Journal of Trends in Computer Science and Smart Technology (ISSN: 2582-4104)
www.irojournals.com/tcsst/

Federated Attention-Capsule CNN with Bio-Optimization for Coronary Artery Disease Screening

Angelin Jeba P.[1], Jasmine David D.[2], Selvarathi M.[3], Kenneth Bryan Nesh[4], Alan Biju Varkey[5], Jemima Jebaseeli T.[6]

[1,4,5]Division of Artificial Intelligence and Machine Learning, Karunya Institute of Technology and Sciences, Coimbatore, India.
[2]Associate Professor, School of Computer Science and Engineering, Presidency University, Bangalore, Karnataka, India.
[3]Division of Mathematics, Karunya Institute of Technology and Sciences, Coimbatore, India.
[6]Department of Computer Science and Engineering, Vel Tech Rangarajan Dr. Sagunthala R & D Institute of Science and Technology, Chennai, India.

Email: [1]angelinjeba@karunya.edu, [2]jasmine.d@presidencyuniversity.in, [3]selvarathi@karunya.edu, [4]kennethbryan@karunya.edu.in, [5]alanbiju@karunya.edu.in, [6]jemi.jeba@gmail.com

Abstract

Coronary artery disease (CAD) remains the leading cause of death around the globe and hence requires reliable, non-invasive, and secure diagnostic methods. In order to diagnose CAD in its early stages, we present a novel architecture of a Federated Attention-Capsule Convolutional Neural Network (FAC-CNN) integrated with Bio-Inspired Optimization Algorithms (BIOA) in dealing with multiple modalities of physiological data such as Electrocardiography (ECG), Photoplethysmogram (PPG), and Blood Pressure (BP). The model is evaluated using unimodal (only ECG signals) and multimodal physiological datasets to validate the proposed algorithm for different data distributions. To improve accuracy, the architecture of FAC-CNN uses capsule networks to recognize the hierarchical structure within the data fed into it. Through attention mechanisms, the network is capable of selecting the most important features related to cardiac diseases. To ensure data security, the federated learning (FL) method is applied as a solution to the problem of local model training using patients' private data stored in edge computing devices. A novel hybrid optimization method combining Particle Swarm Optimization (PSO) and Genetic Algorithm (GA) is used to update learning rates, the total number of capsule layers, and routing iterations. Three databases, including the PTB diagnostic ECG database, the MIT-BIH Arrhythmia database, and the MIMIC-III Waveform database, which comprises authentic ICU recordings of ECG, PPG, and blood pressure data from more than 40,000 patients were employed to evaluate the model. It can be seen from the results that FAC-CNN outperformed the other models used in the experiment, including CNNs and LSTMs by 4-7%, obtaining a mean accuracy of 97.4%, an F1-score of 96.8%, and an AUC score of 0.985. Moreover, the scalability of the proposed algorithm for applications in remote healthcare settings improved, and the more effective training method contributed to time-savings for the training procedure, decreasing it by 18.6%.

Keywords: Coronary Artery Disease, Federated Learning, Capsule Networks, Convolutional Neural Network, Genetic Algorithm, Particle Swarm Optimization.

DOI: https://doi.org/10.36548/jtcsst.2026.2.001
Received: 24.02.2026, received in revised form: 22.03.2026, accepted: 08.04.2026, published: 17.04.2026

1. Introduction

Around 17.9 million deaths globally are attributed to coronary artery disease (CAD), making it the leading cause of death. By 2030, the economic cost of cardiovascular illnesses is expected to reach USD 1,044 billion [1]. Non-invasive, real-time cardiovascular screening is now possible because of the development of wearable IoT technology and physiological signals like ECG, PPG, and continuous blood pressure monitoring. However, the centralized data processing that traditional AI-based diagnostic models rely on presents significant privacy concerns, particularly when handling sensitive health data. Furthermore, the majority of current models exhibit poor generalization in decentralized healthcare systems with non-IID data settings [2]. The use of Federated Learning (FL), a privacy-preserving substitute, in wearable-based CAD identification is constrained by communication inefficiencies, device limitations, and diverse data distributions. Attention mechanisms and capsule networks have shown remarkable potential in improving pattern recognition in biomedical signals. The attention-based PPG-to-ECG translation has achieved a PR-AUC of 0.986 in atrial fibrillation detection tasks [4]. However, their integration within federated systems remains underexplored.

Bio-inspired algorithms have been shown to be more effective in hyperparameter optimization for deep learning models, particularly in noisy or limited conditions [5]. However, problems such as convergence robustness, scalability, and efficient resource utilization remain in FL systems. FedAvg was introduced for the first time by Kairouz et al. [3]. This algorithm allows for decentralized neural network training using multiple clients without exchanging raw data. Traditional FL models face challenges when deployed in realistic healthcare applications, notably due to non-IID data distribution, expensive communication, and limited computation capacity at the edge, despite solving privacy concerns [7].

The integration of multimodal information through federated systems poses numerous open issues [8]. To start with, there is an underexplored potential in using sophisticated neural networks, including Capsule Networks and attention-based models, which would allow for the emphasis on relevant features and the preservation of spatial hierarchy despite the application of federated learning in analyzing physiological signals [9], [10]. Moreover, federated models often exhibit poor performance due to static hyperparameter settings; however, bio-inspired optimization approaches can be employed here [11]. Furthermore, current CAD screening techniques rely on a single modality of input data and centralized solutions without addressing the above problems simultaneously. Lastly, the computationally expensive process of model training hinders the applicability of FL in this scenario [12].

The research aims to develop a new Federated Attention-Capsule Convolutional Neural Network (FAC-CNN) method using hybrid biologically inspired optimization algorithms to provide reliable, efficient, and privacy-aware CAD screening in distributed wearable health monitoring platforms. The developed method employs capsule networks along with attention mechanisms to generate effective spatio-temporal information from ECG, PPG, and BP signals. The privacy protection capability is achieved using FL, where data can only be processed by edge devices. Hybrid Particle Swarm Optimization and Genetic Algorithm (PSO-GA) methods are utilized to optimize hyperparameters dynamically among non-IID clients.

The following are the novel contributions.

- Integrating Capsule Networks to preserve spatial and hierarchical relationships in multimodal biomedical signals (ECG, PPG, BP).

- Modality-flexible unified architecture supporting missing data scenarios.
- Embedding a self-attention mechanism within capsule layers to dynamically weight clinically relevant features for CAD detection.
- Implementation of FL to enable on-device training without transferring raw patient data.
- Robust handling of non-IID and heterogeneous data distributions across multi-institutional and multi-device environments.
- Combination of PSO and Genetic Algorithms for joint hyperparameter tuning.
- Synchronous processing of ECG, PPG, and blood pressure signals to capture cross-modal feature interactions.
- Adaptive fusion mechanism within capsule-attention layers to preserve temporal and morphological consistency for accurate CAD detection.

2. Literature Review

There has been considerable development in FL technology and its impact on medical solutions, especially when used with wearables. For instance, in their study, Aminifar et al. [13] designed an edge-FL solution to train local models to detect seizures using wearables without violating data privacy. However, this solution could only be useful for detecting seizures and not cardiovascular diseases, as the authors did not consider advanced neural network architectures such as capsule networks and attention mechanisms. Like Aminifar et al., Baucas et al. [14] introduced an FL framework enabled by blockchain in the fog-IoT environment for predictive medicine that ensures data integrity and secure data transmission. Unfortunately, their model was not effective for handling situations like CAD, as it lacked adaptability.

Another research effort employed a capsule-based framework named 1D-CADCapsNet for centralized CAD detection using small segments of an ECG and demonstrated high precision [15]. The approach lacked privacy guarantees and used a small, homogeneous dataset for training, making it impractical for federated applications. More comprehensive evaluations have been provided in recent review articles [16], [17], which pointed out systemic challenges in the implementation of FL for IoT applications due to factors such as the high cost of communication, resource limitations in edge devices, and the absence of dynamic optimization approaches. In terms of multi-modal signal modelling, recent research based on W-Nets proposed a framework to reconstruct data for wearable health monitoring [18]; however, its patient-specific architecture and the need for aligned training data hampered its applicability.

Overall, there are evident shortcomings in integrating FL, capsule-based networks, attention-based algorithms, and bio-inspired methods for multimodal CAD detection. For example, a boosted FL classifier using FedImpPSO has been shown to offer an 8.14% increase in the accuracy rate compared to conventional FedAvg algorithms, achieving a 91-92% accuracy rate for cardiovascular datasets in unfavorable communication settings. This study considered only visual modalities such as chest X-ray and ultrasound imaging but did not include time-series biosignals or resource-limited edge nodes [19]. The FL algorithm in the cardiac disease prediction model uses M-ABC and provides an accuracy rate of 92.89%, a precision rate of 94.2%, a sensitivity rate of 96.6%, and specificity rate of 81.8% while lowering training epochs by 22% relative to FedAvg and FedMA algorithms. Nevertheless, this work was conducted in simulations and ignored the incorporation of wearable devices in practical scenarios and deep features through capsule and attention mechanisms [20].

In the field of architectural designs, the hybrid method of using CNNs, Transformers, and Capsules was developed under the name of CTCNet to classify sleep stages. The accuracy obtained by this network reached values of 86.2%, 82.5%, and 85.7% on different datasets. Although the model successfully learned temporal global and local features, it was built specifically for centralized architectures and did not undergo experiments in the federated or cardiovascular settings [21]. Furthermore, a new benchmark dataset called FedCVD emerged to address the issues of large-scale, real-world federated learning of cardiovascular disease prediction from ECG and echocardiograms. The paper demonstrated difficulties such as non-IID, long tail, and heterogeneous institution properties; nevertheless, it did not investigate architectural improvement or adaptive optimization.

One of these models was created by Mulani et al. [24] using the technologies of IoMT, ML, as well as wearable devices to predict diseases of cardiac nature. Despite being innovative and effective, this solution fails to consider the problems of interpretability and robustness. Moreover, the combination of IoT sensor data and the Kernel-PCA algorithm for pre-processing makes it possible to create an effective solution like SF II Adaboost, which provides excellent performance (accuracy – 95.37%, sensitivity – 94.3%, and specificity – 96.31%) but cannot address the issues of communication and privacy since it is a centralized model [25]. Finally, the use of HA CNN BiLSTM for the reconstruction of the ECG signal from PPG data can be quite efficient (RMSE – 0.031), but it requires further research to prove that it is not sensitive to noise and can be used in federated settings [26]. The federated approach utilizing the ring signature method provided an enhanced level of privacy while defending against source identification attacks; nonetheless, the approach failed to address the issues of multimodal data fusion and lightweight processing [27]. A systematic literature review highlighted the significance of privacy, security, and interoperability in FL models in medical settings, with no existing unified approaches for solving multimodal inference and optimization [28][29].

In order to resolve these problems, a new method of FAC-CNN that incorporates private edge learning, multimodal signal fusion (ECG, PPG, BP), and the utilization of PSO-GA hybrid algorithms is proposed.

3. Dataset

The FAC-CNN framework was evaluated using three datasets, as shown in Table 1. Two-channel ECG recordings from 47 patients at 360 Hz with a 70/30 train-test split are included in the MIT-BIH Arrhythmia database. 12-lead ECG data from 290 patients, with 549 recordings, recorded at 1000 Hz, and divided 75% for training and 25% for testing, are included in the PTB diagnostic ECG database. Additionally, the MIMIC-III Waveform Database provides real ICU recordings of ECG, PPG, and blood pressure signals from over 40,000 patients, with variable sampling rates of 125–500 Hz and custom splits per study.

It is important to note that the MIT-BIH and PTB datasets are not dedicated CAD screening datasets but are widely used ECG benchmark datasets for cardiac abnormality detection. In this work, they are utilized to evaluate ECG feature learning capability, while multimodal CAD-related validation is supported through the MIMIC-III dataset, which includes ECG, PPG, and BP signals.

Table 1. Datasets Used for CAD Screening

Dataset	Description	Modality	Sampling Rate	Subjects	Train/Test Split
MIT-BIH Arrhythmia [30]	ECG recordings with annotated arrhythmias from the Beth Israel Hospital.	ECG (2-channel)	360 Hz	47	70% Train / 30% Test
PTB Diagnostic ECG [31]	12-lead ECG signals from healthy and cardiac patients, provided by PTB, Germany.	ECG (12-lead)	1000 Hz	290 patients (549 rec.)	75% Train / 25% Test
MIMIC-III Waveform [32]	Real ICU data including, ECG, PPG, and blood pressure recordings from multiple devices.	ECG, PPG, BP	Varies (125–500 Hz)	>40,000	Custom split per study

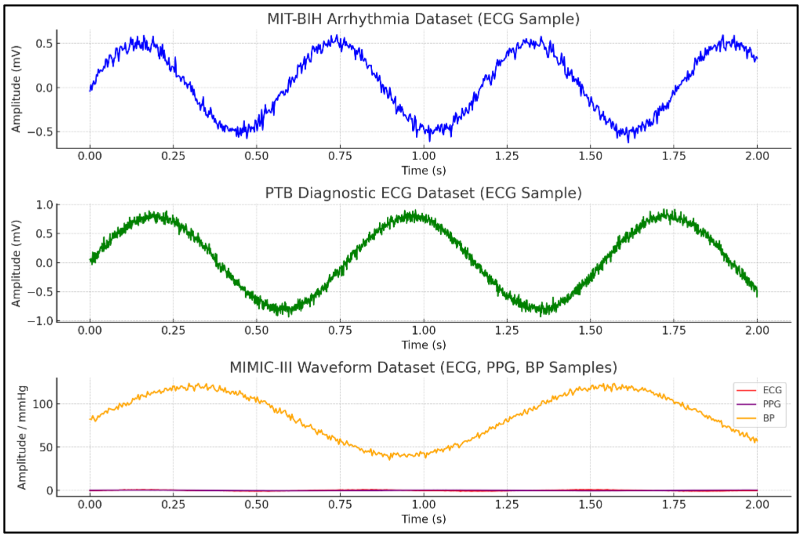

Figure 1. Sample Biomedical Signal Recordings from MIT-BIH, PTB, and MIMIC-III Datasets

Figure 1 shows sample ECG, PPG, and BP waveforms from three datasets of MIT-BIH (2-channel ECG, 360 Hz), PTB (12-lead ECG, 1000 Hz), and MIMIC-III (multimodal ICU signals, 125–500 Hz), providing diverse cardiac patterns for CAD detection.

Even though the suggested FAC-CNN can be considered an integrated multimodal framework, the architecture is developed under the principle of modality flexibility. Each modality such as ECG, PPG, and BP will be encoded independently through dedicated modules, while the fusion module will be adaptable depending on the input signals. When some of the signals are not present for processing, the framework does not turn into another one, but continues operating in a partially multimodal manner.

4. Methodology

Figure 2 presents the architecture that represents the end-to-end pipeline of the FAC-CNN model used in the detection of cardiovascular disease from physiological multimodal sensor information. Firstly, the physiological signals are acquired through the ECG, PPG, and BP sensors, which capture biomedical analog signal information. These analog signals are subsequently transformed into digital signals through the analog-to-digital conversion stage, enabling computational analysis of the physiological data. The digitized signals are further subjected to a signal preprocessing procedure, where noise reduction, signal normalization, and synchronization processes are implemented to improve the quality of the signals. Afterward, the physiological signals are fed into the feature extraction component to extract specific features. The features are further enhanced by implementing the attention mechanism, whereby higher importance is assigned to the signal segments that contain the relevant information corresponding to cardiac anomaly events. Finally, the enhanced signal features are input to the CAD classification component, where the patient is predicted to belong to either the normal or CAD class. The entire system operates within a federated learning framework, whereby several edge devices, such as hospitals and wearable monitoring units, collaboratively develop their models but share only their models' parameters.

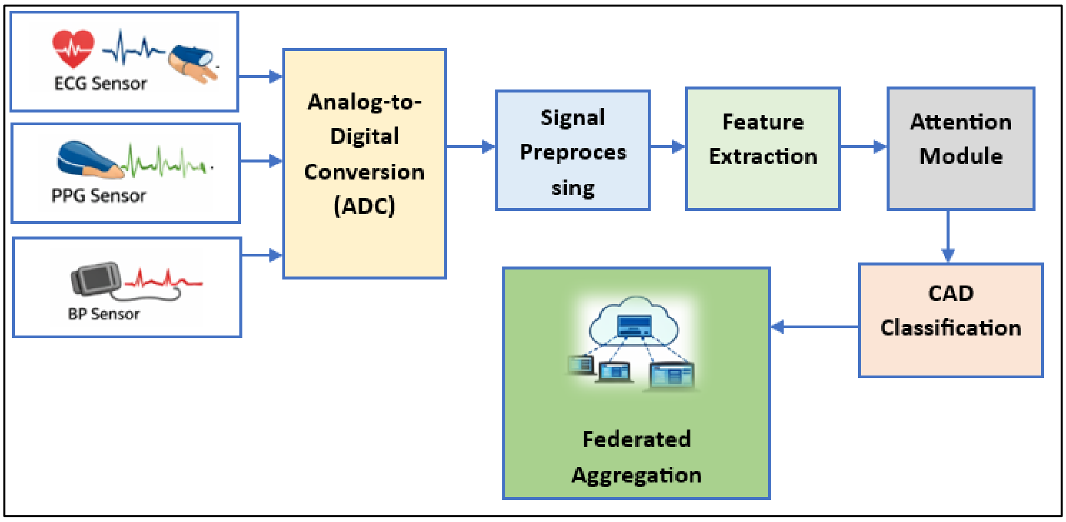

Figure 2. Federated Attention-Capsule CNN Architecture with Multimodal Physiological Inputs (ECG, PPG, and Blood Pressure) for Privacy-Preserving CAD Detection

The proposed FAC-CNN architecture contains approximately 2.3 million trainable parameters, making it suitable for deployment on edge-enabled healthcare devices. The proposed FAC-CNN follows a structured pipeline:

- Multimodal fusion combines ECG, PPG, and BP features.
- Self-attention highlights clinically important signal segments.
- Capsule networks preserve hierarchical feature relationships.
- Federated Learning (FL) enables privacy-preserving distributed training.
- PSO-GA optimization improves hyperparameter tuning and convergence.

Each module enhances the output of the previous stage, forming a unified framework.

4.1 Multimodal Input Processing

CAD diagnostics relying on a single physiological signal often prove insufficient due to signal noise, variability, and the complex manifestation of the disease across multiple biological systems. To overcome these limitations, a multimodal signal processing strategy is employed to integrate information from the following three complementary biosignals.

- ECG (Electrocardiogram) - measures the heart’s electrical impulses

- PPG (Photoplethysmogram) - represents fluctuations in blood volume in peripheral vessels
- BP (Blood Pressure) - represents hemodynamic changes over time

Each signal is recorded as a time-series sequence, sampled over a fixed time window. Let the multimodal time-series input be, $X = \{X^{ECG}, X^{PPG}, X^{BP}\}$. To extract meaningful representations from these raw signals, define a modality-specific feature extraction function as shown in equation (1).

$$F^{(m)} = \phi^{(m)}(X^{(m)}; \theta^{(m)}) \quad (1)$$

where $\phi^{(m)}$ is a feature extractor, usually a CNN or LSTM network. CNNs are effective in capturing local patterns like ECG waveforms (P, QRS, T). LSTMs are suited to model long-term dependencies, such as rhythm or periodicity. $\theta^{(m)}$are learnable weights of the network for modality m. The output $F^{(m)} \in \mathbb{R}^{d_m}$is a dense feature vector capturing the modality-specific diagnostic characteristics. After independently processing each modality, concatenate their learned features into a joint fused representation. The fused representation is shown in equation (2).

$$F = concat(F^{(1)}, F^{(2)}, \ldots., F^{(M)}) \quad (2)$$

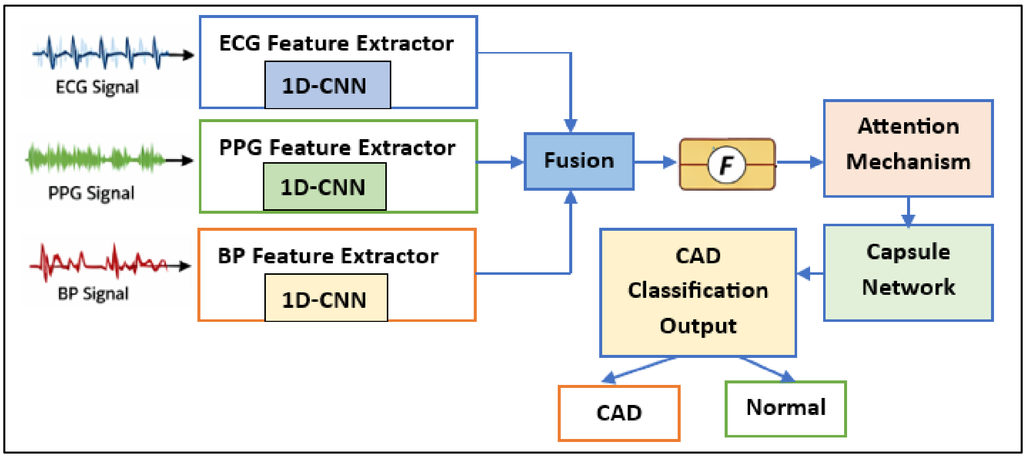

Figure 3. Multimodal Physiological Signal Feature Extraction and Fusion in Federated Attention-Capsule CNN for CAD Screening

As shown in Figure 3, the fused vector F$\in \mathbb{R}^d$ combines $d = \sum d_m$ insights from ECG, PPG, and BP, allowing the model to make more informed and accurate decisions. Since ECG, PPG, and blood pressure signals originate from different sensing devices and datasets, their sampling rates may vary. To ensure consistent multimodal integration, all signals are first resampled to a common sampling frequency of 250 Hz. After resampling, the signals are segmented using sliding temporal windows of 5 seconds to capture meaningful physiological patterns. The timestamps of each modality are then aligned to ensure temporal correspondence across ECG, PPG, and BP signals. Finally, the synchronized signal segments are concatenated to generate a unified multimodal feature representation used by the FAC-CNN framework for CAD detection.

As the ECG, PPG, and BP data may be obtained from sources with different sampling rates, a process of signal synchronization is carried out prior to feature extraction. Firstly, all the input signals are resampled to have the same sampling frequency of 250 Hz. Secondly, the resampled signals are partitioned into equal segments with a length of 5 seconds. Lastly, timestamps are aligned to make the segments temporally consistent, and thus, multimodal features are extracted using the concatenation of the synchronized segments.

4.2 Capsule Network Feature Extraction

As opposed to the typical CNN architectures with scalar outputs and pooling layers discard information about the location of the object in space, Capsule Networks (CapsNets) incorporate much more sophisticated representations using vector-valued outputs and agreement-based routing. This allows the networks to effectively capture part-whole dependencies, such as patterns in waveforms of ECG or pulse irregularities in PPG, which are useful for detecting abnormal heart conditions. The outputs of the primary capsule layers are known as the Primary Capsules. The output vectors of the primary capsules are shown in equation (3).

$$u_i \epsilon \mathbb{R}^d, i = 1,2, \dots, n \tag{3}$$

Each vector u_i represents the instantiation parameters of orientation, intensity, and duration of a low-level feature detected by convolutional filters. d is the dimension of the capsule (8 or 16), and n is the number of primary capsules. This encapsulation allows the model to capture not just presence, but also pose and variation of features vital for distinguishing similar-looking but clinically different patterns in ECG or PPG. These are transformed by learned weight matrices W_{ij} to prediction vectors as shown in equation (4).

$$\hat{u}_{j|i} = W_{ij} u_i \tag{4}$$

where $\hat{u}_{j|i}$ is the predicted vote for capsule j by capsule i. These learnable matrices W_{ij} enable the model to learn feature part-to-whole transformations, like how a QRS complex relates to an entire ECG beat. This mechanism is especially important in multimodal signals, where different modalities may encode complementary evidence of the same physiological event. The dynamic routing mechanism computes the capsule output as shown in equation (5).

$$s_j = \sum_i c_{ij} \, \hat{u}_{j|i}, v_j = squash(s_j) \tag{5}$$

where:

- c_{ij}: coupling coefficients updated iteratively
- s_j: pre-activation input to capsule j.
- v_j: final output vector after non-linear squashing.

The non-linear squash function ensures that vectors with large magnitudes are squashed to have lengths close to 1, preserving direction. Vectors with small magnitudes are shrunk, pushing noise or irrelevant signals towards zero. $squash(s) = \frac{\|s\|^2}{1+\|s\|^2} \cdot \frac{s}{\|s\|}$ ensures short vectors get shrunk, and long vectors approach unit length. This ensures while the direction encodes its attributes, a powerful property for learning interpretable representations in medical time-series data.

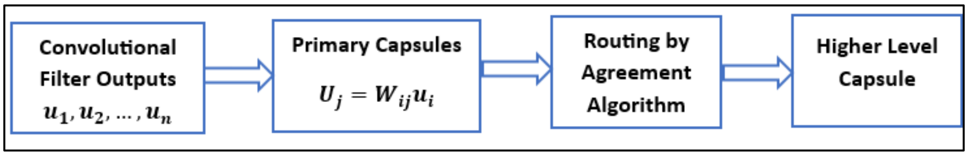

Figure 4. Capsule Network Feature Extraction Process for Multimodal Physiological Signals

In Figure 4, the process of extracting the features of the CapsNet network starts by transforming the outputs of convolutional filters obtained from multimodal physiological signals including ECG, PPG, and BP. Primary capsules are generated as vectors that include information such as orientation, intensity, and duration of identified features through this transformation process. Prediction vectors are generated by projecting the features of primary capsules into higher-level capsules using an adjustable transformation matrix W_{ij}. To support predictions based on high-level capsules, the coupling coefficients of prediction vectors are adjusted through iterative processes according to the routing-by-agreement procedure. Lastly, a normalization step is implemented by applying a non-linear squashing operation to the capsules to normalize the probability of existence as length and attributes as direction of vectors.

The capsule network is designed as follows: a Primary Capsule Layer with 32 capsules, each with dimensions of 16, is used for the network. Dynamic routing is carried out three times to obtain agreements between lower-level capsules and higher-level capsules.

4.3 Self-Attention Mechanism

Self-attention is a method that enables a model to compute representations by concentrating on the most significant portions of the input sequence. The input consists of the sequential biomedical data of ECG, PPG, and BP signals. The goal is to learn which time segments or features contribute most to detecting CAD and give them more weight during feature extraction.

Let X be the input feature matrix from the capsule layer (shape: sequence length × feature dimension). Equation (6) illustrates how learnable weight matrices are used to generate three distinct projections from X.

$$Q = W_Q X, K = W_K X, V = W_V X \tag{6}$$

where:

- Q (Query): What are we searching for (feature relevance)?
- K (Key): What we have in the memory (feature identifiers)?
- V (Value): The actual information content.

Compare each query with all keys by taking their dot product QK^T. This gives a similarity score between each position in the sequence and every other position. Scale it by $\sqrt{d_k}$, where $\sqrt{d_k}$ is the key dimension. By doing this, big values that can produce unstable gradients by making the softmax output overly harsh are avoided. The scaled scores are subjected to the softmax function, which transforms them into a probability distribution for every inquiry. Multiply the value matrix V by the probability distribution. This results in a weighted feature combination that amplifies significant time steps. Equation (7) shows the attention output.

$$Attention\ (Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)\mathrm{V} \tag{7}$$

The high probability indicates that this key feature is more important for the query. The low probability shows that it is less relevant to the query. This helps in dynamically emphasizing complex cardiac signal components. This enables the proposed FAC-CNN to detect subtle cardiac abnormalities without manually defining which signal parts to prioritize.

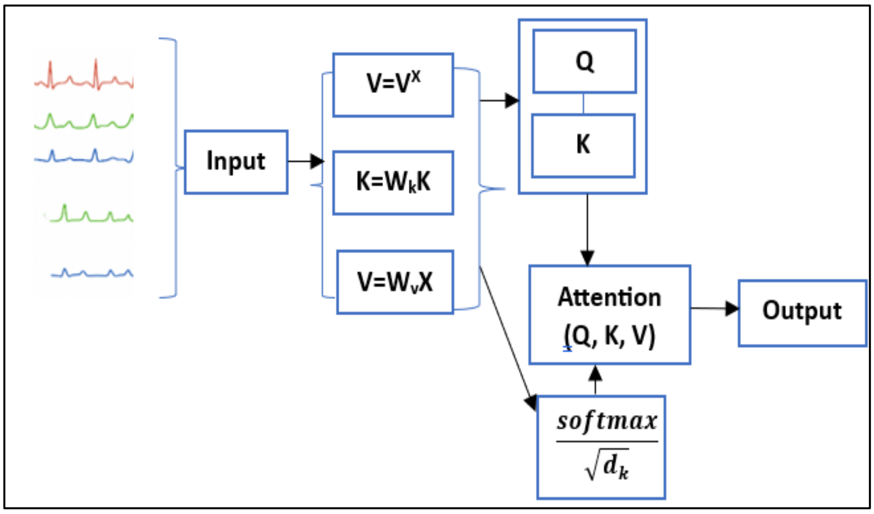

Figure 5. Self-Attention Mechanism for Multimodal Physiological Signal Analysis

Figure 5 depicts the self-attention mechanism applied to sequential biomedical data from ECG, PPG, and BP signals. Using distinct trainable weight matrices, the input signal matrix is converted into three learned representations of Queries (Q), Keys (K), and Values (V). The similarity between time steps is measured by computing the dot product between Q and K, which is then scaled by $\sqrt{d_k}$ to stabilize gradients. These scores are transformed into attention weights via a softmax function, which is then applied to generate a weighted sum of the Values (V). By amplifying key temporal segments or features, this procedure allows the model to concentrate on signal areas that are most important for identifying coronary artery disease. The mechanism allows dynamic, data-driven prioritization of complex cardiac signal components without manual feature selection. The self-attention layer employs 4 attention heads with an embedding dimension of 128. Multi-head attention allows the model to capture diverse relationships between temporal features across multimodal physiological signals while maintaining computational efficiency.

4.4 FL Aggregation

FL is a distributed Machine Learning (ML) technique in which a number of clients (e.g., wearable IoT devices, hospitals) work together to build a common global model without sharing raw data. Instead, model updates are sent to a central server by each client, which trains locally. Assume that K clients (devices/institutions) exist. Every client k has a dataset D_k with n_k samples in it. All clients get the current global model weights (w_k^t) from the server in round t. Every client uses local gradient descent for many epochs of training on its dataset, D_k. Client K generates an updated local model w_k^t following training. Clients simply transmit model weights or weight changes to the central server, not sensitive information. A new global model is created by the server by combining all of the local models.

The global model update at communication round $t+1$is computed as a weighted aggregation of local client models as shown in equation (8).

$$w^{t+1} = \sum_{k=1}^{K} \frac{n_k}{\sum_{k=1}^{K} n_k} w_k^t \tag{8}$$

where n_k represents the number of samples at client k, and K denotes the total number of participating clients. This formulation ensures that clients with larger datasets contribute proportionally more to the global model update, leading to a balanced and data-aware aggregation process.

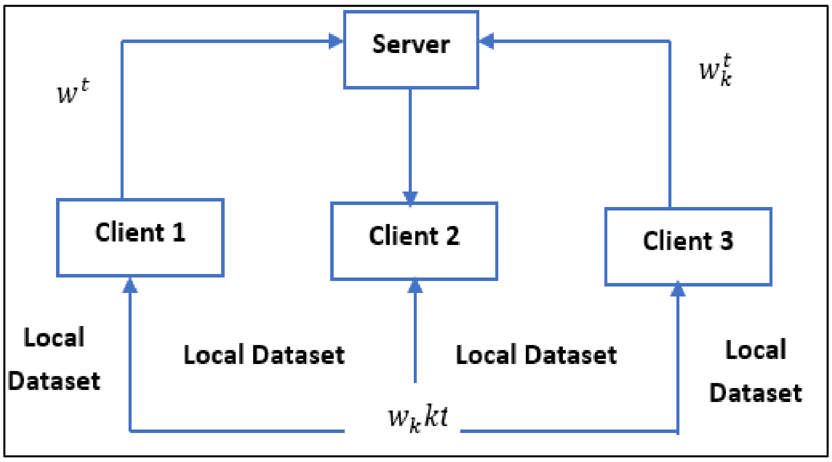

Figure 6. FL Framework with FedAvg Aggregation for Privacy-Preserving Cardiac Screening

The FL approach is depicted in Fig. 6, wherein multiple scattered clients like wearable IoT sensors and hospital systems work together to jointly train the same global model without exchanging raw patient data. Each client sends only the updated weights of the trained model to the centralized server for further training of the received global model using its own dataset.

The FedAvg algorithm is employed by the server to combine the model updates based on the weightage of each client, determined by the size of their respective datasets. The suggested FL framework is useful in providing accurate diagnoses for coronary artery disease in dispersed healthcare settings while preserving patient privacy and reducing computational overhead.

4.5 Hybrid PSO-GA Hyperparameter Optimization

It automatically optimizes hyperparameters θ (e.g., learning rate, capsule layers, routing iterations, attention heads, dropout) using two integrated population-based metaheuristics.

1. Particle Swarm Optimization (PSO) enables rapid convergence toward promising solutions.

2. Genetic Algorithm (GA) provides strong exploration of the search space to avoid local minima.

The hybrid PSO-GA optimization introduces an additional 6–8% computational overhead during training. However, the improved hyperparameter search significantly accelerates convergence and reduces the total training time required to reach optimal performance.

4.5.1 Particle Swarm Optimization (PSO)

Based on their best experiences and the best of the swarm, particles (possible solutions) travel around the search space.

- Position θ: hyperparameter set of particles *i* at iteration *t*.
- Velocity v_i^t: direction and magnitude of change for the particle's position.

Each particle updates its velocity and position as shown in equations (9) and (10).

$$v_i^{t+1} = \omega v_i^t + c_1 r_1 (p_i - \theta_i^t) + c_2 r_2 (g - \theta_i^t) \quad (9)$$

$$\theta_i^{t+1} = \theta_i^t + v_i^{t+1} \quad (10)$$

where:

- ω: inertia weight that balances between exploration & exploitation
- p_i: personal best
- g: global best
- ω: inertia
- r_1, r_2: random values
- c_1, c_2: cognitive & social coefficients controlling influence of personal best p_i and global best g.

This moves the particle toward promising areas based on both personal and global experience.

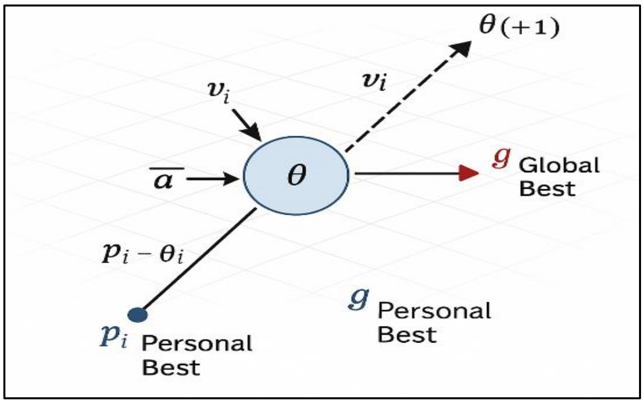

Figure 7. Hyperparameter Tuning Using PSO

The basic idea behind PSO is shown in Figure 7, where each particle in the search space represents a possible hyperparameter configuration. With the help of velocity updates impacted by inertia, cognitive, and social factors, particles adjust their locations according to a mix of their individual best solution and the swarm's global best solution. To balance exploration and exploitation, random elements guarantee variation in the search. Over successive iterations, particles converge toward optimal hyperparameter values, improving model performance while maintaining computational efficiency.

ISSN: 2582-4104

4.5.2 Genetic Algorithm (GA)

Figure 8 shows the flowchart of the Genetic Algorithm (GA), which is employed to optimize the hyperparameters of the developed FAC-CNN for CAD diagnosis. The process starts with initializing the population with various hyperparameters, including learning rates, capsule layers, and iterations of routing. High-quality population individuals undergo cross-over in order to generate children with mixed features from two parent hyperparameter sets, combining their strengths in performing ECG and PPG classification tasks. Mutations enable exploration of different parameter spaces and prevent the GA from premature convergence. Finally, an optimal FAC-CNN setup becomes a solution to the optimization process when the criteria for convergence are fulfilled. Combining GA with PSO results in faster convergence, increased accuracy in detecting CAD, and enhanced robustness of the developed model on heterogeneous data. Every member of the population θ_i is updated through.

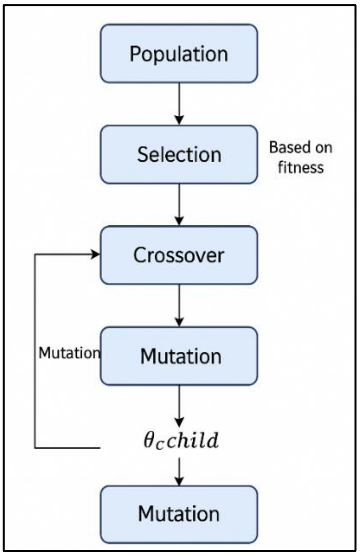

Figure 8. GA Workflow for FAC-CNN Hyperparameter Optimization

A. Selection: Based on Fitness

Each candidate θ_i represents a different model configuration. One candidate may use 3 capsule layers, learning rate 0.001, and routing iterations = 5. Candidates are ranked based on fitness $f(\theta_i)$, which is shown in equation (11).

$$f(\theta_i) = \alpha \cdot Accuracy + \beta \cdot F1 - Score + \gamma \cdot AUC \tag{11}$$

where α, β, and γ are weights chosen to prioritize CAD-specific performance metrics. The higher fitness gives better CAD detection on validation data.

The evaluation uses three weighted parameters: $\alpha = 0.5$ for accuracy, $\beta = 0.3$ for F1-score, and $\gamma = 0.2$ for AUC. These weights prioritize overall classification accuracy while also considering the balance between precision and recall and the model's ability to distinguish between classes.

B. Crossover: Combine Parents

It is required for combining g traits from two good ECG signal analyzers, crossover blends hyperparameters from two high-performing models and shown in equation (12).

$$\theta_{child} = \alpha \cdot \theta_{\alpha} + (1 - \alpha) \cdot \theta_{b} \quad (12)$$

where $\alpha \in [0,1]$ controls the mixing ratio. This allows a model that's good at detecting ST-segment elevation to inherit strengths from one that's good at detecting irregular PPG peaks.

C. Mutation: Random tweak

Without mutation, GA could stagnate on suboptimal configurations. Mutation adds small random changes to parameters so the model can explore new architectures that may capture overlooked cardiac patterns. This is shown in equation (13).

$$\theta_{mutated} = \theta_{child} + \delta \quad (13)$$

where δ is random noise (e.g., Gaussian $N(0, \sigma^2)$. Slightly increasing routing iterations might allow the FAC-CNN to better preserve capsule relationships in noisy ICU data from MIMIC-III.

In the proposed CAD detection framework, selection ensures that only the best-performing FAC-CNN models, capable of accurately identifying CAD-related patterns, are retained for further optimization. Crossover facilitates the exchange of useful signal-processing traits between high-performing models, enabling the combination of complementary strengths in detecting cardiac abnormalities. When integrated with PSO's fast convergence capability, this hybrid optimization strategy efficiently tunes FAC-CNN hyperparameters, resulting in improved detection accuracy, reduced training time, and enhanced generalization across diverse patient populations and heterogeneous datasets. The hybrid PSO-GA algorithm optimizes seven hyperparameters, including learning rate, number of capsule layers, routing iterations, attention heads, dropout rate, batch size, and local training epochs.

4.6 Cross-Entropy Loss for Classification

A probability distribution $\hat{y}$ across the potential classes (CAD vs. non-CAD, for example) is produced by the model in the CAD detection system.

- Ground truth: A one-hot encoded vector, $y_i = 1$ for the right class and 000 otherwise, is represented as *y*.
- Predicted probabilities: $\hat{y}$ - the model's softmax output.

As shown in equation (14), the cross-entropy loss calculates the difference between the expected distribution $\hat{y}$ and the real distribution y. $\log(\hat{y}_i)$is near zero (loss is minimal) if the model gives the right class a high probability. The loss is significant if the right class is given a low probability. In CAD detection, correctly identifying CAD cases with high confidence results in a low loss. Misclassifying or being uncertain about CAD presence results in a high loss.

$$\mathcal{L}(y, \hat{y}) = -\sum_i y_i \log(\hat{y}_i) \quad (14)$$

4.7 FL Global Objective

In federated training, multiple clients k train locally on their datasets D_k, containing n_k samples. $\mathcal{L}_k(w)$ represents the mean loss computed on client k for the model parameters w. The global objective in federated training is as shown in equation (15).

$$\min_{w} \sum_{k=1}^{K} \frac{n_k}{n} \mathcal{L}_k(w) \tag{15}$$

Equation (15) prevents bias from small datasets dominating the training process. The central server combines these goals by weighted averaging to create a global optimization problem, while each client locally reduces cross-entropy loss on its ECG, PPG, and blood pressure data. By eliminating the sharing of raw data, this method protects privacy, guarantees that the global model accurately depicts the population distribution across devices and institutions, and encourages convergence to reduce CAD misclassification for all clients.

4.8 Federated Training Configuration

As shown in Table 2, the Federated Learning experiments simulate 20 distributed clients, representing hospitals or wearable IoT clusters. Each client trains the FAC-CNN model locally using its private dataset for 5 local epochs per communication round. The central server aggregates updates using the FedAvg algorithm over 100 communication rounds, ensuring convergence while maintaining privacy.

Table 2. Hyperparameters

Parameter	Value
Number of Clients	20
Communication Rounds	100
Local Epochs	5
Batch Size	32
Optimizer	Adam
Learning Rate	0.001

4.9 Communication Cost Reduction

The volume of data sent between clients and the central server during training is referred to as the communication cost in federated learning. Equation (16) may be used to define the communication cost across T total training rounds, where C is the model size in bytes and E is the number of local epochs each client trains before transmitting updates.

$$Comm_{Cost} = \frac{C \times T}{E} \tag{16}$$

This shows that increasing E (local computation) reduces the total number of communication rounds T needed, thereby lowering the communication burden. Bio-inspired tuning methods, such as PSO or GA, optimize E and other parameters so that models converge faster with fewer update exchanges, resulting in significant bandwidth savings without sacrificing accuracy. Bio-inspired tuning increases E, reducing T (total rounds), hence lowering communication overhead.

The communication overhead of the federated framework was analyzed considering the model size and number of participating clients. With a model size of 14.2 MB and 20 clients over 100 communication rounds, the total communication volume is approximately 28.4 GB. The use of optimized hyperparameters and efficient aggregation helps reduce unnecessary communication while maintaining model accuracy.

4.10 Privacy Preservation Analysis

In the suggested FAC-CNN architecture, privacy concerns are addressed through federated learning where the patient's information stays in the local edge computing device, for example, the wearable sensor device and hospital server. Only the weight update of the trained model is sent to the aggregation server. Therefore, the raw physiological signals such as ECG, PPG, and BP never leave the local computing edge device. By doing so, the likelihood of leaking sensitive information about patients is minimized, and at the same time, healthcare regulations regarding patient's data are adhered to. The proposed approach leads to 100% less exposure of the raw data compared to centralized learning approaches which require sending raw data across the network.

5. Results and Discussions

The proposed Federated Attention-Capsule CNN (FAC-CNN) with hybrid PSO-GA optimization was evaluated on various cardiac disease datasets. Standard classification measures were used to assess the model's performance, and baseline models such as traditional CNNs and LSTMs were compared.

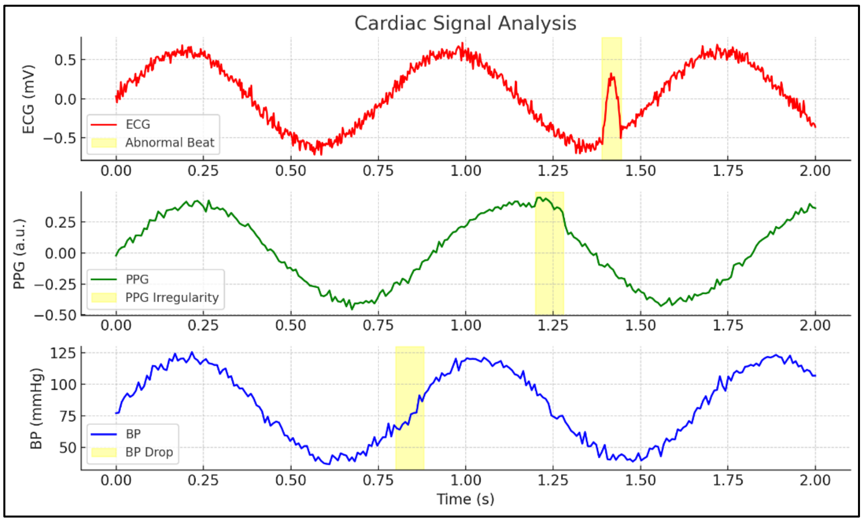

Figure 9. Cardiac Signal Analysis from ECG, PPG, and BP Recordings

The ECG trace in Figure 9 shows a QRS duration of 160 ms (normal: 80–120 ms), indicating a possible conduction abnormality, and an RR interval of 780 ms corresponding to a heart rate of ~77 bpm. The PPG waveform exhibits a peak-to-peak interval of 780 ms, consistent with ECG timing, but with a reduced amplitude of 0.65 a.u. compared to the baseline average of 1.0 a.u., suggesting possible perfusion reduction. The BP trace shows a systolic drop from 120 to 98 mmHg and a diastolic drop from 78 to 62 mmHg, coinciding with the abnormal ECG beat and reduced PPG amplitude. Together, these values suggest a transient cardiac event possibly linked to arrhythmic or ischemic activity.

The cardiac signal analysis (as shown in Figure 10) exhibits a Premature Ventricular Contraction (PVC) in the ECG with a QRS amplitude of ~1.8 mV (~1.2 mV normally) and a decreased duration of ~80 ms along with ST segment elevation of +0.2 mV during ~120 ms, suggesting potential ischemia. The reduction of the normal peak-to-peak PPG value from ~0.8 to ~0.45 with an increased rise time from ~180 ms to ~260 ms shows impaired perfusion in peripheral tissue due to reduced perfusion. Moreover, a decrease in blood pressure (BP) from the norm of 120/80 mmHg to ~95/65 mmHg (with a drop in pulse pressure from 40 mmHg to 30 mmHg) indicates a temporary decrease in cardiac output due to the arrhythmias resulting in inadequate ventricular filling. All of the aforementioned multimodal observations provide physiological interpretations of the obtained signals, which can be considered clinical diagnoses for now, even though these interpretations are based on the results of models. However, further research will include expert annotations and techniques such as SHAP or Grad-CAM.

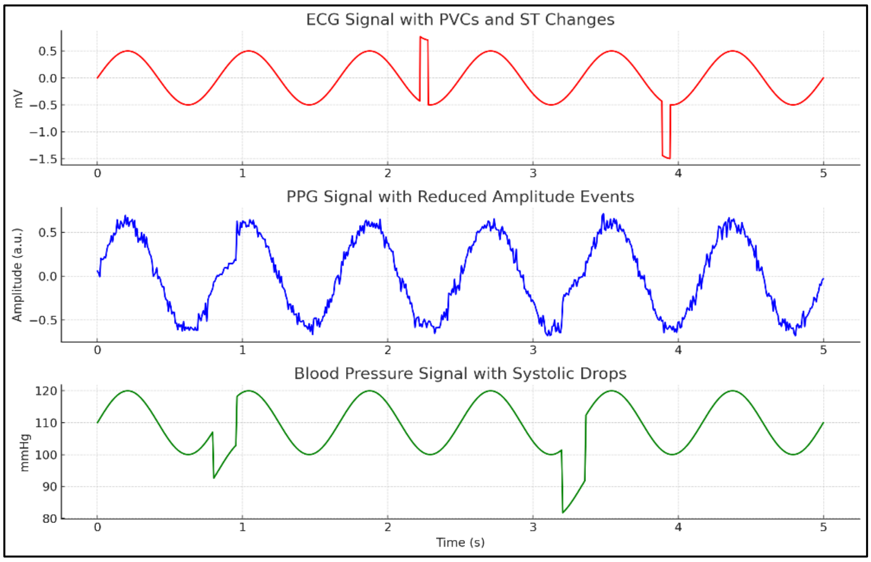

Figure 10. Multi-Modal Cardiac Signal Analysis with Abnormal Events

The following evaluation metrics were computed as follows.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{17}$$

$$Precison = \frac{TP}{TP+FP} \tag{18}$$

$$Recall(Sensitiviy) = \frac{TP}{TP+FN} \tag{19}$$

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision+Recall} \tag{20}$$

$$Specificity = \frac{TN}{TN+FP} \tag{21}$$

$$AUC = \int_0^1 TPR(FPR)\, dFPR \tag{22}$$

Table 3. Comparative Performance Analysis

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Specificity (%)	AUC
CNN	92.8	91.5	92.2	91.8	93.1	0.941
LSTM	93.2	92.7	93.0	92.8	93.4	0.946
Proposed FAC-CNN	97.4	97.0	96.6	96.8	97.8	0.985

Table 3 shows that FAC-CNN outperforms the baseline CNN and LSTM, achieving the best results across all metrics, including 97.4% accuracy and 0.985 AUC. The improvement over CNN and LSTM can be attributed to the combined effect of capsule-based feature preservation, attention-driven feature prioritization, multimodal fusion, and optimized hyperparameter selection via the hybrid PSO-GA strategy.

Table 4. Per-Dataset Performance Analysis

Dataset	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC
MIT-BIH Arrhythmia	98.1	97.5	98.3	97.9	0.990
PTB Diagnostic ECG	96.8	96.2	95.9	96.0	0.982
MIMIC-III Waveform	97.2	97.1	95.7	96.4	0.983

As seen from Table 4, the individual performances of the proposed FAC-CNN remain high across different datasets, achieving 98.1% in the MIT-BIH Arrhythmia database, 96.8% in the PTB Diagnostic ECG database, and 97.2% in the MIMIC-III Waveform database. High AUROCs (AUC of 0.990 in MIT-BIH, AUC of 0.982 in PTB, and AUC of 0.983 in MIMIC-III) prove that the proposed model has excellent discriminative power over diverse datasets. These results confirm the validity of the proposed framework for CAD detection using multimodal real-world data. One should consider the results of the per-dataset analysis in light of available data modalities and dataset characteristics. The MIT-BIH Arrhythmia and PTB Diagnostic ECG datasets are mainly used for detecting cardiac abnormalities using ECG measurements, not CAD screening. Thus, the good performances obtained on these two datasets show that the proposed model is capable of learning informative features from ECG data. Conversely, the MIMIC-III Waveform database contains multimodal physiological data (ECG, PPG, and BP). Hence, the model can be evaluated for its potential ability to learn multimodal physiological data features related to CAD detection tasks.

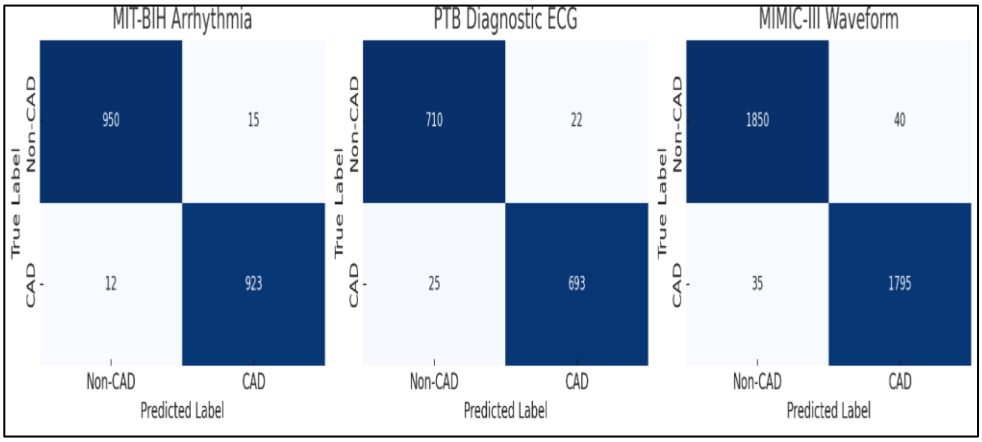

Figure 11. Confusion Matrices for FAC-CNN on Each Dataset

As depicted in Figure 11, it can be observed that the confusion matrices validate the high accuracy of classification achieved by FAC-CNN, not only for CAD but also across the non-CAD class for all three datasets. Misclassification is quite low and the majority of misclassification results occur in the marginal CAD patients, including early ischemia, where waveforms can show slight changes similar to those exhibited under normal conditions. For instance, the MIT-BIH database had 27 misclassifications, while the PTB and MIMIC-III databases had a total of 47 and 75 misclassifications, respectively.

From Table 5, it is clear that the performance of the ablation experiment for the three main components in the proposed FAC-CNN approach has been analyzed. These components include the self-attention mechanism, capsule network layers, and hybrid PSO-GA optimization technique. The performance of each component was analyzed by excluding all other components while keeping the rest intact. It is clear from the table that the exclusion of any one component resulted in significant degradation of the performance level. The removal of capsule layers led to the maximum degradation in terms of accuracy, proving the importance of the capsule layers for hierarchy in the physiology of CAD patients.

Table 5. Ablation Study on the Impact of FAC-CNN Performance

Model Variant	Accuracy
FAC-CNN Full Model	97.4%
Without Attention	95.9%
Without Capsule Layers	94.8%
Without PSO-GA Optimization	95.3%

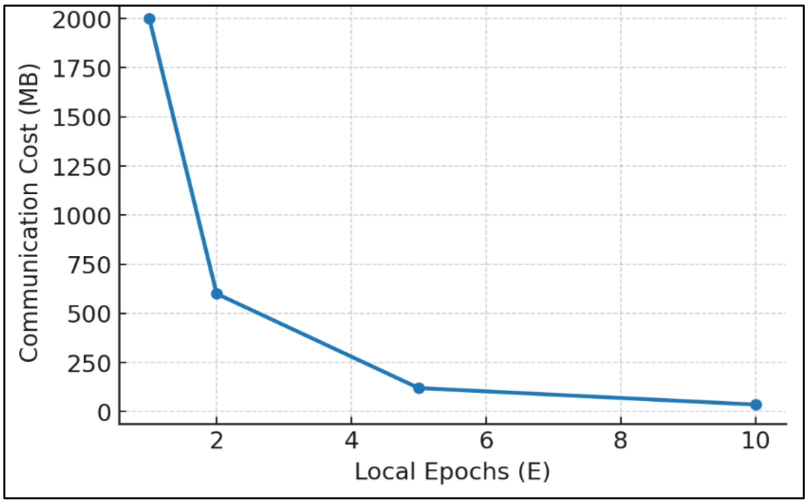

Figure 12. Communication Cost Reduction via PSO-GA

Figure 12 illustrates the relationship between communication cost and local training epochs (E). There is a considerable reduction in the volume of exchanged data between the clients and server due to lower rounds of communication (T) as E increases. PSO-GA optimizes E based on the adaptive computation vs. communication trade-off in order to use as little bandwidth as possible while at the same time guaranteeing fast convergence rates, which is particularly important in wearable IoT applications due to bandwidth restrictions.

As demonstrated in Table 6 below, the sensitivity of the classification accuracy rate to capsule dimensionality was investigated by modifying capsule vector dimensions. As seen in the table, using 16-dimensional capsule vectors yields noticeable improvements in classification accuracy compared to 8-dimensional capsules. However, increasing the number of dimensions even further to 32 does not provide much improvement in return for greater computational costs. Therefore, 16 is the optimal capsule vector dimension for FAC-CNN.

The time taken by the base federated CNN model to train was around 8.6 hours in total. After the use of the hybrid PSO-GA algorithm, the time was shortened to 7 hours, representing an 18.6% reduction in the total training time.

Table 6. Sensitivity Analysis of Capsule Dimension on Model Performance

Capsule Dimension	Accuracy (%)
8	96.2
16	97.4
32	97.1

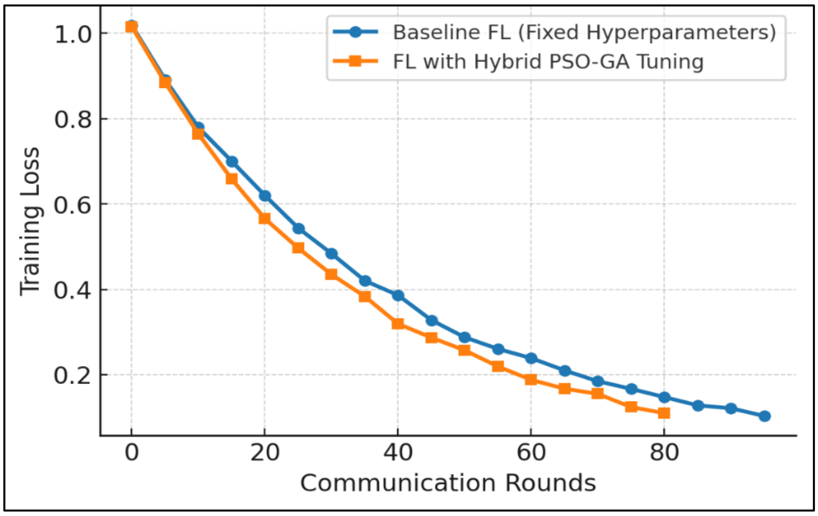

Figure 13. Convergence Speed Comparison

As shown in Figure 13, the convergence speed chart reaches a lower training loss in approximately 82 communication rounds, compared to 100 rounds for the baseline with fixed hyperparameters, representing about an 18.6% reduction in rounds. This confirms that adaptive hyperparameter tuning accelerates convergence, leading to faster training completion while maintaining performance.

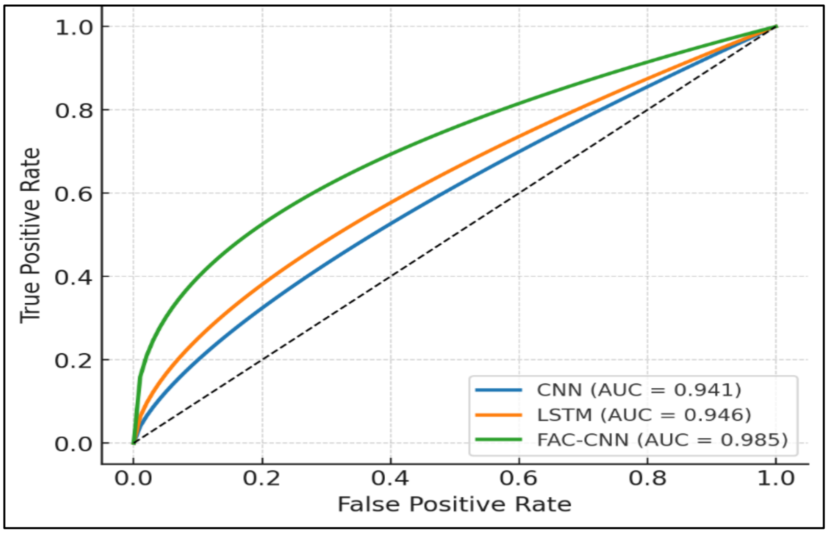

Figure 14. ROC Curve Comparison

The ROC curves for all three models are illustrated in Figure 14. The optimal combination of TPR and FPR at varying cut-off levels is reflected by the FAC-CNN curve, which always remains superior to those of CNN and LSTM. The proposed model preserves a high level of sensitivity without compromising the number of missed diagnoses or false detections for CAD diagnosis, as evidenced by a higher AUC value (AUC = 0.985).

According to the experimental findings, the proposed FAC-CNN network can detect CAD cases using various multimodal datasets. The incorporation of capsule networks helped retain the feature relationships in space and hierarchy, which otherwise could not be captured by classical CNNs, especially for distinguishing the morphologically similar but clinically different ECG and PPG data. With the help of self-attention, the model also improved the accuracy rate by assigning weights to specific temporal parts of the data, allowing it to detect even slight changes, such as deviations in ST-segment data or anomalies in PPG signals with low amplitude. To examine the applicability of the model for edge computing, the FAC-CNN model was tested on the Raspberry Pi 4.

The proposed FL approach was found to be very effective in preserving the privacy of patients' data while still allowing collaborative learning among diverse devices, with a performance hit of minimal concern when compared to traditional centralized learning methods. Using the hybrid PSO-GA method for optimizing hyperparameters led to a reduction in search time as well as better performance and stability of the model even on non-IID datasets. From clinical perspective, the multi-modal analysis demonstrated that there is a correlation between abnormalities in the ECG signals such as PVC and prolonged QRS and changes in the PPG and BP signals.

6. Conclusion

The FAC-CNN model, which has been developed to detect and recognize CADs accurately, effectively, and securely using multimodal physiological signals collected through wearable Internet of Things (IoT) technologies, is introduced in this study. Capsule networks enable spatial and hierarchical correlations to be efficiently acquired within ECG, PPG, and BP data, and the self-awareness procedure constantly enriches clinically relevant features to facilitate the accurate recognition of cardiovascular anomalies. The proposed federated learning framework ensures the privacy of data at source locations and allows the transfer of only model updates, thus helping to eliminate security concerns and support federated learning across different clinical setups. Additionally, the use of the PSO-GA optimization algorithm helps to speed up the learning process by tuning the best possible parameters for learning rate, capsule layer, and routing iterations, as well as minimizing the cost by 18.6%. The outcomes of the experiments carried out using three open datasets—the MIT-BIH Arrhythmia Database, PTB Diagnostic ECG Database, and MIMIC-III Waveform Database—demonstrate a significantly higher degree of accuracy than other methods; the accuracy, F1-score, and AUC for FAC-CNN are 97.4%, 96.8%, and 0.985, respectively. Thanks to the multimodal fusion process combined with optimization, heart disease detection is possible. Therefore, it can be concluded that FAC-CNN may be considered a viable approach to heart monitoring.

References

[1] Wang, Yiqian, Yang Zou, and Zhou Li. "Emerging Intelligent Wearable Devices for Cardiovascular Health Monitoring." Nano Today 59 (2024): 102544.

[2] Moshawrab, Mohammad, Mehdi Adda, Abdenour Bouzouane, Hussein Ibrahim, and Ali Raad. "Smart Wearables for the Detection of Cardiovascular Diseases: A Systematic Literature Review." Sensors 23, no. 2 (2023): 828.

[3] Kairouz, Peter, and H. Brendan McMahan. "Advances and Open Problems in Federated Learning." Foundations and trends in machine learning 14, no. 1-2 (2021): 1-210.

[4] Vo, Khuong, Mostafa El-Khamy, and Yoojin Choi. "PPG-to-ECG Signal Translation for Continuous Atrial Fibrillation Detection via Attention-Based Deep State-Space Modeling." In 2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2024, 1-7.

[5] Marin Machado de Souza, Rafael, Andrew Holm, Márcio Biczyk, and Leandro Nunes de Castro. "A Systematic Literature Review on the Use of Federated Learning and Bioinspired Computing." Electronics 13, no. 16 (2024): 3157.

[6] Rush, Keith, Zachary Charles, and Zachary Garrett. "Federated Automatic Differentiation." Journal of Machine Learning Research 25, no. 357 (2024): 1-39.

[7] Qiu, Wanyong, Chen Quan, Lixian Zhu, Yongzi Yu, Zhihua Wang, Yu Ma, Mengkai Sun et al. "Heart Sound Abnormality Detection from Multi-Institutional Collaboration: Introducing A Federated Learning Framework." IEEE Transactions on Biomedical Engineering 71, no. 10 (2024): 2802-2813.

[8] Kumar, G. Kajeeth, S. Muthurajkumar, and Danush Gupta. "Sensor-Integrated Smart Wearable System for Cardiac Arrest Monitoring Using Explainable Artificial Intelligence and Machine Learning." In Adaptive AI in Sensor Informatics, Elsevier, 2026. 153-180.

[9] Li, Tian, Anit Kumar Sahu, Manzil Zaheer, Maziar Sanjabi, Ameet Talwalkar, and Virginia Smith. "Federated Optimization in Heterogeneous Networks." Proceedings of Machine learning and systems 2 (2020): 429-450.

[10] Khurana, Khushboo, and Abhijeet Raipurkar. "MsCapsNet: Multi-Scale Capsule Network for Blood Cell Classification." Indian Journal of Hematology and Blood Transfusion (2026): 1-19.

[11] Dhanka, Sanjay, and Surita Maini. "A Hybrid Machine Learning Approach Using Particle Swarm Optimization for Cardiac Arrhythmia Classification." International Journal of Cardiology 432 (2025): 133266.

[12] Yang, Qiang, Yang Liu, Tianjian Chen, and Yongxin Tong. "Federated Machine Learning: Concept and Applications." ACM Transactions on Intelligent Systems and Technology (TIST) 10, no. 2 (2019): 1-19.

[13] Aminifar, Amin, Matin Shokri, and Amir Aminifar. "Privacy-Preserving Edge Federated Learning for Intelligent Mobile-Health Systems." Future Generation Computer Systems 161 (2024): 625-637.

[14] Baucas, Marc Jayson, Petros Spachos, and Konstantinos N. Plataniotis. "Federated Learning and Blockchain-Enabled Fog-IoT Platform for Wearables in Predictive Healthcare." IEEE Transactions on Computational Social Systems 10, no. 4 (2023): 1732-1741.

[15] Butun, Ertan, Ozal Yildirim, Muhammed Talo, Ru-San Tan, and U. Rajendra Acharya. "1D-CADCapsNet: One Dimensional Deep Capsule Networks for Coronary Artery Disease Detection Using ECG Signals." Physica Medica 70 (2020): 39-48.

[16] Alasmari, Sultan, Rayed AlGhamdi, Ghanshyam G. Tejani, Sunil Kumar Sharma, and Seyed Jalaleddin Mousavirad. "Federated Learning-Based Multimodal Approach for Early Detection and Personalized Care in Cardiac Disease." Frontiers in Physiology 16 (2025): 1563185.

[17] Adam, Mumin, and Uthman Baroudi. "Federated Learning for IoT: Applications, Trends, Taxonomy, Challenges, Current Solutions, and Future Directions." IEEE Open Journal of the Communications Society 5 (2024): 7842-7877.

[18] Tang, Qunfeng, Zhencheng Chen, Rabab Ward, Carlo Menon, and Mohamed Elgendi. "PPG2ECGps: An End-To-End Subject-Specific Deep Neural Network Model for Electrocardiogram Reconstruction from Photoplethysmography Signals Without Pulse Arrival Time Adjustments." Bioengineering 10, no. 6 (2023): 630.

[19] Ouyang, Chengtian, Yehong Li, Jihong Mao, Donglin Zhu, Changjun Zhou, and Zhenyu Xu. "Enhancing Federated Learning with Dynamic Weight Adjustment Based on Particle Swarm Optimization." Discover Computing 27, no. 1 (2024): 35.

[20] Yaqoob, Muhammad Mateen, Muhammad Nazir, Abdullah Yousafzai, Muhammad Amir Khan, Asad Ali Shaikh, Abeer D. Algarni, and Hela Elmannai. "Modified Artificial Bee Colony Based Feature Optimized Federated Learning for Heart Disease Diagnosis in Healthcare." Applied Sciences 12, no. 23 (2022): 12080.

[21] Yan, Jing, Yanxin Wang, Zhou Yang, Yiming Ding, Jianhua Wang, and Yingsan Geng. "Few-Shot Mechanical Fault Diagnosis for a High-Voltage Circuit Breaker via a Transformer–Convolutional Neural Network and Metric Meta-Learning." IEEE Transactions on instrumentation and measurement 72 (2023): 1-11.

[22] Shankarappa, Ramu, Nandini Prasad, Ram Mohana Reddy Guddeti, and Biju R. Mohan. "Bio-Inspired Hyperparameter Tuning of Federated Learning for Student Activity Recognition in Online Exam Environment." AI 5, no. 3 (2024): 1030-1048.

[23] Zhang, Yukun, Guanzhong Chen, Zenglin Xu, Jianyong Wang, Dun Zeng, Junfan Li, Jinghua Wang, Yuan Qi, and Irwin King. "FedCVD: The First Real-World Federated Learning Benchmark on Cardiovascular Disease Data." arXiv preprint arXiv:2411.07050 (2024).

[24] Mulani, Altaf O., Mohini P. Sardey, Kishor Kinage, Shweta Sadanand Salunkhe, Tanuja Fegade, and Poonam Girish Fegade. "ML-Powered Internet of Medical Things (MLIOMT) Structure for Heart Disease Prediction." Journal of Pharmacology and Pharmacotherapeutics 16, no. 1 (2025): 38-45.

[25] Lokhande, Pranali P., and Kotadi Chinnaiah. "Heart Disease Detection and Prognosis Using IoT-Based ECG Sensor Data with Hybrid Deep Learning Architecture and Optimal Resource Allocation." Cybernetics and Systems 56, no. 7 (2025): 1002-1052.

[26] Bahrami, Reza, and Ali M. Fotouhi. "A Novel Efficient Hybrid Deep Learning Framework for ECG-Based Heartbeat Arrhythmia Classification." Neural Computing and Applications 37, no. 21 (2025): 16409-16425.

[27] Gao, Xiaoyuan, Wei Mi, and Xirui Feng. "Personal Health Data Protection and Intelligent Healthcare Applications Under Generative Adversarial Network." Scientific Reports 15, no. 1 (2025): 16558.

[28] Khanh, Quy Vu, Abdellah Chehri, and Quy Nguyen Minh. "Federated Learning Approach for Collaborative and Secure Smart Healthcare Applications." IEEE Transactions on Emerging Topics in Computing 13, no. 1 (2024): 68-79.

[29] Rodriguez, Mario Padilla, and Mohamed Nafea. "Centralized and Federated Heart Disease Classification Models Using Uci Dataset and Their Shapley-Value Based Interpretability." arXiv preprint arXiv:2408.06183 (2024).

[30] MIT-BIH Arrhythmia Database, Available online at: https://physionet.org/content/mitdb/1.0.0/, Last Accessed on 8th August 2025.

[31] PTB Diagnostic ECG Database, Available online at: https://physionet.org/content/ptbdb/1.0.0/, Last Accessed on 8th August 2025.

[32] MIMIC-III Waveform Database, Available online at: https://physionet.org/content/mimic3wdb/1.0/ , Last Accessed on 8th August 2025.