

Two-Dimensional ECG-based Cardiac Arrhythmia Classification using DSE-ResNet

Jiahao Li

Qilu University of Technology (Shandong Academy of Science)

Shao-peng Pang (✉ pang_shao_peng@163.com)

Qilu University of Technology (Shandong Academy of Science)

Fangzhou Xu

Qilu University of Technology, Shandong Academy of Science)

Shuwang Zhou

Qilu University of Technology (Shandong Academy of Sciences), Shandong Artificial Intelligence Institute

Minglei Shu

Qilu University of Technology (Shandong Academy of Sciences), Shandong Artificial Intelligence Institute

Article

Keywords:

Posted Date: April 22nd, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1550001/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Two-Dimensional ECG-based Cardiac Arrhythmia Classification using DSE-ResNet

Jiahao Li¹, Shao-peng Pang^{1,*}, Fangzhou Xu^{2,+}, Shuwang Zhou^{3,+}, and Minglei Shu^{3,**}

¹School of Electrical Engineering and Automation, Qilu University of Technology (Shandong Academy of Science), Jinan, Shandong Province, 250353, China

²School of Electronic and Information Engineering (Department of Physics), Qilu University of Technology (Shandong Academy of Science), Jinan, Shandong Province 250353, China

³Qilu University of Technology (Shandong Academy of Sciences), Shandong Artificial Intelligence Institute, Jinan 250014, China

*First Corresponding author email: pang_shao_peng@163.com

**Second Corresponding author email: shuml@sdas.org

+these authors contributed equally to this work

ABSTRACT

Electrocardiogram (ECG) is mostly used for clinical diagnosis of cardiac arrhythmia due to its simplicity, non-invasiveness, and reliability. Recently, many models based on the Deep Neural Networks (DNNs) have been applied to the automatic classification of cardiac arrhythmia with great success. However, most models independently extract the internal features of each lead in the 12-lead ECG during the training phase, resulting in a lack of inter-lead features. Here, we propose a general model based on the two-dimensional ECG and ResNet with detached squeeze-and-excitation modules (DSE-ResNet) to realize the automatic classification of 9 cardiac arrhythmias. The original 12-lead ECG is spliced into a two-dimensional plane like a grayscale picture. DSE-ResNet is used to simultaneously extract the internal and inter-lead features of the two-dimensional ECG. Furthermore, an orthogonal experiment method is used to optimize the hyper-parameters of DSE-ResNet and a multi-model voting strategy is used to improve classification performance. Experimental results based on the test set of China Physiological Signal Challenge 2018 (CPSC2018) show that our model has average $F_1 = 0.817$ for classifying normal category and 8 cardiac arrhythmias. Meanwhile, compared with the state of the art model in CPSC2018, our model achieved the best F_1 in 2 sub-abnormal types. This shows that the model based on the two-dimensional ECG and DSE-ResNet has advantage in detecting some cardiac arrhythmias and have the potential to be used as an auxiliary tool to help doctors perform cardiac arrhythmias analysis.

Introduction

The ECG¹ records the electrical signals of the human heart, and is mostly used for clinical diagnosis of cardiac arrhythmias. More than 300 million ECGs are obtained worldwide every year². The huge diagnostic workload leads to inefficiency and misdiagnosis of cardiac arrhythmias based on ECG. So the combination of extensive digitization of ECG data and automatic classification algorithms has attracted more and more attention.

In the early research on the automatic classification of cardiac arrhythmia, most algorithms based on machine learning are usually divided into two parts: feature engineering and classification. Specifically, researchers first manually extracted a large number of ECG features with medical meaning, such as wavelet features³, P-QRS-T composite features⁴⁻⁶, heart rate variability statistical feature⁷, RR-related statistical features^{8,9}, higher order statistical features¹⁰ and morphological features¹¹⁻¹⁴. Meanwhile, the principal component analysis^{15,16} and independent component analysis^{17,18} use mathematical methods to extract ECG features from high-dimensional space to low-dimensional space. After feature engineering, support vector machine¹⁹⁻²¹, self-organizing map²², clustering²³ and other machine learning algorithms are used to analyze artificial features, and give the prediction result. Although machine learning have achieved good performance in the classification of cardiac arrhythmia, there are still some problems that need to be solved. For example, feature engineering based on subjective factors leads to the elimination of some potentially important features, which may affect the final classification performance.

In recent years, DNNs have greatly improved the technical level of speech recognition, image classification, strategy games and medical diagnosis with their powerful feature extraction capabilities. Different from machine learning methods, DNNs can recognize patterns and learn useful features from raw input data without requiring a lot of manual rules and feature engineering, making them particularly suitable for interpreting ECG data. Some studies have been inspired to use DNNs for the automatic classification of cardiac arrhythmia based on single-lead or multi-lead ECG. For example, Ullah et al.²⁴ converted single-lead ECG into two dimension spectral images as input for learning. Hannun et al.²⁵ develop a DNN to classify 12 rhythm classes

based on single-lead ECG. Compared with single-lead, multi-lead ECG contains more valuable information^{2,26}, which is more conducive to the automatic classification of cardiac arrhythmia. Wang et al.²⁷ proposed a method based on multi-scale feature extraction and 12-lead ECG cross-scale information complementation to capture the abnormal state in ECG. Chen et al.²⁸ proposed a neural network that combines convolutional neural networks (CNNs), recurrent neural networks and attention mechanisms for cardiac arrhythmias classification. Their model won the state of the art in CPSC2018²⁹. Zhang et al.² proposed a interpretable DNN for automatic diagnosis based on 12-lead ECG. Ribeiro et al.³⁰ presented a DNN model trained in a dataset with more than 2 million labeled exams, and found that the model outperformed cardiology resident medical doctors in recognizing 6 types of cardiac arrhythmias. In addition, Zhao et al.³¹ showed that adding clinical data such as gender and age as auxiliary features can enhance classification performance.

In this paper, we propose a general model based on the two-dimensional ECG and DSE-ResNet to realize the automatic classification of normal rhythm and 8 cardiac arrhythmias. Specifically, the original 12-lead ECG is spliced into a two-dimensional plane like a grayscale picture, where each column represents the time series of one lead, and each 'pixel' represents a voltage value of ECG. DSE-ResNet contains ResNet for extracting internal and inter-lead features and DSE for extracting global features of the two-dimensional ECG. Meanwhile, gender and age are introduced as auxiliary features into DSE-ResNet for training. In terms of hyper-parameter adjustment, we use a hyper-parameter selection method based on orthogonal experimental design, which greatly reduces the time for parameter adjustment. In the evaluation model stage, we use ensemble learning based on voting strategy to achieve better classification performance. The experimental results show that the combination of the two-dimensional ECG and DSE-ResNet may provide substantial improvement for the automatic classification of cardiac arrhythmia.

Materials and Method

Problem Definition

This paper aims to realize the automatic classification of normal rhythm and 8 cardiac arrhythmias based on the 12-lead ECG records. The input x of the proposed model includes two-dimensional ECG signals and basic information of the patients, and the output is the predicted labels corresponding to the normal rhythm and 8 cardiac arrhythmias. The inputs and reference label y constitute the training set $X = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$. The training goal of our model is to minimize the softmax cross-entropy loss function within a finite number of training epochs, where the softmax cross-entropy loss function is

$$\text{Loss}(X) = -\frac{1}{n} \sum_{i=1}^n \log \frac{\exp(p(x_i, y_i))}{\sum_j \exp(p(x_i, y_j))}, \quad (1)$$

where $p(x_i, y_i)$ and $p(x_i, y_j)$ represent the probability that the model predicts input x_i to the reference label y_i and the other label y_j , respectively.

Two-Dimensional ECG

Data Sources

The 12-lead ECG dataset³² used in this paper comes from CPSC2018, which is sampled at 500Hz and collected from 11 hospitals. It has 9831 samples, where 6877 (female: 3178 and male: 3699) samples are released for training and 2954 samples are kept private for testing. Each sample contains the 12-lead ECG signals, basic information of the patient (gender and age) and the reference label, where label corresponds to 9 categories: Normal rhythm, Atrial fibrillation (AF), First-degree atrioventricular block (I-AVB), Left bundle brunch block (LBBB), Right bundle brunch block (RBBB), Premature atrial contraction (PAC), Premature ventricular contraction (PVC), ST-segment depression (STD) and ST-segment elevated (STE). More details of the data sources are shown in Table 1.

Two-Dimensional ECG

In practice of clinical medicine, cardiologists usually need a multi-lead ECG as a basis for detection of cardiac arrhythmias. For example, the ECG abnormalities of patients with PAC are usually manifested in the leads V1, II, and aVF, and the typical abnormal ECG of patients with LBBB is mainly appeared in the leads I, V1, V2, V5, V6 and aVR³³. The detection of different cardiac arrhythmias requires the comprehensive information of 12-lead ECG, which means that both internal and inter-lead features play an important role in the classification of cardiac arrhythmia.

In order to extract the internal and inter-lead features of the 12-lead ECG at the same time, we perform two-dimensional processing on the 12-lead ECG. Specifically, the leads $a \in \mathbb{R}^{L \times 1}$ are spliced together to form a matrix $A \in \mathbb{R}^{L \times 12}$, where L is the length of leads. As shown in Fig. 1, the original 12-lead ECG is spliced and concatenated into a two-dimensional plane like a grayscale picture, where each column represents the time series of one lead, and each 'pixel' represents a voltage value of ECG.

Table 1 Data profile.

No.	Type	Record	Time length			Training set	Small test set	Hidden test set
			Mean	Min	Max			
1	Normal	918	15.43	10	64	895	69	394
2	AF	1098	15.04	9	74	1112	79	466
3	I-AVB	704	14.27	10	54	695	45	295
4	LBBB	207	14.94	9	65	203	15	97
5	RBBB	1695	14.62	10	118	1691	124	756
6	PAC	574	19.43	9	74	546	47	250
7	PVC	653	20.92	6	144	826	44	276
8	STD	826	15.50	8	138	825	58	340
9	STE	202	17.15	10	60	216	19	80
	Total	6877	15.95	6	144	7117	500	2954

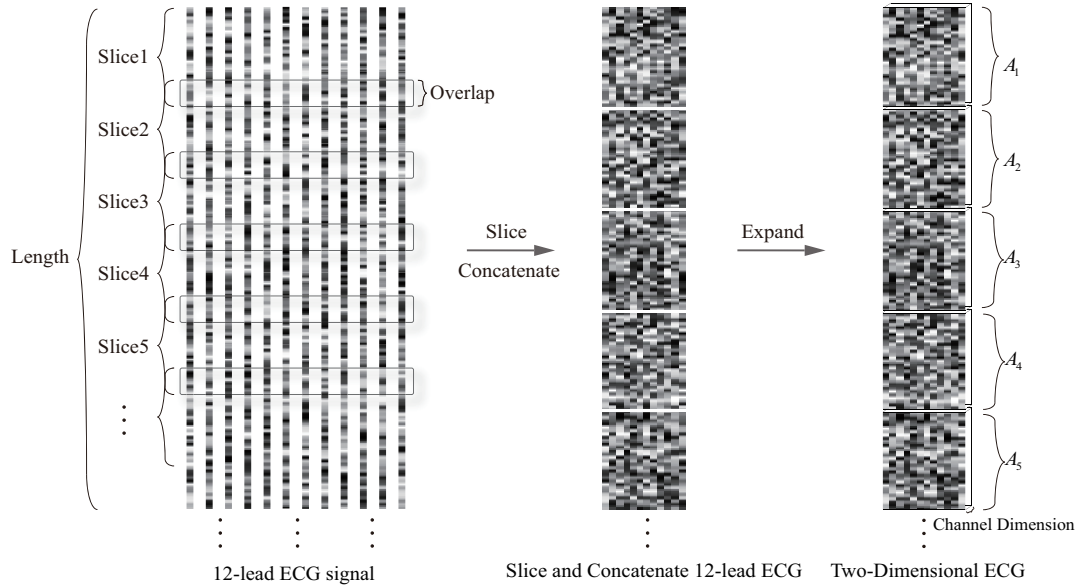


Fig. 1 Two-dimensional and expand dimension process.

Slicing

It can be seen from the Table 1 that the number of records in normal rhythm and different cardiac arrhythmias is quite different, and the time length of the original 12-lead ECG is also different. In order to make full use of the data and unify the length of the 12-lead ECG, we slice the two-dimensional ECG.

The ECG dataset of CPSC2018 contains 6877 training signals. We selected 500 ECG signals as a small number of test set to test the generalization ability of our model. The 12-lead ECG in the remaining 6377 ECG signals are sliced and used for training. The specific steps of slicing are as follows:

- (1) If the length of a two-dimensional ECG A is $L < 8192$, the length of A is filled with zeros to $L = 8192$.
- (2) If the length of A satisfies $8192 \leq L < 1.5 \times 8192$, A is cut off the extra data at the tail to $L = 8192$.
- (3) If the length of A satisfies $L \geq 1.5 \times 8192$, A is sliced into n pieces. The slice length is 8192, and the overlap length between slices is 4096. The number of slices is $n = \lfloor \frac{2L}{8192} \rfloor - 1$, where $\lfloor x \rfloor$ represents the largest integer less than x .

The slicing process effectively extends the training set from 6377 to 7117. The details of the training set after slicing and the small number test set are shown in Table 1.

Dimension Expansion

We added a dimension to the two-dimensional ECG, so that the dimension of the 12-lead signal satisfies the requirements of 2D-convolution (Conv2D) layer for the dimension of the input data. We call the newly added dimension the channel dimension. The two-dimensional ECG $A \in \mathbb{R}^{8192 \times 12}$ is expanded into $A \in \mathbb{R}^{8192 \times 12 \times 1}$, where the length is 8192, the lead number is 12 and the channel number is 1. During the training process, channel number of the output feature map of each convolutional layer changes synchronously with the number of convolution kernels. Fig. 1 shows the process of slicing, concatenating and expanding dimension of the original 12-lead ECG.

DSE-ResNet

As we all know, abnormal ECG signals are mainly manifested as changes in waveform shape and periodic rhythm³⁴. Some abnormal ECG signals are periodic and appear in almost every waveform cycle, other abnormal ECG signals are sporadic and only occur in a few heartbeat cycle. Meanwhile, patients of different ages and genders may have different ECG signals for same cardiac arrhythmia. Therefore, DSE-ResNet contains ResNet for extracting the internal and inter-lead features and DSE for extracting global features of two-dimensional ECG. Furthermore, we introduce the gender and age as auxiliary features for training.

Fig. 2 shows the overall structure of DSE-ResNet. Residual blocks are commonly used in CNNs to improve gradient flow through the networks and enable training of deeper networks. The ResNet in our model is composed of 1 residual block-1 and 9 residual block-2. Every Residual block has 2 Conv2D layers for extracting local features. The entire residual block has 20 Conv2D layers, where the size of the convolution kernel is (32, 1). The first and last 4 Conv2D layers have 12 and 192 convolution kernels respectively, and the number of convolution kernels is doubled for every 4 Conv2D layers in between. Activate Relu is used to increase the non-linear ability of the model. Batch Normalization and Dropout³⁵ play a good role in improving the training speed and preventing overfitting. Shortcut connection is used to complete the identity mapping of features and prevent the phenomenon of gradient disappearance and explosion. The 2D maximum pooling layer in each shortcut connection is used to adjust the dimension of features.

Squeeze-and-excitation (SE) module^{36,37} can squeeze features in the channel dimension and excite features to a higher-dimensional feature space, which has a global receptive field in a sense. The Detached SE (DSE) module in our model is independent of any residual block. It uses 2D global average pooling layer to extract the global features of the two-dimensional ECG from the channel dimension. Then the 4 dense layers in the DSE module map the extracted global features to a new feature space. Although the addition of the DSE module will increase the computational complexity of the entire model, it can increase the nonlinearity of the DSE-ResNet and establishes the correlation between channels. Patients with different age and gender may show different waveform states on the same type of cardiac arrhythmia. Therefore, gender and age are introduced into training as auxiliary features, which is helpful for the DSE-ResNet to capture the influence of basic information of the patient on cardiac arrhythmias.

Orthogonal Experiment

Excellent hyper-parameters can improve the performance and effect of model learning. We use Orthogonal Experimental Design (OED) to select combination of hyperparameter values.

OED is a design method for studying multi-factor and multi-level problems. It selects some representative points with uniform dispersion and neatness characteristics from the entire test point for testing based on orthogonality. The process of selecting representative points is often realized by constructing an orthogonal table. Based on relevant research experience, we selected the hyper-parameters that need to be adjusted and gave a corresponding set of estimated values. These estimated values constitute the entire test point of the orthogonal table. We use pairwise independent combinatorial testing (PICT)³⁸ to

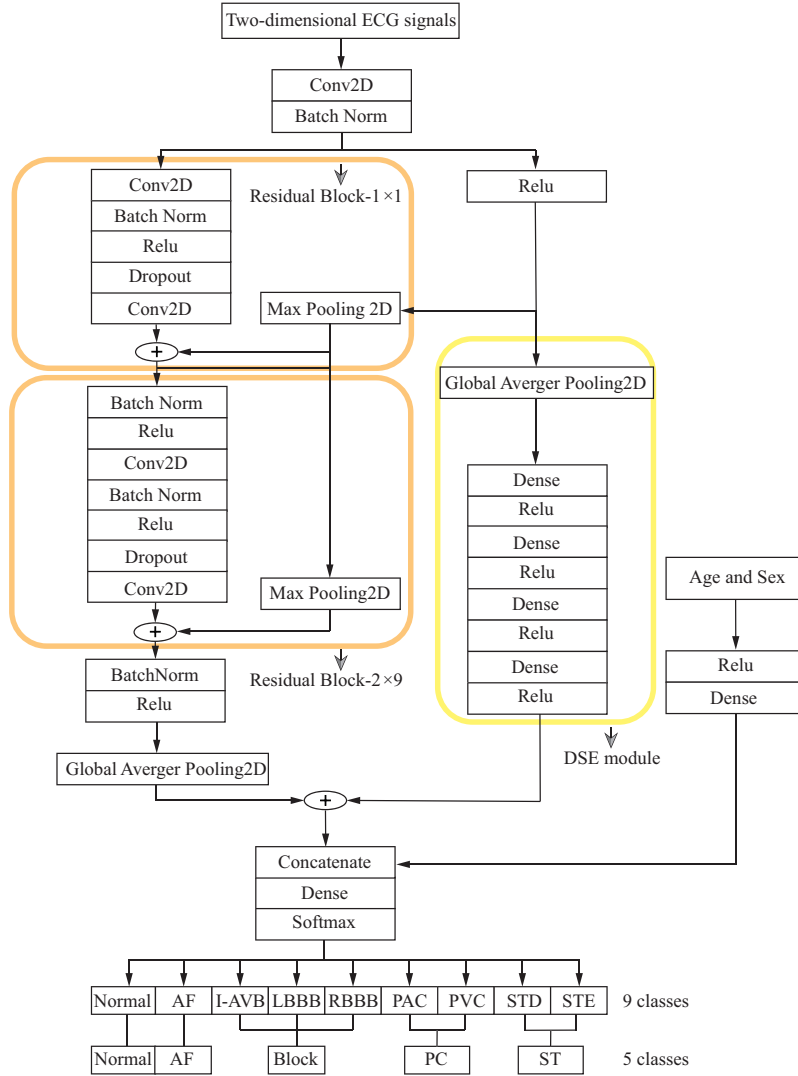


Fig. 2 Structure of DSE-ResNet.

construct an orthogonal table for the selected hyper-parameter values to obtain a representative combination of multiple sets of hyper-parameters.

Ensemble Model

Ensemble model accomplishes learning tasks by constructing and combining multiple learners³⁹. Compared with the classification performance of a single model, ensemble model can often achieve better classification performance and generalization ability⁴⁰. We use ensemble model to reduce the overall error of our model.

The ensemble model contains multiple learners, and each learner is the optimal DSE-ResNet trained based on a representative combination of hyper-parameters. A individual learner is called the single optimal model in this paper. The ensemble model uses a voting strategy to integrate all single optimal models. Specifically, each single optimal model will give a prediction value for the same test sample. Based on the multi-model voting strategy that the minority obeys the majority, the ensemble model takes the predicted value with the most votes as the final output value. Although the use of ensemble model increases the computational complexity, it can effectively improve the classification performance and fault tolerance of the model.

Experimental details

Software and hardware environment

The proposed model is built and trained using the Keras framework. All experiments are run on a server with Quadro P2200 video card and 5G video memory.

Data Pre-processing

Denoising

The original 12-lead ECG is affected by muscle noise, power-line noise and baseline wander. In order to reduce the possible impact of these noises on the learning process, we use a 1-order Butterworth bandpass⁴¹ filter to remove noises outside of 0.5Hz to 49Hz.

The results of the waveform of A1001 before and after denoising are shown in Fig. 3. An obvious result is that the high-frequency noise is suppressed, the overall waveform becomes smooth and the fluctuations are obvious.

Min-Max Normalization

Time series data can take a wide range of values in some cases, so it needs to be scaled to a fixed value interval to speed up the learning process⁴².

The amplitude of the voltage value in the original 12-lead ECG is $[-20.9\text{mV}, 20.7\text{mV}]$, and the amplitude difference between leads is large. It can be seen from Fig. 3 that the maximum and minimum amplitudes of the original 12-lead ECG are distributed in a symmetrical interval. Therefore, we use Min-Max Normalization to scale the amplitude of the voltage value of the two-dimensional ECG to the symmetrical interval $[-3\text{mV}, 3\text{mV}]$, which is

$$A_{ij} = R_{\min} + \frac{(R_{\max} - R_{\min})(A_{ij} - A_{\min})}{A_{\max} - A_{\min}}, \quad (2)$$

where $R_{\max} = 3\text{mV}$ and $R_{\min} = -3\text{mV}$ represent the boundary value of the normalized interval, A_{ij} is the voltage value in the i -th row and j -th column of the two-dimensional ECG, A_{\max} and A_{\min} respectively represent the maximum and minimum voltage value in the two-dimensional ECG. Fig. 3 shows the normalized result.

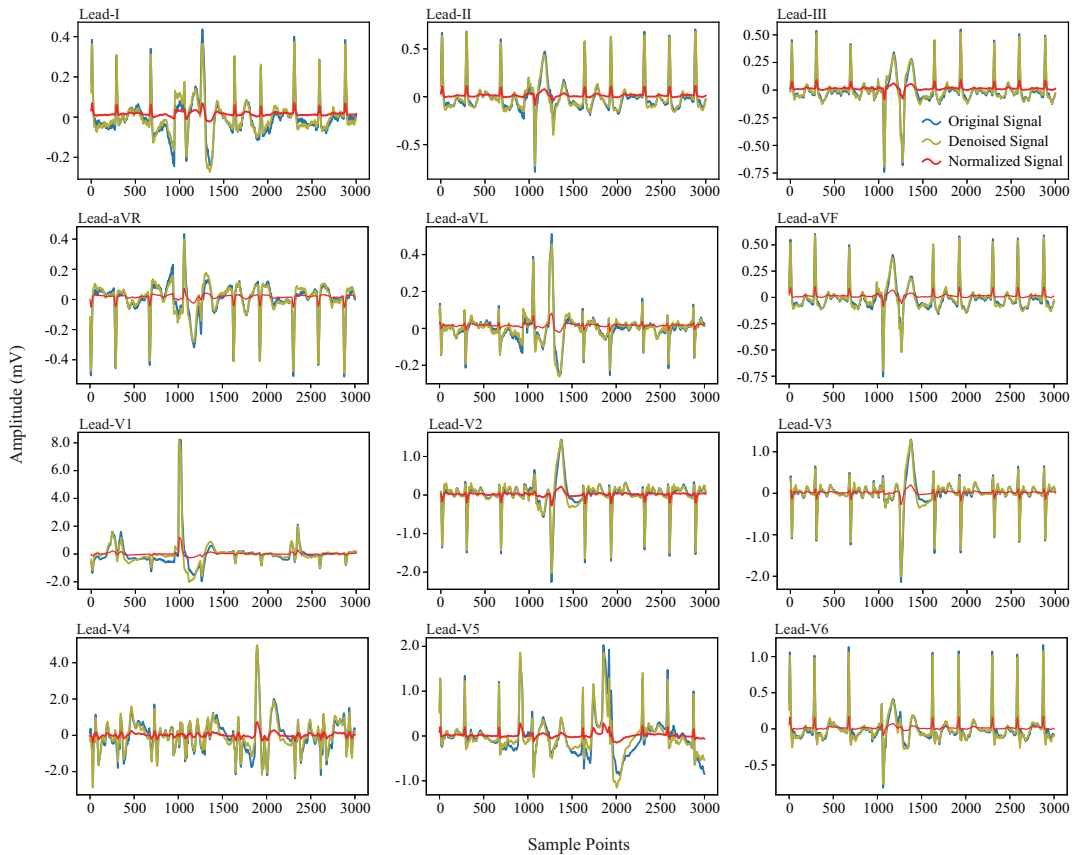


Fig. 3 Data pre-processing.

Orthogonal Experimental Design

We use OED to determine the combination of hyper-parameters values. Firstly, the batch-size is controlled to the maximum limit that the experimental machine can withstand. Secondly, three hyper-parameters are selected for orthogonal experiment,

Table 2 Orthogonal table.

No	Learning rate	Dropout	Momentum
1	0.1	0.8	0.5
2	0.15	0.3	0.9
3	0.05	0.5	0.5
4	0.15	0.8	0.7
5	0.05	0.3	0.7
6	0.1	0.5	0.9
7	0.15	0.3	0.5
8	0.05	0.8	0.9
9	0.1	0.3	0.7
10	0.15	0.5	0.7

including learning rate, dropout and momentum. According to the experience of the previous experiments, the value set of learning rate is $[0.05, 0.1, 0.15]$, the value set of dropout is $[0.3, 0.5, 0.8]$, and the value set of momentum is $[0.5, 0.7, 0.9]$. We use PICT to construct an orthogonal table to combine and match preset values. Table 2 shows the combination of preset values of hyper-parameters configured through PICT.

Results

Performance Metric

For multi-classification tasks, the average F_1 score⁴³ is an important indicator to measure classification performance. The F_{1i} score of the i th cardiac arrhythmia is the harmonic average of precision F_P and recall F_R , where F_P describes how many of the predicted positive samples are true positive samples, F_R describes how many true positive samples are picked out. Specifically, the F_{1i} score is defined as:

$$F_{1i} = \frac{2(F_P \times F_R)}{F_P + F_R}, \quad (3)$$

where $F_P = TP/(TP + FP)$ and $F_R = TP/(TP + FN)$, TP is the number of positive samples that are classified to be positive, FP is the number of negative samples that are classified to be positive, and FN is the number of positive samples that are classified to be negative. The average F_1 score among types is a comprehensive evaluation indicator for evaluating the overall performance of the model, which is defined as:

$$F_1 = \frac{1}{9} \sum_{i=1}^9 F_{1i}. \quad (4)$$

In addition, we also calculate the F_1 scores of 4 sub-abnormal types, i.e., the AF, block, premature contraction (PC) and ST-segment change (ST), where block consists of I-AVB, LBBB and RBBB, PC consists of PAC and PVC, and ST consists of STD and STE.

Performance on the small number of test set

We compared the F_1 scores of the single optimal models and the ensemble model based on the small number of test sets (500 ECG samples), where each single optimal model is an optimal model trained based on a representative combination of hyper-parameters in Table 2, and the ensemble model is based on the voting strategy to integrate all single optimal models.

Table 3 shows the F_1 scores of single optimal models and the ensemble model in the small number of test sets. An important result is that compared to the single optimal models, the ensemble model achieves good performance in multiple categories of tests, and achieves the highest F_1 scores in LBBB, PAC, STE and PC. More importantly, the average $F_1 = 0.843$ of the ensemble model is greater than that of the single optimal models. This result shows the advantage of ensemble model, it can effectively improve classification performance.

Table 3 Comparison of F_1 scores between the ensemble model and the single optimal models.

No.	Average F_1	Normal rhythm and 8 cardiac arrhythmias									4 sub-abnormal types			
		F_{Normal}	F_{AF}	$F_{\text{I-AVB}}$	F_{LBBB}	F_{RBBB}	F_{PAC}	F_{PVC}	F_{STD}	F_{STE}	F_{AF}	F_{Block}	F_{PC}	F_{ST}
1	0.783	0.739	0.962	0.846	0.786	0.912	0.742	0.854	0.659	0.545	0.962	0.885	0.793	0.629
2	0.816	0.821	0.963	0.845	0.938	0.922	0.683	0.830	0.750	0.595	0.963	0.903	0.761	0.709
3	0.810	0.745	0.955	0.821	0.970	0.928	0.682	0.860	0.730	0.595	0.955	0.905	0.773	0.695
4	0.776	0.738	0.933	0.824	0.938	0.915	0.692	0.806	0.742	0.400	0.933	0.896	0.747	0.688
5	0.835	0.787	0.954	0.876	0.938	0.941	0.744	0.907	0.763	0.606	0.954	0.926	0.826	0.728
6	0.824	0.783	0.919	0.851	0.938	0.932	0.738	0.892	0.761	0.600	0.919	0.912	0.814	0.727
7	0.817	0.763	0.969	0.830	0.941	0.938	0.736	0.876	0.762	0.541	0.969	0.911	0.807	0.704
8	0.780	0.743	0.938	0.804	0.811	0.902	0.710	0.870	0.686	0.556	0.938	0.867	0.789	0.652
9	0.832	0.760	0.938	0.857	0.941	0.931	0.742	0.914	0.779	0.629	0.938	0.914	0.824	0.743
10	0.828	0.787	0.942	0.901	0.909	0.919	0.742	0.864	0.778	0.606	0.942	0.914	0.802	0.738
EM	0.843	0.787	0.949	0.870	0.970	0.935	0.764	0.897	0.748	0.667	0.949	0.922	0.830	0.729

Performance on the hidden test set

By submitting our model to the competition official of CPSC2018, we get the test results based on the hidden test set (2954 ECG records). Fig. 4 shows the visual confusion matrix. For the sub-abnormal type ST, 53 samples with STD label and 27 samples with STE label are predicted to be Normal, and 19 samples with the Normal label are predicted to be STD. DSE-ResNet is not sensitive to changes in the ST, which may be due to the scarce number of samples of the ST and the highly similar waveform structure between ST and Normal. Furthermore, doctors disagree on the diagnosis of ST⁴⁴, leading to incorrect labeling of samples, which may also be one of the reasons. For sub-abnormal types AF and Blocks, the proposed model has excellent performance.

Table 4 shows the average F_1 and the F_1 of 4 sub-abnormal types of our model and the top five models with the highest average F_1 in CPSC2018. Note that the test results of the models in Table 4 are based on the same hidden test set. Tests show that the proposed model has the average $F_1 = 0.817$, which is only 0.02 behind the first-ranked model. It is worth noting that the proposed model achieves the best test results in 2 sub-abnormal types, which are $F_{\text{AF}} = 0.944$ and $F_{\text{Block}} = 0.913$, respectively. At the same time, the excellent test results obtained based on the hidden test set indicate that the proposed model has strong generalization ability and robustness. In summary, compared with the top five models in CPSC2018, the proposed model has an excellent overall performance and achieves an obvious improvement on identifying 2 sub-abnormal types, indicating that it has certain advantages in detecting some cardiac arrhythmias.

Table 4 Comparison of F_1 scores with the top five models in CPSC2018.

Model	Average F_1	F_{AF}	F_{Block}	F_{PC}	F_{ST}
Our model	0.817	0.944	0.913	0.786	0.738
Chen et al. ²⁸	0.837	0.933	0.899	0.847	0.779
Cai et al. ²⁹	0.830	0.931	0.912	0.817	0.761
He et al. ⁴⁵	0.806	0.914	0.879	0.801	0.742
Yu et al. ²⁹	0.802	0.918	0.890	0.789	0.718
Yan et al. ²⁹	0.791	0.924	0.882	0.779	0.709

Conclusion

In this paper, we propose a general model based on the two-dimensional ECG and DSE-ResNet to realize the automatic classification of normal rhythm and 8 cardiac arrhythmias. The two-dimensional processing method combines the original

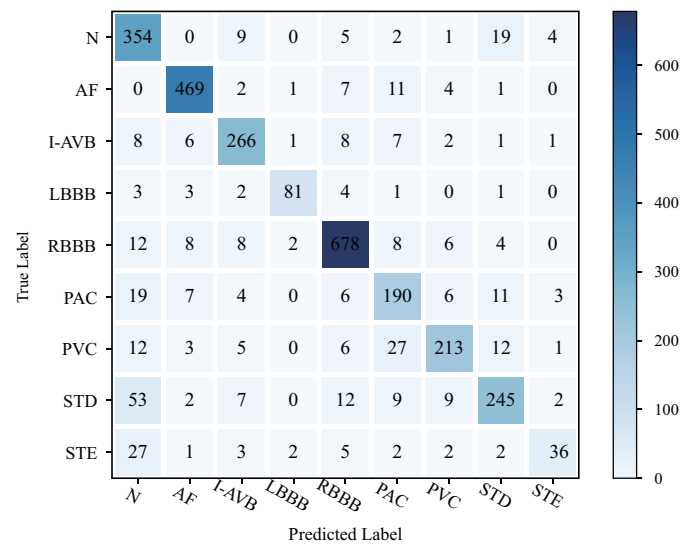


Fig. 4 Confusion matrix.

12-lead ECG into the same two-dimensional space, so that DSE-ResNet can simultaneously extract the internal and inter-lead features of the 12-lead ECG. Orthogonal experiment instead of grid search to select hyper-parameters reduces the computational complexity. Furthermore, the ensemble learning model based on voting strategy is used to improve classification and generalization performance. Experiments based on the small number of test set show that the classification performance of the ensemble learning model is much better than that of single models. Then we submitted our model to the competition official of CPSC2018 and got the test results based on the hidden test set. The comparison with the results of the top five models in the CPSC2018 shows that our model is excellent in the average F_1 value, and achieves the best test results in 2 sub-abnormal types. This shows that the combination of the two-dimensional ECG and DSE-ResNet is reasonable and effective, and can provide substantial improvements for the automatic classification of cardiac arrhythmia.

Our results not only provide a new perspective on the automatic classification of cardiac arrhythmia based on the 12-lead ECG, but also raise several questions. Based on the two-dimensional ECG, future research directions include exploring how to further improve the accuracy of prediction, how to reduce the prediction time, how to find redundant leads in the 12-lead ECG, and so on.

References

1. Van Mieghem, C., Sabbe, M. & Knockaert, D. The clinical value of the ecg in noncardiac conditions. *Chest* **125**, 1561–76, DOI: <https://doi.org/10.1378/chest.125.4.1561> (2004).
2. Zhang, D., Yang, S., Yuan, X. & Zhang, P. Interpretable deep learning for automatic diagnosis of 12-lead electrocardiogram. *iScience* **24**, 102373, DOI: <https://doi.org/10.1016/j.isci.2021.102373> (2021).
3. Wang, J. B., Wang, P. & Wang, S. P. Automated detection of atrial fibrillation in ecg signals based on wavelet packet transform and correlation function of random process. *Biomed. Signal Process. Control.* **55**, 101662, DOI: [ARTN101662](https://doi.org/10.1016/j.bspc.2020.101662) (2020).
4. Tsipouras, M. G., Fotiadis, D. I. & Sideris, D. An arrhythmia classification system based on the rr-interval signal. *Artif. Intell. Med* **33**, 237–50, DOI: <https://doi.org/10.1016/j.artmed.2004.03.007> (2005).
5. Exarchos, T. P. *et al.* A methodology for the automated creation of fuzzy expert systems for ischaemic and arrhythmic beat classification based on a set of rules obtained by a decision tree. *Artif. Intell. Medicine* **40**, 187–200, DOI: <https://doi.org/10.1016/j.artmed.2007.04.001> (2007).
6. Haseena, H. H., Mathew, A. T. & Paul, J. K. Fuzzy clustered probabilistic and multi layered feed forward neural networks for electrocardiogram arrhythmia classification. *J Med Syst* **35**, 179–88, DOI: <https://doi.org/10.1007/s10916-009-9355-9> (2011).
7. Mondejar-Guerra, V., Novo, J., Rouco, J., Penedo, M. G. & Ortega, M. Heartbeat classification fusing temporal and morphological information of ecgs via ensemble of classifiers. *Biomed. Signal Process. Control.* **47**, 41–48, DOI: <https://doi.org/10.1016/j.bspc.2018.08.007> (2019).

8. Lin, C. C. & Yang, C. M. Heartbeat classification using normalized rr intervals and morphological features. *Math. Probl. Eng.* **2014**, DOI: [Artn712474](https://doi.org/10.1155/2014/12474) (2014).
9. Afkhami, R. G., Azarnia, G. & Tinati, M. A. Cardiac arrhythmia classification using statistical and mixture modeling features of ecg signals. *Pattern Recognit. Lett.* **70**, 45–51, DOI: <https://doi.org/10.1016/j.patrec.2015.11.018> (2016).
10. Martis, R. J. *et al.* Application of higher order statistics for atrial arrhythmia classification. *Biomed. Signal Process. Control.* **8**, 888–900, DOI: <https://doi.org/10.1016/j.bspc.2013.08.008> (2013).
11. de Oliveira, L. S., Andreão, R. V. & Sarcinelli-Filho, M. Premature ventricular beat classification using a dynamic bayesian network. In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 4984–4987, DOI: <https://doi.org/10.1109/IEMBS.2011.6091235> (IEEE).
12. Zeng, X. D., Chao, S. & Wong, F. Ensemble learning on heartbeat type classification. In *Proceedings 2011 International Conference on System Science and Engineering*, 320–325, DOI: <https://doi.org/10.1109/ICSSE.2011.5961921> (IEEE).
13. De Chazal, P., O'Dwyer, M. & Reilly, R. B. Automatic classification of heartbeats using ecg morphology and heartbeat interval features. *IEEE transactions on biomedical engineering* **51**, 1196–1206, DOI: <https://doi.org/10.1109/TBME.2004.827359> (2004).
14. Li, H. *et al.* Classification of electrocardiogram signals with waveform morphological analysis and support vector machines. *Med Biol Eng Comput.* **60**, 109–119, DOI: <https://doi.org/10.1007/s11517-021-02461-4> (2022).
15. Ince, T., Kiranyaz, S. & Gabbouj, M. A generic and robust system for automated patient-specific classification of ecg signals. *IEEE Trans Biomed Eng* **56**, 1415–26, DOI: <https://doi.org/10.1109/TBME.2009.2013934> (2009).
16. Wang, J. S., Chiang, W. C., Hsu, Y. L. & Yang, Y. T. C. Ecg arrhythmia classification using a probabilistic neural network with a feature reduction method. *Neurocomputing* **116**, 38–45, DOI: <https://doi.org/10.1016/j.neucom.2011.10.045> (2013).
17. Martis, R. J., Acharya, U. R. & Min, L. C. Ecg beat classification using pca, lda, ica and discrete wavelet transform. *Biomed. Signal Process. Control.* **8**, 437–448, DOI: <https://doi.org/10.1016/j.bspc.2013.01.005> (2013).
18. Martis, R. J., Acharya, U. R., Prasad, H., Chua, C. K. & Lim, C. M. Automated detection of atrial fibrillation using bayesian paradigm. *Knowledge-Based Syst.* **54**, 269–275, DOI: <https://doi.org/10.1016/j.knosys.2013.09.016> (2013).
19. Ye, C., Kumar, B. V. K. V. & Coimbra, M. T. Heartbeat classification using morphological and dynamic features of ecg signals. *Ieee Transactions on Biomed. Eng.* **59**, 2930–2941, DOI: <https://doi.org/10.1109/Tbme.2012.2213253> (2012).
20. Osowski, S., Hoai, L. T. & Markiewicz, T. Support vector machine-based expert system for reliable heartbeat recognition. *IEEE Trans Biomed Eng* **51**, 582–9, DOI: <https://doi.org/10.1109/TBME.2004.824138> (2004).
21. Li, H. Q. *et al.* Novel ecg signal classification based on kica nonlinear feature extraction. *Circuits Syst. Signal Process.* **35**, 1187–1197, DOI: <https://doi.org/10.1007/s00034-015-0108-3> (2016).
22. Lagerholm, M., Peterson, C., Braccini, G., Edenbrandt, L. & Sornmo, L. Clustering ecg complexes using hermite functions and self-organizing maps. *IEEE Transactions on Biomed. Eng.* **47**, 838–848, DOI: <https://doi.org/10.1109/10.846677> (2000).
23. Guo, G., Wang, H., Bell, D., Bi, Y. & Greer, K. Knn model-based approach in classification. In *OTM Confederated International Conferences" On the Move to Meaningful Internet Systems"*, 986–996, DOI: https://doi.org/10.1007/978-3-540-39964-3_62 (Springer).
24. Ullah, A., Anwar, S. M., Bilal, M. & Mehmood, R. M. Classification of arrhythmia by using deep learning with 2-d ecg spectral image representation. *Remote. Sens.* **12**, 1685, DOI: <https://doi.org/10.3390/rs12101685> (2020).
25. Hannun, A. Y. *et al.* Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nat. medicine* **25**, 65–69, DOI: <https://doi.org/10.1038/s41591-018-0268-3> (2019).
26. Chen, B. *et al.* A study of deep feature fusion based methods for classifying multi-lead ecg. *arXiv preprint arXiv:1808.01721* DOI: <https://doi.org/10.48550/arXiv.1808.01721> (2018).
27. Wang, R., Fan, J. & Li, Y. Deep multi-scale fusion neural network for multi-class arrhythmia detection. *IEEE J Biomed Heal. Inf.* **24**, 2461–2472, DOI: <https://doi.org/10.1109/JBHI.2020.2981526> (2020).
28. Chen, T. M., Huang, C. H., Shih, E. S. C., Hu, Y. F. & Hwang, M. J. Detection and classification of cardiac arrhythmias by a challenge-best deep learning neural network model. *iScience* **23**, 100886, DOI: <https://doi.org/10.1016/j.isci.2020.100886> (2020).
29. The china physiological signal challenge 2018: Automatic identification of the rhythm/morphology abnormalities in 12-lead ecgs (2018).

30. Ribeiro, A. H. *et al.* Automatic diagnosis of the 12-lead ecg using a deep neural network. *Nat. communications* **11**, 1–9, DOI: <https://doi.org/10.1038/s41467-020-15432-4> (2020).
31. Zhao, Z. *et al.* Adaptive lead weighted resnet trained with different duration signals for classifying 12-lead ecgs. In *2020 Computing in Cardiology*, 1–4, DOI: <https://doi.org/10.22489/CinC.2020.112> (IEEE).
32. Liu, F. F. *et al.* An open access database for evaluating the algorithms of electrocardiogram rhythm and morphology abnormality detection. *J. Med. Imaging Heal. Informatics* **8**, 1368–1373, DOI: <https://doi.org/10.1166/jmihi.2018.2442> (2018).
33. Surawicz, B., Childers, R., Deal, B. J. & Gettes, L. S. Aha/accf/hrs recommendations for the standardization and interpretation of the electrocardiogram: part iii: intraventricular conduction disturbances a scientific statement from the american heart association electrocardiography and arrhythmias committee, council on clinical cardiology; the american college of cardiology foundation; and the heart rhythm society endorsed by the international society for computerized electrocardiology. *J. Am. Coll. Cardiol.* **53**, 976–981 (2009).
34. Zhang, J. *et al.* Ecg-based multi-class arrhythmia detection using spatio-temporal attention-based convolutional recurrent neural network. *Artif. Intell. Medicine* **106**, 101856, DOI: [ARTN101856](https://doi.org/10.1016/j.artmed.2020.101856) (2020).
35. Dahl, G. E., Sainath, T. N. & Hinton, G. E. Improving deep neural networks for lvsr using rectified linear units and dropout. In *2013 IEEE international conference on acoustics, speech and signal processing*, 8609–8613, DOI: <https://doi.org/10.1109/ICASSP.2013.6639346> (IEEE).
36. Zhu, J., Zhang, Y. & Zhao, Q. Atrial fibrillation detection using different duration ecg signals with se-resnet. In *2019 IEEE 21st International Workshop on Multimedia Signal Processing (MMSP)*, 1–5, DOI: <https://doi.org/10.1109/MMSP.2019.8901729> (IEEE).
37. Zhu, Z. *et al.* Classification of cardiac abnormalities from ecg signals using se-resnet. In *2020 Computing in Cardiology*, 1–4, DOI: <https://doi.org/10.22489/CinC.2020.281> (IEEE).
38. McCaffrey, J. D. Generation of pairwise test sets using a genetic algorithm. In *2009 33rd annual IEEE international computer software and applications conference*, vol. 1, 626–631, DOI: <https://doi.org/10.1109/COMPSAC.2009.91> (IEEE).
39. Hong, S. *et al.* Encase: An ensemble classifier for ecg classification using expert features and deep neural networks. In *2017 Computing in cardiology (cinc)*, 1–4, DOI: <https://doi.org/10.22489/CinC.2017.178-245> (IEEE).
40. Sodmann, P., Vollmer, M., Nath, N. & Kaderali, L. A convolutional neural network for ecg annotation as the basis for classification of cardiac rhythms. *Physiol. Meas.* **39**, 104005, DOI: [ARTN104005](https://doi.org/10.1088/1361-6578/aa9005) (2018).
41. Daud, S. & Sudirman, R. Butterworth bandpass and stationary wavelet transform filter comparison for electroencephalography signal. In *2015 6th international conference on intelligent systems, modelling and simulation*, 123–126, DOI: <https://doi.org/10.1109/ISMS.2015.29> (IEEE).
42. Bhanja, S. & Das, A. Impact of data normalization on deep neural network for time series forecasting. *arXiv preprint arXiv:1812.05519* DOI: <https://doi.org/10.48550/arXiv.1812.05519> (2018).
43. Yao, Q. H., Wang, R. X., Fan, X. M., Liu, J. K. & Li, Y. Multi-class arrhythmia detection from 12-lead varied-length ecg using attention-based time-incremental convolutional neural network. *Inf. Fusion* **53**, 174–182, DOI: <https://doi.org/10.1016/j.inffus.2019.06.024> (2020).
44. McCabe, J. M. *et al.* Physician accuracy in interpreting potential st-segment elevation myocardial infarction electrocardiograms. *J. Am. Hear. Assoc.* **2**, e000268, DOI: <https://doi.org/10.1161/JAHA.113.000268> (2013).
45. He, R. *et al.* Automatic cardiac arrhythmia classification using combination of deep residual network and bidirectional lstm. *IEEE Access* **7**, 102119–102135, DOI: <https://doi.org/10.1109/ACCESS.2019.2931500> (2019).

Ethics statement and consent to participate

The database used in the study was an open access database , <https://doi.org/10.1166/jmihi.2018.2442>. It can be obtained in <https://physionet.org/content/challenge-2020/1.0.1/> or <http://2018.icbeb.org/Challenge.html>. Therefore, no ethics statement and informed consent is required for this study. All methods in this study were carried out in accordance with relevant guidelines and regulations. This study was carried out in compliance with the Declaration of Helsinki.

Acknowledgements

This work was supported by National Nature Science Foundation of China under Grant No. 61903208, Young doctorate Cooperation Fund Project of Qilu University of Technology (Shandong Academy of Sciences) under Grant No. 2019BSHZ0014, Program for Youth Innovative Research Team in the University of Shandong Province in China under Grant No. 2019KJN010, and Graduate Education and Teaching Reform Research Project of Qilu University of Technology in 2019 under Grant No. YJG19007.

Author contributions statement

J.L. contributed mainly to data processing, experimental design, and writing the paper. S.P. was responsible for the overall idea and data analysis, and revising the paper. F.X. and S.Z. was responsible for data processing and experimental guidance. M.S. contributed to the revision of the article.

Competing interests

The authors declare no competing interests.

Data availability

The train datasets used during the current study available in the The China Physiological Signal Challenge 2018, <http://2018.icbeb.org/Challenge.html>

The test datasets used during the current study are not publicly available for scoring purposes, but test scores can be obtained by submitting the model to The China Physiological Signal Challenge 2018.

The datasets generated and analysed during the current study are available from the corresponding author on reasonable request.