Research on Intelligent Detection Technology for Bundle Branch Conduction Block

Hongwei Fang and Songyi Dian*
College of Electrical Engineering and Information Technology, Sichuan University, Chengdu, China
*scudiansy(at)mail.scu.edu.cn

Abstract. To help doctors diagnose heat disease effectively, an Intelligent detection method of bundle branch conduction block(BBB) heart rhythm is proposed in this paper. Firstly, continuous wavelet transform is used to preprocess experimental data obtained from MIT/BIH arrhythmia database, and then, some features of ECG signals are extracted from the time domain, frequency domain, and wavelet domain. After normalization, these features are fed into the BP neural network optimization algorithm based on genetic algorithm(GABPNN) to train the network. The training error, testing error, and F_1 score are evaluation criterions for classification in this paper. Compared with BP neural network algorithm(BPNN), GABPNN algorithm behaves better in accuracy. The accuracy of 99.502% was achieved, which is better than the results of some literature.

1. Introduction
Electrocardiogram (ECG) is critical for doctors to diagnose heart disease in patients. And a complete heart beat in ECG includes P, Q, R, S, T, U, and so on several parts, each part reflects the physiological information. Doctors can diagnose whether the patient is ill or what the disease is according to the abnormal information in each part. Doctors’ interpretation of ECG is often based on their experience, and their diagnosis rely on a description of some causes in ECG signal based on the standard of time, and such a standard is derived from the cardiovascular system of physiological, doctors can have a good understanding of this standard[1]. In order to help doctors to make quick judgment and improve the efficiency of patients’ treatment, this paper has done some research on the intelligent identification of BBB in abnormal ECG.

2. Data acquisition and preprocessing
ECG signals in this paper are obtained from the MIT-BIH arrhythmia database in physionet. In order to facilitate extraction of features, those signals should be well preprocessed.

2.1. Data acquisition
MIT-BIH arrhythmia database contains discontinuous numbers from 100 to 124 and from 200 to 234 of 48 group of ECG signals, that includes more than 10 types of ECG, the sampling frequency is 360 Hz, each set of data is obtained from two lead channels. In this paper, four kinds of ML II lead ECG signals are selected, the normal (N), left bundle branch block (L), right bundle branch block (R), and pacemaker heartbeat (P), pacemaker heartbeat exists as interference category. The four ECG signals
are labeled from 1 to 4 in turn, then the total number is divided by 30% into the training set and the rest as a test set. The obtained data is shown in Table 1.

| ECG Type | MIT-BIH data file | Training Set | Test Set | Total |
|----------|-------------------|--------------|----------|-------|
| N        | 100-101-103-112-115-117-122-205-213-219-234 | 218          | 94       | 312   |
| L        | 109-111-214       | 97           | 41       | 138   |
| R        | 118-124-212       | 71           | 31       | 102   |
| P        | 102-107-217       | 101          | 43       | 144   |
| Total    | 20 files          | 487          | 209      | 696   |

2.2. Preprocessing
There are some interference factors with raw data, power frequency interference, baseline drift, and myoelectricity interference. Studies have shown that the interpretation of ECG features in the computer-assisted manner has more than 90% accuracy in the underlying current and hemodynamic causes, so appropriate treatment can be provided to patients[1]. Signal processing technology is such a manner that can detect the small changes human cannot notice. In this paper, the original ECG signals is processed by wavelet analysis[2]. The Mexican hat function is similar to ECG signals, so it is chosen as the mother wavelet in the wavelet transform. And an attempt is made to obtain a better result when the transformation scale is 6.

This processing works well, before preprocessing, serious baseline drift and other interference can be seen in left image of Figure 1. But after preprocessing, ECG signal’s baseline alignment to the location of 0 on the longitudinal axis, and ECG becomes smoother.

![Figure 1. ECG signal preprocessing image](image)

3. Feature extraction
The features of ECG signals are extracted from time domain, frequency domain and wavelet domain, so that data characteristics can be detected from various aspects so as to accurately classify.

3.1. Feature extraction in time domain
In time domain, one of the most important features of ECG is the cardiac cycle, it is defined as the time span from one R wave to the next, the corresponding amplitude of R wave is another feature. Heart rate can be computed with cardiac cycle, it refers to the number of heart beats per minute in the
quiet state of a normal person, is usually 60 to 100 times per minute. It can be calculated as 
\[ \text{Heart rate} = \frac{60}{R_R}, \] 
where \( R_R \) is cardiac cycle.

Different types of ECG have a certain difference in the visual waveform, so some morphological features can be extracted according to their waveform. The main energy in a heartbeat cycle concentrated in QRS complex. Before extracting the morphological features, each heartbeat cycle is intercepted by capturing the 26 points before R wave and 44 points after R wave, such a total of 71 data points will represent a cardiac cycle. Then the standard deviation, the peak factor, the waveform factor, and the impulse factor of a heartbeat cycle are calculated as the morphological features of ECG.

3.2. Feature extraction in frequency domain

ECG morphological features and time-domain features associated with interphase cannot reflect the characteristics of the frequency domain, while the information in frequency domain can sometimes reflect the differences in signals that cannot be expressed in the time domain. In order to extract ECG features in frequency domain, the FFT transform of ECG signal is necessary. In frequency domain, the relation between frequency and amplitude can be obtained. The QRS waves in the frequency domain are located in the high frequency region, while P wave and T wave are located in the low frequency region. Normal ECG and other abnormal ECG have great difference in the frequency components. In this paper, the spectral mean, standard deviation, spectrum deviation and frequency kurtosis are mainly four features selected in the frequency amplitude spectrum[3].

3.3. Feature extraction in wavelet domain

In the time-frequency analysis of ECG signal, the one-dimensional signal can be expressed as a 2-d joint function of time and frequency, and the spectral composition of each moment can be described and analysed in the time frequency, which can easily reflect the change rule of features in frequency domain over time. In this paper, the wavelet packet decomposition method is used to analyse ECG signal, and the wavelet domain features are extracted. The wavelet packet can analyse the ECG signal in fine detail. It divides the frequency band into multiple layers, and then decomposes the high frequency parts again that are not decomposed in the general wavelet analysis. After a series of experiments, this paper selects the "Shannon" entropy type, takes "db6" as the mother wavelet, the ECG signal is decomposed by 4 layers decomposition, and the fourth layer wavelet packet coefficient is used to extract the features. Singular value, standard deviation, maximum value[4] of the fourth layer of the wavelet packet coefficient are calculated as three kinds of feature, and due to there are 16 band in the fourth layer, a total of 48 features can be obtained.

3.4. Feature processing and analysis

A total of 59 features are obtained in those three domains. If the data set is directly fed to the learning algorithm, the weight value will oscillate in the process of convergence, which is because the orders of magnitude vary greatly from feature to feature, and it is easy to converge to the local optimal result, thus resulting in poor classification results. To avoid such a result, the features should be scaled and normalized. Suppose that the acquired feature values are \( Feature_i \), where \( i = 1,2, \ldots, M \), \( M \) is the total number of features, and the processing process is as follows:

\[
\begin{align*}
\mu_i &= \frac{\sum_{i=1}^{M} Feature_i}{M} \\
S_i &= \text{Max}(Feature_i) - \text{Min}(Feature_i) \\
Feature_{Normal_i} &= \frac{Feature_i - \mu_i}{S_i}
\end{align*}
\] (1)

Where \( Feature_{Normal_i} \) is the processed \( Feature_i \).

After being processed, each sample of the four ECG signals is taken at random, and as is shown in Figure 2, 59 features in the 4 samples are obviously different in some places, and 4 samples can be easily distinguished from each other. Therefore, the features extracted in this paper are reasonable.
4. BP neural network optimization algorithm based on genetic algorithm

BP neural network (BPNN) algorithm is a multi-layer feedforward network based on error inverse propagation algorithm, which is one of the most widely used neural network models. BPNN has the ability to learn and store a large number of input-output mode mapping relationships without having to disclose the mathematical equations describing the mapping relationship in advance. Its learning rule is to use the gradient descent method to continuously adjust the weights and thresholds of the network through the reverse propagation, so as to minimize the error sum of the network.

In the initial iteration of BP neural network algorithm, the weight and threshold value of each layer are randomly generated, and the weight and deviation are constantly updated in the direction of decreasing the cost function gradient in the subsequent iteration process. But there exists a problem, the updation of weight and threshold may be trapped in local optimum. If the weight and threshold of the first generation are optimized so that the error at this time is minimized, then the weight and threshold values will be updated in a better direction, and then a smaller error and better classification results will be obtained. Genetic algorithm is chosen to optimize the initial rights and threshold of BP neural networks. The optimized BP neural network is a better predictor, and the flowchart of GABPNN algorithm is shown in Figure 3.

![Flowchart of GABPNN algorithm](image)

**Figure 3.** Flowchart of GABPNN algorithm
5. Simulation results
The BP neural network adapted in this paper with 59 neurons in the input layer, 8 neurons in hidden layer, 1 neuron in output layer. The training error, test error, and F_1 score are the evaluation criteria of algorithm. These two errors are calculated as two-norm, the smaller test error and train error, the better the prediction of the algorithm and the stronger the generalization. F_1 score is defined as the weighted average of precision and recall. The higher it is, the more effective the method is, and the performance of the algorithm is better.

In GABPNN algorithm, the number of individuals in a population is 40, the largest genetic algebra is 8, in the process of every genetic, each individual in the population as the weights and thresholds are fed into BPNN to train the network and calculate the test error as a fitness value.

Compared to other works in the field, the work in this paper shows better results, which is shown in Table 2.

Table 2. Performance of several methods for ECG arrhythmias classification in comparison with the proposed method

| Method                                      | Arrhythmia types | Training error | Test error | F_1 score | Accuracy % |
|---------------------------------------------|------------------|----------------|------------|-----------|------------|
| Wavelet/neural network[5]                   | 4 types of Arrhythmia | -              | -          | -         | 97         |
| DFT, DCT and wavelet / neural network[6]    | 4 types of Arrhythmia | -              | -          | -         | 95         |
| Maximum margin clustering/ Immun evasionary | 5 types of Arrhythmia | -              | -          | 0.9245    | 95.9       |
| algorithm[7]                                |                  |                |            |           |            |
| PCA and SVM[8]                              | 4 types of Arrhythmia | -              | -          | -         | 99.17      |
| BPNN in this paper                          | 4 types of Arrhythmia | 2.119          | 2.1959     | 0.9755    | 98.01      |
| GABPNN in this paper                        | 4 types of Arrhythmia | 2.2674         | 1.4953     | 0.9956    | 99.502     |

6. Conclusion and prospect
As shown in Table 2, based on the features extracted from the method herein, a better result is achieved in the BPNN algorithm and the GABPNN algorithm on the recognition of the BBB rhythm. Compared with BPNN, GABPNN optimizes the weight and threshold of the initial training neural network, so that it can get a better network, thus increasing the correct recognition rate of ECG signals. This algorithm recognizes that the improvement of accuracy is at the cost of training time. In the GA part, every individual in the population will be used as weight and threshold to be trained in neural network during each genetic process, so the algorithm time is greatly increased.

This paper is the author's first exploration of the ECG signal recognition, and the way that the ECG signal features are extracted is not likely to have a good effect on all the ECG signals, and it's also possible that there are some features that are not found more easily to recognize. In future work, other cardiac arrhythmical features will be tried, more research will also be done on more categories of ECG signal intelligence.

Acknowledgments
This work is supported by Fundamental Research Special Funds for Central Universities (20826041A4133).
References

[1] Najarian K, Splinter R. Biomedical Signal and Image Processing. [M]. Boca Raton: CRC press, 2012: 190-191.

[2] Zhang F D. Wavelet Analysis in Matlab. [M]. Beijing: Mechanical industry press, 2012: 53-58.

[3] Chen Y. Study on Abnormal Heart Rate Classification Algorithm for ECG Signals. [D]. College of Mechanical Engineering of Chongqing University, 2016: 38-40.

[4] Yuan D Y. Study on The Classification of ECG Signals Based on Wavelet Packet and Neural Network. [D]. Tianjin Polytechnic University, 2017: 13-19.

[5] Kianisarkaleh A, Fereshteh P, Ali K, et al. ECG Arrhythmias Detection Using a New Intelligent System Based on Neural Networks and Wavelet Transform. [J]. 2012, 9: 33-39.

[6] Belkheiri M, Douidi Z, Belkheiri A. ECG Beats Extraction and Classification Using Radial Basis Function Neural Networks. [M]. Proceedings of the Fourth International Conference on Signal and Image Processing 2012 (ICSIP 2012). Springer India, 2013: 127-136.

[7] Zhu B, Ding Y, Hao K. A novel automatic detection system for ECG arrhythmias using maximum margin clustering with immune evolutionary algorithm. [J]. Computational and Mathematical Methods in Medicine, 2013, (2013-4-18), 2013, 2013(2): 453402.

[8] Zhang H, Zhang L Q. ECG analysis based on PCA and Support Vector Machines. [C]. International Conference on Neural Networks and Brain, 2005. Icnn&b. IEEE, 2005: 743-747.