Proceedings of the 40th Chinese Control Conference
July 26-28, 2021, Shanghai, China

Classification of myocardial infarction based on ECG signals and multi-network stacking model

Tianqi Zhao[1], Muqing Deng[2], Peng Lin[1], Jianzhong Wang[1], Jiuwen Cao[1]
1. School of Automation(School of Artificial Intelligence), Hangzhou Dianzi University, Zhejiang 310018, China
E-mail: 739174523@qq.com
2. School of Automation, Guangdong University of Technology, Guangdong 510006, China
E-mail: mqdeng@gdut.edu.cn

Abstract: Myocardial infarction is one of the common cardiovascular diseases, and it also is a major cause of death and disability worldwide every year. Therefore, early detection of myocardial infarction is vital for patients. Deep learning algorithms have been used successfully for the recognition of myocardial infarction in recent years, but most of them are based on the segmentation of heartbeat signals and the single deep learning model. In this paper, we proposed a new myocardial infarction classification method based on an improved ECG feature extraction scheme and the multi-network stacking model. Firstly, the mean amplitude spectrum (MAS) is used in feature extraction to elaborate the frequency distribution difference of each wave among ECG signals. Secondly, due to the lack of time-domain features in the MAS map, we proposed a novel time-frequency feature fusion method. Finally, the time-frequency feature is fed to the multi-network stacking model for the classification of myocardial infarction. To improve the diversity of feature learning, this paper researches to stack different networks for learning the characteristic of the time-frequency feature. Meanwhile, we used the improved training methods to solve the problem of data imbalance. Experiments using the PTB database show that this feature extraction scheme and multi-network stacking model are effective for the classification of myocardial infarction.

Key Words: ECG, Myocardial infarction, Feature extraction, Deep learning

1 Introduction

Myocardial infarction (MI) [1], a common cardiovascular disease, is one of the leading causes of death worldwide. In fact, many patients don't realize that they have symptoms of myocardial infarction and most of them miss a lot of valuable time for treatment. Persistence of cardiac inflammation may cause irreversible damage to heart and can even lead to death. Therefore, early diagnosis and treatment of myocardial infarction is vital for patients.

Electrocardiogram (ECG) records the electrical activity of each cardiac cycle of heart and is the traditional basis for diagnosis in clinical practice of myocardial infarction. One of the main characteristics of myocardial infarction in ECG is the ST-segment and T-wave changes [2]. However, due to the weakness of the ECG signal which can be buried by various types of noise, characteristics of myocardial infarction may not be obvious or even undetectable with the naked eye. In addition, ECG signals of some patients with old myocardial infarction will show normal waveform. These phenomena can cause situations such as missed diagnosis or misdiagnosis. Therefore, the diagnosis of myocardial infarction cannot be effectively accomplished only based on morphological features in ECG signals.

With the development of signal processing technology and artificial intelligence algorithms, ECG detection and analysis of myocardial infarction has become a hot topic in the field of medical and artificial intelligence. More and more information that can't be detected by the naked eye is mined from ECGs. In recent years, deep learning [3] has revolutionized several fields. And it also has shown remarkable potential in the field of medical applications [4]. These techniques provide new ideas for recognition algorithms of myocardial infarction. Lui et al. [5] used convolutional recurrent neural network for the detection of myocardial infarction using data from standard I lead. Baxt et al. [6] used an artificial neural network to identify 351 patients with myocardial infarction and obtained a sensitivity of 97.2% and specificity of 96.2%. Reasat et al. [7] implemented shallow convolutional neural network approach which used varying filter sizes in the same convolution layer for myocardial infarction detection. Acharya et al. [8] used an 11-layer CNN for myocardial infarction detection of lead II ECG signals and achieved a relatively high accuracy of 95.22% even with noise in ECG data. Baloglu et al. [9] proposed a 10-layer deep CNN model without any handcrafted feature extraction based on 12-lead ECG signals, and it achieved a better classification rate for myocardial infarction detection.

However, most of these methods utilize only raw ECG signals in the time domain and one single deep learning structure, which can capture only shallow features. Moreover, there are some limitations of these methods. First, the segmentation of heartbeat need a lot of calculational time. Second, the complexity of the model requires tremendous computational power. Aiming at these problems, we proposed an improved ECG feature extraction scheme and multi-network stacking model. Firstly, the signal is preprocessed to remove interference. According to the characteristics of myocardial infarction, we used the mean amplitude spectrum feature map and single-lead ECG signal for feature learning, and the stacked model is used to increase the diversity of feature learning. In the training phase, the improved training method and loss function are used to solve the problem of data imbalance. The results indicated that our model worked better than the single feature or the single model.

This work is supported by National Natural Science Foundation (NNSF) of China under Grant 61803133, U1909209.

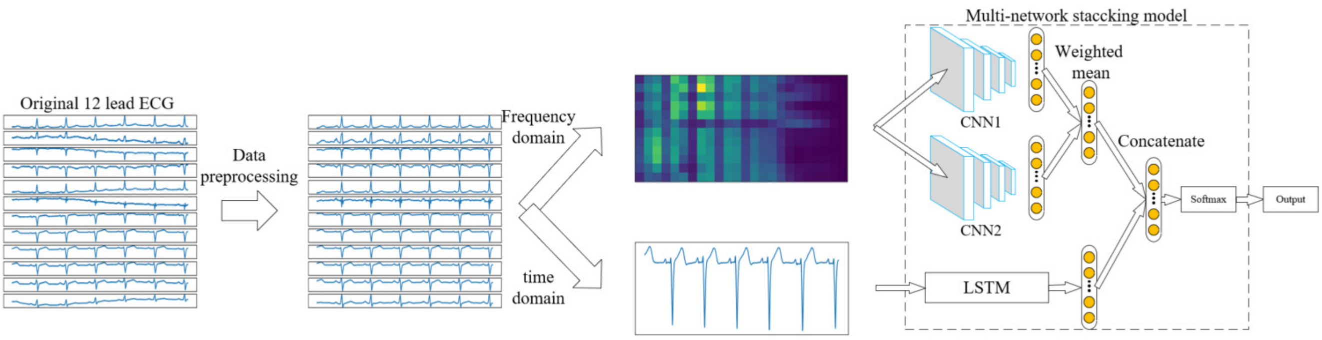

Fig. 1: The proposed algorithm

2 The proposed algorithm

The proposed algorithm is shown in Figure 1. In this section, the algorithm can be divided into four parts: (1) Data preprocessing; (2) Time-frequency domain feature extraction; (3) Construction of the multinetwork stacking model. In the training phase, the original ECG signal is preprocessed to remove the baseline wandering, high-frequency interference and power line interference. Then the mean amplitude feature map and time-domain feature are extracted as training database. Then we bulit the multi-network stacking model for feature learning and input the training data into it. The combination of networks enables the proposed model to extract complementary features. In the testing phase, the pre-trained multi-network stacking model is used to determine the category of the testing data.

2.1 Data pre-processing

In this paper, median filtering [10] is used to eliminate baseline drift we choose a median filter with a window width of 250. The Butterworth low-pass filter of order 5 with cutoff frequencies 70 Hz was applied to remove high-frequency interference. Using the Butterworth band-stop filter of order 5 (stop-band: 49-51 Hz) to remove power line interference. Figure 2 shows the original signals and preprocessed signals for ECG signals.

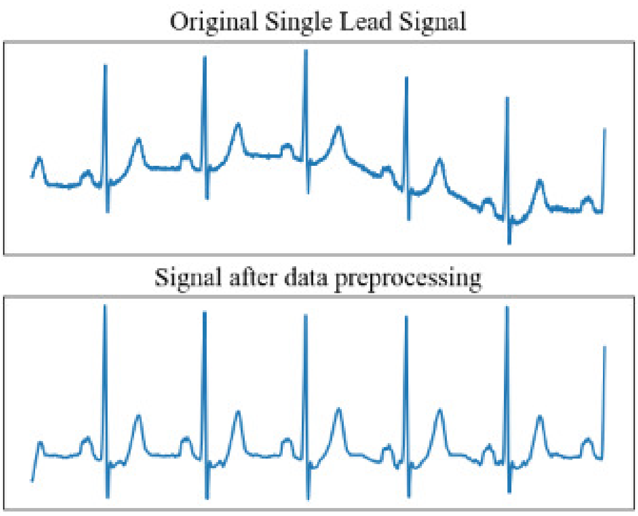

Fig. 2: The original signals and preprocessed signals

2.2 Mean amplitude spectra feature map extraction

To obtain an effective feature representation for classification of myocardial infarction in ECG, we used the amplitude spectrums [11] on the frequency bandwidth of different waves for feature learning. The bandwidth of the ECG signals ranges from 0-58±19 Hz. Among the entire cardiac cycle, the bandwidth of P-wave, QRS-wave and T-wave are 0-8±3 Hz, 0-55±19 Hz and 0-11±2 Hz, respectively [12].

For each of the 12 leads, the original ECG signal is transformed using FFT to frequency components at first, and components with a bandwidth of 0-70 Hz are then used for feature extraction. Then the bandwidth of 0-70 Hz is divided into 24 frequency sub-bands. To avoid the influence of power line frequency interference, the bandwidth of 49-51 Hz should be removed. The frequency and amplitude of the T-wave are important indicators for diagnosis of myocardial infarction, so the bandwidth of T-wave is divided into more detailed sub-band. Each lead is processed as above, a 12×24 feature map is constructed as the feature for each ECG signal. The specific calculation is as follows.

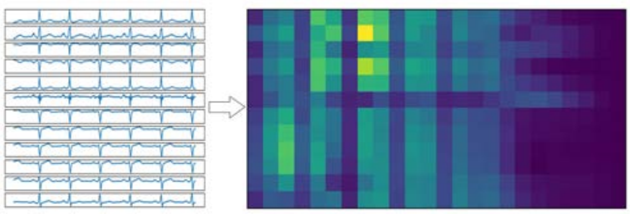

Fig. 3: MAS map

Firstly, we performed feature extraction based on Fast Fourier transform (FFT). FFT is a fast computational form of the Discrete Fourier transform (DFT). For each lead, the Discrete Fourier transform on a frame of ECG signal is:

$$X_k = \sum_{n=0}^{N-1} x(n) e^{-\frac{2\pi jkn}{N}} \tag{1}$$

where N is length of the signal and $k = 0, 1, 2, \ldots, N-1$. The amplitude spectrum can be calculated as:

$$P(k) = |X_k| \tag{2}$$

For the T-wave and the other waves, the ranges of k are 0-13, 13-70, respectively. For the T-wave, as an example, the frequency band 0-13 Hz is divided into 13 parts on average.

And 13-70 Hz is divided into 11 parts on average. To characterize the difference of each wave frequency distribution, the 24-band MAS is computed by:

$$MAS_i = mean(P(k), k \in K_i) \quad (3)$$

where $i = 1, 2, \ldots, 24$ represents the i-th frequency subband and K_i denotes the set of values of k of the i-th subband. This calculation transforms each lead in the 12-lead ECG into a vector with 24 points, and a two-dimensional matrix of 12×24 is obtained by computing all 12 leads. The MAS map is shown in Figure 3.

2.3 Time-domain feature representation

MAS includes frequency domain information but not morphological features. However, morphological changes which are directly perceived through the senses are also important. This shallow feature is the first basis for the diagnosis of myocardial infarction. Therefore, we extract the single lead signal as the time-domain feature for training of model.

Changes due to posterior, anterior, and septal myocardial infarction are reflected in lead V2 data [13]. It can express many kinds of ECG signals of myocardial infarction. Hence, this lead can reflect morphological changes caused by infarction in many heart regions. V2 are chosen after data preprocessing. The sampling rate of ECG signal is 1000 Hz, and the bandwidth of the entire cardiac cycle mainly 0-70 Hz. Therefore, we use downsampling technology to reduce the sampling rate of ECG signal to 200 Hz, which will not cause loss of ECG information. Finally, we obtain the single lead signal after downsampling as one of the features.

2.4 Multi-network stacking model

The previous algorithms of myocardial infarction are mostly based on a single deep learning model. However, it seems to only study the feature learning capability of a basic model, which may not be rigorous as the feature diversity by deep learning models with different structures and kinds is not well explored. In this paper, the model collects the mean amplitude feature map samples and V2 single-lead time domain signal samples in the input layer. In the multi-network stacking model, two CNN of different structures are implemented that take as input the MAS maps, and the V2 single lead signal after pre-processing and downsampling is used as the input of LSTM.

For the first CNN in multi-network staking model, it includes two convolutional layers, with 256 and 512 convolutional kernels. The kernel size in each convolutional layer is 3×3. To extract the features more precisely and reduce model parameters, a max-pooling layer with the size of 2×2 is added after the convolution layer. Then we added a fully connected (FC) layer with the size of 128 for feature learning. For each convolutional layer and FC layer, we used the rectified linear unit (RELU) [14] as the activation function. RELU is defined as:

$$f = max(0, x) \quad (4)$$

For the second CNN, we have made the following changes based on the structure of the first CNN. The average pooling layers were applied to replace the max-pooling layers after convolution layers. In addition, we replace RELU with LRELU [15]. LRELU is defined as:

$$f = max(0, x) + leak * min(0, x) \quad (5)$$

LRelu assigns a non-zero slope to all negative values, where leak is a very small constant so that some negative axis values are preserved and not all information of the negative axis is lost. In this paper, we set $leak = 0.1$. For the LSTM in multi-network staking model, it comprises two LSTM layers, each having 128 cells, then there is a FC layer with the size of 128. The first feature vector is obtained by weighted average of FC layers of CNNs, and the second feature vector is obtained by LSTM. Next, we concatenate the feature vectors. Finally, the new hybrid feature vector with the size of 256 is acted as the input of the softmax function for classifying.

In this section, The multi-network stacking model achieves discriminative feature learning by training multiple diverse networks in parallel. This method improves the performance of multi-network stacking model by combining the structure of the different network. The different structure of convolution neural network makes them have different strategies for feature extraction. And different kinds of neural networks can extract different types of features. For example, LSTM is more sensitive to time-domain signal, while CNN is more suitable for processing image data. Therefore, We use multi-network stacking model to increase the comprehensiveness of feature extraction. Figure 4 shows the structure and parameters of multi-network stacking model.

3 Experiment

3.1 Dataset

In this paper, 148 myocardial infarction patients and 52 healthy patients were selected from the PTB ECG database [16], and each patient selected 12 lead ECG signals with the 20s. From the 200 samples, 40 healthy samples and 110 myocardial infarction samples were randomly selected as the training set, and the remaining 50 samples were used as the test set.

3.2 Experiment-set

All experiments were implemented using windows 10 OS, NVIDIA GeForce RTX 2060 GPU, AMD Ryzen 5 3600 CPU@3.60Hz, and 16GB RAM. The program was implemented by deep learning framework Tensorflow 2.0.0 with Python 3.7.

The model training process is completed by backpropagation. Adam optimizer is used in this paper, and the learning rate is set to 1e-4. We propose to improve the class imbalance by using weight cross-entropy loss such that it downweights the loss assigned to MI samples. The weight cross-entropy is defined as:

$$L = w * y_{true} * log(y_{pre}) + (1 - w) * (y_{true}) log(1 - y_{pre}) \quad (6)$$

where w is the weight of the loss function when the label is 0. For each batch fed into the model, we set that the positive and negative samples are roughly balanced. Model was trained for 128 epochs, and a batch size of 80 samples which randomly selected include 40 healthy samples and 40 MI samples. This training method and the weighted cross-entropy loss function work together to solve the problem of unbalanced myocardial infarction data to some extent.

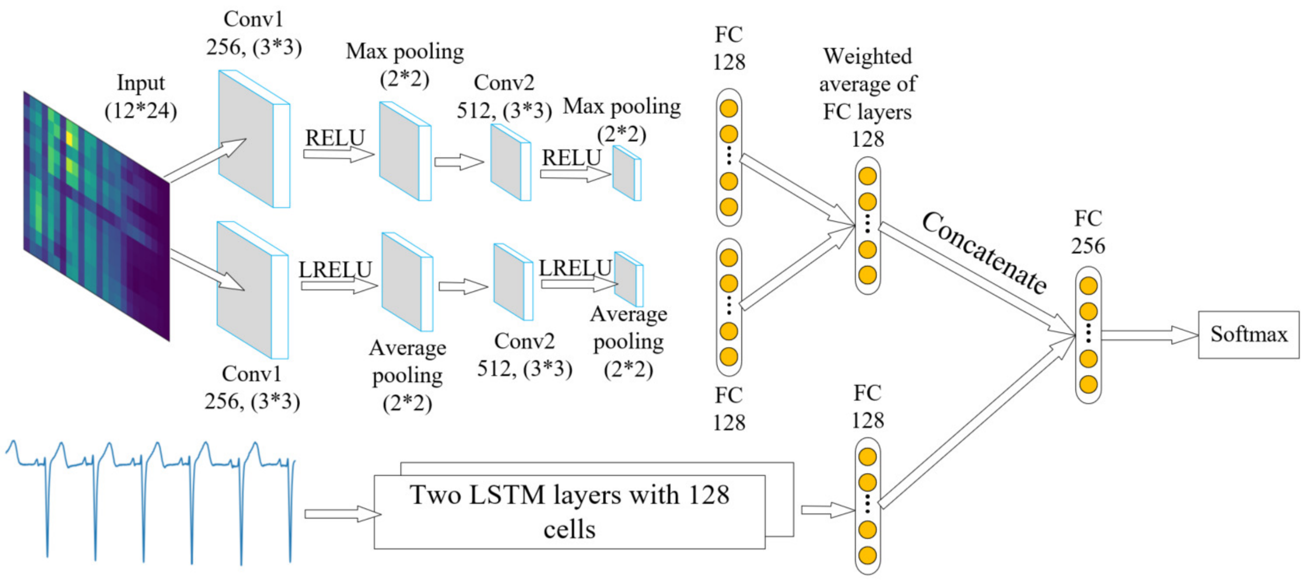

Fig. 4: The structure and parameters of multi-network stacking model

3.3 Evaluation and analysis

In order to accurately evaluate the classification performance of the algorithm, a confusion matrix is used in this paper to evaluate the classification results, which mainly includes three evaluation indexes: accuracy, sensitivity, and specificity. Evaluation indexes are defined as follows:

$$Accuracy(Acc) = \frac{T_p + T_n}{T_p + F_n + T_n + F_p} \tag{7}$$

$$Sensitivity(Sen) = \frac{T_p}{T_p + F_n} \tag{8}$$

$$Specificity(Spe) = \frac{T_n}{T_n + F_p} \tag{9}$$

where, T_p is the number of true positive results, F_p is the number of false positive results, T_n is the number of true negative results, and F_n is the number of false negative results.

In this section, we discuss the four methods (CNN + MAS, Stacked CNN + MAS, LSTM + V2, and Multi-network stacking model) results that have been tested on the testing set. The results are represented in Table 1.

Table 1: Results

Model	Feature	Acc	Sen	Spe
CNN	MAS	0.84	0.625	0.941
Stacked CNN	MAS	0.9	0.769	0.943
LSTM	V2	0.82	0.6	0.914
Multi-network	MAS + V2	0.94	0.909	0.949

Depending on the and type of features and models in the classification experiment, different results can be achieved. Only using CNN + MAS or LSTM + V2 for classification does not yield well results. As a contrast, stacked CNN can slightly improve the performance of the model. It means that stacking can increase the generalization ability of the model. It is observed that the multi-network stacking model can achieve a much better performance than the others.

A large number of heartbeat data are segmented to compose the data set in most of the previous methods. But it can not extract the long-term dependence of the relationship from ECG signals. Aiming at these problems, we extract MAS features map according to the characteristics of myocardial infarction and uses LSTM to extract long-term dependency. Meanwhile, using the long-term ECG signal will reduce sample sizes and makes each sample can extract more information. Features are targeted and stacked model increase the diversity of feature learning. It means that a large amount of data is not required and that well results can be achieved.

4 Conclusion

In this paper, the myocardial infarction classification based on the improved ECG features extraction scheme and multi-network stacking model has been studied. Data preprocessing has been done such as filtering the noise from the ECG signals at first. The single feature description can not accurately represent ECG signals. Therefore, we extract the MAS feature map with the size of 12 $\times$ 24 and fuse time-domain features for model training process. The hybrid features are fed into the multi-network stacking model and the classification result is gained by working together. In this model, we set two CNN with different structures and one LSTM to stack. It takes advantage of the local characteristics by CNN and the long-term dependencies captured by LSTM. Meanwhile, the use of loss function and training method working together improves the problem of data imbalance. It has been observed that the test accuracy (ACC), sensitivity (Sen) and specificity (Spe) are 94%, 90.9% and 94.9%, respectively. In addition, to validate the efficiency of our algorithms, we have experimented on single feature and single model to contrast. The experiment shows that the feature extraction scheme and multi-network stacking model are effective for classification. The proposed algorithm as a complementary diagnostic tool may assist the doctors in some situations.

References

[1] Thygesen K, Alpert J S, Jaffe A S, et al. Third universal definition of myocardial infarction[J]. Circulation, 2012, 113 (2):2020–2035.

[2] Lee D C, Albert C M, Narula D, et al. Estimating Myocardial Infarction Size With a Simple Electrocardiographic Marker Score[J]. Journal of the American Heart Association, 2020, 9(3):e014205.

[3] Lecun Y, Bengio Y, Hinton G. Deep learning[J]. Nature, 2015, 521(7553):436.

[4] Litjens G, Kooi T, Bejnordi B E, et al. A survey on deep learning in medical image analysis[J]. Medical Image Analysis, 2017.42: 60–88.

[5] Lui H W, Chow K L. Multiclass classification of myocardial infarction with convolutional and recurrent neural networks for portable ECG devices - ScienceDirect[J]. Informatics in Medicine Unlocked, 2018, 13:26-33.

[6] Baxt WG. Use of an artificial neural network for the diagnosis of myocardial infarction. Ann Intern Med. 1991 Dec 1;115(11):843-8. doi: 10.7326/0003-4819-115-11-843. Erratum in: Ann Intern Med 1992 Jan 1;116(1):94. PMID: 1952470.

[7] Reasat T, Shahnaz C. Detection of Inferior Myocardial Infarction using Shallow Convolutional Neural Networks[C]// IEEE, 2017.

[8] Acharya U R ,Fujita H ,Oh S L ,et al. Application of Deep Convolutional Neural Network for Automated Detection of Myocardial Infarction Using ECG Signals[J]. Information Sciences, 2017, 415.

[9] A U B B , A M T, A O Y, et al. Classification of myocardial infarction with multi-lead ECG signals and deep CNN[J]. Pattern Recognition Letters, 2019, 122:23-30.

[10] Hao W, Chen Y, Xin Y. ECG baseline wander correction by mean-median filter and discrete wavelet transform. Annu Int Conf IEEE Eng Med Biol Soc. 2011;2011:2712-5. doi: 10.1109/IEMBS.2011.6090744. PMID: 22254901.

[11] Hu W, Cao J, Lai X, et al. Mean amplitude spectrum based epileptic state classification for seizure prediction using convolutional neural networks[J]. Journal of Ambient Intelligence and Humanized Computing, 2019.

[12] He W, Chen L. Frequency Distribution and Effective Band Widths of Electrocardiac Signal and Its Components[J]. Journal of Biomedical Engineering, 1996, 13(4):336-336. (in Chinese)

[13] Sadhukhan D, Pal S, Mitra M. Automated Identification of Myocardial Infarction Using Harmonic Phase Distribution Pattern of ECG Data[J]. IEEE Transactions on Instrumentation & Measurement, 2018:1-11.

[14] Glorot X, Bordes A, Bengio Y. Deep Sparse Rectifier Neural Networks[J]. Journal of Machine Learning Research, 2011, 15:315-323.

[15] Zhang X, Zou Y, Shi W. Dilated convolution neural network with LeakyReLU for environmental sound classification[C]//2017 22nd International Conference on Digital Signal Processing (DSP). IEEE, 2017.

[16] Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. Circulation. 2000 Jun 13;101(23):E215-20. doi: 10.1161/01.cir.101.23.e215. PMID: 10851218.