Journal of Medical Systems (2020) 44:35
https://doi.org/10.1007/s10916-019-1511-2

MOBILE & WIRELESS HEALTH

An Improved Convolutional Neural Network Based Approach for Automated Heartbeat Classification

Haoren Wang [1] · Haotian Shi [1] · Xiaojun Chen [1] · Liqun Zhao [2] · Yixiang Huang [1] · Chengliang Liu [1]

Received: 8 June 2019 / Accepted: 20 November 2019

Abstract

With age, our blood vessels are prone to aging, which induces cardiovascular disease. As an important basis for diagnosing heart disease and evaluating heart function, the electrocardiogram (ECG) records cardiac physiological electrical activity. Abnormalities in cardiac physiological activity are directly reflected in the ECG. Thus, ECG research is conducive to heart disease diagnosis. Considering the complexity of arrhythmia detection, we present an improved convolutional neural network (CNN) model for accurate classification. Compared with the traditional machine learning methods, CNN requires no additional feature extraction steps due to the automatic feature processing layers. In this paper, an improved CNN is proposed to automatically classify the heartbeat of arrhythmia. Firstly, all the heartbeats are divided from the original signals. After segmentation, the ECG heartbeats can be inputted into the first convolutional layers. In the proposed structure, kernels with different sizes are used in each convolution layer, which takes full advantage of the features in different scales. Then a max-pooling layer followed. The outputs of the last pooling layer are merged and as the input to fully-connected layers. Our experiment is in accordance with the AAMI inter-patient standard, which included normal beats (N), supraventricular ectopic beats (S), ventricular ectopic beats (V), fusion beats (F), and unknown beats (Q). For verification, the MIT arrhythmia database is introduced to confirm the accuracy of the proposed method, then, comparative experiments are conducted. The experiment demonstrates that our proposed method has high performance for arrhythmia detection, the accuracy is 99.06%. When properly trained, the proposed improved CNN model can be employed as a tool to automatically detect different kinds of arrhythmia from ECG.

Keywords Electrocardiogram (ECG) · Heartbeat classification · Signal processing · Convolutional neural networks · MIT database

Introduction

According to the report by the United Nations in 2015 [1], the world is now facing the challenge of the aging population. As we grow older, our cardiovascular system is prone to be more vulnerable and more dysfunctional. Cardiovascular disease (CVD) may be the major cause of abnormal heart rhythms. Abnormalities in electrical impulses during any conductive process can result in arrhythmia. For the pre-diagnosis of arrhythmia, many methods are used for early monitoring and control, which have ideal effects. However, partial arrhythmias have quite complicated causes, which may occur suddenly and lead to mortality. Therefore, the real-time monitoring of arrhythmia is of significant importance. Some automated detection algorithms have been proposed for arrhythmias diagnosis [2, 3], and these systems can reduce the workload of doctors through an automatic diagnosis system.

Cardiac arrhythmia is defined as an irregular heartbeat, and the detection of electrocardiograph (ECG) is a practical and effective method for cardiac arrhythmia diagnosis. There is a great relationship between electrocardiogram and potential changes in cardiomyocytes. Also, the electrical impulse conducts downward, forming a QRS complex in the left and right ventricles, which is highly related to ventricular diseases.

This article is part of the Topical Collection on *Mobile & Wireless Health*

✉ Chengliang Liu
chlliu@sjtu.edu.cn

1 School of Mechanical Engineering, Shanghai Jiao Tong University, 800 Dongchuan Road, Shanghai 200240, People's Republic of China

2 Department of Cardiology, Shanghai First People's Hospital Affiliated to Shanghai Jiao Tong University, 100, Haining Road, Shanghai 200080, People's Republic of China

Published online: 18 December 2019

The selection of features is the key to computer-aided diagnosis (CAD). Statistic features [4] such as maximum, minimum, kurtosis, skewness and higher-order statistic feature [5] of certain periods are commonly used in the previous literature. Other features included morphological features [6], position, amplitude, and interval features [7], wavelet packet entropy [8], abstract features [9], auto-encoder [10], entropy features (state space correlation entropy [11], and dispersion entropy [12]) are added in recent work.

For classification, some methods have been discussed in previous literature. For traditional machine learning method, K-nearest neighbor(KNN) [13], decision tree [4], and support vector machine(SVM) [14]. Some researchers tried novel classifiers, such as random forest (RF) [15], and extreme gradient boosting (XGBoost) [16]. With the rapid development of artificial intelligence, neural networks have become an interdisciplinary field. Therefore, methods based on artificial neural networks (ANN) have been applied in the arrhythmia classification. These methods include multilayer perceptron (MLP) [17], fuzzy neural network [18], radial basis function (RBF) neural network [19], self-organized neural networks [20], and restricted Boltzmann machine (RBM) [21].

Although many classification methods have been proposed to improve the quality of features, there still are some defects in these algorithms. On the one hand, the effect of classification is highly related to selected features. Therefore, the choice of features is susceptible to subjective factors. On the other hand, these methods often suffer from overfitting [22]. For comparison, deep learning architecture shows advantages over traditional methods.

In recent deep learning networks, many novel neural networks with inherent features, which are characterized by automatic feature extraction and classification steps, are introduced. Convolutional neural networks (CNN) [23, 24], long short-term memory network [25, 26] (LSTM), and recurrent neural network (RNN) [27] are proposed separately. For each convolutional layer, the kernel size is fixed. However, using different kernel sizes of each convolutional layer will be prone to increase the effect of features in a certain dimension, which is not mentioned in previous work.

In this paper, an improved CNN method is proposed. Primarily, the ECG data is preprocessed to obtain usable heartbeats. Afterward, the processed heartbeats are processed as the input vectors to the first layer of the convolutional neural network. No feature extraction process is conducted in the preprocessing stage. In the convolutional layers, different kernel sizes are applied in each signal convolution layer. In the subsequent pooling layers, the max-pooling operation is applied to all the pooling layers. The fully connected neural network is the final output layer of the convolutional neural network (CNN), the final feature vectors are merged and set as the input vectors of the first fully connected layers. Finally, the results are processed by the last fully connected layer with the Softmax function. To verify the effectiveness of the proposed structure, a comparative experiment is conducted with a traditional CNN structure.

The main innovativeness of this article is the improvement of the traditional CNN. To our knowledge, it is the first application of CNNs with multiple kernel sizes and variable stride length in ECG analysis. Secondly, the comparison of four different configurations is also the contribution of the articles. Addition experiments with different ratios of training sets are also introduced in our experiments. The results show that the larger training set tends to have better results. Moreover, cross-validation experiments are also added to verify the stability of our proposed system.

The remainder of this paper is organized as follows. The public ECG database is introduced for verification in Section 2. The proposed arrhythmia diagnosis method is presented in Section 3. The architecture of the proposed CNN framework is also introduced in detail. Section 4 is the results and discussion. Finally, the conclusion is drawn, and future work prospects are shown in Section 5.

Data Used

For the diagnosis of arrhythmia, the main difficulty is that different people have different forms of heartbeat waveform. The classification of heartbeats between different individuals is the key to disease diagnosis. For different kinds of diseases, the electrocardiogram varies far. To classify different types of heartbeats, the Advancement of Medical Instrumentation (AAMI) [28] proposed a standard to evaluate the performance of algorithms. The heartbeats in our experiments are firstly divided according to the AAMI standard. Four records, which contained paced beats, are removed in this study. The heartbeats in the rest 44 records are divided into five groups, including Non-ectopic (N), Supraventricular ectopic (S), Ventricular ectopic (V), Fusion (F), and Unclassified beats (Q). For the heartbeats division, three different division strategies are considered for data division. The details of the heartbeats division are shown in table 1.

As a universal standard for the diagnostic algorithm of arrhythmia, the Massachusetts Institute of Technology-Beth Israel Hospital (MIT-BIH) arrhythmia database [29] is introduced for verification. There are 48 lead II records from 47 individuals in which 25 men aged 32–89, and 22 women aged 23–89 were recorded in this program. For each recording, the sampling frequency is 360 Hz, and the recording time lasts for about 30 min. All the records are labeled by heartbeats and annotated by experts in the ECG field.

Methodology

The flowchart of the proposed method is shown in fig. 1. The arrhythmia diagnosis system contains three major stages: data preprocessing, datasets division, and heartbeats classification.

Table 1 Mapping the MIT-BIH arrhythmia database heartbeat types

AAMI heartbeat classes	N	S	V	F	Q
Description	Any heartbeat not in the S, V, F or Q classes	Supraventricular ectopic beat	Ventricular ectopic beat	Fusion beat	Unknown beat
MIT-BIH heartbeat types	Normal beat (NOR)	Atrial premature beat (AP)	Premature ventricular contraction (PVC)	Fusion of ventricular and normal beat (fVN)	Paced beat(P)
	Left bundle branch block beat (LBBB)	Aberrated atrial premature beat (aAP)	Ventricular escape beat (VE)		Fusion of paced and normal beat (fPN)
	Right bundle branch block beat (RBBB)	Nodal (junctional) premature beat (NP)			Unclassified beat(U)
	Atrial escape beats (AE)	Supraventricular premature beat (SP)			
	Nodal (junctional) escape beat (NE)				

Data Preprocessing

Our system is aiming at classifying a single heartbeat. Due to the annotation algorithm have sufficient accuracy in the MIT-BIH arrhythmia database, R-points are used as the fiducial points (FP). The division of heartbeats is on the basis of the FPs, and the Pan-Tompkins algorithm [30] is used to determine the position of the FPs. In our experiment, heartbeats are determined 0.4 s before and 0.5 s after the FPs.

Dataset Division

In dataset division stage, the heartbeats are divided by three data division strategies to verify the robustness of the model, different proportions of heartbeats are divided into corresponding folds. To verify the impact of filtering on classification results, after 5:5 training to test division, the first fold is filtered for further processing. The unfiltered fold and filtered fold are both used as the input of the proposed improved CNN model.

Heartbeats Classification

CNN is a deep neural network structure [22] consisting of the input and output layers and multiple hidden layers. The convolutional neural network constructs the visual perception mechanism of the creature, which can perform both supervised learning and unsupervised learning. The convolutional kernel parameters sharing the sparseness of the inter-layer connection of the hidden layer, which enables the convolutional neural network to smaller computational weights for latticed (grid-like topology) features, have a stabilizing effect for feature extraction of data. Due to the dimension of ECG signals, the network for ECG heartbeat is different from the CNNs used in many image processing. One-dimensional (1-D) convolutional neural network is introduced in this study. There are three basic layers containing in CNNs, including convolutional layers, pooling layers, and fully connected layers with active functions.

In one convolution layer of traditional CNN architecture, fixed kernels with the same size are employed. All features are trained and convoluted on the same scale. To increase the perception of different scales, the diversity of features will be increased by different kernel sizes. Therefore, discrimination can be better reflected in different kernel sizes.

The structure of the improved CNN is shown in Fig. 2. 7 layers in this network are added in the proposed CNN: 2 convolution layers, 2 pooling layers, and 3 fully connected layers. The input data is firstly reshaped to fit the interface of our model, and processed by the following architecture. Layer 1 and layer 3 are convolution layers. Kernels of three different sizes are adjusted in each convolution layer to fit different scales. For the first convolution layer, kernels with three different sizes (4, 6, and 8) are used. The number of kernels for each size is 8. The sizes of the kernels for the second convolution layer are 6, 8, and 10 respectively. This layer contained 24 kernels. The stride for the two convolution layers is set at 1. After each convolution layer, a max-pooling of size 2 is applied to the obtained feature maps. The pooling layers also provide three sizes of outputs. The output of the

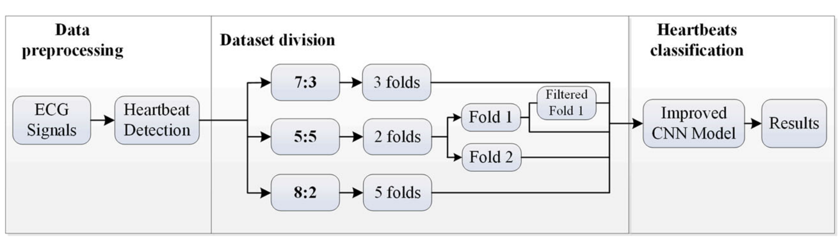

Fig. 1 The flowchart of the proposed arrhythmia diagnosis method

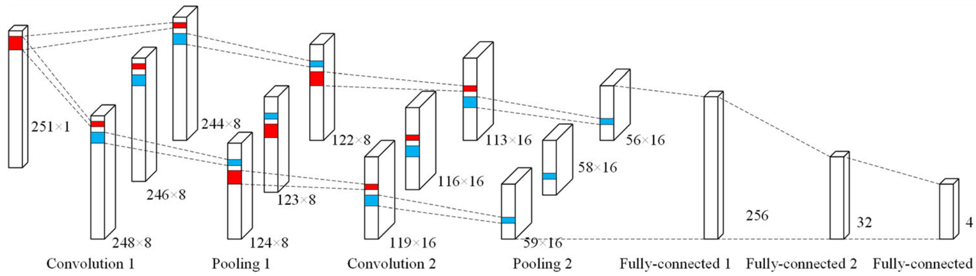

Fig. 2 The structure of the proposed CNN

fourth layer is combined and inputted to the layer 5. The number of neurons in the three fully connected layers (layers 5, 6, and 7) is 256, 32, and 4. For the two convolution layers and the first two fully-connected layers, the rectified linear units (RELU) are taken as the activation functions. The Softmax function is taken as the activation function of layer 7. The four neurons classified by layer 7 corresponded to class N, S, V, and F. The CNN is trained with a batch size of 64. The learning rate is 0.01. The total training process is carried out in 100 epochs.

The Results and Discussion

The Results of Experiments

A controlled experiment by different division strategies

As described above, the proposed algorithm is verified by the MIT-BIH arrhythmia database. The overall accuracy (Acc) is selected for evaluating the performance of the classifier in our research. However, the overall accuracy is easily influenced by the imbalanced dataset. The sensitivity (Sen), and positive predictive value (PPV) are added for comparison.

$$Se = \frac{TP}{TP + FN} \times 100\% \tag{1}$$

$$+P = \frac{TP}{TP + FP} \times 100\% \tag{2}$$

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \tag{3}$$

Where TP is the true positive beats (correctly detected heartbeats) value, FN means false negative beats (the total number of falsely missed heartbeats), and FP is the false positive beats (falsely classified positive heartbeats).

In order to better demonstrate the robustness of the model, cross-validation experiments are designed. The k-fold cross-validation reduces the variance by averaging the results of k different training groups, thus the performance is less sensitivity to the partition of the dataset. Three different data division strategies are conducted in our experiment. As a typical training to test set ratio, the training and test dataset is firstly divided by 7:3 in Experiment 1. In this ratio, 3-fold cross-validation is conducted. The cross-validation diagram for three strategies is shown in Fig. 3 and the steps are as follows:

Step 1: Raw data is randomly divided into 10 shares without repeat sampling.

Step 2: The last 3 shares are used as the test set, and the first 7 shares are used as the training set in the first iteration.

Step 3: Each share has an opportunity to serve as a test set and the remaining shares serve as a training set. Step 2 is repeated 2 times and each share will only be taken as test set once.

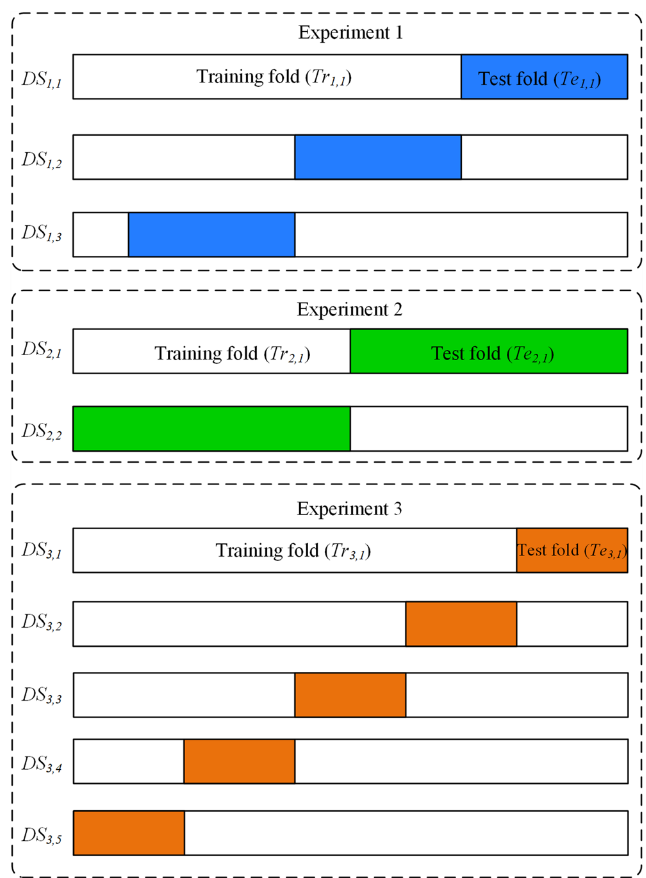

Figure 3. the data partition diagram of three experiments.

Step 4: After training on each training set, a model is obtained, which is tested on the corresponding test set. The evaluation results of each model are saved.

Step 5: The average value of the 3-group test results is calculated as an estimate model accuracy, which is used as the performance criterion of the current 3-fold cross-validation model.

In order to test different proportions of training and test sets, 5:5 and 8:2 are also added for comparison. 2-folds and 5-folds cross-validation are conducted for verification separately. The flow of these two experiments is similar to Experiment 1. Fig. 3 presents the data partition diagram of all three experiments. The exact number of heartbeats and classification results by different strategies are listed in Table 2. $DS_{i,j}$ denotes the dataset used in the j^{th} iteration of experiment i. $Tr_{i,j}$ and $Te_{i,j}$ are the corresponding training set and test set.

The results show that experiment 1 (7:3 training to test ratio) achieves the best average accuracy of 98.93%. The accuracy of experiment 2 (5:5 training to test ratio) is 98.67%, while for experiment 3 (8:2 training to test ratio), the accuracy is 98.88%. Compare experiment 2 with experiment 3 (or experiment 1), when the ratio of training to test increases, the accuracy is improved due to the better-trained model. Because the data ratio is close, the results of Experiment 1 and 3 are basically the same.

Our work used python as the programming language with the scikit-learn library and the TensorFlow library. For training the model, the software is applied on a computer equipped with an Intel Core i7 6700K CPU and an NVIDIA GTX 1070 GPU. The memory size of the computer was 16GB. For example, the average training time was 115 μs per step in experiment 1 on each epoch. The test time costed 0.032 s, and the average test time for a single heartbeat was 0.644 μs.

A controlled experiment with traditional CNN

To prove that the perception of different scales can be better reflected by multiple kernel sizes, an additional controlled experiment is added with a traditional CNN. In traditional CNN, each convolution layer only contained kernels with a fixed size. The total kernels of the contrast CNN are the same as the improved CNN in fig. 2, and other parameters in the two networks are all the same. The detail parameters of the contrast CNN are shown in table 3. Moreover, the effect of filtering for classification is also discussed. In the contrast experiment, the dataset in the first iterations of Experiment 2 is chosen to compare the effects of filtering.

Once the dataset $DS_{2,1}$ in experiment 2 is obtained, noises and unrelated frequency components are removed by a

Table 2 the classification results by three different division strategies

	Iterations	Datasets Name	Training and Test Set	N	S	V	F	Total number	Accuracy	Average
Experiment 1 (7:3)	1	$DS_{1,1}$	$Tr_{1,1}$	62,597	2413	4393	777	70,180	99.06	98.93
			$Te_{1,1}$	27,445	366	2614	25	30,450		
	2	$DS_{1,2}$	$Tr_{1,2}$	63,207	1779	5105	462	70,553	98.87	
			$Te_{1,2}$	26,835	1000	1902	340	30,077		
	3	$DS_{1,3}$	$Tr_{1,3}$	63,204	1746	5122	481	70,553	98.85	
			$Te_{1,3}$	26,838	1033	1885	321	30,077		
Experiment 2 (5:5)	1	$DS_{2,1}$	$Tr_{2,1}$	45,824	943	3788	414	50,969	98.67	98.66
			$Te_{2,1}$	44,218	1836	3219	388	49,661		
	2	$DS_{2,2}$	$Tr_{2,2}$	44,218	1836	3219	388	49,661	98.64	
			$Te_{2,2}$	45,824	943	3788	414	50,969		
Experiment 3 (8:2)	1	$DS_{3,1}$	$Tr_{3,1}$	71,987	2228	5641	648	80,504	98.74	98.88
			$Te_{3,1}$	18,055	551	1366	154	20,126		
	2	$DS_{3,2}$	$Tr_{3,2}$	72,085	2215	5567	637	80,504	98.96	
			$Te_{3,2}$	17,957	564	1440	165	20,126		
	3	$DS_{3,3}$	$Tr_{3,3}$	72,023	2231	5598	652	80,504	98.93	
			$Te_{3,3}$	18,019	548	1409	150	20,126		
	4	$DS_{3,4}$	$Tr_{3,4}$	72,096	2203	5571	634	80,504	98.87	
			$Te_{3,4}$	17,946	576	1436	168	20,126		
	5	$DS_{3,5}$	$Tr_{3,5}$	71,987	2228	5641	648	80,504	98.92	
			$Te_{3,5}$	18,065	540	1356	165	20,126		

Annotations: $DS_{i,j}$ denotes the original dataset (without filtering) used in the j^{th} iteration of experiment i. $Tr_{i,j}$ and $Te_{i,j}$ are the corresponding training set and test set.

Table 3 The parameters of the contrast CNN

No.	Layer	Kernel Size	Stride	No. of Output
0	Input	–	–	251 × 1
1	Convolution	4	1	248 × 24
2	Pooling	2	2	124 × 24
3	Convolution	6	1	119 × 48
4	Pooling	2	2	59 × 48
5	Fully connected	–	–	256
6	Fully connected	–	–	32
7	Fully connected	–	–	4

wavelet de-noising method. The datasets are firstly processed by the Db6 wavelet. The sampling frequency is 360 Hz, therefore, the frequency range of the sixth level approximation sub-band coefficient is 0–5.625 Hz. The frequency range of the third level detail sub-band frequency is 45-90 Hz, which contains the key information and frequency components of the original signals. The baseline drift always exists in the frequency components below 5 Hz, and high-frequency noises exist in the range above 90 Hz. Filtering is firstly carried out by replacing wavelet coefficients below 5 Hz and above 90 Hz to zeros, only the coefficients between the 3rd and 6th detail sub-bands are reserved and used for reconstruction. The dataset without filtering is $DS_{2,\ 1}$, and the dataset processed by the wavelet de-noising method is defined as $DS'_{2,1}$.

Moreover, two types of kernel size (single or multiple) and two types of preprocessing of the dataset (with or without filtering) are listed for comparison. Thus, there are a total of four configurations. Table 4 lists the detail of four different configurations. The performance of each configuration is presented in table 5.

Annotations: $DS_{2,\ 1}$ denotes the original dataset (without filtering) used in the 1st iteration of experiment 2 in table 2. $DS'_{2,1}$ denotes the filtered dataset used in the 1st iteration of experiment 2 in table 2.

Table 4 Four different configurations for comparison

Configuration	Kernel size	Dataset
I	single	$DS_{2,\ 1}$ (Raw)
II	single	$DS'_{2,1}$ (Filtered)
III	multiple	$DS_{2,\ 1}$ (Raw)
IV	multiple	$DS'_{2,1}$ (Filtered)

Annotations: $DS_{2,\ 1}$ denotes the original dataset (without filtering) used in the 1st iteration of experiment 2 in table 2. $DS'_{2,1}$ denotes the filtered dataset used in the 1st iteration of experiment 2 in table 2

The overall accuracies, sensitivities, and positive predictive values of the four configurations are illustrated in table 5. The accuracies of the four configurations are 98.21%, 98.10%, 98.67%, and 98.64% separately. The signals without filtering (Configuration I, III) perform better results than filtered ones (Configuration II, IV). Moreover, the signals processed by multiple kernel size (Configuration III, IV) perform better than the single kernel size (Configuration I, II). Configuration III (signals without filtering and multiple kernel sizes) obtains the highest accuracy. The worst results are provided by Configuration II (filtered signals and single kernel sizes). Filtering may filter out a lot of key information. Therefore, the signals without filtering do improve the accuracy of the final results, and the filtering processing will slightly affect the accuracy of classification.

The confusion matrix and the receiver operating characteristics curve

To present the best classification results of the proposed method, a confusion matrix is presented in table 6. The signals without filtering and having a larger training set tend to have better results, the dataset $DS_{1,\ 1}$ (the dataset set used in the first iterations of Experiment 1) is chosen. The confusion matrix is listed in table 7, and the accuracy of this confusion matrix is 99.06%. Most of the heartbeats are correctly classified. However, there are two major mistakes. 238 heartbeats of class S are classified into class N, and 60 ventricular heartbeats are also recognized to be normal. For class S and class F, the sensitivities are not very high. It is due to the fewer samples of these two classes. CNN always needs a large number of parameters to be trained. Hence, more ECG heartbeats are required, especially the classes with a small number of heartbeats.

To reflect the generalization performance of the model, the receiver operating characteristics (ROC) curve is introduced to evaluate performance for different types of diseases. Figure 4 presents the ROC curves for class N, S, V and F, and the area under the ROC curve (AUC) is also calculated. The dataset we used in Fig. 4 is $DS_{1,\ 1}$ in table 2. The results show that the AUC of each class is very close to 1 and the AUC of class V performs better than other classes.

Discussion

Comparisons between previous methods

Table 7 collects recent high precision classification results from the literature. In table 7, the dataset division strategies, selected features and classification algorithms are presented. For heartbeats classification, all methods use the same database (the MIT-BIH arrhythmia database, mitdb) and the same heartbeats division standard (the AAMI standard). Due to the different data processes, each method removes a certain

Table 5 the classification results of four different configurations (5:5 training to test ratio)

Experiments	Acc	N		S		V		F	
		Sen	+P	Sen	+P	Sen	+P	Sen	+P
Configuration I	98.21	99.61	98.62	70.22	92.23	95.08	95.10	63.59	92.05
Configuration II	98.10	98.94	99.18	86.73	74.77	95.28	95.31	67.08	93.08
Configuration III	98.67	99.81	98.89	77.07	96.56	95.37	97.09	73.32	91.02
Configuration IV	98.64	99.72	98.99	80.24	93.52	94.37	97.15	77.31	86.11

number of heartbeats, such as heartbeats mixed with glitches and noises. Therefore, the total number of heartbeats varies by different methods. Our proposed method uses the same database and follows the same standard, having a relatively high accuracy over previous methods.

Advantages of our method and application scenario

The benefit of this improved CNN model are as follows:

i. The neural network is automatic classification, there is no need for hand-selected features.
ii. The approach is insensitive to the quality of ECG signals, thus, filtering can be deployed before processing without significant loss of classification accuracy.
iii. The implementation of cross-validation experiments proves the robustness of the model.

In Table 2, the effect of different dataset division strategies is discussed. Both 7:3 and 8:2 divisions can train the model adequately and have good results. In practical applications, it is more inclined to use 7:3 as the division of training and test set. In this ratio, the training data is sufficient, and the data of the test set is not too small to affect the classification results.

In Table 5, the classification results of different configurations are discussed. For theoretical research, the quality of the original signals is high. Therefore, filtering is not necessary and unfiltered original signals can be directly used for classification. Configuration III is recommended for the theoretical verification of standard databases (e.g. the MIT-BIH arrhythmia database). For most practical situations, filtering removes glitches from the original signals, which makes the signals better reflect actual waveforms. Configuration IV is recommended for actually collected signals.

The network with multiple kernel size has better classification effects than the network with a single kernel regardless of filtering or not. However, because the structure of the neural network with a single kernel size is simple, Configuration I or II can be used as a benchmark to verify whether the proposed neural network is valid. Based on the results of Configuration I or II, the structure of the neural network is optimized to obtain better results according to Configuration III or IV.

Our method can be applied in the mobile diagnosis system. The ECG signal can be obtained from the portable ECG signal receiving devices. Then, the data will be transfer to the cloud. Our proposed algorithm can be deployed in the cloud server and trained in advance. Then, received heartbeats will be processed by a pre-trained model, and the conclusion will be drawn in seconds.

Table 6 A confusion matrix of our proposed method

Real	Prediction			
	n	s	v	f
N	45,027	31	24	12
S	238	1143	6	0
V	60	3	3359	11
F	46	0	37	318

Annotation: the dataset of the confusion matrix is $DS_{1,\ 1}$ in table 2

Conclusions

In order to automated classify ECG heartbeat, the method based on an improved CNN is proposed in this paper. As an effective feature extractor, CNN can extract implicit information without any hand-selected feature extraction. As a high precision classifier, the ECG heartbeats can be inputted into the network directly. An improved 7 layers CNN is proposed in this study. In the proposed CNN structure, kernels with different sizes are used in each convolution layer. Then a max-pooling layer follows. The outputs of the last pooling layer are merged and as the input to fully connected layers.

Three different data division strategies are proposed in this paper, the proposed method provided the highest accuracy of 99.06% by the 7:3 training to test ratio and no filtering process. Compared with an experiment based on a traditional CNN, the results verified the effectiveness of using multiple kernel sizes. The influence of filtering is also discussed. It is demonstrated that our proposed method is insensitive to noise, and filtering can be applied before the method. In future work, we are planning to study the effect of convolution neural networks with different structures. Moreover, we will collect and

Table 7 Comparisons between our method and previous methods

Author	Year	Class	Database	Data Size	Approach	Performance
Huang et al. [31]	2012	8	mitdb	9800	Sparse repression-based classifier	acc: 98.35%
					Independent component correlation algorithm (ICA) and RR interval	
					5:5 training to test ratio	
Elhaj et al. [32]	2016	5	mitdb	110,094	Neural network classifier (NN) and support vector machine (SVM)	acc: 98.90%
					Linear method based wavelets and principal component analysis (PCA)	
					Nonlinear method-based high order statistic and multivariate analysis	
					Ten folds cross-validation	
Kiranyaz et al. [23]	2016	5	mitdb	100,389	An adaptive 1-D convolutional neural network (CNN)	acc: 98.90%
					The back-propagation (BP) training method	sen: 99.00%
					Dividing training and test sets according to the AAMI standard	spec: 93.9%
Chen et al. [33]	2017	5	mitdb	90,808	A SVM model	acc: 98.46%
					Weighted RR interval and project matrix	sen: 98.46%
					5:5 training to test ratio	spec: 98.46%
Acharya et al. [22]	2017	5	mitdb	109,949	A 9 layers CNN	Set A:
					Two sets of experiment (with and without noise removal)	acc: 93.47%
					Generation of synthetic data	sen: 96.01%
					Ten folds cross-validation	spec: 91.64%
						Set B:
						acc: 94.03%
						sen: 96.71%
						spec: 91.54%
Our method	2019	5	mitdb	100,630	An improved 7 layers CNN	acc: 99.06%
					Four sets of the experiment (with and without noise removal, single and multiple kernel size)	
					Cross-validation experiments by three data division strategies	

Annotation: acc: accuracy, sen: sensitivity, spec: specificity

annotate more ECG heartbeats to improve the balance of the dataset.

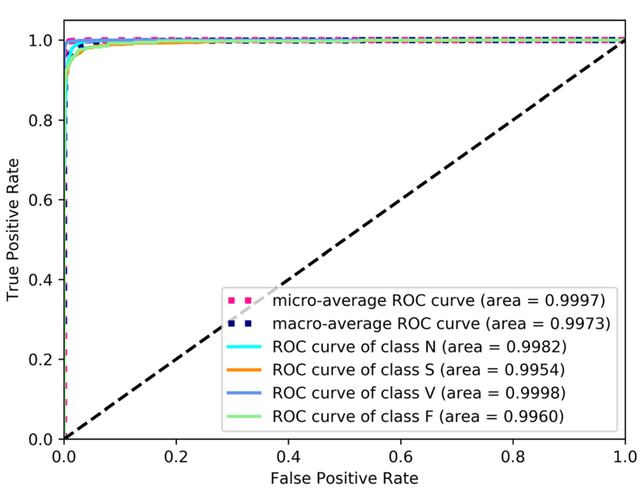

Figure 4 the ROC curves for four different classes (Annotation: the dataset of the ROC curves is $DS_{1,1}$ in table 2).

Acknowledgments Our research is supported by the National Key R&D Program of China (No. 2018YFB1307005), and the major project from Shanghai Municipal Commission of Health and Family Planning (No. 2018ZHYL0226).

Compliance with Ethical Standards

Conflict of Interest All authors declare that there is no conflict of interest in this work.

Ethical Approval This article does not contain any studies with human participants or animals performed by any of the authors.

Informed Consent As the article does not involve Human participants, thus, there is no such informed consent.

References

1. United Nations. Department of economic and social affairs population division. World population aging 2015. New York, 2015.

2. Tekeste, T. et al., Ultra-low power QRS detection and ECG compression architecture for IoT healthcare devices. IEEE Transactions on Circuits and Systems I: Regular Papers 66(2):669–679, 2019.
3. Beach, C. et al., An ultra low power Personalizable wrist worn ECG monitor integrated with IoT infrastructure. IEEE Access 6:44010–44021, 2018.
4. Martis, R. J. et al., Computer aided diagnosis of atrial arrhythmia using dimensionality reduction methods on transform domain representation. Biomedical Signal Processing and Control 13(1):295–305, 2014.
5. Martis, R. J. et al., Application of higher order statistics for atrial arrhythmia classification. Biomedical Signal Processing and Control 8(6):888–900, 2013.
6. De Chazal, P., O'Dwyer, M., and Reilly, R. B., Automatic classification of heartbeats using ECG morphology and heartbeat interval features. IEEE Transactions on Biomedical Engineering 51(7): 1196–1206, 2004.
7. de Lannoy, G., Francois, D., Delbeke, J., and Verleysen, M., Weighted conditional random fields for supervised interpatient heartbeat classification. IEEE Trans Biomed Eng 59(1):241–247, 2012.
8. Li, T. and M. Zhou, ECG classification usingwavelet packet entropy and random forests. Entropy, **18**(8), (2016).
9. Teijeiro, T., Felix, P., Presedo, J., and Castro, D., Heartbeat classification using abstract features from the Abductive interpretation of the ECG. IEEE J Biomed Health Inform 22(2):409–420, 2018.
10. Luo, K., et al., Patient-specific deep architectural model for ECG classification. Journal of Healthcare Engineering, 2017. (2017).
11. Tripathy, R. K., Deb, S., and Dandapat, S., Analysis of physiological signals using state space correlation entropy. Healthcare Technology Letters 4(1):30–33, 2017.
12. Rostaghi, M., and Azami, H., Dispersion entropy: A measure for time-series analysis. IEEE Signal Processing Letters 23(5):610–614, 2016.
13. Faziludeen, S., and Sankaran, P., ECG beat classification using evidential K -nearest Neighbours. Procedia Computer Science 89: 499–505, 2016.
14. Raj, S., Ray, K. C., and Shankar, O., Cardiac arrhythmia beat classification using DOST and PSO tuned SVM. Computer Methods and Programs in Biomedicine 136:163–177, 2016.
15. Clifford, G.D., et al. AF classification from a short single lead ECG recording: The PhysioNet/computing in cardiology challenge 2017. in Computing in Cardiology, (2017).
16. Sodmann, P., et al., A convolutional neural network for ECG annotation as the basis for classification of cardiac rhythms. Physiological Measurement. **39**(10), (2018).
17. Mar, T. et al., Optimization of ECG classification by means of feature selection. IEEE Transactions on Biomedical Engineering 58(8):2168–2177, 2011.
18. Shyu, L. Y., Wu, Y. H., and Hu, W., Using wavelet transform and fuzzy neural network for VPC detection from the Holter ECG. IEEE Transactions on Biomedical Engineering 51(7):1269–1273, 2004.
19. Belkheiri, M., Z. Douidi, and A. Belkheiri, ECG beats extraction and classification using radial basis function neural networks, in Lecture Notes in Electrical Engineering. p. 127–136, (2013).
20. Chen, Y., and Yang, H., Self-organized neural network for the quality control of 12-lead ECG signals. Physiological Measurement 33(9):1399–1418, 2012.
21. Polanía, L. F., and Plaza, R. I., Compressed sensing ECG using restricted Boltzmann machines. Biomedical Signal Processing and Control 45:237–245, 2018.
22. Acharya, U. R. et al., A deep convolutional neural network model to classify heartbeats. Computers in Biology and Medicine 89:389–396, 2017.
23. Kiranyaz, S., Ince, T., and Gabbouj, M., Real-time patient-specific ECG classification by 1-D convolutional neural networks. IEEE Trans Biomed Eng 63(3):664–675, 2016.
24. Acharya, U. R. et al., Automated detection of arrhythmias using different intervals of tachycardia ECG segments with convolutional neural network. Information Sciences 405:81–90, 2017.
25. Tan, J. H. et al., Application of stacked convolutional and long short-term memory network for accurate identification of CAD ECG signals. Comput Biol Med 94:19–26, 2018.
26. Andersen, R. S., Peimankar, A., and Puthusserypady, S., A deep learning approach for real-time detection of atrial fibrillation. Expert Systems with Applications 115:465–473, 2019.
27. Limam, M. and F. Precioso. Atrial fibrillation detection and ECG classification based on convolutional recurrent neural network. in Computing in Cardiology, (2017).
28. ANSI/AAMI, Testing and reporting performance results of cardiac rhythm and ST segment measurement algorithms, Association for the Advancement of Medical Instrumentation (AAMI), 2008, American National Standards Institute, Inc. (ANSI), 2008 ANSI/AAMI/ISO EC57, 1998-(R).
29. Goldberger, A. L., Amaral, L. A., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., Mietus, J. E., Moody, G. B., Peng, C. K., and Stanley, H. E., PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation 101(23):E215–E220, 2000.
30. Pan, J., and Tompkins, W. J., A real-time QRS detection algorithm. IEEE transactions on biomedical engineering 3:230–236, 1985.
31. Huang, H. F., Hu, G. S., and Zhu, L., Sparse representation-based heartbeat classification using independent component analysis. Journal of Medical Systems 36(3):1235–1247, 2012.
32. Elhaj, F. A., Salim, N., Harris, A. R., Swee, T. T., and Ahmed, T., Arrhythmia recognition and classification using combined linear and nonlinear features of ECG signals. Computer Methods and Programs in Biomedicine 127:52–63, 2016.
33. Chen, S. et al., Heartbeat classification using projected and dynamic features of ECG signal. Biomedical Signal Processing and Control 31:165–173, 2017.