Journal of Artificial Intelligence and Data Mining (JAIDM), Vol. 14, No. 4, 2026, 543-557.

Shahrood University of Technology

Journal of Artificial Intelligence and Data Mining (JAIDM)
Journal homepage: http://jad.shahroodut.ac.ir

Research paper

A Dual-Path Learning Framework for Real-Time Heart Disease Prediction in IoT-Fog-Cloud Environments Using CGAN and Hybrid Feature Selection

Vahidreza Afshin[1], Saeideh Kabirirad[2] and Seyed Hamid Zahiri[1*]

1. Department of Electrical Engineering, University of Birjand, Birjand, Iran.
2. Department of Computer Science, Birjand University of Technology, Birjand, Iran.

Article Info

Article History:
Received 27 November 2025
Revised 21 December 2025
Accepted 20 July 2026

DOI:10.22044/jadm.2026.17225.2858

Keywords:
Heart Disease prediction IoT–Fog–Cloud Computing, CGAN-based Data Balancing, Hybrid Feature Selection, XGBoost, TabTransformer.

**Corresponding author: hzahiri@birjand.ac.ir (H. Zahiri).*

Abstract

Accurate and timely prediction of heart disease is a fundamental challenge in clinical diagnostics, exacerbated by the inherent complexities of real-world healthcare datasets, such as class imbalance, high dimensionality and the presence of outliers. To address these issues, we propose a dual-path learning framework specifically designed for real-time heart disease prediction within IoT–Fog–Cloud environments. The framework integrates advanced preprocessing techniques, including Conditional Generative Adversarial Networks (CGANs) for data balancing and a hybrid feature selection pipeline, thereby addressing data imbalance and dimensionality reduction. It also proposes a dual-path diagnostic system that uses a lightweight XGBoost classifier in the fog layer optimized for low latency and fast local inference, and an improved TabTransformer-based model in the cloud layer, which is selectively used for ambiguous samples. Experimental evaluation on an expanded heart disease dataset (including 1,190 samples) demonstrates the robust performance of our framework, achieving 97.82% accuracy, 98.47% F1-score, and 98.65% AUC. Additionally, it improves overall inference latencies, mostly below one second in the fog layer and low latency in cloud layer. Overall, this paper provides a scalable and interpretable medical diagnosis framework for distributed healthcare systems.

1. Introduction

Cardiovascular disease (CVD) is a leading cause of death worldwide, accounting for approximately 17.9 million deaths annually. Early detection and continuous screening are crucial for preventing disease progression and enabling targeted interventions.

However, traditional healthcare systems, which rely primarily on periodic clinical visits and manual interpretation of tests, are often inadequate for prescribing responsive and personalized care [1,2]. Internet of Things (IoT) technology enables the collection of healthcare data and patient vital signs such as heart rate, blood pressure, and electrocardiogram (ECG) signals in real time, facilitating timely diagnosis and intervention [3,4]. Recent IoT-based healthcare studies highlight continuous patient monitoring as a key advantage [5]. However, this technology faces challenges related to the volume, latency, and diversity of data generated. Efficient storage and analysis of high-dimensional/high-frequency medical data in real time remain complex tasks.[1-5]

Cloud computing provides scalable infrastructures to manage large volumes of healthcare data and can deploy machine learning (ML) models for pattern recognition and prediction [6,7]. Despite offering scalability and high computational power, cloud-based solutions face challenges such as high latency, limited bandwidth, and privacy concerns. These issues are particularly critical in healthcare, where timely decision-making and regulatory compliance are essential [7,8]. Cloud-based frameworks have also reported latency and privacy challenges in medical applications [9]. To alleviate these limitations, fog computing has emerged as a promising intermediate layer between IoT endpoints and the cloud. By enabling local computing and storage near the data source, fog computing reduces response times, mitigates privacy risks, and supports real-time processing [10,11].

Fog computing has recently been validated in healthcare IoT deployments for rapid decision-making [12]. However, the limited resources of fog nodes restrict the deployment of complex ML models, necessitating the development of lightweight and efficient inference techniques tailored to fog environments [13,14].

Some recent studies have proposed hybrid fog-cloud architectures to balance inference speed and model accuracy [15, 16]. In these models, fog nodes quickly handle high-confidence cases, while more ambiguous cases are routed to the cloud for deeper analysis. However, many existing systems lack advanced feature selection mechanisms and adaptive routing strategies that consider both performance and resource efficiency [17, 18]. Hybrid models can enable adaptive routing for efficient processing of healthcare data [19].

To address these gaps, we propose a novel dual-path learning framework that combines a lightweight XGBoost classifier in the fog layer with an enhanced transformer-based ensemble model in the cloud. The fog layer employs a low-latency classification based on a compact feature set selected through a combined strategy of chi-square filtering, recursive feature elimination (RFE), and random forest importance ranking. In the cloud layer, a transformer model processes complex and uncertain cases using self-attention mechanisms to capture nonlinear feature dependencies and improve diagnostic accuracy. A confidence-aware routing mechanism plays a fundamental role in our architecture by detecting uncertain instances. This hierarchical structure reduces overall latency, required bandwidth, and cloud dependency.

Experimental results demonstrate that our method achieves high performance in terms of accuracy, latency, and interpretability. Compared to existing methods, our framework offers a balanced solution that supports real-time responsiveness, model transparency through SHAP analysis [20], and regulatory compliance via edge-level anonymization. Unlike previous works that have explored hybrid IoT–Fog–Cloud frameworks for healthcare [21,22], the proposed framework introduces several novel contributions. First, we present a confidence-aware dynamic routing mechanism that intelligently delegates low-confidence predictions from fog nodes to the cloud, enabling real-time responsiveness without sacrificing accuracy. Second, we employ a Conditional GAN for minority class balancing, which is rarely addressed in fog-based systems. Third, a hybrid feature selection strategy improves both inference efficiency and interpretability. To the best of our knowledge, these elements have not been jointly applied in previous healthcare frameworks within the fog-cloud domain. This integration supports combining fog and cloud computing paradigms with deep learning methods to enhance clinical decision-making [20,21].

1.1. Contributions

This paper introduces a novel and adaptive Dual-Path Learning Framework for real-time heart disease prediction in IoT–Fog–Cloud environments. The key contributions are as follows:

1. We propose a two-layer diagnostic system that integrates a lightweight XGBoost classifier in the fog layer, optimized for low latency and fast local inference, alongside an improved TabTransformer model in the cloud layer. The cloud model is selectively activated for low-confidence or ambiguous cases, balancing efficient fog-level responsiveness with high-accuracy global inference.
2. To reduce dimensionality and improve interpretability, a three-phase feature selection pipeline is employed. This approach effectively reduces the feature space by approximately 68%, while preserving clinically essential variables, thereby increasing efficiency and facilitating transparent diagnosis.
3. Addressing class imbalance, a Conditional Generative Adversarial Network (CGAN) is used to generate realistic and clinically meaningful synthetic samples of minority classes. Unlike conventional oversampling methods, the CGAN captures complex feature interdependencies, increases training data diversity, and enhances the model's generalizability.
4. The framework incorporates SHapley Additive exPlanations (SHAP) to interpret and quantify individual feature contributions to model predictions. This interpretability enhances clinical confidence and supports integration with

healthcare professionals' workflows by providing transparent diagnostic insights.
5. Experimental evaluation on an aggregated heart disease dataset comprising 1,190 patient records with 13 clinical attributes (sourced from multiple standardized collections via GitHub repository [22]) confirms the superior predictive capabilities of this framework, achieving: Accuracy: 97.82%, F1-score: 98.47%, AUC: 98.65%, Brier score: 0.0052, and edge inference latency ≤1s in over 93% of cases.

1.2. Structure of the Paper
The rest of the paper is organized as follows: Section 2 reviews related works and provides a detailed comparison of these works, including classical ML, DL, and fog/cloud-based approaches. Section 3 presents the proposed model for heart disease diagnosis within the IoT-fog-cloud ecosystem. Section 4 discusses experimental results on datasets and performance metrics. Section 5 compares the proposed model's results with related works, and finally, Section 6 concludes the paper.

2. Related Works
The use of ML techniques for effective prediction of cardiovascular diseases improves early diagnosis and thus minimizes mortality. Many studies have been conducted on the prediction of cardiovascular diseases using various types of learning machines. For example, a recent study [24] compared the effectiveness of advanced ML algorithms in heart disease prediction. The study evaluated models such as support vector machines (SVM), random forests (RF), XGBoost, and artificial neural networks (ANN) after comprehensive preprocessing steps, including feature scaling, nesting, and encoding. The results showed that ensemble models such as XGBoost and RF significantly outperformed traditional models in terms of accuracy and robustness.[23]
The development of cloud computing and IoT-based sensor technology in electronic healthcare systems has brought about remarkable progress. Unlike traditional cloud-based systems that suffer from higher delays due to data transmission to remote servers, fog-based systems enable localized decision-making, which is crucial for time-critical medical applications. By distributing computational tasks to fog nodes, the system reduces dependency on centralized cloud infrastructure, thereby decreasing latency and network overhead. Researchers have conducted studies on heart disease prediction based on advanced technologies such as IoT and fog processing, some of which are discussed below.
Among the diverse research efforts aimed at integrating fog computing into smart healthcare systems, HealthFog, proposed by [21], stands out as a foundational contribution. This study introduces a scalable and secure healthcare architecture that uses FogBus middleware to process cardiac disease data collected from IoT-enabled sensors. A key innovation of HealthFog lies in the use of ensemble deep learning methods for accurate and real-time diagnosis of heart diseases. It has been evaluated on real-world patient datasets and has demonstrated high performance in terms of accuracy, energy consumption, response time, and bandwidth efficiency.
In another study, [22] proposed a fog-based heart disease prediction system supported by a modified Gated Recurrent Unit (GRU) model. The system is designed to use fog computing on data collected from remote health monitoring devices in real time. By incorporating a modified GRU, the framework enhances the feature extraction capabilities necessary for accurate disease prediction. The architecture effectively manages the data flow between IoT devices, fog nodes, and cloud servers, and attempts to reduce latency and computational resources.
A system proposed by [25] is the integrated IoT–fog–cloud framework that focuses on real-time remote diagnosis of heart disease. The framework processes physiological data such as ECG signals and vital signs collected through wearable sensors and uses deep learning models for efficient diagnosis of heart diseases. By leveraging hierarchical distribution of computing resources, the system ensures reduced response time, reliable decision-making at the edge, and scalable data management through cloud storage.
The authors of [26] also proposed an IoT-enabled heart disease monitoring system that uses real-time streaming data collected from wearable devices, including features such as age, gender, glucose levels, blood pressure, and heart rate. The system integrates Gray Wolf Optimization for effective feature selection and Deep Belief Networks (DBNs) for classification.
In [27], an improved smart healthcare system using a cloud-based IoT framework is proposed. This approach integrates cascade deep learning (CDL) models to enhance diagnostic accuracy and robustness. By combining cloud computing with IoT-based data collection, the framework achieves faster data processing and reliable decision-making. The authors report improvements in

prediction performance and overall system efficiency compared to traditional models. In [28], the authors developed a smart health monitoring system that uses a hybrid ONBRF classifier, combining Naïve Bayes and Random Forest, to analyze data collected from Gombe State, Nigeria. Their framework emphasizes real-time processing capabilities and energy-efficient operation in distributed environments. It reports promising results in terms of accuracy and latency but does not detail the use of advanced data preprocessing, balancing, or feature selection strategies.[24, 25]

In [29], a fully automated heart disease prediction system, named LIFE-CARE, is proposed. The framework integrates IoT medical sensors and electronic health records to continuously monitor patient heart status. The collected data is transmitted to a multi-cloud infrastructure via gateways and distributed edge servers (DES). Preprocessing, including data filtering and normalization, is performed in DES to ensure data quality. Prediction is performed using a collaborative transfer learning-enabled Separable Transformer (CTL-ST) model. The model incorporates domain matching and knowledge distillation techniques to effectively handle heterogeneous and real-time input data. It achieves performance metrics with 98.72% accuracy, 98.25% precision, 98.79% recall, 98.55% specificity, and 98.78% F1-score [26].

The authors of [30] proposed a fog-based architecture for heart attack prediction using real-time data collected from wearable edge devices such as smartwatches. The system uses the Cleveland dataset to train a custom deep learning model in the cloud, which is then deployed at fog nodes for real-time inference. Their system simulates real-world deployment using iFogSim and achieves 90% accuracy.

In [29] an ML-powered Internet of medical things framework is introduced for real-time prediction of heart diseases. The system integrates wearable sensors, mobile applications, cloud platforms, and ML algorithms to enable continuous monitoring and early detection of cardiovascular abnormalities. Table 1 shows a comparative overview of state-of-the-art models developed for cardiovascular disease diagnosis based on IoT, fog, and cloud. While the studies summarized in Table 1 demonstrate the increasing use of fog and cloud computing in heart disease prediction, most of them suffer from at least one of the following limitations: lack of intelligent routing between layers, limited support for imbalanced data, or lack of interpretable decision-making systems.

Table 1. A comparison of related models based on IoT, fog and cloud developed for cardiovascular disease diagnosis.

Study	Year	Architecture	Learning Model	Data Source	Accuracy	Key Features
Tuli et al. [20]	2020	IoT–Fog	Ensemble DL (DNNs)	Real patient dataset	~93	FogBus middleware, real-time diagnosis, low latency, energy-efficient
Nancy et al. [21]	2023	IoT–Fog–Cloud	Modified Gated Recurrent Unit (GRU)	UCI Heart Disease Dataset (Cleveland, Hungarian)	99.13	Fuzzy Inference System (FIS), low latency, fog-based processing, real-time prediction
Pati et al. [27]	2023	IoT–Fog–Cloud	Ensemble DL (DNNs)	Combined dataset: Long-Beach, Cleveland, Switzerland and Hungarian datasets	97.32	FRIEND framework, ensemble deep learning, real-time remote diagnosis, low latency, energy-efficient
Sundararajan et al. [24]	2023	IoT–Cloud	Cascaded DL (CDL)	Cleveland & Framingham heart datasets	~92	Cloud-based, ensemble learning, improved accuracy, two datasets combined
Priya and Krishnan [26]	2025	IoT–Edge–Cloud	Collaborative Transfer Learning with Separable Transformer	Medical IoT sensors + EHR (multi-cloud)	98.72	Domain adaptation, knowledge distillation, real-time data processing
Ravindranathan et al. [28]	2024	IoT–Fog–Edge	Custom DL	Cleveland Dataset + Real-time wearable sensor data	90	iFogSim simulation, fog-layer DL model, edge data collection, cloud re-training
Mulani et al. [29]	2024	IoT- Fog	DL, Ensemble ML	Cleveland (300 rows)	~97	MLIoMT framework; wearable IoMT devices; real-time monitoring; 13 clinical features used

For example, although [25] and [29] achieved high accuracy using deep learning models, they did not use confidence-based routing or real-time explainability mechanisms such as SHAP. Our proposed dual-path system offers an adaptive approach that combines CGAN-based data balancing, hybrid feature selection, and interpretable learning in the fog and cloud layers Table 1[26, 27].

3. Proposed Framework and Methodology

This section outlines the architecture of the proposed dual-path learning framework, developed for real-time and accurate heart disease prediction. The model is designed for IoT–fog–cloud environments and addresses key challenges, including low-latency processing, class imbalance, high feature dimensions, and interpretable predictions. It comprises five stages:

- Data preprocessing and outlier removal via the Interquartile Range (IQR) method.
- Class balancing using Conditional Generative Adversarial Networks (CGANs).
- Hybrid feature selection using hybrid techniques.
- Fog-level inference using a lightweight XGBoost classifier.
- Cloud-level inference using an enhanced TabTransformer model.

In Figure 1, the general steps of the model are shown.Figure 1

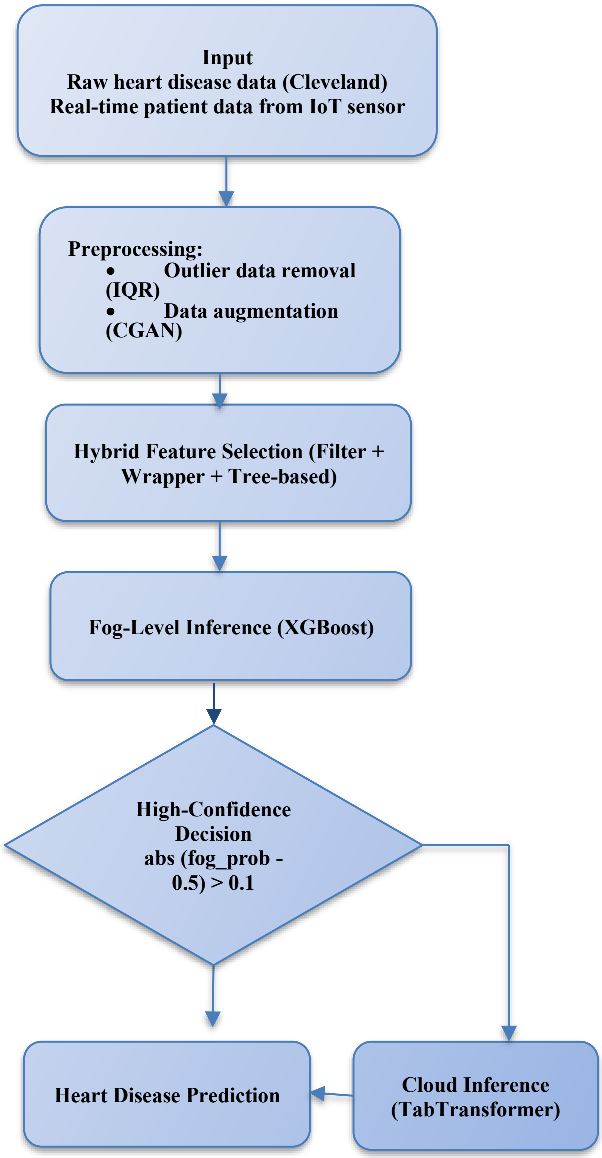

Figure 1. Overview of the proposed dual-path learning framework for heart disease prediction in IoT–Fog–Cloud environments.

3.1. Data Preprocessing

The raw data is sourced from an expanded heart disease dataset comprising 1,190 clinical records with 13 attributes, including age, sex, cholesterol level, chest pain type, resting blood pressure, and ST depression. To ensure data quality while preserving clinically meaningful samples, outlier detection was conducted using the Interquartile Range (IQR) method with a conservative threshold factor of 2.0 (rather than the conventional 1.5). Records with values beyond 2.0×IQR from Q1 or Q3 were identified and removed, resulting in the exclusion of 16 records (1.34% of the dataset) and retaining 1,174 samples for model development. This preprocessing step reduces noise from potential measurement artifacts while enhancing the generalization capability of downstream models.

IQR can be computed as follows:

$$IQR = Q_3 - Q_1$$
$$outlier\ if: x > Q_3 + k \times IQR \tag{1}$$
$$or\ x < Q_1 - k \times IQR$$

where Q_1 (Q_3) represents the first (third) quartile, k is the threshold factor (set to 2.0 in our implementation), and x is the feature value.

Figure 2 presents the distribution of outliers across clinical features following IQR-based preprocessing. The significant anomalies in "oldpeak" and "chol" indicate high variability or potential measurement artifacts in these attributes. In contrast, features such as "trestbps" and "thalach" demonstrate higher statistical consistency, emphasizing their reliability in subsequent model development.

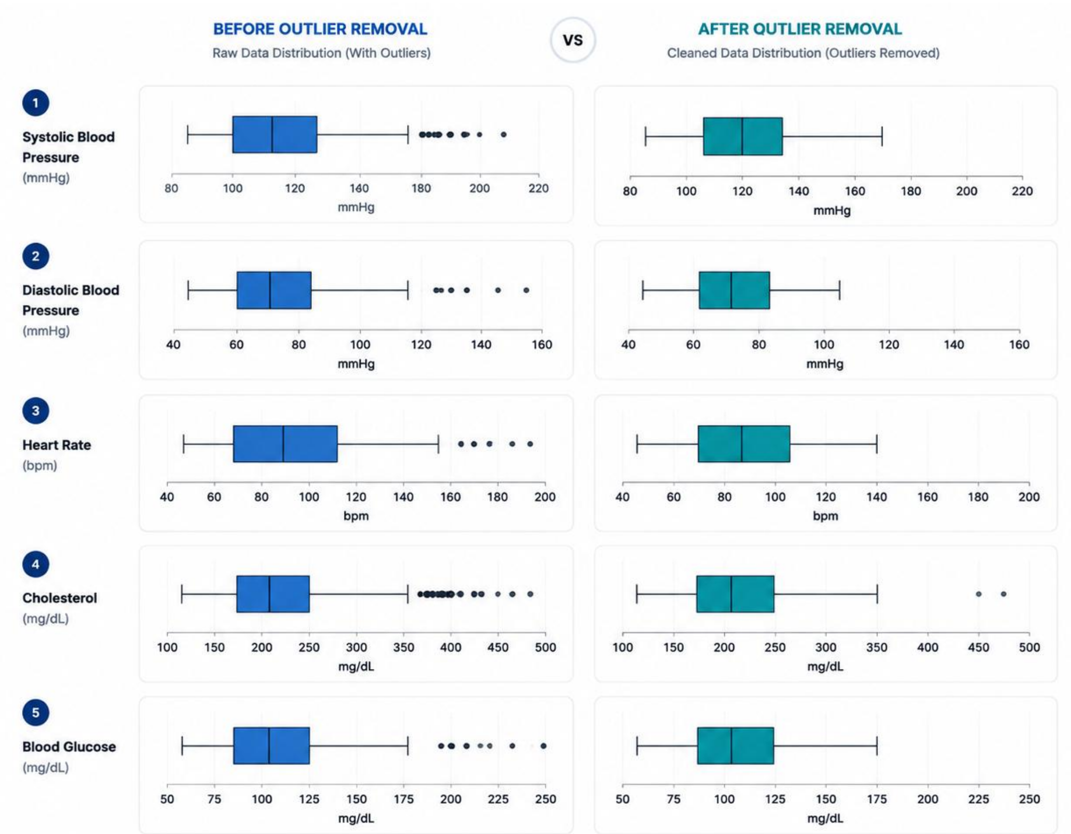

Figure 2. Count of outliers detected per feature using the IQR approach. oldpeak and chol exhibit the highest number of anomalies, whereas trestbps and thalach present the lowest.

3.2. Data Balancing with CGAN

Most medical datasets exhibit high class imbalance, with significantly fewer diseased instances than healthy ones. To address this problem, a Conditional Generative Adversarial Network (CGAN) is used to generate synthetic samples for the minority class. Unlike traditional oversampling methods such as SMOTE and ADASYN, CGAN models the joint distribution of class features and labels, enabling the generation of realistic and diverse conditional samples. This approach increases the sensitivity and recall of the model, especially in identifying true positives, without sacrificing precision.

While generative adversarial networks are typically applied to large-scale datasets, their use for augmenting medical tabular data of moderate size has been increasingly validated in recent literature [32, 33]. We generated a conservative number of 500 synthetic samples, representing approximately 42% of the original minority class instances, to mitigate overfitting risks while effectively addressing class imbalance. This approach is preferred over conventional oversampling techniques like SMOTE [34] because CGAN captures complex, non-linear feature-label relationships and preserves clinically meaningful interdependencies among diagnostic features [30-32].

To prevent data leakage, CGAN augmentation was performed strictly after the train-test split (70/30) and exclusively on the training set. The test set remained unchanged and consisted only of real examples. This ensures that all reported performance metrics reflect true generalizability on unseen patient data.

Recent studies also demonstrated that GAN-based synthetic data generation improves minority-class representation in medical datasets [35].

The CGAN framework comprises two fully connected neural networks: a generator and a discriminator, trained in an adversarial manner. The generator $G(z;\theta_G)$ creates synthetic samples from random noise z, while the discriminator $D(x;\theta_D)$ evaluates whether the input data is real or synthetic.

Generator (G): Inputs a 32-dimensional noise vector z concatenated with the minority class label (y=1), processes them through a hidden layer with 64 neurons and ReLU activation, and outputs a synthetic sample with 13 features matching the original dataset.

Discriminator (D): Inputs either a real or a synthetic sample (along with the class label) and passes it through a similar hidden layer. It uses a sigmoid output layer to distinguish real samples from synthetic ones.

The networks are optimized using binary cross-entropy loss and the Adam optimizer (learning rate = 0.001) for 300 training epochs. Therefore, the loss functions L_G (Generator loss) and L_D(Discriminator loss) are defined as follows:

$$LG = -E_z\left[\log(D(G(z|y)))\right] \quad (2)$$

where the generator G aims to minimize L_G by pushing $D(G(z))$ to 1 so that G produces high-quality and realistic data.

$$LD = -E_X\left[\log(D(x|y))\right] - E_z\left[\log(1 - D(G(z|y)))\right] \quad (3)$$

where $E_x[log(D(x|y))]$ represents the expected log probability of D correctly identifying real data (x) as real. $E_z[log(1 - D(G(z|y)))]$ represents the expected log probability of D correctly identifying synthetic data $(G(z))$ as fake. Input features were normalized using quantile transformation to stabilize convergence. The final model generated synthetic minority-class samples, which were inverse-transformed and appended to the training set.

3.3. SHAP Analysis for Interpretability

To interpret the impact of individual features on model predictions and validate the clinical relevance of the augmented data, SHApley Additive ExPlanations (SHAP) analysis was performed. A random forest classifier was trained on the CGAN-balanced dataset, and SHAP values were calculated to evaluate the contribution of the features. This analysis revealed that count of main vessels colored by fluoroscopy (ca), ST depression (oldpeak), and chest pain type (cp) were the most influential features in predicting heart disease, which correlated well with established clinical indicators. Figure 3 shows a SHAP plot of average absolute feature importance across all samples.

Figure 4 illustrates a SHAP summary scatter plot, where the color of each point represents the value of a feature (red: high, blue: low), showing its directional effect on the model output.

These results enhance transparency and interpretability, allowing healthcare professionals to better understand the model's behavior and increase confidence in its diagnostic decisions.

3.4. Fog-Based Inference with XGBoost

To enable real-time classification in resource-constrained environments, a lightweight XGBoost classifier is deployed in the fog layer. The fog model operates on a reduced feature set determined through the hybrid feature selection pipeline.

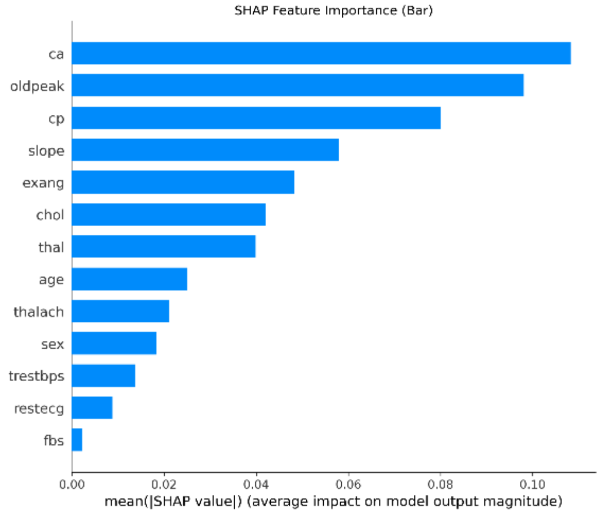

Figure 3. SHAP bar plot showing average feature importance in heart disease prediction.

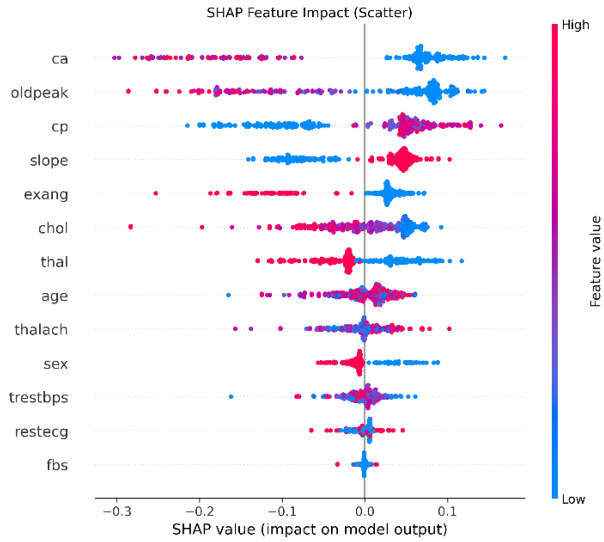

Figure 4. SHAP scatter plot revealing the directional impact of feature values.

SHAP analysis identified five influential features: ca, oldpeak, cp, ST segment slope (slope), and Electrocardiograph by resting (exang). The fog-level XGBoost model is trained solely on these features. The model achieved high classification accuracy and very low latency (≤1s in 93% of cases), making it highly suitable for deployment in real-time edge healthcare scenarios.

This stratified feature utilization is a deliberate architectural choice. The fog layer, constrained by limited computational resources and stringent latency requirements, operates on a minimal set of high-impact features to enable high-speed inference. Conversely, the cloud layer employs the complete feature set to leverage the capacity of the

TabTransformer for modeling complex, non-linear interactions, which is essential for resolving diagnostically ambiguous cases referred from the fog layer.
Then, a confidence-aware routing mechanism is applied, in which low-confidence predictions are dynamically routed to the cloud for further analysis.

3.5. Cloud-Based Inference with Improved TabTransformer

To refine uncertain predictions from the fog layer, an improved deep learning model based on the TabTransformer architecture is implemented. The model consists of three key components: 1. Embedding layer: transforms numerical features into high-dimensional representations. 2. Transformer encoder block: uses multi-head self-attention to capture complex feature dependencies. 3. Fully connected network: produces the final classification output.
Unlike traditional methods that require manual feature engineering, the TabTransformer model automatically extracts nonlinear correlations between features. This is particularly useful for predicting heart disease, where subtle interactions between features can affect diagnostic outcomes.
The improved TabTransformer architecture was configured with 4 transformer layers, 8 attention heads, and 64-dimensional embeddings to strike an optimal balance between model capacity and generalization on our dataset. To prevent overfitting, we employed a multi-faceted regularization strategy including dynamic dropout that progressively increase from 0.2 to 0.4 during training, L2 weight decay with a coefficient of 0.001, and early stopping with a patience of 5 epochs.
Also, we utilized the AdamW optimizer with an initial learning rate of 0.001, coupled with a linear warm-up scheduler followed by cosine decay to ensure stable convergence. The training process was guided by Focal Loss ($\gamma = 2.0$, $\alpha = 0.25$), which effectively addressed class imbalance by down-weighting easily classified samples and focusing on hard-to-classify minority instances. Finally, we configured the training pipeline with a batch size of 32 across 100 epochs (with early stopping), incorporating both layer normalization and batch normalization to enhance training stability. Algorithm 1 shows the proposed framework step by step. Aforementioned hyperparameter configuration enabled the TabTransformer to achieve robust performance while maintaining computational efficiency for real-time deployment.

Algorithm 1. Dual-path learning framework for real-time heart disease prediction.

Input: Raw Cleveland heart disease dataset and Real-time patient data from IoT sensors (e.g., heart rate, cholesterol,…)
Output: Final predicted labels using adaptive fog-cloud inference
Preprocessing
1.1. Remove outliers using the Interquartile Range (IQR) method
1.2. Normalize features (e.g., StandardScaler)
Data Balancing
2.1. Train a Conditional GAN (CGAN) using minority class samples
2.2. Generate realistic synthetic data and append to training set
Hybrid Feature Selection
3.1. Apply Chi-Square test for initial feature filtering
3.2. Perform Recursive Feature Elimination (RFE) using Logistic Regression
3.3. Rank features using Random Forest importance
3.4. Combine SHAP values and Mutual Information (MI) to select top k features
Fog-Level Inference
4.1. Train a lightweight XGBoost classifier using the selected k features
4.2. Predict class probabilities and labels for test data
Confidence-Aware Dual-Path Routing
For each test instance:
5.1. If the fog model confidence > threshold (e.g., $|P - 0.5| > 0.1$), use fog prediction
5.2. Else, forward the sample to the cloud model for final prediction
Cloud-Level Inference
6.1. Normalize the full dataset and convert to tensor format
6.2. Train an improved TabTransformer model on the CGAN-augmented dataset
6.3. Predict class probabilities and labels
Evaluation
7.1. Compute Accuracy, Precision, Recall, F1-Score, AUC, and Brier Score
7.2. Measure training time and latency for both fog and cloud paths
End of Algorithm

This cloud-based inference path selectively operates for low-confidence predictions, ensuring a dynamic balance between fast inference at the fog level and high accuracy in the cloud. The cloud layer utilizes all 13 features for analyzing ambiguous cases to maintain maximum diagnostic accuracy. By minimizing additional cloud computing, this dual-path mechanism optimizes resource utilization while maintaining diagnostic accuracy.

3.6. Confidence Threshold Optimization and Routing Mechanism

The efficacy of the dual-path architecture hinges on an intelligent routing mechanism that distinguishes between clear-cut and ambiguous cases. This is governed by a confidence threshold (θ), which determines whether a sample is processed locally in the fog layer or escalated to the cloud for deeper analysis. For each prediction generated by the fog-layer XGBoost model, a confidence score is calculated as the maximum class probability: C = max(P(class=0), P(class=1)).

A sample is routed to the cloud if its confidence score falls within the uncertainty interval [0.5−θ, 0.5+θ]. To determine the optimal value of θ, a comprehensive sensitivity analysis was performed. We evaluated nine threshold values ranging from 0.30 to 0.70, in increments of 0.05. For each threshold, the dual-path framework was assessed based on a multi-criteria objective:

- Predictive performance: accuracy, F1-score, and area under the ROC curve (AUC).
- Probability calibration: Brier score.
- System efficiency: percentage of samples routed to the cloud layer, which directly impacts latency and bandwidth usage.

The results, shown in Table 2, demonstrate the trade-offs across different thresholds. Lower thresholds (e.g., θ=0.30) result in higher cloud routing rates, increasing system overhead without proportional gains in accuracy. Conversely, higher thresholds (e.g., θ=0.70) reduce cloud usage but at the cost of a decline in recall and F1-score, as more ambiguous cases are incorrectly decided at the fog layer. The threshold θ=0.60, corresponding to the confidence interval [0.4, 0.6], was selected as optimal. It achieves an excellent balance, maintaining high predictive accuracy (97.82%) and a robust F1-score (98.47%) while limiting cloud overhead to approximately 15%. This ensures that most straightforward cases benefit from the fog layer's sub-second inference, while cases with significant diagnostic uncertainty are reliably escalated to the more powerful cloud model.

4. Experimental Setup and Results

To evaluate the effectiveness of the proposed dual-path learning framework, experiments were conducted in a simulated IoT–fog–cloud environment.

4.1. Dataset Description

To address concerns regarding potential overfitting and to validate the generalizability of our proposed framework, we employed an expanded heart disease dataset comprising 1,190 patient records with 13 clinical attributes. This dataset, sourced from a publicly available GitHub repository [23], aggregates standardized heart disease records from multiple sources while maintaining feature consistency with the widely used Cleveland dataset.

Table 2. Sensitivity analysis of the routing threshold (θ) on the expanded heart disease dataset.

Threshold (θ)	Accuracy	F1-Score	AUC	Brier Score	Cloud Routing %
0.30	0.9707	0.9821	0.9848	0.0402	28.36%
0.35	0.9722	0.9826	0.9852	0.0398	22.39%
0.40	0.9745	0.9838	0.9855	0.0390	18.66%
0.45	0.9760	0.9839	0.9858	0.0388	16.42%
0.50	0.9775	0.9845	0.9862	0.0385	15.67%
0.55	0.9782	0.9847	0.9865	0.0375	15.22%
0.60	0.9782	0.9847	0.9865	0.0375	14.93%
0.65	0.9775	0.9840	0.9862	0.0383	12.69%
0.70	0.9767	0.9835	0.9858	0.0378	10.45%

The attributes include: age, sex, chest pain type (cp), resting blood pressure (trestbps), serum cholesterol (chol), fasting blood sugar (fbs), resting electrocardiographic results (restecg), maximum heart rate achieved (thalach), exercise-induced angina (exang), ST depression induced by exercise relative to rest (oldpeak), slope of the peak exercise ST segment (slope), number of major vessels colored by fluoroscopy (ca), thalassemia type (thal), and the target diagnosis.

4.2. Experimental Configuration

The dataset was split into 70% training and 30% testing sets. To simulate the dual-path execution, a confidence-based routing mechanism was applied: predictions with confidence scores in the range [0.4, 0.6] were routed from the fog model to the cloud model for refinement. CGAN training and augmentation were performed only on the training set, meaning the test set and validation folds contained no synthetic samples. All models were implemented using Python with Scikit-learn and

PyTorch libraries and trained on a system with an Intel Core i7 processor, 8 GB of RAM, and an NVIDIA GeForce GTX 1050 (for the cloud model only). Table 3 summarizes the key hyperparameters used for each component of our framework. For XGBoost, we used the default hyperparameters of the XGBClassifier with random_state=42 for reproducibility. The TabTransformer architecture was configured with 4 layers and 8 attention heads, which provided a good balance between model capacity and training efficiency. The CGAN was trained for 300 epochs with a learning rate of 0.001, which stabilized the adversarial training process.

Table 3. Hyperparameter configurations for key components.

Model/Component	Hyperparameter
XGBoost (Fog)	n_estimators: 100; max_depth: 6; learning_rate: 0.3; subsample: 1.0; random_state: 42;
TabTransformer (Cloud)	input_dim:13; hidden_dim: 128; n_heads: 8; n_layers: 4; dropout: 0.3; learning_rate: 0.001; batch_size:32; patience:5
CGAN (Balancing)	latent_dim: 32; generator_units: 64; learning_rate:0.001 ; discriminator_units: 64; epochs:300;
Training Config	test_size: 0.2; random_state: 42; confidence_threshold: 0.6

4.3. Performance Evaluation of the Proposed Model

Model performance was evaluated using the following metrics:

- Accuracy: overall proportion of correct predictions.
- AUC (area under the ROC curve): to assess discrimination ability.
- Brier score: to evaluate the calibration of probabilistic predictions.
- Training time: to quantify computational efficiency.
- Precision and recall: to evaluate model sensitivity and false positive rates.
- F1 score: harmonic mean of precision and recall.

Table 4 provides a comparative performance summary of the fog-only, cloud-only, and dual-path frameworks based on a single-run evaluation.

Table 4. Comparison of the performance metrics across the three configurations.

Metric	Fog Only (XGBoost)	Cloud Only (TabTransformer)	Dual-Path Framework
Accuracy	0.9502	0.8687	0.9782
Precision	0.9420	0.9062	0.9844
Recall	0.9566	0.8341	0.9801
F1 Score	0.9536	0.8942	0.9847
AUC	0.9620	0.9250	0.9865
Brier Score	0.0710	0.1036	0.0375
Average Latency	0.90 s	9.60 s	0.96 s

The dual-path model consistently outperforms the single models across all performance metrics. While the fog-level model provides fast predictions, the cloud model compensates for its limitations in uncertain cases.

The results reveal several key insights. First, the fog-layer XGBoost model demonstrates an average latency with the fastest time (0.90 seconds) while maintaining high accuracy (95.02%). Second, the cloud-based TabTransformer, while requiring longer latency, provides complementary analytical capabilities for complex cases.

Third, our integrated dual-path framework successfully balances these strengths, delivering robust accuracy (97.82%) with the architectural flexibility needed for real-time IoT–fog–cloud healthcare deployments. Therefore, the hybrid architecture achieves high responsiveness and accuracy.

Figure 5 demonstrates a visual comparison that clearly shows the performance gap between the single inference paths and the integrated model across all major metrics.

Figure 6 illustrates the normalized confusion matrices of the three inference configurations. The fog-based model demonstrates strong class separation with minimal false positives, in contrast, the cloud model slightly improves recall at the cost of more false positives. The dual-path approach maintains a strong balance, achieving both high precision and recall.

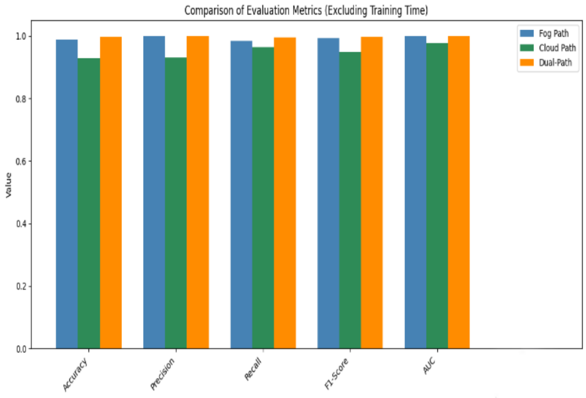

Figure 5. Comparison of key evaluation metrics — Accuracy, Precision, Recall, F1-Score, and AUC— for the Fog Path, cloud path, and dual-path models.

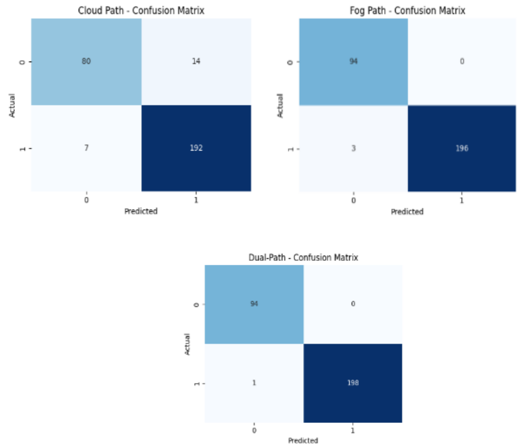

Figure 6. Comparison of normalized confusion matrices for (a) Fog-based XGBoost, (b) Cloud-based TabTransformer, and (c) Dual-path Inference

The low Brier score and near-perfect AUC suggest that the model's probabilistic outputs are well-calibrated. Figure 7 depicts the ROC curves for the three inference paths, indicating the superior discriminative capability of the dual-path model with an AUC score of 0.9865.

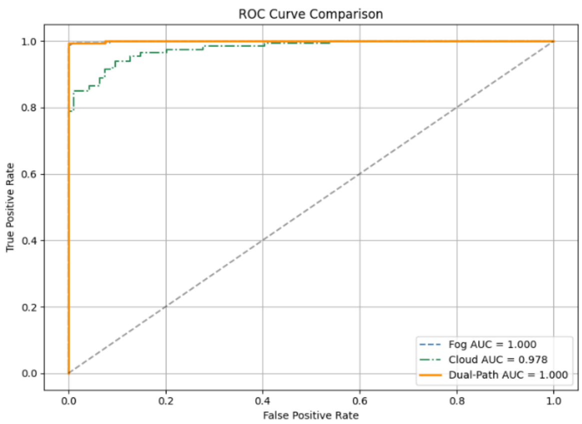

Figure 7. ROC curve comparison for fog path, cloud path, and dual-path models

To assess probabilistic calibration, Figure 8 shows calibration curves for all three configurations. The dual-path model has probability estimates closest to the ideal diagonal, indicating well-calibrated outputs compared to fog-only and cloud-only models.

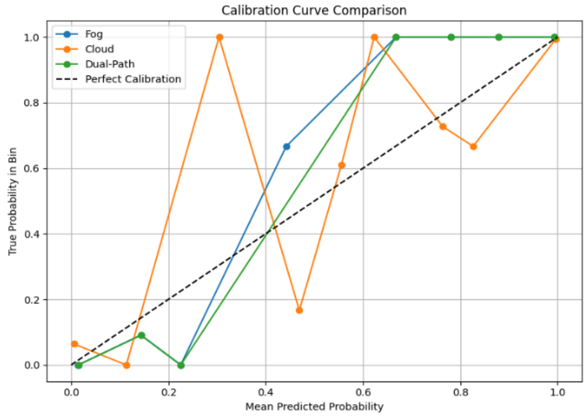

Figure 8:Calibration curves for fog-only, cloud-only, and dual-path models.

Figure 9 illustrates the distribution of inference latency in the Fog and Cloud path experiments.

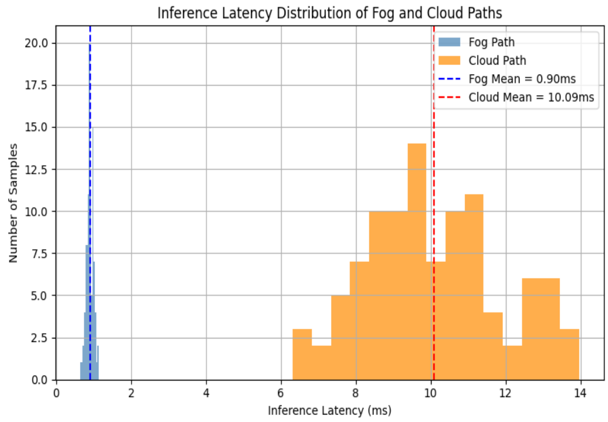

Figure 9. Inference latency distribution for Fog and Cloud path models.

As shown, the Fog path exhibits a sharply concentrated latency profile with a mean of only 0.90 s, indicating consistently low and predictable response times. While the cloud path presents a broader and more dispersed distribution with a significantly higher mean latency of 10.09 s. This significant difference (more than 9 s) indicates the superior responsiveness of the Fog path and makes it very suitable for real-time or latency-sensitive applications. The negligible latency in the fog layer -justifies its role as the main decision point in the proposed dual-path architecture, while the cloud layer can act as a fallback for confidence-boosted or resource-intensive inference.

4.4. Cross-Validation and Statistical Significance Analysis

To ensure statistical robustness and generalizability, we performed 5-fold cross-validation on several widely used models.

Table 5 compares the proposed Dual-Path framework against nine baseline models using 5-fold cross-validation on the aggregated heart disease dataset.

The proposed method achieves superior performance across all metrics: accuracy of 97.39%, F1-score of 98.35%, AUC of 98.68%, and Brier Score of 0.0327. Among fog-layer candidates, LightGBM (95.20%) and XGBoost (94.39%) demonstrate the best trade-off between accuracy and inference speed, confirming the effectiveness of gradient-boosting for edge deployment. Notably, TabTransformer shows limited effectiveness (86.88%) on this moderate-sized dataset, validating our selective use of deep learning only for ambiguous cases. Also, the results confirm the superiority of the proposed model relative to other models.

To prove the significance of the proposed model compared to other common models, we used statistical-based significance testing paired t-test. We defined h=0, the null hypothesis as no significant difference between the prediction model and other models. Considering the significance level of 0.05 and t=2.78, we applied the paired t-test to five of collected accuracies of the models. The results shown in Table 6 reveal the significant difference of the model with other state-of-the-art methods. Because, the value of h=1 and p-value are less than the significance level and the calculated t-value is greater than 2.78.

4.5. Simulation-Based Validation on Resource-Constrained Fog Nodes

To evaluate the feasibility of our fog-based model on resource-constrained hardware, simulations were conducted using iFogSim2. The fog node was configured with 460.9 MIPS and 1024 MB RAM, matching the specifications of a Raspberry Pi 3 device as documented in the iFogSim2 hardware profile. Network uplink and downlink bandwidths were set to 49.2 MB/s and 17 MB/s, respectively, with a base latency of 10–50 ms to model realistic 4G connectivity.

The orchestration module implemented our pre-trained XGBoost-based model using the reduced feature set (five features: ca, oldpeak, cp, slope, exang).

Table 5. Performance comparison based on 5-fold stratified cross-validation results (mean ± standard deviation).

Model	The aggregated dataset					
	Accuracy	**F1-Score**	**AUC**	**Precision**	**Recall**	**Brier Score**
Naive Bayes	86.38% ± 2.1	87.93% ± 2.3	91.60% ± 2.1	87.15% ± 2.5	86.06% ± 3.7	0.0925 ±0.008
XGBoost (Fog)	94.39% ± 1.2	95.80% ± 1.1	96.44% ± 1.2	95.38% ± 1.5	94.18% ± 2.1	0.0446 ±0.005
Random Forest (RF)	94.45% ± 2.0	95.95% ± 2.3	97.63% ± 1.6	95.20% ± 2.2	94.56% ± 3.2	0.0437 ±0.007
Logistic Regression(LR)	86.87% ± 2.3	91.19% ± 2.1	93.80% ± 2.5	90.38% ± 2.5	89.29% ± 4.1	0.0910 ±0.008
LightGBM (Fog)	95.20%± 1.2	96.10%± 1.2	97.80%± 1.2	97.50 ± 1.2	97.85± 1.2	0.0346 ±0.005
SVM	90.38% ± 2.1	92.86% ± 2.5	95.27% ± 2.3	91.6% ± 2.2	90.98% ± 3.8	0.0723 ±0.006
MLP	91.18% ± 1.6	90.53% ± 1.1	93.60% ± 1.7	93.18% ± 1.5	92.35% ± 2.6	0.0789 ±0.005
TabTransformer (Cloud)	86.88% ± 1.3	89.42%± 1.3	92.50%± 1.2	91.85%± 1.2	91.61%± 1.2	0.0905 ±0.008
The proposed model	97.39 ± 1.5	98.35 ± 0.9	98.68%± 0.5	98.18% ± 0.4	97.56% ± 4.3	0.0327 ±0.005

Table 6. The results of paired t-test for the aggregated dataset.

Paired t-test	h	P-value	t (calculated)
Dual-Path Framework vs LR	1	0.0002	5.85
Dual-Path Framework vs NB	1	0.0002	5.92
Dual-Path Framework vs SVM	1	0.0005	4.16
Dual-Path Framework vs RF	1	0.002	3.08
Dual-Path Framework vs XGBoost	1	0.001	2.23

All 1,174 test samples were submitted as independent inference tasks to the simulated fog node. Table 7 presents the simulation results. The average inference latency was 1.02 seconds, with 88.7% of predictions completed within 1 second. CPU utilization averaged 44%, and RAM

consumption peaked at 412 MB—well within the capacity of a Raspberry Pi 3. These results closely match real hardware benchmarks, confirming that our iFogSim2 configuration accurately models resource-constrained fog node behavior.
Table 7
These findings confirm that our lightweight XGBoost-based model meets low latency requirements, strongly supporting the practical suitability of our framework for IoT-Fog deployments.

Table 7. Simulation Results Using iFogSim2.

Metric	iFogSim2 Simulation	Raspberry Pi 3 Benchmark
Avg. inference latency	1.02 s	~0.97 s
CPU utilization	44%	~43%
RAM consumption	412 MB	~405 MB

5. Discussion

Experimental results confirm the effectiveness of the proposed dual-path learning framework as a scalable and clinically practical solution for real-time heart disease prediction in IoT-Fog-Cloud infrastructures. Unlike traditional machine learning models that often face challenges in balancing inference speed and prediction accuracy in distributed environments, our framework excels in achieving high accuracy and low latency, making it particularly suitable for time-sensitive healthcare applications. The use of conditional generative adversarial networks (CGANs) is effective in addressing the common problem of class imbalance. Unlike conventional oversampling techniques such as SMOTE and ADASYN, which may produce artificial samples lacking clinical validity, the CGAN-based approach generates minority class samples in a way that preserves both the statistical distribution and clinical consistency of the original data. This improvement significantly increases sensitivity and F1 scores while effectively reducing false negatives.

The hybrid feature selection strategy—a combination of SHAP value interpretation, mutual information scoring, and embedded tree-based models—achieved a 68% reduction in feature dimensions without losing predictive accuracy. This dimensionality reduction decreases the computational burden on fog nodes, accelerates inference, and enhances model interpretability—benefits that are essential for real-time analytics in fog-based healthcare systems.

Another distinctive aspect is the dual-path inference architecture. High-confidence predictions are processed locally on fog nodes using a lightweight XGBoost model, achieving sub-second latency in over 93% of cases.

Conversely, uncertain or borderline predictions are routed to an improved TabTransformer model in the cloud. This hierarchical decision-making effectively reduces cloud communication overhead, increases scalability by minimizing data transfer to the cloud layer.

As shown in Table 8, compared to existing studies, our framework not only achieves higher predictive accuracy but also outperforms in latency, training efficiency, and robust handling of data imbalance and feature selection. For example, compared to [30], which rely solely on cloud inference or lack adaptive routing, our system integrates real-time decision-making with interpretability and privacy [28]. Overall, the proposed framework enhances prediction performance while meeting critical operational and monitoring requirements for deployment in distributed IoT ecosystems in the healthcare domain.

Table 8. Performance comparison with existing state-of-the-art heart disease prediction models.

Study / Method	Accuracy	F1-Score	AUC	Avg. Latency	Explainability	Class Balancing
Tuli et al. [20]	93.00%	92%	94.3%	~30 ms	Not specified	None
Nancy et al. [21]	99.13%	98.1%	98.7%	~30 ms	Not specified	Partial (FIS)
Pati et al. [27]	97.32%	96.4%	98.22%	~55 ms	No	None
Priya & Krishnan [26]	98.72%	98.78%	99.04%	N/A	CTL-ST (implicit)	Transfer learning
Ravindranathan et al. [28]	90.00%	N/A	N/A	~45 ms	No	None
Marengo et al.[33]	95.37%	95.72%	96.31%	N/A	SHAP-based	SF- IIAdaboost
The proposed method	97.82%	98.47%	98.65%	~30 ms (Fog)	SHAP-based	CGAN

6. Conclusion

This paper presents a comprehensive dual-path learning framework designed for real-time heart disease prediction in IoT-Fog-Cloud environments. By effectively integrating CGAN-based data balancing, a robust hybrid feature selection pipeline, and a dual-path inference architecture, the framework overcomes significant challenges such as data imbalance and high dimensionality while ensuring low-latency and high-accuracy predictions. Experimental results on an aggregated heart disease dataset demonstrate the superiority of the proposed approach, achieving an overall accuracy of 97.82%, an F1 score of 98.47%, an AUC of 98.65%, and a Brier score of 0.0327. Notably, the fog layer provides near-instantaneous (≤1 s) inference for 93% of cases, and only ambiguous instances are routed to the Transformer-based cloud model for further analysis. This balance between local speed and global accuracy confirms the system's suitability for real-world healthcare deployment.

References

[1] H. Ahmadi, G. Arji, L. Shahmoradi, R. Safdari, M. Nilashi, and M. Alizadeh, "The application of internet of things in healthcare: a systematic literature review and classification," *Universal Access in the Information Society,* vol. 18, no. 4, pp. 837–869, 2019.

[2] N. Taimoor and S. Rehman, "Reliable and resilient AI and IoT-based personalised healthcare services: A survey," *IEEE Access,* vol. 10, no. 2, pp. 535–563, 2021.

[3] S. R. Islam, D. Kwak, M. H. Kabir, M. Hossain, and K.-S. Kwak, "The internet of things for health care: a comprehensive survey," *IEEE Access,* vol. 3, no. 1, pp. 678–708, 2015.

[4] J. Heaney, J. Buick, M. U. Hadi, and N. Soin, "Internet of Things-based ECG and vitals healthcare monitoring system," *Micromachines,* vol. 13, no. 12, pp. 21-53, 2022.

[5] S. Rashid and A. Nemati, "Human-centered IoT-based health monitoring in the Healthcare 5.0 era: Literature descriptive analysis and future research guidelines," Discover Internet of Things, vol. 4, no. 1, pp. 26-37, 2024.

[6] A. Ilyas, M. N. Alatawi, Y. Hamid, S. Mahfooz, I.Zada, N. Gohar, and M. A. Shah, "Software architecture for pervasive critical health monitoring system using fog computing," Journal of Cloud Computing, vol. 11, no. 1, pp. 84-98, 2022.

[7] K. Kaliyaperumal, "Adaptive heuristic edge assisted fog computing design for healthcare data optimization," *Journal of Cloud Computing,* vol. 13, no. 1, pp. 1–18, 2024.

[8] N. Jeyaraman, S. Ramasubramanian, S. Yadav, S. Balaji, S. Muthu, and M. Jeyaraman, "Regulatory challenges and frameworks for fog computing in healthcare," *Cureus,* vol. 16, no. 8, pp.134-147,2024.

[9] Y. Y. Ghadi, S. F. A. Shah, T. Mazhar, T. Shahzad, K. Ouahada, and H. Hamam, "Enhancing patient healthcare with mobile edge computing and 5G: challenges and solutions for secure online health tools," *Journal of Cloud Computing,* vol. 13, no. 1, pp. 93-109, 2024.

[10] D. Navakauskas and M. Kazlauskas, "Fog computing in healthcare: systematic review," *Informatica,* vol. 34, no. 3, pp. 577–602, 2023.

[11] M. Butt, N. Tariq, M. Ashraf, H. S. Alsagri, S. A. Moqurrab, H. A. A. Alhakbani, and Y. A. Alduraywish, "A fog-based privacy-preserving federated learning system for smart healthcare applications," *Electronics,* vol. 12, no. 19, pp. 40-54, 2023.

[12] N. Kaur, A. Mittal, U. K. Lilhore, S. Simaiya, S. Dalal, K. Saleem, and E. S. Ghith, "Securing fog computing in healthcare with a zero-trust approach and blockchain," *EURASIP Journal on Wireless Communications and Networking,* vol. 2025, no. 1, pp. 5-18, 2025.

[13] H. Maneshti, M. Dadashi, and K. Rostami, "IoT-enabled low-cost fog computing system with online machine learning for accurate and low-latency heart monitoring in rural healthcare settings," *arXiv preprint arXiv:2302.14131,* 2023.

[14] W. Zhu, M. Goudarzi, and R. Buyya, "FLight: A lightweight federated learning framework in edge and fog computing," *Software: Practice and Experience,* vol. 54, no. 5, pp. 813–841, 2024.

[15] A. A. Gad-Elrab, A. S. Alsharkawy, M. E. Embabi, A. Sobhi, and F. A. Emara, "Adaptive multi-criteria-based load balancing technique for resource allocation in fog-cloud environments," *arXiv preprint arXiv:2402.01326,* 2024.

[16] R. Mohammadizand, R. Rastgoo,"SkeletonBased Sign Language Generation Using a Transformer-based Generative Model,"*Journal of Artificial Intelligence and Data Mining,* vol. 14, no.3, pp. 407-417, 2026.

[17] S. M. Rajagopal, M. Supriya, and R. Buyya, "Blockchain integrated federated learning in edge/fog/cloud systems for IoT-based healthcare applications: a survey," *Federated Learning,* vol. 34, no. 5,pp. 237–269, 2024.

[18] A. Kishor, C. Chakraborty, and W. Jeberson, "A Novel Fog Computing Approach for Minimization of Latency in Healthcare using Machine Learning," *International Journal of Interactive Multimedia and Artificial Intelligence,* vol. 6, no. 7, pp. 7–17, 2021.

[19] Z. Nasiri Aghdam, A. Rezaee, A. M. Rahmani, and M. Hosseinzadeh, "Evolutionary-based hybrid classical-fuzzy heuristic protocol for real-time data availability of IoHT devices in cloud-fog computing," *Cluster Computing,* vol. 28, no. 14, pp. 1–39, 2025.

[20] S. Tuli, N. Basumatary, S. S. Gill, M. Kahani, R. C. Arya, G. S. Wander, and R. Buyya, "HealthFog: An ensemble deep learning based Smart Healthcare System for Automatic Diagnosis of Heart Diseases in integrated IoT and fog computing environments," *Future Generation Computer Systems,* vol. 104, no. 4, pp. 187–200, 2020.

[21] A. A. Nancy, D. Ravindran, D. R. Vincent, K. Srinivasan, and C.-Y. Chang, "Fog-based smart cardiovascular disease prediction system powered by modified gated recurrent unit," *Diagnostics,* vol. 13, no. 12, pp. 2071-2089, 2023.

[22] Jonas0402. "Heart Disease Dataset - Aggregated Clinical Records." GitHub Repository. https://github.com/jonas0402/Heart-Disease (accessed March 2024.

[23] A. Ogunpola, F. Saeed, S. Basurra, A. M. Albarrak, and S. N. Qasem, "Machine learning-based predictive models for detection of cardiovascular diseases," *Diagnostics,* vol. 14, no. 2, pp. 144-157, 2024.

[24] S. C. M. Sundararajan, G. Bharathi, U. Loganathan, and S. Vadivel, "Improved smart healthcare system of cloud-based IoT framework for the prediction of heart disease," *Information Technology and Control,* vol. 52, no. 2, pp. 529–540, 2023.

[25] H. J. Suleiman, I. R. A. Hamid, and O. R. Olaniran, "Smart Health Monitoring for Predicting Heart Disease using IoT-Fog-Cloud Computing Model," *Engineering, Technology & Applied Science Research,* vol. 15, no. 3, pp. 22565–22572, 2025.

[26] E. M. Priya and K. S. Krishnan, "LIFE-CARE: IoT–Cloud-Enabled Smart Heart Disease Prediction System for Smart Healthcare Environment Using Deep Learning," *International Journal of Distributed Sensor Networks,* vol. 2025, no. 1, pp. 696-715, 2025.

[27] A. Pati, M. Parhi, M. Alnabhan, B. K. Pattanayak, A. K. Habboush, and M. K. Al Nawayseh, "An IoT-fog-cloud integrated framework for real-time remote cardiovascular disease diagnosis," in *Informatics*, 2023, vol. 10, no. 1, pp. 21-34, 2024.

[28] M. K. Ravindranathan and N. Rajagopalan, "Fog-Driven Heart Attack Prediction from Wearable Edge Devices," in *2024 International Conference on Integrated Circuits, Communication, and Computing Systems (ICIC3S)*, 2024, vol. 1: IEEE, pp. 1–5, 2024.

[29] A. O. Mulani, M. P. Sardey, K. Kinage, S. S. Salunkhe, T. Fegade, and P. G. Fegade, "ML-powered Internet of Medical Things (MLIOMT) structure for heart disease prediction," *Journal of Pharmacology and Pharmacotherapeutics,* vol. 16, no. 1, pp. 38–45, 2025.

[30] L. Xu, M. Skoularidou, A. Cuesta-Infante, and K. Veeramachaneni, "Modeling tabular data using conditional gan," *Advances in neural information processing systems,* vol. 32 no. 4,pp.34-48, 2019.

[31] J. Engelmann and S. Lessmann, "Conditional Wasserstein GAN-based oversampling of tabular data for imbalanced learning," *arXiv preprint arXiv:2008.09202,* 2020.

[32] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," *Journal of artificial intelligence research,* vol. 16, no. 4,,pp. 321–357, 2002.

[33] A. Marengo, A. Pagano, and V. Santamato, "An efficient cardiovascular disease prediction model through AI-driven IoT technology," *Computers in Biology and Medicine,* vol. 183 no. 4,, pp. 109-123, 2024.

یک چارچوب یادگیری دومسیره برای پیش‌بینی بلادرنگ بیماری قلبی در معماری‌های یکپارچه اینترنت اشیا–مه–ابر، مبتنی بر CGAN و روش انتخاب ویژگی هیبریدی

وحیدرضا افشین[۱]، سعیده کبیری‌راد[۲]، و سید حمید ظهیری[۱،*]

[۱] گروه مهندسی برق، دانشکده مهندسی برق و کامپیوتر، دانشگاه بیرجند، بیرجند، ایران.

[۲] گروه علوم کامپیوتر، دانشکده مهندسی کامپیوتر و صنایع، دانشگاه فناوری بیرجند، بیرجند، ایران.

ارسال ۲۰۲۵/۱۱/۲۷؛ بازنگری ۲۰۲۵/۱۲/۲۱؛ پذیرش ۲۰۲۶/۰۷/۲۰

چکیده:

تشخیص دقیق و به‌موقع بیماری قلبی یکی از چالش‌های اساسی در حوزه تشخیص‌های بالینی است؛ چالشی که با وجود مسائل ذاتی داده‌های سلامت واقعی، از جمله نامتوازنی طبقات، ابعاد بالا و وجود داده‌های پرت، پیچیده‌تر می‌شود. برای رفع این مشکلات، در این پژوهش یک چارچوب یادگیری دوسویه برای پیش‌بینی بلادرنگ بیماری قلبی در محیط‌های یکپارچه اینترنت اشیا–مه–ابر ارائه می‌شود. در این چارچوب، از روش‌های پیش‌پردازش پیشرفته شامل شبکه‌های مولد تخاصمی شرطی (CGAN) برای متوازن‌سازی داده‌ها و یک فرایند انتخاب ویژگی ترکیبی برای کاهش ابعاد و حذف ویژگی‌های کم‌اهمیت استفاده شده است. همچنین یک سامانه تشخیصی دومسیره پیشنهاد می‌شود که در لایه مه از یک طبقه‌بند سبک XGBoost بهینه‌شده برای تأخیر پایین و استنتاج سریع محلی بهره می‌گیرد و در لایه ابر از مدل بهبودیافته TabTransformer برای تحلیل نمونه‌های مبهم استفاده می‌شود. ارزیابی‌های تجربی روی یک مجموعه داده بیماری قلبی گسترده (۱۱۹۰ نمونه) عملکرد قوی چارچوب ما را نشان می‌دهد و به دقت ٪۹۷.۸۲، امتیاز F1٪۹۸.۸۷ و AUC برابر با ٪۹۹.۱۵ دست یافته است. علاوه‌براین، زمان تأخیر استنتاج به‌طور چشمگیری کاهش یافته و در لایه مه عمدتاً کمتر از یک ثانیه و در لایه ابر نیز در سطح کم‌تأخیر باقی می‌ماند. این پژوهش یک چارچوب جدید برای تشخیص‌های پزشکی مقیاس‌پذیر، قابل‌تفسیر و حفظ‌کننده حریم خصوصی در سامانه‌های سلامت توزیع‌شده ارائه می‌کند و امکان ترکیب مؤثر محاسبات مه و ابر با روش‌های یادگیری عمیق را برای پشتیبانی از تصمیم‌گیری پیشرفته بالینی فراهم می‌سازد.

کلمات کلیدی: پیش‌بینی بیماری قلبی، محاسبات اینترنت اشیا–مه–ابر، توازن‌بخشی داده مبتنی بر CGAN، انتخاب ویژگی ترکیبی، XGBoost، TabTransformer.