Computer Modeling in Engineering & Sciences
DOI:10.32604/cmes.2020.010798
Article

PDNet: A Convolutional Neural Network Has Potential to be Deployed on Small Intelligent Devices for Arrhythmia Diagnosis

Fei Yang[1,2,#], Xiaoqing Zhang[1,*,#] and Yong Zhu[3]

[1]Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen, China
[2]Cooperative Innovtion Center of Internet Healthcare, Zhengzhou University, Zhengzhou, 450052, China
[3]College of Computer Science and Technology, Huaqiao University, Xiamen, 361021, China
[*]Corresponding Author: Xiaoqing Zhang. Email: dbozhang@foxmail.com
[#]Fei Yang and Xiaoqing Zhang contributed equally to this work
Received: 21 March 2020; Accepted: 20 July 2020

Abstract: Heart arrhythmia is a group of irregular heartbeat conditions and is usually detected by electrocardiograms (ECG) signals. Over the past years, deep learning methods have been developed to classify different types of heart arrhythmias through ECG based on computer-aided diagnosis systems (CADs), but these deep learning methods usually cannot trade-off between classification performance and parameters of deep learning methods. To tackle this problem, this work proposes a convolutional neural network (CNN) model named PDNet to recognize different types of heart arrhythmias efficiently. In the PDNet, a convolutional block named PDblock is devised, which is comprised of a pointwise convolutional layer and a depthwise convolutional layer. Furthermore, an improved loss function is utilized to improve the results of heart arrhythmias classification. To verify the proposed CNN model, extensive experiments are conducted on public MIT-BIH ECG databases. The experimental results demonstrate that the proposed PDNet achieves an accuracy of 98.2% accuracy and outperforms state-of-the-art methods about 2%.

Keywords: Electrocardiograms; heart arrhythmia; convolutional neural network; PDblock; loss

1 Introduction

Cardiovascular diseases (CVDs) are the leading cause of death globally. It is reported that about 17.7 million people died in 2015 due to CVDs worldwide, 82% of which are from developing countries [1]. In 2015, China also reported approximately 290 million people who were suffering from cardiovascular diseases, which means that CVDs are parts of the priority diseases and should be taken seriously. According to recent research [2], it is reported that over 90% of CVDs may be preventable with prevention including improving risk factors (e.g., exercise). Hence, it is meaningful and important to predict CVDs as soon as possible through developing advanced CADs. Arrhythmia (abnormal heartbeats) is one of the most common types of CVDs, which is usually detected through ECG signals. In clinic, arrhythmias contain a broad range of irregular heart rhythms, and can be classified into two groups of non-life-threatening and life-threatening. According to the Association for the Advancement of Medical Instrumentation (AAMI) [3], non-life-threatening arrhythmias can be loosely divided into five types: unknown (Q), supraventricular ectopic (S), ventricular ectopic (V), non-ectopic (N), and fusion (F). The ECG signals have been standard testing for arrhythmia diagnosis over the years. In clinical practice, arrhythmia diagnosis is performed by cardiologists or doctors via special ECG monitor devices, which can collect ECG signals from the patients. However, due to the differences in experiences of cardiologists and doctors,the actual diagnosis of arrhythmia is time-consuming, subjective, laborious, and error-prone. In addition, a serious problem in poor countries is that experienced cardiologists and expensive medical equipment are very scarce [4].

In recent years, a lot of small intelligent devices have been developed to collect various data from human and environments like smartwatches, smartphones, sensors, and ECG monitors. With the development of Internet of Thing (IoT), small intelligent devices have become an essential part of daily life in both developing countries and developed countries. Many mobile applications have been deployed on small intelligent devices to provide different service such as health monitoring, face recognition and sport monitoring. Researchers [5–7] have committed to developing CAD systems for biomedical signal processing and ECG analysis based on small intelligent devices, which presents a potential method to provide health service for people and is possible to be used for early CVD prevention.

To detect different types of heart arrhythmias accurately, it is necessary and important to propose advanced methods for ECG CAD systems. Heart arrhythmia detection can be treated as a classical pattern recognition problem. In the past decades, many machine learning methods have been used to analyze ECG signals ranging from traditional machine learning methods to deep learning methods, like support vector machine (SVM), convolutional neural networks (CNNs), and recurrent neural networks (RNNs). Due to the strong feature representation abilities of deep learning methods, recent works have gradually used deep learning methods for heart arrhythmia detection and have achieved expected performance. However, the classification performance of deep learning methods for heart arrhythmia detection heavily depends on the number of parameters or the depth of deep learning methods. Therefore, deploying deep learning models on small intelligent devices and resource-constrained platforms has become a bottleneck.

In general, with the rapid development of artificial intelligence technique (AI), Internet of things (IoT), wearable sensors and small intelligent medical equipment in recent years, it is likely to collect and monitor ECG signals in the remote. It presents new challenges and great potentials for the development of intelligent ECG CAD systems.

This work is different from most previous works which are focused on improving heart arrhythmia detection results through devising complicated deep learning methods. It is an extension of our previous work [8] in 2nd ICHSE. The core aim of this work is to apply deep learning methods to heart arrhythmia detection, which can not only achieve high heart arrhythmia detection performance, but also does not require a huge number of parameters. To achieve the aim, a novel convolutional neural network (CNN) model called PDNet is built on a novel convolutional block named PDblock. The PDblock is comprised of a pointwise convolutional layer and a depthwise convolutional layer, in which the pointwise convolution operation is followed by the depthwise convolution operation. According to related literature [9,10], both depthwise convolution operation and pointwise convolution operation are efficient methods to reduce the parameters of deep learning methods. To further improve the accuracy of heart arrhythmia detection, an improved loss is devised, inspired by Lee et al. [11] proposed in literature. The improved loss utilizes label replication methods to make the dense layers of the PDNet generate errors, thus it enables the PDNet to learn more feature information from data. To validate the effectiveness of the proposed methods, the experiments are conducted on MIT-BIH arrhythmia databases [12]. The experiment results demonstrate that the proposed PDNet outperforms strong baselines and state-of-the-art previous methods. Compared with advanced CNN models such as AlexNet [13], VGGNet, and MobileNet [14], the PDNet trades off better between heart arrhythmia detection results and the number of parameters in CNN models. Furthermore, the results also show that the improved loss achieves better classification performance than cross entropy loss and focal loss. Therefore, the PDNet is likely to be embedded in intelligent ECG CAD systems to assist cardiologists to screen common heart arrhythmias and provide health service to people on intelligent ECG-based CAD systems and devices. Fig. 1 presents a framework of intelligent ECG CAD systems with the proposed PDNet, which is likely to be deployed on small intelligent devices to detect different types of heart arrhythmias. The main contributions of this work are constructed as follows.

(1) A CNN model named PDNet is proposed to detect different types of heart arrhythmias. Moreover, the comparison to state-of-the-art CNN models shows that it is capable of trading off accuracy and parameters of deep learning methods efficiently through comparison to state-of-the-art CNN models.
(2) An improved loss function is utilized to improve the heart arrhythmia results based on the label replication method.
(3) Extensive experiments have been carried out on public MIT-BIH arrhythmia databases, and the proposed methods outperform strong baselines and previous methods on evaluation measures such as accuracy, precision, and recall.

The rest of this work is organized as follows: we review previous works of heart arrhythmia detection based on CAD and recent advances in CNN in Section 2. Section 3 elucidates the PDNet architecture and the improved loss. MIT-BIH Arrhythmia databases and data preprocessing are introduced in Section 4. In Section 5, we introduce the experiment setup and result analysis. Conclusion and future work are presented in Section 6.

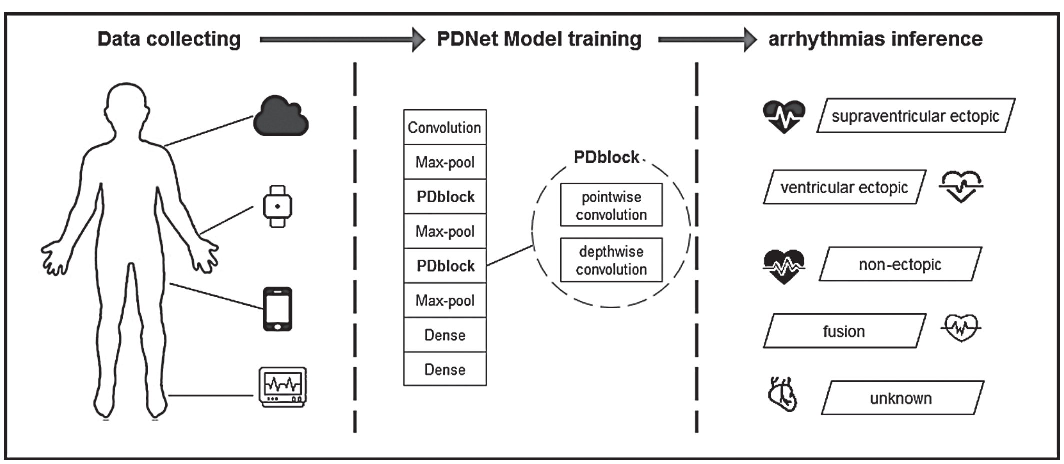


Figure 1: The framework of intelligent ECG CAD systems with the proposed PDNet deployed on small intelligent devices to detect different types of heart arrhythmias

2 Related Work

In this section, previous works of heart arrhythmia detection based on CAD are surveyed from traditional machine learning methods and deep learning methods, and followed by recent advances in CNN.

ECG CAD. Heart arrhythmia detection through ECG signals can be taken as a classical pattern recognition problem. In the past decades, researchers have developed lots of machine learning methods to recognize different types of heart arrhythmias through ECG signals can be loosely split into two classes as traditional machine learning methods and recent deep learning methods. Traditional machine learning methods comprise three common processing procedures: data preprocessing, feature extraction, modeling, and classification. Literature [15–19] use traditional machine learning methods to classify different types of arrhythmias based on ECG signals and achieve good classification results. E.g., Alonso-Atienza et al. [20] used a personalized feature selection method and support vector machine (SVM) to detect arrhythmias and achieved over 90% accuracy. In literature [21], researchers developed a hierarchical classification method for heartbeat classification based on weighted extreme gradient boosting (XGBoost) and feature selection methods. The result showed that the proposed method improved the classification performance.

Due to powerful feature representation learning ability [22], recent works have gradually deep learning methods for heart arrhythmia detection based on ECG signals [7,20,23–31] and have achieved good classification performance in both public ECG datasets and self ECG datasets. Hannun et al. [28] used a 34-layer convolutional neural network (CNN) to detect arrhythmias. To alleviate the gradient vanishing problem, two shortcut connections were used. The results showed the proposed CNN model outperformed the board-certified cardiologists, which also confirmed deep learning has potential in ECG based CAD systems. Acharya et al. [32] utilized a deep CNN for arrhythmia diagnosis on publicly available arrhythmia database. The results showed the proposed CNN was able to achieve similar accuracies of arrhythmia diagnosis on noise and noise-free ECG dataset. He et al. [33] also proposed a CNN-based method to diagnose arrhythmias and achieved over 95% accuracy. Wang et al. [34] proposed a dual fully-connected neural network for different types of heartbeat classification and achieves excellent classification performance. Shi et al. [35] proposed a hybrid deep learning framework for automated heartbeat classification which is comprised of a CNN and a long short-term memory (LSTM) network. In their work, the combination of automatic features and handcraft features as inputs were used to improve the overall classification performance, and results showed that the proposed model achieved an accuracy of 99.26%.

In recent years, many m-health devices and mobile applications are applied to monitor the activities and health of human [36]. The Computing in Cardiology [12] organized a challenge to devise advanced methods to classify arrhythmias accurately through mobile devices, which presented a promising perspective in Mobile Health (m-health). Gradl et al. [37] developed a mobile system for arrhythmia detection automatically based on mobile devices, which can be used to process ECG signals and generates arrhythmia prediction results in real-time. The results showed that the proposed mobile system achieved a sensitivity of 89.5% and a specificity of 80.6%, respectively.

Recent Advances in CNN. Considering the real requirements of intelligent mobile devices and resource-constrained platforms, it is necessary and urgent to construct light-weight deep learning models. Hence, many convolution methods were proposed to devise advanced and light-weight

CNN architectures [9,38,39] such as group convolution method [10,40], depthwise separable convolution method, depthwise convolution method, and pointwise convolution method [41]. In the depthwise separable convolution method, a standard convolution operation is divided into a pointwise convolution operation and a depthwise convolution operation. It is difficult to construct powerful deep learning models in the medical field, and researchers [42] have gradually used the transfer learning strategy to train the deep neural network for achieving excellent performance by loading pre-trained network structures. However, pre-trained network structures are trained on two-dimensional (2D) images, which is not suitable for one-dimensional (1D) time-series ECG signals.

3 Method

In this section, the proposed CNN architecture based on the devised PDblock is introduced in detail and followed by the improved loss function, which is introduced to further enhance the classification results of arrhythmia diagnosis.

3.1 PDblock

Convolutional neural network (CNN) is one of the most used deep learning models and has achieved great success in many tasks. It is usually comprised of convolutional layers, pooling layers, and dense layers (fully-connected layers). In this work, a novel CNN model named PDNet is constructed to classify heart arrhythmias, which is built on the proposed convolutional block called PDblock. The PDblock is comprised of a pointwise convolutional layer and a depthwise convolutional layer based on a pointwise convolution method and a depthwise convolution method. It can be treated that a standard convolutional layer is split into two above types of convolutional layers. Generally, a pointwise convolution is a 1×1 standard convolution that projects the output space of each channel to a new channel space. While a depthwise convolution method applies a single convolution operation to each channel of a pointwise channel output correspondingly. Fig. 2 presents an example of how to factorize a standard convolutional layer Fig. 2a into a pointwise convolutional layer (Fig. 2b) and a depthwise convolutional layer (Fig. 2c). The pointwise convolutional layer is a 1×1 standard convolutional layer, which is capable of learning correlated feature representations from the previous layer. The depthwise convolutional layer can get feature representations of each feature map respectively via one-to-one learning.

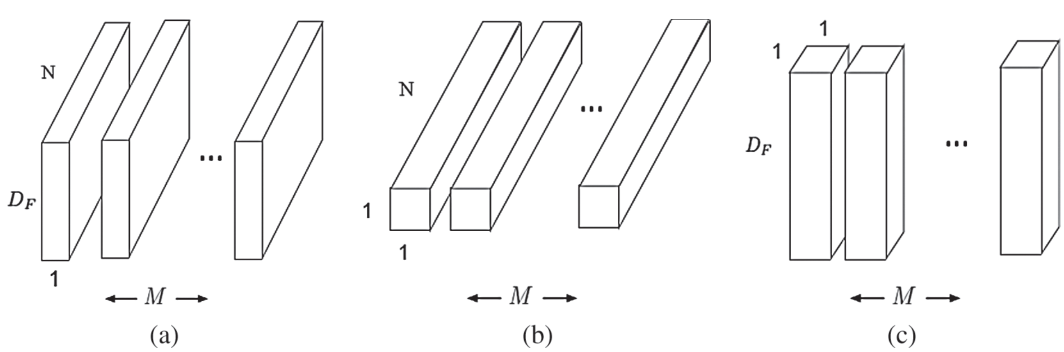


Figure 2: The standard convolutional layer (a) are replaced by two layers: a 1×1 pointwise convolutional layer (b) and a depthwise convoluional layer (c)

Considering time-series ECG signals are one-dimensional (1D) data, hence, the 1D convolution method is adopted in this work. To understand the working mechanism of the proposed PDblock easily, we make a comparison between the PDblock and the standard convolutional layer theoretically. For example, a standard convolutional layer of the CNN which uses the feature map N with $1 * D_k * N$ as the input and generates a feature map M with $1 * D_k * M$, where $1 * D_k$ is the length of feature map N and output feature map M, N is the number of input channels, and M is the number of output channels.

The number of parameters of convolutional kernel F in a standard convolutional layer is $1 * D_F * N * M$ and $1 * D_F$ is the size of the convolution kernel, N, M denotes the number of input channels and output channels as previously defined. Hence, the computational cost of a standard convolutional layer can be computed as follows.

$$1 * D_F * N * M * 1 * D_k. \tag{1}$$

where the computational cost highly depends on the size of convolution kernel $1 * D_F$, the number of input channels N, the number of output channels M, and the length of feature map $1 * D_k$ of previous convolutional layer. To reduce the computational cost of the standard convolutional layer, a convolution block named PDblock is introduced and is comprised of a pointwise convolutional layer and a depthwise convolutional layer. Pointwise convolutional layer is applied to cluster correlated feature representation from the input channels and followed by a depthwise convolutional layer, which maps a single convolution kernel to each input channel of the previous pointwise convolutional layer. Eq. (2) presents the computational cost of a depthwise convoultional layer:

$$1 * D_F * M * 1 * D_k, \tag{2}$$

F denotes the weight parameters of depthwise convolution kernel $1 * D_F * M$, $1 * D_F$ is the size of the convolution kernel, M is the number of output channels. Thus, compared with standard convolution operation, depthwise convolution operation is able to reduce computational cost efficiently. Pointwise convolution layer can reduce computational cost through adopting 1×1 convolution kernel. The computational cost of the PDblock can be represented as follows.

$$M * N * 1 * D_k + 1 * D_F * M * 1 * D_k, \tag{3}$$

which is the sum of pointwise convolution operation (left) and depthwise convolution operation (right). $1 * D_k$, $1 * D_k$, N, and M denote the length of the input feature map, the length of the output feature map, the number of input channels, and the number of output channels, respectively. Compared with a standard convolutional layer, the PDblock can get a reduction of computational cost as follows.

$$\frac{1 * D_F * M * 1 * D_k + M * N * 1 * D_k}{1 * D_F * M * N * 1 * D_k} = \frac{1}{N} + \frac{1}{D_F} \tag{4}$$

If depthwise convolution kernel sizes with 1×3 or 1×4 are used for ECG analysis in this work, reducing 3 to 4 times less computation cist than a standanard convolutional layer.

3.2 Convolutional Neural Network Architecture

In this work, we utilize the PDblock as the backbone to construct the PDNet, as shown in Fig. 3. It contains an input layer, a convolutional layer, three max-pooling layers, two PDblock layers, two fully-connected (dense) layers, and the output layer (softmax layer). The input layer is used as the input for time-series ECG signals and the output of the softmax layer is used for multi-classification. Max-pooling layers can reduce the computational cost between two layers and select useful feature representations. A convolutional layer and two PDblock layers are used to learn high-level feature representations from original ECG features. Based on learned local high-level feature representations of the convolutional layer and PDblock layers, we apply two fully-connected layers to learn the relationship from local high-level feature representations, which can help the PDNet achieve good classification performance.

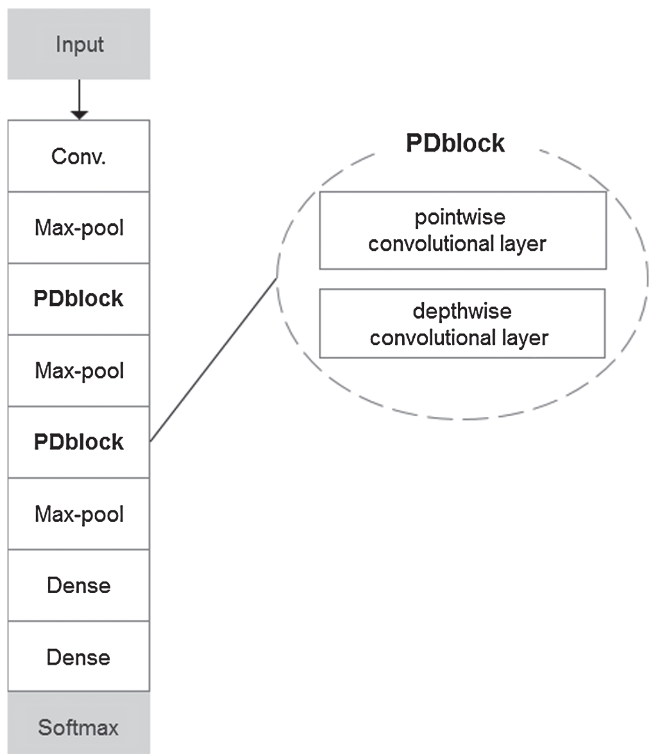


Figure 3: The PDNet architecture

Tab. 1 summarizes the specific architecture of the proposed PDNet. In the PDNet, the first convolutional layer uses convolution kernel size 1×4 and generates five feature maps after the convolution operation and the activation operation. Two PDbolck layers convolve with the same kernel size 1×1 and 1×3 and generate 10 and 20 feature maps correspondingly. Every convolutional layer, followed by a max-pooling layer is able to lower the computational cost without reducing the performance. Two dense layers have 30 and 20 neurons respectively. The softmax layer outputs 5 predicted results according to types of arrhythmias.

The activation function is an important factor to affect the performance of deep learning methods. To achieve good arrhythmia classification performance in this work, we apply several activation functions like tanh, sigmoid, rectified linear unit, leaky rectified linear unit (leakyRelu) [43] activation functions to the proposed PDNet model respectively, and the leakyRelu is adopted as activation function based on the arrhythmia classification performance. Batch normalization (BN) [44] layer has become an important component of modern CNN models due to its powerful function. In this work, we also used it after every convolutional layer for getting the expected classification performance and accelerating the convergence of the proposed CNN model.

Table 1: Construction of the proposed CNN model

	Layer type	Number of feature maps	Kernel size	Stride
	Convolutional layer	5	1×4	1
	Max-pooling layer		1×2	2
PDblock	Pointwise convolutional layer	10	1×1	1
	Depthwise convolutional layer	10	1×3	1
	Max-pooling layer		1×2	2
PDblock	Pointwise convolutional layer	20	1×1	1
	Depthwise convolutional layer	20	1×3	1
	Max-pooling layer		1×2	2
	Dense layer	30		
	Dense layer	20		
	Softmax layer	5		

3.3 Improved Loss and Adam Optimizer

The training process of the CNN model is implemented by the backpropagation method [45]. Loss function is an important factor for classification performance. For deep learning methods, it is standard to use cross entropy (CE) loss as loss function and can be expressed as following equation:

$$CE = -\frac{1}{N}\left[\sum_{i=1}^{N}\sum_{j=1}^{K} 1\{y^{(i)} = j\} \log p_j^{(i)}\right]. \tag{5}$$

y, p denotes ground truth and predicted probability of each label, respectively, N and K denote the number of instances and the number of arrhythmia types.

Getting inspiration from [11,46], an improved loss is introduced based on the label replication method and CE. In the improved loss, we make every hidden layer generate predicted outputs and compute losses of each output. We use P to represent the sum of losses of each hidden layer as you can see in Eq. (6):

$$P = \alpha \sum_{j=1}^{n} CE_j. \tag{6}$$

α is a hyper-parameter and set for 0.02 based on the experiment results, n represents the number of layers and is set at 2, which means that two dense layers all generate error signals by replicating true labels to them except for softmax layer in the training phase. According to the previous introduction, the improved loss (L) is a companion optimization objective function and can be computed as follows:

$$L = CE_{final} + P. \tag{7}$$

P denotes the auxiliary loss of hidden layers, and CE_{final} denotes the loss of the class-fication layer. The proposed improved loss function considers the different high-level feature representation stage, which can make the PDNet learn good feature representations in the training.

Furthermore, we use Adam [47] optimizer as the optimizer for the proposed CNN, which is a first-order gradient-based descent optimizer of stochastic objective functions. It is very easy to implement and work efficiently.

4 Dataset

ECG datasets used in this work are from the public MIT-BIH Arrhythmia database [12]. The database contains 48-half hour-long ECG records from 47 objects, which sampling rate is 360 Hz. Two types of ECG sets are collected in this work: set A and set B. Set A with a sampling rate of 260 Hz and set B with a sampling rate of 360 Hz. 107375 ECG records were extracted from the MIT-BIH database, and the number of N, S, V, F, and Q is 90592, 2781, 802, 7235, and 5965 respectively. Given real-world applications of arrhythmias detection into consideration, two data sets of ECG dataset are original data without denoising. Fig. 4 gives the distribution of five types of non-life-threatening arrhythmias (F, N, Q, S, and V), and the class imbalance as large as 112-fold between N and F. Hence, to keep the label balanced, the synthetic data strategy is used to synthesize ECG data. After synthesizing data, the total number of ECG segments including N, S, V, F, and Q types of arrhythmia is 465000 and all types of heart arrhythmias have the same number of data. Furthermore, the Z-score method is used to preprocess ECG data, which can alleviate the amplitude scaling problem in an ECG heartbeat. To verify the performance of the proposed PDNet, both balanced and imbalanced data sets are used in experiments. Additionally, a ten-fold cross-validation method is applied to the training and testing, which means 90% of the ECG data are used for training, the rest is used for testing every time.

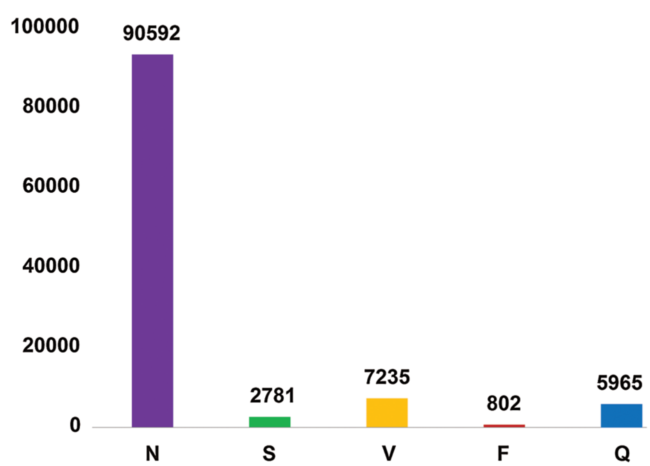


Figure 4: The distribution of five types of arrhythmias (N, S, V, F, and Q)

5 Experiment

5.1 Experiment Setup

All methods are implemented on the TensorFlow platform [48] and python. To verify the performance of the proposed PDNet comprehensively, state-of-the-art convolutional neural networks were used like AlexNet and MobileNet. The number of convolution kernels of the MobileNet is the same as the PDNet, and AlexNet uses the same parameter settings as the model in literature [32], which [32] can achieve good classification performance. Thus, this work follows it to set similar parameters such as the size of the convolution kernel and the number of convolution kernels. Other advanced CNN models (VGGNet [49] and GoogleNet [38]) are also used for comparison. The batch size for the training is set for 64, and we manually set the learning rate at 0.0035 and initialize the weights of the proposed PDNet randomly. Network models are trained on a server with six Intel Xeon (R) 2.60 GHz (E5-2650) processors and 64 GB RAM. It takes about four hours to complete ten training epochs for the PDNet, the testing time of PDNet is in 5 ms, which is very short. However, it can not reflect the real testing time of PDNet, because we do not deploy it on real applications.

Common evaluation measures like accuracy, sensitivity, and recall (positive predictive value) are used to evaluate the performance of CNN models. The main task of this work is to find a good balance between the computational cost and accuracy of CNN models. Parameter count (PC) is used as an evaluation measure, which is the sum of parameters on convolutional layers (not including dense layers). Parameter count is an important factor which is related to computational complexity, memory demanding, and classification performance.

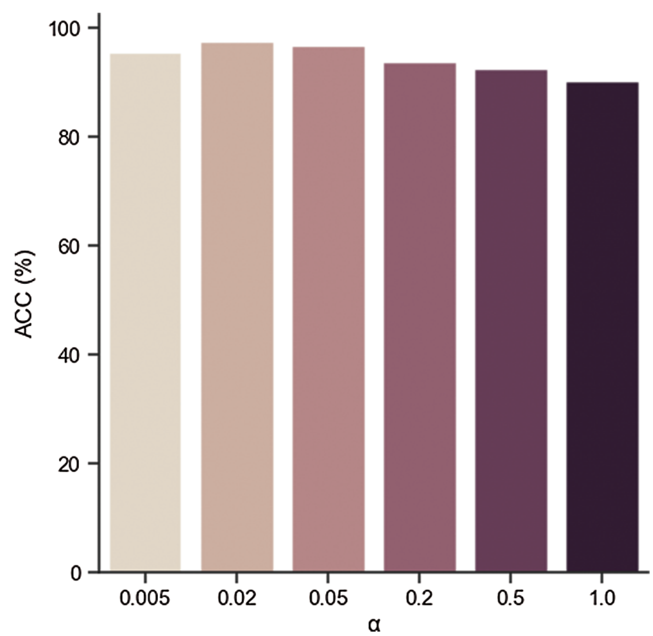


Figure 5: Hyper-parameter selection results of α

5.2 Result Analysis and Discussion

Fig. 5 shows hyper-parameter selection results of α, on set A and the PDNet achieves the best accuracy of 97.71% when α is set as 0.02. When α is set for 1.0, the model achieves the worst performance, because the auxiliary loss dominates the gradient in the training phase. Furthermore, the value of α is small, which has little effect on the total loss. Hence, it is hard to set the value of α and we set it for 0.02 in this work.

Tabs. 2 and 3 present the classification performance of Adam optimizer and SGD optimizer based on the PDNet. Adam optimizer outperforms SGD optimizer about 2.11% of accuracy. SGD optimizer takes five times of epochs than Adam optimizer in the training. Because Adam optimizer gets better classification performance and requires less training time than SGD optimizer, thus, Adam optimizer is used as the optimizer for all following experiments.

Table 2: Experimental results of Adam and SGD on set A

Opt	Rec	Se	Acc
Adam	99.34%	99.37%	97.71%
SGD	95.89%	98.57%	95.60%

Table 3: Experimental results of Adam and SGD on set B

Opt	Rec	Se	Acc
Adam	99.36%	99.46%	98.20%
SGD	96.25%	98.89%	95.74%

Table 4: Experimental results of five CNN models on set A

Model	Rec	Se	Acc	PC
This work	99.34%	99.37%	97.71%	395
AlexNet	95.80%	97.28%	94.78%	805
MobileNet	95.78%	97.33%	92.96%	345
VGGNet	98.16%	98.85%	96.33%	1535
GoogLeNet	98.99%	99.12%	97.53%	1276

Table 5: Experimental results of five CNN models on set B

Model	Rec	Se	Acc	PC
This work	99.36%	99.48%	98.20%	395
AlexNet	95.78%	98.43%	95.49%	805
MobileNet	95.75%	97.98%	93.91%	345
VGGNet	96.93%	98.89%	96.92%	1535
GoogLeNet	99.05%	99.12%	97.98%	1276

According to Tabs. 4 and 5, it can be seen that PDNet and MobileNet have the approximate number of parameters, but AlexNet, VGGNet, and GoogleNet use two-fold parameters than the

PDNet at least. The PDNet model achieves 97.71% and 98.20% on accuracy over two sets and outperforms other CNN models. Although MobileNet uses 10% the number of parameters less than the PDNet, PDNet improves 4.75% of accuracy than MobileNet. Hence, the PDNet model trades off better than other CNN models between accuracy and model size. For the length of ECG signals, it also can be inferred that the length of ECG signals is longer can result in higher accuracy because long ECG sample contains a complete heartbeat sample of ECG in set B, but set A only contains partial heartbeat sample in one-second.

Tabs. 6 and 7 show a comparison between the improved loss, CE, and focal loss (FL) on set A and set B, respectively. The improved loss increases approximately 0.5% of accuracy compared to FL and CE, which confirms the effectiveness of the improved loss.

Table 6: Experimental results of the improved loss, CE and FL on set A

Loss	Rec	Se	Acc
IL	99.34%	99.37%	97.71%
CE	98.77%	99.45%	97.26%
FL	98.80%	99.36%	97.28%

Table 7: Experimental results of the improved loss, CE and FL on set B

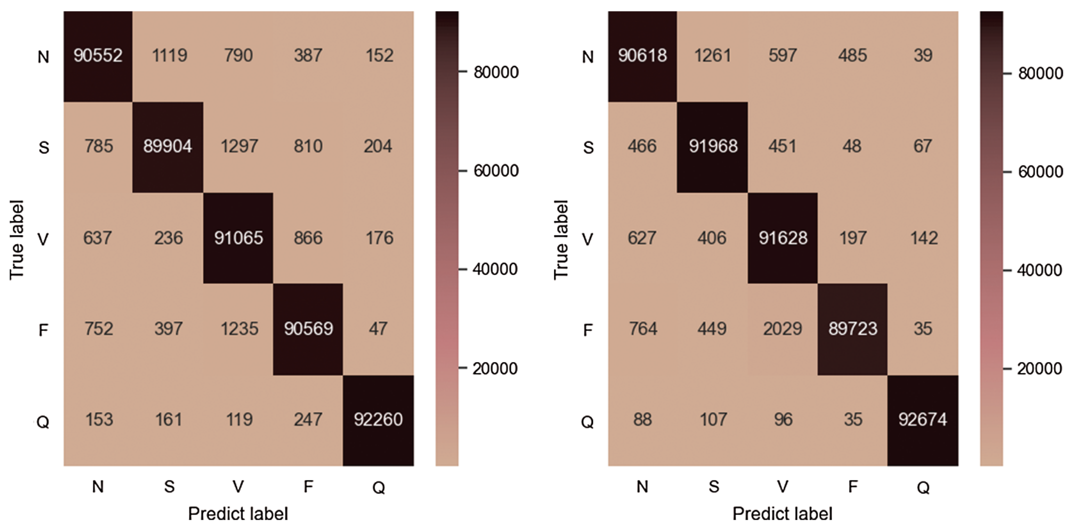
Loss	Rec	Se	Acc
IL	99.36%	99.48%	98.20%
CE	98.77%	99.25%	97.57%
FL	98.83%	99.31%	97.62%

Table 8: Comparison of original and noise free ECGs on five types of arrhythmias

Reference	Type	Method	Acc	Rec	Se
This work	Original	CNN	98.20%	99.36%	99.48%
This work	Original	CNN	97.71%	99.34%	99.37%
[32]	Original	CNN	93.47%	96.01%	97.87%
[32]	Noise free	CNN	94.03%	97.86%	96.71%
[7]	Noise free	CNN	99.00%	98.90%	93.90%
[50]	Noise free	CNN + LSTM	98.10%	98.70%	97.50%
[51]	Noise free	LSTM	99.39%		
[34]	Noise free	DNN	93.4%		
[35]	Noise free	CNN-LSTM	99.26%		
[52]	Noise free	LS-SVM	93.76%	99.13%	99.76%
[52]	Noise free	PCA + NN	94.52%	99.36%	98.61%
[53]	Noise free	DCT + PCA + PNN	99.58%	99.79%	98.69%
[54]	Noise free	DCT + PCA + PNN	94.61%	99.73%	94.67%
[15]	Noise free	SVM	98.39%	99.87%	99.69%

Tab. 8 presents the results of heart arrhythmia detection on five- class among the PDNet and previous advanced methods. It shows that the PDNet works better previous CNN models on original ECG signals. It is obvious that the proposed PDNet can achieve competitive performance on original ECG dataset compared with other methods on noise-free ECG dataset.

To make the overall classification results of the PDNet model on the balanced datasets understandable, Figs. 6a and 6b provide the five-classification results of confusion matrices of the PDNet on dataset A and dataset B. According to Figs. 6a and 6b, the type of F (fusion) achieves the worst results of recall with 96% and 97% respectively. V (ventricular ectopic) gets the worst results of sensitivity (precision) with 96% and 97% on set A and set B, respectively.



(a) (b)

Figure 6: Confusiuon matrices for S, N, V, F and Q five types of arrhythmias on balanced dataset A and B respectively

Table 9: Experimental results of the PDNet model on imbalanced sets

Set	Acc	Rec	Se
Set A	95.95%	79.49%	96.54%
Set B	96.82%	82.38%	98.26%

The PDNet model is capable of achieving excellent classification performance on balanced datasets, but the class distribution of ECG data is skewed in real applications. Hence, it is important and necessary to train CNN models on the imbalanced dataset and achieve classification results for comparison. Tab. 9 shows the classification performance of the PDNet on imbalanced datasets. Alhough the PDNet model obtains accuracy with 95.95% and 96.82% on two sets respectively, it only gets 79.49% and 82.36% on recall, because recall is an important evaluation measure to verify the classification performance of a method in arrhythmia detection. Figs. 7a and 7b provide classification performance of each class on imbalanced set A and imbalanced set B respectively through two confusion matrices. The recalls of F type on two sets are 31%

and 59%, and S class obtains the same recall with 41% on imbalanced set A and imbalanced set B respectively as you can see from Fig. 7. The PDNet reduces recall of F class by as much as 65% and 38% on the imbalanced sets. One explanation for the results is that the number of F class is the minimum and the classifier easily classifies true F class into other four classes (N, S, V, Q). In the future, it is necessary to develop methods to improve arrhythmia detection results on imbalanced ECG datasets, and we also would combine the time-frequency analysis methods [55,56] with deep learning methods to enhance the stability and interpretability of deep learning methods in medical fields.

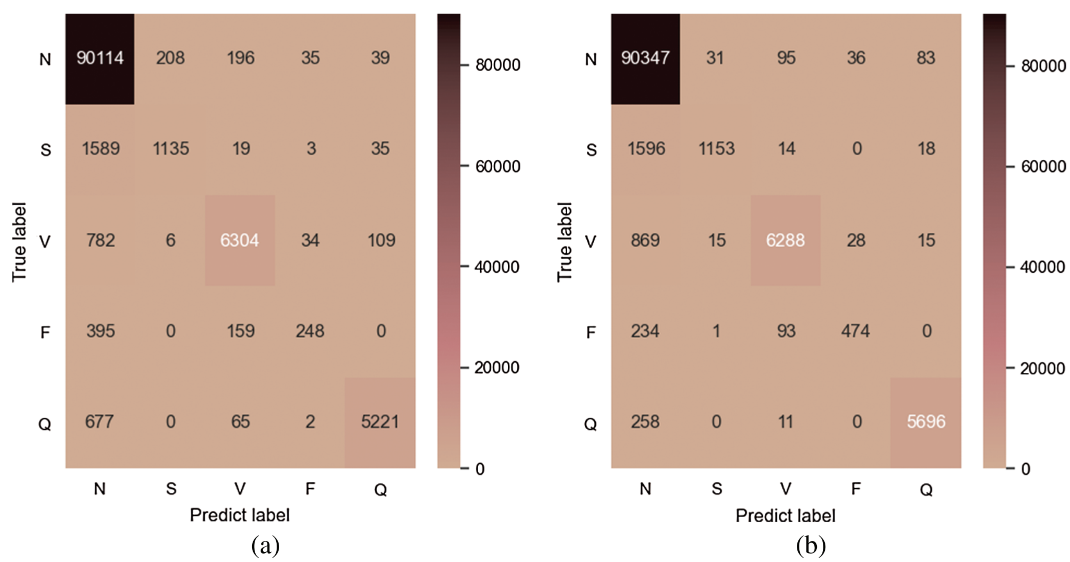


Figure 7: Confusiuon matrices for S, N, V, F and Q five types of arrhythmias on imbalanced dataset A and B respectively

6 Conclusion and Future Work

This work presents a convolutional neural network (CNN) named PDNet to detect different types of heart arrhythmias automatically. We propose a convolutional block called PDblock based on a pointwise convolution method and a depthwise convolution method. To further improve the heart arrhythmia detection results, an improved loss function is utilized. The extensive experiments are conducted on the public MIT-BIH arrhythmia database. The results show that the proposed methods achieve 97.71% and 98.20% of accuracy on two balanced datasets and outperform previous state-of-the-art methods. The PDNet also achieves 95.95% and 96.82% on original imbalanced datasets, which demonstrates that the PDNet is insensitive to the original ECG signals on MIT-BIH arrhythmia database. Compared with state-of-the-art CNN models, the PDNet gets a better balance between accuracy and the number of parameters. Hence, the proposed PDblock model has the potential to be embedded in the ECG-based CAD system on small intelligent devices, which can be utilized to diagnose heart arrhythmias and reduce subjective errors of clinicians in the diagnosis.

In the future, we would develop an ECG based CAD system and deploy the proposed PDNet on it to test the effectiveness of the PDNet based on real-time collected ECG signals through small intelligent devices.

Funding Statement: The authors received no specific funding for this study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

1. World Health Organization (2017). *Cardiovascular diseases (CVDS) fact sheet*. World Health Organization.
2. O'Donnell, M. J., Chin, S. L., Rangarajan, S., Xavier, D., Liu, L. et al. (2016). Global and regional effects of potentially modifiable risk factors associated with acute stroke in 32 countries (interstroke): a case-control study. *Lancet, 388(10046),* 761–775. DOI 10.1016/S0140-6736(16)30506-2.
3. AAMI EC57 (1998). *Testing and reporting performance results of cardiac rhythm and st segment measurement algorithms*. Arlington, VA: Association for the Advancement of Medical Instrumentation.
4. Silva, I., Moody, G. B., Celi, L. (2011). Improving the quality of ECGs collected using mobile phones: the physionet/computing in cardiology challenge 2011. *Computing in Cardiology,* 273–276.
5. Eskofier, B., Kugler, P., Melzer, D., Kuehner, P. (2012). Embedded classification of the perceived fatigue state of runners: towards a body sensor network for assessing the fatigue state during running. *2012 Ninth International Conference on Wearable and Implantable Body Sensor Networks,* 113–117.
6. Hadjem, M., Salem, O., Nat-Abdesselam, F. (2014). An ECG monitoring system for prediction of cardiac anomalies using WBAN. *IEEE 16th International Conference on e-Health Networking, Applications and Services (Healthcom),* 441–446.
7. Kiranyaz, S., Ince, T., Gabbouj, M. (2015). Real-time patient-specific ECG classification by 1-d convolutional neural networks. *IEEE Transactions on Biomedical Engineering, 63(3),* 664–675. DOI 10.1109/TBME.2015.2468589.
8. Li, R., Zhang, X., He, Z., Shi, D., Zhao, H. et al. (2019). A lightweight neural network to detect arrhythmias. In: Wu, C. Q.; Chyu, M. C.; Lloret, J.; Li, X., (eds.), *Proceedings of the 2nd International Conference on Healthcare Science and Engineering*. Singapore: Springer Singapore.
9. Chollet, F. (2017). Xception: deep learning with depthwise separable convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,* 1251–1258.
10. Zhang, T., Qi, G. J., Xiao, B., Wang, J. (2017). Interleaved group convolutions. *Proceedings of the IEEE International Conference on Computer Vision,* 4373–4382.
11. Lee, C. Y., Xie, S., Gallagher, P., Zhang, Z., Tu, Z. (2015). Deeply-supervised nets. *Artificial Intelligence and Statistics, 2015,* 562–570.
12. Goldberger, A. L., Amaral, L. A., Glass, L., Hausdorff, J. M., Ivanov, P. C. et al. (2000). Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. *Circulation, 101(23),* e215–e220.
13. Krizhevsky, A., Sutskever, I., Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems, 2012,* 1097–1105.
14. Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W. et al. (2017). Mobilenets: efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv: 1704.04861.
15. Sahoo, S., Kanungo, B., Behera, S., Sabut, S. (2017). Multiresolution wavelet transform based feature extraction and ECG classification to detect cardiac abnormalities. *Measurement, 108,* 55–66. DOI 10.1016/j.measurement.2017.05.022.
16. Li, H., Yuan, D., Wang, Y., Cui, D., Cao, L. (2016). Arrhythmia classification based on multi-domain feature extraction for an ECG recognition system. *Sensors, 16(10),* 1744. DOI 10.3390/s16101744.
17. Martis, R. J., Acharya, U. R., Min, L. C. (2013). ECG beat classification using PCA, LDA, ICA and discrete wavelet transform. *Biomedical Signal Processing and Control, 8(5),* 437–448. DOI 10.1016/j.bspc.2013.01.005.

18. Martis, R. J., Acharya, U. R., Mandana, K., Ray, A. K., Chakraborty, C. (2012). Application of principal component analysis to ECG signals for automated diagnosis of cardiac health. *Expert Systems with Applications, 39(14),* 11792–11800. DOI 10.1016/j.eswa.2012.04.072.
19. Kang, C., Yu, X., Wang, S., Guttery, D., Pandey, H. et al. (2020). A heuristic neural network structure relying on fuzzy logic for images scoring. *IEEE Transactions on Fuzzy Systems,* 1. DOI 10.1109/TFUZZ.91.
20. Alonso-Atienza, F., Morgado, E., Fernandez-Martinez, L., Garca-Alberola, A., Rojo-Alvarez, J. L. (2013). Detection of life-threatening arrhythmias using feature selection and support vector machines. *IEEE Transactions on Biomedical Engineering, 61(3),* 832–840. DOI 10.1109/TBME.2013.2290800.
21. Shi, H., Wang, H., Huang, Y., Zhao, L., Qin, C. et al. (2019). A hierarchical method based on weighted extreme gradient boosting in ECG heartbeat classification. *Computer Methods and Programs in Biomedicine, 171,* 1–10. http://www.sciencedirect.com/science/article/pii/S016926071831900X.
22. LeCun, Y., Bengio, Y., Hinton, G. (2015). Deep learning. *Nature, 521(7553),* 436–444.
23. Zubair, M., Kim, J., Yoon, C. (2016). An automated ECG beat classification system using convolutional neural networks. *6th International Conference on IT Convergence and Security,* 1–5.
24. Acharya, U. R., Fujita, H., Oh, S. L., Raghavendra, U., Tan, J. H. et al. (2018). Automated identification of shockable and non-shockable life-threatening ventricular arrhythmias using convolutional neural network. *Future Generation Computer Systems, 79,* 952–959. DOI 10.1016/j.future.2017.08.039.
25. Acharya, U. R., Fujita, H., Oh, S. L., Hagiwara, Y., Tan, J. H. et al. (2019). Deep convolutional neural network for the automated diagnosis of congestive heart failure using ECG signals. *Applied Intelligence, 49(1),* 16–27. DOI 10.1007/s10489-018-1179-1.
26. Oh, S. L., Ng, E. Y., San Tan, R., Acharya, U. R. (2019). Automated beat-wise arrhythmia diagnosis using modified u-net on extended electrocardiographic recordings with heterogeneous arrhythmia types. *Computers in Biology and Medicine, 105,* 92–101. DOI 10.1016/j.compbiomed.2018.12.012.
27. Dinakarrao, S. M. P., Jantsch, A., Shafique, M. (2019). Computer-aided arrhythmia diagnosis with bio-signal processing: a survey of trends and techniques. *ACM Computing Surveys, 52(2),* 1–37. DOI 10.1145/3297711.
28. Hannun, A. Y., Rajpurkar, P., Haghpanahi, M., Tison, G. H., Bourn, C. et al. (2019). Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nature Medicine, 25(1),* 65–69. DOI 10.1038/s41591-018-0268-3.
29. Tan, J. H., Hagiwara, Y., Pang, W., Lim, I., Oh, S. L. et al. (2018). Application of stacked convolutional and long short-term memory network for accurate identification of CAD ECG signals. *Computers in Biology and Medicine, 94,* 19–26. DOI 10.1016/j.compbiomed.2017.12.023.
30. Yıldırım, Ö., Pławiak, P., Tan, R. S., Acharya, U. R. (2018). Arrhythmia detection using deep convolutional neural network with long duration ECG signals. *Computers in Biology and Medicine, 102,* 411–420.
31. Acharya, U. R., Fujita, H., Oh, S. L., Hagiwara, Y., Tan, J. H. et al. (2017). Application of deep convolutional neural network for automated detection of myocardial infarction using ECG signals. *Information Sciences, 415,* 190–198. DOI 10.1016/j.ins.2017.06.027.
32. Acharya, U. R., Oh, S. L., Hagiwara, Y., Tan, J. H., Adam, M. et al. (2017). A deep convolutional neural network model to classify heartbeats. *Computers in Biology and Medicine, 89,* 389–396. DOI 10.1016/j.compbiomed.2017.08.022.
33. He, Z., Zhang, X., Cao, Y., Liu, Z., Zhang, B. et al. (2018). Litenet: lightweight neural network for detecting arrhythmias at resource-constrained mobile devices. *Sensors, 18(4),* 1229. DOI 10.3390/s18041229.
34. Wang, H., Shi, H., Lin, K., Qin, C., Zhao, L. et al. (2020). A high-precision arrhythmia classification method based on dual fully connected neural network. *Biomedical Signal Processing and Control, 58,* 101874.
35. Shi, H., Qin, C., Xiao, D., Zhao, L., Liu, C. (2020). Automated heartbeat classification based on deep neural network with multiple input layers. *Knowledge-Based Systems, 188,* 105036. DOI 10.1016/j.knosys.2019.105036.

36. Lehnert, T., Heider, D., Leicht, H., Heinrich, S., Corrieri, S. et al. (2011). Health care utilization and costs of elderly persons with multiple chronic conditions. *Medical Care Research and Review, 68(4),* 387–420. DOI 10.1177/1077558711399580.
37. Gradl, S., Kugler, P., Lohmüller, C., Eskofier, B. (2012). Real-time ECG monitoring and arrhythmia detection using android-based mobile devices. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society,* 2452–2455.
38. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z. (2016). Rethinking the inception architecture for computer vision. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,* 2818–2826.
39. Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A. et al. (2018). Recent advances in convolutional neural networks. *Pattern Recognition, 77,* 354–377. DOI 10.1016/j.patcog.2017.10.013.
40. Zhang, X., Zhao, H., Zhang, S., Li, R. (2019). A novel deep neural network model for multi-label chronic disease prediction. *Frontiers in Genetics, 10,* 3532. DOI 10.3389/fgene.2019.00351.
41. Lin, M., Chen, Q., Yan, S. (2013). Network in network. arXiv preprint arXiv: 1312.4400.
42. Lu, S., Lu, Z., Zhang, Y. D. (2019). Pathological brain detection based on alexnet and transfer learning. *Journal of Computational Science, 30,* 41–47. DOI 10.1016/j.jocs.2018.11.008.
43. He, K., Zhang, X., Ren, S., Sun, J. (2015). Delving deep into rectifiers: surpassing human-level performance on imagenet classification. *Proceedings of the IEEE International Conference on Computer Vision,* 1026–1034.
44. Wang, S. H., Muhammad, K., Hong, J., Sangaiah, A. K., Zhang, Y. D. (2020). Alcoholism identification via convolutional neural network based on parametric ReLU, dropout, and batch normalization. *Neural Computing & Applications, 32(3, SI),* 665–680. DOI 10.1007/s00521-018-3924-0.
45. Bouvrie, J. (2006). Notes on convolutional neural networks. http://cogprints.org/5869/1/cnn_tutorial.pdf.
46. Wang, S., Zhang, Y. D., Yang, M., Liu, B., RamÃrez, J. et al. (2019). Unilateral sensorineural hearing loss identification based on double-density dual-tree complex wavelet transform and multinomial logistic regression. *Integrated Computer-Aided Engineering, 26,* 1–16.
47. Kingma, D. P., Ba, J. (2014). Adam: a method for stochastic optimization. arXiv: 1412.6980. DOI corr abs/1412.6980.
48. Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z. et al. (2016). Tensorflow: large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv: 1603.04467.
49. Simonyan, K., Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. *Computing in Cardiology, International Conference on Learning Representations,* arXiv preprint arXiv: 1409.1556.
50. Oh, S. L., Ng, E. Y., Tan, R. S., Acharya, U. R. (2018). Automated diagnosis of arrhythmia using combination of CNN and LSTM techniques with variable length heart beats. *Computers in Biology and Medicine, 102,* 278–287. DOI 10.1016/j.compbiomed.2018.06.002.
51. Yildirim, Ö. (2018). A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification. *Computers in Biology and Medicine, 96,* 189–202. DOI 10.1016/j.compbiomed.2018.03.016.
52. Martis, R. J., Acharya, U. R., Lim, C. M., Mandana, K., Ray, A. K. et al. (2013). Application of higher order cumulant features for cardiac health diagnosis using ECG signals. *International Journal of Neural Systems, 23(04),* 1350014. DOI 10.1142/S0129065713500147.
53. Martis, R. J., Acharya, U. R., Lim, C. M., Suri, J. S. (2013). Characterization of ECG beats from cardiac arrhythmia using discrete cosine transform in PCA framework. *Knowledge-Based Systems, 45,* 76–82. DOI 10.1016/j.knosys.2013.02.007.
54. Li, T., Zhou, M. (2016). ECG classification using wavelet packet entropy and random forests. *Entropy, 18(8),* 285. DOI 10.3390/e18080285.

55. Tao, J., Qin, C., Xiao, D., Shi, H., Ling, X. et al. (2020). Timely chatter identification for robotic drilling using a local maximum synchrosqueezing-based method. *Journal of Intelligent Manufacturing, 31(5),* 1243–1255.
56. Tao, J., Qin, C., Xiao, D., Shi, H., Liu, C. (2019). A pre-generated matrix-based method for real-time robotic drilling chatter monitoring. *Chinese Journal of Aeronautics, 32(12),* 2755–2764. DOI 10.1016/j.cja.2019.09.001.