IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received June 17, 2019, accepted July 5, 2019, date of publication July 11, 2019, date of current version July 26, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2928017

ECG Arrhythmia Classification Using STFT-Based Spectrogram and Convolutional Neural Network

JINGSHAN HUANG[1,2], BINQIANG CHEN[1,2], BIN YAO[1,2], AND WANGPENG HE[3] (Member, IEEE)

[1]School of Aerospace Engineering, Xiamen University, Xiamen 361005, China
[2]Shenzhen Research Institute of Xiamen University, Shenzhen 518000, China
[3]School of Aerospace Science and Technology, Xidian University, Xi'an 710071, China

Corresponding author: Bin Yao (yaobin@xmu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 51605403, in part by the Natural Science Foundation of Guangdong Province, China, under Grant 2015A030310010, and in part by the Natural Science Foundation of Fujian Province, China, under Grant 2016J01012.

ABSTRACT The classification of electrocardiogram (ECG) signals is very important for the automatic diagnosis of heart disease. Traditionally, it is divided into two steps, including the step of feature extraction and the step of pattern classification. Owing to recent advances in artificial intelligence, it has been demonstrated that deep neural network, which trained on a huge amount of data, can carry out the task of feature extraction directly from the data and recognize cardiac arrhythmias better than professional cardiologists. This paper proposes an ECG arrhythmia classification method using two-dimensional (2D) deep convolutional neural network (CNN). The time domain signals of ECG, belonging to five heart beat types including normal beat (NOR), left bundle branch block beat (LBB), right bundle branch block beat (RBB), premature ventricular contraction beat (PVC), and atrial premature contraction beat (APC), were first transformed into time-frequency spectrograms by short-time Fourier transform. Subsequently, the spectrograms of the five arrhythmia types were utilized as input to the 2D-CNN such that the ECG arrhythmia types were identified and classified finally. Using ECG recordings from the MIT-BIH arrhythmia database as the training and testing data, the classification results show that the proposed 2D-CNN model can reach an averaged accuracy of 99.00%. On the other hand, in order to achieve optimal classification performances, the model parameter optimization was investigated. It was found when the learning rate is 0.001 and the batch size parameter is 2500, the classifier achieved the highest accuracy and the lowest loss. We also compared the proposed 2D-CNN model with a conventional one-dimensional CNN model. Comparison results show that the 1D-CNN classifier can achieve an averaged accuracy of 90.93%. Therefore, it is validated that the proposed CNN classifier using ECG spectrograms as input can achieve improved classification accuracy without additional manual pre-processing of the ECG signals.

INDEX TERMS Electrocardiogram (ECG); arrhythmia classification; short-time Fourier transform (STFT); convolutional neural network; model parameter optimization.

I. INTRODUCTION

Cardiovascular disease is one of the major diseases that threaten human life. According to reports by the world health organization, cardiovascular diseases (CVDs) mortality ranks first in all causes of death today. Over 17.7 million people died from CVDs, which is an about 31 percentages of all deaths. More than 75 percentages of these deaths occurred in developing countries. What's more, the prevalence and mortality of cardiovascular disease (CVD) are still growing [1]. Therefore, regular monitoring of heart rhythm has become an increasingly important and necessary matter so as to manage and prevent the CVDs.

Arrhythmia is an important group of diseases in cardiovascular disease. Arrhythmia can occur on its own or with other cardiovascular diseases. The diagnosis of arrhythmia mainly depends on the ECG (electrocardiogram). ECG (electrocardiogram) is an important modern medical tool that records the process of cardiac excitability, transmission, and recovery. Automatic detection of irregular heart rhythms from ECG

The associate editor coordinating the review of this manuscript and approving it for publication was Yongqiang Cheng.

signals is a significant task for the automatic diagnosis of cardiovascular disease.

Traditionally, the classification of ECG signals usually needs to be divided into two steps, i.e., feature extraction and pattern classification. Studies on detection concentrate on the problem of determining heartbeat within the ECG data, and some technologies have been applied in heartbeat detection, including threshold-based methods [2], digital filter-based methods [3], [4], and wavelet transform (WT) [5]–[7] and so on. The pattern classification of detected ECG signals is another step. Wavelet transform is one of the commonly used methods for obtaining the features of ECG signals. Li used the wavelet packet decomposition (WPD) technique to obtain representative features for the detection of five different heartbeat types, and calculate entropy from decomposed coefficients with WPD [8]. Özbay computed detail and approximation wavelet coefficients of the ECG features to generate feature vectors [9]. Li proposed a novel method based on genetic algorithm-back propagation neural network (GA-BPNN) for classifying ECG signals with feature extraction using wavelet packet decomposition [10]. Other methods were also employed for extracting features from ECG signals. Elhaj extracted the linear and nonlinear features from the signals for automatic ECG beat recognition [11]. Yeh used a complex ECG feature extraction technique to classify and distinguish differences between normal heartbeats and abnormal heartbeats [12]. Alickovic proposed an autoregressive (AR) model in the feature extraction module for diagnosing heart diseases [13].

After the step of feature extraction of ECG signals, classification is traditionally conducted using state-of-the-art classifiers such as support vector machines (SVM), neural networks (NN), cluster analysis (CA), random forests (RF), optimum-path forest (OPF) and some other classification tools [14]. R.J. Martis automatically classified five types of ECG beats using HOS features (higher order cumulants) using three layer feed-forward Neural Network (NN) and Least Square- Support Vector Machine (LS-SVM) classifiers [15]. Varatharajan.R presented a big data classification approach using LDA with an enhanced SVM method for ECG signals in cloud computing [16]. Yun-Chi Yeh proposed a method of analyzing ECG signal to diagnose cardiac arrhythmias utilizing the cluster analysis (CA) method, which could accurately classify and distinguish the difference between normal heartbeats (NORM) and abnormal heartbeats [12]. Taiyong L proposed a method to classify ECG signals using wavelet packet entropy (WPE) and random forests (RF) following the Association for the Advancement of Medical Instrumentation (AAMI) recommendations and the inter-patient scheme [17]. VHC De Albuquerque introduced the Optimum-Path Forest (OPF) classifier to automatic arrhythmia detection in ECG patterns [18]. In our approach ,we have used SVM (Support Vector Machine) for classification purpose. P.Raman proposed an approach for classification of Heart Diseases based on ECG analysis using type2-fuzzy c-means (FCM) and SVM Methods [19]. P.Pławiak presentd an innovative research methodology that enables the efficient classification of cardiac disorders (17 classes) based on ECG signal analysis and an evolutionary-neural system [20].

In recent years, deep learning techniques have shown their outstanding performances in pattern recognition applications [21]. Owing to this, researchers and engineers have shifted their attentions on ECG classification studies based on deep learning related techniques. Many scholars have made a large amount of efforts on the study of ECG classification using deep learning techniques. In the literature, it have been reported that several methods are effective for ECG arrhythmia classification. Salloum proposed the using of recurrent neural networks (RNNs) to develop an effective solution to identification and authentication in electrocardiogram (ECG)-based biometrics [22]. Mostayed proposed a recurrent neural network which consists of two bi-directional long-short-term-memory layers to detect pathologies in 12-lead ECG signals [23]. Zhang proposed a novel patient-specific electrocardiogram classification algorithm based on the recurrent neural networks (RNN) to learn time correlation from ECG signal samples and to classify ECG beats with different heart rates [24]. Kiranyaz proposed a real-time patient-specific ECG classification approach based on the 1D convolutional neural networks, which can be solely used to classify long ECG records of patients [25]. Li proposed a 1D-CNN based method to realize the classification of 5 typical types of arrhythmia signals, i.e., normal, left bundle branch block, right bundle branch block, atrial premature contraction and ventricular premature contraction [26]. Yin proposed an ECG monitoring system integrated with the Impulse Radio Ultra Wideband (IR-UWB) radar based on CNN [27]. Jun proposed an effective ECG arrhythmia classification method using a deep two-dimensional convolutional neural network which recently shows outstanding performance in the field of pattern recognition [28]. Salem proposed an ECG arrhythmia classification method using transfer learning from 2D deep CNN features, and the method was applied in the identification and classification of four ECG patterns [29].

Studies mentioned above show that a deep neural network can automatically learn complex representative features directly from the data adaptively so that we can reduce excessive dependencies on manual feature extractions, and create end-to-end learning systems that take ECG signals as input and arrhythmia class prediction as output, while extracting the "deep features" automatically [29]. However, as compared to the traditional classification approaches, DNNs require a considerable amount of data as the training set. This problem creates a gap between the dataset size and deep features, since the datasets that are publicly available in this domain lack in volume [30].

In this paper, we propose an ECG arrhythmia classification method based on a deep two-dimensional convolutional neural network on classification of five different rhythms. The input one-dimensional ECG signals in the time domain signals are transformed into two-dimensional time-frequency

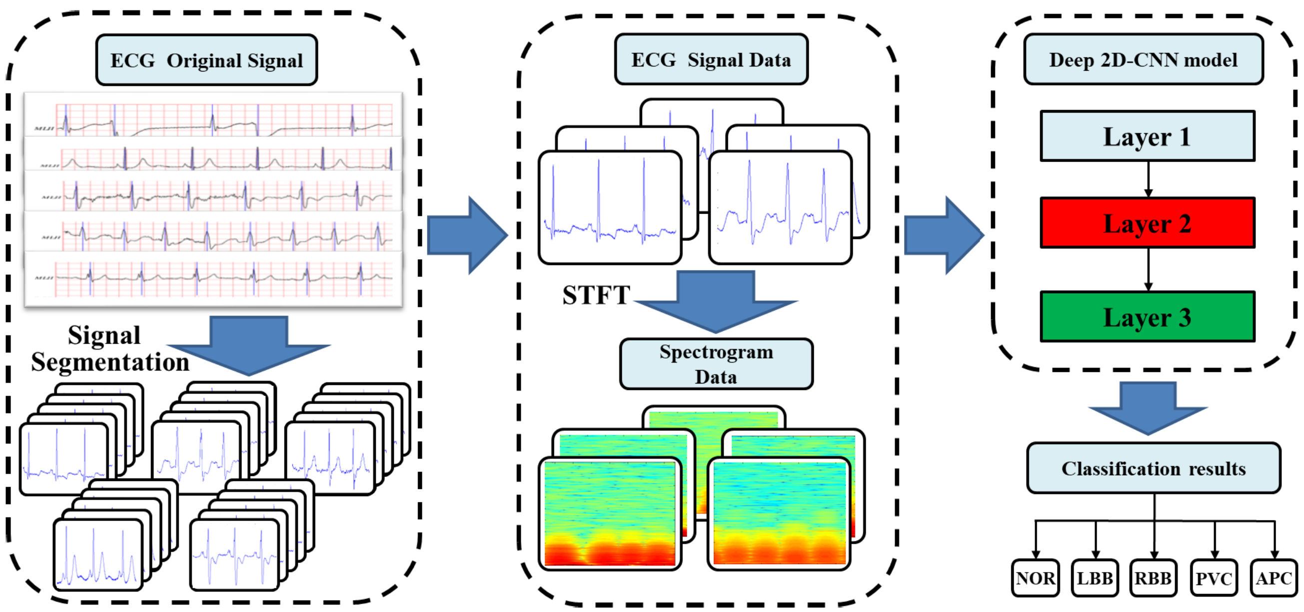


FIGURE 1. Overall procedures in ECG arrhythmia classification based on proposed 2D-CNN.

spectrograms. Nevertheless, in this step noise filtering and manual feature extraction are no longer required. In addition, training data are obtained through augmenting of the derived ECG images, which can result in higher classification accuracy. The segmented 2D time-frequency spectrograms are fed as input of the convolutional neural network. The 2D CNN model can automatically suppress the measurement noises and extract relevant feature maps throughout the convolutional and pooling layer intelligently. Thus, the propose method can be applied to the ECG signals from various ECG acquisition devices with different sampling rates, and it is beneficial in precise identifications of ECG arrhythmia [28]. In all, the proposed ECG arrhythmia classification method consists of the following steps, e.g., ECG signals data acquisition, ECG signals pre-processing, and 2D-CNN classifier.

The rest of this paper is organized as below. In Section 2, we explain the methodologies used for the ECG arrhythmia classification, including ECG data pre-processing, and convolutional neural network classifier. In Section 3, Numerical evaluation and experimental results of ECG arrhythmia classification are shown. In Section 4, we give some discussions and major conclusions of the paper.

II. METHODOLOGY

A. METHOD OVERVIEW

The overall procedures of the proposed ECG arrhythmia classification model is shown in Figure 1. The original ECG signals were shared by the MIT-BIH arrhythmia database [31]. The input ECG signals were divided into data recordings with an identical duration of 10 seconds. The one-dimensional ECG time domain signals, there are five different classes of arrhythmia, based on the recordings annotations which made by two or more cardiologists independently. Afterward, each ECG signal record is transformed into an image of time-frequency spectrogram by using the short time Fourier transform (STFT). The ECG spectrogram images are fed into the proposed deep two-dimensional convolutional neural network (CNN) model. With these obtained ECG spectrogram images, classification of the five ECG types is performed in the 2D-CNN classifier automatically and intelligently. The five ECG types are normal beat (NOR), left bundle branch block beat (LBB), right bundle branch block beat (RBB), premature ventricular contraction beat (PVC), atrial premature contraction beat (APC).

TABLE 1. The ECG data downloaded from MIT-BIH database.

Arrhythmia Type	*MIT-BIH*	*Training set*	*Testing set*
NOR	100,105,215	450	90
LBB	109,111,214	450	90
RBB	118,124,212	450	90
PVC	106,223	300	60
APC	207,209,232	450	90

B. DATA ACQUISITION AND SELECTION

In this subsection, we conducted experimental analysis in order to evaluate the performance of the proposed method. Five types of ECG signals, of which the sampling rate was uniformly set as 360 Hz, were obtained from the MIT-BIH arrhythmia database [31].

The original ECG signal data contains heartbeat information of patients with cardiovascular disease, and each record of the ECG signal contains 1 hour excerpts of two-channel ambulatory recordings. For each type of the ECG signal, a segment of 10 seconds were selected. Furthermore, the signals were divided into 2520 samples for ECG classification. The related information of the employed data from the MIT-BIH arrhythmia database is listed in Table 1.

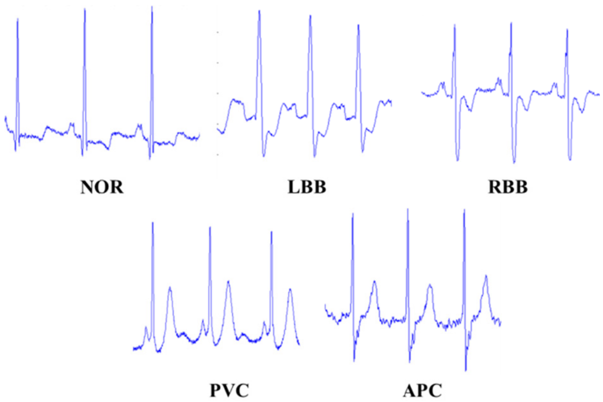


FIGURE 2. Waveforms (V5 lead) of normal beat and that of the other four ECG arrhythmia diseases.

Samples of NOR were obtained from records 100, 105 and 215. Samples of LBB were derived from records 109, 111 and 214. Samples of RBB were obtained from records 118, 124 and 212. Samples of APC were obtained from records 207, 209 and 232. For the above four types, Each type of ECG signals has 450 samples for the training set and 90 samples for the testing set. Samples of PVC were obtained from records 106 and 233. The type of PVC has 300 samples for the training set and 60 samples for the testing set. Waveforms of the ECG signals are shown in Figure 2.

C. ECG DATA PRE-PROCESSING

Within the proposed 2D-CNN, the input data is required to be of the type of image. Therefore, the ECG signals in the time domain, which belong to five heart disease types, were firstly transformed into 2D time-frequency spectrograms using short-time Fourier transform (STFT).

The ECG signal is a nonstationary data which instantaneous frequency varies according to the time. Therefore, the properties of the changes cannot be fully described by merely using information in the frequency domain. The STFT is an enhanced mathematical methodology, derived from the discrete Fourier transform (DFT), to explore the instantaneous frequency as well as the instantaneous amplitude of localized waves with time-varying characteristics.

In the analysis of a non-stationary signal, it is assumed that it is approximately stationary within the span of a temporal window of finite support [32]. For a discretized digital signal, its time-frequency spectrogram is given as

$$STFT\{x[n]\} = X(m, \omega) = \sum_{n=-\infty}^{\infty} x[n]w[n-m]e^{-j\omega n} \quad (1)$$

where $x[n]$ represents the ECG signal which sampling rate was 360 Hz, and $w[n]$ is the window function. In this proposed method, we adopt the Hanning window, whose definition is given as below. And the window size is 512.

$$w(n) = \begin{cases} 0.5\left[1-\cos\left(\frac{2\pi n}{M-1}\right)\right], & 0 \le n \le M-1 \\ 0, & otherwise \end{cases} \quad (2)$$

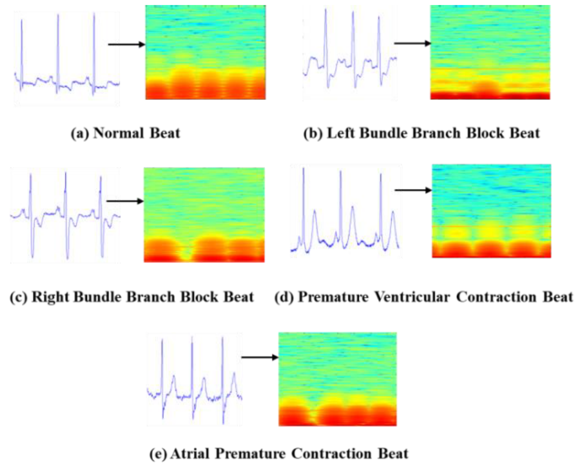


FIGURE 3. ECG spectrograms of a sample data instance belonging to each type of arrhythmia.

TABLE 2. The explanations for the applied functions in the proposed 2D-CNN model.

Function	*Explanations*
Convolution2D	Convolutional layer, sliding window convolution to 2-dimensional input information;
MaxPooling2D	Maximum pooling layer, imposing a maximum pooling on the spatial domain signal;
RELU	Rectified Linear Unit, which performs linear rectification activation on the input vector of the upper layer neural network and outputs nonlinear results.
Flatten	The Flatten layer is used to translate the multidimensional input information into one-dimensional information.
Dropout	It is an regularization layer to prevent overfitting;
Softmax	It is an activation function for multi-class neural network output.

Therefore, we transformed ECG time domain signals into ECG spectrums images by plotting each ECG data recording as an individual 256 × 256 pixel image. A sample of each class's spectrogram is shown in Figure 3.

D. ECG ARRHYTHMIA CLASSIFIER

In this paper, CNN is adopted as the ECG arrhythmia classifier. CNN was first introduced by LeCun [33] and was developed through a project to recognize handwritten zip codes. With the advent of the CNN model, correlation of spatially adjacent pixels can be extracted by applying a nonlinear filter and by applying multiple filters, and it is capable of extracting various local features of the image.

2D convolutional and pooling layers are more suitable for filtering the spatial locality of ECG images [28]. Therefore, to facilitate 2D-CNN in ECG signal classification, we convert ECG signals in the time domain into 2D spectrograms in time-frequency representations. The structure of the 2D-CNN is illustrated in Figure 4.

The explanations for the applied functions in the 2D-CNN model are shown in Table 2.

Firstly, each ECG data recording was transformed into an ECG spectrums images which size is 256 × 256 pixel. In the first hidden layer, the Convolution2D layer with

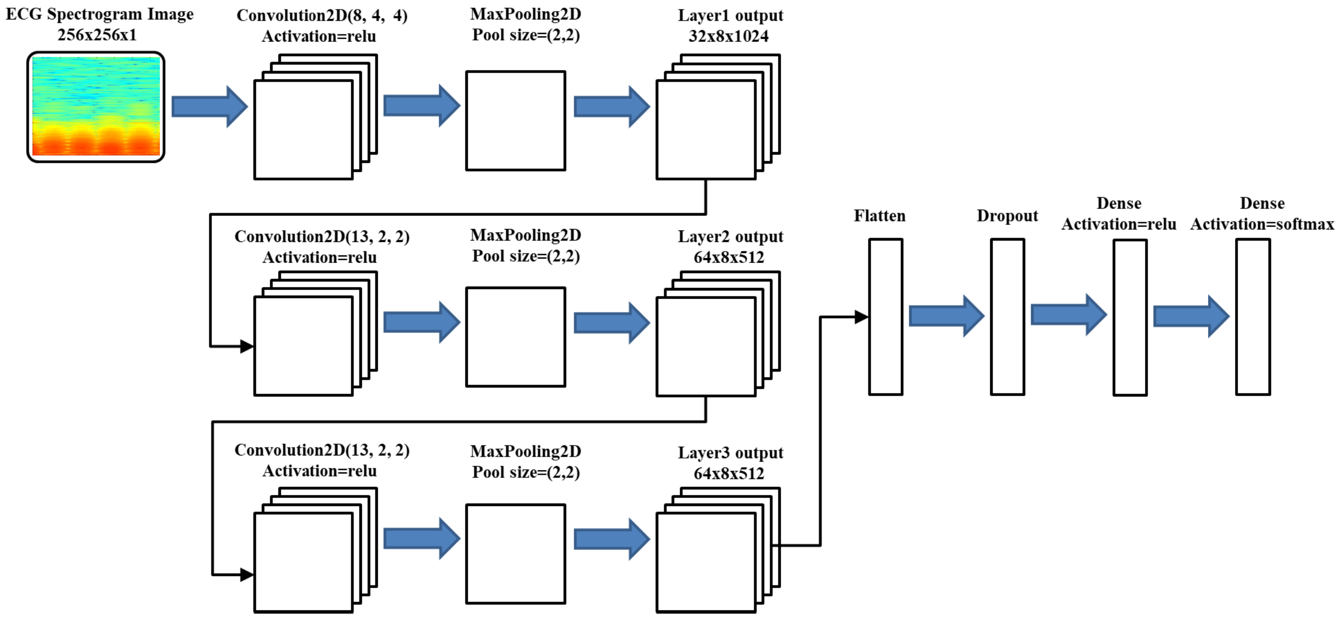


FIGURE 4. The architecture of the proposed 2D-CNN model.

8 convolution kernels and the kernels size of 4×4 was applied, and the activation function we choose is RELU (Rectified Linear Unit). Afterwards, a MaxPooling2D which pool size is (2, 2) was added. Then, the output shape of the first Layer is $32 \times 8 \times 1024$. In the second hidden layer, the Convolution2D layer with 13 convolution kernels and the kernels size of 2×2 was applied, and the activation function is RELU. Then, a MaxPooling2D with the pool size of (2, 2) was added, and the output shape of the second layer is $64 \times 8 \times 512$. In the third hidden layer, the Convolution2D layer with 13 convolution kernels and the kernels size of 2×2 was applied, and the activation function is RELU. Next, a MaxPooling2D with the pool size of (2, 2) was added, and the output shape of the third Layer is $64 \times 8 \times 512$ finally.

III. RESULTS

A. EVALUATION METRICS

In this section, we attempt to evaluate the classification performance using two metrics, e.g., the accuracy and the loss.

The indicator of the accuracy is the ratio between the number of correctly classified samples and that of the whole test samples. Its mathematical expression is defined as

$$Accuracy(\%) = \frac{TP + TN}{TP + TN + FP + FN} \times 100, \quad (3)$$

where TP stands for true positive, meaning the correct classification as arrhythmia; TN stands for true negative, meaning correct classification as normal; FP stands for false positive, meaning incorrect classification as arrhythmia; FN represents false negative, meaning incorrect classification as normal [27].

As for the metric of loss, it is defined as the difference between the predicted value of the model and the true value for a specific sample. This metric has several distinct types of mathematical expressions. In this study, we choose the function of categorical cross entropy loss.

$$loss = -\frac{1}{n}\sum_{i=1}^{n} \hat{y}_{i1} \ln y_{i1} + \hat{y}_{i2} \ln y_{i2} + \ldots + \hat{y}_{im} \ln y_{im}, \quad (4)$$

where n represents the number of samples; m represents the number of categories; $\hat{y}$ represents the predictive output value; and y represents the actual value.

B. MODEL PARAMETER OPTIMIZATION

In this proposed 2D-CNN model, there are two main parameters, that is, the learning rate and the batch size. In order to achieve the best classification performance of ECG heart rhythm abnormalities, the step of model parameter optimization is indispensable.

To evaluate the importance of the learning rate and the batch size within the proposed 2D-CNN model, a series of contrast experiments with different parameter sets were conducted. On one hand, we tested the contrast experiments with different learning rates when keeping the value of the batch size unchanged. The detailed parameter set is shown in Table 3. On the other hand, we attempted the contrast experiments with different batch sizes when keeping the value of the learning rate unchanged. The detailed parameter sets are shown in Table 5.

We set the number of iteration steps as 100. Figure 5 represents the accuracy value curves for the 7 data sets with the batch size of 2500. Table 3 shows the average accuracies for the 7 data sets. From the figures above, we can find that with the same batch size parameters, the average accuracies of the 7 data sets are similar. However, the accuracy curve shows different fluctuation at different learning rates. When

TABLE 3. Average accuracies for seven data sets with the batch size of 2500.

Batch Size	*Learning Rate*	*Average Accuracy*
2500	0.001	0.9900
2500	0.0025	0.9893
2500	0.005	0.9878
2500	0.0075	0.9885
2500	0.1	0.9882
2500	0.15	0.9887
2500	0.2	0.9885

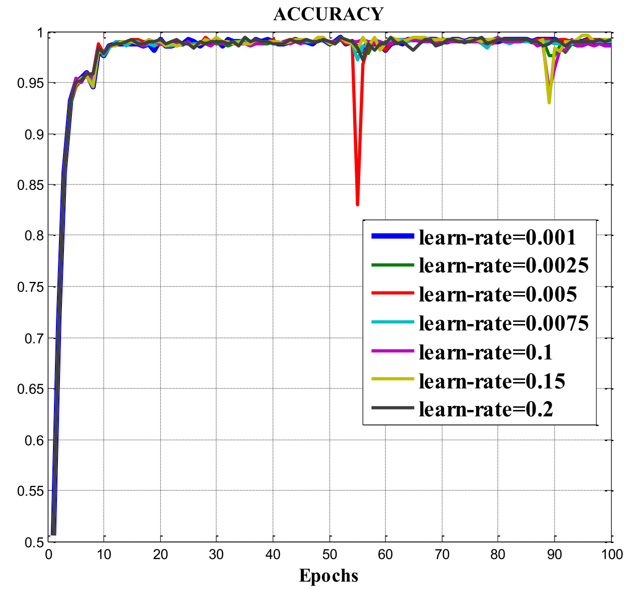


FIGURE 5. Accuracy value curves for 7 data sets with the Batch Size of 2500.

TABLE 4. Average losses for 7 data sets with the batch size of 2500.

Batch Size	*Learning Rate*	*Average Loss*
2500	0.001	0.0414
2500	0.0025	0.0447
2500	0.005	0.0506
2500	0.0075	0.0441
2500	0.1	0.0470
2500	0.15	0.0444
2500	0.2	0.0419

the learning rate is 0.001, as the number of iteration steps increases, the accuracy curve exhibits a convergence trend close to the value of 1, and it maintains a relatively stable state during convergence. When the learning rate is 0.0025, as the number of iteration steps increases, the accuracy curve still exhibits a convergence trend, but several relatively large fluctuations are found during the process of convergence. As the value of the learning rate parameter increases from 0.0025 to 0.2, the fluctuation during the convergence process of the accuracy curve becomes larger.

Figure 6 shows the loss value curves for the 7 data sets with the Batch Size of 2500. Table 4 shows the average losses for the 7 data sets. From the illustrations above we can find that with the same batch size parameters of 2500, the average losses of the 7 data sets are similar. However, at different learning rates, the loss curve shows different fluctuation.

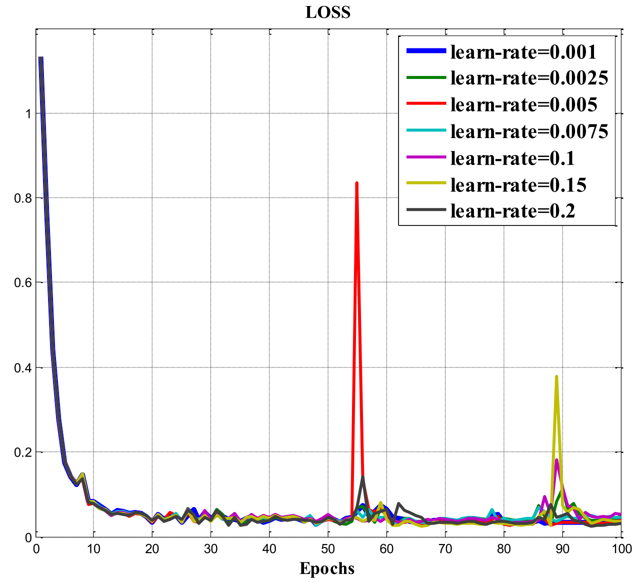


FIGURE 6. The loss value curves for seven data sets with the batch size of 2500.

TABLE 5. Average accuracies for 8 data sets with the Learning Rate of 0.001.

Batch Size	*Learning Rate*	*Average Accuracy*
2500	0.001	0.9900
2000	0.001	0.9894
1500	0.001	0.9887
1000	0.001	0.9901
750	0.001	0.9869
500	0.001	0.9891
250	0.001	0.9895
100	0.001	0.9898

When the learning rate is 0.001, as the number of iteration steps increases, the accuracy curve exhibits a convergence trend close to the value of 0, and it maintains a relatively stable state during convergence. When the learning rate is 0.0025, as the number of iteration steps increases, the loss curve still exhibits a convergence trend, but there are several relatively large fluctuations in the process of convergence. As the learning rate parameter increases from 0.0025 to 0.2, the fluctuations during the convergence of the loss curve become larger.

Table 4 shows detailed parameter set for the contrast experiments with different parameters of batch size when keeping the value of the learning rate unchanged.

Figure 7 represents the accuracy value curves for the 8 data sets with a common learning rate of 0.001. Table 5 shows the average accuracies for the 8 data sets. From the illustrations above we can find that with the same learning rate parameters, the average accuracies of the 8 data sets are similar. However, the accuracy curve shows different fluctuation at different batch sizes. When the batch size is 2500, as the number of iteration steps increases, the accuracy curve exhibits a convergence trend close to the value of 1, and it maintains a relatively stable state during convergence. When the batch size is set as 2000, as the number of iteration steps increases,

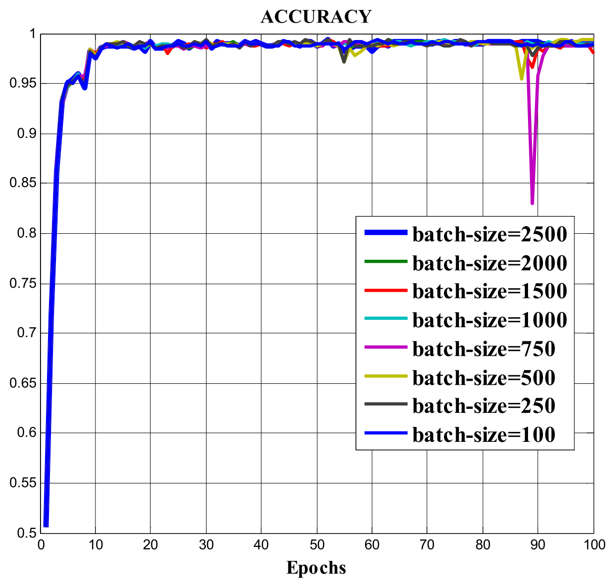


FIGURE 7. Accuracy curves for 8 data sets with the learn-rate of 0.001.

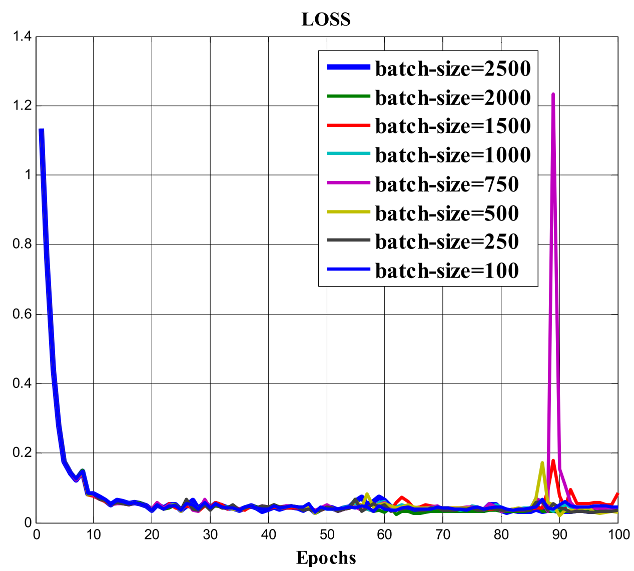


FIGURE 8. The loss value curves for 8 data sets with the learn-rate of 0.001.

TABLE 6. Averaged losses for 8 data sets with the Learning Rate of 0.001.

Batch Size	*Learning Rate*	*Average Loss*
2500	0.001	0.0414
2000	0.001	0.0385
1500	0.001	0.0482
1000	0.001	0.0440
750	0.001	0.0605
500	0.001	0.0445
250	0.001	0.0410
100	0.001	0.0437

the accuracy curve still exhibits a convergence trend, but there are several relatively large fluctuations in the process of convergence. With the batch size parameter gradually decreased, from 2000 to 100, the fluctuations during the convergence of the accuracy curve become larger.

TABLE 7. The averaged accuracies and losses of 1D-CNN model and the proposed 2D-CNN model.

MODEL	*Average Accuracy*	*Average Loss*
1D-CNN	0.9093	0.1610
Proposed 2D-CNN	0.9900	0.0414

Figure 8 represents the loss value curves for the 8 data sets with a common learning rate of 0.001. Table 6 shows the average losses for the 8 data sets. From the illustrations above we can find that with the same parameter of the learning rate, the average losses of the 8 data sets are similar. However, the loss curve shows different fluctuation at different batch sizes. When the batch size is set as 2500, as the number of iteration steps increases, the loss curve exhibits a convergence trend close to the value of 0, and maintains a relatively stable state during convergence. When the batch size is 2000, as the number of iteration steps increases, the loss curve still exhibits a convergence trend, but there are several relatively large fluctuations in the process of convergence. With the batch size parameter gradually decreased, from 2000 to 100, the fluctuation during the convergence of the accuracy curve becomes larger.

From the experimental comparisons demonstrated above, we can conclude that when the learning rate is 0.001 and the batch size parameter is 2500, the 2D-CNN model show the best accuracy as well as the lowest loss. In addition, the accuracy curve and the loss curve show the best convergence performance.

C. COMPARISON WITH 1D-CNN MODEL

In this subsection, we show a contrast experiment between the 1D-CNN model and the proposed 2D-Model. In the 1D-CNN model, the learning rate is 0.001 and the batch size parameter is **250**0. Evaluation results of 1D-CNN model and the proposed 2D-Model are summarized in Table 6. From the table, we can find that the 1D-CNN model achieved an average accuracy of 90.03% and an average loss of 16.10%. In contrast, the proposed 2D-CNN model achieved an average accuracy of 99.00% and an average loss of 4.14%.

Figure 9 represents the accuracy value curves of the 1D-CNN and the proposed 2D-CNN. From the graphs above we can find that, with the same learning rate and batch size, the accuracy value curve convergence rate of the proposed 2D-CNN model is faster than that of 1D-CNN model, and the final accuracy convergence value of 2D-CNN model is also much higher than that of 1D-CNN model.

Figure 10 represents the loss value curves of the 1D-CNN and that of the proposed 2D-CNN. From this figure, we can observe that the loss value curve convergence rate of the proposed 2D-CNN model is faster than that of 1D-CNN model, and the final loss convergence value of 2D-CNN model is also much lower than that of 1D-CNN model From these results, we can conclude that the proposed 2D-CNN model achieves a higher average accuracy with lower loss than the 1D-CNN model based on the classification results of five different

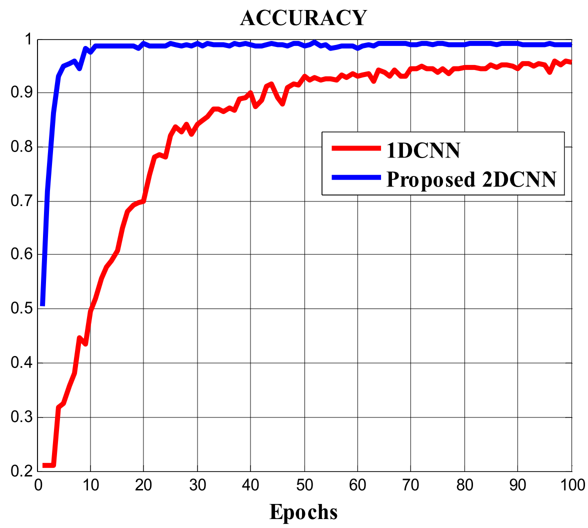


FIGURE 9. Accuracy value curves of the 1D-CNN and the proposed 2D-CNN.

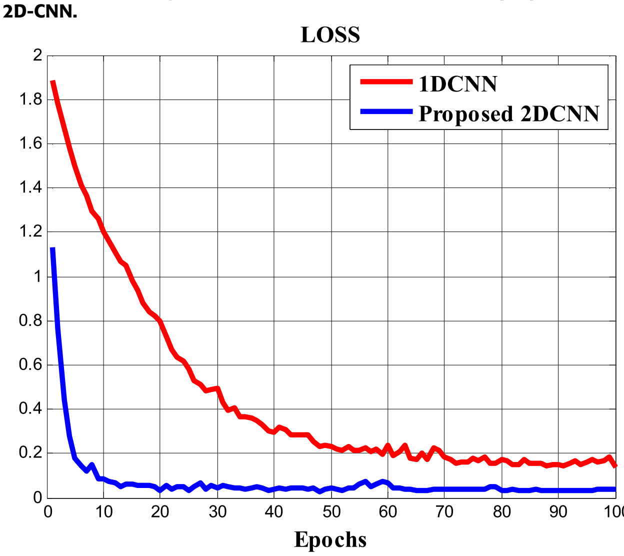


FIGURE 10. The loss value curves of the 1D-CNN and the proposed 2D-CNN.

arrhythmia types. The proposed 2D-CNN model outperforms the 1D-CNN model in the ECG arrhythmia classification application.

D. COMPARISON WITH OTHER EXISTING APPROACHES

We compared the performance of the proposed 2D-CNN model with previous ECG arrhythmia classification works, including SVM(Support Vector Machine), RNN(Recurrent Neural Network), RF(Random Forest), K-NN(K Nearest Neighbor). Since these works have a different number of the test set and types of arrhythmia, it is unfair to directly compared with accuracy itself. However, our proposed CNN model achieved successful performance compared to other previous works while introducing the different approach of classifying ECG arrhythmia using STFT-based spectrogram and convolutional neural network. Table 8 presents performance comparison with previous works. From the Table 8, we can find that the proposed method achieved the best results in average accuracy.

TABLE 8. Comparison with other existing approaches.

Model	*Work*	*Type*	*Test set*	*Average Accuracy*
2D-CNN	**Proposed**	**5**	**2520**	**99.00%**
FFNN	Guler et al.[35]	4	360	96.94%
	Yu et al.[36]	8	4900	98.71%
SVM	Melgani [37]	6	40438	91.67%
	Dutta et al.[38]	3	93246	95.82%
RNN	Ubeyli et al.[39]	4	360	98.06%
RFT	Kumar etal.[40]	3	154	92.16%
K-NN	Park et al.[41]	17	10972	97.00%

TABLE 9. Comparison with feature-extraction-pattern classification approaches.

Method	*Work*	*Average Accuracy*
STFT+2D-CNN	**Proposed**	**99.00%**
KPCA + SVM	Maya Kallas [41]	97.34%
WT&AR + SVM	Qibin Zhao et al.[42]	99.68%

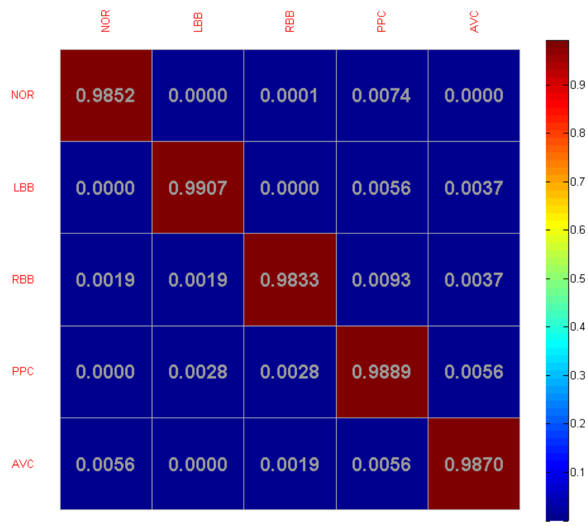


FIGURE 11. The normalized confusion matrix of the best result achieved.

We also compared the classification performance of the proposed 2D-CNN model with feature extraction–pattern classification approaches. From the TABLE 9, we can observe that the classification accuracy of the two feature-extraction–pattern classification approach is similar to that of the proposed method in this paper.

Maya Kallas used Kernel Principal Component Analysis (KPCA) for feature extraction of ECG signals, and apply the Support Vector Machines (SVM) classification, to diagnose heartbeat abnormalities [41]. Qibin Zhao used the wavelet transform and autoregressive modelling(AR) to extract the features of each ECG segment, and the support vector machine(SVM) with Gaussian kernel is used to classify different ECG heart rhythm [42]. The two feature-extraction–pattern classification approaches are comprised of three components including data preprocessing, feature extraction and classification of ECG signals. Compared with the proposed 2D-CNN classifier, the feature-extraction processing of feature-extraction-pattern approaches is much more complex.

In summary, the proposed method is a simple and efficient method with high classification accuracy.

The confusion matrix for the final result in the proposed 2D-CNN classifier is shown in Fig. 11.

IV. CONCLUSION

In this paper, we proposed an ECG arrhythmia classification method based on deep learning techniques. ECG signals, belonging to five different types, were obtained from the MIT-BIH arrhythmia database. The ECG signals were segmented into records of the duration of 10 seconds. 2520 records were selected for ECG classification.

In the procedure of the proposed method, the ECG signals in the time domain were transformed into two-dimensional time-frequency ECG spectrograms by short-time Fourier transform. The resultant ECG spectrograms were used as input to the proposed method. The ECG arrhythmia was identified and classified using the CNN. The results show that the classification of ECG signals based on two-dimensional convolution neural network can reach an averaged accuracy of 99.00%.

In addition, in order to achieve the best classification performance, a series of contrast experiments with different parameter sets were made. We found that the classifier based on the proposed 2D-CNN model has the highest accuracy and the lowest loss when the learning rate is 0.001 and the batch size is 2500.

Finally, we compared the performance of the proposed 2D-CNN model with that of the 1D-CNN model. ECG recordings of five arrhythmia types, shared by the MIT-BIH arrhythmia database, were utilized for evaluations of the classification performance. As a result, the 2D-CNN classifier achieved an averaged accuracy of 99.00%, while the 1D-CNN classifier achieved an averaged accuracy of 90.93%. The ECG arrhythmia classification experimental results have successfully validated that the proposed 2D-CNN can achieve better classification accuracy without manual pre-processing of the ECG signals such as noise filtering, feature extraction, and feature reduction.

REFERENCES

[1] World Health. (2017). *Cardiovascular Diseases (CVDs)*. Accessed: Apr. 18, 2018. [Online]. Available: http://www.who.int/mediacentre/factsheets/fs317/en/

[2] V. S. Chouhan and S. S. Mehta, "Threshold-based detection of P and T-wave in ECG using new feature signal," *Int. J. Comput. Sci. Netw. Secur.*, vol. 8, no. 2, pp. 144–153, 2008.

[3] S. A. Mahmoud, A. Bamakhramah, and S. A. Al-Tunaiji, "Six order cascaded power line notch filter for ECG detection systems with noise shaping," *Circuits, Syst., Signal Process.*, vol. 33, no. 8, pp. 2385–2400, 2014.

[4] E. Pasolli and F. Melgani, "Active learning methods for electrocardiographic signal classification," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 6, pp. 1405–1416, Nov. 2010.

[5] Z. Zidelmal, A. Amirou, M. Adnane, and A. Belouchrani, "QRS detection based on wavelet coefficients," *Comput. Methods Programs Biomed.*, vol. 107, pp. 490–496, Sep. 2012.

[6] M. A. Kabir and C. Shahnaz, "Denoising of ECG signals based on noise reduction algorithms in EMD and wavelet domains," *Biomed. Signal Process. Control*, vol. 7, no. 5, pp. 481–489, Sep. 2012.

[7] U. Raghavendra, H. Fujita, A. Gudigar, R. Shetty, K. Nayak, U. Pai, J. Samanth, and U. R. Acharya, "Automated technique for coronary artery disease characterization and classification using DD-DTDWT in ultrasound images," *Biomed. Signal Process. Control*, vol. 40, pp. 324–334, Feb. 2018.

[8] T. Li, and Z. Min, "ECG classification using wavelet packet entropy and random forests," *Entropy*, vol. 18, no. 8, p. 285, 2016.

[9] Y. Ozbay, R. Ceylan, and B. Karlik, "Integration of type-2 fuzzy clustering and wavelet transform in a neural network based ECG classifier," *Expert Syst. Appl.*, vol. 38, pp. 1004–1010, Jan. 2011.

[10] H. Li, D. Yuan, X. Ma, D. Cui, and L. Cao, "Genetic algorithm for the optimization of features and neural networks in ECG signals classification," *Sci. Rep.*, vol. 7, Jan. 2017, Art. no. 41011.

[11] F. A. Elhaj, N. Salim, A. R. Harris, T. T. Swee, and T. Ahmed, "Arrhythmia recognition and classification using combined linear and nonlinear features of ECG signals," *Comput. Methods Programs Biomed.*, vol. 127, pp. 52–63, Apr. 2016.

[12] Y. C. Yeh, C. W. Chiou, and H. J. Lin, "Analyzing ECG for cardiac arrhythmia using cluster analysis," *Expert Syst. Appl.*, vol. 39, pp. 1000–1010, Jan. 2012.

[13] E. Alickovic and A. Subasi, "Effect of multiscale PCA de-noising in ECG beat classification for diagnosis of cardiovascular diseases," *Circuits, Syst., Signal Process.*, vol. 34, no. 3, pp. 513–533, 2015.

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

[15] R. J. Martis, U. R. Acharya, C. M. Lim, K. M. Mandana, A. K. Ray, and C. Chakraborty, "Application of higher order cumulant features for cardiac health diagnosis using ECG signals," *Int. J. Neural Syst.*, vol. 23, no. 4, 2013, Art. no. 1350014.

[16] R. Varatharajan, G. Manogaran, and M. K. Priyan, "A big data classification approach using LDA with an enhanced SVM method for ECG signals in cloud computing," *Multimedia Tools Appl.*, vol. 77, no. 8, pp. 10195–10215, 2018.

[17] L. Taiyong and M. Zhou, "'ECG classification using wavelet packet entropy and random forests," *Entropy*, vol. 18, no. 8, p. 285, 2016.

[18] V. H. C. de Albuquerque, T. M. Nunes, D. R. Pereira, E. J. D. S. Luz, D. Menotti, J. P. Papa, and J. M. R. S. Tavares, "Robust automated cardiac arrhythmia detection in ECG beat signals," *Neural Comput. Appl.*, vol. 29, pp. 679–693, Feb. 2018.

[19] P. Raman and S. M. Ghosh, "Classification of heart diseases based on ECG analysis using FCM and SVM methods," *Int. J. Eng. Sci.*, vol. 6, pp. 6739–6744, Jun. 2016.

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

[21] X.-C. Cao, B.-Q. Chen, B. Yao, and W.-P. He, "Combining translation-invariant wavelet frames and convolutional neural network for intelligent tool wear state identification," *Comput. Ind.*, vol. 106, pp. 71–84, Apr. 2019.

[22] R. Salloum and C.-C. J. Kuo, "'ECG-based biometrics using recurrent neural networks," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2017, pp. 2062–2066.

[23] A. Mostayed, J. Luo, X. Shu, and W. Wee, "Classification of 12-lead ECG signals with Bi-directional LSTM network," 2018, *arXiv:1811.02090*. [Online]. Available: https://arxiv.org/abs/1811.02090

[24] C. Zhang, G. Wang, J. Zhao, P. Gao, J. Lin, and H. Yang, "Patient-specific ECG classification based on recurrent neural networks and clustering technique," in *Proc. 13th IASTED Int. Conf. Biomed. Eng. (BioMed)*, Feb. 2017, pp. 63–67.

[25] 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. 664–675, Mar. 2016.

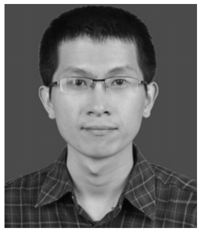
[26] D. Li, J. Zhang, Q. Zhang, and X. Wei, "Classification of ECG signals based on 1D convolution neural network," in *Proc. IEEE 19th Int. Conf. E-Health Netw.*, Oct. 2017, pp. 1–6.

[27] W. Yin, X. Yang, L. Zhang, and E. Oki, "ECG monitoring system integrated with IR-UWB radar based on CNN," *IEEE Access*, vol. 4, pp. 6344–6351, 2016.

[28] T. J. Jun, H. M. Nguyen, D. Kang, D. Kim, D. Kim, and Y.-H. Kim, "ECG arrhythmia classification using a 2-D convolutional neural network," 2018, *arXiv:1804.06812*. [Online]. Available: https://arxiv.org/abs/1804.06812

[29] M. Salem, S. Taheri, and J.-S. Yuan, "ECG arrhythmia classification using transfer learning from 2- dimensional deep CNN features," in *Proc. IEEE Biomed. Circuits Syst. Conf.*, Oct. 2018, pp. 1–4.
[30] P. Rajpurkar, A. Y. Hannun, M. Haghpanahi, C. Bourn, and A. Y. Ng, "Cardiologist-level arrhythmia detection with convolutional neural networks," 2017, *arXiv:1707.01836*. [Online]. Available: https://arxiv.org/abs/1707.01836
[31] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH arrhythmia database," *IEEE Eng. Med. Biol. Mag.*, vol. 20, no. 3, pp. 45–50, May/Jun. 2001.
[32] S. Haykin and B. V. Veen, *Signals and Systems*. Hoboken, NJ, USA: Wiley, 1999.
[33] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," *Neural Comput.*, vol. 1, no. 4, pp. 541–551, 1989.
[34] I. Güler and E. D. Übeyli, "ECG beat classifier designed by combined neural network model," *Pattern Recognit.*, vol. 38, no. 2, pp. 199–208, 2005.
[35] S.-N. Yu and K.-T. Chou, "Integration of independent component analysis and neural networks for ECG beat classification," *Expert Syst. Appl.*, vol. 34, no. 4, pp. 2841–2846, 2008.
[36] F. Melgani and Y. Bazi, "Classification of electrocardiogram signals with support vector machines and particle swarm optimization," *IEEE Trans. Inf. Technol. Biomed.*, vol. 12, no. 2, pp. 667–677, Sep. 2008.
[37] S. Dutta, A. Chatterjee, and S. Munshi, "Correlation technique and least square support vector machine combine for frequency domain based ECG beat classification," *Med. Eng. Phys.*, vol. 32, no. 2, pp. 1161–1169, 2010.
[38] E. D. Übeyli, "Combining recurrent neural networks with eigenvector methods for classification of ECG beats," *Digit. Signal Process.*, vol. 19, no. 2, pp. 320–329, 2009.
[39] R. Kumar and Y. S. Kumaraswamy, "Investigating cardiac arrhythmia in ECG using random forest classification," *Int. J. Comput. Appl.*, vol. 37, no. 2, pp. 31–34, 2012.
[40] J. Park, K. Lee, and K. Kang, "Arrhythmia detection from heartbeat using k-nearest neighbor classifier," in *Proc. IEEE Int. Conf. Bioinf. Biomed.*, 2013, pp. 15–22.
[41] M. Kallas, C. Francis, L. Kanaan, D. Merheb, P. Honeine, and H. Amoud, "Multi-class SVM classification combined with kernel PCA feature extraction of ECG signals," in *Proc. IEEE Int. Conf. Telecommun.*, Apr. 2012, pp. 1–5.
[42] Q. Zhao and L. Zhang, "'ECG feature extraction and classification using wavelet transform and support vector machines," in *Proc. Int. Conf. Neural Netw. Brain*, Oct. 2005, pp. 1089–1092.

JINGSHAN HUANG was born in Quanzhou, China. He received the bachelor's degree of engineering from Xiamen University, in 2017, where he is currently pursuing the master's degree of engineering. His research interests include advanced manufacturing technologies, intelligent equipment and smart production line, and signal processing.



BINQIANG CHEN was born in Fuqing, China, in 1986. He received the bachelor's degree in mechanical engineering from the School of Manufacturing Science and Technology, Sichuan University, in 2008, and the Ph. D degree in mechanical engineering from the School of Mechanical Engineering, Xi'an Jiaotong University, in 2013. He is currently an Assistant Professor with the School of Aerospace Engineering, Xiamen University, China. His main research interests include intelligent equipment and smart manufacturing, structural health monitoring of equipment, and applied harmonic analysis.

BIN YAO was born in Luoyang, Henan, China, in 1963. He received the Ph.D. degree from Xi'an Jiaotong University, in 2003. He is currently a Full Professor with the School of Aerospace Engineering, Xiamen University, China. His main research interests include shape of curved surface and NC technology, detection technology, and intelligent machining equipment.

WANGPENG HE was born in Yulin, China. He received the B.S. and Ph.D. degrees in mechanical engineering from Xi'an Jiaotong University, Xi'an, China, in 2007 and 2016, respectively. In 2014, he was appointed as a Visiting Scholar with the New York University, USA. He joined the School of Aerospace Science and Technology, Xidian University, where he is currently an Assistant Professor. His research interests include signal processing, machine vision, sparsity-based signal processing, and machinery fault diagnosis.