Accepted Manuscript

Automatic recognition of arrhythmia based on principal component analysis network and linear support vector machine

Weiyi Yang, Yujuan Si, Di Wang, Buhao Guo

PII: S0010-4825(18)30221-X

DOI: 10.1016/j.compbiomed.2018.08.003

Reference: CBM 3038

To appear in: *Computers in Biology and Medicine*

Received Date: 16 May 2018

Revised Date: 2 August 2018

Accepted Date: 2 August 2018

Please cite this article as: W. Yang, Y. Si, D. Wang, B. Guo, Automatic recognition of arrhythmia based on principal component analysis network and linear support vector machine, *Computers in Biology and Medicine* (2018), doi: 10.1016/j.compbiomed.2018.08.003.

Automatic recognition of arrhythmia based on principal component analysis network and linear support vector machine

Weiyi Yang[a], Yujuan Si[a,b,*], Di Wang[a], Buhao Guo[a]

[a]*College of Communication Engineering, Jilin University, Changchun, 130012, China*
[b]*Zhuhai College of Jilin University, Zhuhai, 519041, China*

ABSTRACT

Electrocardiogram (ECG) classification is an important process in identifying arrhythmia, and neural network models have been widely used in this field. However, these models are often disrupted by heartbeat noise and are negatively affected by skewed data. To address these problems, a novel heartbeat recognition method is presented. The aim of this study is to apply a principal component analysis network (PCANet) for feature extraction based on a noisy ECG signal. To improve the classification speed, a linear support vector machine (SVM) was applied.In our experiments, we identified five types of imbalanced original and noise-free ECGs in the MIT-BIH arrhythmia database to verify the effectiveness of our algorithm and achieved 97.77% and 97.08% accuracy, respectively. The results show that our method has high recognition accuracy in the classification of skewed and noisy heartbeats, indicating that our method is a practical ECG recognition method with suitable noise robustness and skewed data applicability.

Keywords: principal component analysis network; arrhythmia recognition; noise robustness; deep learning; cardiovascular diseases.

1. Introduction

In recent years, the mortality rate of cardiovascular diseases has continued to increase worldwide [1]. According to the World Health Organization (WHO), cardiovascular diseases, accounting for one-third of global deaths, take more than 10 million lives each year [2]. Furthermore, an American Heart Association (AHA) report stated that 80% of cardiovascular deaths occurred in low-income and middle-income countries with almost equal mortality rates in men and women in 2017 [3]. As a collective term for a heterogeneous group of conditions including abnormal cardiac electrical activity, arrhythmia comprises a set of important cardiovascular diseases. Since 1980, the American College of Cardiology (ACC) has made multiple recommendations to reduce the incidence of arrhythmia [4]. To predict the occurrence of arrhythmia, an electrocardiogram (ECG) is generally used by doctors to identify the condition of the patient. However, the early prevention of such diseases is challenging due to subtle ECG variations [5] and common occasional similarities in ECG changes [6]. Hence, to accurately detect the abnormalities of the heart in advance, computer-assisted detection and classification of arrhythmia are expected to play major roles.

At present, artificial intelligence and machine learning are widely used in this field, with features extracted from ECGs and used for classifier training and classification [7]. In general, the complete process of classification includes three primary steps: (i) pre-processing, (ii) feature extraction, and (iii) classification. Currently, most mainstream classifiers generally identify a heartbeat by means of

training and classifying the P-QRS-T waveform. Several specific methods have achieved high recognition effectiveness. For example, Acharya et al. proposed using the DCT to extract seven features; the k-nearest neighbors (KNN) method was used as the classifier. That group identified normal, coronary artery disease (CAD)-related, and myocardial infarction (MI)-related heartbeats in the Physionet open access databases and obtained an accuracy of 98.5%, a sensitivity of 99.7% and a specificity of 98.5% [8]. Li et al. used the wavelet packet entropy and RR intervals as heartbeat features and then applied the random forest (RF) method for heartbeat classification; over 94% accuracy was achieved in MI-related and normal heartbeats [9]. Prasad et al. extracted features by combining higher-order spectra (HOS) and independent component analysis (ICA) and used the KNN method for different arrhythmia heartbeat types. Their accuracy rate in identifying normal, atrial fibrillation and atrial flutter ECG beats reached 97.65% [10].

Great effects have been achieved in the classification of ECG events; nevertheless, certain drawbacks exist. Since the original ECG signals generally contain noise, a denoising algorithm has to be applied in the preprocessing stage. However, several problems arise after noise removal under normal conditions.In general, the denoising process inevitably removes useful information from the source signals. For example, the frequency components of the baseline wander are typically within the range of 0-0.5 Hz [11], while the ventricular fibrillation ECG signal lies in the frequency range of 0.3-30 Hz [12]. Due to the presence of overlapping frequencies, it is not possible to completely remove baseline wander without eliminating the useful components of the ECG signal. In addition, due to the diverse environments in which the signal is collected in practice, common differences exist in the signal-to-noise ratio (SNR) and noise types. For various SNRs and noise types, we may need different denoising algorithms and multiple iterations to obtain the best parameters in the noise removal process. If only simple noise removal or no denoising is needed, we can keep the relative integrity of the information in ECG signals and also reduce the time overhead. Hence, a feature extraction scheme robust to noise needs to be implemented to avoid excessive dependence on the denoising processing. Relative to traditional models, certain deep learning-based frameworks are less disturbed by the ECG noise during classification. The SVM-classifier-based evolutionary neural network proposed by Paweł Pławiak classified 17 types of noisy heartbeats and reached 98.85% accuracy, 90.8% sensitivity and 99.39% specificity [13]. However, they used only a total of 1,000 samples, and most of the heartbeats were used for training; thus, the results of their experiments are not universal. A deep convolutional neural network (CNN) developed by Acharya et al. identified MI occurrence in the heartbeats and achieved an accuracy of 93.53% without signal noise removal [14]. Subsequently, that group proposed a 7-layer CNN. Accuracies of 93.45% and 94.03% for nondenoising and denoising, respectively, that balanced the 5 types were obtained. However, when the CNN was trained with highly imbalanced data (original data), the accuracy of the CNN was reduced to 89.07% and 89.3% in two ECG sets [15].

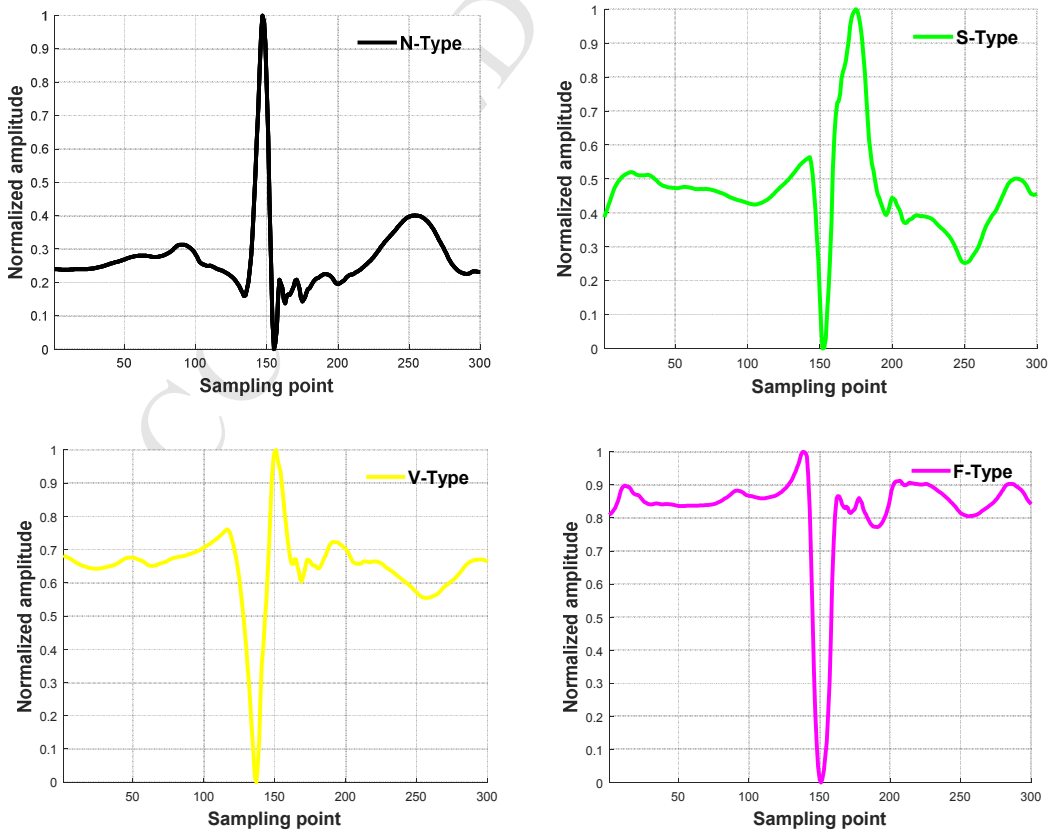
To mine more useful components from noisy heartbeats, we propose to use PCANet as the feature extraction tool. PCANet is a model that uses a convolutional deep neural network structure. So fiducial point detection is no need.Since the PCA filter is used as a convolution kernel, some noise is directly filtered out during the convolution process. Furthermore, the feature dimension containing the classification information is greatly increased. Therefore, sufficient effective features can be obtained, even for a small-scale data. High-dimensional features ordinarily lead to slow classification, and so we use a linear SVM for faster recognition speed in the final classification.

Table 1

Heartbeat type according to ANSI/AAMI EC57; 2012 standard

N	Normal beat (NOR)
	Left bundle branch block beat (LBBB)
	Right bundle branch block beat (RBBB)
	Atrial escape beat (AE)
	Nodal (junctional) escape beat (NE)
S	Atrial premature beat (AP)
	Aberrated atrial premature beat (aAP)
	Nodal (junctional) premature beat (NP)
	Supraventricular premature beat (SP)
V	Premature ventricular contraction (PVC)
	Ventricular escape beat (VE)
F	Fusion of ventricular and normal beat (fVN)
Q	Paced beat (P)
	Fusion of paced and normal beat (PN)
	Unclassifiable beat (U)

According to the Association for the Advancement of Medical Instrumentation (AAMI), non-life-threatening arrhythmia can be divided into 5 main classes [15]: normal or bundle branch block beat (N), supraventricular anomalous beat (S), ventricular abnormal beat (V), fusion beat (F) and unknown beat (Q) [16][17]. ***Table 1*** shows the specific divisions. The waveforms for all types of heartbeats are shown in ***Fig. 1***.



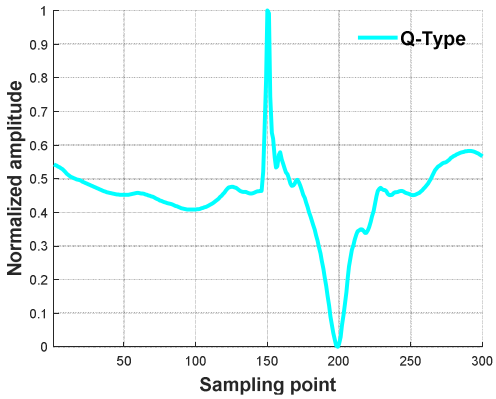


Fig. 1 Example ECG events with different categories of heartbeats

The structure of this paper is introduced in this part. The second section introduces the structure of principal component analysis network (PCANet). The third section provides the source of the data we used. Our specific method is shown in the fourth section. Then, we describe the results of our experiment and discuss them in the fifth section. The sixth section and seventh section present the conclusion obtained through our research and the references we cited, respectively.

2. Principal Component Analysis Network

2.1 Introduction

As a deep learning framework, Principal Component Analysis Network (PCANet) is mainly used to extract high-dimensional feature vectors from samples [18][19]. In terms of structure, this network can be taken as a simplified version of the (CNN), which has an essential hierarchical and cascaded structure [20]. Several algorithms similar to CNN have been developed such as PCANet. However, since the extracted PCA filters is directly used as convolution kernels for PCANet , it does not need to calculate convolution kernels like CNN through a complicated iterative process. At the same time, PCANet has only a few hyperparameters to adjust compared to traditional neural network models.These allow PCANet to have faster processing of data. In addition, PCANet has significant noise robustness and less demand for cleanliness of the data due to the denoising function of the PCA filters. Currently, PCANet has shown great value in deep learning and has achieved excellent effect in visual image recognition.

2.2 Structure

The structure of PCANet adopted in this paper consists of three layers: the input layer, the convolution layer and the output layer. ***Fig. 2*** specifically shows this acquisition process of the PCA filters and the feature extraction process .

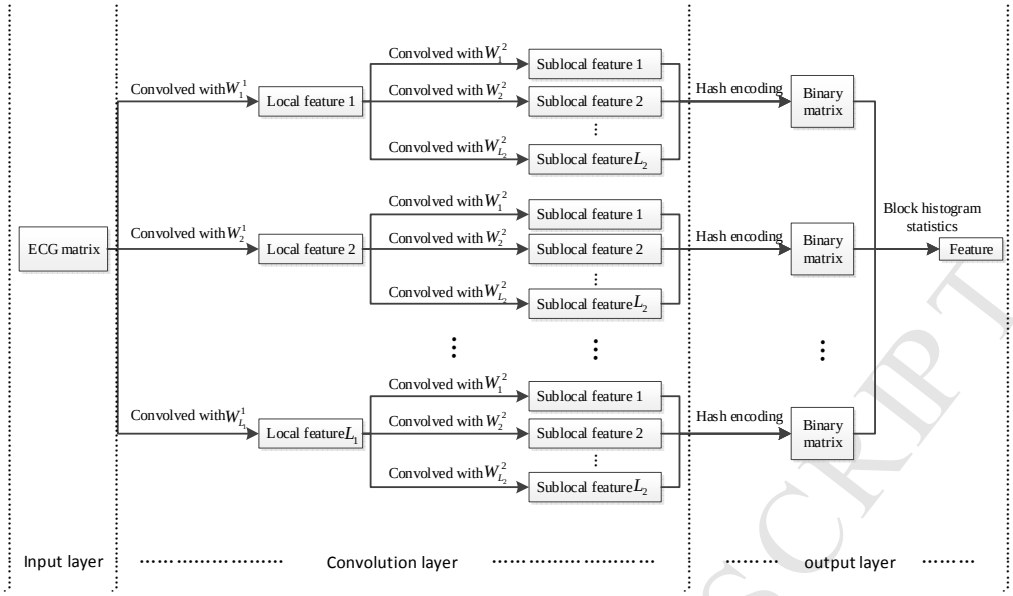


Fig. 2 The structure of PCANet

2.2.1 Input layer

To match our PCANet algorithm, folding these heartbeat vectors into heartbeat matrices suitable for convolution processing is the main function of input layer. Equation (1) shows the specific process.

$$\left[x_1, x_2, \cdots, x_{mn}\right] \rightarrow \begin{bmatrix} x_1 & x_2 & \cdots & \cdots & x_n \\ x_{n+1} & x_{n+2} & \cdots & \cdots & x_{2n} \\ \vdots & \vdots & & & \vdots \\ \vdots & \vdots & & & \vdots \\ x_{(m-1)n} & x_{(m-1)n+1} & \cdots & \cdots & x_{mn} \end{bmatrix} \tag{1}$$

where x_k is the amplitude of the k-th sampling point of the heartbeat and mn is the number of sampling points.

2.2.2 Convolution layer

a) First stage filter extraction

A patch of size $k1 \times k2$ is used to scan with a step of 1 across the i-th heartbeat matrix, and then are collected $m \times n$ patch matrices, i.e., $x_{i,1}, x_{i,2}, \cdots, x_{i,mn} \in \mathbb{R}^{k_1 k_2}$. Here, we remove the mean of each patch matrix and and expand the patch matrices into $m \times n$ column vectors, i.e., $\bar{x}_{i,1}, \bar{x}_{i,2}, \cdots, \bar{x}_{i,mn} \in \mathbb{R}^{k_1 k_2}$, where $\bar{x}_{i,1}$ is the 1st mean-removed column vector for i-th heartbeat. We combine these column vectors to acquire $\bar{X}_i = \left[\bar{x}_{i,1}, \bar{x}_{i,2}, \cdots, \bar{x}_{i,mn}\right]$ and connect each $\bar{X}_i$ to obtain the $\bar{X} = \left[\bar{X}_1, \bar{X}_2, \cdots, \bar{X}_N\right] \in \mathbb{R}^{k_1 k_2 \times 300N}$.

We then obtain the PCA filters based on principal component analysis (PCA). As we know, the purpose of the PCA algorithm is to find a standard orthogonal matrix that can minimize the reconstruction errors by the below equation (2) [21].

$$\min_{V \in \mathbb{R}^{k_1 k_2 \times L_1}} \left\| \bar{X} - VV^T \bar{X} \right\|_F^2, \quad \text{s.t. } V^T V = I_{L_1} \tag{2}$$

Here, L_1, V and I_{L_1} are the number of filters, the standard orthogonal matrix and the unit orthogonal matrix, respectively.

We perform PCA on the $\overline{X}$ to obtain the principal eigenvectors which are ordered based on the decrement of the corresponding eigenvalues. Then we reconstruct the first L_1 principal eigenvectors as the PCA filters.

$$W_l^1 = mat_{k_1,k_2}\left(q_l\left(\overline{X}\overline{X}^T\right)\right) \in \mathbb{R}^{k_1 \times k_2}, l = 1,2,\cdots,L_1 \tag{3}$$

where W_l^1 is the l-th PCA filter for first convolution, $\overline{X}\overline{X}^T$ is the covariance matrix of $\overline{X}$, $q_l(\ )$ extracts the principal eigenvectors of the $\overline{X}\overline{X}^T$, $mat_{k_1,k_2}(\ )$ converts the vectors in parentheses into matrices.

b) First convolution

The convolution layer is the main building block of the PCANet. In this step, matrix sets consisting of several feature matrices are generated by convolving the PCA filters with each heartbeat matrix. Here, Υ_i is used to represent the i-th heartbeat matrix.

$$\Upsilon_i^l = \Upsilon_i * W_l^1, i = 1,2,\cdots,N,\ \ l=1,2,\cdots,\ \ L_1 \tag{4}$$

Where Υ_i^l is the l-th feature matrix of Υ_i.Based on first convolution process of all heartbeats, we obtain N matrix sets containing $N \times L_1$ feature matrices according to ***Fig. 2***.

c) Second stage filter extraction

Here, the operation is similar to the first stage filters extraction.Instead of the heartbeat matrices, the feature matrices are the objects to be processed.We first sample the patch matrices from all the feature matrices in each matrix set. And by removing the mean from each patch matrix and expanding the patch matrices into column vectors, we obtain $\overline{Y}_i^l = \left[\overline{y}_{i,l,1}, \overline{y}_{i,l,2}, \cdots, \overline{y}_{i,l,mn}\right] \in \mathbb{R}^{r_1 r_2 \times mn}$,where $\overline{y}_{i,l,j}$ is the j-th mean-removed column vector of Υ_i^l .Then we combine each $\overline{y}_{i,l,j}$ to acquire $\overline{Y}^l = \left[\overline{Y}_1^l, \overline{Y}_2^l, \cdots, \overline{Y}_N^l\right] \in \mathbb{R}^{k_1 k_2 \times Nmn}$, $l = 1,2,\cdots,L_2$ and $\overline{Y} = \left[\overline{Y}^1,\ \overline{Y}^2, \cdots,\ \overline{Y}^{L1}\right] \in \mathbb{R}^{k_1 k_2 \times L_1 Nmn}$.

At this stage, we perform PCA on the $\overline{Y}$ to obtain L_2 PCA filters.

$$W_\lambda^2 = mat_{k_1,k_2}\left(q_\lambda\left(\overline{Y}\overline{Y}^T\right)\right) \in \mathbb{R}^{k_1 \times k_2}, \lambda = 1,2,\cdots,L_2 \tag{5}$$

Where W_λ^2 is the λ-th PCA filter for second convolution, $\overline{Y}\overline{Y}^T$ is the covariance matrix of $\overline{Y}$, the usages of $q_\lambda(\ )$ and $mat_{k_1,k_2}(\ )$ are the same as them in Equation (3).

d) Second convolution

We convolve the PCA filters with all the feature matrices in all the matrix sets to produce submatrix sets according to the equation (6).

$$O_i^l = \left\{\Upsilon_i^l * W_\lambda^2\right\}_{\lambda=1}^{L_2}, i = 1,2,\cdots,N, \lambda = 1,2,\cdots,L_2 \tag{6}$$

Where O_i^l is obtained by convolving Υ_i^l with all the W_λ^2 and is regarded as a submatrix set which consists of L_2 subfeature matrices .

2.2.3 Output Layer

a) Hash encoding

In this step, each submatrix set is operated separately. Binarization is implemented for all subfeature

matrices in a submatrix set to become binarized matrices. Then, these binary matrices are converted to decimal matrices by means of hash encoding.Finally, these decimal matrices are summed to obtain a new decimal matrix..

$$T_i^l = \sum_{\lambda=1}^{L_2} 2^{\lambda-1} H\left(\Upsilon_i^l * W_\lambda^2\right), \quad l = 1, 2, \cdots, L_1 \tag{7}$$

The function $H(\ )$ converts the sub feature matrices matrices in l-th submatrix set to $0 \sim 2^{\lambda\text{-}1}$, and T_i^l is the l-th decimal matrix for i-th heartbeat.

b) Block histogram statistics

Block histogram statistics are performed through equation (8).

$$f_i = \left[\text{Bhist}\left(T_i^1\right), \text{Bhist}\left(T_i^2\right), \cdots, \text{Bhist}\left(T_i^{L_1}\right)\right]^T \in \square^{\left(2^{L_2}\right)L_1 B} \tag{8}$$

We obtain the B blocks by dividing the T_i^l with a overlap rate. For all B sampling blocks, we calculate the histograms, which are then connected into a vector. The above operations are gathered in $\text{Bhist}(\ )$. Finally, we concatenate the L_1 vectors into a feature vector f_i.

3. Data used

The ECG signals used to verify the validity of the algorithm are from the PhysioBank MIT-BIH arrhythmia database [22], which contains 48 groups of MLII and V5 leads at a 360Hz sampling frequency. Each signal consists of a half-hour length of ECG recording containing more than 20 different types of recorded heartbeats.The labels of all heartbeats are sequentially recorded in the database based on each R-peak. In our experiment, they are reallocated into five types according to ***Table 1***. The number of various types is shown in ***Table 2***.

Table 2

The number of various types of Heartbeat

Type	Number of heartbeats
N	90,411
S	2,778
V	7,227
F	802
Q	5,950
Total	107,168

4. Methodology

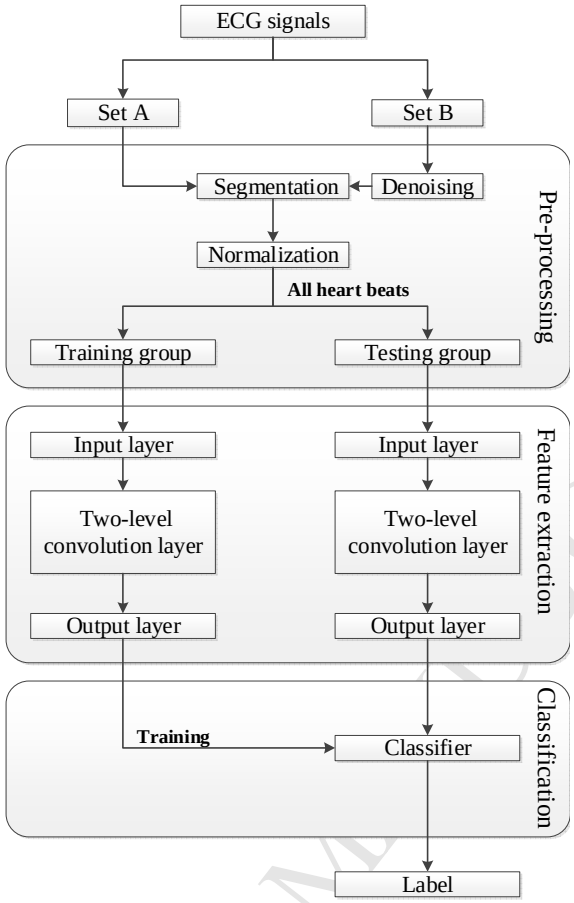


Fig. 3 The proposed system

The process of our work includes preprocessing, feature extraction and classification.The specific process is shown in ***Fig. 3***. To confirm the resistance of our approach to noise, we establish two sets that contains all ECG signals and are marked as set A and set B.The associated experiments are carried out separately.

4.1 Pre-processing

4.1.1 Denoising

In this work, the noise of the ECG signals downloaded from the MIT–BIH arrhythmia database was removed by a denoising algorithm because of the presence the presence of baseline wander and high-frequency noise, which are caused by several factors such as the DC bias drift of the ECG amplifier and the change in the resistance of the electrodes during the recording of the ECG signals [8]. In general, these factors cause a certain degree of distortion in ECG wave forms, which makes heartbeat recognition challenging. The frequency of baseline drift is typically less than 0.5 Hz [11], and most of the information contained in an ECG signals is at a frequency of less than [23]. Here, to retain signal components that facilitate classification, Daubechies wavelet 8 filters were used to remove these interference components from the ECG signals in set B [24]. Set A was left unprocessed.

4.1.2 Heartbeats segmentation

At this stage, we split 48 ECG signals into heartbeats. The Pan-Tompkins algorithm was adopted for

the R-peak point detection of ECG signals [25]. Based on the R-peak of each heartbeat, we eliminated a certain number of points according to the 360 Hz sampling frequency of each original signal. Then, we resampled all the heartbeats to contain 300 points. In this work, 107,168 heartbeat samples were acquired in each set.

4.1.3 Normalization

We used min-max normalization as a heartbeat normalized method [26]. The calculation process is as follows.

$$x_i^* = \frac{x_i - \min}{\max - \min} \tag{9}$$

Where x_i is the value of a point in a heartbeat, $\min$ and $\max$ respectively represent the highest and lowest points in a waveform, respectively; and x_i^* is the normalized value of the point.

Fig. 4 shows the preprocessing results for the same heartbeat in set A and set B.

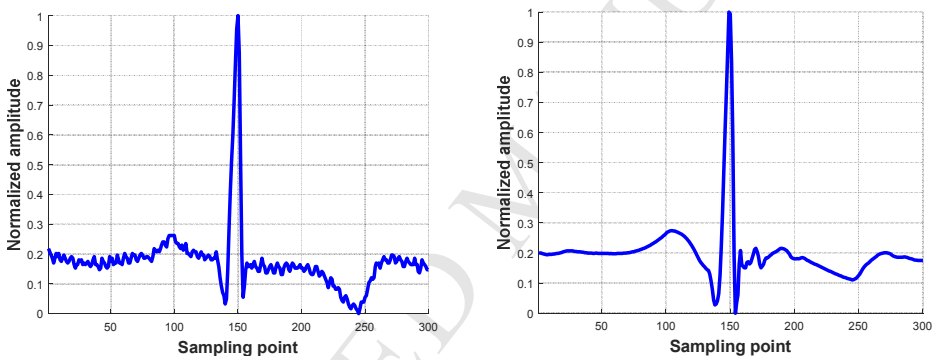


Fig. 4ECG with noise and without noise

4.2 Feature extraction

In this section, the features of heartbeats are extracted by the PCANet algorithm which includes the two-level convolution layer. The associated parameters, which were obtained through multiple experiments, are shown in ***Table 3***.

Table 3

Parameter settings

Layers	Project	Parameter
Input	Heartbeat matrix size	15×20
Convolution	Patch size	7×7
	The number of filters for first convolution	9
	The number of filters for second convolution	9
Output	Histogram block size	7×7
	Block overlap ratio	0.5

Fig. 5 shows a heartbeat and its extracted features.

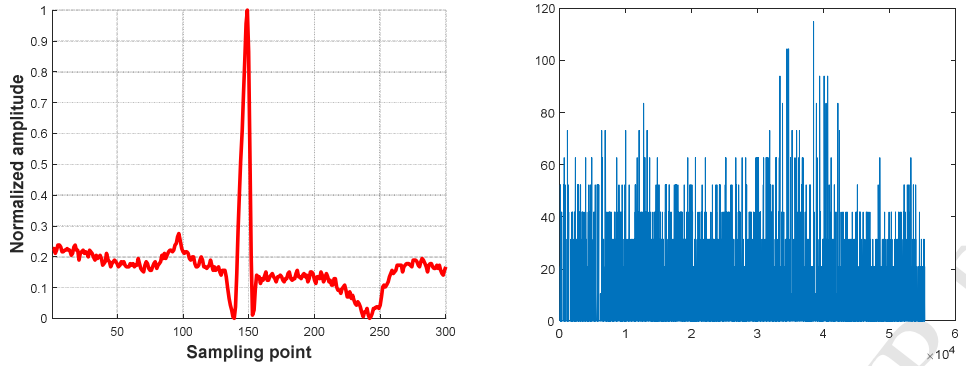


Fig. 5Example ECG segment and its feature vectors extracted by PCANet

4.3 Classification

In this section, four different classifiers were used to classify the heartbeats based on the PCANet feature from the two groups: the linear support vector machine (SVM), KNN, the back propagation neural network (BP-NN) and the RF method. The following is a brief explanation of these classifiers.

4.3.1 Linear SVM

The linear SVM classifier is implemented specifically for massive levels of data and features; this classifier is supported by liblinear [27]. The decision hyperplane that is calculated is used to classify samples into different categories. The selection of the error penalty factor, which expresses the tolerance to error, significantly affects the precision of the linear SVM. In our experiment, we used the SVM with the linear kernel function, and the parameter C was set to 1.

4.3.2 K-nearest neighbors

The KNN classification, also known as a nonparametric lazy algorithm, is based on the nearest training samples in the feature space. The k-nearest neighbors are collected by calculating the Euclidean distance between the training sample and the testing sample. Then, the category that is the most ordinary neighbor for the new sample is ultimately acquired. Here, the best effect of classification was achieved for K = 5.

4.3.3 Back propagation Neural Network (BP-NN)

As one of the most widely used neural network models at present, the BP-NN is a kind of multilayer feed forward network trained according to error back propagation. This approach was proposed by a team of scientists led by Rumelhart et al. [28]. The algorithm uses the steepest descent method and continuously adjusts the weights and thresholds of the network using back propagation to minimize the squared error of the network. We used a BP-NN with two hidden layers.

4.3.4 Random Forest (RF)

The RF classifier, which was first proposed by Leo Breiman, uses multiple trees to train and predict samples [29]. Here, we used cart classification trees, which selected features based on the Gini index, to form an RF. The number of trees is an important parameter of the forest, and the best classification results occurred when the quantity was 500.

4.4 K-fold cross-validation

To obtain a reliable and stable model, we used 10-fold cross-validation in our experiment. Specifically, the experimental data was divided into 10 equal parts by random extraction. In each experiment, we selected one of the parts for training and the remaining data for testing. For each classifier, we repeated this method ten times by rotating the training data and averaged the overall results of all experiments to obtain our final result.

4.5 Evaluation index

The evaluation indicators used in our study is based on the confusion matrix.

$$ACC = \frac{TP + TN}{TP + FP + TN + FN} \times 100\% \tag{10}$$

$$PPV = \frac{TP}{TP + FP} \times 100\% \tag{11}$$

$$SN = \frac{TP}{TP + FN} \times 100\% \tag{12}$$

$$SPC = \frac{TN}{TN + FP} \times 100\% \tag{13}$$

where ACC, PPV, SN, and SPC are the accuracy, precision, sensitivity and specificity, respectively. To acquire these parameters, the true positive (TP), true negative (TN), false positive (FP) and false negative (FN) are calculated by gathering the statistics of the experimental results.

$$Average\ Accuracy = \frac{\sum_{i=1}^{5} TP_i}{TP_1 + FP_1 + TN_1 + FN_1} \times 100\% \tag{14}$$

Where TP_i, TN_i , FP_i , FN_i are the TP, TN, FP and FN of i-th class of heartbeat, respectively.

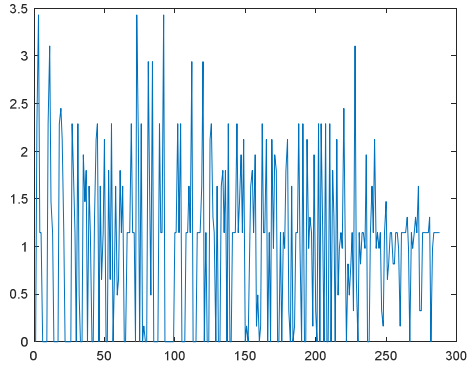
5. Result and Discussion

We implemented the experimental procedures with four 2.60 GHz Intel® Core ™ i5-3230 CPUs and 4 GB of RAM. Windows 10 and MATLAB 2017 (A) were used as the operating system and the platform, respectively. The average time spent on the training processes of set A and set B was 1532.22 seconds and 1480.47 seconds, respectively. The confusion matrices (***Table 4*** and ***Table 5***) summarize the specific data of the experimental results.

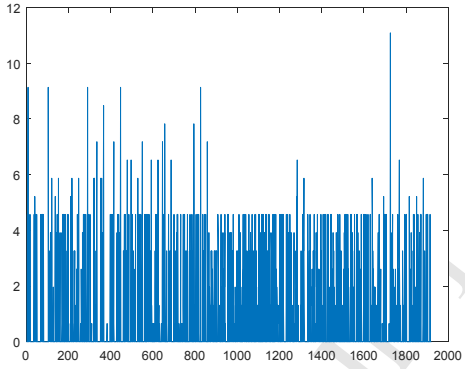
5.1 Selection of the number of convolution kernels

In our experiment, the selection of the amount of filtering for the two-level convolution of PCANet can have a significant impact on the performance.The features, which were extracted by the PCANet algorithm with different numbers of convolution kernels, are shown in ***Fig. 6***. The number of convolution kernels has significant effects on the length and amplitude of the feature vector, which are

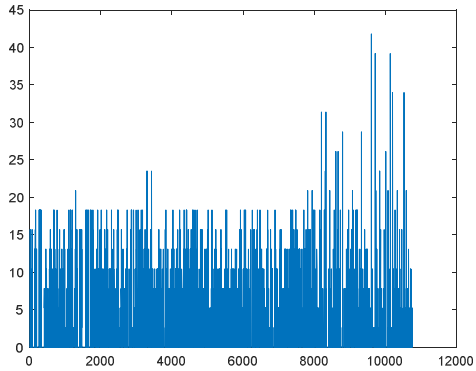
positively related.



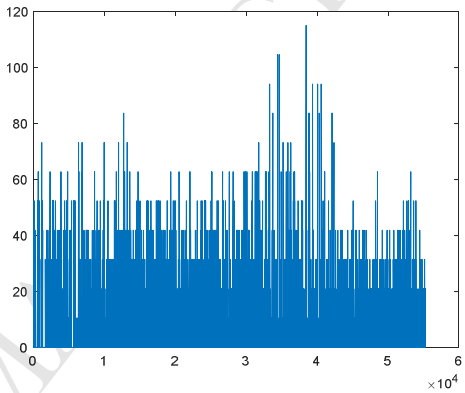
Three kernels for each convolution



Five kernels for each convolution



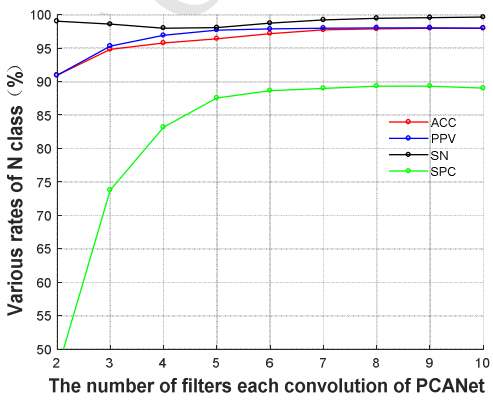
Seven kernels for each convolution

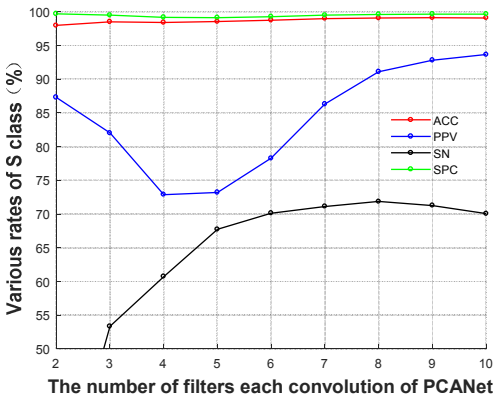


Nine kernels for each convolution

Fig. 6 Feature vectors extracted by PCANet with different numbers of convolution kernels

Then, we used the linear SVM classifier to classify the heartbeats multiple times in set A and recorded the line charts of all the effects. ***Fig. 7*** shows the degree of correlation between the number of kernels and the performance. For the S, V, F and Q types of heartbeats, the differences in the kernel quantity strongly influenced the precision and sensitivity of their classification. However, the accuracy, precision and sensitivity were less affected by these differences for the N and Q types. As the kernels in each convolution level increased from 2 to 9, almost all classification evaluation indicators tended to increase. When the kernel quantity exceeded 9, some of the curves began to decline, and others became stable, indicating that only within a certain range is there a positive correlation between the recognition effect and the complexity of the PCANet features. Hence, the classification effect of our method with the abscissa of 9 is ideal.





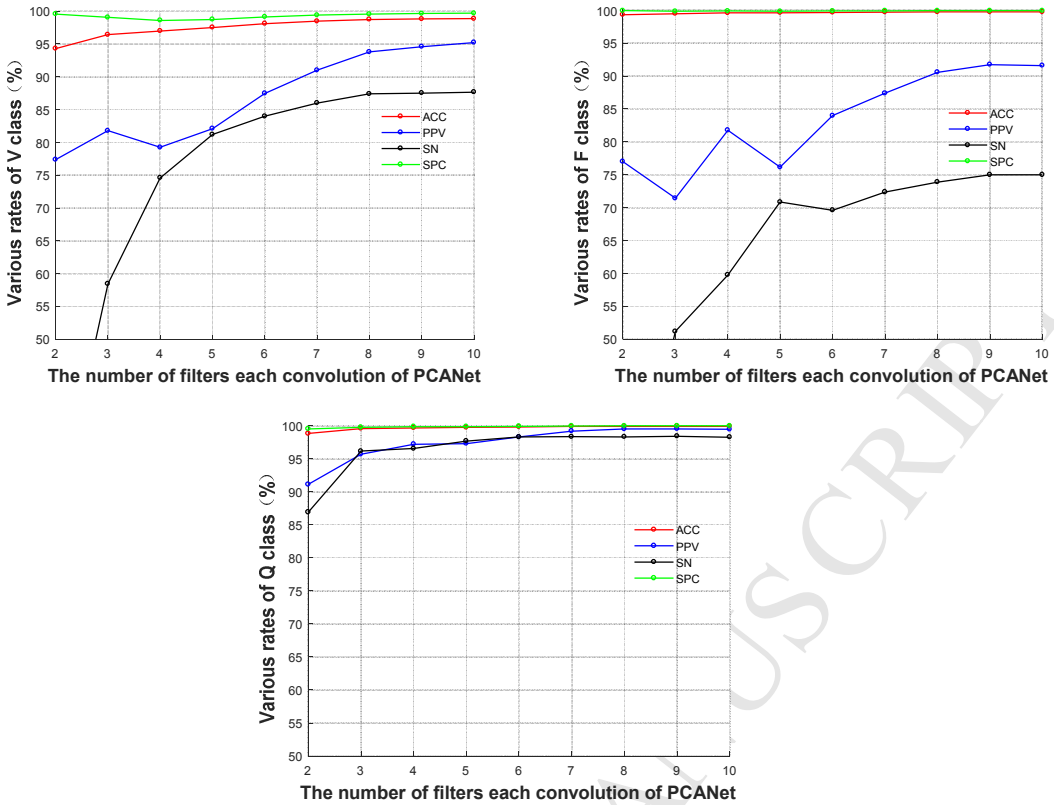


Fig. 7 Line chart for classification of different categories of heartbeats

5.2 Reliability

All of the heartbeats used in our study are from the MIT-BIH arrhythmia database.The ANSI/AAMI EC57 standard that involves all types of heartbeats in the database was used, and the incomplete heartbeats of the waveform at both ends of each ECG recording were removed. Then, our experiment was conducted on the two sets separately. To suppress the influence of chance on our results, we used 10-fold cross-validation. The records in ***Table 4*** and

Table 5 show the overall results across 10-fold sampling (the standard deviations of their accuracies are about 0.0005 and 0.0022, respectively). Regardless of the amount of data or the method of validation, our experiment is relatively reliable and can be reproduced.

Table 4

Confusion matrix for the classification result of set A across all 10-folds

Linear SVM		Predicted N	S	V	F	Q	Acc (%)	Ppv (%)	St (%)	Sf (%)
Original	N	810166	1034	2304	107	88	97.97	98.05	99.56	89.33
	S	6832	17819	308	30	13	99.11	92.80	71.27	99.85
	V	7261	330	56934	354	164	98.82	94.59	87.53	99.64
	F	1226	19	560	5413	0	99.76	91.75	74.99	99.95
	Q	756	10	85	0	52689	99.88	99.50	98.39	99.97
Average Accuracy (%)							97.77			

Table 5

Confusion matrix for the classification result of set B across all 10-folds

Linear SVM		Predicted					Acc (%)	Ppv (%)	St (%)	Sf (%)
		N	S	V	F	Q				
Original	N	807858	1552	3737	180	372	97.31	97.57	99.28	86.66
	S	7842	16660	461	10	27	98.94	89.76	66.63	99.80
	V	8703	320	55446	403	173	98.49	91.77	85.21	99.45
	F	1907	17	614	4671	9	99.69	89.79	64.71	99.94
	Q	1672	15	161	0	51702	99.75	98.89	96.59	99.94
Average Accuracy (%)							97.08			

5.3 High accuracy and precision (Important advantage)

Table 4 and
Table 5 show that the overall accuracy reaches 97.77% and 97.08% in heartbeat classification. In particular, the corresponding minimal accuracy recorded for both sets is attributed to the detection of the N type and is 97.97% and 97.31%. The advantage of our experimental accuracy is remarkable. In terms of precision, we achieved a prediction success rate exceeding 91% for each type and exceeding 94% for the N, V and Q types by experimenting on set A. Relative to the results of the previous methods in ***Table 6***, our accuracy and precision are at a high level.

5.4 Applicability to skewed data (Important advantage)

Our experiments were conducted on five beat types that were extremely imbalanced in quantity. The materials contained over 90,000 N beats. However, only 2778 S and 802 V beats were recorded in these ECG signals. Nonetheless, our recognition effects are still acceptable according to ***Table 4*** and
Table 5. This finding suggests that our algorithm can extract high-dimensional features that are sufficiently favorable for classification from small-scale data. In the case of classifying imbalanced data, the 11-layer deep convolutional neural network proposed by Acharya et al. exhibited only approximately 17.80% and 31.25% precision in classifying the F and S beats, respectively [15]. We can conclude that our method offers beneficial classification effects on skewed data.

5.5 Robustness to noise (Important advantage)

Fig. 8 shows the consequences of the different classifiers performed on set A and set B; the evaluation indicators for the two sets are all prominent. Several classifiers were used to classify the heartbeats in set A, and all achieved an accuracy exceeding 96%. Although each classifier has its own characteristics, the recognition accuracy of set A is higher than that of set B by 0.7-1.1 percentage points. Among other indicators in ***Table 4*** and
Table 5, the classification for set A also performed better. Therefore, the proposed model can detect noisy ECG heartbeats reasonably well [15] because the PCA filters are used in the convolution layer and the PCA filter itself can remove unwanted noise from the ECG matrix.

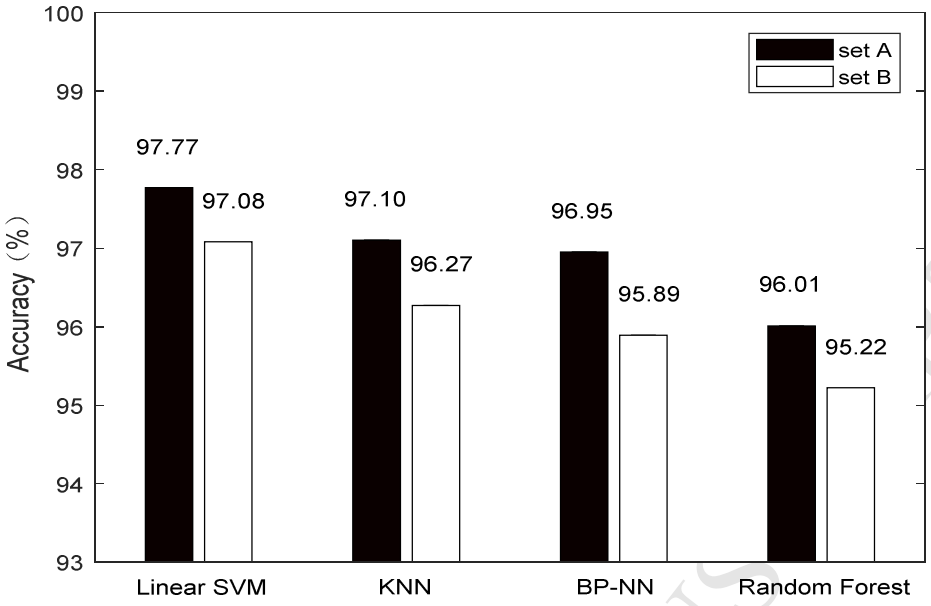


Fig. 8 Comparison of effects of different classifiers

5.6 Comparison of linear SVM, KNN, BP-NN and Random Forest

In this paper, four classifiers were involved in heartbeat classification. PCANet can extract high-dimensional features from a sample. Although the original heartbeat dimensions were increased by two orders of magnitude, the linear SVM was used for the classification. Relative to the KNN, BP-NN and RF methods, the linear SVM has faster training and classification speed because the linear SVM is a linear classifier well suited for high-dimensional features. As shown in ***Fig. 8***, the highest classification accuracy is 97.77%, whereas the recognition accuracy of the KNN, BP-NN and RF are only 97.10%, 96.95% and 96.01%, respectively. Hence, the linear SVM has prominent advantages over the other three classifiers in our experiment.

Table 6

Two-class			
Author	Method	Experience	Best performance
Inan et al. [30]	WT + timing interval NN classifier	Classification with a denoising process	Accuracy:96.82%
Giri et al. [31]	WT + ICA GMM classifier	Classification with a denoising process	Accuracy:96.80%
Sharma et at. [32]	Wavelet transform of multilead ECG Multiscale energy Multiscale eigenspace analysis RBF + KNN classifier	Classification with a denoising process	Accuracy:96.00% Sensitivity:93.00% Specificity:99.00%
Polat et al. [33]	PCA LS-SVM classifier	Classification without a denoising process	Accuracy:100%

Four-class			
Author	Method	Experience	Best performance
Oster et al. [34]	X-factor Switching Kalman filters	Classification without a denoising process	F1-score: 98.3%
Five-class			
Author	Method	Experience	Best performance
Ince et al. [35]	WT+PCA M-PSO	Classification with a denoising process	Accuracy:95.58%
Martis et al. [36]	HOS + PCA □ LS-SVM classifier	Classification with a denoising process	Accuracy:93.48%
Martis et al. [37]	PCA □ LPC + PCA □ DWT + PCA	Classification with a denoising process	Accuracy:98.11%
Zhang et al. [38]	RR-intervals + morphological features + ECG-intervals and segments SVM classifier	Classification with a denoising process	Accuracy:86.66%
Knvps et al. [39]	EEMD-IMF SMO-SVM classifier	Classification with a denoising process	Accuracy:99.20%
Our method	PCANet Linear SVM	Classification without a denoising process	Accuracy:97.94%
Seven-class			
Author	Method	Experience	Best performance
Homaeinezhad et al. [40]	Virtual QRS image-based geometrical Neuro-SVM-KNN hybrid classifier	Classification without a denoising process	Accuracy:98.06%
Osowski et al. [41]	HOSA Hybrid fuzzy NN classifier	Classification with a denoising process	Accuracy:96.06%

6. Conclusion

The features are one of the most crucial factors in ECG heartbeat classification systems. Here, a deep learning framework was proposed to automatically identify arrhythmia, which is significant for the diagnosis of cardiovascular disease. In this study, we classified 107,164 heartbeats from the MIT-BIH database by using the convolutional deep learning network PCANet and the linear SVM. To evaluate the effectiveness of our algorithm, the experiment was performed on two skewed data sets containing noisy and noise-free ECGs. Ultimately, our classification of the noisy heartbeats achieved an average recognition accuracy of 97.77%, which was better than that of most existing ECG recognition algorithms. Nevertheless, only 97.08% accuracy was obtained on the other set. Therefore, our experimental results show that the proposed method achieves high recognition accuracy in heartbeat recognition. Moreover, the method offers notable robustness to noise, and skewed data has less of a

negative impact on the classification effect.

Regarding future work, we plan to take measures to improve our methods. The recognition accuracy can be potentially improved by using another solution, such as a dimension reduction method, data augmentation that uses the synthetic minority over-sampling technique (SMOT) or other techniques. This approach is expected to contribute to preventing patients from having arrhythmia and saving a significant number of lives.

Acknowledge

This work was supported by the Key Scientific and Technological Research Project of Jilin Province under Grant Nos. 20150204039GX and 20170414017GH; the Natural Science Foundation of Guangdong Province under Grant No. 2016A030313658; the Innovation and Strengthening School Project (provincial key platform and major scientific research project) supported by Guangdong Government under Grant No. 2015KTSCX175; the Premier-Discipline Enhancement Scheme Supported by Zhuhai Government under Grant No. 2015YXXK02-2; the Premier Key-Discipline Enhancement Scheme Supported by Guangdong Government Funds under Grant No. 2016GDYSZDXK036.

References

[1] Anand, S. S., & Yusuf, S. (2011). Stemming the global tsunami of cardiovascular disease. Lancet, 377(9765), (pp. 529-32).

[2] World Health Organization (2012), Cardiovascular diseases (CVDs). Fact sheet No317, 2012, (pp.4–7).

[3] Benjamin EJ, M.J. Blaha, S.E. Chiuve, M. Cushman, S.R. Das, R. Deo, S.D. de Ferranti, et al. (2013). American Heart Association Statistics Committee and Stroke Statistics Subcommittee Heart disease and stroke statistics-2017 update: a report from the American Heart Association, Circulation, 135 (10), (pp.e146-603).

[4] Nishimura, R. A., Otto, C. M., Bonow, R. O., Carabello, B. A., Erwin, J. P., & Fleisher, L. A., et al. (2017). 2017 aha/acc focused update of the 2014 aha/acc guideline for the management of patients with valvular heart disease. Journal of the American College of Cardiology, 70(2), (pp.252-289).

[5] Acharya, U. R., Joseph, K. P., Kannathal, N., Lim, C. M., & Suri, J. S. (2006). Heart rate variability:a review. Medical & Biological Engineering & Computing, 44(12), (pp.1031-1051).

[6] Altan, G., & Kutlu, Y. (2015). ECG based Human identification using Logspace Grid Analysis of Second Order Difference Plot. IEEE Signal Processing and Communications Applications Conference (pp.1288-1291).

[7] Liu, T., Si, Y., Wen, D., Zang, M., & Lang, L. (2016). Dictionary learning for vq feature extraction in ecg beats classification. Expert Systems with Applications, 53, 129-137.

[8] Acharya, U. R., Fujita, H., Adam, M., Lih, O. S., Sudarshan, V. K., & Tan, J. H., et al. (2016). Automated characterization and classification of coronary artery disease and myocardial infarction by decomposition of ecg signals: a comparative study. Information Sciences, 377, (pp.17-29).

[9] Li, T., & Zhou, M. (2016). Ecg classification using wavelet packet entropy and random forests. Entropy, 18(8), (pp.1-16).

[10] Prasad, H., Martis, R. J., Acharya, U. R., Min, L. C., & Suri, J. S. (2013). Application of higher order spectra for accurate delineation of atrial arrhythmia. Engineering in Medicine and Biology Society (pp.57-57). IEEE.

[11] Jane, R., Laguna, P., Thakor, N. V., & Caminal, P. (1992). Adaptive baseline wander removal in the ecg: comparative analysis with cubic spline technique. Computers in Cardiology Proceedings, 143 - 146.

[12] Strohmenger, H. U., Lindner, K. H., & Brown, C. G. (1997). Analysis of the ventricular fibrillation ecg signal amplitude and frequency parameters as predictors of countershock success in humans. Chest, 111(3), (pp.584-9).
[13] Pławiak, P. (2017). Novel methodology of cardiac health recognition based on ecg signalsand evolutionary-neural system. Expert Systems with Applications, 92C(2018), (pp.334-349).
[14] Acharya, U. R., Fujita, H., Lih, O. S., Hagiwara, Y., Tan, J. H., & Adam, M. (2017). Application of deep convolutional neural network for automated detection of myocardial infarction using ecg signals. Information Sciences, 415, (pp.190-198).
[15] Acharya, U. R., Oh, S. L., Hagiwara, Y., Tan, J. H., Adam, M., & Gertych, A., et al. (2017). A deep convolutional neural network model to classify heartbeats. Computers in Biology & Medicine, 89, (pp.389-396).
[16] ANSI/AAMI EC57: Testing and Reporting Performance Results of Cardiac Rhythm and ST Segment Measurement Algorithms (AAMI Recommended Practice/American National Standard). Available: http://www.aami.org, Order Code: EC57-293, 1998.
[17] Jiang, W., & Kong, S. G. (2007). Block-based neural networks for personalized ecg signal classification. IEEE Transactions on Neural Networks, 18(6), (pp.1750-1750).
[18] Shi, J., Wu, J., Li, Y., Zhang, Q., & Ying, S. (2017). Histopathological image classification with color pattern random binary hashing based pcanet and matrix-form classifier. IEEE J Biomed Health Inform, PP(99), (pp.1-1).
[19] Luz, E. J. D. S., Schwartz, W. R., Cámara-Chávez, G., & Menotti, D. (2016). Ecg-based heartbeat classification for arrhythmia detection: a survey. Computer Methods & Programs in Biomedicine, 127(C), (pp.144-164).
[20] Chan, T. H., Jia, K., Gao, S., Lu, J., Zeng, Z., & Ma, Y. (2015). Pcanet: a simple deep learning baseline for image classification?. IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society, 24(12), 5017.
[21] Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: a review and recent developments. Philos Trans A Math Phys Eng Sci, 374(2065).
[22] Mark, R., Moody, G., (1997). May.MIT-BIH Arrhythmia Database [Online]. Available:http://ecg.mit.edu/dbinfo.html.
[23] Poornachandra, S., & Kumaravel, N. (2008). A novel method for the elimination of power line frequency in ecg signal using hyper shrinkage function. Digital Signal Processing, 18(2), (pp.116-126).
[24] Singh, B. N., & Tiwari, A. K. (2006). Optimal selection of wavelet basis function applied to ecg signal denoising. Digital Signal Processing, 16(3), (pp.275-287).
[25] Pan, J., & Tompkins, W. J. (1985). A real-time qrs detection algorithm. IEEE Trans Biomed Eng., 32(3), (pp.230-236).
[26] Pollesch, N. L., & Dale, V. H. (2016). Normalization in sustainability assessment: methods and implications. Ecological Economics, 130, (pp.195-208).
[27] Fan, R. E., Chang, K. W., Hsieh, C. J., Wang, X. R., & Lin, C. J. (2008). Liblinear: a library for large linear classification. Journal of Machine Learning Research, 9(9), (pp.1871-1874).
[28] Rumelhart, David, E., Hinton, Geoffrey, E., Williams, & Ronald, J. (1986). Learning representations by back-propagating errors. Readings in Cognitive Science, 323(6088), (pp.399-421).
[29] Breiman, L. (2001). Random forests. Machine Learning, 45(1), (pp.5-32).
[30] Inan, O. T., Giovangrandi, L., & Kovacs, G. T. (2006). Robust neural-network-based classification of premature ventricular contractions using wavelet transform and timing interval features. IEEE Transactions on Biomedical Engineering, 53(12), (pp.2507-2515).
[31] Giri, D., Acharya, U. R., Sree, S. V., Sree, S. V., Lim, T. C., & Suri, J. S. (2013). Automated diagnosis of coronary artery disease affected patients using lda, pca, ica and discrete wavelet transform. Knowledge-Based Systems, 37(2), (pp.274-282).
[32] Sharma, L. N., Tripathy, R. K., & Dandapat, S. (2015). Multiscale energy and eigenspace approach to detection and localization of myocardial infarction. IEEE Trans Biomed Eng, 62(7), (pp.1827-1837).
[33] Polat, & Gunes. (2007). Detection of ecg arrhythmia using a differential expert system approach based on principal component analysis and least square support vector machine. Applied Mathematics & Computation,186(1), (pp.898-906).
[34] Oster, J., Behar, J., Sayadi, O., Nemati, S., Johnson, A., & Clifford, G. (2015). Semi-supervised ecg ventricular beat classification with novelty detection based on switching kalman filters. IEEE Trans Biomed Eng,62(9), (pp.2125-2134).

[35] Ince, T., Kiranyaz, S., & Gabbouj, M. (2009). A generic and robust system for automated patient-specific classification of ecg signals. IEEE Trans Biomed Eng, 56(5), (pp.1415-1426).
[36] Martis, R. J., Acharya, U. R., Mandana, K. M., Ray, A. K., & Chakraborty, C. (2013). Cardiac decision making using higher order spectra. Biomedical Signal Processing &Control, 8(2), (pp.193-203).
[37] Martis, R. J., Acharya, U. R., Mandana, K. M., Ray, A. K., & Chakraborty, C. (2012).Application of principal component analysis to ecg signals for automated diagnosis ofcardiac health. Expert Systems with Applications,39(14), (pp.11792-11800).
[38] Zhang, Z., Dong, J., Luo, X., Choi, K. S., & Wu, X. (2014).Heartbeat classification using disease-specific feature selection. Computers in Biology & Medicine, 46(1), (pp.79-89).
[39] Knvps, R., & Dhuli, R. (2017). Classification of ecg heartbeats using nonlinear decomposition methods and support vector machine. Computers in Biology & Medicine, (pp.87, 271).
[40] Homaeinezhad, M. R., Atyabi, S. A., Tavakkoli, E., Toosi, H. N., Ghaffari, A., & Ebrahimpour, R. (2012). Ecg arrhythmia recognition via a neuro-svm-knn hybrid classifier with virtual qrs image-based geometrical features. Expert Systems with Applications, 39(2), (pp.2047-2058).
[41] Osowski, S., & Linh, T. H. (2001). Ecg beat recognition using fuzzy hybrid neural network. IEEE Transactions on Biomedical Engineering, 48(11), (pp.1265-1271).

Vitae

Weiyi Yang was born in Liaoning, China. He received his bachelor's degree in 2017 from Jilin University. He is currently working toward his master's degree in signal processing at Jilin University. His research interests include biomedical signal processing and recognition.

Yujuan Si is a professor in the Institute of Communication Engineering at Jilin University. She received her master's degree and PhD in engineering in 1988 and 1996, respectively, from the Jilin University of Technology. Her research interests include embedded systems and biomedical signal processing and recognition.

Di Wang was born in Henan, China. He is a doctoral candidate in the Institute of Communication Engineering at Jilin University. He received his master's degree in signal and information processing in 2016 from Jilin University in Changchun, China. His research interests include biomedical signal processing and recognition.

Buhao Guo was born in Shanxi, China. He received his bachelor's degree in 2017 from the Shandong University of Technology. He is currently working toward his master's degree in signal processing, his primary interest, at Jilin University.

1. Our method is robust to noise. Compared with the noiseless heart beat, experimenting with the noisy heartbeat achieves better results.
2. Our results are better than most other methods.
3. Compared to traditional neural networks, our method has only a few hyperparameters and do not require iteration.
4. Fiducial point detection is no need.
5. Our approach is applicable for skewed data.