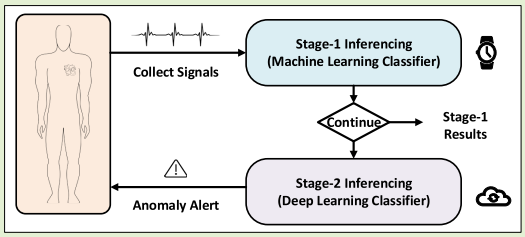
A Two-Stage ECG Classifier for Decentralized Inferencing Across Edge-Cloud Continuum

Li Xiaolin, *Graduate Student Member, IEEE*, Barry Cardiff, *Senior Member, IEEE*, and Deepu John, *Senior Member, IEEE*

***Abstract*—In this article, we propose a multistage electrocardiogram (ECG) classifier for distributed machine learning (ML) inferencing across the edge-cloud continuum for wearable systems. Traditionally, biomedical data acquired from wearable systems are processed in one step, using a single-stage classifier deployed either on a cloud or on an edge device. Though there are merits, both approaches have limitations that relate to computational complexity, network connectivity, and so on. In this work, we propose a multistage, cascaded, ensemble classifier that aims to address these limitations by decentralizing the processing steps, while achieving good overall performance. We employed low-complexity, explainable boosting machines (EBMs) and convolutional neural networks (CNNs) to develop a multistage distributed ECG classifier, which achieves a high-sensitivity binary classification on the edge device and a more comprehensive multiclass classification on the cloud. In standalone performance evaluation using the MIT-BIH Arrhythmia database, the Stage-1 EBM classifier and Stage-2 CNN classifier achieved a maximum accuracy, sensitivity of 96.71%, 96.76%, and 99.49%, 98.19%, respectively. Furthermore, the distributed multistage classifier achieved a maximum cumulative binary classification accuracy, sensitivity of 99.64%, 99.01%, and multiclass classification accuracy, sensitivity of 99.56%, 98.79% when DtC equals 40%. Furthermore, we evaluated the use of the EBM classifier threshold as a control parameter to dynamically vary the system performance and network traffic based on real-time conditions. We verified the feasibility of the model, calculated the energy consumption, and estimated the latency. When streaming only 40% of the data to the cloud, it will result in 60% latency saving. With the proposed technique, energy consumption is reduced by approximately three times.**



***Index Terms*— Arrhythmia, decentralized inferencing, distributed computing, edge-cloud network, electrocardiogram (ECG) classification, Internet of Health Things (IoHT).**

I. Introduction

CARDIAC arrhythmias are a leading cause of sudden deaths and have a higher incidence rate than stroke, cancer, and so on [1]. Examining a subject's electrocardiogram (ECG) is the most common way clinicians use to diagnose arrhythmia. Since many arrhythmias occur only intermittently, it is challenging for clinicians to manually review long periods of ECG for diagnosis as it is a laborious and error-prone task. Wearable Internet of Things (IoT) devices make it possible to continuously acquire physiological signals like the ECG for enhancing treatment outcomes for chronic conditions, including cardiovascular diseases (CVDs). With the advent of wearable, low-cost data collection equipment, ECG can be acquired continuously for a long duration [2], [3]. Several existing works are reported in the literature for automated ECG classification using artificial intelligence (AI) methods. These include traditional machine learning (ML) classifiers and deep neural networks (DNNs) that can achieve good performance results using extracted features and/or raw samples.

Traditionally, data acquired by a wearable device is first transmitted to an intermediate gateway and then sent to a cloud server for analysis. The automated data processing algorithms are deployed as cloud-native applications for continuous and long-term analysis of ECG data. Many DNN algorithms used for ECG classification are inherently complex, with many floating-point operations (FLOPs), weight parameters that necessitate large storage space and need significant computational resources that may result in large power consumption [4], [5], [6], [7], [8], [9], [10]. However, since computing in the cloud is highly scalable, the models deployed are mostly exempt from the constraints of complexity and resources

Manuscript received 10 April 2024; revised 25 May 2024; accepted 26 May 2024. Date of publication 4 June 2024; date of current version 16 July 2024. This work was supported in part by China Scholarship Council, in part by the Microelectronic Circuits Center Ireland, and in part by the Irish Research Council. The associate editor coordinating the review of this article and approving it for publication was Dr. Theerawit Wilaiprasitporn. *(Corresponding author: Li Xiaolin.)*

The authors are with the School of Electrical and Electronics Engineering, University College Dublin, Dublin, D04 V1W8 Ireland (e-mail: xiaolin.li@ucdconnect.ie; barry.cardiff@ucd.ie; deepu.john@ucd.ie).

Digital Object Identifier 10.1109/JSEN.2024.3406780

and therefore can attain high performance. Many arrhythmias occur intermittently and usually are time sensitive [11]. Therefore, wearable devices must continuously acquire and transmit ECG to the cloud for processing in a cloud-based processing approach. Wearable sensors often have modest batteries, and wireless transmission depletes the battery quickly and limits the overall utility of the system [12]. In addition, the entire process from signal collection and transmission to processing in the cloud can result in high communication latency (or complete disruption) depending on the network and channel conditions. In such cases, the system cannot offer a guaranteed response time for critical events, which limits the utility of the system.

Due to the limitations of fully cloud-based model deployments for wearable applications, several works are aimed at deploying the models fully on the wearable sensor, that is, natively on the edge device. This addresses concerns about: 1) power consumption due to the continuous wireless transmission; 2) concerns of response times due to poor network connectivity or outages; and 3) data privacy. However, this introduces fresh concerns regarding: 1) limited computational and memory resources [13] and 2) high power consumption due to the model complexity [14]. Several model compression and optimization techniques have been proposed to address the model complexity issue [15], [16], [17], [18]. However, such optimizations invariably limit the model performance, due to quantization or other losses involved. Some works also report the development of lightweight models for edge devices [19], [20], which reduces power consumption, lowers latency, and enables immediate alerts of abnormal signals without syncing data to the cloud. However, constrained by limited resources, only simpler models can be implemented on edge devices, limiting overall accuracy. To prolong the battery life, Keskes et al. [21] evaluated ten resampling techniques for signal quality assessment to decrease the amount of data to be transferred. Satija et al. [22] introduced an automated, low-complexity technique for signal quality assessment, which prolongs battery life by excluding unusable beats during transmission.

In this work, we validated the feasibility of a two-stage, distributed ML inferencing model that addresses the challenges in edge-native and cloud-native deployments of biomedical signal classification tasks, and subsequently in future work, we aim to conduct validation for models involving multiple stages. We propose to distribute the processing over multiple nodes (edge, gateway, cloud) in the system such that each individual processing stage can be deployed on an individual node, generating results that are useful by themselves and can be incrementally aggregated with the results from previous nodes for achieving more comprehensive processing and outcomes. We propose a generalized, multistage, cascaded, ensemble classification model for this. This generalized model can be applied to a multitude of biomedical use cases based on the data type, volume, classification criteria, number of nodes/stages available, and so on.

The proposed model is applied to ECG arrhythmia classification to develop a multistage, edge-cloud distributed ML inferencing model. For the Stage-1 classifier to be deployed on the edge device, we propose using a relatively simple binary classifier such as an explainable boosting machine (EBM) [16] tuned for high detection sensitivity. The sensor wireless transmission is to be gated using the classifier output such that only abnormal data is transmitted to the cloud, thus enabling power reduction [19]. Also, Stage-1 model response times will be unaffected by network conditions or latency. For the Stage-2 classifier which is to be deployed on the cloud, we use a multiclass convolutional neural networks (CNNs) classifier for a more comprehensive analysis of data [4]. We prioritize the sensitivity of the first stage as a crucial performance metric, aiming for high sensitivity to minimize false negatives at the first stage to ensure almost all abnormal beats are sent to the next stage. Any incorrectly classified (false-positives) data from Stage-1, will be correctly reclassified in Stage-2 ensuring high system accuracy. In standalone operation, our Stage-1 and Stage-2 classifiers achieved an accuracy, and sensitivity of 96.71%, 96.76%, and 99.49%, 98.19%, respectively, when tested using the MIT-BIH Arrhythmia database [23]. The cumulative performance of both stages for binary classification and multiclass classification is evaluated and found to be 99.64%, 99.01%, and 99.56%, 98.79%, respectively, using the same database. Further, we evaluated the use of EBM-threshold as a control parameter to dynamically alter the system performance and network traffic based on real-time conditions.

The main contributions of this work are listed below.

1) A generalized, multistage classification approach is proposed for distributed and decentralized inferencing of biomedical signals over multiple nodes in the edge-cloud continuum. The proposed approach addresses the limitations of edge-native and cloud-native approaches.
2) A distributed, multistage classifier for ECG arrhythmia detection, with a cumulative sensitivity and accuracy of 99.64% and 99.01%, is presented. The proposed approach has been verified to reduce the amount of data to be transmitted to 30% enabling power savings while not compromising on response times.
3) For the distributed classifier, we evaluated the EBM-threshold as a control parameter to vary the system performance dynamically.
4) If there is a network failure, a traditional cloud-deployed classifier may not work and could lead to a complete disruption of monitoring. For the proposed method, the classifier at the edge can provide initial assessments and quick responses, which enables the user to seek proactive help, and keeps the functionality of the system.

Section II illustrates some background work related to ECG classification. We introduce the distributed multistage cascaded ensemble processing architecture for decentralized inferencing in Section III. Section IV shows the simplified multistage edge-cloud classifier for ECG classification and Section V introduces the dataset and the methods applied for data preprocessing; the experimental results obtained and the comparison with state-of-the-art works are in Section VI. Section VII concludes our work on the proposed distributed decentralized multistage computing architecture.

II. Related Work

Several reported studies used traditional ML algorithms and deep-learning (DL) techniques for automated arrhythmia detection and classification from ECG. Paper [9], [20] and [24] used ML classifiers in their work. Bhattacharyya et al. [24] introduced an ensemble model combining random forest (RF) and support vector machines (SVMs), for classifying ECG into five classes (N, S, V, F, and Q). Although this can achieve 98.21% overall accuracy, the sensitivity of some abnormal categories is very low, S: 74.20%, F: 73.21%, and Q: 0%, which limits the feasibility of this model to be used in actual clinical settings. Pal et al. [25] proposed a two-step classifier aimed at achieving higher accuracy and lower latency in detecting 15 classes of cardiac rhythms using 13 features. Initially, the signals are categorized into three groups, followed by further classification using an RF-based classifier into the 15 specified classes. Paper [6], [8], [26], [27], [28] and [29] employed CNN models in their respective studies. In [28], notable success was achieved through a new approach for hyperparameter optimization, leading to commendable performance. Evren et al. [28] employed a metaheuristic algorithm known as the memory-enhanced artificial hummingbird algorithm and presented a novel fitness function that accounts for both the accuracy rate and the total number of parameters in each candidate network. Experimental evaluations were performed on raw ECG samples from the MIT-BIH arrhythmia database. The proposed method demonstrates a classification accuracy of 98.87%.

Meanwhile, [30] introduced an automated diagnostic system that combines CNN and long short-term memory (LSTM) for diagnosing cardiac arrhythmias. The system utilizes variable-length ECG segments sourced from the MIT-BIT arrhythmia database. The proposed model exhibits strong classification performance, achieving an accuracy of 98.10%, and sensitivity of 97.50% through a tenfold cross-validation strategy.

Based on the long duration 10-s ECG segment, YECG seg et al. [26] classified cardiac arrhythmia into 17 classes, reaching an overall classification accuracy of 91.33% and sensitivity of 83.91%. Although the article claims that the model may be utilized in edge devices and cloud computing, there are still layout issues in edge devices because the 16-layer deep network has a significant number of parameters based on the network's volume. Also, this work is not based on the widely accepted Association for the Advancement of Medical Instrumentation (AAMI) standards, and because the model is not sensitive enough, positive patients may not be identified in time, which could have catastrophic repercussions.

Hannun et al. developed a CNN model to classify ECG into 12 classes. They achieved an F1 score of 83.7%, which exceeded that of average cardiologists [6]. This model uses 33 convolutional layers, and therefore requires a large number of computations and memory, making it difficult to use in edge devices.

Acharya et al. proposed a CNN model to classify ECG single beat into five classes, achieving an accuracy of 93.47% and sensitivity of 96.01% using synthetic de-skewed data [27]. The model contains nine convolutional layers resulting in a large number of model parameters, needing to be deployed on the cloud for execution.

Niu et al. [29] presented a method using symbolic representations and multiperspective CNN (MCNNs). The model can achieve an overall accuracy of 96.4% classifying ECG into three classes, N, S, and V. However, the sensitivity of abnormal classes is quite low, S: 76.5% and V: 86.7%. In addition, this system must perform a significant preprocessing for feature extraction. Thus, due to MCNN and feature extraction steps, the overall model complexity remains high and therefore is better suited for cloud-native implementations.

However, none of these models proved suitable for distributed processing. These existing models often face obstacles during the classification process, making them less than ideal for distributed environments. As the demand for distributed computing continues to grow, there is a pressing need for such distributed processing.

III. Distributed Multistage Computing

Deploying the ML inferencing process exclusively on any single node (e.g., edge, fog, and cloud) introduces challenges regarding ML model performance, communication bandwidth, sensor power consumption, computational/storage resources, and so on. While cloud-native model deployments are optimal for high model performance with limited computational/storage resource restrictions, they increase sensor power consumption and response times and consume significant channel bandwidth. On the contrary, edge-native model deployments result in lower model performance and face sensor computational/storage resource limitations while resolving latency and sensor power consumption issues. To address these issues and get the best of both worlds, we propose a multistage model that distributes inferencing over multiple nodes on the Edge-Fog-Cloud continuum. To formulate such a model, we propose using a cascaded ensemble approach, originally proposed in [31] and [32]. This is illustrated in Fig. 1. Here, each processing stage will generate partial results and will be deployed on any single node, such that the results are useful in themselves and can be incrementally aggregated with previous results for improved cumulative performance. Such a design ensures the robustness and resilience of the overall system. In the event of cellular network disruptions, data from wearable devices can still be transmitted to smartphones via BLE. In the three-stage model, models in edge and fog stages work collaboratively to produce more accurate results without waiting for network restoration compared to the two-stage model. Furthermore, there is a reduced risk of data loss due to the absence of a cellular network, as smartphones have greater storage capacity compared to wearable devices. This system can also function as an anomaly detector (Stage-1 on-edge device) even if the smartphone (or BLE) is not available, enabling the user to seek medical help themselves. Therefore, there is a need to generalize to a multistage model to ensure the robustness of the system.

A. Architecture

Model size and limited performance are major constraints in the edge-native deployment of biomedical signal processing

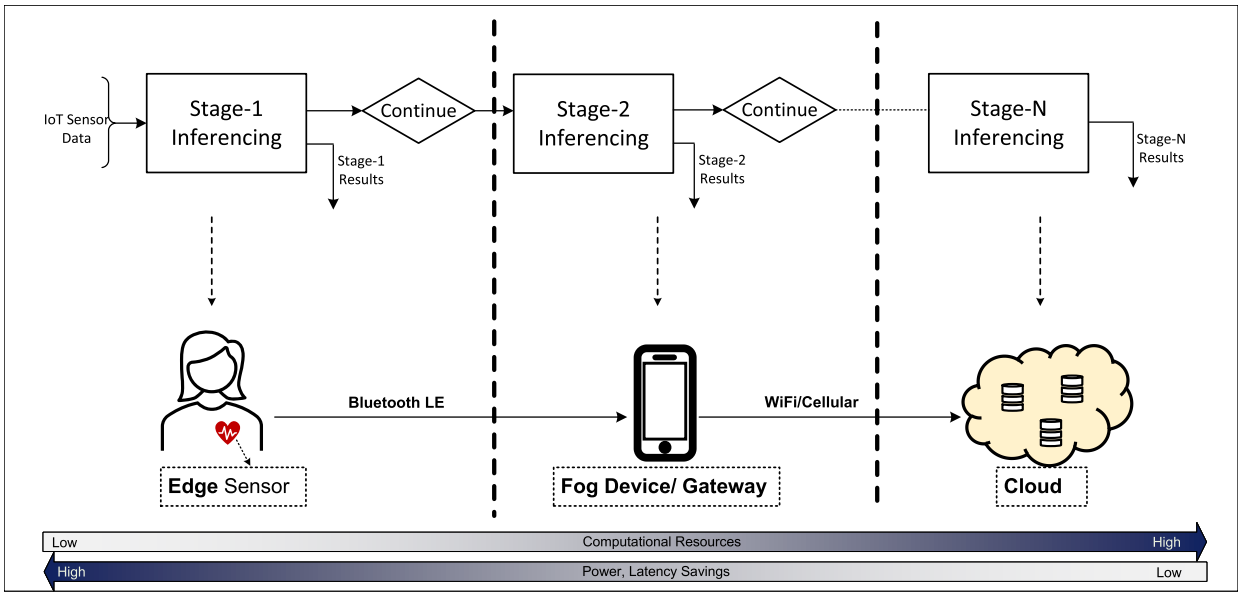


Fig. 1. Distributed, multistage, and cascaded ensemble processing architecture for decentralized inferencing.

tasks. On the other hand, communication bandwidth and latency, power consumption in the edge sensor, and system robustness are major challenges in cloud-native model deployments. We propose a cascaded, ensemble-based processing approach to address these issues. Here, the overall data processing is split into multiple stages, where each stage is capable of processing the data independently, obtaining intermediate or final results, and is deployed as a standalone node in the edge-cloud continuum as illustrated in Fig. 1.

Each stage is a relatively simple model capable of making a high-quality decision on: 1) all input samples for a more simplified classification problem and 2) on a subset of the input samples for the original classification problem. In both cases, inferencing is deemed complete at each stage for a subset of the input samples, and further processing is terminated. The multiple inferencing stages are cascaded in such a way that the remaining smaller subset of input samples from each stage is fed into a more capable succeeding stage for further (re)classification and thus results in an ensemble processing. This approach reduces the communication bandwidth and transmission overheads between different stages (or nodes). In addition, the processing latency and response times can be improved as inferencing can be terminated at earlier nodes (edge) for most input samples. By minimizing the amount of data sent to succeeding stages, edge devices can reduce energy consumption and achieve a longer lifespan [19]. The proposed approach also increases the robustness and fault tolerance of the system as incorrect results from earlier stages will be re-looked by the more capable succeeding stages.

The cumulative results (in terms of model performance) at any stage have to be better than that of individual prior stages. In addition, the cumulative result at the final stage should achieve high processing accuracy, sensitivity, and so on, that is, the overall decentralized model should achieve high performance like that of a large model deployed entirely in the cloud without any complexity constraints. The initial stages should achieve high detection sensitivity (in the case of a simplified classifier) while maintaining low computational complexity. The high detection sensitivity ensures that no critical events are missed, and the low complexity ensures the initial stages can be deployed on resource-limited edge sensors. Any incorrectly classified samples from the initial stages will be passed on to a succeeding stage (or node) with more computational resources to re-process and reclassify the input more accurately.

The decentralized multistage inferencing architecture can be applied to various complex scenarios with multiple positive classes and can be leveraged to classify broad categories in the initial few stages, followed by sending the results to a cloud-based server for further refinements. Thus, the selection of the number of stages used in constructing a system can be determined by the quantity of data available, the performance of the models, and classification requirements. Elaboration on the advantages of the proposed system is provided below.

B. Advantages

1) Low System Latency: The time consumed for the acquisition, transmission, and inferencing of the data is referred to as system latency. For time-sensitive processes, like ECG arrhythmia detection, a high system latency may result in serious and irreversible consequences [33], [34]. In the proposed multistage distributed processing, signals can be classified as abnormal (or normal) in the early stages implemented on edge devices. This reduces latency in transmission and cloud-based processing and thus reduces the overall system latency.

2) High Robustness: The proposed architecture offers high robustness due to its distributed nature. Different stages in the model can detect anomalous signals with high sensitivity and increasingly high accuracy. As abnormal signals are likely to be picked up at various stages, the possibility of missing such signals is low, resulting in high reliability. Since there is no requirement for a large cloud-based model to react to abnormal signals, a simple Stage-1 algorithm on the edge device can generate alerts for abnormal data, enhancing robustness in

TABLE I
COMPLEXITY COMPUTATION OF STAGE-1 STANDALONE TREE-BASED CLASSIFIERS

	DT	RF	XGBoost	LGBM	EBM
Multiplication	Average depth	Average depth	Max depth*Boosted rounds	Max depth*Boosted rounds	N+K
Addition	None	None	Boosted rounds-1	Boosted rounds-1	N+K-1

TABLE II
ACCURACY AND SENSITIVITY OF K-FOLD ($K = 5$) CROSS-VALIDATION OF STAGE-1 STANDALONE TREE-BASED CLASSIFIERS

	Decision Tree	Random Forest	XGBoost	LightGBM	EBM
Mean accuracy	96.65	98.10	98.14	95.45	95.91
Standard accuracy	0.0016	0.0014	0.0004	0.0012	0.0022
Mean sensitivity	90.19	90.21	93.41	96.05	95.99
Standard sensitivity	0.0006	0.0003	0.0006	0.0007	0.0010

emergencies. The proposed framework can continue to operate effectively even in the temporary absence of network connectivity without compromising the core functionality.

3) Low Cost: The proposed approach reduces data storage and communication bandwidth requirements for IoT wearable sensing. A significant portion of biomedical data acquired using wearables is normal signals. Therefore, wireless transmission and storage of such data do not contribute to improving the metrics in biosensing and increase the cost of transmission and storage. Preprocessing data on the IoT edge using a low-complexity Stage-1 classifier reduces the quantity of data to be transmitted and stored in the cloud.

4) Low Power Consumption on the Edge: The power consumption on edge devices comprises two main components: signal acquisition and data transmission. Efforts aimed at decreasing power consumption on edge devices are predominantly focused on mitigating the energy-intensive process of data transmission, which constitutes a substantial proportion of power usage. Consequently, implementing the proposed approach can minimize the volume of data transmissions, which significantly alleviates power consumption burdens on edge devices.

IV. TWO-STAGE DISTRIBUTED ECG CLASSIFIER

While the distributed multistage classification architecture is compatible with multiple stages, we chose to implement a multistage classifier (as shown in Fig. 2) for experimental illustration due to its simplicity. Here, we propose to use a low-complexity, high-sensitivity binary classifier for Stage-1 inferencing. In Stage-1, ECG is classified as either Normal or Abnormal class. Only the Abnormal signals are then sent to the Stage-2 classifier, which is implemented as a CNN-based high-accuracy multiclass classifier, to be deployed on the cloud. This multistage classification significantly reduces the amount of data transmitted to the cloud, sensor power, and inference latency and achieves a good balance of performance and complexity.

A. Stage-1 Inferencing

In Stage-1 inferencing, it is desirable to use a low-complexity classifier as it is to be implemented in a resource-limited edge device. Toward this, we propose using a simple, binary classifier that classifies ECG into Normal or Abnormal classes. In addition, high sensitivity is desirable to reduce false negatives. Compared to DL approaches, a traditional feature-based ML classification method with lower complexity is deemed more appropriate for Stage-1. In literature, various ML methods have been employed for classifying abnormalities.

[35] employed an SVM classifier for arrhythmia classification, resulting in a classification accuracy of 94.45% and a sensitivity of 70.30%. The sensitivity performance is poor, and the accuracy is not very satisfactory either. Moreover, SVM is difficult to interpret, and features may need normalization and scaling. Logistic regression is characterized by its transparency and efficiency. However, its performance deteriorates significantly when handling a large number of variables, and it does not possess the capability to automatically capture interactions [36]. Naïve Bayes is distinguished by its speed, scalability, and ability to operate with limited data. Nevertheless, it lacks direct interpretability and assumes variable independence [37].

Using tree-based models is a commonplace approach for addressing classification tasks, predominantly owing to their explainability and remarkable predictive performance [38], [39], [40]. Model complexity is estimated by counting the number of arithmetic operations, such as multiplications and additions. Table I shows the approach used for estimating arithmetic operations for all tree-based algorithms. Each decision of each branch of these tree-based models is considered a multiplication. For decision tree (DT) and RF, we traverse the tree using the depth-first search (DFS) algorithm to calculate the average tree depth, which gives the number of multiplication operations. The complexity of XGBoost and light gradient boosting machine (LGBM) is estimated using the number of boosting rounds and maximum depth. The complexity of an EBM classifier can be estimated using the number of features (N) and interactions (K).

To evaluate the models, we conducted a thorough comparison of the receiver operating characteristic (ROC) curve presented in Fig. 3. It can be observed that the DT has the lowest area under the ROC Curve (AUC), while the AUC for the other models all exceeds 0.99.

Since the data for biomedical AI is very scarce, K-fold ($K = 5$) cross-validation is conducted, and the mean and standard deviation of accuracy and sensitivity across the folds are reported in Table II to provide a comprehensive understanding of the models' generalization capabilities. Based on the results in Table II, XGBoost achieves the highest mean accuracy and relatively high mean sensitivity, which is 98.14% and 93.41%. EBM achieves 95.91% and 95.99% for mean accuracy, and sensitivity separately. Those high mean values of the metrics and minimal standard deviation values reflect a remarkably high consistency in performance across different cross-validation folds. The result enhances our confidence in the reliability of the models and also indicates that our models are less influenced by specific data distributions, showcasing improved generalization capabilities.

To determine whether the model is appropriate for Stage-1 inferencing, Table III displays the performance of four

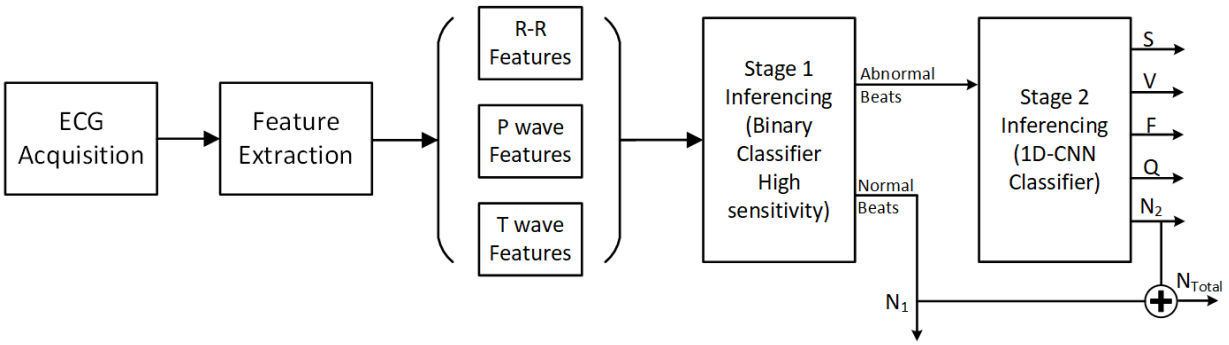


Fig. 2. Block diagram of the multistage distributed ECG classifier.

TABLE III
BINARY ECG CLASSIFICATION PERFORMANCE OF STAGE-1 STANDALONE TREE-BASED CLASSIFIERS

Method	Acc(%)	Sen(%)	Spe(%)	F1(%)	DtC(%)	Multi	Add
DT	97.07	91.48	98.24	91.53	17.29	18	None
RF	98.38	91.41	99.84	95.13	15.95	2549	None
XGBoost	98.3	93.73	99.25	95.01	16.84	501	166
LGBM	95.82	96.18	95.75	88.81	20.09	501	166
EBM	96.84	96.83	91.38	96.84	19.37	132	131

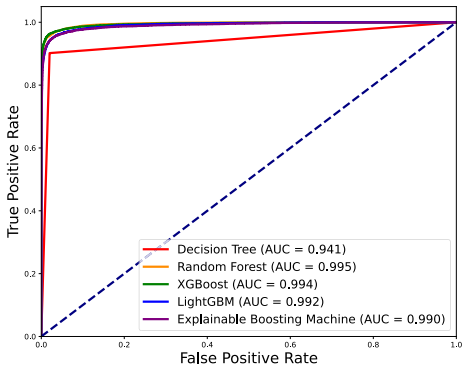


Fig. 3. ROC curve of Stage-1 standalone tree-based classifiers.

tree-based algorithms, calculated using equations in [15] and (2). Table III shows EBM has a significantly lower level of complexity than any other evaluated model, except for DT. On the other hand, EBM can reach significantly higher sensitivity than other models, which is why EBM is chosen as the Stage-1 inferencing model in the proposed architecture because the sensitivity of Stage-1 must be as high as possible. In addition, most of the existing classification techniques reported in the literature use black-box approaches that lack interpretability. Omitting interpretability in clinical decision support systems poses a threat to core ethical values in medicine and may have detrimental consequences for individual and public health [41]. Considering the importance of interpretability, we propose to employ EBM as the Stage-1 inferencing classifier.

EBM is a tree-based, cyclic gradient boosting generalized additive model (GAM) with automatic interaction detection [42]. Boosting is an ensemble learning method that combines a set of weak learners into strong learners to improve performance. EBMs are often as accurate as state-of-the-art

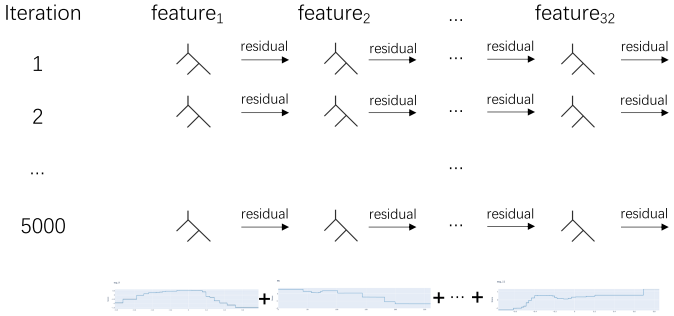


Fig. 4. Illustration of how EBM algorithm works.

black-box models while remaining completely interpretable, and the number of leaf nodes on the EBM tree used in boosting can be varied for performance tuning. Equation (1) shows how EBM learns from the combination of original features and their interactions

$$g(E[y]) = \Sigma f_i(x_i) + \Sigma f_{i,j}(x_i, x_j) \tag{1}$$

where g is the link function that adapts the GAM to classification, i, j varies from 0 to the number of features, (x_i, x_j) means the pairwise interactions between features.

Fig. 4 displays the process of EBMs generating trees and making predictions. A small tree is trained for each feature, and the residual is computed. This procedure is repeated for each feature, and a round-robin approach is used. After the first iteration, boosting rounds follow, with each feature generating a graph showing the prediction results of all trees. The final model is the sum of scores generated from a series of graphs.

B. Stage-2 Inferencing

DL has gained immense popularity in recent years due to its exceptional performance in a wide range of applications and the capability to handle large volumes of complex data [43], [44], [45], [46]. CNN is a powerful DL tool, which is increasingly being used in practical applications. The efficacy of the 1-D CNN model shown in Fig. 5 has been verified through validation on the MIT-BIH Arrhythmia dataset, exhibiting a high degree of accuracy and sensitivity. The hyperparameters used in this 1-D CNN model are as follows: Seed = 1, Learning Rate = 0.001, Batch Size = 128, and Epochs = 100. This ten-layer 1-D CNN consists of three

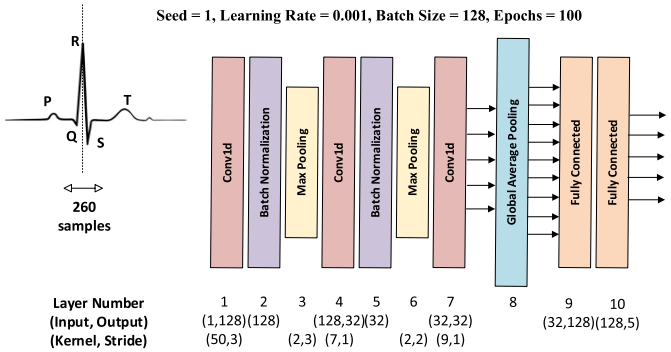


Fig. 5. Baseline CNN architecture for classification.

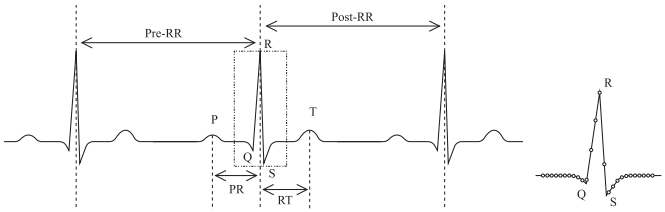


Fig. 6. (a) Features extracted from ECG Signals and (b) 26 uniform sampling points from QRS complex.

convolutional layers, two batch normalization layers, two max-pooling layers, one global average pooling layer, and two fully connected layers. The input of the network is a single ECG beat with 260 samples, and the outputs will be the different classes. For each convolutional layer, we use the rectified linear unit (ReLU) as the activation function. A dropout with a probability of 0.5 was applied in layers 4, 7, and 9 to reduce overfitting.

V. Dataset and Preprocessing

The ECG records from the MIT-BIH Arrhythmia database are used in this work [23]. This dataset is the most commonly used dataset for evaluating ECG classifiers [24], [26], [27], [47], [48], [49], [50], [51], [52]. Comparing methodologies that utilize the same dataset can ensure the validity of our results. It consists of ambulatory ECG obtained from 48 subjects. Noise in ECG signals can significantly impact classification tasks [21], [53]. It's crucial to ensure the accuracy and reliability of diagnosis of different types of arrhythmias under noisy ECG recording environments [54]. When the noise level is too high, sending data to a cloud server can result in the data being unusable for diagnostic applications and automated analysis [55]. In the MIT-BIH dataset [23], signals have been filtered through a bandpass to remove noise, with a passband ranging from 0.1 to 100 Hz, and then digitized at a rate of 360 Hz per signal, ensuring elimination of unwanted frequencies beyond this range. Thus, in this case, we don't do filtering in this work. Each record is ∼30 min long and sampled at 360 Hz.

We extracted time intervals between fiducial points (such as RR, PR, and RT Intervals), peak amplitudes (for P and T waves) and polarity of ECG wave segments, overall signal shape, and so on, for feature-based ECG Stage-1 classifier [16]. The overall morphology of ECG could be used for distinguishing ECG beats [56]. It was shown that ECG morphology could be represented using 260 samples centered around the R peak [27]. To reduce complexity, we employed a 10× downsampled variant (i.e., 26 samples) in this study as an additional feature. Finally, appending all individual features, a feature vector of size 32 was created as the Stage-1 classifier input. As seen in Fig. 6(a) and (b), this consists of 26 ECG samples indicating the compressed morphology, four time intervals, and two peak amplitudes/polarities. For the Stage-2 classifier, we used ECG segmented into separate beats (260 samples [27]) as input. The segment is centered on the R peak as shown in Fig. 5. Based on the AAMI standards, ECG is mapped into five classes, that is, N, SVEB, VEB, F, Q [57].

TABLE IV
Number of Heartbeats in Training, and Test Set on the Edge

	Normal	Abnormal	Total	Abnormal Rate
Training	63401	13188	76589	17.22%
Test	27167	5658	32825	17.24%
Total	90568	18846	109414	17.22%

TABLE V
Number of Heartbeats in Training, Validation, and Test Set in the Cloud

	N	SVEB	VEB	F	Q	Total
Training original	63417	1965	5067	548	5618	76615
Validation	13619	394	1101	129	1175	16418
Test	13557	422	1067	125	1247	16418

VI. Evaluation and Result

In the evaluation process for the first stage, the feature extraction is done by using "ecgpuwave" tool. The second stage of validation was completed using the PyTorch framework. As shown in Table IV, 70% of the total ECG beats were used for training and the remaining for testing the EBM classifier. The "Abnormal Rate" shown in the table indicates the proportion of abnormal beats in the overall dataset. The system is designed to send the abnormal beats to Stage-2 inferencing for further testing. The multistage distributed ECG classification system expects that DtC will include all the abnormal beats with assured high sensitivity.

The model training and testing are done using the PyTorch platform with a Tesla T4 GPU. Table V displays the number of ECG signals used in the process of evaluating the ten-layer 1-D CNN model. The training set consists of 70% of the whole dataset, 15% of the whole dataset is used for validation, and another 15% of the whole dataset for testing. The number of FLOPs consumed in each layer is shown in Table VI.

A. Feasibility

1) Edge: It is assumed that a microcontroller nRF5340 series low energy system-on-chip (SoC) device running at 64 MHz is employed on the edge. The cycles consumed

TABLE VI
FLOATING POINT OPERATIONS FOR EACH LAYER

Layer	Name	# FLOPs
1	Conv1d	71×50×128×2
2	Batch Normalization	71×128
3	Max Pooling	0
4	Conv1d	18×7×32×2
5	Batch Normalization	18×32
6	Max Pooling	0
7	Conv1d	1×9×32×2
8	Global Average Pooling	0
9	Fully Connected	32×128×2
10	Fully Connected	128×5×2
Total		936576

TABLE VII
PARAMETER SETTINGS AND BINARY CLASSIFICATION PERFORMANCE OF EBM ALGORITHM

Leaf nodes	Interactions	Acc(%)	Sen(%)	Spe(%)	F1(%)	DtC(%)
2	0	94.28	89.42	95.3	84.4	19.36
	100	96.71	96.76	96.70	91.05	19.47
	300	97.53	97.18	97.60	93.16	18.8
3	0	94.32	89.60	95.31	84.53	19.38
	100	96.75	96.69	96.76	91.15	19.41
	300	97.54	97.10	97.64	93.19	18.75
5	0	94.36	89.56	95.36	84.60	19.33
	100	96.76	96.69	96.78	91.18	19.39
	300	97.48	97.06	97.57	93.02	18.81

by the EBM classifier can be estimated using the number of features (N) and interactions (K), which equals $N + K$ multiplications and $N + K + 1$ additions, that is, 265 clock cycles are taken by the classifier. Divided by the max clock speed, the execution will be achieved within 2.47 ms.

B. Performance

Various figures of merit are used for measuring performance, such as accuracy, sensitivity, specificity, and F1 score [15]. DtC is given below

$$\text{DtC} = \frac{\text{Number of Abnormal}}{\text{Number of Total}} \times 100\% \tag{2}$$

where DtC represents the amount of data sent to the cloud (as a result of being identified as Abnormal) as a percentage of total data in the test set.

1) Stage-1 Standalone: To evaluate the standalone performance of the Stage-1 EBM classifier, we experimented with various parameter configurations of leaf nodes and interactions. Table VII displays the results for different configurations. It can be seen that more leaf nodes do not significantly improve performance, and hence we kept only two leaf nodes to reduce complexity. It is noted that a higher number of pairwise interactions can improve performance and decrease DtC at the expense of complexity.

Fig. 7 shows how detection sensitivity and complexity vary with the number of interactions. The complexity here indicates the total number of arithmetic operations. When increasing the feature interactions from 100 to 300, detection sensitivity

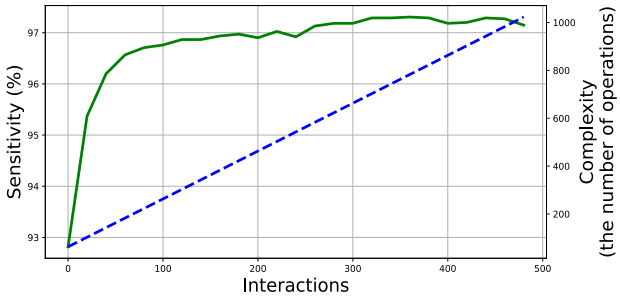


Fig. 7. Analysis of sensitivity, complexity with respect to interactions.

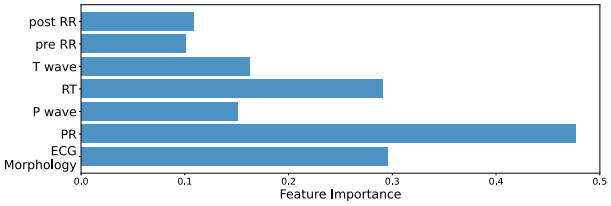


Fig. 8. Feature importance analysis for Stage-1 EBM classifier.

TABLE VIII
CLOUD PERFORMANCE FOR EACH CLASS AND THE OVERALL DATASET

	N	SVEB	VEB	F	Q	Total
Accuracy	99.49%	99.74%	99.79%	99.82%	99.97%	99.49%
Sensitivity	99.81%	91.73%	98.74%	81.82%	99.67%	98.19%
Specificity	97.91%	99.96%	99.86%	99.94%	99.99%	99.83%
F1 Score	99.69%	94.84%	98.31%	85.71%	99.79%	98.69%

increases a very small 0.82%, while the interaction increases to 400. Thus, 100 interactions may be a good trade-off between complexity and sensitivity. In our experiment, we applied an interaction number of 100 and a leaf node number of 2.

Compared to other traditional black-box approaches, the EBM-based glass-box model can not only achieve high performance with very low complexity but also add explainability to the classifier output. Fig. 8 shows the individual feature importance for the EBM binary ECG classifier. The processes of training and testing the model are not conducted independently for each subject. Therefore, the importance of each feature remains the same for all subjects. ECG-Morphology in the figure represents the average feature score of all the compressed morphology samples. EBM is an additive model where the contribution of each feature to the result is straightforward, which offers explainability and makes it easy for clinicians to interpret the potential cause of an ECG beat classified as abnormal.

2) Stage-2 Standalone: The Stage-2 model was trained, validated, and tested using the PyTorch platform in a standalone fashion. During the training phase, an augmented training set is employed, while approximately 15% of the total dataset is used for testing. The various parameters for evaluating model performance are reported for individual classes and the overall model. Table VIII displays the results for individual classes and the overall dataset.

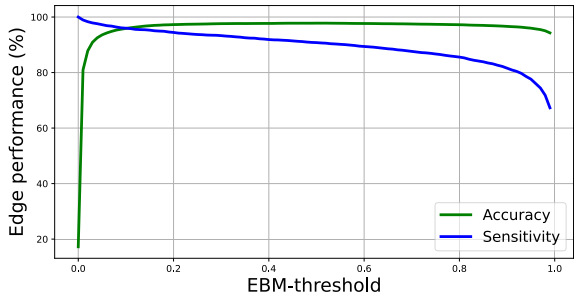


Fig. 9. Analysis of accuracy, sensitivity on the edge with respect to EBM-threshold.

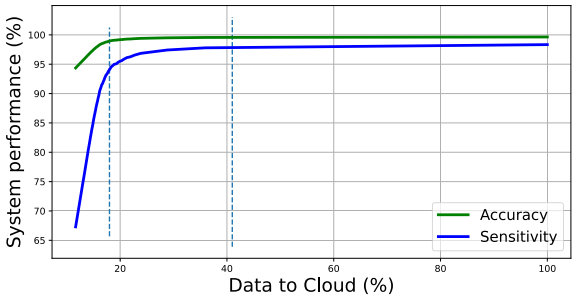


Fig. 10. Analysis of system accuracy, sensitivity with respect to DtC.

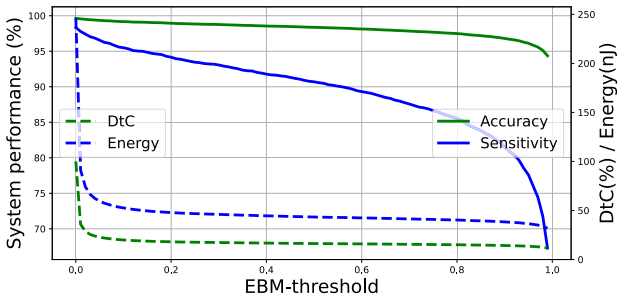


Fig. 11. Analysis of system accuracy, sensitivity, DtC, energy with respect to EBM-threshold.

3) Multistage Classifier: In multistage classification, an EBM classifier is employed in Stage-1 and a CNN classifier in Stage-2. It's important to note that while EBM is our current choice for Stage-1, our proposed distributed classifier is agnostic of the exact models used. We remain open to exploring and incorporating other models based on evolving needs and objectives. The ECG heartbeats deemed abnormal by Stage-1, the entire heartbeat consisting of 260 samples, will be sent to Stage-2 for reclassification by the Stage-2 classifier. If the logistic regression value exceeds the EBM threshold, the ECG beat is classified as abnormal, whereas a lower value is classified as normal. The EBM-threshold is used to regulate the amount of data retained at the endpoint and transmitted to the next stage, and this experiment is carried out 100 times with EBM-thresholds evenly distributed between 0 and 1. All signals consisting of 260 samples identified as abnormal in the first stage will be forwarded to the second stage model for classification.

In the experiments, we used the "ecgpuwave" package to extract those features, which is a package of the waveform database (WFDB) software package designed specifically for processing ECG signals [58]. Using this package, users can accurately detect and label heartbeat positions in ECG signals, extract crucial features related to heartbeat waveforms, and calculate time intervals between adjacent heartbeats.

The edge accuracy and sensitivity with regard to the EBM threshold are plotted in Fig. 9. Accuracy increases rapidly as the EBM threshold increases from 0 to 0.1, after which it stays at about 98%, while sensitivity decreases as the threshold increases. In Stage-1, it is important not to miss abnormal beats. Therefore we propose to prioritize optimizing for sensitivity by using a lower EBM-threshold at the expense of Stage-1 accuracy. Any beat incorrectly labeled abnormal will be reclassified by the Stage-2 CNN classifier with higher accuracy. For a standalone Stage-1 classification, an EBM-threshold between 0.1 and 0.2 is an ideal trade-off as it achieves a good balance between accuracy and sensitivity performance.

Fig. 10 plots multistage system accuracy and sensitivity with respect to DtC. The plot shows that by transmitting a mere 41% of data to the cloud (EBM-threshold set to 0.01), an accuracy and sensitivity of 98% can be achieved. This performance is comparable to a high-performance CNN classifier deployed fully in the cloud, as shown in Table VIII while saving nearly 60% of communication bandwidth and other resources. The proposed two-stage system allows a graceful degradation in performance and achieves 98% accuracy and 95% sensitivity when DtC is at a mere 18% (EBM-threshold set to 0.2). In practical scenarios, DtC selection should be based on real-time system requirements. The fact that the proposed multistage system achieves a comparable performance of a high-performance classifier while requiring much less data transfer and system resources verifies the utility of our approach.

Setting a lower EBM-threshold results in a higher DtC value, leading to higher sensor energy consumption. However, this yields the highest values for both system accuracy and sensitivity. The system sensitivity is significantly influenced by the EBM-threshold, while the system accuracy tends to vary more gradually. In practical scenarios, if the device can provide sufficient energy, a smaller EBM-threshold is preferred to achieve higher performance. However, if the sensor energy or communication bandwidth (DtC) is limited, a tradeoff between energy, DtC, and performance needs to be considered when determining the EBM-threshold.

Compared to the single-stage multiclass classifier in Table IX, our proposed approach not only outperforms it in terms of accuracy and sensitivity but also uses less sensor power, requires less overall computations, and achieves inferencing more quickly. The system performance of our model in both two-class and five-class classifications are displayed in Table IX. The five-class classification results vary depending on the DtC value. In [8], [20], [48], and [59] have reported high performance. However, [20] doesn't take the sensitivity into account for the evaluation and test the performance using proprietary data, [8] has high complexity in feature extraction, [48] classify SVEB and VEB, and [59] classify as ventricular and nonventricular beats, instead of normal and abnormal. Lai et al. [59] combines four databases rather than drawing conclusions based on a single database, and allocating only 10% of the entire dataset as the test set may lead

TABLE IX
PERFORMANCE COMPARISON OF THE PROPOSED MULTISTAGE CLASSIFIER AND OTHER MULTICLASS APPROACHES

Authors	Datasets	No. of classes	Features	Methods	Accuracy(%)	Sensitivity(%)
Venkatesan *et al.* [20]	MIT-BIH, Proprietary	2 classes	HRV	KNN	97.5	-
Raghu *et al.* [8]	MIT-BIH	2 classes	Hybrid features	DNN	98.3	98.6
Yan *et al.* [47]	MIT-BIH	2 classes	Wavelet and RR features	CNN	81.09	86.80
Tang *et al.* [48]	MIT-BIH	2 classes	Morphology, timing features	SVM	99.1	89.90
Lai *et al.* [59]	AHA, CUDB, MIT-BIH, VFDB	2 classes	ECG segment (3 s)	CNN	98.82	95.05
This work (Ours)	MIT-BIH	2 classes	Raw ECG data, extracted features	EBM, CNN	99.64	99.01
Bhattacharyya [24]	MIT-BIH	5 classes	Spectral, statistical, temporal features	SVM, RF	98.21	68.23
Acharya *et al.* [27]	MIT-BIH	5 classes	ECG segmentation, synthetic data	CNN	94.03	96.71
Oh *et al.* [30]	MIT-BIH	5 classes	Variable lengths	CNN-LSTM	98.1	97.5
Ozal [49]	MIT-BIH	5 classes	360 samples (1 s)	DBLSTM-WS3	99.39	-
Yang *et al.* [50]	MIT-BIH	5 classes	300 samples (0.83 s)	Linear SVM	97.43	84.42
Yang *et al.* [51]	MIT-BIH	5 classes	Multidimensional feature set	MSDPS, WRelifF-GA-SVM	99.74	99.42
Hannun *et al.* [6]	Proprietary	12 classes	Raw ECG data	DNN	-	75.22
Plawiak [52]	MIT-BIH	13 classes	3600 samples (10 s)	SVM	94.6	94.6
Ozal *et al.* [26]	MIT-BIH	17 classes	ECG fragment (10 s)	CNN	91.33	83.91
				EBM, CNN, DtC=20%	99.17	95.49
This work (Ours)	MIT-BIH	5 classes	Raw ECG data, extracted features	EBM, CNN, DtC=30%	99.49	98.42
				EBM, CNN, DtC=40%	99.56	98.79

to the possibility of overfitting. Yan et al. [47] exclude four records before the evaluation, so the reliability of the results is uncertain. Also, the performance presented is impractical for real-world applications, given its low accuracy and sensitivity. Hannun et al. [6] utilized a proprietary dataset and achieved a low level of sensitivity. The training and testing of [6], [26], [52], and [51] do not follow AAMI standards. In [26], [30], and [52] are primarily focused on rhythm-based analysis rather than beat-by-beat classification. In [49], [52], and [24] achieve an accuracy that is approximately 1.5% higher than or comparable to ours, while the sensitivity is lower by 2% to around 30% than us. While [27], [50] achieve relatively high performance, our results surpass theirs.

C. Latency

The latency of a standalone cloud-based CNN classifier is denoted as L_c, including the time for acquisition, transmission, and inferencing of the data in the cloud classifier, and the latency of a standalone edge-based EBM classifier is denoted as L_e, including the time for acquisition and inferencing of the data in the edge classifier. Upon the implementation of our proposed distributed classifier, the average latency of the proposed multistage classifier with different DtC is calculated by the following equation:

$$L = \text{DtC} \cdot L_c + L_e. \tag{3}$$

Due to the significant contribution of transmission latency, it is plausible to disregard L_e for analysis. According to the selected data points in Fig. 10, when the DtC is 41%, an average reduction of 59% in latency can be achieved. Similarly, when the DtC is 18%, an average reduction of 82% in latency can be attained. As the DtC decreases, a greater reduction in latency can be achieved, resulting in faster inferencing.

D. Energy Estimation

In this section, we estimate the energy consumption of the Stage-1 classifier, assuming it is implemented as a standard cell ASIC in 28 nm FD-SOI technology. The total energy consumption on the edge is estimated by assuming that the energy required for a 16-bit multiplication accumulation (MAC) operation is 0.39 pJ [60], and for a 16-bit adder is around 20 fJ [61]. The QRS detector consumes an estimated 19 nJ per beat [62]. EBM classifier consumes 51.5 pJ for each beat classification since it does $N + K$ multiplications and $N + K + 1$ additions, which equals $N + K$ MAC and 1 addition. A total of 0.63 mJ is used in the preliminary binary ECG classification on the edge for our test set's 32 825 beats.

For wirelessly transmitting data to the cloud, it is estimated that the energy required is approximately 143 nJ/bit when using a Bluetooth transceiver [63]. The quantity of data that needs to be transferred to the cloud is dramatically decreased after passing the EBM classifier, leaving the majority of the normal ECG beats at the edge. For all the MIT-BIH records, the total amount of data to be transmitted wirelessly is approximately 34.3 MB in the cloud-native system, which requires an estimated total of 41.15 J. With our proposed technique, only 11.6 MB of data has to be transmitted to the cloud for processing after stage-1 EBM classification. Consequently, an estimated 3× reduction in energy consumption is achieved, as only one-third of the data is required to be sent to the cloud. Fig. 11 illustrates more energy can be saved if the EBM-threshold is increased, as fewer data need to be transmitted to the cloud for further classification.

VII. CONCLUSION

This article presents a novel, multistage ECG classifier for distributed inferencing over multiple processing in the edge-cloud continuum. The proposed system is capable of obtaining an extremely high accuracy and sensitivity of 99.64% and 99.01% for binary classification. The experimental results indicate that when the classification stages are staggered across edge and cloud, the average inference latency can be reduced. When only 40% of the data is sent, the average inferencing latency is reduced by ∼60%. Furthermore, the latency can be reduced by as much as ∼80% when sending 20% of the data. By only transmitting ECG beats that are abnormal to the cloud, an estimated 3× reduction in energy consumption is achieved using the proposed distributed approach, consequently leading to extended battery life.

REFERENCES

[1] D. Mozaffarian et al., "Executive summary: Heart disease and stroke statistics—2015 update: A report from the American Heart Association," *Circulation*, vol. 131, no. 4, pp. 434–441, 2015.

[2] D. L. T. Wong et al., "An integrated wearable wireless vital signs biosensor for continuous inpatient monitoring," *IEEE Sensors J.*, vol. 20, no. 1, pp. 448–462, Jan. 2020.

[3] J. Li, A. Ashraf, B. Cardiff, R. C. Panicker, Y. Lian, and D. John, "Low power optimisations for IoT wearable sensors based on evaluation of nine QRS detection algorithms," *IEEE Open J. Circuits Syst.*, vol. 1, pp. 115–123, 2020.

[4] L. Xiaolin, B. Cardiff, and D. John, "A 1D convolutional neural network for heartbeat classification from single lead ECG," in *Proc. 27th IEEE Int. Conf. Electron., Circuits Syst. (ICECS)*, 2020, pp. 1–2.

[5] L. Xiaolin, F. Xiang, R. C. Panicker, B. Cardiff, and D. John, "Classification of ECG based on hybrid features using CNNs for wearable applications," in *Proc. IEEE 5th Int. Conf. Artif. Intell. Circuits Syst. (AICAS)*, Jun. 2023, pp. 1–4.

[6] A. Y. Hannun et al., "Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network," *Nature Med.*, vol. 25, no. 1, pp. 65–69, 2019.

[7] M. K. Gautam and V. K. Giri, "A neural network approach and wavelet analysis for ECG classification," in *Proc. IEEE Int. Conf. Eng. Technol. (ICETECH)*, Mar. 2016, pp. 1136–1141.

[8] R. Nanjundegowda, J. University, and V. Meshram, "Arrhythmia detection based on hybrid features of T-wave in electrocardiogram," *Int. J. Intell. Eng. Syst.*, vol. 11, no. 1, pp. 153–162, Feb. 2018.

[9] A. Vishwa, M. K. Lal, S. Dixit, and P. Vardwaj, "Clasification of arrhythmic ECG data using machine learning techniques," *Int. J. Interact. Multimedia Artif. Intell.*, vol. 1, no. 4, p. 67, 2011.

[10] M. Saeed et al., "Evaluation of level-crossing ADCs for event-driven ECG classification," *IEEE Trans. Biomed. Circuits Syst.*, vol. 15, no. 6, pp. 1129–1139, Dec. 2021.

[11] M. Carrington et al., "Monitoring and diagnosis of intermittent arrhythmias: Evidence-based guidance and role of novel monitoring strategies," *Eur. Heart J. Open*, vol. 2, no. 6, 2022, Art. no. oeac072.

[12] A. Natarajan, V. Krishnasamy, and M. Singh, "Occupancy detection and localization strategies for demand modulated appliance control in Internet of Things enabled home energy management system," *Renew. Sustain. Energy Rev.*, vol. 167, Oct. 2022, Art. no. 112731.

[13] J. Shao and J. Zhang, "Communication-computation trade-off in resource-constrained edge inference," *IEEE Commun. Mag.*, vol. 58, no. 12, pp. 20–26, Dec. 2020.

[14] L. Deng, G. Li, S. Han, L. Shi, and Y. Xie, "Model compression and hardware acceleration for neural networks: A comprehensive survey," *Proc. IEEE*, vol. 108, no. 4, pp. 485–532, Apr. 2020.

[15] L. Xiaolin, R. C. Panicker, B. Cardiff, and D. John, "Multistage pruning of CNN based ECG classifiers for edge devices," in *Proc. 43rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Nov. 2021, pp. 1965–1968.

[16] L. Xiaolin, W. Qingyuan, R. C. Panicker, B. Cardiff, and D. John, "Binary ECG classification using explainable boosting machines for IoT edge devices," in *Proc. 29th IEEE Int. Conf. Electron., Circuits Syst. (ICECS)*, Oct. 2022, pp. 1–4.

[17] S. Han, H. Mao, and W. J. Dally, "Deep compression: Compressing deep neural networks with pruning, trained quantization and Huffman coding," 2015, *arXiv:1510.00149*.

[18] Z. Liu, J. Li, Z. Shen, G. Huang, S. Yan, and C. Zhang, "Learning efficient convolutional networks through network slimming," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2736–2744.

[19] G. Sivapalan, K. K. Nundy, S. Dev, B. Cardiff, and D. John, "ANNet: A lightweight neural network for ECG anomaly detection in IoT edge sensors," *IEEE Trans. Biomed. Circuits Syst.*, vol. 16, no. 1, pp. 24–35, Feb. 2022.

[20] C. Venkatesan, P. Karthigaikumar, and R. Varatharajan, "A novel LMS algorithm for ECG signal preprocessing and KNN classifier based abnormality detection," *Multimedia Tools Appl.*, vol. 77, no. 8, pp. 10365–10374, Apr. 2018.

[21] N. Keskes, S. Fakhfakh, O. Kanoun, and N. Derbel, "Representativeness consideration in the selection of classification algorithms for the ECG signal quality assessment," *Biomed. Signal Process. Control*, vol. 76, Jul. 2022, Art. no. 103686.

[22] U. Satija, B. Ramkumar, and M. S. Manikandan, "Robust cardiac event change detection method for long-term healthcare monitoring applications," *Healthcare Technol. Lett.*, vol. 3, no. 2, pp. 116–123, Jun. 2016.

[23] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH arrhythmia database," *IEEE Eng. Med. Biol. Mag.*, vol. 20, no. 3, pp. 45–50, May/Jun. 2001.

[24] S. Bhattacharyya, S. Majumder, P. Debnath, and M. Chanda, "Arrhythmic heartbeat classification using ensemble of random forest and support vector machine algorithm," *IEEE Trans. Artif. Intell.*, vol. 2, no. 3, pp. 260–268, May 2021.

[25] D. Pal, S. Mukhopadhyay, and R. Gupta, "Two-stage classifier for resource constrained on-board cardiac arrhythmia detection," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–10, 2023.

[26] Ö. Yıldırım, P. Pławiak, R.-S. Tan, and U. R. Acharya, "Arrhythmia detection using deep convolutional neural network with long duration ECG signals," *Comput. Biol. Med.*, vol. 102, pp. 411–420, Nov. 2018.

[27] U. R. Acharya et al., "A deep convolutional neural network model to classify heartbeats," *Comput. Biol. Med.*, vol. 89, pp. 389–396, Oct. 2017.

[28] E. Kıymaç and Y. Kaya, "A novel automated CNN arrhythmia classifier with memory-enhanced artificial hummingbird algorithm," *Expert Syst. Appl.*, vol. 213, Mar. 2023, Art. no. 119162.

[29] J. Niu, Y. Tang, Z. Sun, and W. Zhang, "Inter-patient ECG classification with symbolic representations and multi-perspective convolutional neural networks," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 5, pp. 1321–1332, May 2020.

[30] S. L. Oh, E. Y. K. Ng, R. S. Tan, and U. R. Acharya, "Automated diagnosis of arrhythmia using combination of CNN and LSTM techniques with variable length heart beats," *Comput. Biol. Med.*, vol. 102, pp. 278–287, Nov. 2018.

[31] N. García-Pedrajas, D. Ortiz-Boyer, R. del Castillo-Gomariz, and C. Hervás-Martínez, "Cascade ensembles," in *Computational Intelligence and Bioinspired Systems*, J. Cabestany, A. Prieto, and F. Sandoval, Eds. Berlin, Germany: Springer, 2005, pp. 598–603.

[32] P. Zhang, T. D. Bui, and C. Y. Suen, "A novel cascade ensemble classifier system with a high recognition performance on handwritten digits," *Pattern Recognit.*, vol. 40, no. 12, pp. 3415–3429, Dec. 2007.

[33] S. B. Karch, J. Graff, S. Young, and C.-H. Ho, "Response times and outcomes for cardiac arrests in las Vegas casinos," *Amer. J. Emergency Med.*, vol. 16, no. 3, pp. 249–253, May 1998.

[34] M. P. Larsen, M. S. Eisenberg, R. O. Cummins, and A. P. Hallstrom, "Predicting survival from out-of-hospital cardiac arrest: A graphic model," *Ann. Emergency Med.*, vol. 22, no. 11, pp. 1652–1658, Nov. 1993.

[35] V. Mondéjar-Guerra, J. Novo, J. Rouco, M. G. Penedo, and M. Ortega, "Heartbeat classification fusing temporal and morphological information of ECGs via ensemble of classifiers," *Biomed. Signal Process. Control*, vol. 47, pp. 41–48, Jan. 2019.

[36] A. K. Feeny et al., "Artificial intelligence and machine learning in arrhythmias and cardiac electrophysiology," *Circulat., Arrhythmia Electrophysiol.*, vol. 13, no. 8, Aug. 2020, Art. no. e007952.

[37] J. Rahul, M. Sora, L. D. Sharma, and V. K. Bohat, "An improved cardiac arrhythmia classification using an RR interval-based approach," *Biocybern. Biomed. Eng.*, vol. 41, no. 2, pp. 656–666, Apr. 2021.

[38] M.-C. Kang, D.-Y. Yoo, and R. Gupta, "Machine learning-based prediction for compressive and flexural strengths of steel fiber-reinforced concrete," *Construct. Building Mater.*, vol. 266, Jan. 2021, Art. no. 121117.

[39] M. M. Ghiasi, S. Zendehboudi, and A. A. Mohsenipour, "Decision tree-based diagnosis of coronary artery disease: CART model," *Comput. Methods Programs Biomed.*, vol. 192, Aug. 2020, Art. no. 105400.

[40] S. S. Yadav and S. M. Jadhav, "Detection of common risk factors for diagnosis of cardiac arrhythmia using machine learning algorithm," *Expert Syst. Appl.*, vol. 163, Jan. 2021, Art. no. 113807.

[41] J. Amann, A. Blasimme, E. Vayena, D. Frey, and V. I. Madai, "Explainability for artificial intelligence in healthcare: A multidisciplinary perspective," *BMC Med. Informat. Decis. Making*, vol. 20, no. 1, pp. 1–9, Dec. 2020.

[42] Y. Lou, R. Caruana, J. Gehrke, and G. Hooker, "Accurate intelligible models with pairwise interactions," in *Proc. 19th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2013, pp. 623–631.

[43] G. Cheng, X. Xie, J. Han, L. Guo, and G.-S. Xia, "Remote sensing image scene classification meets deep learning: Challenges, methods, benchmarks, and opportunities," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 3735–3756, 2020.

[44] A. Gupta, Anjum, S. Gupta, and R. Katarya, "InstaCovNet-19: A deep learning classification model for the detection of COVID-19 patients using chest X-ray," *Appl. Soft Comput.*, vol. 99, Feb. 2021, Art. no. 106859.

[45] H. Jahangir, S. Lakshminarayana, C. Maple, and G. Epiphaniou, "A deep learning-based solution for securing the power grid against load altering threats by IoT-enabled devices," *IEEE Internet Things J.*, vol. 10, no. 2, pp. 10687–10697, Jun. 2023.

[46] K. Muhammad, S. Khan, J. D. Ser, and V. H. C. D. Albuquerque, "Deep learning for multigrade brain tumor classification in smart healthcare systems: A prospective survey," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 2, pp. 507–522, Feb. 2021.

[47] Z. Yan, J. Zhou, and W.-F. Wong, "Energy efficient ECG classification with spiking neural network," *Biomed. Signal Process. Control*, vol. 63, Jan. 2021, Art. no. 102170.

[48] X. Tang et al., "A near-sensor ECG delineation and arrhythmia classification system," *IEEE Sensors J.*, vol. 22, no. 14, pp. 14217–14227, Jul. 2022.

[49] Ö. Yıldırım, "A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification," *Comput. Biol. Med.*, vol. 96, pp. 189–202, May 2018.

[50] W. Yang, Y. Si, D. Wang, and B. Guo, "Automatic recognition of arrhythmia based on principal component analysis network and linear support vector machine," *Comput. Biol. Med.*, vol. 101, pp. 22–32, Oct. 2018.

[51] J. Yang and R. Yan, "A multidimensional feature extraction and selection method for ECG arrhythmias classification," *IEEE Sensors J.*, vol. 21, no. 13, pp. 14180–14190, Jul. 2021.

[52] P. Pławiak, "Novel methodology of cardiac health recognition based on ECG signals and evolutionary-neural system," *Expert Syst. Appl.*, vol. 92, pp. 334–349, Feb. 2018.

[53] S. Luo and P. Johnston, "A review of electrocardiogram filtering," *J. Electrocardiol.*, vol. 43, no. 6, pp. 486–496, Nov. 2010.

[54] U. Satija, B. Ramkumar, and M. S. Manikandan, "A new automated signal quality-aware ECG beat classification method for unsupervised ECG diagnosis environments," *IEEE Sensors J.*, vol. 19, no. 1, pp. 277–286, Jan. 2019.

[55] A. John, R. C. Panicker, B. Cardiff, Y. Lian, and D. John, "Binary classifiers for data integrity detection in wearable IoT edge devices," *IEEE Open J. Circuits Syst.*, vol. 1, pp. 88–99, 2020.

[56] P. de Chazel and R. B. Reilly, "A comparison of the ECG classification performance of different feature sets," in *Proc. Comput. Cardiol.*, vol. 27, 2000, pp. 327–330.

[57] P. De Chazal, M. O'Dwyer, and R. B. Reilly, "Automatic classification of heartbeats using ECG morphology and heartbeat interval features," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 7, pp. 1196–1206, Jul. 2004.

[58] I. Silva and G. B. Moody, "An open-source toolbox for analysing and processing PhysioNet databases in MATLAB and octave," *J. Open Res. Softw.*, vol. 2, Sep. 2014, Art. no. e27.

[59] D. Lai, X. Fan, Y. Zhang, and W. Chen, "Intelligent and efficient detection of life-threatening ventricular arrhythmias in short segments of surface ECG signals," *IEEE Sensors J.*, vol. 21, no. 13, pp. 14110–14120, Jul. 2021.

[60] R. Taco, I. Levi, M. Lanuzza, and A. Fish, "An 88-fJ/40-MHz [0.4 V]–0.61-pJ/1-GHz [0.9 V] dual-mode logic 8×8 bit multiplier accumulator with a self-adjustment mechanism in 28-nm FD-SOI," *IEEE J. Solid-State Circuits*, vol. 54, no. 2, pp. 560–568, Feb. 2019.

[61] R. Taco, I. Levi, M. Lanuzza, and A. Fish, "Evaluation of dual mode logic in 28nm FD-SOI technology," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2017, pp. 1–4.

[62] A. Gautier, M. Dael, R. Benarrouch, B. Larras, and A. Frappé, "An ultra-low-power integrated heartbeat detector for wearable sensors," in *Proc. 15th EAI Int. Conf. BODYNETS*. Tallinn, Estonia: Springer, Oct. 21, 2020, pp. 199–211.

[63] C. Crispin-Bailey, C. Dai, and J. Austin, "A 65-nm CMOS lossless bio-signal compression circuit with 250 FemtoJoule performance per bit," *IEEE Trans. Biomed. Circuits Syst.*, vol. 13, no. 5, pp. 1087–1100, Oct. 2019.

Li Xiaolin (Graduate Student Member, IEEE) received the B.E. (Electronics) degree in Internet of Things from both the University College Dublin, Dublin, Ireland, and Beijing University of Technology, Beijing, China, in 2019. She is currently pursuing the Ph.D. degree with the School of Electrical and Electronic Engineering, University College Dublin.

She is the Ambassador of IEEE Women in Engineering U.K. and Ireland region. Her research interests include developing a distributed biomedical signal classification strategy to address the issues of slow inferencing, high requirements of computation, and energy.

Dr. Xiaolin has received multiple IEEE CASS Student Grants and she was recognized as an IEEE Young Professional by IEEE flagship conference.

Barry Cardiff (Senior Member, IEEE) received the B.Eng., M.Eng.Sc., and Ph.D. degrees in electronic engineering from University College Dublin, Dublin, Ireland, in 1992, 1995, and 2011, respectively.

He was a Senior Design Engineer or a Systems Architect for Nokia, from 1993 to 2001, moving to Silicon and Software Systems (S3 group), Dublin thereafter as a Systems Architect in their Research and Development division focused on wireless communications and digitally assisted circuit design. Since 2013, he has been an Assistant Professor with University College Dublin. He holds several U.S. patents related to wireless communication. His research interests include digitally assisted circuit design and signal processing for wireless and optical communication systems.

Deepu John (Senior Member, IEEE) received the B.Tech. degree in electronics and communication engineering from the University of Kerala, Thiruvananthapuram, India, in 2002, and the M.Sc. and Ph.D. degrees in electrical engineering from the National University of Singapore, Singapore, in 2008 and 2014, respectively.

From 2014 to 2017, he was a Post-doctoral Researcher with the Bio-Electronics Laboratory, the National University of Singapore. Previously, he was a Senior Engineer with Sanyo Semiconductors, Gifu, Japan. He is currently an Assistant Professor with the School of Electrical and Electronics Engineering, University College Dublin, Dublin, Ireland. His research interests include IoT/wearable sensing, biomedical circuits and systems, and edge computing.

Dr. John served as a member of the Organizing or Technical Committee for several IEEE conferences, such as ISCAS, BioCAS, NorCAS, ICECS, AICAS, MWSCAS, TENCON, ASICON, and ICTA. He has received several awards, including the Institution of Engineers Singapore Prestigious Engineering Achievement Award, the Best Design Award at the Asian Solid-State Circuit Conference in 2013, and the IEEE Young Professionals, Region Ten Individual Award. He is a reviewer of several IEEE journals and conferences. He served as the Guest Editor for IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS-I and IEEE OPEN JOURNAL OF CIRCUITS AND SYSTEMS. He is an Associate Editor of IEEE TRANSACTIONS ON BIOMEDICAL CIRCUITS AND SYSTEMS, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS-II, *IEEE IoT Magazine*, and Wiley *International Journal of Circuit Theory and Applications*.