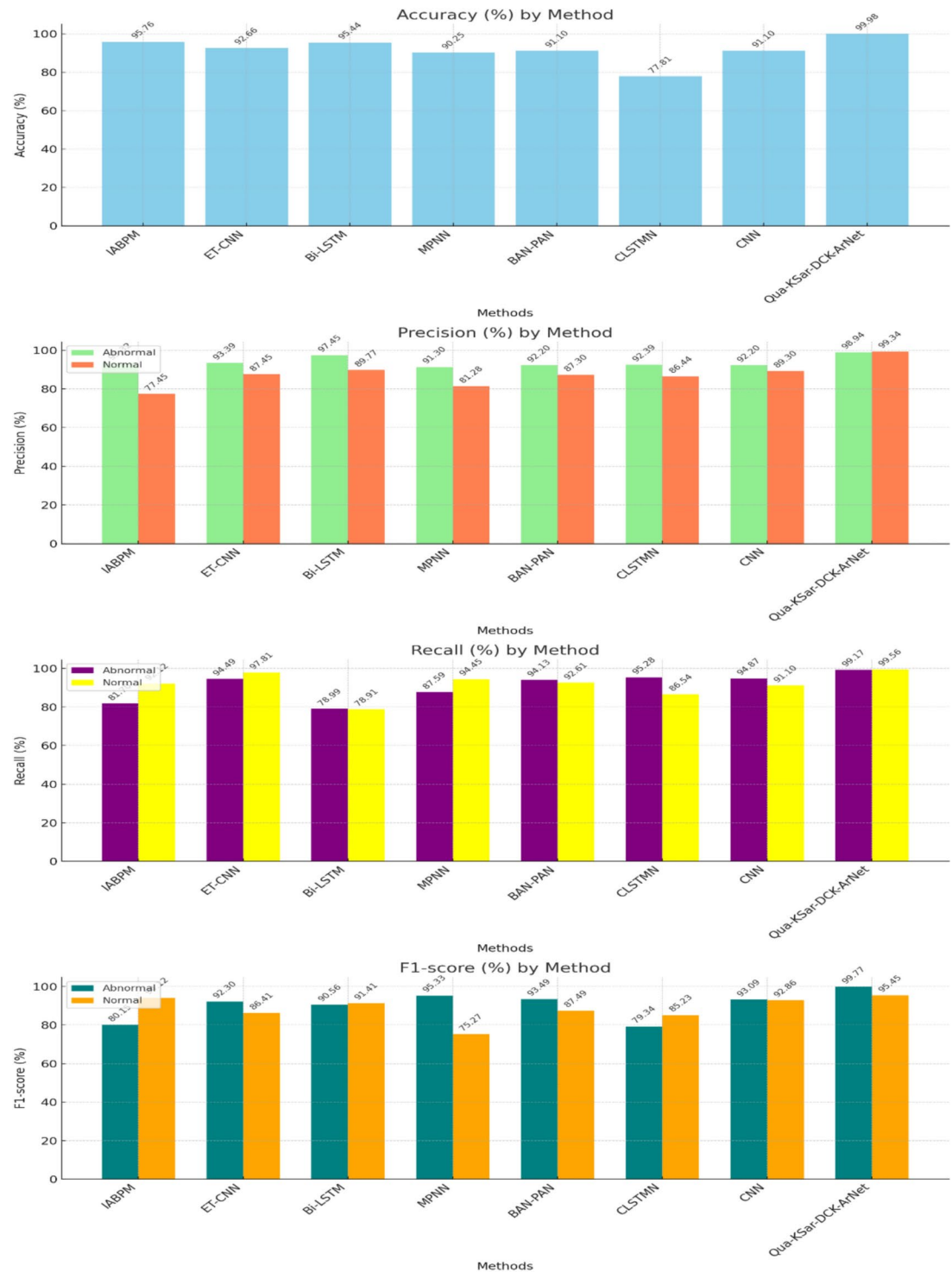
Methods	Accuracy		Precision		Recall		F1-score	
	Abnormal	Normal	Abnormal	Normal	Abnormal	Normal	Abnormal	Normal
IABPM	95.76	93.32	77.45	81.70	92.12	80.15	94.12	91.25
ET-CNN	92.66	93.39	87.45	94.49	97.81	92.30	86.41	95.20
Bi-LSTM	95.44	97.45	89.77	78.99	78.91	90.56	91.41	93.23
MPNN	90.25	91.30	81.28	87.59	94.45	95.33	75.27	90.17
BAN-PAN	91.1	92.2	87.30	94.13	92.61	93.49	87.49	96.49
CLSTMN	77.81	92.39	86.44	95.28	86.54	79.34	85.23	67
CNN	91.10	92.2	89.30	94.87	91.10	93.09	92.86	95.76
Qua-KSar-DCK-ArNet (Proposed)	99.98	98.94	99.34	99.17	99.56	99.77	95.45	96.34

**Table 3.** Precision, F1-score Accuracy and Recall for abnormal and normal.

The Shapiro–Wilk test, or SW test, can be used to determine a sample's normal distribution. It computes a W statistic to compare the sample distribution to a normal distribution. A p-value of less than 0.05 indicates statistical significance and suggests that the data may not follow a normal distribution.

#### b. WSR Test

The Wilcoxon Signed-Rank Test (WSR), a non-parametric statistical test, compares matching data. In cases when the data are not regularly distributed, it ascertains whether the median difference between paired observations is zero. It rates them to see if there is a substantial deviation of the absolute differences from zero.



**Fig. 9.** Accuracy, Precision, Recall and F1-Score Comparison of existing methods.

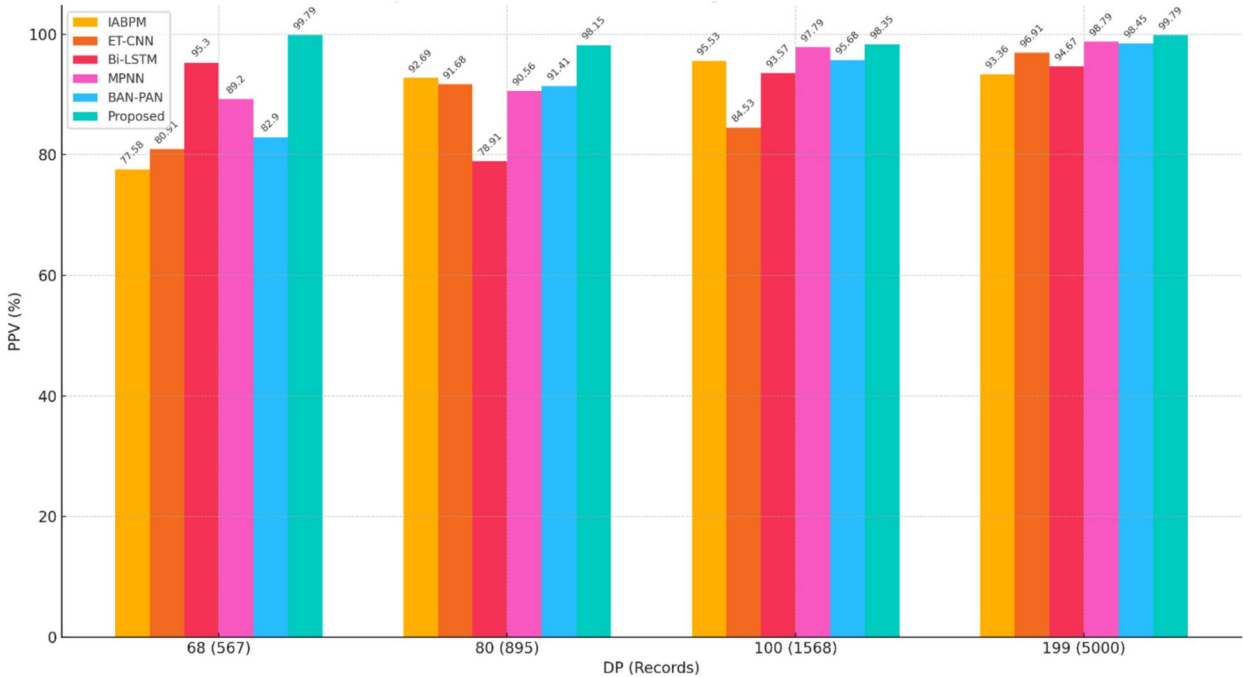
#### c. Friedman Test (FT)

A non-parametric statistical tool called the Friedman technique (FT) can be used to find differences in treatment between test runs. It is used in matched-subject designs and repeated measures to determine whether there are statistically significant differences in rankings between related groups when the data are ordinal or non-normally distributed.

#### d. Kruskal–Wallis H-test

DP	Records	PPV (%)					
		IABPM	ET-CNN	Bi-LSTM	MPNN	BAN-PAN	Qua-KSar-DCK-ArNet (Proposed)
68	567	77.58	80.91	95.3	89.2	82.9	99.79
80	895	92.69	91.68	78.91	90.56	91.41	98.15
100	1568	95.53	84.53	93.57	97.79	95.68	98.35
199	5000	93.36	96.91	94.67	98.79	98.45	99.79

**Table 4.** Analysis of disease prevalence using positive predicted values (PPV).



**Fig. 10.** Analysis of disease prevalence using positive predicted values (PPV).

DP	Records	NPV (%)					
		IABPM	ET-CNN	Bi-LSTM	MPNN	BAN-PAN	Qua-KSar-DCK-ArNet (Proposed)
68	567	93.55	90.47	91.68	94.57	95.67	98.95
80	895	90.47	91.36	92.65	95.30	96.44	99.90
100	1568	92.99	92.58	93.45	96.79	97.58	98.55
198	5000	95.79	93.22	94.00	97.30	98.45	97.70

**Table 5.** Analysis of disease prevalence using Negative Predicted values (NPV).

Based on ranking data, the Kruskal–Wallis H-test is a nonparametric statistical procedure used to evaluate whether there is a statistically significant difference between three or more independent groups. It is appropriate for ordinal or nonnormally distributed data since it considers whether the group median ranks fluctuate.

e. Kolmogorov–Smirnov Test (KS Test)

The Kolmogorov–Smirnov Test (KS) determines whether a sample conforms to a given distribution by comparing the empirical distribution function to the theoretical distribution. It establishes the degree to which these roles vary from one another. A high KS statistic suggests that the sample does not fit the expected distribution and indicates a significant difference.

Table 8 and Fig. 14 present the statistical analysis of the proposed Qua-KSar-DCK-ArNet model, showcasing its superior performance compared to existing methods. The p-values below 0.001 in tests like the F-test and Wilcoxon Signed-Rank indicate high statistical significance. The model also has the lowest Variance Inflation



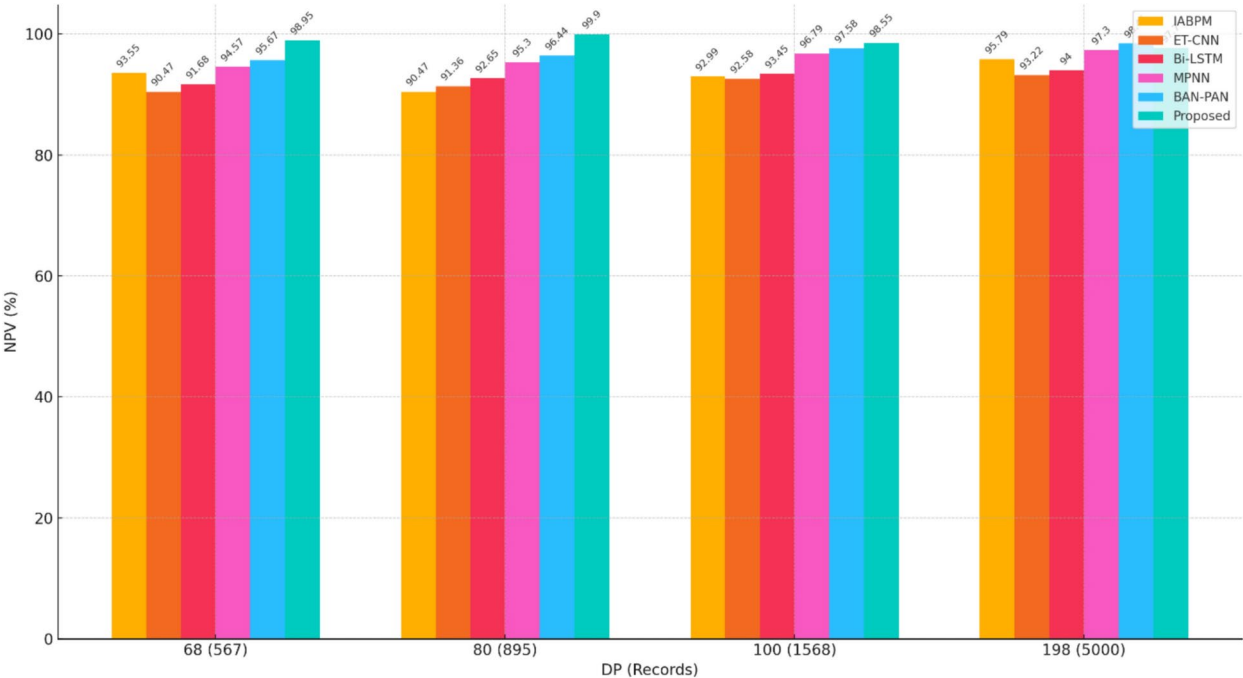


Fig. 11. Analysis of disease prevalence using Negative Predicted values (NPV).

Metrics	IABPM <sup>16</sup>	ET-CNN <sup>17</sup>	Bi-LSTM <sup>18</sup>	MPNN <sup>19</sup>	BAN-PAN <sup>20</sup>	CLSTMN <sup>21</sup>	CNN <sup>22</sup>	Qua-KSar-DCK-ArNet (Proposed)
Accuracy	98.95	95.28	94.42	95.14	96.80	96.65	97.58	99.97
Recall	95.1	85.36	80.47	75.74	85.39	93.73	91.53	98.47
Precision	64.83	75.65	85.73	85.64	75.84	91.27	92.46	99.59
Specificity	86.50	89.23	85.84	80.36	75.16	88.27	77.43	99.27
F1-score	90.66	88.53	87.47	85.35	85.40	92.28	94.76	99.46
MSE	0.06	0.05	0.04	0.05	0.03	0.07	0.02	0.01
MAE	0.05	0.05	0.02	0.06	0.04	0.05	0.02	0.01
RMSE	0.75	0.90	0.85	0.80	0.80	0.81	0.86	0.65
AAE	0.07	0.06	0.02	0.05	0.05	0.04	0.03	0.014
Computational cost	0.25	0.30	0.33	0.65	0.55	0.47	0.49	0.015
Computational time	0.54	0.95	0.74	0.67	0.59	0.30	0.84	0.023

Table 6. Overall analysis of the introduced technique related to prevailing approaches.

Factor (1.001), suggesting minimal multi-collinearity. Its mean performance (61,429.83) and moderate standard deviation (3,725.78) further highlight its reliability and stability, outperforming other models.

Ablation study of the proposed method

An ablation study is a methodical experiment to examine the effects of specific elements or characteristics of a suggested approach. An ablation study for the proposed method (Qua-KSar-DCK-ArNet) would include methodically eliminating or adjusting particular aspects or techniques to regulate how they affect the model's overall presentation. Table 9 shows the ablation study of the proposed method compared with previous Siamese networks.

Table 9 and Fig. 15 comparisons of different configurations of the model and their performance on datasets. The proposed method provides better results by adding DKAN and Q2CNN. With these networks, this method would have more accuracy. The configuration DKAN exhibits a slight drop in accuracy and attains. 94.78% for the first dataset and 96.87% for the second dataset. Like that, the newly developed variant Q2CNN has slightly improved from the DKAN CNN with the following accuracy: first dataset 96.86%, second dataset 93.45%. Therefore, the accuracy obtained from the proposed Qua-KSar-DCK-ArNet for the first dataset is 99.95% and for the second dataset is 99.97%.

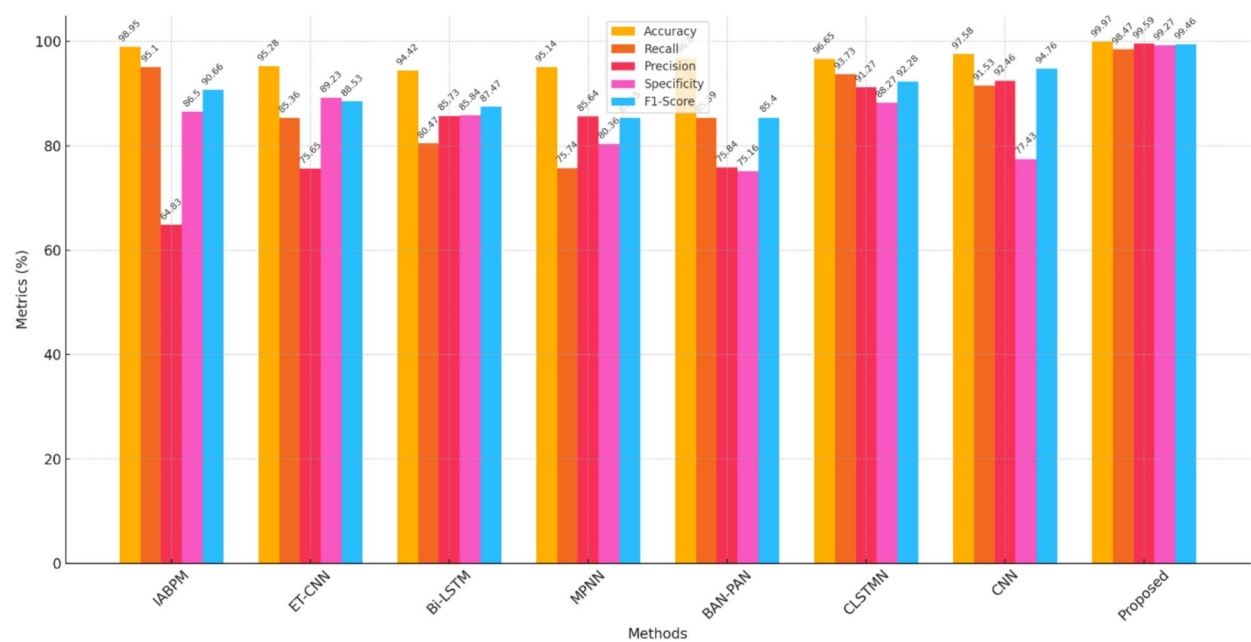


Fig. 12. Overall analysis of the introduced technique related to prevailing approaches.

Optimization technique	Fitness (Training accuracy in (%))	Time in seconds
SHOA <sup>33</sup>	93.59	0.2
GAOA <sup>34</sup>	96.35	0.4
GGOA <sup>35</sup>	94.84	0.3
TTAOA <sup>36</sup>	95.26	0.6
GTOA <sup>37</sup>	97.60	0.5
Qua-KSar-DCK-ArNet (Proposed)	99.9	0.1

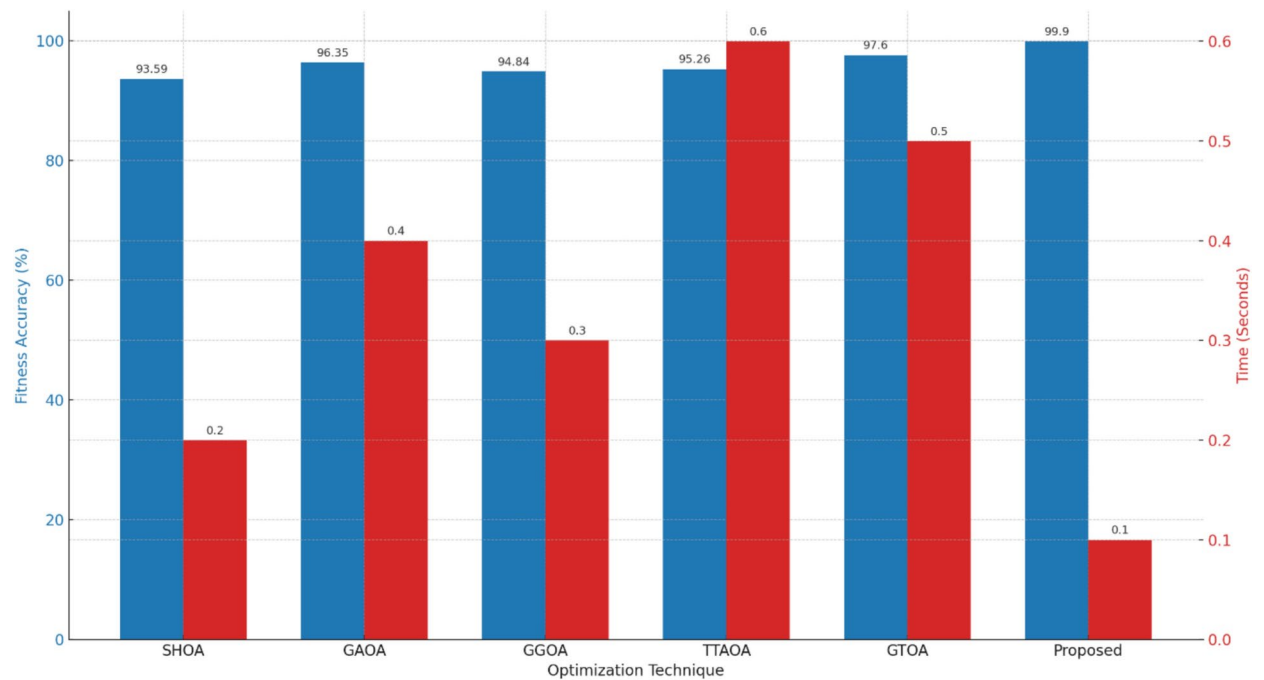
Table 7. Fitness of fittest agents.

Discussion

The proposed Qua-KSar-DCK-ArNet system demonstrates significant potential in enhancing heart disease prediction through IoT-based smart health monitoring. The integration of quantum clustering, advanced preprocessing, and deep learning-based attention mechanisms ensures high accuracy and efficient data processing. Compared to traditional methods, our approach achieves superior predictive performance while enabling real-time diagnosis, which is crucial for early intervention. However, challenges such as data security, computational complexity, and system scalability must be addressed for widespread implementation. Future improvements will focus on optimizing resource utilization, integrating additional physiological parameters, and enhancing model interpretability for clinical applications.

Conclusions

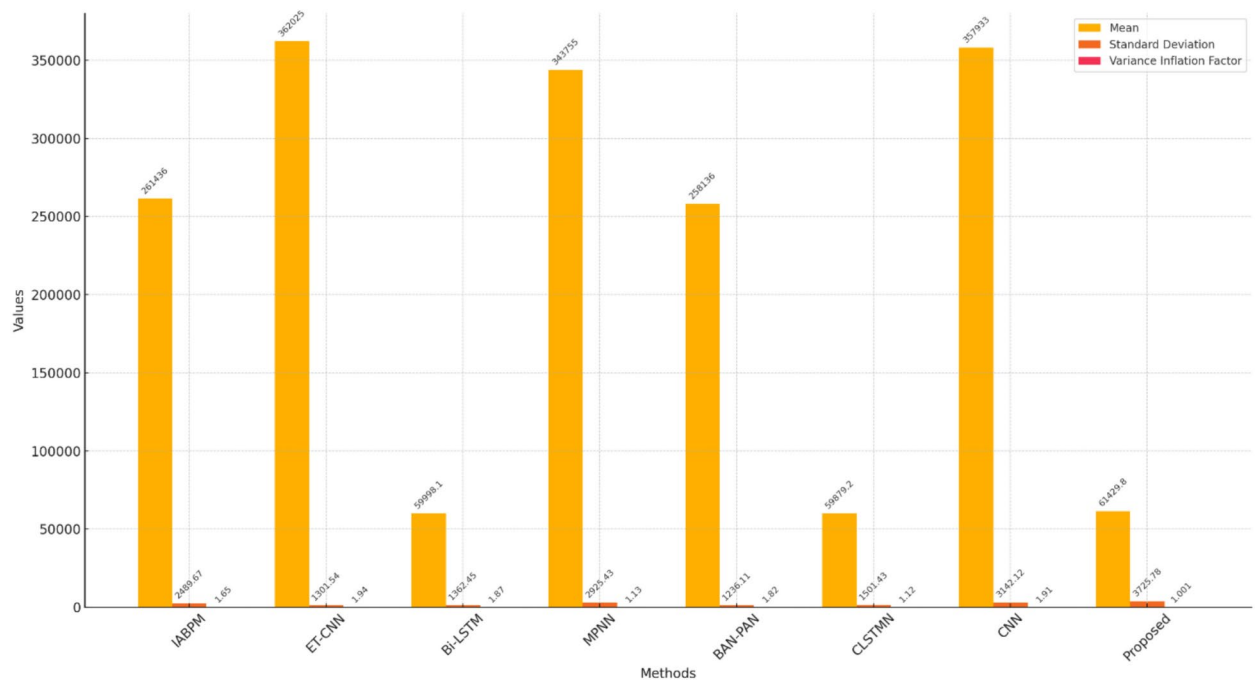
This study presents the Qua-KSar-DCK-ArNet system, an IoT-based smart health monitoring framework for heart disease prediction. By leveraging quantum clustering, advanced preprocessing, and a diffusion attention network, the system achieves an impressive 99% accuracy, enabling real-time diagnosis and improved patient outcomes. The findings highlight its potential for quick and precise heart disease detection, particularly in resource-limited settings. Future work will focus on integrating additional physiological data and enhancing deep learning techniques to further improve prediction quality and scalability. Moreover, addressing challenges related to data security, computational efficiency, and seamless cloud integration will be crucial for its broader adoption in clinical practice. Next, we will explore the feasibility of using RTA for, for instance, anticipating and preventing heart disease cases, and we aspire to advance the system’s applicability to as many facilities, especially those in LMICs, as possible.



**Fig. 13.** Fitness of fittest agents.

Methods	SW test <i>p</i> -value	WSR test/U-test <i>p</i> -value	H-test <i>p</i> -value	KS test <i>p</i> -value	FT <i>p</i> -value	Mean	Standard deviation	Variance inflation factor
IABPM <sup>16</sup>	0.035	0.045	0.37	0.029	0.027	261,435.7	2489.67	1.65
ET-CNN <sup>17</sup>	0.225	0.265	0.047	0.031	0.042	362,024.6	1301.54	1.94
Bi-LSTM <sup>18</sup>	0.256	0.193	0.36	0.024	0.035	59,998.12	1362.45	1.87
MPNN <sup>19</sup>	0.347	0.231	0.31	0.021	0.048	343,754.9	2925.43	1.13
BAN-PAN <sup>20</sup>	0.396	0.286	0.35	0.042	0.032	258,135.8	1236.11	1.82
CLSTMN <sup>21</sup>	0.262	0.312	0.23	0.025	0.078	59,879.24	1501.43	1.12
CNN <sup>17</sup>	0.229	0.412	0.28	0.038	0.051	357,932.7	3142.12	1.91
Qua-KSar-DCK-ArNet (Proposed)	<0.001	<0.001	<0.001	<0.001	<0.001	61,429.83	3725.78	1.001

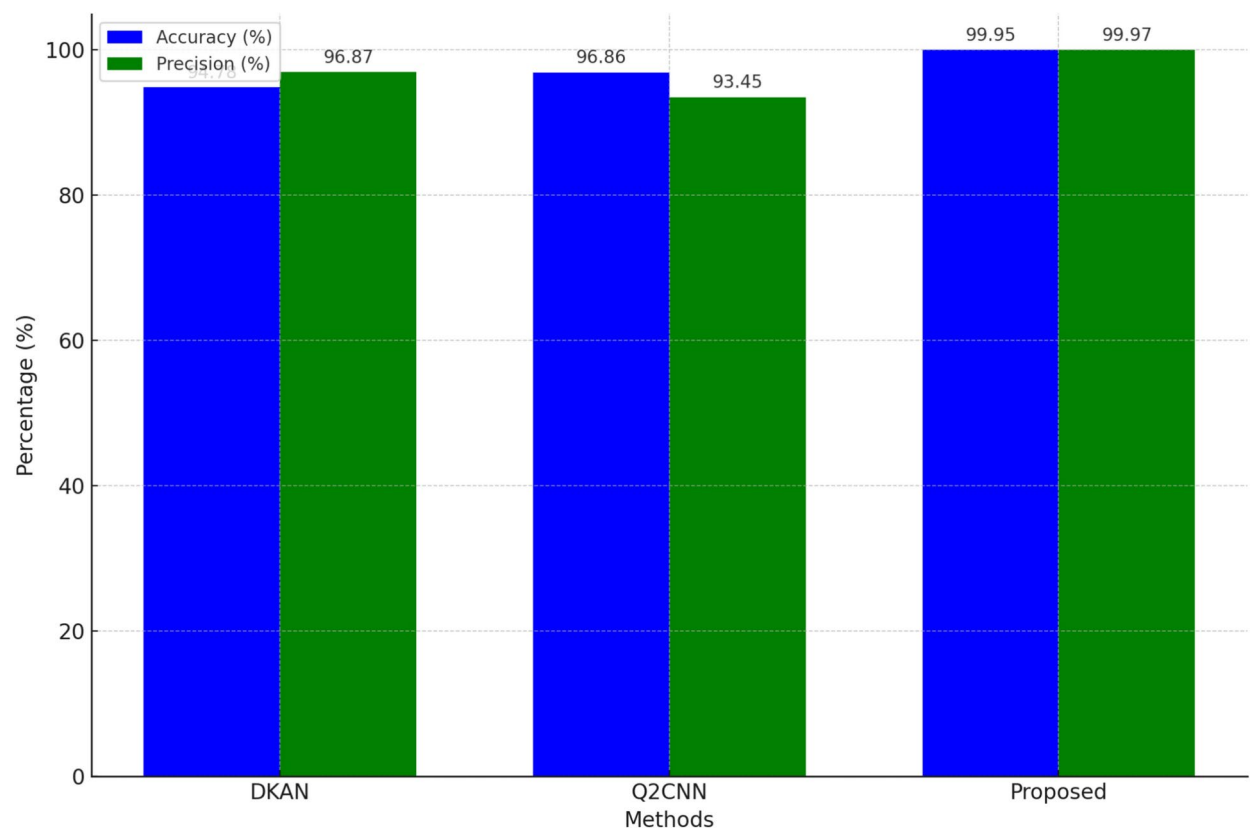
**Table 8.** Statistical Analysis of the proposed method versus existing methods.



**Fig. 14.** Statistical analysis of the proposed method versus existing methods.

Methods	IoT-based smart health monitoring sensors	
	Accuracy (%)	Precision (%)
DKAN	94.78	96.87
Q2CNN	96.86	93.45
Qua-KSar-DCK-ArNet (Proposed)	99.95	99.97

**Table 9.** Ablation study.



**Fig. 15.** Ablation study.

## Data availability

The data that supports the findings of this study are available within the article.

Received: 22 February 2025; Accepted: 24 April 2025

Published online: 19 May 2025

## References

- Ramkumar, G., Seetha, J., Priyadarshini, R., Gopila, M. & Saranya, G. IoT-based patient monitoring system for predicting heart disease using deep learning. *Measurement* **218**, 113235 (2023).
- Mistry, J. & Ganesh, A. An analysis of IoT-based solutions for congenital heart disease monitoring and prevention. *J. Xidian Univ.* **17**(7), 325–334 (2023).
- Arthy, S. & Prasanth, A. An intelligent heart disease prediction system using IoT and adaptive deep convolution neural network, in *AIP Conf. Proc.*, Vol. 2802, No. 1 (2024).
- Kumar, V. et al. Development of medical IoT system for the prediction of heart disease, in *2024 4th Int. Conf. Adv. Comput. Innov. Technol. Eng. (ICACITE)* 951–956 (2024).
- Yashudas, A., Dinesh Gupta, G. C., Prashant, A. D., Dokhyl AlQahtani, A. & Reddy, S. K. DEEP-CARDIO: Recommendation system for cardiovascular disease prediction using IoT network. *IEEE Sens. J.* **24**(9), 14539–14547. <https://doi.org/10.1109/JSEN.2024.3373429> (2024).
- Amune, A. et al. Integrated healthcare system for saline monitoring, patient communication, and heart disease prediction using IoT and machine learning algorithms, in *2024 Int. Conf. Emerg. Smart Comput. Inform. (ESCI)* 1–6 (2024).
- Rajkumar, G., Devi, T. G. & Srinivasan, A. Heart disease prediction using IoT-based framework and improved deep learning approach: Medical application. *Med. Eng. Phys.* **111**, 103937 (2023).
- Mistry, J., Ganesh, A., Ramakrishnan, R. & Logeshwaran, J. An IoT-based congenital heart disease prediction system will amplify authentication and data security by using cloud computing. *Eur. Chem. Bull.* **12**, 7201–7213 (2023).
- Krishnamoorthy, R. et al. An intelligent IoT-based smart healthcare monitoring system using machine learning. In *5G-Based Smart Hospitals and Healthcare Systems: Evaluation, Integration, and Deployment* (eds Kumar, A. et al.) 230–247 (CRC Press, 2023).
- Safa, M. et al. Real-time healthcare big data analytics model for improved quality of service in cardiac disease prediction with IoT devices. *Health Technol.* **13**(3), 473–483 (2023).
- Nancy, A. A. et al. Fog-based innovative cardiovascular disease prediction system powered by modified gated recurrent unit. *Diagnostics* **13**(12), 2071 (2023).
- Vishnu, K., Sebastian, T., Binu, P. K. Smart healthcare systems: A survey of IoT and AI technologies for heart disease prediction, in *2024 Int. Conf. Adv. Power, Commun. Intell. Syst. (APCI)* 1–7 (2024).
- Janarthanan, V., Annamalai, T. & Arumugam, M. Enhancing healthcare in the digital era: A secure e-health system for heart disease prediction and cloud security. *Expert Syst. Appl.* **255**, 124479 (2024).
- Bebortta, S. et al. Fedehr: A federated learning approach towards the prediction of heart diseases in IoT-based electronic health records. *Diagnostics* **13**(20), 3166 (2023).
- Mishra, J. & Tiwari, M. IoT-enabled ECG-based heart disease prediction using three-layer deep learning and meta-heuristic approach. *Signal Image Video Process.* **18**(1), 361–367 (2024).



16. Yenurkar, G. et al. Revolutionizing chronic heart disease management: The role of IoT-based ambulatory blood pressure monitoring system. *Diagnostics* **14**(12), 1297 (2024).
17. Umer, M. et al. Heart failure patients monitoring using IoT-based remote monitoring system. *Sci. Rep.* **13**(1), 19213 (2023).
18. Alzakari, S. A. et al. Enhanced heart disease prediction in remote healthcare monitoring using IoT-enabled cloud-based XGBoost and Bi-LSTM. *Alex. Eng. J.* **105**, 280–291 (2024).
19. Almujaally, N. A. et al. Monitoring acute heart failure patients using internet-of-things-based smart monitoring system. *Sensors* **23**(10), 4580 (2023).
20. Akhbarifar, S. et al. A secure remote health monitoring model for early disease diagnosis in cloud-based IoT environment. *Pers. Ubiquitous Comput.* **27**(3), 697–713 (2023).
21. Islam, M. R. et al. Deep learning-based IoT system for remote monitoring and early detection of health issues in real-time. *Sensors* **23**(11), 5204 (2023).
22. Pachiyannan, P. et al. A novel machine learning-based prediction method for early detection and diagnosis of congenital heart disease using ECG signal processing. *Technologies* **12**(1), 4 (2024).
23. Mazhar, T. et al. A novel expert system for the diagnosis and treatment of heart disease. *Electronics* **11**(23), 3989 (2022).
24. Khan, M. A. et al. Optimal feature selection for heart disease prediction using modified Artificial Bee colony (M-ABC) and K-nearest neighbors (KNN). *Sci. Rep.* <https://doi.org/10.1038/s41598-024-78021-1> (2024).
25. Mazhar, T. et al. Analysis of integration of IoMT with blockchain: issues, challenges and solutions. *Discov. Internet Things* <https://doi.org/10.1007/s43926-024-00078-1> (2024).
26. Ghadi, Y. Y. et al. Enhancing patient healthcare with mobile edge computing and 5G: challenges and solutions for secure online health tools. *J. Cloud Comput.* <https://doi.org/10.1186/s13677-024-00654-4> (2024).
27. Poggiali, A. et al. Quantum clustering with k-means: A hybrid approach. *Theor. Comput. Sci.* **992**, 114466 (2024).
28. Sholeh, M. & Nurnawati, E. K. Comparison of Z-score, min-max, and no normalization methods using support vector machine algorithm to predict student's timely graduation, in *AIP Conf. Proc.*, Vol. 3077, No. 1 (2024).
29. Park, C., Jeong, Y., Cho, M., & Park, J. Fast point transformer, in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.* 16949–16958 (2022).
30. Matic, A. et al. Quantum-classical convolutional neural networks in radiological image classification, in *2022 IEEE Int. Conf. Quantum Comput. Eng. (QCE)* 56–66 (2022).
31. He, K., Cai, Y., Peng, S. & Tan, M. A diffusion model-assisted multiscale spectral attention network for hyperspectral image super-resolution. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **17**, 8612–8625. <https://doi.org/10.1109/JSTARS.2024.3386702> (2024).
32. Soltani, N. & Maybodi, N. A. Image thresholding based on sailfish optimizer algorithm, in *2024 3rd Int. Conf. Distrib. Comput. High Perform. Comput. (DCHPC)* 1–10 (2024).
33. Zhao, S. et al. Sea-horse optimizer: A novel nature-inspired meta-heuristic for global optimization problems. *Appl. Intell.* **53**(10), 11833–11860 (2023).
34. Dehghani, M., Trojovský, P. & Malik, O. P. Green anaconda optimization: A new bio-inspired metaheuristic algorithm for solving optimization problems. *Biomimetics* **8**(1), 121 (2023).
35. Sharma, A. et al. Multi-objective greylag goose optimization, in *2024 2nd Int. Conf. Adv. Comput. Comput. Technol. (InCACCT)* 374–379 (2024).
36. Zeidan, M. A. et al. Enhancement of a hybrid electric shipboard microgrid's frequency stability with triangulation topology aggregation optimizer-based 3DOF-PID-TI controller. *IEEE Access* **12**, 66625 (2024).
37. Elymany, M. M., Enany, M. A. & Elsonbaty, N. A. Hybrid optimized-ANFIS based MPPT for hybrid microgrid using zebra optimization algorithm and artificial gorilla troops optimizer. *Energy Convers. Manag.* **299**, 117809 (2024).

## Acknowledgements

The authors extend their appreciation to the Deanship of Research and Graduate Studies at King Khalid University for funding this work through Large Research Project under grant number RGP2/441/46.

## Author contributions

Conceptualization, D.K.H, S.N.T; Writing—Review and Editing, C.N.R, D.D; Methodology, Q.N.N, S.K; Supervision, M.K, A.F.E.

## Declarations

## Competing interests

The authors declare no competing interests.

## Ethical approval

This article does not contain any studies with human participants or animals performed by any of the author.

## Additional information

**Correspondence** and requests for materials should be addressed to D.K.H. or A.F.E.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025