Advances in Computer, Signals and Systems (2022)
Clausius Scientific Press, Canada

DOI: 10.23977/acss.2022.060506
ISSN 2371-8838 Vol. 6 Num. 5

A Classification Scheme for ECG Signals Based on Bidirectional LSTM Model

Shuoxuan Zhang, Xinmi Zhang, Yuanyuan Huang, Zhenlin Dai, Jianhua Wang

Sydney Smart Technology College, Northeastern University at Qinhuangdao, Tangshan, 066004, China

Keywords: Bidirectional LSTM, ECG signal, Deep learning, ECG, Feature extraction

Abstract: The application of ECG to diagnose cardiovascular diseases is a common method in clinical medicine, so the use of deep learning tools to achieve automatic analysis and classification of ECG has been a research direction for a wide range of researchers. This paper proposes a classification model for ECG signals based on a bidirectional LSTM model which is trained and tested using the dataset used for the PhysioNet 2017 computational cardiology challenge. The data are normalized and then processed by feature extraction. After passing a bidirectional LSTM layer, a fully connected layer, a softmax layer, and a classification layer in the model, and finally achieve the binary classification of normal signals and atrial fibrillation signals. In this process, the feature of bidirectional LSTM that can integrate contextual information is fully utilized. The experiments show that the classification accuracy of the model reaches 94.1%, demonstrating a good classification result.

1. Introduction

The number of patients with cardiovascular disease has been increasing in recent years, posing a serious threat to human health. Arrhythmia is a common type of cardiovascular disease with a long latency period and is not easily detected, therefore it is crucial to detect it. [1] ECG is an important tool in the diagnosis of arrhythmic diseases. It reflects the electrical signal changes of cardiomyocytes from excitation to resting state [2], and is an vital tool for medical professionals to assist in diagnosis and pathological analysis. The application of computer technology for automatic recognition and classification of ECG is of high value in preventing misdiagnosis and leakage, improving efficiency, and reducing medical costs[3]. Many researchers have made different attempts on deep learning algorithms. WANG Guanjun et al.[4] used K-nearest neighbor algorithm, random forest, logistic regression and support vector machine algorithm for binary classification. Respectively, among which support vector machine outperformed the other three in model prediction. But support vector machine has lower diagnostic sensitivity due to low recall rate and higher risk of misclassification due to low precision rate. SUN Bo et al.[5] proposed a hybrid model based on CNN+SVM for ECG signal identification and judgment, with CNN for feature extraction and SVM for classification. The combination of which has better stability and improves the accuracy in small data sets. LIU Shouhua et al.[6] used a one-dimensional convolutional ResNet network based on a multilead two-dimensional structure to automatically identify a total of 34 heart rate types, including sinus rhythm. But the network structure parameters of their model can be further optimized. XIE Sheng Long et al.[7] proposed a one-dimensional convolutional neural network-based intelligent diagnosis method for atrial fibrillation, which achieved good diagnostic accuracy. WEI Xiaoling et al.[8] proposed an atrial fibrillation detection algorithm using a combination of fused multiple features and convolutional neural network. Through deep feature extraction and decision-level fusion of CNN, it solves the problems of obtaining discriminative features and improving the performance of the algorithm. This wider application to patient groups and stronger generalization ability. MA Zhiyi et al.[9] constructed the 12-lead ECG signal as a multi-channel viewable, mapped it to a complex network by inter-layer mutual information. They also introduced thresholding to construct an undirected weighted network, extracted two parameters of weight degree and weighted clustering coefficient, which made the joint multi-channel analysis more flexible and efficient. However, there are still limitations in identifying intrinsic features of single-dimensional time series. Liu et al.[10] proposed a neural network combining CNN and Bi-LSTM to classify myocardial infarction and normal ECG in PTB ECG database with 99.9% accuracy.

This paper proposes an automatic ECG classification method based on a bidirectional LSTM model, where the ECG signal is dichotomized (distinguishing between normal and atrial fibrillation (AFib) signals). And it is validated by the dataset used for the PhysioNet 2017 computational cardiology challenge.[11] By extending and optimizing the network architecture of the LSTM model to include a two-layer bidirectional LSTM network, as well as combining the time-frequency characteristics of ECGs, the advantages of recurrent neural networks are maximized. This approach not only shortens the training time, but also improves the accuracy of classification and achieves better classification results.

2. Principle

2.1. Electrocardiogram recognition

The activity of cardiomyocytes is conducted in the form of electrical signals, and their electrical activity can be depicted by the surface of the instrument connector and shown on the electrocardiogram. Related cardiovascular diseases can be diagnosed according to the regularity of ECG signals. The differences between normal ECG signals and atrial fibrillation (AFib) ECG signals are mainly as follows. First, the atrial fibrillation signal does not have P waves, but rather flat and irregular small F waves with a frequency of about 350-600 beats per minute. Second, patients with atrial fibrillation have a rapid heart rate and an extremely irregular ventricular rate. Third, due to the occurrence of excessive ventricular rate, the QRS cluster will widen and deform.

2.2. Bidirectional LSTM

2.2.1. Long short-term memory network model (LSTM)

Compared with general recurrent neural networks, LSTM has the advantages of solving long-term dependence problems, efficiently processing sequence problems, and having better long-term memory function. As shown in Figure 1, the core structure of LSTM consists of four parts: forget gate, input gate, update gate and output gate. [12]

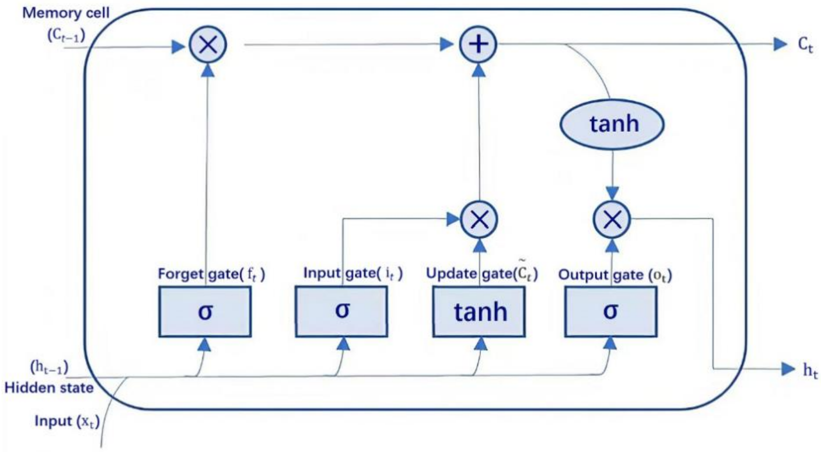

Figure 1: Structure diagram of LSTM

(1) Forget gate

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{1}$$

In Formula (1), W_f represents the weight of the forget gate, b_f represents the deviation of the forget gate, and f_t represents the forget gate.

The forget gate will selectively forget the incoming information and only retain the needed information. The forget gate value is calculated by h_{t-1} and x_t, which represents the degree of forget the input information and acts on the cell state of the previous layer. Then, b_f matrix is used to adjust it into the same dimension as the hidden layer at time t, and then a G bias is added to make the classification between 0 and 1 through Sigmoid function. When the value of a bit of f_t is 0, the information of the corresponding bit of C_{t-1} is completely forgotten. When the value is (0, 1), part of the information about the corresponding bit is preserved. Only when the value is 1, the corresponding information will be completely preserved.

(2) Update gate

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{2}$$

$$\tilde{C_t} = tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{3}$$

In Formula (2), W_i represents the weight of the update gate, and b_i represents the deviation of the update gate.

The update gate is divided into two parts, i_t is the retained new input information after filtering out the unnecessary information, and $\tilde{C_t}$ is the information brought by the new input. The tanh activation function is used to normalize the content to -1 to 1. Through this structure, how much information can be saved to the unit state of the current moment in the input of the network. Unlike classical RNN, LSTM obtains the current state.

(3) Cell status update

$$C_t = f_t * C_{t-1} + i_t * C_t \tag{4}$$

As shown in Formula (4), the result of the multiplication of the forgetting gate value and memory cell C_{t-1} obtained in the previous step is added to the result of the multiplication of the input gate value and the unupdated C_t processed by tanh function, so as to update the state. The result is input to the next node as the updated C_t as the new state.

(4) Output gate

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \tag{5}$$

$$h_t = o_t * tanh(C_t) \quad (6)$$

In Formula (5) and (6), W_o represents the weight of the output gate, b_o represents the deviation of the output gate, and o_t represents the output gate.

The output gate is used to control how much useful information can be stored in the total memory of the current unit state. This gate can be used to generate the hidden state h_t, which will act on the updated state C_t. After being activated by tanh, it will become part of the input of the node in the next time as the hidden state.

2.2.2. Bidirectional long short-term memory network model(Bi-LSTM)

Bi-LSTM model is an extension of LSTM model. Bi-LSTM neural network structure model is composed of two independent LSTM models, including forward LSTM and backward LSTM. The establishment of bidirectional sequence can not only improve the long-term dependence of LSTM model, but also continue the calculation advantages of LSTM model, so as to further improve the accuracy of identification. [13]

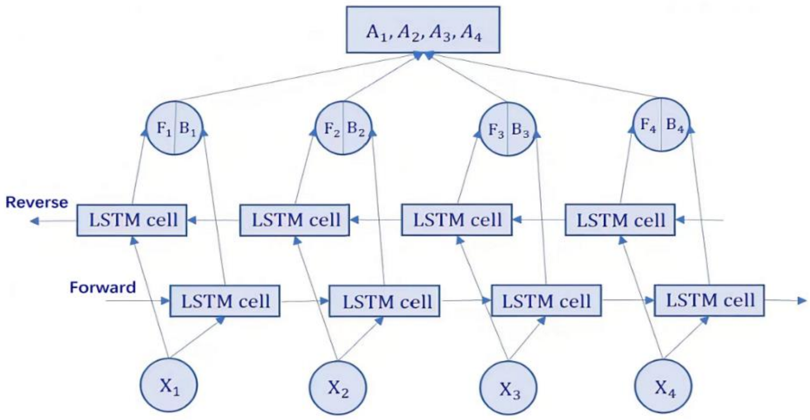

Figure 2: Schematic diagram of Bi-LSTM

As shown in Figure 2, BI-LSTM propagates the input quantities X_1, X_2, X_3 and X_4 in two LSTM models to obtain F_1, F_2, F_3 and F_4 . And back propagation gives you B_1, B_2, B_3, B_4. Finally, A_1, A_2, A_3 and A_4 are obtained by pairwise splicing of the information combination obtained before and after propagation.[14]

2.2.3. Signal processing and Bi- LSTM model

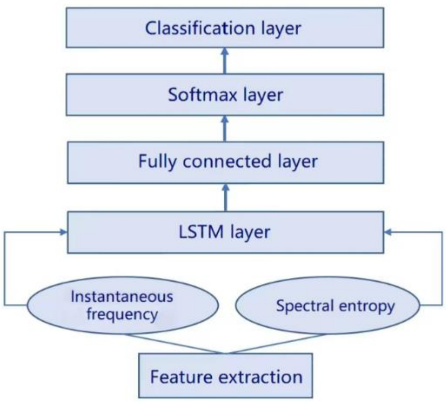

Figure 3: The process of implementing classification

Using the bidirectional LSTM model is shown in Figure 3. First, feature extraction is performed. This process can map high-dimensional raw data containing a lot of redundant information to low-dimensional feature space, so as to extract effective information, greatly improve training efficiency, and reduce computing time. Calculate the spectrogram of the normal signal and atrial fibrillation ECG signal, visualize the instantaneous frequency and spectral entropy and normalize the mean value, and input them as two one-dimensional features to the LSTM layer, respectively. After the forward and reverse LSTM training, the data obtained from the two features is input into the fully connected layer. The network nodes of the fully connected layer are complex and have the function of multi-feature fusion. The output results of the two features are synthesized and input to the softmax layer.

$$p = \frac{e^{v_i}}{\sum_{j=1}^{k} e^{v_j}} \tag{7}$$

The softmax function formula is shown in formula (7). k represents the number of multiple outputs or categories of the neural network, and v represents the output vector, which v_j is v the value of the first output or category in the middle j and i indicates the category that currently needs to be calculated. Processed by the softmax function, the calculation result is between 0 and 1, and the sum of all categories is 1, reflecting the probability of the signal being normal or atrial fibrillation signal. Calculated by the softmax layer, and finally the classification layer selects the category with a higher probability, and outputs the classification result.

3. Experiment procedure

3.1. Data preprocessing

In order to facilitate training, the signal of the existing data set is first processed. The label column is added to mark the existing data. The normal signal is marked as " N ", and the atrial fibrillation signal is marked as “A”. According to statistics, there are 738 AFib signals and 5050 normal signals in this data set. Comparing the two ECG signal visualization diagrams shown in Figure 4, it can be found that the interval of normal signals is basically the same, and the P wave and QRS wave are clear and regular. In contrast, atrial fibrillation signals are irregularly spaced and lack P waves.

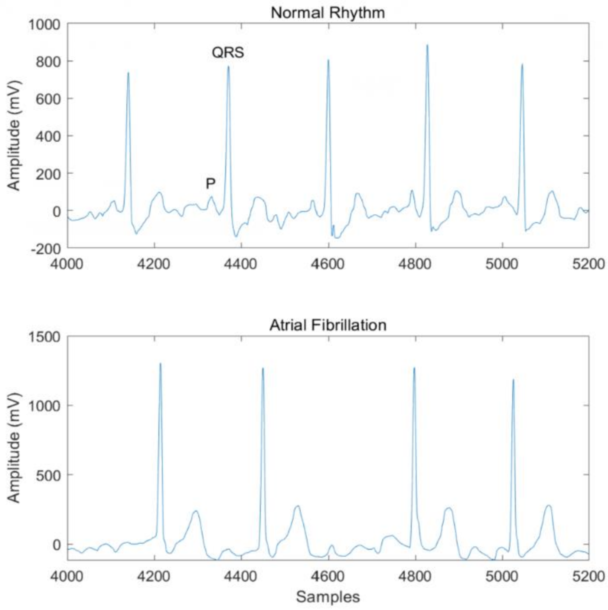

Figure 4: Visualization of the two signals

Plotting a histogram of signal lengths found that most of the signals were 9000 samples long. Use the segmentSignals function to ignore signals with fewer than 9000 samples, decompose signals with more than 9000 samples into as many 9000-sample segments as possible, and ignore the rest. At this time, all signals are 9 000 samples long, with a total of 718 AFib signals and 4937 normal signals, with a ratio of 1:7.

Then, use the DividerRand function to randomly divide into a training set and a test set, so that 646 AFib signals and 4443 normal signals are used for training, and 72 AFib signals and 494 normal signals are used for testing. In order to prevent the classification of the dataset from being unbalanced and cause the classifier to be biased, and to ensure that the two categories of the normal signal and the atrial fibrillation signal have the same number of signals, the oversampling method is used to replicate the atrial fibrillation signal. That is, the first 634 AFib signals for training and the first 70 AFib signals for testing were repeated seven times using the repmat function, respectively. Finally, 4 438 normal signals and AFib signals for training and 4 90 normal signals and AFib signals for testing were finally determined.

3.2. Feature extraction

A spectrogram shows the relationship between frequency and amplitude in a signal. The time domain function is calculated and transformed to generate the corresponding spectrograms of the two signals, as shown in Figure 5.

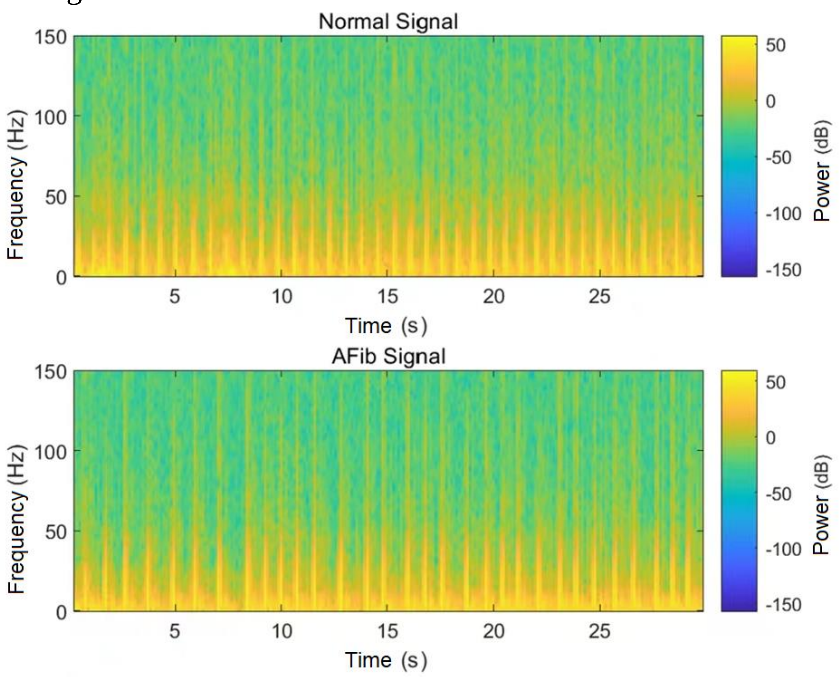

Figure 5: Spectrograms of the two signals

Time-frequency analysis is performed on the spectrogram, selecting two features of instantaneous frequency and spectral entropy. The instantaneous frequency is the reciprocal of the signal phase and reflects the rotational speed of the vector argument. In order to apply the features to the one-dimensional signal input to the LSTM model, the short-time Fourier transform method was used to visualize the instantaneous frequencies of the two signals using 225-time windows as shown in Figure 6.

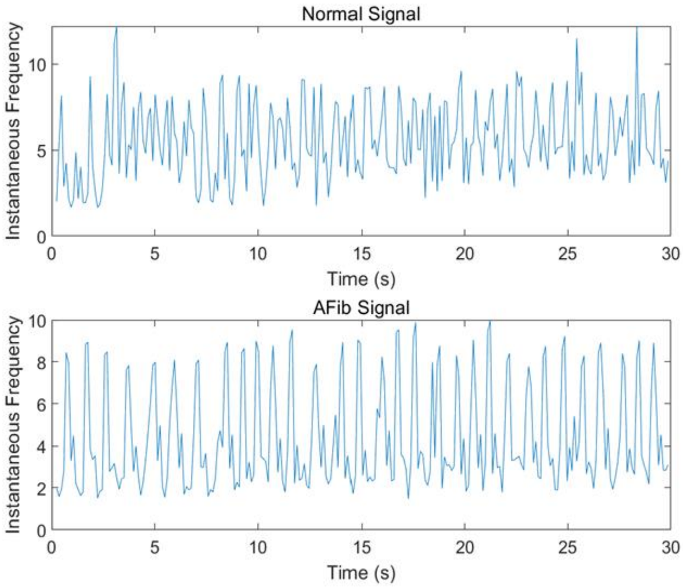

Figure 6: Instantaneous frequency diagram of two signals

The spectral entropy reflects the relationship between the power spectrum and the entropy rate. When the signal waveform is sharper, the spectral entropy value is smaller, and when the signal waveform is flatter, the spectral entropy value is larger. Similarly, 225 time Windows were used to process the data, and the spectral entropy of the two signal spectrograms was visualized as shown in Figure 7.

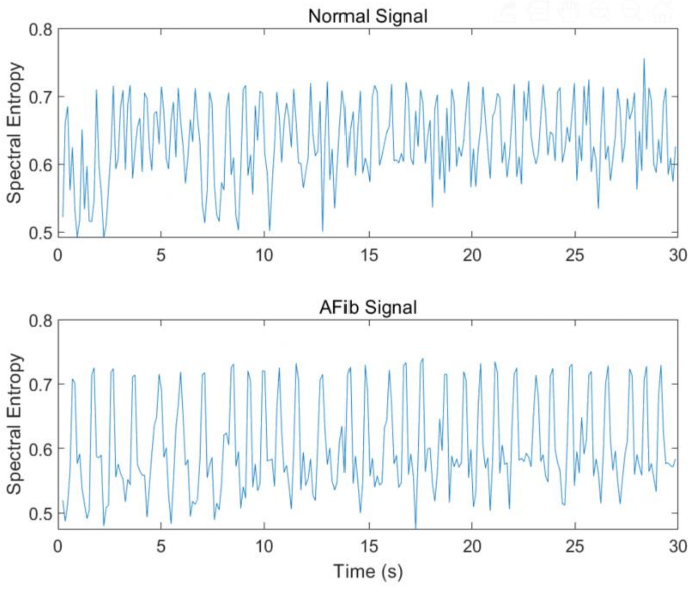

Figure 7: Spectral entropy diagram of the two signals

By connecting the two features, each unit in the new training set and test set is upgraded to two dimensions. At this time, each unit has changed from one signal with 9000 samples length to two signals with 255 samples length. However, the average value of the two features is very different, which may lead to the inability of LSTM to learn effectively, and the large input value may also reduce the learning speed and convergence speed. Therefore, it is necessary to standardize the data of training set and test set, adjust the mean value and standard deviation of instantaneous frequency and spectral entropy to the same level, and finally complete the data preparation.

3.3. LSTM model construction and experimental results

Each signal has two dimensions after feature extraction processing, so the network architecture input sequence size is specified as 2. At the same time, a bidirectional LSTM model was constructed, and the output size of the LSTM layer was set to 100, and the size of the fully connected layer was set to 2. Then, the softmax layer and two class classification layers were constructed. The training period was set as 10 rounds, 20 rounds and 30 rounds for the experiment. The experimental results under different training cycles obtained from the training progress chart and confusion matrix are shown in Table 1.

Table 1: Experimental results of different training cycles

Experimental result	Training cycle		
	10rounds	20rounds	30rounds
Training time	19 minutes 40 seconds	42 minutes 24 seconds	66 minutes 09 seconds
Total accuracy	90.1%	94.1%	80.9%
Positive accuracy	92.9%	94.3%	87.1%
Negative accuracy	87.3%	93.9%	74.7%
Final loss rate	25%	9%	47%

By analyzing the experimental results in Table 1, it can be found that the training duration increased from 10 to 30 rounds, from 19 minutes and 40 seconds to 66 minutes and 9 seconds. The total accuracy, positive accuracy and negative accuracy reached the highest when the training cycle was 20 rounds, reaching 91.1%, 94.3% and 93.9%, respectively. The loss rate of the training process was as low as 9% at 20 rounds and as high as 47% at 30 rounds. The loss rate reflects the loss of cross-entropy. With the success of training, it usually converges to 0. Therefore, the lower the loss rate at the end of training, the smoother the training is. When the training rounds are insufficient, the training of the model is not sufficient, and the training accuracy does not reach the maximum value that can be achieved by the model. In this case, the classification effect of the model is different from the best state. When there are too many training rounds, data overfitting means that the performance on the test set is far worse than that on the training set, and the generalization performance is poor. On the contrary, the accuracy rate will decrease and the loss rate will increase. Therefore, according to the comparison and analysis of the experimental results, the training period is selected as 20 rounds.

4. Conclusion

For a doctor at the time of arrhythmia disease diagnosed by ECG signal, it may appear misdiagnosis or missed diagnosis and the diagnosis efficiency is low. This paper proposes a model based on two-way LSTM combining feature extraction scheme. It can realize the normal signal and atrial fibrillation signal automatic binary classification, and will optimize the ECGsignal identification in the field of deep learning model with reference value. During the experiment, the instantaneous frequency and spectral entropy of ECG signals were extracted, and the features were merged to construct grids of different layers of the LSTM model. The optimal training period was adjusted to improve the accuracy and classification speed of the model step by step. At the same time, a two-way LSTM is introduced to add a reverse model on the basis of the forward model, which can not only input the past information, but also take into account the subsequent information, so that the generalization ability of the proposed algorithm is improved and the training time is greatly reduced. Through the validation of PhysioNet 2017 dataset, the classification accuracy of the model constructed in this paper can reach 94.1%, which is suitable for the detection of AF.

References

[1] Chen, Xie, Yuan, et al. Research on a Real-Time Monitoring Method for the Wear State of a Tool Based on a Convolutional Bidirectional LSTM Model[J]. Symmetry, 2019, 11(10):1233.

[2] Pahar M , Miranda I , Diacon A , et al. Automatic non-invasive Cough Detection based on Accelerometer and Audio Signals[J]. 2021.

[3] Qian Kun. Dynamic electrocardiogram in diagnosis of arrhythmia in patients with coronary heart disease Comparative analysis with conventional ECG [J]. Journal of Clinical Medicine in Practice, 2017, 21(09): 191-192.

[4] Zhang Aiai, ZHANG Pengxiang, WANG Xiaoyuan, LI Feixing, LI Fangjiang. Assessment on efficacy of remote monitoring with artificial intelligence electrocatdiogram in ECG critical value processing [J]. Chin J Evid Based Cardiovasc Med, 2022, 14(04): 431-433.

[5] Wang Guanjun, WU Ting, WANG Long, TANG Zusheng. Research of ECG diagnosis based on machine learning [J]. Journal of Practical Electrocardiology, 2020, 29(04): 262-268+297. DOI: 10.13308/j.issn.2095-9354.2020.04.007.

[6] Sun Bo, Yang Lei, Guo Xiumei et al. A Novel Method for ECG Signal Recognition Based on CNN and SVM Hybrid Mode [J]. Journal of Shandong Agricultural University (Natural Science Edition), 2020, 51(02): 283-288.

[7] Liu Shouhua, Wang Xiaosong, LIU Yu. Clinical Electrocardiogram Classification Algorithm Based on Deep Learning [J]. Computer and modernization, 2021(08): 52-57.

[8] Xie Shenglong, Zhang Weimin, LU Yujun Et Al. Research on Intelligent Diagnosis of Atrial Fibrillation Based on One-dimensional Convolution Neural Network [J]. Journal of Metrology, 2020, 41(05): 620-626.

[9] Wei Xiaoling, Liu Ming, Yuan Xin et al. Atrial fibrillation detection based on multi-feature fusion and convolution neural network [J]. Laser Journal, 2017, 38(05): 176-179. DOI: 10.14016/j.cnki.jgzz.2017.05.176.

[10] Ma Zhiyi, Yang Xiaodong, He Aidong et al. Complex network recognition of ECG signals in healthy and myocardial infarction patients based on multi view [J]. Acta Phys. Sin., 2022, 71(05): 48-57.

[11] Liu W, Wang F, Huang Q Et Al. MFB-CBRNN: A hybrid network for MI detection using 12-leadECGs [J]. IEEE journal of biomedical and health informatics, 2019, 24(2): 503-514

[12] Cheng Shi, Luo Xiaoning, Li Dongcheng, Yao Chenhong. A speech emotion recognition model based on BLSTM [J]. Changjiang information communication, 2022, 35(07): 19-22.

[13] Gao Jie, Yu Yali, Yu Zijie, Cao Zilong. Research on multiple exchange rate forecasting based on Bi LSTM and transfer learning [J]. Computer programming skills and maintenance, 2022(07): 39-41+45. DOI: 10.16184/j.cnki. comprg.2022.07.006.

[14] Fang Xiaofan, Xv Zeyu, Zhang Ying. Aeronautical Engine Life Prediction Based on Bi- LSTM [J]. Agricultural equipment & vehicle engineering, 2022, 60(07): 130-133.