A Low-Power Co-Processor to Predict Ventricular Arrhythmia for Wearable Healthcare Devices

Meenali Janveja, Rushik Parmar, Srichandan Dash, Jan Pidanic, *Senior Member, IEEE*, and Gaurav Trivedi, *Member, IEEE*

***Abstract*— Ventricular arrhythmia (VA) is the most critical cardiac anomaly among all arrhythmia beats. Thus, it becomes imperative to predict the occurrence of VA to avoid sudden casualties caused by these arrhythmia beats. In the past, only a few hardware designs have been proposed to predict VA using various features derived from electrocardiogram (ECG) signals and processed using machine learning classifiers. However, these designs are either complex or need more prediction accuracy. Therefore, a deep neural network (DNN)-based co-processor for arrhythmia prediction is proposed in this article. It can predict VA at least 15 min before its occurrence with 91.6% accuracy. Co-processor architecture for arrhythmia prediction (CoAP) uses an optimal feature vector extracted from the ECG signal and an optimized DNN, using a novel approximate multiplier (AM). CoAP operates at 12.5 kHz and consumes 4.69 μW when implemented using SCL 180-nm bulk CMOS technology. The low power realization of the proposed design and its higher accuracy, compared with well-known state-of-the-art methods, make it suitable for wearable devices.**

***Index Terms*— Arrhythmia, application specific integrated circuit (ASIC), deep neural network (DNN), electrocardiogram (ECG), low power.**

I. Introduction

IN THE medical field, wearable devices can connect doctors, patients, and other parties to understand to understand changes in conditions, alleviate pain, treat diseases, and facilitate the collection of a large sample of case data. This information is helpful for the development of national epidemiology strategies and preventive medicine. Wearable devices are used mainly in health and safety monitoring, chronic disease management, disease diagnosis and treatment, and rehabilitation. With the rise of mobile medicine, the development of new technologies such as smart sensing and the popularization of personalized health concepts have developed rapidly in recent years. Among them, medical wearable devices have become one of the most promising fields. With wearable technology, medical professionals can monitor and evaluate patients' vitals, which can help doctors better manage patients' health and safety. It also aids clinicians in monitoring the compliance of medicines and treatments. Healthcare has become more efficient and convenient due to the advancements in wearable technology. It also enables doctors to connect with their patients via telehealth solutions and allows them to engage 24 $\times$ 7. Thus, wearable devices in healthcare offer numerous benefits to patients and healthcare providers. However, there are several hurdles, such as power consumption, portability, patient safety, secure data delivery, and integration which inhibit making wearable devices from a concept to a profitable product.

Cardiovascular diseases (CVDs) are a major cause of mortality globally, accounting for about 17 million casualties annually [33]. Among the chronic CVDs, malignant ventricular arrhythmia (VA)—specifically ventricular tachycardia (VT) and ventricular fibrillation (VF)—are highly prevalent and responsible for sudden casualties. A VT is characterized by a rhythm of more than three consecutive beats originating from the ventricles at a rate exceeding 100 beats/min [1], while VF is a chaotic activation of ventricles, leading to immediate cessation of blood circulation and degeneration into a pulseless or flat ECG signal, indicating no cardiac electrical activity. VT/VF that lasts more than 30 s can result in hemodynamic compromise and sudden cardiac fatality [2], underscoring the criticality of predicting these arrhythmic beats before they occur.

Wearable healthcare devices incorporating continuous ECG monitoring and cardiovascular disease detection are at the forefront of medical consumer electronics. Wearable ECG devices' ability to continuously monitor ECG, along with machine learning algorithms, can aid in the timely diagnosis of VA events to reduce fatalities. Incorporating prediction models for VT/VF into wearable healthcare devices represents a significant leap forward in managing CVDs. This article proposes a co-processor architecture for arrhythmia prediction (CoAP) to overcome the shortcomings of the state-of-the-art methods reported in the literature. The accurate VT/VF prediction is imperative, and it signifies the need to develop a co-processor for arrhythmia prediction. The design of CoAP

Manuscript received 17 December 2023; revised 17 May 2024; accepted 2 June 2024. This work was supported in part by the Inter-Excellence Funds of the Ministry of Education, Youth and Sports, Czech Republic, "Artificial Intelligence Enabled Smart Contactless Technology Development for Smart Fencing" under Project LTAIN19100; and in part by the Electronics and ICT Academy, Indian Institute of Technology Guwahati, India. *(Corresponding author: Meenali Janveja.)*

Meenali Janveja is with the Department of Electronics and Electrical Engineering, Indian Institute of Technology at Guwahati, Guwahati, Assam 781039, India (e-mail: meena176102001@alumni.iitg.ac.in).

Rushik Parmar, Srichandan Dash, and Gaurav Trivedi are with the Department of Electronics and Electrical Engineering, Indian Institute of Technology at Guwahati, Guwahati, Assam 781039, India (e-mail: rushik_parmar@iitg.ac.in; srichand@iitg.ac.in; trivedi@iitg.ac.in).

Jan Pidanic is with the Department of Electrical Engineering, University of Pardubice, 532 10 Pardubice, Czech Republic (e-mail: jan.pidanic@upce.cz).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TVLSI.2024.3413584.

Digital Object Identifier 10.1109/TVLSI.2024.3413584

TABLE I
COMPARISON OF STATE-OF-THE-ART METHODS FOR ARRHYTHMIA CLASSIFICATION

Paper	Approach (Detection/Prediction)	Model Used	Implementation Platform	Number of Classes	Accuracy	Limitations
[3]	Detection	CNN	Software	9	NS	a) Classification is class based b) Cannot predict arrhythmia
[4]	Detection	NEO-CNN	Microcontroller	5	97.83	a) Classification is class based b) Cannot predict arrhythmia c)No ASIC implementation
[6]	Detection	ANN	ASIC	5	98	a) Classification is patient dependent b) Cannot predict arrhythmia
[7]	Detection	Clinicians Based	ASIC	5	99.8	a) Classification is class based b) Cannot predict arrhythmia
[8]	Detection	TNN	ASIC	13	99.3	a) Classification is class based b) Cannot predict arrhythmia
[12]	Detection	NS	ASIC+Smartphone	2	95.8	a) Classification is class based b) No ASIC implementation
[14]	Detection	DNN	ASIC	5	91.6	a) Cannot predict arrhythmia
[15]	Detection	DNN	ASIC	5	98	a) Classification is patient dependent b) Cannot predict arrhythmia
[17]	Prediction	Naive-Bayes	ASIC	2	86	a) Low prediction accuracy
[18]	Prediction	Decision Tree	Software	2	NS	a) Complex ECG feature set b) Not suitable for hardware implementation

proposed in this article uses a deep neural network (DNN) classifier, which is trained to perform the prediction of VA rather than detecting it.

The proposed design uses an optimal feature extraction algorithm and DNN-based classifier which reuses the hardware resources and while being capable of operating on low frequency. This maintains high performance in predicting VT/VF compared with the contemporary methods making our design suitable for wearable or consumer electronic healthcare devices. Incorporating our design as a co-processor in typical healthcare devices can save lives by predicting potentially fatal arrhythmic events before they occur.

The rest of this article is organized as follows. Section III describes the arrhythmia co-processor in detail. The results of the proposed CoAP are discussed in Section IV. The proposed future work is presented in Section VI. Finally, the article is concluded in Section V.

II. PREVIOUS WORKS

A wide range of methods using machine and deep learning techniques have been developed to detect VA using ECG. An ECG-based consumer electronic device in the healthcare domain to detect arrhythmia is proposed [9]. Furthermore, Chen et al. [3] and Sabor et al. [4] present the algorithms for detecting five types of arrhythmia which were later implemented using microcontrollers. Although [3] and [4] exhibit high accuracy, but use computationally complex convolutional neural networks, which are unsuitable for application specific integrated circuit (ASIC) implementation. Furthermore, Cai et al. [6], Xu et al. [7], Abubakar et al. [8], Janveja et al. [11], and Zhao et al. [12] use different types of DNNs for the classification of different types of arrhythmia. These designs have low area and power requirements but can only detect the occurrence of arrhythmia and do not predict it.

Furthermore, Sammani et al. [32] and Kolk et al. [34] present a detailed survey of the latest techniques used for detection and prediction of cardiac arrhythmia The prediction of VA, particularly VF, is a critical task as it can lead to mortality within a few minutes of its occurrence. Therefore, accurately predicting the onset of VT/VF beats is of paramount importance in avoiding casualties. The main focus of this study is to predict arrhythmia rather than detect it. However, the literature reveals only a few methods that can predict the occurrence of arrhythmic beats before they happen. Bayasi et al. [14] depict a low-power ASIC using Naive Bayes, but its accuracy for predicting arrhythmia is only 86%. Mandala et al. [15] illustrate an algorithm that can predict the VT/VF 15 min before its occurrence using decision trees. These two methods have advantages in discriminating between normal and abnormal conditions. However, [14] has low accuracy in predicting VT/VF, and [15] needs to be simplified for the hardware implementation. Furthermore, the method reported in [16] predicts the occurrence of arrhythmia. Nevertheless, it uses the data from the occurrence of the arrhythmia rather than taking ECG beats before the occurrence of VT for classification, which identifies it as a detection methodology instead of a prediction. Table I presents the comparison of the state-of-the-art methodologies mentioned above, along with their limitations.

III. PROPOSED WORK

A typical wearable system for ECG processing has two significant operation blocks, as shown in Fig. 1.

The first is the analog front-end responsible for acquiring the ECG signal from the on-body sensors and converting them into digital samples. The second block is the co-processor block responsible for processing the acquired data and further detecting the abnormalities present in the ECG signal.

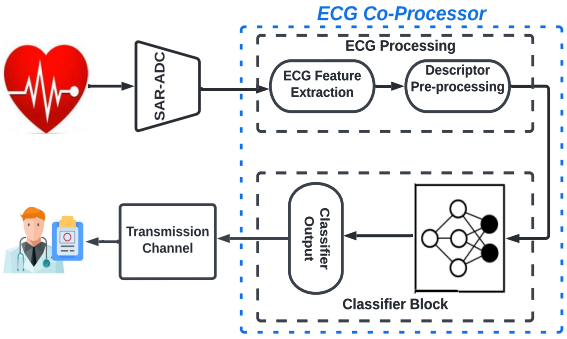


Fig. 1. Typical ECG processing system of a wearable device.

The conventional systems acquire the ECG signals from the analog front-end and transmit them over a channel for further analysis. Whereas CoAP is an integrated co-processor that acquires ECG signals from the sensors, processes them, and performs ON-chip classification without any external interaction, as shown in Fig. 1. Such design methodology immediately alerts an individual, which is essential in critical conditions. Moreover, processing and classifying the data locally on a chip reduces the amount of data transmitted through a data channel, optimizing power consumption.

A. Brief Description of CoAP

CoAP consists of three primary blocks: the ECG feature extraction, descriptor processing, and classifier block, as shown in Fig. 1. First, the ECG feature extraction block is responsible for extracting the fiducial points of ECG. The raw ECG signal obtained from the ADC is first filtered using a bandpass filter of band 0.5–40 Hz. In the proposed work, filtered ECG excerpts of 3-s windows are considered for processing at a time. Later, the $P-QRS-T$ waves of ECG signals are extracted using a novel ECG feature extraction algorithm, which is also described in our previous work [17]. The second block is the descriptor processor block, which is responsible for extracting the critical information from the variation in ECG waves. In the proposed work, the features required as an input feature vector for the classifier block are termed as "descriptors."

Researchers have conducted many studies to discover various ECG features that could reflect the underlying VT/VF [18]. Some ECG descriptors have predictive values for arrhythmic events, such as heart rate variability (HRV), T-peak T-end (TpTe), and QT wave intervals. Mandala et al. [15] have used descriptors, such as the mean of QRS intervals, to detect VT/VF. They have also used standard and root-mean-square deviations of Q, S, and R waves to predict the occurrence of VT/VF using decision trees on the software platform. However, the complex square root and division operations required to estimate these descriptors might significantly add to the hardware utilization of the co-processor. Furthermore, Bayasi et al. [14] used seven unique sets of interval-based descriptors to predict the VT/VF, which does not require any complex mathematical operations. However, to estimate these intervals, both peaks as well as onset and offsets of $P-QRS-T$ waves need to be delineated. Besides, the accuracy obtained after using these intervals as descriptors is only 86%, which is less than [15]. Therefore, we propose a novel descriptor set that balances the tradeoff between accuracy and hardware resource utilization in this work. This is crucial for enhancing the robustness and acceptability of the proposed method for wearable healthcare devices.

These descriptors are fed as input feature vectors to the final classifier block of the ECG co-processor. In the proposed work, a DNN is used as a classifier to predict the excerpts susceptible to VA. The novelties of the proposed design are summarized as follows.

1) *Prediction Over Detection:* The co-processor designed in this study emphasizes prediction over detection. It can predict the occurrence of VT/VF with significantly higher accuracy than the current state-of-the-art ASIC implementations up to 15 min before the onset of these conditions.
2) *Novel Descriptor Set:* The classifier block of the co-processor for arrhythmia prediction uses a unique descriptor set, which is optimized to achieve maximum accuracy with low power and area requirements. Better accuracy is achieved using a novel feature set with clinical relevance that eliminates the need to extract the boundaries of $P-QRS-T$ waves, which is critical for accurate feature extraction. In addition, the estimation of mean and variance is optimized using simple shift operations, further enhancing the efficiency of the hardware.
3) *Optimized Hardware Implementation of Activation Function:* Traditionally, a DNN for two-class classification uses only one output node with a sigmoid activation function. However, our work introduces a novel approach using two output nodes, each with a sigmoid activation function in the output layer for training. This approach assists in simplifying the hardware implementation of the sigmoid function after training using only a simple comparator.
4) *Optimized Multiplier:* We propose an optimized multiplier design that uses operand decomposition and Mitchell's algorithm to enable low-power implementation of pretrained DNN on hardware. Our design maintains high accuracy while reducing power consumption, making it a promising solution for efficient DNN hardware implementation.

Sections III-B–III-D presents the architecture of CoAP.

B. ECG Feature Extraction

The architecture of CoAP is broadly divided into two blocks which are controlled using the main control unit (MCU). Furthermore, each layer of the DNN is controlled by the DNN control unit, as shown in Fig. 2. An electrocardiogram (ECG) excerpt of 3 s, with each sample represented in 16-bit 2's complement format sampled at a frequency of 250 Hz, is considered at a particular time for processing. An ECG is a bio-signal, which is used by medical practitioners all over the world to monitor electrical activity of the heart. A human heart contains four chambers; two atria and two ventricles, whose rhythmic depolarization and repolarization

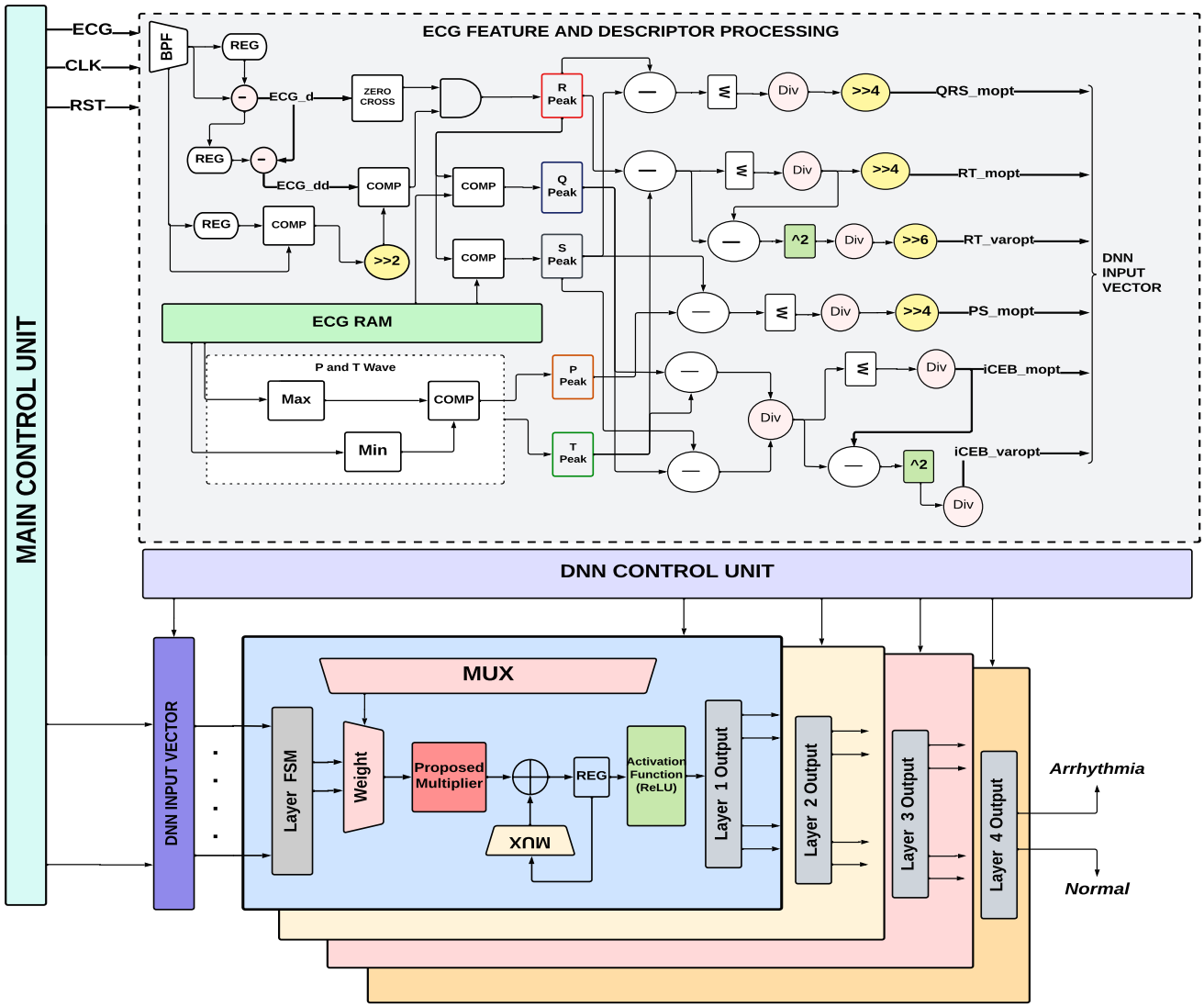


Fig. 2. Architecture of the proposed co-processor.

is captured with the help of ECG. As we know, an ECG has three primary features, namely, P wave, QRS complex, and T wave, as shown in Fig. 3. Atrial depolarization and ventricular repolarization is represented by a P wave and a T wave, respectively. The QRS wave characterizes rapid depolarization of the ventricles. Due to the large muscle mass of ventricles when compared with atria, QRS complex has a larger amplitude than the P wave. Any variation in the standard values of ECG features and derived intervals aids cardiologists to diagnose cardiac abnormalities. Hence, accurate and efficient algorithms are required that can process the acquired ECG signal to extract fiducial points. The features extracted from an ECG are used individually or in a group for different analysis and classification of heart abnormalities.

Initially, $P-QRS-T$ peaks are estimated using our previously described algorithm [17]. However, architecture for only peak detection modules is used in this work as shown in Fig. 2. The feature extraction algorithm is validated using the ECG excerpts from several datasets available on Physionet [13]. The ECG feature extraction algorithm exhibits average accuracy of 98.36%, 99.86%, and 99.16% accuracy in delineating P, T, and QRS waves, respectively, which is more or equivalent to the algorithm proposed in the literature. Moreover, the proposed ECG feature extraction design is computationally simple and optimized compared with other state-of-the-art methods making it resource- and power-efficient. Once the $P-QRS-T$ peaks are delineated for a 3-s window, these peaks are input to the descriptor processing block.

C. Descriptor Processing and Optimization

After a rigorous study of different ECG characteristics associated with VT/VF, we developed a novel set of six descriptors, which considerably help in the prediction of VT/VF. These descriptors are the mean and variance of RT interval (RT_m), mean of PS intervals (PS_m), mean of the duration of QRS complexes, and mean and variance of the index of cardiac electrophysiological balance (iCEB_m and iCEB_{var}). The definition of these features is given in Table II. Conventionally, RT interval, PS interval, QRS duration, and QT intervals are the difference between the boundaries of the respective $P-QRS-T$ waves. For instance, RT is estimated as the

TABLE II
DETAILS OF OPTIMIZED DESCRIPTORS

Descriptor		Details	Original Definition	Optimized Definition
QRS Duration	mean	QRS complex duration represents ventricular depolarization.	$QRS_m = \frac{\sum(QRS_{off} - QRS_{on})}{n}$	$QRS_{m_{opt}} = \frac{1}{16}(\frac{\sum(S_{peak} - Q_{peak})}{n})$
RT Interval	mean	It depicts the time from the beginning of ventricular depolarization to the end of ventricular repolarization. It includes all of the electrical events that take place in the ventricles.	$RT_m = \frac{\sum(T_{off} - QRS_{on})}{n}$	$RT_{m_{opt}} = \frac{1}{16}\frac{\sum(T_{peak} - R_{peak})}{n}$
	variance		$RT_{var} = \frac{\sum(RT - RT_m)^2}{n}$	$RT_{var_{opt}} = \frac{1}{64}(\frac{\sum(RT_{opt} - RT_{m_{opt}})^2}{n})$
PS Interval	mean	It represents both atrial depolarization and ventricular depolarization.	$PS_m = \frac{\sum(P_{on} - QRS_{off})}{n}$	$PS_{m_{opt}} = \frac{1}{16}\frac{\sum(P_{peak} - S_{peak})}{n}$
iCEB	mean	It signifies balance and imbalance of depolarization (QRS duration) and repolarization (QT) of the cardiac electrophysiology. $iCEB = \frac{QT}{QRS}$	$iCEB_m = \frac{\sum(\frac{(T_{off} - QRS_{on})}{(QRS_{off} - QRS_{on})}}{n}$	$iCEB_{m_{opt}} = \frac{\sum(\frac{(T_{peak} - Q_{peak})}{(Q_{peak} - S_{peak})}}{n}$
	variance		$iCEB_{var} = \frac{\sum(iCEB_m - iCEB)^2}{n}$	$iCEB_{var_{opt}} = \frac{\sum(iCEB_{m_{opt}} - iCEB_{opt})^2}{n}$

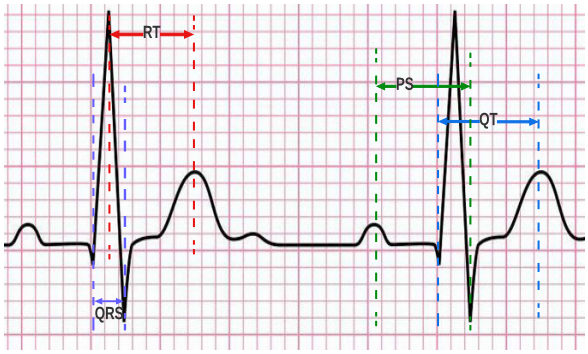


Fig. 3. Optimized ECG intervals.

difference between QRS_{on} and T_{off}. However, estimation of these features requires delineating both peaks and boundaries of the $P-QRS-T$ waves. However, estimating the boundaries of $P-QRS-T$ waves is critical due to the varying morphology of the ECG waves and requires separate modules for its implementation. This significantly increases resource utilization. Therefore, in this article, features are optimized by taking the RT interval as the difference between the R peak and the T peak (Fig. 3) instead of the boundaries which avoids the need to incorporate boundary detection modules of ECG feature extraction algorithm [17]. Other features are optimized similarly, as shown in Table II.

Furthermore, it was observed that using the interval features might cause overflow while implementing a DNN on hardware. Thus, the number of bits required to represent data needs to be increased to get high classification accuracy to accommodate the values of estimated features, increasing hardware resources. Therefore, the features are further optimized by dividing them with a specific threshold described in Table II to avoid increasing the number of data bits without reducing accuracy. Initially, the 16-bit values of $P-QRS-T$ features are zero-extended to 32 bits for the intermediate operations. Later, this is truncated to 16 bits as required.

As evident from Table II, to estimate the descriptors, we mainly need subtraction, squaring, and division operations. However, it is observed that a 3-s ECG signal sampled at 250 Hz has four ECG beats at most. Therefore, the number of peaks, i.e., "n," is less than or equal to four. Thus, the summation of every feature needs to be divided by an integer less than or equal to four to find the mean and variance of an ECG excerpt. Therefore, the descriptor processor is implemented using simple case statements, which become functional as per the number of R peaks. Thus, division by "n" operation is converted into right shift by 1 or 2, if n is 2 or 4, respectively. However, if the value of n is 3, then dividing by three is approximated using the following equation:

$$\text{div\_out} = (x \gg 4 + x \gg 2 + x \gg 5). \tag{1}$$

The complete descriptor block is implemented as an FSM. In the first state, the difference operations are executed. Once the difference values of peaks as per the optimized definition of descriptors are obtained, the summation of these values is performed in the second state. In the third state, the mean values are estimated to find $QRS_{m_{\text{opt}}}$, $RT_{m_{\text{opt}}}$, $PS_{m_{\text{opt}}}$, and $\text{iCEB}_{m_{\text{opt}}}$. Once the mean values are obtained, these values are divided by the respective thresholds, as shown in Table II. As evident from this table, the thresholds are taken so that they can be implemented using simple right shift operations, which avoids the need for division operations. Similar computations are done to find the variance of RT and iCEB, once the mean of the respective descriptors is estimated.

D. DNN Classifier

A conventional architecture of DNN is as shown in Fig. 4. The DNN has input layer, hidden layers, and output layer. Each node of the network is connected to the nodes of the next layer with some weights. The main control FSM of the DNN classifier is enabled once the estimated descriptors are fed as input to the classification block. In this article, we have used a low-power implementation of a DNN trained on software using an Intel[1] Xeon[1] CPU E5-2630 machine.

[1]Registered trademark.

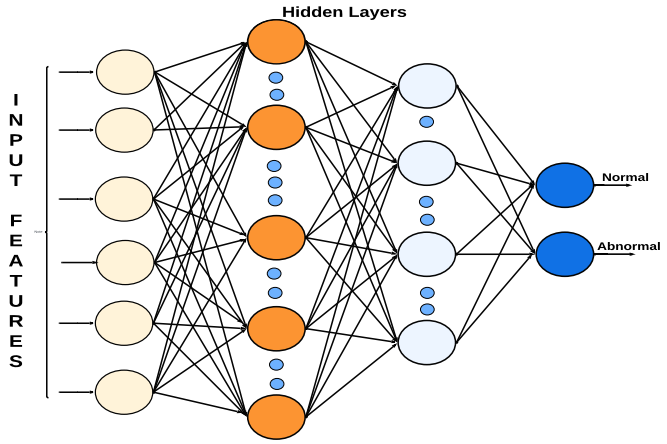


Fig. 4. Conventional DNN architecture.

To achieve this, a set of six optimized features are used, including RT_m, PS_m, mean duration of QRS complexes, and iCEB_m and iCEB_{var}. The training parameters of the DNN are carefully selected to optimize the model's performance. The activation functions used in DNN are ReLU and sigmoid, commonly used in deep learning applications. The number of epochs is determined using early stopping to prevent overfitting of the model. The Adam optimizer, a computational and memory-efficient gradient-based optimization algorithm, is used in the proposed design. The sparse categorical cross-entropy loss function is used in DNN training. It is ideal for multiclass classification problems where the classes are not mutually exclusive, and multiple nodes represent the output. It calculates the cross-entropy loss between the predicted and true class probabilities and is particularly useful when the number of classes is large or unknown. Note that two output nodes are used instead of one to optimize the hardware implementation of the trained DNN. The size of the DNN is set to $32 \times 16 \times 8 \times 2$, where each number represents the number of neurons in the respective layer. This architecture is chosen after several rounds of experimentation and testing to balance model complexity and accuracy.

During implementation, each descriptor is represented in 2's complement format using 16-bits. It is worth mentioning that the classification accuracy has no distinct drop after limiting the bit length of the descriptors to 16 bits, where 10 bits are used to represent the integer part, and 6 bits are for the fractional part. Furthermore, the weights and biases are presented in a similar fixed-point format. As the weights obtained after training are less than 1, the neuron output is also truncated to 16 bits without reducing accuracy. Overall, the selected DNN training parameters, and the optimized set of features, achieve high accuracy, sensitivity, and specificity in classifying ECG signals into normal and arrhythmia.

1) Operation of DNN Classifier: The operational principle of the proposed DNN classifier is presented in Fig. 2. Once the descriptor processing block computes the descriptors, the top FSM is activated. This further enables the FSM of the first DNN layer, which reads the input from the descriptor block and weights of the first neuron. Next, these weights are sequentially multiplied by the inputs. Once the multiplication and accumulation (MAC) of weights and input are completed, a bias is added to the MAC output. Furthermore, this output is passed through a multiplexer (MUX) of ReLU to obtain an output of the first neuron. The same procedure is repeated for all the neurons of a given layer. The output of neurons acts as input to the second layer, which is fed through a serial-in parallel-out (SIPO) register. Once the operation of layer-1 is completed, the top FSM passes the control to the FSM of the second layer, repeating the same process. The MAC unit and MUX are reused in successive layers, which minimizes area and power. In addition, folding the hardware mentioned above also optimizes leakage power. Furthermore, VLSI architecture of the proposed DNN classifier is optimized using a novel low-power multiplier and optimizing the implementation of activation function used in the output layer which are explained below.

2) Optimized Multiplier: For low-cost and compact wearable devices, power and area are the primary considerations over speed. Therefore, a parallel architecture of DNN is unnecessary as it increases leakage power and uses a larger silicon area. Therefore, as explained above, a single multiplier and accumulated unit is used in the proposed work, which is reused by neurons of every layer. In our previous work [11], we proposed a reduced register sequential multiplier for the efficient implementation of a MAC unit to optimize the power requirements of a DNN. However, the sequential operation of the proposed multiplier requires a higher clock frequency for real-time operation, which leads to higher power dissipation. Therefore, in this article, we propose an area and power optimal approximate multiplier (AM) which is also used in our previous work [24]. Despite this minimal error introduced in multiplication due to approximations, the accuracy of DNN is not affected when compared with the software implementation. Moreover, using the AM enabled us to maintain low-frequency operation for real-time processing of CoAP. The architecture is described in detail below.

In [24], we used an AM based on the Mitchell algorithm, which approximates the complete product. However, to further increase the efficiency of the algorithm, we use multiplication by splitting the multiplier and multiplicand into two parts, "I" and "F." As the "F" part is very small, the error percentage in logarithm approximation is smaller when compared with the error percentage of complete product [24]. The proposed AM uses shifting operations, which reduces hardware computation costs. The operation and implementation of the proposed multiplier are described below.

As we know, the product of any number with 2^n can be reduced as a left shift operation ($\ll n$). Therefore, to calculate the approximate product, binary numbers $a = a_{n-1}\ a_{n-2}\ \cdots\ a_0$ and $b = b_{n-1}\ b_{n-2}\ \cdots\ b_0$ are first represented as follows:

$$\begin{aligned} a &= 2^{ka}(1 - x_a) = 2^{ka} - 2^{ka}x_a \\ b &= 2^{kb}(1 - x_b) = 2^{kb} - 2^{kb}x_b. \end{aligned} \tag{2}$$

Let 2^{ka} and 2^{kb} be represented as A_i and B_i. Furthermore, let $2^{ka}x_a$ and $2^{kb}x_b$ be represented as A_f and B_f, respectively. Therefore, the product of two binary numbers a and b is

approximated as follows:

$$\begin{aligned} a \times b &= \left(A_i - A_f\right)\left(B_i - B_f\right) \\ &= A_i \times B_i - A_f B_i + A_f B_f - A_i B_f \\ &\approxeq A \times 2^{kb} + B \times 2^{ka} - 2^{k_a+k_b} + x_a x_b \times 2^{k_a+k_b}. \end{aligned} \quad (3)$$

$A \times 2^{kb}$, $B \times 2^{ka}$, and $2^{k_a+k_b}$ are calculated using simple shift operations, which would not introduce any errors, whereas $x_a x_b \times 2^{k_a+k_b}$ is estimated using a modified Mitchell algorithm proposed in [24].

Several accurate and AMs are proposed in the literature. However, AMs are widely used for applications where speed and area are prime considerations. Among all the available methods of approximate multiplications, Mitchell's algorithm [25] is the most simplified approach for integer and fixed-point multiplication.

Mitchell has proposed an efficient AM [25], which uses the log multiplication property. It estimates approximate multiplication by linearly approximating log and antilog values. First, the approximate characteristic part of log is obtained by finding the position of the leading one, i.e., the leftmost one in the binary sequence. Later, the remaining value is used as an approximate fractional part of the log. Furthermore, an approximate antilog operation is performed by adding the two operands, which generates an approximate product.

Let an N bit integer be Y with bits $Y_{n-1}, Y_{n-2}, \ldots, Y_0$, which is represented as $Y = \sum_{i=0}^{N-1} 2^i y_i$. Assuming the leading one occurs at position K, where $(N-1) <= K <= 0$, Y can be presented as follows without any loss in accuracy:

$$Y = 2^k\left(1 + \sum_{i=0}^{k-1} 2^{i-k} y_i\right). \quad (4)$$

Equation (4) is further represented as $Y = 2^k(1+x)$, where $x = \sum_{i=0}^{k-1} 2^{i-k} y_i$. Therefore, $\log_2$ of Y can be presented as follows, where K being an integer represents the characteristic value of $\log_2$ and $\log_2(1+x)$ is the fractional part:

$$\log_2 Y = k + \log_2(1+x). \quad (5)$$

As $0 <= x < 1$, therefore, the linear approximation of $\log_2 Y$ can be presented as follows, where function $\widetilde{\log}$ represents approximate $\log_2$:

$$\widetilde{\log_2} Y = k + x. \quad (6)$$

Furthermore, let us assume that there are two N-bits binary numbers, Y_1 and Y_2, with leading ones at k_1 and k_2. These numbers can be presented as $Y_1 = 2^{k_1}(1+x_1)$ and $Y_2 = 2^{k_2}(1+x_2)$. Therefore, the approximate product $(\widetilde{P})$ is estimated as follows:

$$\widetilde{\log_2}\left(\widetilde{P}\right) = k_1 + k_2 + x_1 + x_2. \quad (7)$$

Next, to calculate approximate antilog of (7), 1 is added to the fractional part (x), which is in range $[0, 1)$ and is scaled with respect to the characteristic part. Since $0 \leq (x_1, x_2) < 1$, thus, in this case $0 \leq x_1 + x_2 < 2$. Furthermore, the approximate product is decomposed into two cases. The first case is $0 \leq x_1 + x_2 < 1$, in which a carry is not generated from the fractional part to the characteristic part,

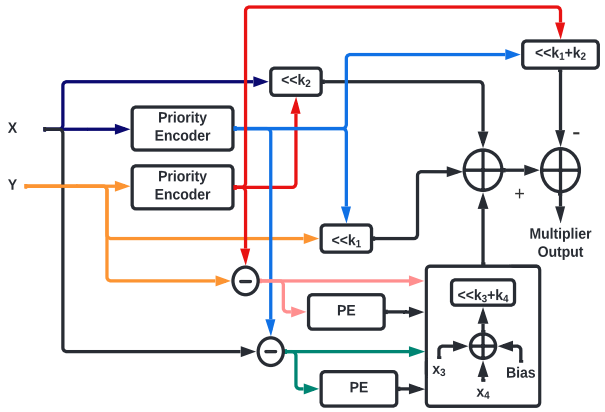


Fig. 5. Proposed optimal multiplier.

whereas the second case is $1 \leq x_1 + x_2 < 2$, in which a carry is generated to the characteristic part as follows:

$$\widetilde{P} = \begin{cases} 2^{k_1+k_2}(x_1 + x_2 + 1), & x_1 + x_2 < 1 \\ 2^{k_1+k_2+1}(x_1 + x_2), & x_1 + x_2 \geq 1. \end{cases} \quad (8)$$

This approximation reduces the multiplication operation to the basic add and shift operations. However, this decomposition introduces certain errors in the product obtained. Thus, an efficient error reduction technique is used for a more accurate multiplication of the weight and bias of a neuron, which considerably reduces the error of the multiplier. A bias is calculated by averaging the error across the entire range of fractional part x, which is added to the approximate product to improve accuracy. Next, an error E in the approximate product is estimated by the following equation, where P is the logarithm product of Y_1 and Y_2, and $\widetilde{P}$ is the approximate product output by Mitchell's algorithm:

$$\begin{aligned} E &= \widetilde{P} - P \\ &= \begin{cases} -2^{k_1+k_2}(x_1 x_2), & x_1 + x_2 < 1 \\ -2^{k_1+k_2}(1 + x_1 x_{12} - x_1 - x_2), & x_1 + x_2 \geq 1. \end{cases} \end{aligned} \quad (9)$$

Furthermore, the average error (bias) is calculated using the following equation:

$$\begin{aligned} E_{\text{avg}} &= \frac{1}{(1-0)(1-0)} \int_0^1 \int_0^1 E dx_2 dx_1 \\ &= \int_0^1 \int_0^{1-x_1} -2^{k_1+k_2}(x_1 x_2) dx_2 dx_1 \\ &\quad + \int_0^1 \int_{1-x_1}^1 -2^{k_1+k_2}(1 + x_1 x_2 - x_1 - x_2) dx_2 dx_1 \\ &= -2^{k_1+k_2}(0.083333). \end{aligned} \quad (10)$$

Note that the average error (bias) is always negative and depends on k_1 and k_2. Bias is shifted and added as per the position of the leading ones in Y_1 and Y_2. The implementation of the proposed multiplier is depicted in Fig. 5. The leading one finder is optimally implemented in the hardware as a simple priority encoder to estimate k_1 and k_2. Furthermore, the binary numbers can be shifted and rearranged in a fixed point representation to get respective fractional parts x_1 and x_2. Finally, the approximate product is obtained by adding and shifting x_1, x_2, k_1, and k_2 according to (8).

Fig. 5 presents the architecture of the proposed multiplier. It was observed that the proposed multiplier reduces the error in the output by 1.69% when compared with the Mitchell multiplier. This reduction in error maintains the accuracy of DNN on a hardware platform when compared with the software.

3) Optimized Implementation of Activation Function: For power-restricted wearable applications, the primary consideration for implementation is the simplicity of the hardware. Therefore, ReLU is chosen as an activation function in the proposed design because it requires only a simple MUX in the implementation

$$ai = \text{ReLU}(x) = \begin{cases} x, & \text{if } x \geq 0 \\ 0, & \text{if } x < 0. \end{cases} \tag{11}$$

Furthermore, as we perform two-class classification, a Sigmoid is used as an activation function for the final output layer. Conventionally, only a single output node with a sigmoid activation function is used for a two-class classification using a DNN. However, in this work, we propose two output nodes instead of one, each with a sigmoid activation function in the output layer while training the network. It aids in optimizing the hardware implementation of the neural network, as explained below.

A sigmoid activation function has an exponent function, as shown in the following equation:

$$\text{Sigmoid}(x) = \frac{1}{1 + e^x}. \tag{12}$$

Hardware implementation of exponential function requires complex algorithms such as CORDIC or series expansion, requiring more area and power resources. To optimize the hardware implementation of the trained neural network, we use two output nodes during training. Using two output nodes instead of a single node does not create any classification errors. Moreover, it enabled us to replace the sigmoid activation function by a simple comparator in hardware implementation of the trained neural network. The comparator compares the output of two final nodes, and the one with the maximum value is chosen as the output. Thus, using two nodes enable us to eliminate the sigmoid function during hardware implementation, optimizing the area and power considerably.

IV. Results and Discussions

This section presents a detailed discussion of the outcome of the proposed design and its validation using various testcases.

A. Dataset Description and Evaluation Metrics

For detecting VA, the features extracted from the ECG segment before the onset of VT/VF is used as training and test data. It is imperative to predict VA rather than detect it to save lives. Therefore, analyzing the feature variations in the ECG segments preceding VA beats is necessary. To incorporate this methodology, we consider ECG segments from the MIT-DB dataset [13]. The MIT-BIH arrhythmia database [26] consists of 48 two-lead ECG recordings, and each is of half-hour duration and is sampled at 360 Hz. The dataset has 108 655 beats, labeled by expert cardiologists into different types of arrhythmia as per the American National Standard (ANSI/AAMI EC57:1998) [27]. The MIT-BIH database has two ECG lead recordings; modified limb lead II (MLII) and modified lead (V1). Only MLII is used for ECG classification because of the prominent QRS complexes in the proposed work. Later, ECG data are downsampled to 250 Hz to optimize the number of samples in an ECG segment. Note that downsampling the data reduces the number of samples per ECG segment without affecting the overall performance of the classifier.

Furthermore, there are different approaches to testing an ECG classifier; the first is a class-oriented approach, in which the complete data are split into test and training data. Note that in this approach, ECG beats from the same patient might or might not be included in both test and training data. This technique is used when limited data are available for classifier training. However, a cross-validation technique can be used to make the classifier robust to the variations in an ECG signal. The second is the subject-oriented approach, in which patients are divided into training and test datasets completely blind to each other. Since, the training data are completely blind to the test data, a large number of training excerpts are required for training such kinds of classifier to achieve a high accuracy. Achieving high accuracy in this case ensures that the classifier can work well for every individual, making it highly reliable and robust in real-life applications.

We use both the approaches to validate the performance of our proposed design in this article. Furthermore, the ECG records taken from the MIT-DB database are processed in the following manner. First, the VT/VF beats are identified. Next, ECG segments of 3 s are considered before this onset of VT/VF beats. The 3-s windows are extended to different time intervals, "T." The data between this "T" and the onset of VT/VF beat are considered abnormal (Class A). The data before this "T" interval are considered normal (Class N). Later, these data are divided into two groups: DT1 (75% of total data), the training set, and DT2 (25% of total data), the test set for the class-oriented approach. Whereas for the subject-oriented approach, we consider ECGs of patients "100," "113," "106," "123," "232," "221," "210," "111," "101," "222," "112," "103," "102," "234," "231," "121," "105," "230," and "228" for training and "115," "116," "118," "122," "124," "209," and "223" for testing from the MIT-BIH dataset. Analyzing the ECG signal in such a way enabled the classifier to predict the VT/VF event 15 min (T) before its occurrence.

B. Performance Investigation

The initial algorithm verification and model optimization is performed using Python. Furthermore, the complete design of CoAP is implemented using Verilog-HDL, whose performance was tested using several testbenches created using ECG excerpts. This section first presents the performance metrics of CoAP. Furthermore, it depicts the outcome of FPGA and ASIC implementations of CoAP. Next, a detailed description of the comparison of the proposed design with the contemporary state-of-the-art designs is given. Finally, we summarize

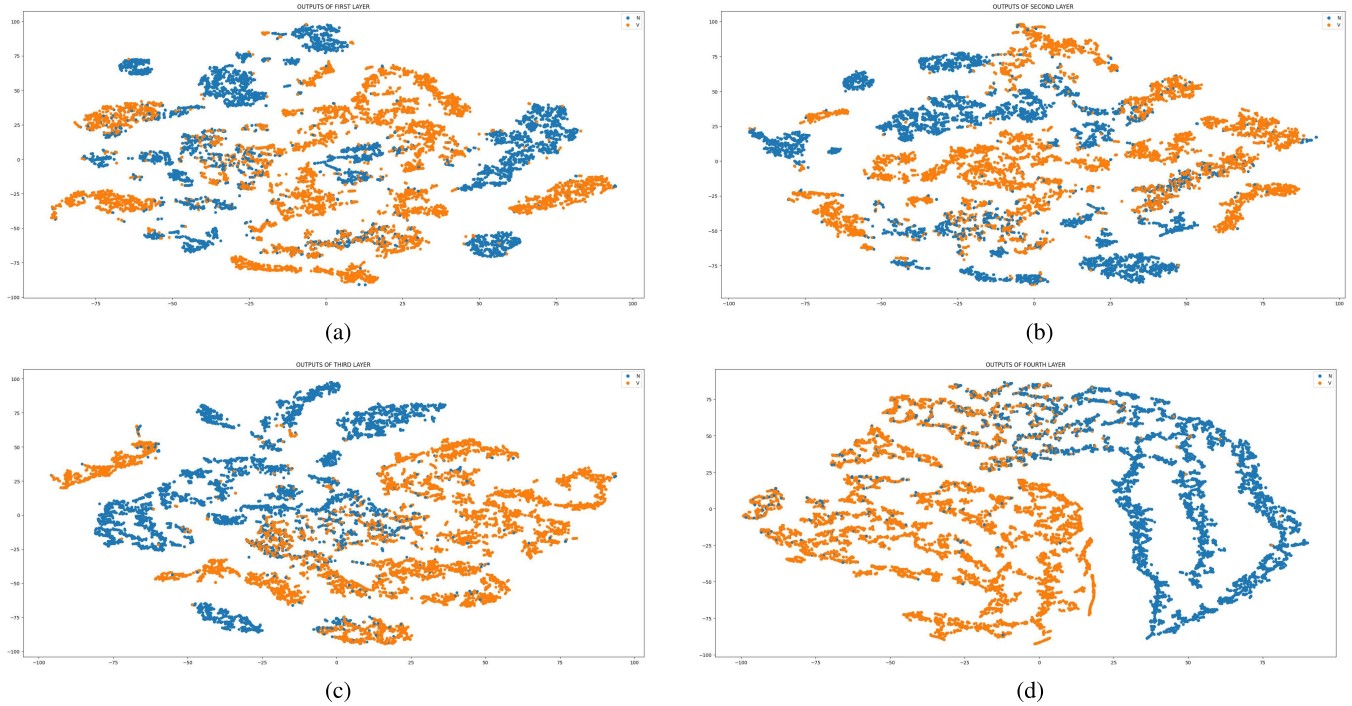


Fig. 6. t-SNE plots of different layers of DNN. (a) Output of layer 1. (b) Output of layer 2. (c) Output of layer 3. (d) Output of layer 4.

the findings and novelties of the proposed design in this section.

1) Evaluation Metrics: Among the performance metrics used for medical diagnosis, sensitivity (Se), specificity (Sp), precision, and F1 score are the crucial measures of performance diagnostics. High sensitivity enables clinicians to rule out disease, whereas a high specificity identifies a particular disease in the patients [28]. Therefore, sensitivity (Se) and specificity (Sp) are used to measure the diagnostic ability of the classifier in Classes N and V. Furthermore, the accuracy, precision, and F1 score of CoAP are also calculated to estimate overall system performance.

The performance of CoAP is tested at different times "T," i.e., from 5 min before the onset to 30 min before the onset of VA beats. Furthermore, a ten-fold cross-validation is performed to better estimate the model's performance compared with a single train–test split. We observe that CoAP can predict arrhythmia 15 min before its occurrence with an average 91.19% accuracy, 94.01% sensitivity, 87.52% specificity, 90.76% precision, and 0.92 F1 score. Considering "T" more than 15 min leads to very low accuracy. Since a wearable device needs to generate minimal false alarms, DNN is implemented on hardware and is trained on the dataset taken 15 min before the onset of VA.

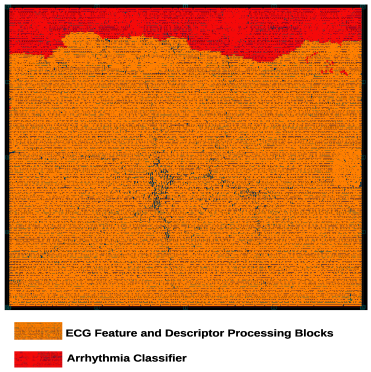
The t-distributed stochastic neighbor embedding (t-SNE) plots are also presented to illustrate the functioning of hidden nodes. The t-SNE [31] is a nonlinear dimension reduction method used to visualize high-dimensional data on a lower dimensional (two–three) space. The outputs of the first, second, and third hidden layers of the DNN are extracted from the input vectors. Note that t-SNE is often used as a dimensionality reduction technique, whereas we use it only as a data visualization technique in the proposed study. We have applied t-SNE on the 6-D input vectors and the hidden layers. Fig. 6 presents the results for Classes N and V. It can be observed from Fig. 6 that there are no separate clusters for the feature vectors [Fig. 6(a)]. As we progressively move forward to the hidden layers [Fig. 6(b) and (c)], the clusters become visible. However, Classes N and V still have multiple small clusters for the first two hidden layers. This shows that nonlinear operations are needed for a complete classification. However, it becomes evident that in the final layer [Fig. 6(d)], large clusters of both the classes are clearly visible. This showcases the discriminative nature of DNN as we move from the bottom layer to the top layer. The hidden layers have disentangled the classes from the considered ECG excerpts, making the final layer representation distinctive.

Furthermore, the proposed model is tested using a subject-oriented approach as explained in Section IV-A. Our model achieves an accuracy of 74.64%, sensitivity of 76.34%, specificity of 71.54%, and F1 score of 0.79. As explained before, that during the subject oriented approach, a large amount of training data which will be completely independent of the test data is required. This approach ensures that the classifier will be more accurate for the real-world applications as it can predict arrhythmia with high accuracy for any individual. Note that even with limited data, our proposed model can predict the VA events with approximately 75% accuracy for a highly accurate patient-independent subject-oriented approach.

2) FPGA Implementation: The proposed architecture is first implemented on Xilinx Virtex-7 FPGA for functional verification. Table III presents the resource utilization of the proposed design. It is observed from Table III that our design uses 0.69% of total available resources, making it resource-efficient.

TABLE III
FPGA RESOURCE UTILIZATION

Resources	Total Available	Resources Utilized	Percentage Utilization(%)
Slice LUT	303600	5257	1.73
Slice REG	607200	2354	1.55
F7 MUX	151800	206	0.13
F8 MUX	75900	64	0.08
DSP	2800	2	0.07
BRAM	1030	1	0.09
IOB	600	21	0.35
BUFGCTRL	32	1	3.125
Total Resources	1142962	7906	0.69



CMOS Process	SCL 180nm
Area	1.8 mm^2
Voltage	1.98V
Frequency	12.5kHz
Dynamic Power	1.8323 μW
Static Power	2.8594 μW

Fig. 7. Layout photograph of the CoAP architecture and its specifications.

3) ASIC Implementation: CoAP is implemented as an ASIC using 180-nm Bulk CMOS Technology available with SCL Chandigarh as shown in Fig. 7. The design is synthesized using Synopsys DC compiler, and placement and routing are performed using Synopsys IC compiler. It is observed that the proposed design, which includes the feature extraction block and the complete classifier block with all its parameters, can operate at a minimum frequency of 12.5 kHz to perform beat-by-beat classification in real-time, consuming 4.69-μW power. The total power dissipation comprises static and dynamic power as shown in Fig. 7.

Despite the ability of the proposed architecture to operate at a higher clock frequency till 1 MHz, it is sufficient to choose 12.5 kHz as an operating frequency to fulfill the low power requirement.

4) Comparison With State-of-the-Art: Table IV presents a comparison of our proposed classifier with well-known state-of-the-art arrhythmia classifiers on the hardware and software platforms. It is observed from Table IV that CoAP exhibits lower accuracy than the methods reported in [9], [10], [12] and [16]. However, the accuracy reported in [9], [10], [12], and [16] is for the detection of arrhythmia and not the prediction, and an alert is sent to an individual only when arrhythmia has occurred.

Nevertheless, our proposed classifier can predict the arrhythmia beats 15 min before their occurrence with an accuracy of 91.19%. Therefore, the most fair comparison of the proposed work will be with [14] and [15]. It can be observed from Table IV that the architecture of CoAP has better or comparable performance compared with [14] and [15], which also report the prediction of VT/VF over detection. First, we consider the performance comparison of the proposed model in predicting the arrhythmia. It can be seen that CoAP has approximately 6% higher accuracy than [14]. Clinically, sensitivity and specificity are important for confirming or excluding disease during screening. Sensitivity is the percentage of persons with the disease who are correctly identified by the test. Specificity is the percentage of persons without the disease who are correctly excluded by the test. The proposed work has high sensitivity and specificity which implies that model has high performance in both accurately detecting the true positives and rejecting the false negatives. However, [14] does not report sensitivity and specificity. On the other hand, CoAP has almost comparable sensitivity to [15], but it is slightly less specific than the design reported in [15]. This means that [15] is better in excluding persons without the disease than the proposed approach. However, the primary aim of the proposed work is to develop a computationally simple VLSI architecture which can have good performance metrics in terms of both accuracy and hardware requirements. The methodology proposed in [15] is implemented only on the software platform using complex input feature vectors and decision trees as a classifier. Although decision tree implementation is not critical in hardware, the complex input features, which lead to a slight increase in the sensitivity considered in [15], require complete ECG feature extraction. These features also need to find the boundaries of $P-QRS-T$ waves. Therefore, the hardware requirements of [15] would be higher if it is implemented as an ASIC. Therefore, it can be stated that the proposed design has the highest performance metrics than the other state-of-the-art methods. Furthermore, the power requirement of the proposed design is less than [9], [12] and [16] but more than [10] and [14]. The designs reported in [10] and [14] are implemented on the lower technology nodes with less operating voltage. Furthermore, the design reported in [10] can only detect the arrhythmia beats, and therefore, it uses simple R peak-based features for classification, which have lower hardware requirements. Moreover, the design proposed in [14] uses a Naive Bayes classifier and interval-based input features. The computation of these features is easier than the statistical features used in the proposed work, and therefore, it used a lower power of 2.69 μW which is approximately 2 μW less than the proposed approach. Therefore, the battery life of the wearable device will be a little longer than the proposed approach. However, it should be noted that the accuracy of the design proposed in [14] is less than CoAP. As it is very crucial for a wearable device to have as high accuracy as possible because a false alarm may cause significant panic due to fatal nature of the proposed disease. Therefore, it is beneficial to trade off battery life with accuracy. Furthermore, our design is implemented on 180 nm compared to the 65-nm implementation of [14]. As we know that the dynamic power is directly proportional to the voltage, it can be anticipated that our proposed design will have even less power requirement when implemented on the lower technology node. Therefore, it can be inferred from Table IV that CoAP is most suitable for low-power wearable electronics devices in the healthcare domain to predict arrhythmia as it has higher accuracy and low power consumption.

TABLE IV
COMPARISON WITH STATE-OF-THE-ART METHODS

Parameter	**Wavelet Based [9]**	**SVM [12]**	**ANN [10]**	**Threshold Based [16]**	**Decision Trees [15]**	**Naive Bayes [14]**	**Proposed**
Approach	Detection	Detection	Detection	Detection	Prediction	Prediction	Prediction
Features Used	Wavelet Based Features	ECG Beats and RR Interval	RR Interval and QRS Area	RR,QS Interval and, T and S Waves	SDNN, RMSSD, mQRSd, mean, standard deviations of Q-R-S points, heart rate, amplitude of Q, R and S Points	RR,PQ,QP, RT, TR, PS, SP Interval	QRS mean, RT mean and variance, PS mean , and iCEB mean and variance
Database	MIT-BIH	MIT-BIH	MIT-BIH	MIT-BIH	NSRDB, VFDB	MIT-BIH , NSRDB	MIT-BIH
Accuracy	95.83%	99.68%	98.3%	97.02%	NA	86%	91.19%
Sensitivity	NA	NA	NA	94.64%	95%	NA	94.01%
Specificity	NA	NA	NA	99.41%	90%	NA	87.52%
Platform	ASIC	ASIC	ASIC	ASIC	Software	ASIC	ASIC
Technology Node (nm)	180	180	40	180	NA	65	180
Voltage (V)	1.2	1.8	1.1	1.8	NA	1	1.98
Frequency (Hz)	NA	25M	10k	1k	NA	10k	12.5k
Area (mm^2)	2.92	0.9246	0.12	NA	NA	0.112	1.8
Power (W)	28.7u	13.34μ	3.76μ	5.04μ	NA	2.78μ	4.69μ

V. CONCLUSION

In this article, a low-power ECG co-processor for predicting cardiac arrhythmia is proposed, which can predict the occurrence of VA and VF 15 min before its occurrence with an accuracy of 91.19%. The proposed design uses an optimal feature set and novel DNN classifier architecture, enabling it to operate at a minimum frequency of 12.5 kHz, consuming 4.69 μW of power. The low power and high classification performance of CoAP enable it to be used efficiently in wearable devices in the healthcare domain.

VI. FUTURE WORK

We have trained our models on a selected population of malignant VAs, not including populations from different demographic locations, which might affect the generalization of the proposed method. Due to their lifestyles, the ECG signal variations may differ in various geographical regions. A classifier trained on a wide variety of ECG data is anticipated to be more accurate for a generic wearable device, predicting the VA above 90% accuracy for a subject-oriented approach. Moreover, we can optimize the feature set more efficiently with extensive data without compromising performance metrics. Furthermore, the proposed methodology can be improved to predict the VA minimum an hour before its occurrence. In addition, the circuit and transistor-level optimization techniques can be explored to make the proposed implementation more efficient, and the power can be optimized using low-power techniques such as clock gating and lower technology nodes.

REFERENCES

[1] C. J. Garratt, C. Statter, and D. Statters, *Mechanisms and Management of Cardiac Arrhythmias*. London, U.K.: BMJ Books London, 2001.

[2] D. P. Zipes et al., "ACC/AHA/ESC 2006 guidelines for management of patients with ventricular arrhythmias and the prevention of sudden cardiac death: A report of the American College of Cardiology/American Heart Association task force and the European society of cardiology committee for practice guidelines (writing committee to develop guidelines for management of patients with ventricular arrhythmias and the prevention of sudden cardiac Death) developed in collaboration with the European heart rhythm association and the heart rhythm society," *Europace*, vol. 8, no. 9, pp. 746–837, Sep. 2006.

[3] T.-M. Chen et al., "SRECG: ECG signal super-resolution framework for portable/wearable devices in cardiac arrhythmias classification," *IEEE Trans. Consum. Electron.*, vol. 69, no. 3, pp. 250–260, 2023.

[4] N. Sabor, G. Gendy, H. Mohammed, G. Wang, and Y. Lian, "Robust arrhythmia classification based on QRS detection and a compact 1D-CNN for wearable ECG devices," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 12, pp. 5918–5929, Dec. 2022.

[5] P. M. Tripathi, A. Kumar, M. Kumar, and R. Komaragiri, "Multilevel classification and detection of cardiac arrhythmias with high-resolution superlet transform and deep convolution neural network," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–13, 2022.

[6] Q. Cai, X. Xu, Y. Zhao, L. Ying, Y. Li, and Y. Lian, "A 1.3 μW event-driven ANN core for cardiac arrhythmia classification in wearable sensors," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 68, no. 9, pp. 3123–3127, Sep. 2021.

[7] X. Xu, Q. Cai, Y. Zhao, G. Wang, L. Zhao, and Y. Lian, "A 2.66 μW clinician-like cardiac arrhythmia watchdog based on P-QRS-T for wearable applications," *IEEE Trans. Biomed. Circuits Syst.*, vol. 16, no. 5, pp. 793–806, Oct. 2022.

[8] S. M. Abubakar, Y. Yin, S. Tan, H. Jiang, and Z. Wang, "A 746 nW ECG processor ASIC based on ternary neural network," *IEEE Trans. Biomed. Circuits Syst.*, vol. 16, no. 4, pp. 703–713, Aug. 2022.

[9] S. Lee, P. Huang, M. Liang, J. Hong, and J. Chen, "Development of an arrhythmia monitoring system and human study," *IEEE Trans. Consum. Electron.*, vol. 64, no. 4, pp. 442–451, Nov. 2018.

[10] Z. Chen et al., "An energy-efficient ECG processor with weak-strong hybrid classifier for arrhythmia detection," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 65, no. 7, pp. 948–952, Jul. 2018.

[11] M. Janveja, R. Parmar, M. Tantuway, and G. Trivedi, "A DNN-based low power ECG co-processor architecture to classify cardiac arrhythmia for wearable devices," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 69, no. 4, pp. 2281–2285, Apr. 2022.

[12] Y. Zhao, Z. Shang, and Y. Lian, "A 13.34 μW event-driven patient-specific ANN cardiac arrhythmia classifier for wearable ECG sensors," *IEEE Trans. Biomed. Circuits Syst.*, vol. 14, no. 2, pp. 186–197, Apr. 2020.

[13] 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, Jun. 2000.

[14] N. Bayasi, T. Tekeste, H. Saleh, B. Mohammad, A. Khandoker, and M. Ismail, "Low-power ECG-based processor for predicting ventricular arrhythmia," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 24, no. 5, pp. 1962–1974, May 2016.

[15] S. Mandala, T. Cai Di, M. S. Sunar, and Adiwijaya, "ECG-based prediction algorithm for imminent malignant ventricular arrhythmias using decision tree," *PLoS ONE*, vol. 15, no. 5, May 2020, Art. no. e0231635.

[16] S. M. Abubakar, W. Saadeh, and M. A. B. Altaf, "A wearable long-term single-lead ECG processor for early detection of cardiac arrhythmia," in *Proc. Design, Autom. Test Eur. Conf. Exhib. (DATE)*, Mar. 2018, pp. 961–966.

[17] M. Janveja, R. Parmar, and G. Trivedi, "MInSC: A VLSI architecture for myocardial infarction stages classifier for wearable healthcare applications," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 70, no. 3, pp. 1159–1163, Mar. 2023.

[18] S. Mandala and T. C. Di, "ECG parameters for malignant ventricular arrhythmias: A comprehensive review," *J. Med. Biol. Eng.*, vol. 37, no. 4, pp. 441–453, Aug. 2017.

[19] M. Thaler, *The Only EKG Book You'll Ever Need*. Philadelphia, PA, USA: Lippincott Williams & Wilkins, 2017.

[20] S. K. Jain and B. Bhaumik, "An energy efficient ECG signal processor detecting cardiovascular diseases on smartphone," *IEEE Trans. Biomed. Circuits Syst.*, vol. 11, no. 2, pp. 314–323, Apr. 2017.

[21] J. M. Bote, J. Recas, F. Rincón, D. Atienza, and R. Hermida, "A modular low-complexity ECG delineation algorithm for real-time embedded systems," *IEEE J. Biomed. Health Informat.*, vol. 22, no. 2, pp. 429–441, Mar. 2018.

[22] T. Tekeste, H. Saleh, B. Mohammad, and M. Ismail, "Ultra-low power QRS detection and ECG compression architecture for IoT healthcare devices," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 66, no. 2, pp. 669–679, Feb. 2019.

[23] G. S. Wagner, *Marriott's Practical Electrocardiography*. Philadelphia, PA, USA: Lippincott Williams & Wilkins, 2001.

[24] R. Parmar, M. Janveja, J. Pidanic, and G. Trivedi, "Design of DNN-based low-power VLSI architecture to classify atrial fibrillation for wearable devices," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 31, no. 3, pp. 320–330, Mar. 2023.

[25] J. N. Mitchell, "Computer multiplication and division using binary logarithms," *IRE Trans. Electron. Comput.*, vol. EC-11, no. 4, pp. 512–517, Aug. 1962.

[26] 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 2001.

[27] *Testing and Reporting Performance Results of Cardiac Rhythm and St Segment Measurement Algorithms*, Standard ANSI/AAMI EC57:2012 (R2020), AAMI Recommended Practice/American National Standard, 1998.

[28] A. Laupacis, "Clinical prediction rules. A review and suggested modifications of methodological standards," *JAMA, J. Amer. Med. Assoc.*, vol. 277, no. 6, pp. 488–494, Feb. 1997.

[29] X. Chen, G. Wang, W. Zhou, S. Chang, and S. Sun, "Efficient sigmoid function for neural networks based FPGA design," in *Intelligent Computing. ICIC 2006* (Lecture Notes in Computer Science), vol. 4113, D. S. Huang, K. Li, and G. W. Irwin., Eds. Berlin, Germany: Springer, 2006.

[30] J. A. Swets, *Signal Detection Theory and ROC Analysis in Psychology and Diagnostics: Collected Papers*. London, U.K.: Psychology Press, 2014.

[31] L. Van der Maaten and G. Hinton, "Visualizing data using t-SNE," *J. Mach. Learn. Res.*, vol. 9, no. 11, pp. 1–27, 2008.

[32] A. Sammani et al., "Life-threatening ventricular arrhythmia prediction in patients with dilated cardiomyopathy using explainable electrocardiogram-based deep neural networks," *EP Europace*, vol. 24, no. 10, pp. 1645–1654, Oct. 2022, doi: 10.1093/europace/euac054.

[33] *Cardiovascular Diseases (CVDs)*, World Health Org., Geneva, Switzerland, 2019. [Online]. Available: https://www.who.int/en/newsroom/fact-sheets/detail/cardiovascular-diseases-(cvds)

[34] M. Z. H. Kolk et al., "Dynamic prediction of malignant ventricular arrhythmias using neural networks in patients with an implantable cardioverter-defibrillator," *eBioMedicine*, vol. 99, Jan. 2024, Art. no. 104937, doi: 10.1016/j.ebiom.2023.104937.