Interpretable Rule Mining for Real-Time ECG Anomaly Detection in IoT Edge Sensors

Gawsalyan Sivapalan, Koushik Kumar Nundy, *Senior Member, IEEE*, Alex James, *Senior Member, IEEE*, Barry Cardiff, *Senior Member, IEEE*, and Deepu John, *Senior Member, IEEE*

***Abstract*—Electrocardiogram (ECG) analysis is widely used in the diagnosis of cardiovascular diseases. This article proposes an explainable rule-mining strategy for prioritizing abnormal class detection in ECG data. The proposed method utilizes a biased-trained artificial neural network (ANN) with input features derived from an ECG beat sequence and formulates a set of rules at each node of an on-demand tree-like search algorithm. The rule base at each node is derived from a linear combination of the most impactful features identified using gradient analysis in an ANN. The final derived model is an explainable rule-based system that detects abnormal heartbeats based on statistical and morphological features from ECG. The model achieves the target sensitivity, and accuracy with a low run time complexity through a comprehensive offline rule-mining process and is trained using the MIT-BIH Arrhythmia Database. The system achieves an accuracy of 93% with only nine nodes and a test sensitivity of 90% and 80%, respectively, for VEB and SVEB beat types, when tested on previously unseen ECG data from the INCART database. The model performance and complexity can be easily adjusted based on the real-time resource constraints of a wearable sensor. The model was deployed on an ARM Cortex M4-based embedded device and is shown to achieve a > 50% reduction in sensor power consumption when only abnormal beats are wirelessly transmitted. That is, RF transmission is gated using the model output and transmission is disabled when the subject's ECG is normal. The proposed technique is highly suited for healthcare applications because of its explainability, lower complexity, and real-time flexibility when deployed in the Internet of Things (IoT)-enabled wearable edge sensors.**

***Index Terms*—Anomaly detection, artificial neural network (ANN), electrocardiogram (ECG), explainability, heart rate variability (HRV), Internet of Things (IoT) edge sensors, rule-based algorithm.**

I. Introduction

ARRHYTHMIAS, such as respiratory sinus arrhythmia, are a natural periodic variation in the heart rhythm and typically do not have negative consequences to the individual's health. However, some other arrhythmia types may indicate a serious problem that may eventually lead to stroke or sudden cardiac death [1]. Continuous and autonomous monitoring for anomalies in electrocardiogram (ECG) signals using an Internet of Things (IoT)-enabled wearable device could mitigate risks due to cardiovascular diseases [2]. However, continuous monitoring of medical-grade ECG signals is not yet a practical reality due to poor usability (high-power consumption) of solutions and interpersonal variability affect the data interpretation quality. To make the detection system reliable, it is required to extract robust features from the ECG signals. It is not feasible to keep individuals under clinical monitoring for prolonged periods. Even though long-term data can be collected using Holter devices, due to the large volume of the data involved, it is difficult for clinicians to manually analyze it effectively. If we use a wearable device and choose to send all data in real-time to a cloud server, it will end up consuming significant power and limit the useful battery life of the sensor. Therefore, a wearable device that can automatically detect anomalous ECG beats and classify them into Normal and Abnormal beats using a low-complexity approach is desirable. This way, only anomalous beats (which occur infrequently in most subjects) need to be transmitted to a cloud server, which can save resources in the edge device. A second-stage classifier using more complex techniques with cardiologist-level accuracy could be implemented in the cloud platform for comprehensive multiclass classification of the anomalous data.

There are several techniques, including rule-based methods and machine-learning methods for the detection of abnormal heartbeats [3], [4], [5], [6]. Many such methods use black-box AI models and any attempt to increase detection accuracy typically results in an increased computational complexity [7]. As complexity increases, the power and space footprints of the wearable platform increases proportionally. Further, the complex data analysis would require cloud services, and introduces communication overheads and high-memory usage. Rule-based algorithms are usually simpler to implement and by definition explainable. However, we have identified common unexplored areas in existing rule-based algorithms [8], [9], [10], [11], [12], [13] as listed follows.

1) Normal and Abnormal ECG beat classes are treated equally and there is no prioritization to detect the critical Anomalous beats.
2) Feature selection is manually done and an expert in the loop is needed to select the perfect combination of features.

Manuscript received 22 August 2022; revised 9 November 2022; accepted 28 February 2023. Date of publication 22 March 2023; date of current version 25 July 2023. This work was supported in part by the Microelectronic Circuit Centre Ireland under Grant MCCI-2018-03, and in part by the JEDAI Project under the Horizon 2020 FET Chist-Era Program. *(Corresponding author: Deepu John.)*

Gawsalyan Sivapalan is with Meta Reality Lab, Redmond, WA 98052 USA (e-mail: gawsalyans@meta.com).

Koushik Kumar Nundy is with Think Biosolution Limited, Dublin, D08 YPP9 Ireland (e-mail: kknundy@thinkbiosolution.com).

Alex James is with the Maker Village, Digital University Kerala, Thiruvananthapuram 695581, India (e-mail: apj@ieee.org).

Barry Cardiff and Deepu John are with the School of Electrical and Electronic Engineering, University College Dublin, Dublin 4, D04 V1W8 Ireland (e-mail: barry.cardiff@ucd.ie; deepu.john@ucd.ie).

Digital Object Identifier 10.1109/JIOT.2023.3260722

3) Structure of the rule-based systems is fixed or rigid, making any real-time improvements or pruning on the edge node a difficult task.
4) Testing is only conducted over the records from the same database used for training the model and, therefore, the performance results are not reliable or reproducible on new or previously unseen data.

To address the aforementioned research gaps and perform computations at edge devices, we propose an easy to interpret rule-based algorithm (*Chain of Rules*) that can detect abnormal beats with high sensitivity and accuracy whilst remaining low in complexity.

A. Chain of Rules Algorithm

The *Chain of Rules* runtime algorithm, as illustrated in Fig. 1, is essentially that of navigating a binary tree where, at each node, a different rule is applied until the end of a branch is reached where the result of the last rule applied is then taken as the final result. The length of each possible branch is not necessarily identical and the set of rules/thresholds applied at each node are all trained parameters that have been chosen so as to achieve a target sensitivity and accuracy through a comprehensive (offline) training process.

For each ECG beat, we compute several features to be used with this algorithm (the exact subset of features used is a trained parameter). In the context of this work, we define a *Rule* as the comparison of a specific feature index/s against a computed (trained) threshold to make a tentative Normal or Abnormal beat classification.

B. Contributions

The main contributions of this work are as follows.

1) The *Chain of Rules* runtime algorithm is presented. Given a training population of ECG beats, we define a systematic procedure to generate an on-demand structure of the tree (the length of each branch) and the set of rules to apply at each node. The algorithm seeks to achieve the desired accuracy and sensitivity (focused approach over the unbalanced nature of the training data and priority class detection).
2) Developed a rule-based model for ECG anomaly detection, whose performance-complexity can change by adding/deleting nodes to the existing *Chain of Rules* by simply following the steps described in Algorithm 3. The model was tested on a previously unseen INCART ECG database and performance metrics are reported.
3) Implementation and experimental verification of the developed model on an ARM Cortex M4 Microcontroller-based Bluetooth SoC, demonstrating reduction in sensor power consumption.

The remainder of this article is organized as follows. Section II explains the currently available solutions and algorithms for ECG classification and anomaly detection using rule mining techniques. Section III reveals the runtime algorithm obtained through our novel rule-mining algorithm. Section IV details the proposed methodology for generation of rule-based algorithm while Section V delineates the obtained final *Chain of Rules* runtime algorithm. The performance analysis of the proposed method is evaluated and presented in Section VI.

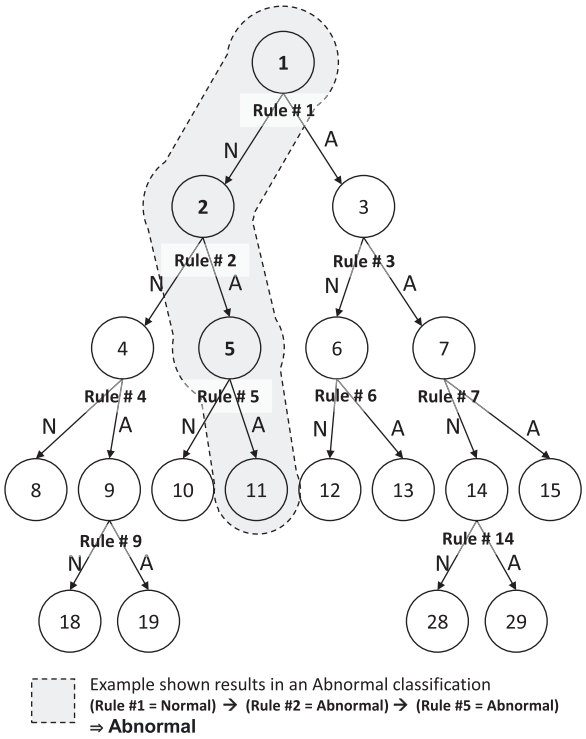

Fig. 1. Illustration of the *Chain of Rules* procedure as in Algorithm 2 and the final derived model.

Appendixes A and B outline the ECG data set used in this study, preparation of the data, and definitions of both singleton and compound features extracted from the data.

II. Related Work

Automatic detection of cardiovascular diseases using continuous and autonomous systems has been the subject of extensive research. Recent advancements in machine learning, the abundance of medical data, and the increase in computing power have paved the way for the development of highly accurate artificial neural networks (ANNs) that can surpass cardiologist-level accuracy in detecting arrhythmia. However, these systems are heavily reliant on the availability of huge amounts of data and computational power due to their complex structures and sometimes even require patient-specific training even though some have recently managed to produce better performance with less annotated data [23] or with reduced complexity for edge implementation [24]. Additionally, these systems are mainly designed as black boxes, and even though they achieve high accuracy in the test data, clinicians may not able to trust the system unless otherwise explained using eXplainable artificial intelligence (XAI) frameworks such as local interpretable model-agnostic explanations [25] (LIMEs). This stops us from being able to guarantee the performance of these models in a practical environment involving actual subjects.

Previously, several researchers have built rule-based expert systems to detect and classify abnormal heartbeat and arrhythmia events from ECG signals. Expert systems is a branch

TABLE I
SUMMARY OF RULE MINING TECHNIQUES USED IN ECG BEAT CLASSIFICATION

Year	Method	Technique	Features	Classification
*2005	Tsipouras et al [14]	Knowledge Base from expert cardiologists	RR Intervals	Normal Sinus, PVC, Ventricular flutter / fibrillation, 2 deg Heart Block
*2006	Exarchos et al [15]	Association Rule Mining, Classification based on predictive association rules (CPAR)	ST, T-wave, Age	Ischemic beat
*2007	Exarchos et al [16]	Fuzzy Expert System (Rules derived from Decision Tree)	ST, T and QT based features according to expert cardiologists	Ischemic beat and arrhythmic beat (Tsipouras et al)
2011	Mariano et al [17]	Floating Feature Selection algorithm (SFFS) [13]	RR, QRS duration, Maximal Vector of QRS loop (VCG_m) and the angle (VCG_ϕ), Peak amplitudes and positions, auto-correlation and inter lead cross correlation from the fourth scale of the DWT, wavelet scale, where the QRS complex is centered for each lead	N, S, V
2018	He et al [18]	Neighbor related information with Earth Movers Distance [19], Refined using SVM and Mix Ensemble	RR, HOS (skewness and kurtosis), level 7 Haar Wavelet decomposition	N, S, V
2021	Bidias'a Mougoufan et al [20]	Ordinal Patterns series (Conditional Entropy of Ordinal Patterns (CEOP [21]) and Permutation Entropy (PE [22]))	RR, QRS complex	N, S, V

*Early researchers does not comply with the AAMI recommended protocol
Refer Table VI for generalized comparison.

of artificial intelligence that makes extensive use of specialized knowledge to solve problems at a level equivalent to a human expert which can be represented by a set of rules [26]. Rules were inferred and obtained from cardiologists, fuzzy expert systems, and simple machine-learning approaches such as decision trees. Most of these decision support systems are well understood by any novice user. In this section, we will explore some of the related work in the domain and the techniques used in the development of these rules. A summary of the respective works is provided in Table I.

Early researchers in the field have generated rules with expert cardiologist knowledge and other rule mining algorithms, such as classification based on predictive association rules (CPARs [8]), first-order inductive learner (FOIL [9]), classification based on associations (CBAs [10]), classification based on multiple-class association rules (CMARs [11]), and apriori-total from partial classification (Apriori-TFPC [12]). However, the comparison and related performance are not clearly available to us as Tsipouras et al. [14] and Exarchos et al. [15], the researchers who have implemented these approaches for ECG analysis have not followed the standards recommended by association for the advancement of medical instrumentation (AAMI) to evaluate the algorithms.

Llamedo and Martínez [17] used a floating feature selection algorithm to obtain the best performing and generalizing models in the training and validation sets for different search configurations. The best model found comprehends eight features and obtained a small and well-performing model, through the sequential floating feature selection algorithm (SFFS) [13]. The SFFS is a combination of a sequential forward selection (SFS) algorithm followed by a sequential backward selection (SBS) algorithm. Starting from a single feature model, SFFS iterates for all model sizes while registering the best performances found for each model size. Each iteration starts with a model size greater than two features. After each SFS step, an SBS step is repeated until the new performance of the model is not further improved than the registered best model for the respective size. This way the algorithm floats forward and backward searching for the path of maximum performance. The algorithm terminates at the specified model size with the registered maximum performance achieved through the search.

Recent abundance of data and large compute power has drawn significant interest toward machine-learning methods. Therefore, some researchers have revisited the old school technique of rule mining with transfer learning obtained from traditional machine-learning approaches. He et al. [18] and Bidias à Mougoufan et al. [20] have proposed models with crisp rules that have been obtained using simple machine-learning techniques, such as SVM and Mixed Ensembles.

He et al. [18] proposed a pyramid-like model. The model distinguishes the classification of Normal and S beats and takes advantage of the neighbor-related information to assist identification of S beats. RR intervals (RR_i and RR_{i+1}), higher order statistics (HOSs) (skewness and kurtosis), and 7 levels of Haar wavelet decomposition coefficients are used as feature sets. The algorithm then looks for a patient from a training group, who has the most similar RR_i values distribution and assigns its threshold value. The function uses the Earth mover's distance (EMD) to measure the dissimilarity of two distributions and the heartbeats are processed by the decision rules in the first pyramid level [19]. The second level uses SVM and Mix Ensemble models to refine the classifications.

Bidias à Mougoufan et al. [20] presented a training-free two-level hierarchical model based on ordinal patterns. The classification rules include morphological and temporal properties of the ECG signal that are compared to R-R and QRS dependent thresholds derived from the beat, conditional entropy of ordinal patterns (CEOPs), or permutation entropy (PE) series [21], [22]. The beat length, entropy, skewness, and other features, such as the beat mean value and the beat standard deviation of RR intervals have been extracted and used for defining quantifiers and thresholds that will be used in the classification process.

Algorithm 1: Basic Structure of a "Rule"

Result: Normal / Abnormal beat tentative decision
if *Feature* $\geq$ *Threshold* **then**
  Normal;
else
  Abnormal;
end

Algorithm 2: Chain of Rules Runtime Algorithm

Result: Normal / Abnormal beat decision
Compute the initial set of features
Set current node index to root: $n = 1$
while *Node n not a leaf node* **do**
  Apply node specific Rule and get tentative decision
  if *tentative decision* == *Normal* **then**
    $n = 2n$
  else
    $n = 2n + 1$
  end
end
if *n is even* **then**
  Result = Normal
else
  Result = Abnormal
end

Algorithm 3: Top-Level GenEIC Algorithm for Building a *Chain of Rules*

Result: Chain of Rules definition
Inputs:
$\mathcal{F}$: the set of all annotations and pre-calculated singleton features for all beats used to train whole network.
θ_{Acc}: Target Accuracy
θ_{Sen}: Target Sensitivity
Initialize:
Let $L = \{1\}$ be the initial set of leaf nodes, i.e. just the the root node.
Let $\mathcal{F}_1 = \mathcal{F}$ be the set of annotated beats and their pre-calculated singleton features used to train the root node.
Let Acc = 0 and Sen = 0
Procedure:
while ($Acc < \theta_{Acc}$ *or* $Sen < \theta_{Sen}$) **do**
  Compute accuracy for all leaf nodes, i.e., $\forall\ n \in L$.
  Let $L' \subset L$ be that subset of leaf nodes where the accuracy is $< \theta_{Acc}$.
  Let n be the index of the largest of maximum achievable incremental scores by correct classification of subspace $\mathcal{F}_{n^*}$: $n = \arg\max_{n^*} \left\{ \begin{matrix} F_{\beta=1}(\mathcal{F}_{n^*}) & \forall\ even\ n^* \in L' \\ F_1(\mathcal{F}_{n^*}) & \forall\ \ odd\ n^* \in L' \end{matrix} \right\}$
  Remove n from L and add it's children instead: $L = \{L \setminus n\} \cup \{2n, 2n+1\}$
  Generate Rule for node n: $[\text{Rule}_n, \mathcal{F}_{2n}, \mathcal{F}_{2n+1}] = \text{GenerateRule}(\mathcal{F}_n, n)$
  Simulate new *Chain of Rules* algorithm and compute accuracy (Acc) and sensitivity (Sen).
end

Although some algorithms perform sufficiently well for the proposed continuous feature domain, the feature selection to optimize the classification of a specific class is not proposed as part of the strategy. The ability to determine the features that contribute to the formulation of the most efficient rules from annotated data is extremely painstaking. On the other hand, these algorithms do not guarantee the detection of difficult beat classes that need a higher degree of attention. Therefore, we have proposed a novel feature selective first-order rule mining methodology to address the problem.

III. Chain of Rules Runtime Algorithm

The *Chain of Rules* algorithm is the repeated application of "Rules" (Algorithm 1) where we evaluate a given feature with a specific threshold, resulting in a tentative Normal or Abnormal beat decision. It is the process of navigating a binary tree (starting from the root node, labeled node 1) and applying a rule associated with that node until a leaf node is reached. The result of the final rule in the tree gives the final classification decision of the ECG beat. Prior to running the *Chain of Rules* algorithm the necessary features for each ECG beat must be extracted. The full set of possible features is described in Appendix B, and the specific set of features that an algorithm instance requires is determined during an off-line training process as defined in Section IV.

The procedure is summarized in Algorithm 2, and as an example, the final model is illustrated in Fig. 1, where the nodes are labeled such that the children of the nth node are $2n$ and $2n + 1$ for Normal and Abnormal tentative decisions, respectively. Accordingly, if a node's label is even then the previous tentative decision was Normal, and likewise, odd nodes were preceded by Abnormal tentative decisions—this fact is used to determine the final Normal/Abnormal decision in step 11 of Algorithm 2. The population of nodes along with its impact on performance scores is presented in Section V.

If we limit the longest path from the root to any leaf node to have $L_{\max}$ Rules, then, there is a maximum of $2^{L_{\max}+1} - 1$ possible nodes (although many of these nodes do not exist as many branches terminate earlier than $L_{\max}$ rules). Accordingly, for each of the possible nodes, we must enumerate the following.

1) A binary flag indicating if this node is a leaf node or not.
2) Nonleaf nodes must define a Rule comprising:
   a) the feature to use;
   b) the comparison threshold.

Any such enumeration defines an instance of the *Chain of Rules* algorithm. The following section provides a training algorithm to create such an enumeration, and a specific instance is given in Section VI. Despite the perceived complexity of the training algorithm, it is worth noting that the runtime procedure as embodied in Algorithm 2 is very simple and can be easily implemented on an embedded device.

IV. Generate Evaluate Identify and Combine (GenEIC): Novel Approach to Generate Rules

The *Chain of Rules* algorithm is trained in an offline iterative process, called GenEIC as described in Algorithm 3. Starting with just one leaf node (the root node labeled as node 1) we continue to add additional leaf nodes until a top-level simulation of the current *Chain of Rules* indicates that the

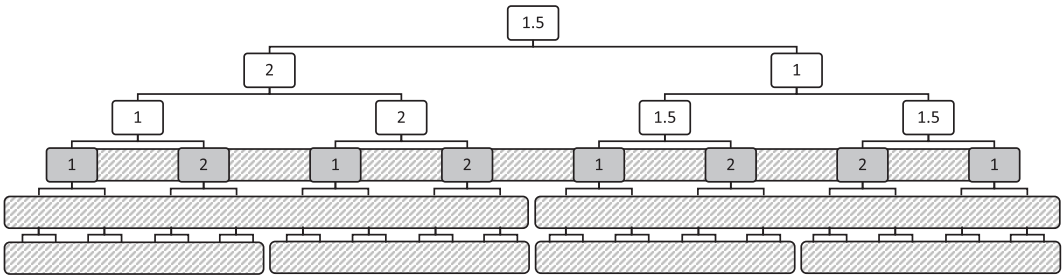

Fig. 2. β Strategy.

Algorithm 4: Generate Rules, Evaluate Them, and Identity the Best One for a Specific Node. [Rule, $\mathcal{F}_{\mathrm{N}}, \mathcal{F}_{\mathrm{A}}$] = GenerateRule($\mathcal{F}, n$)

Results:
Rule: The identified best rule (feature / threshold tuple) for this node.
$\mathcal{F}_{\mathrm{N}}$: That part of input database $\mathcal{F}$ that the Rule determines to be Normal.
$\mathcal{F}_{\mathrm{A}} = \mathcal{F} \setminus \mathcal{F}_{\mathrm{N}}$
Inputs:
$\mathcal{F}$: the set of all annotations and pre-calculated singleton features for all beats used to train this node.
n: The current node's index.
Procedure:
Compute β (using n), the parameter used to compute the F_β according to Fig. 2.
Use a Neural Network to determine **F**, the set of the N most impactful singleton features.
Expand **F** to include compound features *to* **G**.
for $k = 1$ *to* $2N - 1$ **do**
    Generate a possible Rule= $\{G_k, B_k\}$ using eq. (3).
    Evaluate / simulate this rule over the set $\mathcal{F}$ and compute the Acc: Accuracy and Sen: Sensitivity, and hence the F_β score for this Rule.
end
Identify the best Rule as that Rule having largest F_β.
Apply Rule to $\mathcal{F}$ yielding $\mathcal{F}_N$ and $\mathcal{F}_A$ subsets for Normal and Abnormal decisions, respectively.

algorithm achieves the desired target accuracy (θ_{Acc}) and sensitivity (θ_{Sen}). We denote as $\mathcal{F}$ the entire ECG database and their precalculated singleton features. Further, we denote as $\mathcal{F}_n$ as that part of $\mathcal{F}$ passing all the rules in the branch leading to node n, and for completeness, we let $\mathcal{F}_1 = \mathcal{F}$.

The process of adding new leaf nodes is actually that of *growing* the network by changing a leaf into a parent node, called n, and adding two child nodes, labeled $2n$ and $2n + 1$ representing Normal and Abnormal beat classes, respectively. The decision regarding which leaf node to grow is taken by considering those nodes with the largest incremental score[1] (w.r.t. the accurate classification of set $\mathcal{F}_n$)—the idea being that this node has the largest ability to impact the network's overall performance. We note that the accuracy of a node n w.r.t. the set $\mathcal{F}_n$ can be computed by observing that if n is even (or odd) then its decision would be normal (or abnormal) which, coupled with the ground-truth annotations in $\mathcal{F}_n$. To add a child to a node, n, we must determine the *best* Rule, denoted Rule$_n$, for that node and add it to the *Chain of Rules*. This is done by a comprehensive off-line subprocess that is summarized in Algorithm 4. It utilizes a biased training score called F_β (4) parameterized by β to select the best rule from a set of trained candidates' rules. However, this β parameter changes depending on the nodes' position in the *Chain of Rules*. Hence, the top-level algorithm passed the node index n to the lower level subprocess (see step 15 in Algorithm 3).

The acronym *GenEIC* is derived from the keywords.

1) Generate a set of candidate rules.
2) Evaluate the performance of candidate rules.
3) Identify the best rules.
4) Combine these rules into the chain.

The Combine step is already described in Algorithm 3, whereas the other steps are illustrated in Algorithm 4 and are described in the following sections.

A. Generate Set of Candidate Rules

This step involves several smaller substeps.

1) Reduce the full set of possible singleton features down to N most impactful ones.
2) Expand that set by adding some compound features.
3) Use statistical analysis to generate a candidate rule for each feature in this expanded set.

1) Identify Most Impactful Singleton Features: There are many alternative feature selection techniques available, and we used a feed-forward neural network,[2] NN (Fig. 3), with all possible singleton features used as inputs. Note that the input features were normalized across the entire training set such that they all have unity expected values; hereafter we only considered this *normalized* features.

Denote $\boldsymbol{W_h} \in \mathbb{R}^{N' \times N'}$ and $\boldsymbol{b_h} \in \mathbb{R}^{1 \times N'}$ as the weights and biases in the fully connected (FC) hidden layer and likewise $\boldsymbol{W_c} \in \mathbb{R}^{2 \times N}$ and $\boldsymbol{b_h} \in \mathbb{R}^{1 \times 2}$ for the classification layer, where N' is the total number of possible singleton features defined in Appendix B-A. Accordingly, the vector of gradients, $\boldsymbol{\Delta}$, used during the back-propagation can be calculated as

$$\boldsymbol{\Delta} = \boldsymbol{W_h}^T \times \left(\boldsymbol{W_c}^T \times (\boldsymbol{y_c} \cdot (1 - \boldsymbol{y_c}))\right) \cdot \boldsymbol{y_h} \cdot (1 - \boldsymbol{y_h}) \quad (1)$$

where $\boldsymbol{y_c}$ is the softMax output in the classification layer and $\boldsymbol{y_h}$ is the sigmoid output from the hidden layer.

[1] $F1$ score (5) for normal class/ odd nodes and F_β score (4) for abnormal class/ even node.

[2] The NN has two layers of perceptrons and two neurons in the classification final layer to decide in favor of Normal or Abnormal as depicted in Fig. 3.

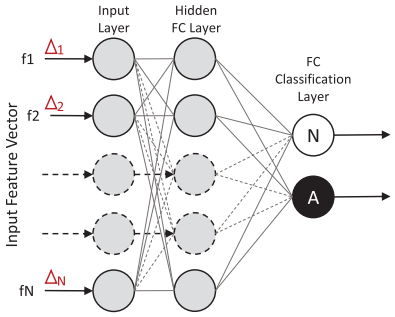

Fig. 3. ANN model used to identify the feature significance.

Record	Beat No.	Class	Feature Vector						Gradient of Feature Vector					
			RR_{pre}	RR_{pos}	RR_{mean}	RR_{std}	RR_{modstd}	$RR_{rtindex}$	ΔRR_{pre}	ΔRR_{pos}	ΔRR_{mean}	ΔRR_{std}	ΔRR_{modstd}	$\Delta RR_{rtindex}$
	7	A(A)	0.66	0.99	0.89	0.14	0.78	0.41	0.35	0.02	0.05	0.02	-0.18	-0.36
100	50	N	0.80	0.84	0.82	0.03	0.10	0.04	0.02	0.00	0.00	0.00	-0.01	-0.02
	100	N	0.78	0.80	0.81	0.02	0.11	0.03	0.02	0.00	0.00	0.00	-0.01	-0.03
	11	A(V)	0.48	0.78	0.65	0.13	0.62	0.48	0.26	0.01	0.04	0.02	-0.13	-0.27
200	12	N	0.78	0.56	0.63	0.13	0.58	- 0.33	0.03	0.00	0.00	0.00	-0.01	-0.03

Fig. 4. Example of gradient calculation for six particular features. When developing our rules, we rank these gradients in order of their magnitudes and take the N largest.

Borrowing a technique from the domain of image Classification [27], we select the N features corresponding to the gradients with the largest magnitudes as being the most impactful features. We denote this pruned feature set as $\mathbf{F}$, and order it such that $|\Delta_k| > |\Delta_{k'}|$ for all $k < k'$. Typically, we set N to be eight or half the total number of possible features (N') whichever is the smallest. We call $\mathbf{F}$ the set of useful *singleton* features as opposed to the *compound* features that are a function of multiple features.

By way of a toy example, if the ANN was trained using just the length $N' = 6$ feature vector

$$\left[\mathrm{RR}_i, \mathrm{RR}_{i+1}, \overline{RR}_i, \mathrm{RR}_{\mathrm{SDNN}_i}, \mathrm{RR}_{\mathrm{wSDNN}_i}, \mathrm{RR}_{\mathrm{Index}_i}\right]$$

using the entire MIT-BIH database we would get a $\sim 94\%$ accuracy and the trained weights and biases would be

$$\boldsymbol{W_h} = \begin{bmatrix} 0.22 & 0.31 & 0.38 & 0.18 & -0.18 & -0.17 \\ -1.24 & -0.17 & -0.17 & 0.12 & 0.53 & 1.40 \\ 0.37 & 0.15 & 0.17 & 0.03 & 0.14 & -0.00 \\ 1.97 & 0.25 & 0.24 & 0.08 & -1.15 & -1.76 \\ -0.08 & 0.07 & 0.08 & 0.21 & -0.20 & 0.21 \\ 1.91 & 0.29 & 0.19 & 0.07 & -1.48 & -1.57 \end{bmatrix}$$

$$\boldsymbol{W_c} = \begin{bmatrix} 0.20 & -1.43 & 0.17 & 2.17 & -0.16 & 2.13 \\ 0.06 & 1.80 & 0.22 & -1.87 & 0.43 & -1.84 \end{bmatrix}$$

$$\boldsymbol{b_h} = \begin{bmatrix} -0.04 & 0.24 & 0.06 & -0.72 & 0.05 & -0.69 \end{bmatrix}$$

$$\boldsymbol{b_c} = \begin{bmatrix} -0.3 & 0.5 \end{bmatrix}$$

and the corresponding gradients (1) are shown in Fig. 4.

2) Expand Set to Include Compound Features: Additionally, we define a size $2N - 1$ super set, $\mathbf{G}$, containing $\mathbf{F}$ plus an extra $N - 1$ features formed by taking linear combinations of the features in $\mathbf{F}$

$$G_k \triangleq \begin{cases} F_k & 1 < k \leq N \\ \sum_{k'=1}^{1+k-N} \mathrm{sgn}(\Delta_{k'}) \mathrm{F}_{k'} & N < k \leq 2N-1. \end{cases} \tag{2}$$

As this contains newly defined features derived from combinations of the features in $\mathbf{F}$, we refer to $\mathbf{G}$ as the set of *compound* features.

TABLE II
COMPARISON OF F_1 (5) AND F_β (4) SCORES FOR MODELS THAT CAN BE CONSIDERED AS BEST UNDER DIFFERENT OBSERVATIONS

Type	Data		Model Performance				Score	
	A	N	TP	FP	TN	FN	F_1	$F_{\beta=1.5}$
Unbalanced	1	99	1	9	90	0	**18%**	**96%**
	99	1	90	0	1	9	95%	91%
Balanced	50	50	45	10	40	5	86%	88%

3) Generate Candidate Rules: Here, the identified $2N - 1$ features in the set $\mathbf{G}$ (2) are analyzed using histograms. Fig. 5 shows the distribution of the individual features separately for Normal and Abnormal (VEB classes) as per the MIT-BIH database. Accordingly, for the kth feature the means $\mu_{N,k}$ and $\mu_{A,k}$ and standard deviations $\sigma_{N,k}$ and $\sigma_{A,k}$ are estimated for the Normal and Abnormal (VEB) classes, respectively. Fig. 5 also superimposes normal probability density function (PDF) curves generated with those estimated parameters for illustration purposes.

Given these estimated means and standard deviations, we define the following decision threshold for the kth feature

$$B_k = \frac{\mu_{N,k}\sigma_{A,k} + \mu_{A,k}\sigma_{N,k}}{\sigma_{N,k} + \sigma_{A,k}} \tag{3}$$

which ensures equal-error probability under certain statistical simplifying assumptions.[3] Each feature in $\mathbf{G}$ together with its threshold (3) comprise a candidate rule according to Algorithm 1 and there will be $2N - 1$ of such rules.

B. Evaluate the Candidate Rules With Priority Class

As a result of previous steps, we have now generated $2N-1$ candidate Rules for the nth leaf node. In this step, we evaluate the performance of each through simulation and rank them according to some measure of *goodness*. The measure used here is based on Rijsbergen's effectiveness measure [28], called the F_β score, as it allows us to attach β times more importance to Sensitivity (Recall) compared to Accuracy.

It is defined as

$$F_\beta \triangleq (1+\beta)\frac{\text{Accuracy} \times \text{Recall}}{\beta \times \text{Accuracy} + \text{Recall}} \tag{4}$$

where $\beta \in \{1, 1.5, 2\}$. The score is better when dealing with unbalanced data with priority minority class as positives compared to the commonly used F_1 score (5) as can be seen in the examples presented in Table II

$$F_1 \triangleq 2\frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}. \tag{5}$$

The choice of β depends on the position of the node within the tree, specifically the three previous Normal/Abnormal decisions in the branch leading up to node n as shown in Fig. 2. The Normal/Abnormal status of a node is determined by the even-/odd-ness of its label n, and a node's parent's labels are determined as $\lfloor n/2 \rfloor$.

[3]Statement is true if the PDFs are such that the distance from the mean in units of standard deviation is proportional to the probability.

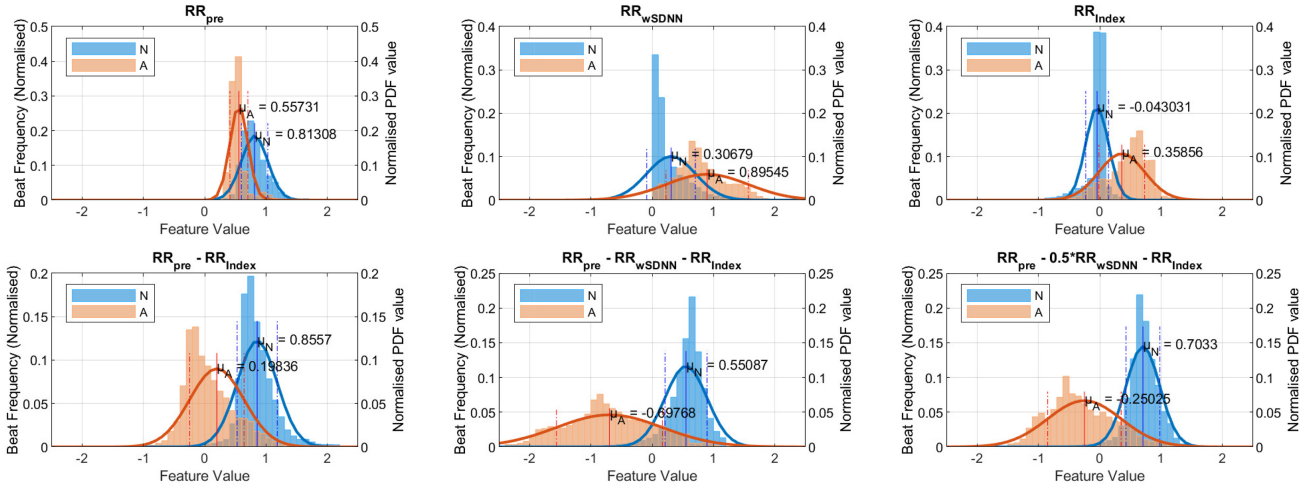

Fig. 5. Generated rules and their distribution in MIT-BIH.

TABLE III
G (2) AT ROOT NODE—SINGLETON AND COMPOUND FEATURE/S PERFORMANCE ON MIT-BIH ARRHYTHMIA DATABASE

Metric	TP	FP	TN	FN	$F_{\beta=1.5}$
Singleton:					
$S1:\ RR_{pre}$	10,029	44,617	45,445	570	74%
$S2:\ QRS_{Energy}$	4,626	21,355	68,707	5,973	52%
$S3:\ PCA$	5,332	27,593	62,469	5,267	56%
$S4:\ SD1$	7,739	18,399	71,663	2,860	75%
$S5:\ QRS_{Sign}$	3,500	3,025	87,037	7,099	44%
$S6:\ wSDNN$	8,299	23,586	66,476	2,300	77%
$S7:\ RRIndex$	7,389	7,836	82,226	3,210	76%
$S8:\ SDNN$	7,889	24,697	65,365	2,710	74%
Compound Rules:					
$C1: RR_{pre} - QRS_{Energy}$	5,583	19,324	70,738	5,016	60%
$C2:\ C1 + PCA$	6,827	10,960	79,102	3,772	71%
$C3:\ C2 - SD1$	7,196	9,946	80,116	3,403	74%
$C4:\ C3 + QRS_{Sign}$	6,822	8,641	81,421	3,777	72%
$C5:\ C4 - wSDNN$	7,331	8,453	81,609	3,268	76%
$C6:\ C5 + RRIndex$	**8,066**	**8,977**	**81,085**	**2,533**	**81%**
$C7:\ C6 - SDNN$	7,888	8,858	81,204	2,711	79%

C. Identify the Best Rule

Once the F_β score has been computed for each candidate Rule, the largest is selected as Rule$_n$—an example is shown in Table III where the 6th compound rule in this was chosen as its F_β was the highest at 81%. This step also helps to ensure the prioritised detection of abnormal classes without compromising accuracy.

Additionally, the supplied database $\mathcal{F}_n$ is divided according to the application of said rule into two subsets $\mathcal{F}_{2n}$ and $\mathcal{F}_{2n+1}$ for Normal and Abnormal decisions, respectively.

V. FINAL CHAIN OF RULES MODEL

In our experiments, we trained the *Chain of Rules* using ECG signals from MIT-BIH arrhythmia database and set $\theta_{\text{Acc}} = \theta_{\text{Sen}} = 90\%$. Starting from the root node (node 1), the process of node population to produce the final model is presented in Fig. 6 along with its performance scores at each step. The scores are not always guaranteed to improve with the

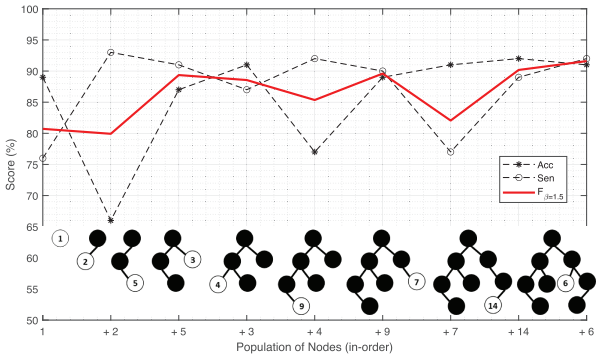

Fig. 6. Node population with change in performance.

growth of the nodes as the choice of the next node to grow is based on the achievable incremental F score found using the accurate classification of the subspace $\mathcal{F}_{n^*} \subset \mathcal{F}_{\lfloor n^*/2 \rfloor} \subset \mathcal{F}_1$. Our algorithm tries to find the perfect balance with the rules populated at even nodes pulling up the overall sensitivity while odd nodes improve the accuracy. The resulting final *Chain of Rules* model is presented in Table IV and is shown in Fig. 1.

We can see that out of the thirty-odd possible singleton features defined in Appendix B only eight are needed to achieve the desired θ_{Acc} and θ_{Sen}. Therefore, this instance of the algorithm has low complexity with minimal preprocessing requirements. Note also that while the eight singleton rules in effect define another seven compound rules, we only use six of them as given in Table III.

Another point to note is that simple RR interval-based features serve as the highest contributing features, and very often decisions can be made when only equipped with just these features. This presents the opportunity for (zero cost) additional power-saving optimization whereby one could limit the computation of the more complex features, e.g., the principal component analysis (PCA) coefficients, to only occasions where they are actually needed, essentially implementing a just-in-time feature calculation strategy. All of the test results presented in this work (in Section VI) are based on this configuration of the chain of rules.

TABLE IV
PROPOSED ECG CLASSIFICATION CHAIN OF RULES AS DEVELOPED USING THE GENEIC ALGORITHM

Node	Leaf	Features	A < / > N	Threshold	FLOPs/beat
1	0	$RR_i - QRS_{Energy,i} + PCA_i - SD1_i + QRS_{Sign,i} - wSDNN_i + RRIndex_i$	<	0.999	845
2	0	$RR_i - \sigma(VS)_i - SD1_i + QRS_{Sign,i} - wSDNN_i$	<	0.549	75
3	0	RR_i	<	0.759	1
4	0	$RR_i - \Sigma QRS_i - SD1_i$	<	−0.641	29
5	0	$RR_i - wSDNN_i$	<	0.537	2
6	0	$\sigma(PCA)_i$	>	2.097	43
7	0	RR_i	<	0.585	1
8	1	No Rule in leaf nodes	–	–	
9	0	$-\Sigma QRS_i + RR_i + \sigma(VS)_i - SD1_i$	<	−0.543	4
10-13	1	No Rule in leaf nodes	–	–	
14	0	$RRIndex_i - \sigma(PCA)_i + \sigma(VS)_i + RR_i + QRS_{Energy,i}$	<	2.544	117
15	1	No Rule in leaf nodes	–	–	
16,17	–	Nodes not reachable	–	–	
18,19	1	No Rule in leaf nodes	–	–	
20-27	–	Nodes not reachable	–	–	
28,29	1	No Rule in leaf nodes	–	–	
30,31	–	Nodes not reachable	–	–	

VI. EXPERIMENTAL SETUP

This section describes the evaluation methodology and performance of the proposed algorithm. The specific chain of rules model used was trained using the MIT-BIH database as described in Section V and shown in Fig. 2 and Table IV. To evaluate its real-world performance, we tested the model using more than 175 000 previously unseen heartbeats from the INCART [29] database. This testing approach is in accordance with the best practice of separating training and testing data sets which will result in robust and reliable performance measures and avoids any overfitting to any specific training data set [4]. Note that, using a different data set for testing necessitated signal resampling (to 250 Hz), as the two data sets were acquired at different sampling rates. The model development was done in MATLAB.

A. Evaluation Method

The evaluation of the classification algorithm is based on the standard performance matrix, including Acc: Accuracy, Sen: Sensitivity, Spe: Specificity, and Ppr: Positive Predictive and the F_β score described in Section IV-B. The correctly classified abnormal class is considered as true positives (TPs) while the correctly classified normal class is considered as true negatives (TNs). Incorrect classification of the abnormal class is represented as false negative (FN) while false positives (FPs) are the beats that are classified as abnormal by the algorithm while they are actually normal.

B. Performance

A simplified set of performance summary is presented in Table V where the records are ranked and clustered (high, good, average, and poor performance) according to their accuracy.

Roughly two-thirds of all test records achieved a very high-average sensitivity of 95% and an average accuracy of 98%. According to our analysis, record I71 is the worst performing record among all with only 11% accuracy but with 97% of sensitivity. It is largely due to the QRS energy concentration (QRS window) as shown in Figs. 9 and 11. Some other records (I54, I09) achieved low accuracy due to poor annotation[4] in the INCART database even after our window of correction is applied as seen in Fig. 9.

The overall training and testing performance with MIT-BIH and INCART databases are shown in Table VII. The accuracy is affected by the high FP as the algorithm is optimized with higher sensitivity toward Abnormal beats to reduce the chances of missing anomalous beats. Also, the higher FP is due to the possible violation of the concept of a normal ECG rhythm. For, e.g., in the MIT-BIH database, record 203 which accounts for a significant portion of FPs, seemingly Abnormal beats with irregular rhythm are annotated as Normal as illustrated in Fig. 12.

The INCART database, which is unseen by the model during training is used to evaluate the real-world performance of the technique [30]. The comparison with other state-of-the-art techniques which have provided sufficient information on their performance on the INCART database for VEB and SVEB classes are given in Table VI. A completely unseen test performance of 93% accuracy and 88% sensitivity was achieved in Table VII. The fact that we used a never seen database for testing purposes, affected the results (Table VI). However, the results reported are more reliable and the model should achieve similar real-world performance in any other new databases without any further training. We attempted to increase the detection sensitivity at the expense of accuracy, which caused the higher FPs in Table V. However, it is worth noting that typically an embedded application will behave as an "ECG recorder" and any episodes containing abnormal beats will be saved for more complex post-processing elsewhere (e.g., in the cloud). In such cases, what is important is the ability of the algorithm to identify episodes where at least one abnormal beat resides. We implemented this and found that if we flag all 5-s duration windows containing at least one abnormal beat, then the accuracy can be increased to 95%

[4]By right these records should be removed from our results but we have elected not to do so to permit fair comparison with other researcher's works.

TABLE V
TEST PERFORMANCE ON INCART DATABASE

Rec No.	TN	FN	TP	FP	Acc	Sen
Group A: High Performance (46/75)					**98%**	**95%**
I52	1,608	0	137	1	100%	100%
I61	1,451	1	0	0	100%	0%
I16	1,516	0	2	3	100%	100%
I75	1,481	10	609	1	99%	98%
I23	2,179	0	13	12	99%	100%
I69	1,988	2	167	10	99%	99%
I45	1,426	3	488	9	99%	99%
I34	1,425	12	524	2	99%	98%
I43	1,077	9	1,113	8	99%	99%
I37	1,996	7	445	12	99%	98%
I39	1,451	5	308	9	99%	98%
I15	2,609	0	3	21	99%	100%
I28	1,696	0	4	15	99%	100%
I51	1,947	8	798	17	99%	99%
I33	1,245	17	573	0	99%	97%
I70	1,522	0	126	16	99%	100%
I24	2,537	0	6	26	99%	100%
I32	1,545	2	55	15	99%	96%
I17	1,626	0	27	18	99%	100%
I48	2,093	3	235	25	99%	99%
I53	1,137	17	1,091	10	99%	98%
I03	2,295	2	125	28	99%	98%
I13	1,765	0	230	27	99%	100%
I41	1,606	5	1	17	99%	17%
I26	1,479	4	7	18	99%	64%
I50	2,951	2	2	42	99%	50%
I14	1,770	2	62	30	98%	97%
I58	2,272	3	9	39	98%	75%
I42	1,502	15	1,547	43	98%	99%
I64	1,864	15	11	21	98%	42%
I07	2,612	27	39	26	98%	59%
I72	1,852	32	362	22	98%	92%
I49	2,058	2	25	61	97%	93%
I44	1,738	10	681	64	97%	99%
I04	2,222	30	124	46	97%	81%
I57	2,757	10	19	83	97%	66%
I21	2,054	60	52	17	96%	46%
I63	1,832	58	89	13	96%	61%
I11	2,025	24	3	51	96%	11%
I66	2,059	11	190	77	96%	95%
I73	1,829	17	85	60	96%	83%
I74	2,072	87	235	9	96%	73%
I25	1,632	1	6	72	96%	86%
I06	2,359	35	22	76	96%	39%
I56	1,613	20	13	58	95%	39%
I47	1,771	5	89	87	95%	95%
Group B: Good Performance (24/75)					**91%**	**81%**
I68	2,352	5	158	128	95%	97%
I46	2,113	17	409	118	95%	96%
I59	2,006	50	31	60	95%	38%
I08	1,675	13	339	100	95%	96%
I05	1,505	83	173	13	95%	68%
I20	2,287	79	211	73	94%	73%
I19	1,159	69	780	53	94%	92%
I22	2,736	109	201	78	94%	65%
I60	2,316	0	0	157	94%	–
I65	2,105	11	380	164	93%	97%
I02	2,294	37	192	149	93%	84%
I62	1,422	164	652	30	91%	80%
I35	3,028	160	315	169	91%	66%
I01	2,185	40	303	227	90%	88%
I38	2069	203	341	83	89%	63%
I30	1516	76	678	189	89%	90%
I12	2491	3	6	304	89%	67%
I36	3200	194	266	234	89%	58%
I10	3156	11	72	433	88%	87%
I55	1890	6	12	256	88%	67%
I27	1561	13	706	323	87%	98%
I67	2150	107	429	282	87%	80%
I29	1560	124	655	275	85%	84%
I18	2323	177	245	336	83%	58%
Group C: Average Performance (3/75)					**79%**	**82%**
I31	1,430	249	1,109	403	80%	82%
I09	2,345	20	29	600	79%	59%
I40	1,981	4	93	586	78%	96%
Group D: Poor Performance (2/75)					**42%**	**86%**
I54	1,495	7	16	844	64%	70%
I71	154	1	34	1,480	11%	97%
INCART	**144,048**	**2,605**	**19,587**	**9,464**	**93%**	**88%**

while achievable maximum accuracy with additional processing of transmitted data is 98.5% making this algorithm now extremely attractive for use in an embedded application.

TABLE VI
GENERALIZED COMPARISON FOR THE INCART

Method	Acc	SVEB	VEB
Proposed GenEIC	93%	80%	**90%**
CEOP, 2021 [20]	95%	62%	72%
PE, 2021 [20]	**95%**	64%	72%
Pyramid, 2018 [18]	90%	79%	87%
SFFS, 2011 [17]	91%	**85%**	82%

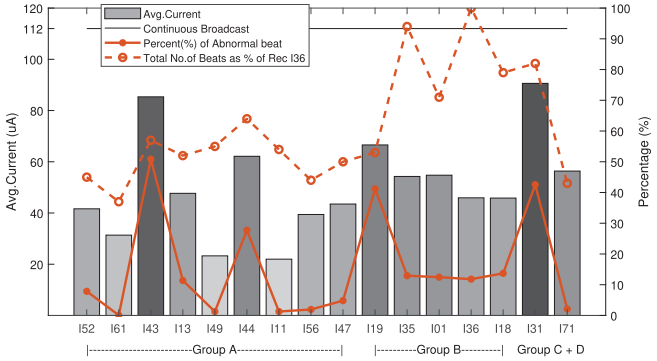

Fig. 7. Measured Avg current on different records during testing.

The robust performance of the proposed system is not only evident by the generalized test performance metrics, rather it is also due to the built-in regularization strategy to reduce the overfitting. The first-order linear combination mapping of ANN learning and use of only the subset of data $\mathcal{F}_n \subset \mathcal{F}_1$ for training similar to bagging (Bootstrap Aggregation learning) at every node ensure it.

C. Complexity and Power Analysis

The final *Chain of Rules* model is reimplemented in C language and ported to Nordic Semiconductor NRF52 Cortex ARM M4F microcontroller-based Bluetooth SoC development kit (NRF52DK). Using a USB to SPI bridge (FTDI4222H), the test signals are transferred to the embedded board from a desktop computer (PC) as shown in Fig. 10. Floating-point operations for feature calculation and rule execution is 964 FLOPs/beat at its max via path (node $1, 3, 7,$ and 14 in Table IV).

Randomly selected INCART test records from each performance group (refer Table V) are fed to the device in real time. Nordic Semiconductor's nRF Power Profile Kit is used to measure the power consumption of the system for each 30-min record and the results are shown in Fig. 7.

Wireless transmission consumes the most power in an IoT wearable sensor and, therefore, we have gated the wireless transmission such that the transmission is enabled only when Abnormal beats are detected [4]. These Abnormal beats could be further analysed in the cloud or evaluated by clinicians, while Normal beats could be ignored or stored locally in the sensor. In the proposed system (Fig. 8), the power consumption depends on the number of times the rule-based algorithm is triggered and the number of times a wireless transmission occurs. The triggering of the algorithm is subject to the detection of beats and, therefore, the records with a higher number of beats will consume more power. On the other hand, since

TABLE VII
OVERALL PERFORMANCE

Database	Total	TP	FP	TN	FN	Acc	Spe	Sen	Ppr	F_β	AAMI Sensitivity		
											VEB	$SVEB$	F
Training	100,661	9,764	8,065	81,997	835	**91%**	91%	**92%**	54%	**92%**	**98%**	87%	62%
Testing	175,704	19,587	9,464	144,048	2,605	**93%**	94%	**88%**	67%	**90%**	**90%**	80%	48%

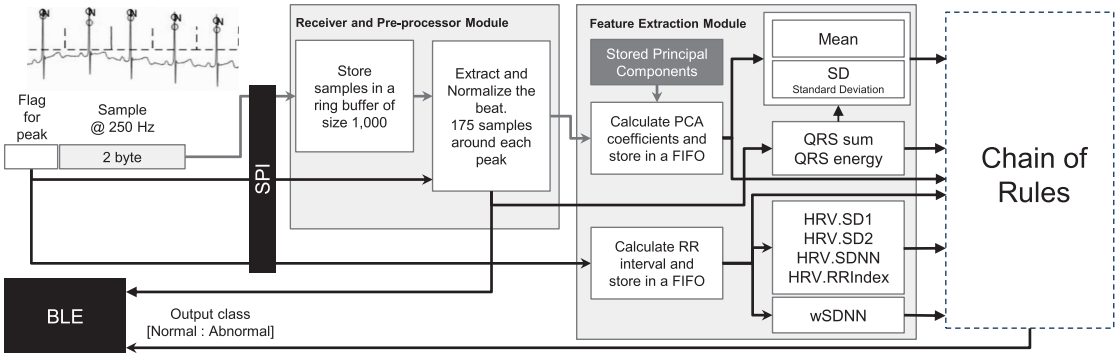

Fig. 8. Implementation flow diagram on embedded environment.

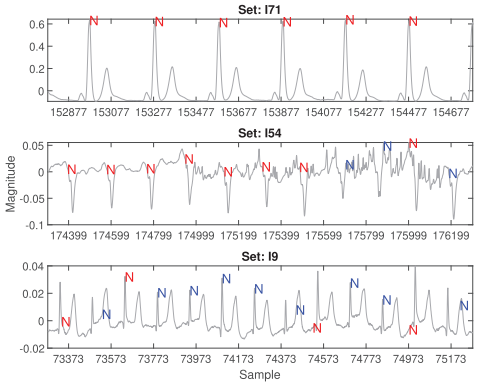

Fig. 9. Highest number of FPs by the proposed rule-based system in records I71, I54, and I09. (Red annotations are the misclassified and blue are the correctly classified beats).

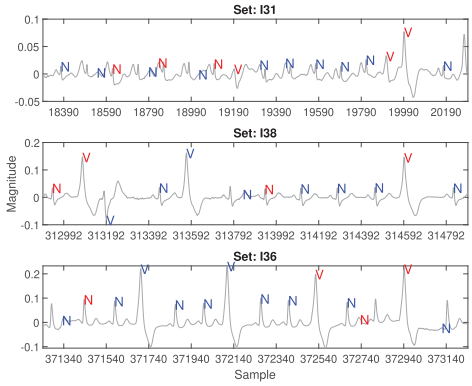

Fig. 11. Highest number of FNs by the proposed rule-based system in records I31, I38, and I36. (Red annotations are the misclassified and blue are the correctly classified beats.)

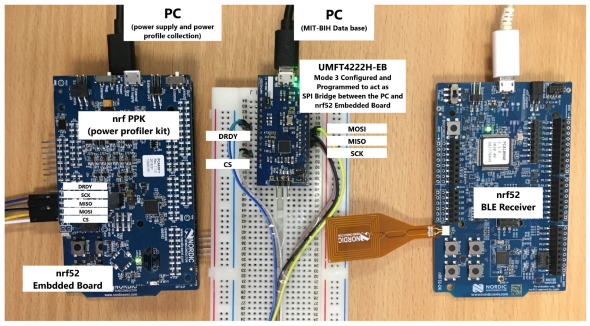

Fig. 10. Test setup for experimental verification.

radio transmission is enabled only for Abnormal beats, the number of such beats influences the power consumption. In the case of most patients, anomalous beats occur only infrequently and, therefore, overall system power is reduced using this approach.

The implementation of the algorithm using fixed-point arithmetic will introduce quantization error at each level. However, unlike in the case of ANNs, these errors will not propagate exponentially through layers and our experiments show that the performance of the fixed-point model is stable. It can be seen from Fig. 7, continuous wireless transmission of ECG data consumes an average current of 112 μA. By gating wireless transmission for Normal beats, the average current can be reduced to < 50 μA when tested using several ECG records from INCART database. That is a reduction of $> 50\%$ power in the wireless transmitter.

Compared to complex deep neural networks, our proposed model relies on an interpretable set of rules, i.e., it falls into the broad category of explainable AI. These rules are simple to implement and are generated offline using statistical knowledge with assistance from an ANN to recognize the most important signal features. The proposed approach has an overall *unseen test accuracy* of 93% and sensitivity is at 88%. Unlike many existing works which train and test a model using the same database, the performance reported in this work is reliable and reflects real-world performance and this may be further improved by adding additional nodes to the binary tree. The univariate coefficients such as PCA may not serve as sufficient for the morphological representation of the beat.

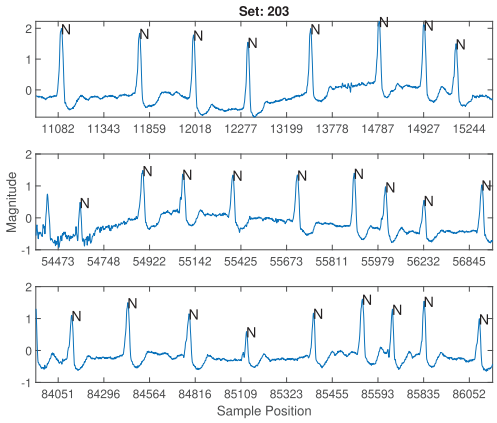

Fig. 12. Example of FPs: abnormal rhythmic activity of beats that annotated as normal in MIT-BIH arrhythmia database record 203 which achieved very low accuracy due to its signal quality and abnormal RR intervals.

TABLE VIII
AAMI STANDARD

AAMI	Description	Physionet	Description
N	Normal beats	L	Left Bundle Branch Block
		N	Normal
		R	Right Bundle Branch Block
		e	Atrial escape beat
		j	Nodal (junctional) escape beat
S	Supra-ventricular ectopic-beat	A	Atrial premature beat
		J	Nodal (junctional) premature beat
		S	Supraventricular premature beat
		a	Aberrated atrial premature beat
V	Ventricular ectopic-beat	E	Ventricular escape beat
		V	Premature ventricular contraction
F	Fusion beat	F	Fusion of Ventricular and normal beat
Q	Unknown beat	P	Paced beat*
		U	Unclassified beat
		f	fusion of paced and normal beat*

*The beats which belong to record no 102, 104, 107 and 217 from MIT-BIH database are rejected in this study as per the recommendation from AAMI.

Considering more coefficients of PCA with respect to lower eigenvalues and different class representations might introduce a huge performance impact on the rules mining. The discarded fiducial point information, such as PR duration, T duration, and QT duration in the first level of analysis may have a significant impact on deeper levels in classification. However, the effective gain anticipated with such features when compared to their calculation complexity is estimated to be insignificant.

VII. CONCLUSION

In this study, we have proposed a novel approach for mining a set of rules using gradient analysis of a trained neural network for ECG beat classification. The mining algorithm is specifically designed to be highly sensitive to a particular priority class in an unbalanced data set. The product is an explainable classification model with low run time complexity and tested in laboratory conditions, simulating real-world conditions. The proposed model achieves high real-world detection performance on previously unseen subjects who were not part of the training set. We have analyzed roughly 30 ECG extracted features (Table III) for developing the model and most of them are based on ventricular rate (RR intervals). Features based on the ECG morphology have not been parameterized in this work, but may be used for improving the model accuracy in future works. The proposed model ensures prolonged battery life for IoT wearable edge devices while maintaining a very high sensitivity to even difficult beat classes such as VEB.

APPENDIX A
ECG DATABASE

In this work, we use the ECG beat classification standard notations as defined by AAMI [31] as in Table VIII. The beat classes supra ventricular ectopic beats (SVEB—S), ventricular ectopic beats (VEB—V), Fusion Beat (F), and Unclassified Beat (Q) are grouped as "Abnormal"—"A" and all other beats are considered "Normal"—"N" beats.

We used the MIT-BIH arrhythmia database [32] and the St. Petersburg Institute of Cardiological Technics (INCART) database [29] for evaluation and testing purposes. These data sets are freely available on Physionet [33] and the important details along with the globally accepted testing standards are summarized as follows.

A. MIT-BIH Arrhythmia Database

The database consists of 360 Hz sampled, 48 two-lead recordings of approximately 30 min, and more than 100 thousand individual heartbeats obtained from 47 subjects studied by the BIH Arrhythmia Laboratory between 1975 and 1979. The first 23 (100–124) recordings were extracted from a set of 4000, 24-h routine ambulatory ECG recordings collected from a mixed population of both in and outpatients at Boston's Beth Israel Hospital, while the remaining 25 (200–234) were selected because of the presence of less common but complex and clinically significant ventricular, junctional, and supraventricular arrhythmias. Both leads are not the same in all recordings, depending on the arrhythmia and physical limitation of the subject's body condition, and mostly modified lead II (MLII) is present in the first lead. The second lead varies a lot with V1 and V5 in most records whilst V2, V4, and MLII lead in very few instances. The annotations provided with the database were used for evaluation purposes, following the recommendations and class labeling of AAMI. The four recordings with paced beats were discarded in this work in accordance with AAMI recommendations [31]. The AAMI Q class (unclassified heartbeats) was considered under arrhythmia class since it will be sent to the cloud for further study according to the proposed architecture.

B. INCART Database

St. Petersburg Institute of Cardiological Technics (INCART) database consists of 75 records sampled at 257 Hz with 12 ECG leads and corresponding annotations. Each record is approximately 30-min long and contains over 175-thousand individual beats. The original records were collected from 17 males and 15 females, aged between 18–80 years, undergoing tests for coronary artery disease. None of the patients had pacemakers and most patients had VEBs. Preference was given to subjects whose ECG readings were consistent with ischemia, coronary artery disease, conduction abnormalities,

and arrhythmias while selecting records to be included in the database. The annotations were produced by an automatic algorithm that generally places beat annotations in the middle of the QRS complex (as determined from all 12 leads) and then corrected manually, following the standard PhysioBank beat annotation definitions. However, the locations have not been manually corrected, and consequently, there are occasional misaligned annotations.

APPENDIX B
FEATURE DEFINITIONS

A. Singleton Features

1) Peak Locations: The location (in time) of the R peak and the P, QRS, and T waveform onset times are the most elementary features that are the basis of many other elaborate features below. A complete implementation would need to estimate these features using existing algorithms [34], [35], [36]. In this work, we assume that *R*-peak detection is done on-chip by the data acquisition analog-front-end circuit, as in the case of commercially available solutions [37]. During the training process, these features are directly obtained from the database annotations. Since the annotations in the records are heavily distorted around the peak, we apply a search window to correct the peak locations.

2) Corrected QT Intervals: The QT interval, for the *i*th beat, QT_i, is defined as the time duration between the *Q* onset and *T* offset for that beat. To compensate for QT variability with respect to heart rate a commonly used measure is the so called "corrected QT interval" [38] which, for the *i*th beat, is defined as

$$\mathrm{QT}c_i = \frac{\mathrm{QT}_i}{\sqrt{\mathrm{RR}_i}} \tag{6}$$

where RR_i is the RR interval between the *i*th and $(i-1)$th beats. Despite first impressions, $\mathrm{QT}c_i$ actually has units of seconds, as explained in [39], and can be understood as that QT interval we would expect this individual to have if their heart rate were exactly 60 beats per minute.

3) HRV Metrics: Heart rate variability (HRV) analyzes the physiological phenomenon is commonly used with the ones initially considered in this article summarized in Table IX. Based on our experimentation in Section IV-A, the HRV measures we found to be of most interest are RRIndex and SD1 and SD2 which are now defined in more detail (for the *i*th beat)

$$\mathrm{RR}_{\mathrm{Index}_i} = 2\frac{\mathrm{RR}_i - \mathrm{RR}_{i-1}}{\mathrm{RR}_i + \mathrm{RR}_{i-1}} \tag{7}$$

$$\mathrm{SD1}_i = \sqrt{\frac{1}{2}\sum\left(\boldsymbol{XR}_i(1,:) - \overline{\boldsymbol{XR}_i(1,:)}\right)^2} \tag{8}$$

$$\mathrm{SD2}_i = \sqrt{\frac{1}{2}\sum\left(\boldsymbol{XR}_i(2,:) - \overline{\boldsymbol{XR}_i(2,:)}\right)^2} \tag{9}$$

where $\boldsymbol{XR}_i$ is 2×3 matrix formed as a 45° rotation of three consecutive points of a Poincaré plot defined as follows:

$$\boldsymbol{XR}_i = \begin{bmatrix} \sin\left(-\frac{\pi}{4}\right) & \cos\left(-\frac{\pi}{4}\right) \\ \cos\left(-\frac{\pi}{4}\right) & -\sin\left(-\frac{\pi}{4}\right) \end{bmatrix} \begin{bmatrix} \mathrm{RR}_{i-2} & \mathrm{RR}_{i-1} & \mathrm{RR}_i \\ \mathrm{RR}_{i-1} & \mathrm{RR}_i & \mathrm{RR}_{i+1} \end{bmatrix} \tag{10}$$

TABLE IX
SELECTED TIME-DOMAIN MEASURE OF HRV

Variable	Description
SDNN	Standard deviation of all RR intervals (RR_{SDNN})
RMSSD	The square root of the mean sum of the squares of differences between adjacent RR intervals
pNN50	Number of pairs of adjacent RR intervals differing by more than 50 ms divided by the total number of all RR intervals
RRIndex	Relative variation of consecutive RR intervals with distance one, which is usually between -20% and +20%
SD2	Standard deviation of the Poincaré plot along the line-of-identity
TRI	Total number of all RR intervals divided by the height of the histogram of all RR intervals measured on a discrete scale with bins of 1/128 sec.
MED	The median distance to the center-point of return map of relative RR intervals
TINN	Baseline width of the minimum square difference triangular interpolation of the highest peak of the histogram of all RR intervals
SD1	Standard deviation of Poincaré plot perpendicular to the line-of-identity
ApEntropy	Measure of complexity using sliding window correlation

and $\boldsymbol{XR}_i(k,:)$ is the *k*th row in $\boldsymbol{XR}_i$. The summations and averages in the definition of $\mathrm{SD1}_i$ and $\mathrm{SD2}_i$ are over all three elements in their respective rows.

We also introduce a new measure called weighted SDNN (wSDNN) by modifying the SDNN metric with a weighted vector as follows

$$\mathrm{RR}_{\mathrm{wSDNN}_i} = \sqrt{\frac{1}{10}\sum_{j=-9}^{+1} W_j\left(\mathrm{RR}_{i+j} - \overline{\mathrm{RR}_i}\right)^2} \tag{11}$$

where

$$W_j = \begin{cases} 10, & j = 0 \\ 1, & \text{otherwise} \end{cases}$$

and $\overline{\mathrm{RR}_i}$ is the average of the *RR* intervals over the same window.

4) Principal Component Analysis: We computed the principal component waveforms based on all the normal beats in the MIT database (i.e., across all individuals). Subsequently, each beat can be decomposed into a weighted sum of these PCA waveforms with PCA coefficients as weights. In this work, we focus on the first PCA coefficient of each beat which we denote as PCA_i. An example is shown in Fig. 13. Note that prior to computation of the PCA components, the ECG signal is normalized so that each beat[5] has zero mean and unit standard deviation. These normalized ECG signals will hereafter be referred to as *x*.

We have also introduced a new measure called "σ(PCA)": standard deviation of PCA value of ten consecutive beats defined as

$$\sigma(\mathrm{PCA})_i = \sqrt{\frac{1}{9}\sum_{j=-9}^{0}\left(\mathrm{PCA}_{i+j} - \overline{\mathrm{PCA}_i}\right)^2} \tag{12}$$

and $\overline{\mathrm{PCA}_i}$ is the average of the *PCA* values over the same window.

[5]At a sampling rate of 250 Hz, a beat is taken as 62 samples before and 112 samples after the *R*-peak.

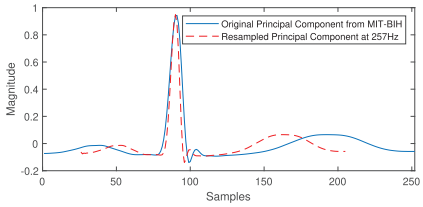

Fig. 13. Principal component (beat) for MIT-BIH and INCART database derived from MIT-BIH samples at 360 Hz and resampled at 257 Hz.

5) Other Features: In addition to the aforementioned features, we have also analysed,

1) Teager energy operator (TEO) defined as:

$$\text{TEO}_i = \sum_n \Big(x[n]^2 - x[n-1]x[n+1] \Big) \qquad (13)$$

where $x[n]$ is the normalized ECG signal defined in Section B-A4 and the summation spans a window of ± 0.25 s (or equivalently ± 62 samples) about the ith peak;
2) $\sigma(\text{TEO})_i$;
3) VS_i: Difference in peak to peak amplitude between the window of ± 0.056 s (or equivalently ± 14 samples) about the ith peak;
4) $\sigma(VS)_i$;
5) ΣQRS_i: The sum of amplitudes in a 0.1s window about the R peak of the ith beat;
6) $|\Sigma\text{QRS}_i|$: Magnitude of ΣQRS_i;
7) $\Sigma\text{QRS}_{\text{Energy}_i}$: Sum of amplitudes squared in a 0.1s window about the R peak of the ith beat;
8) DTW_i Dynamic Time Warping cost with respect to the principal normal beat (Fig. 13);
9) $\sigma(DTW)_i$;
10) $\text{QRS}_{\text{Sign}_i}$ which is "1" for a rising QRS peak and "0" for an inverted QRS peak.

and are used for testing. Refer to Table III for some of the identified features and their associated rules performance in the MIT-BIH arrhythmia database.

B. Compound Features

Denoting all the previous features as *singleton* features, we define a compound feature as being the weighted sum of these singleton features where the weights are either 0 (i.e., that singleton feature is absent from the compound feature) or ± 1. If there are N singleton features, then there are 3^N possible compound features. All compound features are not computed continually, as during the training phase, the reduced set of singleton and compound features required to implement the chain of rules is identified and only these ones are actually computed at model runtime.

REFERENCES

[1] W. Ning et al., "Automatic detection of congestive heart failure based on a hybrid deep learning algorithm in the Internet of Medical Things," *IEEE Internet Things J.*, vol. 8, no. 16, pp. 12550–12558, Aug. 2021.
[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] A. A. Cook, G. Mısırlı, and Z. Fan, "Anomaly detection for IoT time-series data: A survey," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 6481–6494, Jul. 2020.
[4] 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.
[5] 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.
[6] 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.
[7] L. Sun, Y. Wang, Z. Qu, and N. N. Xiong, "BeatClass: A sustainable ECG classification system in IoT-based eHealth," *IEEE Internet Things J.*, vol. 9, no. 10, pp. 7178–7195, May 2022.
[8] X. Yin and J. Han, "CPAR: Classification based on predictive association rules," in *Proc. SIAM Int. Conf. Data Min.*, 2003, pp. 331–335.
[9] J. R. Quinlan and R. M. Cameron-Jones, "FOIL: A midterm report," in *Proc. Eur. Conf. Mach. Learn.*, 1993, pp. 1–20.
[10] B. Liu, W. Hsu, and Y. Ma, "Integrating classification and association rule mining," in *Proc. KDD*, vol. 98, 1998, pp. 80–86.
[11] W. Li, J. Han, and J. Pei, "CMAR: Accurate and efficient classification based on multiple class-association rules," in *Proc. IEEE Int. Conf. Data Min.*, 2001, pp. 369–376.
[12] F. Coenen, *The LUCS-KDD Group*, Dept. Comput. Sci., Univ. Liverpool, Liverpool, U.K, 2004.
[13] P. Pudil et al., "Floating search methods in feature selection," *Pattern Recognit. Lett.*, vol. 15, no. 11, pp. 1119–1125, 1994.
[14] M. G. Tsipouras, D. I. Fotiadis, and D. Sideris, "An arrhythmia classification system based on the RR-interval signal," *Artif. Intell. Med.*, vol. 33, no. 3, pp. 237–250, 2005.
[15] T. P. Exarchos, C. Papaloukas, D. I. Fotiadis, and L. K. Michalis, "An association rule mining-based methodology for automated detection of ischemic ECG beats," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 8, pp. 1531–1540, Aug. 2006.
[16] T. P. Exarchos et al., "A methodology for the automated creation of fuzzy expert systems for ischaemic and arrhythmic beat classification based on a set of rules obtained by a decision tree," *Artif. Intell. Med.*, vol. 40, no. 3, pp. 187–200, 2007.
[17] M. Llamedo and J. P. Martínez, "Heartbeat classification using feature selection driven by database generalization criteria," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 3, pp. 616–625, Mar. 2011.
[18] J. He, L. Sun, J. Rong, H. Wang, and Y. Zhang, "A pyramid-like model for heartbeat classification from ECG recordings," *PLoS One*, vol. 13, Nov. 2018, Art. no. 206593.
[19] Y. Rubner, L. J. Guibas, and C. Tomasi, "The earth mover's distance, multi-dimensional scaling, and color-based image retrieval," in *Proc. ARPA Image Understanding Workshop*, vol. 661, 1997, p. 668.
[20] J. B. Bidias à Mougoufan, J. S. A. E. Fouda, M. Tchuente, and W. Koepf, "Three-class ECG beat classification by ordinal entropies," *Biomed. Signal Process. Control*, vol. 67, May 2021, Art. no. 102506.
[21] C. Bandt and B. Pompe, "Permutation entropy: A natural complexity measure for time series," *Phys. Rev. Lett.*, vol. 88, no. 17, 2002, Art. no. 174102.
[22] J. A. E. Fouda, W. Koepf, and S. Jacquir, "The ordinal Kolmogorov-Sinai entropy: A generalized approximation," *Commun. Nonlinear Sci. Numer. Simul.*, vol. 46, pp. 103–115, May 2017.
[23] F. Yang, G. Wang, C. Luo, and Z. Ding, "Improving automatic detection of ECG abnormality with less manual annotations using siamese network," in *Proc. 43rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, 2021, pp. 1120–1123.
[24] R. Zhao et al., "A novel intrusion detection method based on lightweight neural network for Internet of Things," *IEEE Internet Things J.*, vol. 9, no. 12, pp. 9960–9972, Jun. 2022.
[25] C. Molnar, *Interpretable Machine Learning*, 2nd ed. Munich, Germany: Christoph Molnar, 2022. [Online]. Available: https://christophm.github.io/interpretable-ml-book
[26] G. Riley and J. C. Giarratano, *Expert Systems: Principles and Programming*. Thomson Course Technol., Stamford, CT, USA, 2005.
[27] R. Goebel et al., "Explainable AI: The new 42?" in *Proc. Int. Cross-Domain Conf. Mach. Learn. Knowl. Extraction*, 2018, pp. 295–303.
[28] C. J. Van Rijsbergen, *The Geometry of Information Retrieval*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[29] "St Petersburg INCART 12-lead arrhythmia database." May 2008. [Online]. Available: https://physionet.org/content/incartdb/1.0.0/

[30] E. J. D. S. Luz, W. R. Schwartz, G. Cámara-Chávez, and D. Menotti, "ECG-based heartbeat classification for arrhythmia detection: A survey," *Comput. Methods Programs Biomed.*, vol. 127, pp. 144–164, Apr. 2016.
[31] *ANSI/AAMI EC57, Testing and Reporting Performance Results of Cardiac Rhythm and ST Segment Measurement Algorithms*, Assoc. Adv. Med. Instrum., Arlington, VA, 1998.
[32] "MIT-BIH arrhythmia database." Feb. 2005. [Online]. Available: https://physionet.org/content/mitdb/1.0.0/
[33] A. L. Goldberger et al., "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.
[34] C. Nayak, S. K. Saha, R. Kar, and D. Mandal, "An efficient and robust digital fractional order differentiator based ECG pre-processor design for QRS detection," *IEEE Trans. Biomed. Circuits Syst.*, vol. 13, no. 4, pp. 682–696, Aug. 2019.
[35] 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. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9178390
[36] A. John, S. J. Redmond, B. Cardiff, and D. John, "A multimodal data fusion technique for heartbeat detection in wearable IoT sensors," *IEEE Internet Things J.*, vol. 9, no. 3, pp. 2071–2082, Feb. 2022.
[37] "Ultra-low power, single-channel integrated biopotential (ECG, R to R detection) AFE," Data Sheet MAX30003, Analog Devices, Wilmington, MA, USA, 2021. [Online]. Available: https://www.analog.com/media/en/technical-documentation/data-sheets/max30003.pdf
[38] N. Bekken, *ECG Interpretation Made Incredibly Easy*, Lippincott Williams Wilkins, Philadelphia, PA, USA, 2005.
[39] J. Molnar, J. S. Weiss, and J. E. Rosenthal, "The missing second: What is the correct unit for the Bazett corrected QT interval?" *Amer. J. Cardiol.*, vol. 75, no. 7, pp. 537–538, 1995.

Gawsalyan Sivapalan received the Bachelor of Science degree (Hons.) in electronics and telecommunication engineering from the University of Moratuwa, Moratuwa, Sri Lanka, in 2016, and the master's degree (by research) from the School of Electrical and Electronic Engineering, University College Dublin, Dublin, Ireland, in 2021.

He was a Research Intern with the Singapore University of Technology and Design, Singapore, from October 2014 to March 2015. He co-founded Kairos Sensing, a university spin-off providing wearable-based high-performance capturing solutions to elite clients like Sri Lanka National Cricket from 2016 to 2017. From 2017 to 2018, he worked as a Senior Executive in Corporate Planning and Strategy Development with Dialog Axiata PLC, a super brand and a blue-chip telecommunication arm of Axiata Group Berhad, Colombo, Sri Lanka. He is currently working with Meta Reality Lab, Redmond, WA, USA, as a Characterization and Embedded Engineer. His research interests are in motion capture and machine-learning solutions for wearable devices.

Mr. Sivapalan was the winner of Disrupt 2.0—FutureX Technopreneur Challenge in 2016. He has been a Qualified Member of the Chartered Institute of Management Accountants, U.K., since 2016.

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/Systems Architect with Nokia, U.K. from 1993 to 2001, moving to Silicon and Software Systems (S3 Group) 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 are in digitally assisted circuit design and signal processing for wireless and optical communication systems.

Koushik Kumar Nundy (Senior Member, IEEE) received the Bachelor of Technology degree in electronics and communication engineering from the National Institute of Technology Durgapur, Durgapur, India, in 2010, and the Master of Science degree in telecommunication from The Hong Kong University of Science and Technology, Hong Kong, in 2011.

He received the NTSE Scholarship in 2004 and the NGS Scholarship in 2013. He has worked with multiple research and academic institutes, including the National University of Singapore, Singapore, and the Indian Institute of Science, Bengaluru, India, and has presented his work in multiple conferences and journals. He has been featured in publications, such as the Irish Times, Rochester Business Journal, and Silicon Republic. He has also worked with the Research and Development Team, Altai Technologies Ltd., Hong Kong, and is the Co-Founder of Think Biosolution Limited, Dublin, Ireland.

Dr. Nundy has won multiple awards, including the Roche Unicorn Champion 2020, the Innovation of the Year 2017, and the Start-Up Weekend 2015. He is the Ireland Area Chair for the IEEE Young Professionals.

Alex James (Member, IEEE) received the Ph.D. degree from Griffith University, Brisbane, QLD, Australia, in 2009.

He is currently a Professor and the Dean (Academic) of the Kerala University of Digital Sciences, Innovation and Technology (also known as Digital University Kerala), Thiruvananthapuram, India. He is the Professor-in-Charge of the Maker Village, and the Chief Investigator with the India Innovation Centre for Graphene, Thrissur, India. His research interests include AI—neuromorphic systems (software and hardware), VLSI, and image processing.

Dr. James was awarded the IEEE Outstanding Researcher by IEEE Kerala Section for 2022, the Kairali Scientist Award (Kairali Gaveshana Puraskaram) for Physical Science in 2021, and the Best Associate Editor for IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—PART I: REGULAR PAPERS in 2021. He was the Founding Chair of IEEE CASS Kerala Chapter and a member of IET Vision and Imaging Network, and currently a member of BCS' Fellows Technical Advisory Group. He was an Editorial Board Member of *Information Fusion* from 2010 to 2014, and has been serving as an Associate Editor for IEEE ACCESS since 2017, *Frontiers in Neuroscience* since 2022, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEM—PART I: REGULAR PAPERS from 2018 to 2023, and IEEE OPEN JOURNAL OF CIRCUITS AND SYSTEMS since 2022. He is a member of CASS TC on Nonlinear Circuits and Systems, IEEE CTSoc TC on Quantum in Consumer Technology, TC on Machine Learning, Deep Learning, and AI in CE, IEEE CASS TC on Cellular Nanoscale Networks and Memristor Array Computing, and IEEE CASS SIG on AgriElectronics. He is a Life Member of ACM, a Senior Fellow of HEA, and a Fellow of British Computer Society and IET.

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 National University Singapore, Singapore, in 2008 and 2014, respectively.

He is currently an Assistant Professor with the School of Electrical and Electronics Engineering, University College Dublin, Dublin, Ireland. He was a Postdoctoral Researcher with the Bio-Electronics Lab, National University Singapore from 2014 to 2017. His research interests include low-power biomedical circuit design, energy-efficient signal processing, and edge computing.

Dr. John is a recipient of the Institution of Engineers Singapore Prestigious Engineering Achievement Award in 2011, the Best Design Award at the Asian Solid-State Circuit Conference in 2013, the IEEE Young Professionals, and the Region 10 Individual Award in 2013. He served as a member of the Organizing Committee/Technical Program Committee for several IEEE conferences, including ISCAS, BioCAS, ICECS, TENCON, ASICON, and ICTA. He served as a Guest Editor for IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—PART I: REGULAR PAPERS and IEEE OPEN JOURNAL OF CIRCUITS AND SYSTEMS. He currently serves as an Associate Editor for IEEE TRANSACTIONS ON BIOMEDICAL CIRCUITS AND SYSTEMS, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—PART II: EXPRESS BRIEFS, and *International Journal of Circuit Theory & Applications* (Wiley).