AI Accelerator With Ultralightweight Time-Period CNN-Based Model for Arrhythmia Classification

Shuenn-Yuh Lee, *Senior Member, IEEE,* Ming-Yueh Ku, *Graduate Student Member, IEEE,* Wei-Cheng Tseng, and Ju-Yi Chen

***Abstract*—This work proposes a classification system for arrhythmias, aiming to enhance the efficiency of the diagnostic process for cardiologists. The proposed algorithm includes a naive preprocessing procedure for electrocardiography (ECG) data applicable to various ECG databases. Additionally, this work proposes an ultralightweight model for arrhythmia classification based on a convolutional neural network and incorporating R-peak interval features to represent long-term rhythm information, thereby improving the model's classification performance. The proposed model is trained and tested by using the MIT-BIH and NCKU-CBIC databases in accordance with the classification standards of the Association for the Advancement of Medical Instrumentation (AAMI), achieving high accuracies of 98.32% and 97.1%. This work applies the arrhythmia classification algorithm to a web-based system, thus providing a graphical interface. The cloud-based execution of automated artificial intelligence (AI) classification allows cardiologists and patients to view ECG wave conditions instantly, thereby remarkably enhancing the quality of medical examination. This work also designs a customized integrated circuit for the hardware implementation of an AI accelerator. The accelerator utilizes a parallelized processing element array architecture to perform convolution and fully connected layer operations. It introduces proposed hybrid stationary techniques, combining input and weight stationary modes to increase data reuse drastically and reduce hardware execution cycles and power consumption, ultimately achieving high-performance computing. This accelerator is implemented in the form of a chip by using the TSMC 180 nm CMOS process. It exhibits a power consumption of 122 μW, a classification latency of 6.8 ms, and an energy efficiency of 0.83 μJ/classification.**

***Index Terms*—Arrhythmia, electrocardiography, biosignal preprocessing, advancement of medical instrumentation standard, convolution neural network, artificial intelligence accelerator.**

Manuscript received 20 April 2024; revised 7 June 2024 and 9 July 2024; accepted 18 July 2024. Date of publication 30 July 2024; date of current version 5 February 2025. This work was supported by Taiwan Semiconductor Research Institute and the National Science and Technology Council (NSTC), Taiwan, R.O.C., under Grant NSTC 112-2221-E-006-181 MY3 and Grant NSTC 113-2640-E-006-006. This paper was recommended by Associate Editor A. Cossettini. *(Corresponding authors: Shuenn-Yuh Lee; Ju-Yi Chen.)*

Shuenn-Yuh Lee, Ming-Yueh Ku, and Wei-Cheng Tseng are with the Department of Electrical Engineering, National Cheng Kung University, Tainan 701, Taiwan (e-mail: ieesyl@mail.ncku.edu.tw; myku0814@gmail.com; a0983989571@gmail.com).

Ju-Yi Chen is with the Department of Internal Medicine, National Cheng Kung University Hospital, College of Medicine, National Cheng Kung University, Tainan 701, Taiwan (e-mail: juyi@mail.ncku.edu.tw).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TBCAS.2024.3435718.

Digital Object Identifier 10.1109/TBCAS.2024.3435718

I. Introduction

In recent years, factors, such as high-stress work environments and the dietary culture of processed foods, have led to a dramatic increase in the occurrence of cardiovascular diseases among the modern population. According to surveys by the World Health Organization [1], cardiovascular diseases have become the leading cause of global mortality. This issue reflects the importance of heart health in contemporary society given that conditions like myocardial infarctions can result in sudden death. Early detection and treatment are crucial to prevent such tragic events. Arrhythmias, which are characterized by irregular heartbeats, are common occurrences in daily life for everyone. However, the regularity and specific symptoms associated with arrhythmias can lead to severe cardiovascular diseases and potentially life-threatening situations. Therefore, the timely detection of arrhythmia symptoms is a preventive measure against heart disease. Recognizing the signs of abnormal heart conditions early on allows for prompt intervention and treatment, effectively reducing the likelihood of myocardial infarctions.

An electrocardiogram (ECG) is a diagnostic test that measures and records the electrical activity of the heart over a specific period. It is a valuable tool that is used by healthcare professionals to assess the heart's rhythm and detect any abnormalities or irregularities in its electrical impulses. Arrhythmias can be detected by measuring ECG signals, as commonly illustrated in Fig. 1. Given that the occurrence of arrhythmias cannot be precisely determined, a 24-hour ECG Holter monitoring device is commonly employed to record long-term information. After the monitoring period, the collected data are manually reviewed by medical professionals, who annotate instances of arrhythmic symptoms. The annotated results are then provided to a cardiologist for diagnosis. Finally, a diagnostic report is generated for the patient. However, this diagnostic process is not only time-consuming, it also demands a considerable amount of human resources. The abundance of data increases the risk of misclassification, ultimately compromising the quality of healthcare.

This work utilizes machine learning technology to develop an arrhythmia classification system to address the above issues. It can train deep learning models to classify disease types, achieving rapid and automated classification with high accuracy. This work also designs a web-based system that provides a graphical interface for displaying ECG waveforms. Additionally, an

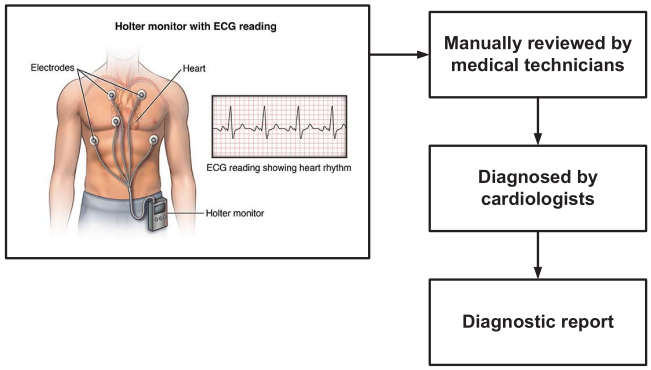


Fig. 1. 24-hour holter electrocardiogram diagnostic process.

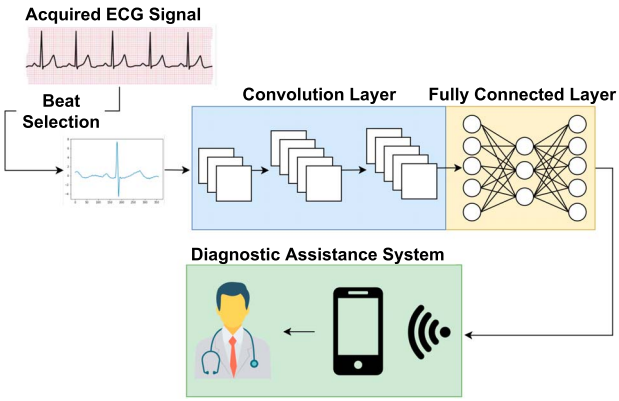


Fig. 2. Arrhythmia classification system for diagnostic assistance.

artificial intelligence (AI) model is integrated with cloud computing to assist cardiologists in diagnosis, thus reducing the time consumption of the diagnostic process and enhancing healthcare quality. Arrhythmia classification for diagnosis is shown in Fig. 2.

With the rapid development of AI in recent years, its applications in the field of medical diagnosis have become increasingly widespread. Relevant studies indicate that convolutional neural networks (CNNs) are suitable for arrhythmia classification [2]. Convolutional layers function as feature extractors [3], effectively handling sequential graphical data, such as ECG data. Fully connected layers serve the purpose of classification by learning the correlations between feature points. CNN models can achieve high accuracy in this context. In recent years, numerous studies have employed AI models to identify heart diseases. However, most of these studies have focused solely on software development without considering issues, such as model size and computational load [4], [5], [6]. They typically used conventional GPUs or CPUs for hardware implementation, leading to the excessive consumption of hardware resources and power, thus making them unsuitable for wearable medical devices. By contrast, this study designed a lightweight model in consideration of practical applications. We proposed a customized AI accelerator to achieve low power consumption and latency in hardware circuits, making the accelerator suitable for wearable devices. This work has considered a hardware-software codesign, aiming to achieve high-accuracy results with limited hardware resources. Additionally, it incorporates additional time information to enhance classification performance. In short, the significant contributions of the study are as follows:

1) An ultralightweight time-period CNN-based model that contains only 921 parameters is proposed to achieve acceptable accuracy. Most importantly, it is friendly to hardware design.
2) A hybrid stationary technique to maximize hardware resource usage is adopted, resulting in a notable energy efficiency of 0.83 μJ/classification.
3) A web-based user interface (UI) that visualizes the classification results and the ECG waveform is developed, enabling doctors and technicians to easily check the classification results and modify the incorrect beats easily.

This paper is organized as follows: In the beginning, Section II introduces two databases we used. Section III introduces the algorithmic framework for the arrhythmia classification system, encompassing ECG data preprocessing and AI model design. Subsequently, Section IV presents the proposed AI hardware accelerator architecture and provides a detailed explanation of the design process. Next, Section V presents the system implementation and measurement results of the proposed AI accelerator. Finally, Section VI provides the conclusion of this study.

II. DataBases

A. MIT-BIH Arrhythmia Database

The MIT-BIH Arrhythmia Database [7] is a collaborative effort between the Massachusetts Institute of Technology (MIT) and Beth Israel Deaconess Medical Center. The team recorded the ECG data of clinical patients between 1975 and 1980, ultimately including data from 48 participants. Each participant's recording consists of 30 minutes of resting-state ECG. The database provides signals from two leads, primarily limb lead II (MLII) and lead V1, with a few participants having signals from leads V2, V4, and V5. The sampling frequency of the ECG signals is 360 Hz with a resolution of 11 bits and voltage ranges of $+5$ to -5 mV. MIT-BIH is internationally recognized as a standard ECG database and is one of the most widely used databases in related research. Therefore, this study used this database for model training and validation.

B. NCKU-CBIC ECG Database

This study utilizes the NCKU-CBIC ECG Database [8] for validation. This database was developed by the Communication and Biological Integrated Circuit Laboratory (CBIC) at National Cheng Kung University (NCKU). The team collected the ECG data of clinical patients from the Ministry of Health and Welfare Tainan Hospital between 2018 and 2022. The data collection process underwent review and certification by the Institutional Review Board of NCKU. The ECG recording device used by the patients is an ECG patch [9] with the lead configuration set to MLII. The sampling frequency of the ECG signals is 400 Hz, and the resolution is 12 bits. The database includes ECG records from six patients, with each patient's data comprising 4-hour

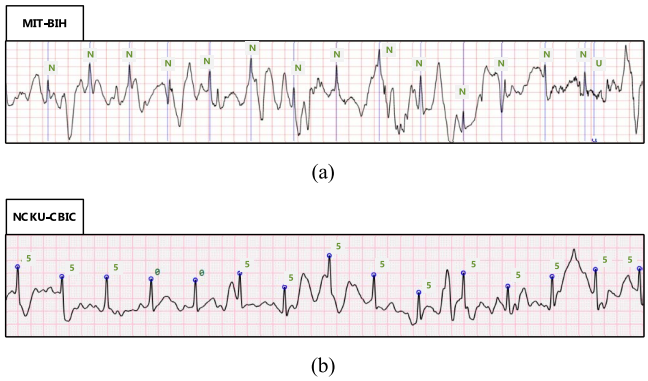


Fig. 3. (a) ECG labeled as a normal beat in MIT-BIH. (b) ECG labeled as a motion artifact in NCKU-CBIC.

TABLE I
NCKU-CBIC DISEASE LABELS AND STATISTICS

Arrhythmia Type	Label	Total Beats
Normal beat	0	66 330
Atrial fibrillation	1	23 674
Supraventricular tachycardia	2	2
Premature ventricular contraction	3	2360
Atrial premature contraction	4	527
Motion artifact	5	6912
Baseline wandering	6	1027
First-degree AV block	7	58
Atrial flutter	8	12 560

ECG recordings at different times of the day. All signals are derived from the 24-hour ECG recordings of clinical patients, reflecting ECG patterns during various daily life activities. In addition, this database includes annotations for motion artifacts and baseline wandering. During wearable ECG acquisition, signal interference from body movement or high-frequency noise is common, causing some signals to be affected by noise and rendering them unreliable for diagnostic purposes. Although the MIT-BIH database contains a few ECG signals affected by noise, they are still labeled as normal heartbeats. By contrast, the NCKU-CBIC database addresses this issue, providing improvements that align with the diagnostic needs of cardiologists. A comparison of the disease annotations by the two databases is illustrated in Fig. 3. The different types of disease labels and their statistics are shown in Table I.

Given the diversity of cardiac arrhythmias, making comparisons across different studies can be challenging. Therefore, this study adopts the standards established by the Association for the Advancement of Medical Instrumentation [10]. These standards categorize heartbeats into normal beats, supraventricular ectopic beats (SVEB), ventricular ectopic beats (VEB), fusion beats, and unknown beats in alignment with regulatory requirements for medical devices. The conversion table for MIT-BIH and NCKU-CBIC is shown in Table II. The model's classification of diseases is unlimited, allowing it to encompass dozens of arrhythmia and making it applicable to real-world scenarios.

TABLE II
MIT-BIH AND NCKU-CBIC CONVERSION INTO THE AAMI STANDARD

AAMI	MIT-BIH	NCKU-CBIC
Normal (N)	N、LBBB、RBBB、e、j	0
Supraventricular ectopic beat (S)	A、a、J、S	1、2、4、8
Ventricular ectopic beat (V)	V、E	3
Fusion beat (F)	F	-
Unknown beat (Q)	/、f、Q	5、6

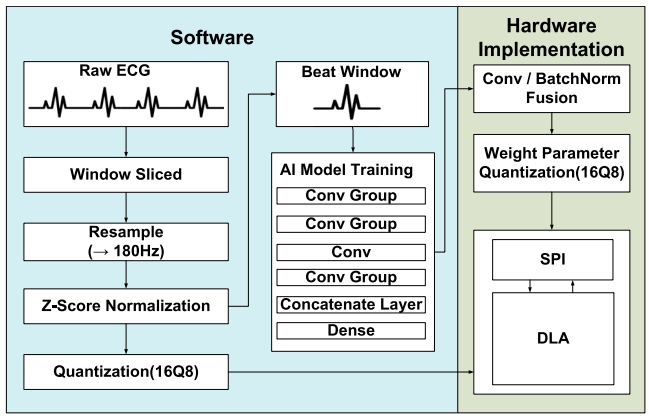


Fig. 4. Arrhythmia classification system for software and hardware implementation.

III. PROPOSED ARRHYTHMIA CLASSIFICATION ALGORITHM

Fig. 4 illustrates the architecture of the arrhythmia classification system we proposed [11], which includes software and hardware implementation. This section focuses on explaining the software algorithm components. First, we use annotation files provided by the MIT-BIH and NCKU-CBIC databases to locate the positions of R-peaks, as this work focuses on hardware-software co-design. The R-peak location algorithm is another important research topic, and there are several methods to achieve it, such as the Pan-Tompkins algorithm [12] and the Hamilton algorithm [13]. Many signal processing techniques, such as the discrete wavelet transform [14], phasor transform [15], and Hilbert transform [16], can also be used. After continuous ECG signals are segmented and mapped into fixed-size windows using our segmentation strategy, signal resampling and normalization are applied, completing the data preprocessing stage. Next, the AI model architecture is designed by utilizing a CNN as the foundation. Auxiliary features are incorporated to address insufficient temporal information, and the model undergoes iterative training until satisfactory classification performance is achieved. Finally, ECG data and model parameters are quantized for utilization in hardware implementation. The algorithm flowchart for preprocessing and AI model design is illustrated in Fig. 5.

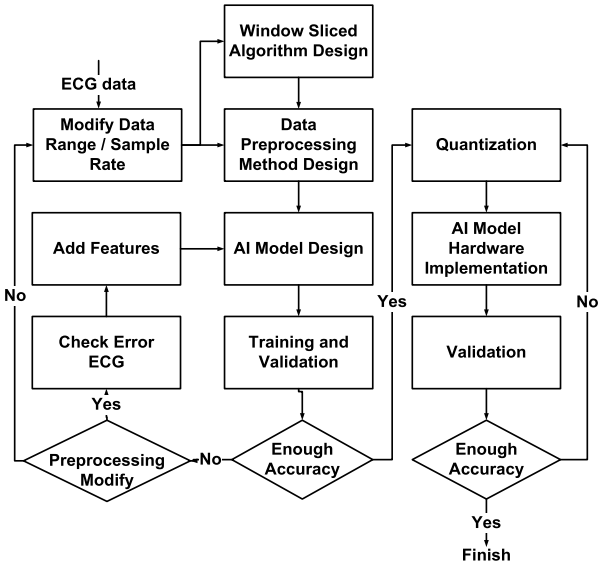


Fig. 5. Algorithm flowchart for preprocessing and AI model design.

A. Proposed General ECG Preprocessing Method

Given the fixed input size of the AI model, segmenting continuous ECG signals is necessary. A common approach in current research is to perform fixed window segmentation with the R-peak as the center and a fixed range selected before and after to isolate a heartbeat [17]. However, individual patients may have different heart rhythms. This variation can lead to situations wherein multiple heartbeats are included within the window, posing a challenge in determining the segmentation range, as shown in Fig. 6(a). Reference [18] sets the midpoint between adjacent R-peaks as the boundary for each heartbeat, ensuring that the segmented range contains only one heartbeat, to address the above issue. This approach is referred to as dynamic window segmentation. After experimentation, this work adjusts the boundary range to be 2/3 of the R-peak interval because using the midpoint of the R-peak results in insufficient data selection and a lack of features, hence leading to poor model training effectiveness. The results of the comparison of different segmentation ranges are shown in Fig. 7.

Dynamic window segmentation results in varying amounts of data for each segmented heartbeat. Signals are further mapped to fixed-size windows, a process referred to as alignment, to standardize data. In this study, the chosen window size is equal to the sampling frequency of the ECG data. For example, a window size of 360 is used for the MIT-BIH database. This method is adaptable to different databases, thus eliminating the need to redesign preprocessing algorithms due to database variations. Fig. 6(b) illustrates the processes of ECG segmentation and alignment.

Given that differences in sampling frequencies among various measurement devices lead to variations in mapped window sizes, either redesigning preprocessing methods or modifying the AI model architecture is necessary, making algorithm implementation challenging. In recent years, numerous studies have also adopted resampling [17], [19] to address this issue. In this work, signal resampling to 180 Hz is chosen because resampling

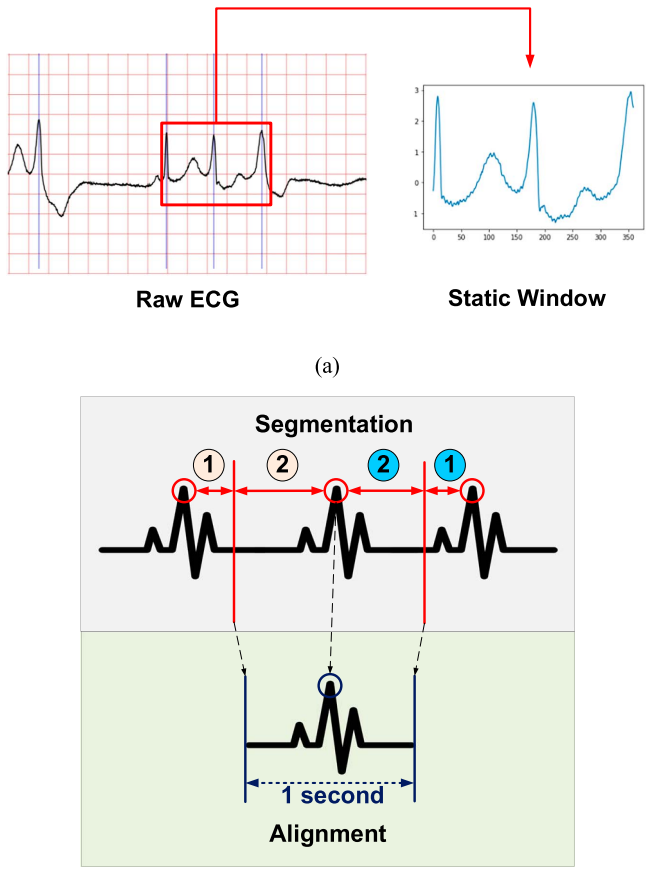


Fig. 6. (a) Multiple heartbeats in a static window. (b) ECG segmentation and alignment ensure that only one heartbeat is present in a window.

Fig. 7. Comparison of different R-peak interval segmentation ranges.

has been proven to be adaptable to various databases. Additionally, the input data are compressed, resulting in a reduction in model parameters and in a lightweight AI model conducive to hardware implementation. Finally, z-score normalization is applied to the data to mitigate data divergence and enhance model training effectiveness.

B. Time-Period CNN-Based Model Design

This work trains and tests the model using the MIT-BIH and NCKU-CBIC databases. We split the dataset as follows: Training set (80%), Testing set (20%), and Validation set (30% of the Training set) during training and adopted 5-fold cross-validation (CV) to ensure reliable training and testing results. Also, the data imbalance issue has been considered, and a simple oversampling technique has been adopted for data augmentation in

TABLE III
MIT-BIH AND NCKU-CBIC TESTING DATA STATISTICS

Class	MIT-BIH	NCKU-CBIC
N	18 046	13 238
S	554	105
V	5772	1988
F	161	-
Q	1601	1582

this work. The distribution of data for each of the five diseases is evenly split, as shown in Table III. Given the emphasis on hardware implementation and considering limited hardware resources, the design focuses on reducing model parameters, computational complexity, and memory storage requirements. Therefore, an ultralightweight AI model is proposed. Fig. 8 illustrates the ultralightweight time-period CNN-based model proposed in this work. In traditional CNN architectures [20], parameters in the fully connected layers typically dominate the total count. This work decreases the depth and number of nodes in the fully connected layers to reduce model size. However, this approach may decrease classification performance and accuracy. In response, this work increases the depth of convolutional layers because convolutional layers excel at feature extraction, particularly in the context of continuous signals, such as ECG. Compared to [20], which contains 3 convolutional layers, 3 fully connected layers, and additional layers, this approach enhances classification performance by extracting important features and, in line with lightweight requirements, reduces the kernel size in convolution layers. Thanks to this strategy, the model maintains accuracy while achieving ultralightweight. The final model architecture comprises four convolution layers and two fully connected layers. This adjustment results in fewer model parameters, making it more suitable for low-power wearable device design. Additionally, the experimental results of different model layer configurations are provided in Table IV. It employs max pooling layers to filter features and reduce the number of feature maps. Additionally, batch normalization layers [21] are incorporated to prevent data divergence and contribute to the training effectiveness of the model. This work also includes a global average pooling layer, which connects convolution layers with fully connected layers. Recent research indicates that this layer contributes to improved classification by filtering out unimportant information [22]. Moreover, the GAP layer aids in compressing a considerable amount of node data, contributing to the lightweight design of the model. The proposed model in this work contains only 921 parameters.

The classification categories of this model are based on the Advancement of Medical Instrumentation (AAMI) standard. In the early stages of training experiments, two categories, SVEB and VEB, exhibited poor training performance, and related research results also indicated the difficulty in recognizing diseases of these two kinds [2]. This study aims to investigate misclassified heartbeats and identify a specific issue. The characteristics of these diseases involve irregular heart rhythms, as illustrated in Fig. 9, which shows consecutive heartbeats with symptoms of atrial premature contraction. Individual heartbeats within a single window may appear normal. However, given that the fixed input size of the CNN model is a single window, the AI model misclassifies this waveform as a normal heartbeat. The limitation of the CNN model's fixed input size prevents the effective capture of long-term temporal information, leading to challenges in accurately diagnosing certain diseases.

The simplest approach to address the above issue would be to expand the input data range to include consecutive heartbeats. However, this method would remarkably increase the size of the model, adding a large number of model parameters and increasing computational complexity, making it challenging for hardware implementation. Therefore, this work proposes adding R-peak intervals as additional features that provide long-term rhythm information. Only the rhythm around a single heartbeat is considered, allowing the model to receive minimal efficient data. After the experiment, this study selects 10 R-peak intervals before and after a single heartbeat and concatenates them to fully connected layers for simultaneous training. There are 3 reasons: First, our design referenced the standard ECG strip, which is typically 8-10 seconds long in real-world practice. This length is comfortable and reasonable for doctors or technicians to classify arrhythmias at a glance. The system latency is also acceptable and meets the requirements for daily ECG monitoring scenarios, as it only takes a few seconds ($<$ 15 seconds). Second, for atrial premature contraction (APC), doctors must examine the preceding and following beats in real-life scenarios. Choosing only 10 previous R-R intervals can be problematic. In this case, the AI model may detect the APC, but the beat could be mis-positioned. Third, since R-R values vary from person to person, using 10 R-R intervals helps the AI model learn the normal R-R interval values and distinguish them from the abnormal R-R intervals that occur during an APC event. Those 10 R-peak intervals are 10 values that represent long-term ECG information, obtained by calculating 11 consecutive R-peak positions. After testing several combinations, we determined that adding 3 additional nodes provided the best balance between software and hardware performance, resulting in an increase of 30 parameters. This technique also effectively improves classification performance, particularly for SVEB and VEB types. Finally, it aligns with the design philosophy of ultralightweight models, making it very suitable for hardware implementation.

Table V presents a comparison of the model's classification performance on the MIT-BIH dataset using common evaluation metrics: accuracy (Acc.), sensitivity (Sen.), specificity (Spe.), and positive predictive value (PPV). The experimental results indicate a remarkable improvement in sensitivity for the classification of SVEB and VEB after the incorporation of long-term information. The model achieves an accuracy of 98.32% on the MIT-BIH database along with a sensitivity of 84.16% for SVEB and 94.50% for VEB under 5-fold CV. The lightweight model achieves an accuracy of 97.1% on the CBIC database, demonstrating its effectiveness.

A. *Hardware–Software Codesign*

This study implements the AI model in hardware, optimizing hardware performance through hardware-software codesign.

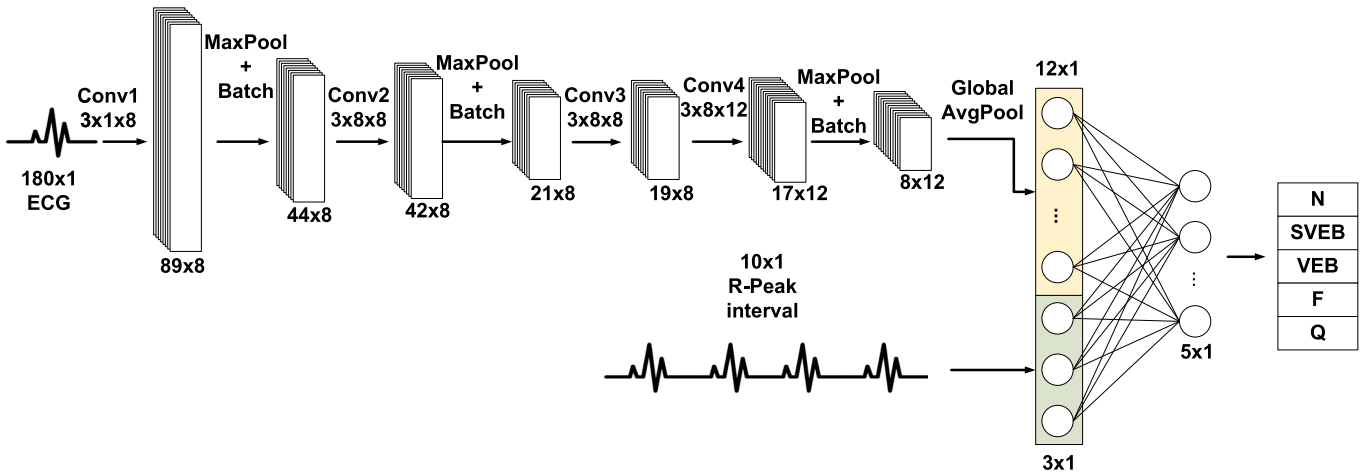


Fig. 8. Structure of the proposed ultralightweight time-period CNN-based model.

TABLE IV
RESULT OF DIFFERENT MODEL LAYER CONFIGURATION

Conv Layer[1]	Transition Layer	FC Layer[2]	Model Accuracy[3]	Parameter Number
Conv1(5,2,16) Conv2(5,2,32)	Flatten	FC1(320) FC2(10) FC3(5)	97% 96.2%	5905
Conv1(5,2,16) Conv2(5,2,16) Conv3(3,1,20)	Flatten	FC1(40) FC2(10) FC3(5)	96.6% 95.4%	2785
Conv1(5,2,16) Conv2(5,2,16) Conv3(3,1,20)	Flatten	FC1(40) FC2(5)	95.7% 95%	2525
Conv1(3,2,16) Conv2(3,2,16) Conv3(3,2,16) Conv4(3,1,20)	Flatten	FC1(20) FC2(5)	96.2% 95.8%	2649
Conv1(3,2,10) Conv2(3,1,10) Conv3(3,1,10) Conv4(3,1,15)	Flatten	FC1(120) FC2(5)	97.5% 96.4%	1825
Conv1(3,2,10) Conv2(3,1,10) Conv3(3,1,10) Conv4(3,1,15)	Global Average Pooling	FC1(15) FC2(5)	97.1% 96.3%	1300
Conv1(3,2,8) Conv2(3,1,8) Conv3(3,1,8) Conv4(3,1,12)	**Global Average Pooling**	**FC1(12) FC2(5)**	**97% 96.1%**	**873**

[1]Conv(kernel size, stride, kernel number).
[2]FC(neuron number).
[3]Accuracy of MIT-BIH and NCKU-CBIC databases.

Convolutional and batch normalization layers are fused. The convolutional layer operation involves multiplying weights (w) by input data (x), and the result is represented as x'. The formula for this computation is as follows:

$$x' = wx. \tag{1}$$

The batch normalization layer follows the convolutional layer, where x' is the input, and y is the result. The operational formula is as follows:

$$\begin{aligned} y &= \gamma \frac{x' - \mu_\beta}{\sqrt{{\sigma_\beta}^2 + \varepsilon}} + \beta, \\ y &= \frac{\gamma}{\sqrt{{\sigma_\beta}^2 + \varepsilon}}\ x' + \left(\beta - \frac{\gamma \mu_\beta}{\sqrt{{\sigma_\beta}^2 + \varepsilon}}\right), \end{aligned} \tag{2}$$

where γ and β represent the scaling factor and bias, respectively. μ_β is denoted as the mean value, and σ_β is the variance. The following formula results from combining and simplifying Equations (1) and (2):

$$y = ax + b \begin{cases} a = \dfrac{\gamma w}{\sqrt{{\sigma_\beta}^2 + \varepsilon}} \\ b = \left(\beta - \dfrac{\gamma \mu_\beta}{\sqrt{{\sigma_\beta}^2 + \varepsilon}}\right) \end{cases}, \tag{3}$$

where a represents the newly merged weights, and b is the newly merged bias. Through the above formula, the parameters of the batch normalization layer are integrated with the convolution layer, effectively reducing the overall parameter count. This fusion results in the reduction of three additional computational layers during hardware execution, thereby optimizing hardware efficiency.

This work quantifies model parameters and ECG data into 16Q8 fixed-point numbers, hence reducing storage space, to reduce hardware resources. Additionally, fixed-point arithmetic is implemented to reduce power consumption.

IV. PROPOSED AI ACCELERATOR HARDWARE DESIGN

The demand for high-speed computation has drastically increased with the rapid development of AI in recent years. Given the large requirements for data storage and highly repetitive model calculations, traditional CPUs are unsuitable for AI models. The current mainstream trend is to design customized hardware architectures based on specific models. This research focuses on practical applications in the field of clinical medicine in consideration of the requirements of biomedical wearable devices. Such devices must be compact, lightweight, and energy efficient to be suitable for patients to carry around for extended

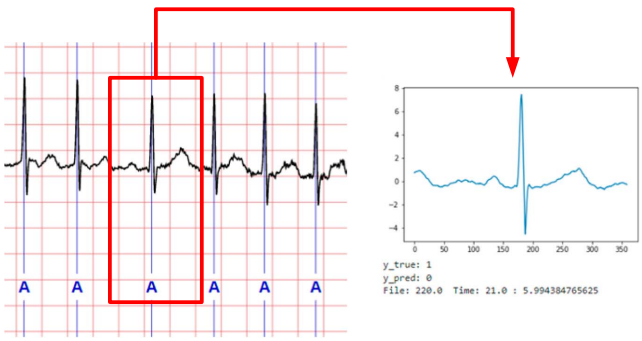


Fig. 9. Segmenting consecutive atrial premature contractions into individual windows leads to misclassification by the model.

TABLE V
COMPARISON OF MODEL PERFORMANCES ON MIT-BIH

	CNN		CNN with R-Peak Intervals	
	SVEB	VEB	SVEB	VEB
Acc.	99.04%	98.1%	99.25%	99.39%
Sen.	75.3%	92.2%	84.16%	94.50%
Spe.	99.6%	99.8%	99.65%	99.74%
PPV	79.4%	99.3%	86.32%	99.28%

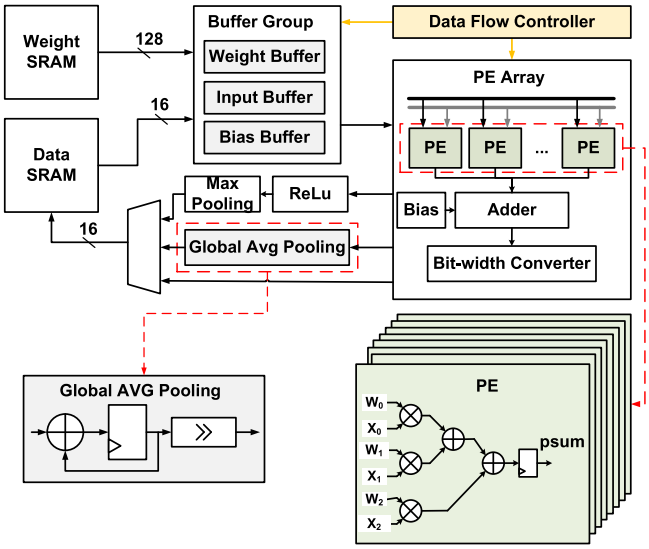


Fig. 10. Proposed AI accelerator architecture and peripheral circuits.

periods of measurement. Therefore, the hardware design emphasizes low power consumption and resource efficiency.

This work has designed a customized AI accelerator based on the proposed AI model to enhance hardware inference efficiency. Fig. 10 illustrates the hardware architecture, which comprises two static random-access memories (SRAMs). The weight SRAM stores the model parameters, and weights from different channels within the kernel are merged and expanded to a width of 128 bits to improve throughput. The data SRAM stores 16-bit ECG data as well as the results of feature maps computed at each layer. This lightweight model requires only 4 KB of storage space, remarkably conserving hardware resources and making it highly suitable for hardware implementation.

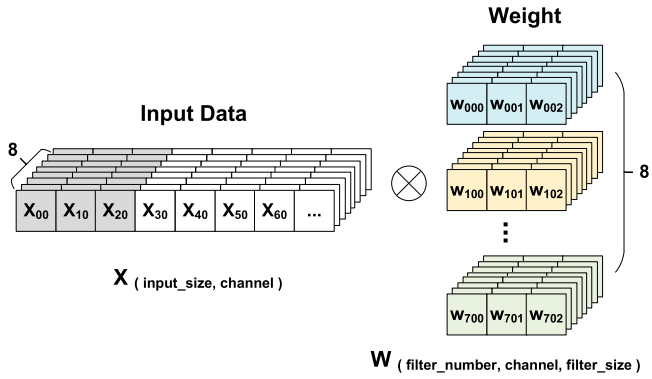


Fig. 11. Analysis of convolution layer behavior.

This design employs a data flow controller to transmit data to the corresponding data buffer, which is then processed by processing elements (PEs). The architecture introduces a parallelized PE array to handle operations related to convolution layers and fully connected layers. It also incorporates a specially designed data flow to implement a hybrid stationary mode, enhancing computational efficiency. Peripheral circuits handle pooling layer operations. In this approach, global average pooling (GAP), which typically requires a division circuit for average value computation, is efficiently implemented by using a shifter due to the cooperative design of the hardware and software in this work. The convolution layer's output size is 8 (number of feature maps) × 12 (channels), resulting in a data quantity of a power of 2, allowing hardware realization through shifters. This approach considerably reduces hardware resources, requiring only simple accumulators and shifters to achieve the GAP layer.

A. Convolution Layer Implementation

The AI model proposed in this work is based on CNN and involves an increase in the number and depth of convolution layers. The hardware is dominated by convolution operations. Therefore, hardware implementation is specifically optimized for convolution. Fig. 11 illustrates the analysis of convolution layer behavior, wherein the computation results at length coordinate l in the output feature map y is given by

$$y[l][k] = \sum_{k=0}^{K-1} \sum_{j=0}^{C-1} \sum_{i=0}^{L-1} (x[l+i][j] \times w[k][j][i]), \tag{4}$$

where w *and* x represent the kernel and input feature map, respectively. L denotes the kernel length, C denotes the number of feature map channels, and K denotes the number of kernels.

The proposed AI model has uniform kernel sizes based on hardware and software co-design. This characteristic is advantageous for hardware implementation. Each PE consists of three multiply–accumulators corresponding to the kernel size. The hardware design includes eight PEs, which correspond to the number of channels in the model's kernels. As observed in Fig. 11, multiple sets of kernels are present in the convolution layer. Therefore, the design concept is to map the PE array to a set of multichannel kernels by adopting a parallelized architecture

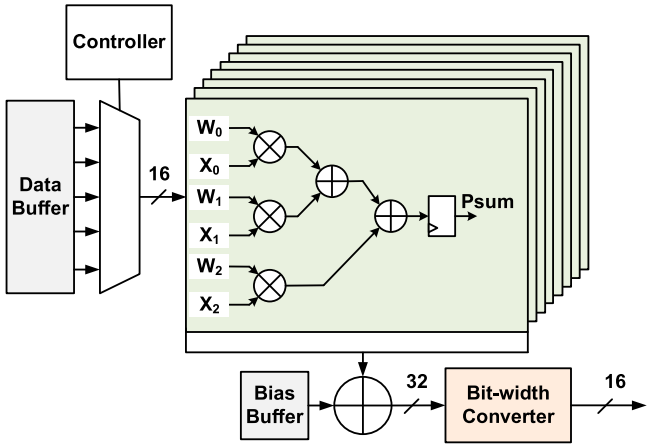


Fig. 12. Proposed parallel PE array architecture.

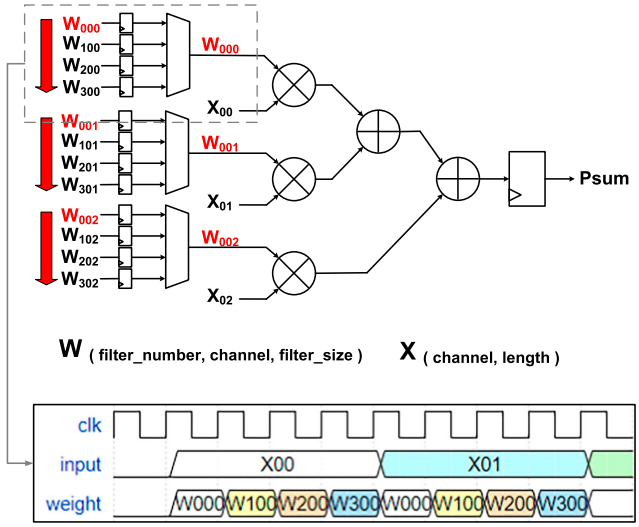


Fig. 13. Data flow in hybrid stationary mode on the proposed PE for convolution.

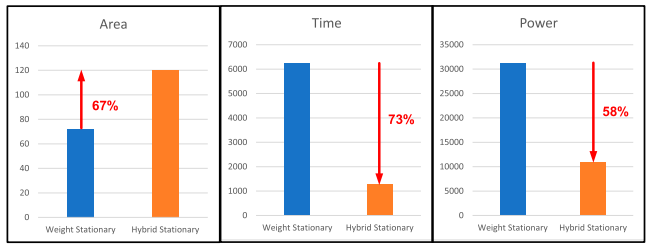


Fig. 14. Hardware performance analysis of different stationary modes.

to complete the computations for this kernel within one cycle, thereby increasing throughput.

Fig. 12 illustrates the parallel PE array architecture. Through the controller's selection of input data and weights, each PE performs multiply-accumulate operations, yielding a 32-bit computation result. The results are then accumulated, added with biases, and finally quantized into 16-bit by using a bit converter before being stored back in SRAM. Furthermore, common weight stationary techniques [23] are employed, fixing weights while performing computations with different input data to enhance weight reuse efficiency.

However, convolution layers consist of multiple sets of multichannel kernels. As a result of limited hardware resources, the PE array can only handle one set of kernel operations at a time. This condition implies that the hardware must sequentially compute each kernel and repeatedly read input data and weights from SRAM. This process results in a remarkable amount of SRAM data access, especially considering the need for the repeated reading and computation of input data, leading to additional power consumption and decreasing computational speed. Although adopting weight stationary enhances weight reuse efficiency, the increasing number of channels makes accommodating all weights challenging for the hardware. This condition, in turn, leads to poor input reuse efficiency due to the repeated computation of input data. Although the expansion of the number of PEs can address this issue, it contradicts the requirements of wearable devices.

Therefore, this study has designed a specialized data buffer that combines a 4-to-1 multiplexer and incorporates a unique data flow design on the PEs, as illustrated in Fig. 13, to address the above problem. Initially, it extracts weights from four different kernels and one input feature map set while maintaining the weight stationary mode. It computes the result for the first channel in the output feature map. Subsequently, it fixes the input feature map and switches weights with the controller to calculate the result for the second channel in the output feature map. This process is repeated until all four sets of weights have been computed, and then the controller shifts the input feature map to start a new iteration. This approach maps one input data set to four sets of weights, thereby improving data reuse efficiency. During computation, the hardware switches between input/weight stationary modes. This process is referred to as the hybrid stationary mode.

The hybrid stationary mode reduces the number of SRAM data accesses, thereby enhancing computational speed. According to relevant research [24], a considerable portion of hardware power consumption arises from memory access. Therefore, the hybrid stationary mode effectively reduces power consumption, hence aligning with the requirements of wearable devices. This work further analyzes the advantages and disadvantages of the hybrid stationary mode, as illustrated in Fig. 14. Assuming the PE array architecture is the same, convolutional layer operations are performed by using the weight and hybrid stationary modes. Important metrics, such as area, speed, and power consumption, are compared. Given that the hybrid stationary mode requires an additional data buffer for completion, it incurs an extra 67% register overhead. However, it reduces memory access times by 73% and decreases data access power consumption by 58%. This analysis demonstrates that the proposed hybrid stationary mode considerably improves hardware efficiency due to the lightweight model design, which is achievable with only 120 registers. The proposed mode presents a lightweight and low-power AI accelerator design.

B. *Implementation of the Fully Connected Layer*

The computation of the fully connected layer is performed by the parallel PE array, thus enhancing hardware utilization.

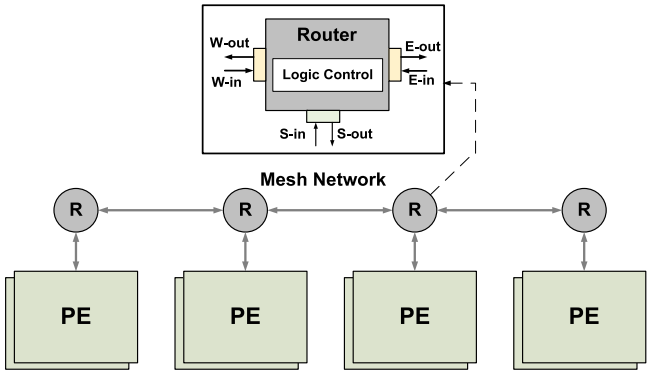


Fig. 15. 1-D mesh network for the PE array.

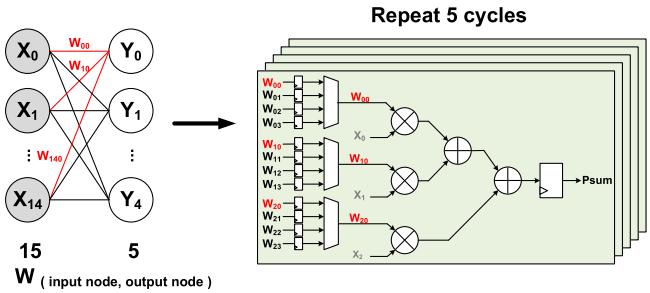


Fig. 16. Input stationary mode for the fully connected layer on PE.

In this work, a 1-D mesh network is designed, as illustrated in Fig. 15. This architecture allows for a bidirectional flow of input data and weights, with the controller selecting whether data are transmitted in parallel to multiple PEs for computation or fixed within a single PE. This flexible structure supports various stationary modes of operation. In the convolution layer, a hybrid stationary mode is adopted, alternating between input and weight processing. However, in the fully connected layer, an input stationary mode is utilized. The computational formula for this layer is

$$y_j = \sum_{j=1}^{m} \left(\left(\sum_{i=1}^{n} x_i \times w_{ij} \right) + b \right), \tag{5}$$

where y and x represent the output and input nodes, respectively. w denotes the weight, and b denotes the bias. m is the number of output nodes, and n is the number of input nodes.

The above formula shows that in the fully connected layer, a substantial number of weight calculations are performed for each input node. In this layer's computation, the hardware adopts the input stationary mode. As a result of the lightweight model design, the number of nodes in the fully connected layer is small enough for the PE array to accommodate all data. Therefore, a parallel node [25] computation approach is chosen. All input nodes are first fixed within the PEs, and one output node value is computed at a time. Subsequently, the weight data are refreshed for repeated computations until all output nodes are processed. Fig. 16 illustrates the data flow in the fully connected layer computation. Fifteen input nodes are initially stored and fixed in the PEs by utilizing a predesigned data buffer. The remaining weights can be stored in additional registers. Five PEs are then activated for multiply-accumulate operations, and the controller switches the weights in a repeating cycle of five periods to complete the computation. The computation in the fully connected layer is straightforward, leveraging the PE array hardware architecture and a customized data buffer. This approach enhances hardware utilization, and the data flow controller switches to the input stationary mode, thereby increasing data reuse efficiency.

TABLE VI
CONFUSION MATRIX FOR MIT-BIH & NCKU-CBIC

MIT-BIH					
	N[1]	S	V	F	Q
N[2]	17934	38	20	39	15
S	72	467	13	2	0
V	53	18	1340	32	0
F	16	2	5	128	0
Q	30	0	2	0	1569
NCKU-CBIC					
	N	S	V	F	Q
N	13191	1	0	0	46
S	13	92	0	0	0
V	13	8	1812	4	46
F	0	0	0	0	0
Q	320	6	27	6	1223

[1]Model classification result.
[2]Ground truth.

V. SYSTEM IMPLEMENTATION AND RESULTS

A. Software Simulation Results

This work proposes an ultralightweight time-period CNN-based model, which is trained and tested by using the TensorFlow [26] module. The model is validated by using the MIT-BIH and NCKU-CBIC databases. The classification confusion matrix results are presented in Table VI. The model achieves accuracies of 98.32% with 5-fold CV and 97.1% on the MIT-BIH and NCKU-CBIC databases, respectively.

This work incorporates the R-peak intervals in place of prolonged heart rhythm information, thus enhancing the classification of VEB and SVEB and reducing model misjudgments. Compared with recently reported models, as shown in Table VII, the proposed model demonstrates excellent performance in disease classification, proving its efficacy in classification. Additionally, as illustrated in Table VIII, the model in this study is extremely lightweight compared with recently reported models. Despite having the fewest parameters among all of the investigated models, this model achieves comparable classification performance, making it suitable for hardware implementation.

B. Web-Based Diagnostic Assistance System

This work implements a web-based assistance diagnostic system, as illustrated in Fig. 17. It designs a web-based ECG waveform display interface wherein captured ECG data can be

TABLE VII
COMPARISON OF THE PERFORMANCES OF THE PROPOSED MODEL AND PREVIOUSLY REPORTED MODELS FOR VEB AND SVEB CLASSIFICATION

	VEB				SVEB			
	Acc	Sen	Spe	Ppr	Acc	Sen	Spe	Ppr
ACCESS'18 [27]	98.6	93.8	99.2	92.4	97.5	76.8	98.7	74
JBHI'19 [18]	N/A	88.8	86.7	N/A	N/A	69.7	86.7	N/A
JBHI'22 [28]	99.7	97.2	99.8	97.5	99.6	86.6	99.9	97.3
TBME'22 [29]	99.04	91.8	99.4	95.7	98	61.7	99.6	94.8
ACCESS'22 [31]	N/A	97.5	N/A	97.88	N/A	92.49	N/A	92.49
ACCESS'23 [32]	N/A	95	N/A	92	N/A	83	N/A	69
This work (5-fold CV)	**99.39**	**94.50**	**99.74**	**96.28**	**99.25**	**84.16**	**99.65**	**86.32**

TABLE VIII
COMPARISON OF STATE-OF-THE-ART AI MODEL DESIGNS

	TENCON '19[30]	ACCESS'18 [27]	ACCESS'22 [31]	ACCESS'23 [32]	**This Work**
Model	Classic MLs	2D-CNN	STFT-CNN	H-PSO-CNN	**Time period CNN**
Parameters	-	360 750	14,741	1 112 869	**921**
Database	MIT-BIH	MIT-BIH	MIT-BIH	MIT-BIH	**MIT-BIH**
Classes	5	5	5	5	**5**
Accuracy	96%	96.05%	99%	98%	**98.32%**

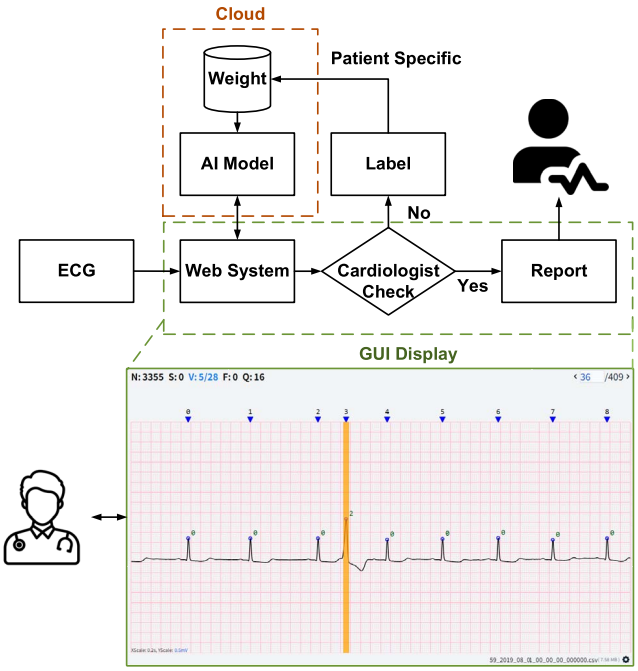
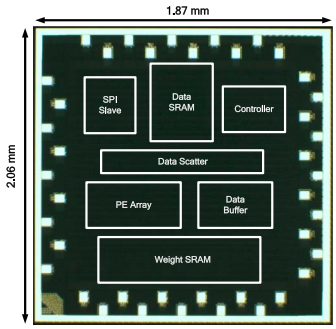


Fig. 17. Structure of the web-based diagnostic assistance system.

uploaded to the web platform for the real-time visualization of ECG waveforms. The AI model proposed in this study is integrated with a cloud backend, enabling the automated execution of AI classification computations. The classification results are promptly presented for physicians and patients to review. By replacing manual review with AI models, the system considerably reduces the waiting time for diagnosis.



Chip Specification	
Techonology	TSMC 0.18 um CMOS
Supply Voltage	3.3 V / 1.8 V (Pad / Core)
Package	SB48
Area	1.87 x 2.06 mm^2
SRAM	4 KB
Frequency	1 MHz
Power Consumption	122 uW
Classification Latency	6.8 ms

Fig. 18. Chip specification and micrograph.

Furthermore, the system provides a disease annotation function. Given that some rare arrhythmia symptoms or patients may exhibit unique waveform patterns, cardiologists can adjust the incorrect classification results based on individual patient cases. The modified ECG signals are saved back to the cloud database and can be confirmed by the doctor to ensure their correctness. The modified label is stored in Amazon Web Services (AWS) S3, and several methods are used to provide personal data security.

The experiments show that the system utilizes cloud computing for AI model inference, processing approximately 1 hour of ECG data in 5–6 min. Through extrapolation based on this result, the system is estimated to be capable of generating a 24-hour ECG report in approximately 2 h. This process exhibits a considerable reduction in patient waiting time compared with the traditional diagnostic process, which typically takes 1–2 days. The application of the proposed system in the medical field enhances healthcare quality by expediting diagnosis.

C. Hardware Implementation Results

This work proposes AI accelerator hardware development. Fig. 18 displays a micrograph of the AI accelerator and chip specifications. This chip is implemented in the TSMC 0.18 μm CMOS process, operates at a supply voltage of 1.8 V, and includes a 4 KB SRAM. It operates at a frequency of 1 MHz with a power consumption of 122 μW, requiring only 6.8 ms per classification. In accordance with the estimated hardware

TABLE IX
COMPARISON OF HARDWARE DESIGNS FOR ARRHYTHMIA CLASSIFICATION

	ACCESS'19 [33]	OJOCAS'21 [34]	TCASI'22 [35]	**This Work**
Technology (nm)	40	40	40	**180**
Supply Voltage (V)	1.1	0.9	1.1	**1.8**
Frequency (MHz)	100	185	2	**1**
SRAM (KB)	186.2	52	48	**4**
Power (mW)	2.13	19.16	0.854	**0.122**
Energy Efficiency (μJ/classification)	2.78	0.477	3.93	**0.83**
Classes	5	2	5	**5**
Accuracy	96.06%	92%	98.99%	**98.32%**

processing time, only 30 s are required to process 1-hour ECG data, and 24-hour ECG data can be processed in just 10–15 minutes. This process represents a remarkable reduction in processing time compared with cloud computing on web platforms. Although diagnostic assistance systems effectively shorten waiting times, they still fall short of achieving real-time classification. Therefore, the design of hardware circuits aims for integration with sensor devices in the future, thus enabling the development of wearable devices for real-time classification.

Table IX illustrates a comparison of hardware efficiencies reported by studies focusing on the classification of arrhythmias and proposing hardware implementations. The table shows that the hardware designed in this work exhibits the lowest power consumption and requires the least storage space among all investigated models, making it particularly suitable for wearable devices. However, the models used in each work are different, and a comparison based solely on power consumption may be unfair. The calculation of energy efficiency provides a reflective measure of hardware performance. Considering process variations, the chip performance in this work surpasses that in other studies, validating the low-power, high-efficiency characteristics of the proposed architecture.

VI. CONCLUSION

This work introduces an arrhythmia classification system designed to assist in diagnosis. It proposes a naive and general ECG preprocessing flow that is easy to apply to various databases. Additionally, it presents an ultralightweight time-period CNN-based model that incorporates R-peak intervals, thus addressing the challenge of insufficient temporal information in CNN models. This enhancement improves the classification results of SVEBs and VEBs. In this work, in consideration of hardware implementation, the model parameters are compressed to only 921, achieving accuracies of 98.32% and 97.1% on the MIT-BIH and NCKU-CBIC databases, respectively. This work designs a web-based diagnostic assistance system with AI computation implemented through a cloud backend. It offers a graphical interface to display ECG waveforms, aiding patients and physicians in reviewing data and reducing the time required for examination. Finally, this study proposes a customized AI accelerator designed with a hybrid stationary mode to enhance computational efficiency. When implemented by using the TSMC 0.18 μm CMOS process, it achieves an energy efficiency of 0.83 μJ/classification, completing one classification in only 6.8 ms. This accomplishment represents a low-power, high-performance hardware solution, with the hope that someday it can be applied in real-time wearable devices.

ACKNOWLEDGMENT

The authors greatly appreciate the support of the Taiwan Semiconductor Research Institute and the National Science and Technology Council, Taiwan.

REFERENCES

[1] "The top 10 causes of death," World Health Organization, Dec. 9, 2020. Accessed Jun. 10, 2023. [Online]. Avaliable: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-ofdeath
[2] S. Kiranyaz, T. Ince, and M. Gabbouj, "Real-time patient-specific ECG classification by 1-D convolutional neural networks," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 3, pp. 664–75, Mar. 2016.
[3] M. Jogin, Mohana, M. S. Madhulika, G. D. Divya, R. K. Meghana, and S. Apoorva, "Feature extraction using convolution neural networks (CNN) and deep learning," in *Proc. IEEE Int. Conf. Recent Trends Electron., Inf. Commun. Technol. (RTEICT)*, 2018, pp. 2319–2323.
[4] M. Kachuee, S. Fazeli, and M. Sarrafzadeh, "ECG heartbeat classification: A deep transferable representation," in *Proc. IEEE Int. Conf. Health. Inf. (ICHI)*, 2018, pp. 443–444.
[5] J. Huang, B. Chen, B. Yao, and W. He, "ECG arrhythmia classification using STFT-based spectrogram and convolutional neural network," *IEEE Access*, vol. 7, pp. 92871–92880, 2019.
[6] L. Meng, K. Ge, Y. Song, D. Yang, and Z. Lin, "Long-term wearable electrocardiogram signal monitoring and analysis based on convolutional neural," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–11, 2021.
[7] 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, 2001.
[8] W.-C. Tseng, T.-S. Zhong, S.-Y. Lee, and J.-Y. Chen, "NCKU CBIC ECG database." *Figshare*, Sep. 26, 2023, doi: 10.6084/m9.figshare.23807286.v1.
[9] S.-Y. Lee, P.-W. Huang, M.-C. Liang, J.-H. Hong, and J.-Y. Chen, "Development of an arrhythmia monitoring system and human study," *IEEE Trans. Consum. Electron.*, vol. 64, no. 4, pp. 442–451, Nov. 2018.
[10] NSI/AAMI, "Testing and reporting performance results of cardiacrhythm and ST segment measurement algorithms," Assoc. Adv. Med. Instrum., no. EC57, 1998.
[11] S.-Y. Lee, W.-C. Tseng, and J.-Y. Chen, "An ultra-lightweight time period CNN based model with AI accelerator design for arrhythmia classification," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 2024.
[12] J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE Trans. Biomed. Eng.*, vol. BME-32, no. 3, pp. 230–236, Mar. 1985.
[13] P. S. Hamilton and W. J. Tompkins, "Quantitative investigation of QRS detection rules using the MIT/BIH arrhythmia database," *IEEE Trans. Biomed. Eng.*, vol. BME-33, no. 12, pp. 1157–1165, Dec. 1986.
[14] S. Sahoo et al., "Multiresolution wavelet transform based feature extraction and ECG classification to detect cardiac abnormalities," *Meas. J. Int. Meas. Confederation*, vol. 108, pp. 55–66, Oct. 2017.
[15] A. Martínez, R. Alcaraz, and J. J. Rieta, "Application of the phasor transform for automatic delineation of single-lead ECG fiducial points," *Physiol. Meas.*, vol. 31, no. 11, pp. 1467–1485, Nov. 2010.
[16] D. S. Benitez et al., "New QRS detection algorithm based on the Hilbert transform," in *Proc. Comput. Cardiol.*, 2000, pp. 379–382.

[17] M. Zubair, J. Kim, and C. Yoon, "An automated ECG beat classification system using convolutional neural networks," in *Proc. 6th Int. Conf. IT Converg. Secur. (ICITCS)*, 26 Sep. 2016, pp. 1–5.
[18] S.-S. Xu, M.-W. Mak, and C.-C. Cheung, "Towards end-to-end ECG classification with raw signal extraction and deep neural networks," *IEEE J. Biomed. Health Inform.*, vol. 23, no. 4, pp. 1574–1584, Jul. 2019.
[19] C. Varon, A. Caicedo, D. Testelmans, B. Buyse, and S. V. Huffel, "A novel algorithm for the automatic detection of sleep apnea from single-lead ECG," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 9, pp. 2269–2278, Sep. 2015.
[20] U. R. Acharya et al., "A deep convolutional neural network model to classify heartbeats," *Comput. Biol. Med.*, vol. 89, pp. 389–396, Oct. 2017.
[21] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. Int. Conf. Mach. Learn.*, PMLR, 2015, pp. 448–456.
[22] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, "Learning deep features for discriminative localization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2016, pp. 2921–2929.
[23] Y.-H. Chen, T. Krishna, J. S. Emer, and V. Sze, "Eyeriss: An energy-efficient reconfigurable accelerator for deep convolutional neural networks," *IEEE J. Solid-State Circuits*, vol. 52, no. 1, pp. 127–138, Jan. 2017.
[24] S. Han, J. Pool, J. Tran, and W. Dally, "Learning both weights and connections for efficient neural network," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 28, pp. 1135–1143, Dec. 2015.
[25] P. N. Whatmough, S. K. Lee, D. Brooks, and G. Y. Wei, "DNN engine: A 28-nm timing-error tolerant sparse deep neural network processor for IoT applications," *IEEE J. Solid-State Circuits*, vol. 53, no. 9, pp. 2722–2731, Sep. 2018.
[26] M. Abadi et al., "TensorFlow: Large-scale machine learning on heterogeneous distributed systems," 2016, *arXiv:1603.04467*.
[27] X. Zhai and C. Tin, "Automated ECG classification using dual heartbeat coupling based on convolutional neural network," *IEEE Access*, vol. 6, pp. 27465–27472, 2018.
[28] J. Xiao et al., "ULECGNet: An UltraLightweight end-to-end ECG classification neural network," *IEEE J. Biomed. Health Inform.*, vol. 26, no. 1, pp. 206–217, Jan. 2022.
[29] J. Malik, O. C. Devecioglu, S. Kiranyaz, T. Ince, and M. Gabbouj, "Real-time patient-specific ECG classification by 1D self-operational neural networks," *IEEE Trans. Biomed. Eng.*, vol. 69, no. 5, pp. 1788–1801, May 2022.
[30] A. Peimankar, M. J. Jajroodi, and S. Puthusserypady, "Automatic detection of cardiac arrhythmias using ensemble learning," in *Proc. IEEE Region 10 Conf. (TENCON)*, Kochi, India, 2019, pp. 383–388.
[31] M. M. Farag, "A self-contained STFT CNN for ECG classification and arrhythmia detection at the edge," *IEEE Access*, vol. 10, pp. 94469–94486, 2022.
[32] F. S. Baños, N. H. Romero, J. C. S. T. Mora, J. M. Marín, I. B. Vite, and G. E. A. Fuentes, "A novel hybrid model based on convolutional neural network with particle swarm optimization algorithm for classification of cardiac arrhythmias," *IEEE Access*, vol. 11, pp. 55515–55532, 2023.
[33] J. Wu, F. Li, Z. Chen, Y. Pu, and M. Zhan, "A neural network-based ECG classification processor with exploitation of heartbeat similarity," *IEEE Access*, vol. 7, pp. 172774–172782, 2019.
[34] Y.-C. Chuang, Y.-T. Chen, H.-T. Li, and A. Y. A. Wu, "An arbitrarily reconfigurable extreme learning machine inference engine for robust ECG anomaly detection," *IEEE Open J. Circuits Syst.*, vol. 2, pp. 196–209, 2021.
[35] J. Lu, D. Liu, X. Cheng, L. Wei, A. Hu, and X. Zou, "An efficient unstructured sparse convolutional neural network accelerator for wearable ECG classification device," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 69, no. 11, pp. 4572–4582, Nov. 2022.

Shuenn-Yuh Lee (Senior Member, IEEE) was born in Taichung, Taiwan, in 1966. He received the B.S. degree from the National Taiwan Ocean University, Keelung, Taiwan, in 1988, and the M.S. and Ph.D. degrees from the National Cheng Kung University, Tainan, Taiwan, in 1994 and 1999, respectively. He is currently a Distinguished Professor with the Department of Electrical Engineering, National Cheng Kung University, Tainan, Taiwan. He served as the Technical Program Chair (TPC) of the 2014/2015 International Symposium on Bioelectronics & Bioinformatics (ISBB), and the 2015 Taiwan and Japan Conference on Circuits and Systems (TJCAS). He served as a General Co-Chair of the 2022 IEEE Biomedical Circuits and Systems Conference (BioCAS) and the 2024 IEEE Asia Pacific Conference on Circuits and Systems, and Organization Chair of the 2022 IEEE Asia Solid-State Circuits Conference (ASSCC). From 2013 to 2016, he served as the Chairman of the IEEE Solid-State Circuits Society Tainan Chapter. From 2016 to 2017, he served as the Vice Chairman of the IEEE Tainan Section. From 2016 to 2023, he served as the Associate Editor of IEEE Transaction on Biomedical Circuits and Systems. His research interests include the design of analog and mixed-signal integrated circuits, biomedical circuits and systems, artificial intelligence integrated circuits, low-power, and low-voltage analog circuits, and RF front-end integrated circuits for wireless communications. Currently, he is a Member of the Circuits and Systems (CAS) Society, Solid-State Circuits Society (SSCS), and Medicine and Biology Society (EMBS) of IEEE. He is also a Member of IEICE.

Ming-Yueh Ku (Graduate Student Member, IEEE) was born in Taichung, Taiwan, in 1998. He received the B.S. degree from the National Cheng Kung University, Tainan, Taiwan, in 2020. He is currently working toward the Ph.D. degree with the Department of Electrical Engineering. His research interests include CMOS image sensor (CIS), computing-in-memory (CIM) circuit, mixed-mode circuit, digital control system, bio-signal acquisition system-on-chip (SoC), and the Internet of Things (IoT).

Wei-Cheng Tseng was born in Taoyuan, Taiwan, in 1999. He received the M.S. degree from the National Cheng Kung University, Tainan, Taiwan, in 2023. His research interests include AI algorithm, digital integrated circuits design, and arrhythmia classification.

Ju-Yi Chen was born in Tainan, Taiwan, in 1974. He received the M.S. degree from Chang Gung University, Taoyuan, Taiwan, in 1999, and the Ph.D. degree from the National Cheng Kung University, Tainan, Taiwan, in 2013. Since 2022, he has been a Professor with the Department of Internal Medicine, National Cheng Kung University. His research interests include the cardiovascular diseases, including arrhythmias, hypertension, arterial stiffness, and cardiac implantable electric devices.