Research on the Diagnosis Algorithm of ECG

Yamei Zhang

Xi'an Peihua Universty

Xi'an, China, 710125

e-mail: 312413420@qq.com

*Abstract—**In this paper, through the research of previous research on ECG diagnosis process, preprocessing method proposed multi order classification algorithm of BP neural network and De-noising and sampling frequency conversion, combined with the waveform feature extraction method, the recognition rate is obviously improved, the diagnosis is more accurate.***

Keywords-ECG; BPNN; Multi-level Classification; Algorithm

I. INTRODUCTION

ECG from the invention, the application has been one hundred years of history. ECGdue to its reliable diagnosis, the method is simple, and the advantages of not damaging the patient, always plays an important role in the process of diagnosis, still can't be replaced.ECG is the use of ECG acquisition instrument record the change of the human body heart potential, and applied to clinical heart disease monitoring, diagnosis, visible graphic record (al Fig1), it is an important means of common heart disease diagnosis. By summarizing and analyzing the previous studies, it can be found that the ECG signal processing algorithm is very rich, but few involve the neural network being applied in the field of electrocardiogram, especially the BP neural network. In this paper, the BP neural network, multistage classification algorithms, based on the full research of ECG diagnosis process proposed De-noising and sampling frequency conversion in advance Methods. Experiments show that the proposed pretreatment scheme has good effect.
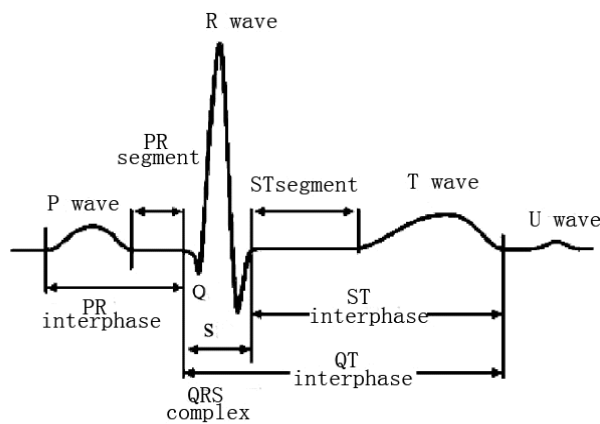

Figure 1. The illustration of a heart beat

II. RESEARCH PROGRESS AND ACHIEVEMENTS OF ECG AT HOME AND ABROAD

Previous studies on ECG were mainly focused on three parts: early ECG signal acquisition and processing, mid-term feature extraction and later classification identification. Good ECGdata obtained is the precondition of classification, and the appropriate feature extraction method is the basis of guarantee the high recognition rate, under the condition of the two to choose good classification method is likely to get a good effect, so the predecessors of the three links done a lot of work. Through the research of ECG diagnosis research papers and found predecessors in the aspect of signal preprocessing mainly use the traditional filter, wavelet transform method, the difference threshold algorithm, mainly including linear classifier in ECG classification identification, artificial neural network and support vector machine (SVM), Bayesian classification and so on.

III. STUDY BACKGROUND AND PRACTICAL SIGNIFICANCE

ECG in the diagnosis of heart disease plays an important role in the process, but because of its complex monitoring equipment, has been unable to applied in daily life. With the development of science and technology, smart phones and various Wearable devices are emerging, which provides great convenience for the acquisition of ECG. People no longer need to go to the hospital to use sophisticated equipment to get the ECG signal is an important disease diagnosis index, now only by a small smart devices can be realized, which greatly saves time and money. The ECG signal can also be processed and analyzed directly to reflect the overall condition of the individual body.

IV. PRETREATMENT OF ECG SIGNALS

The diagnosis first to collect ECG signal and ECG in the process of acquisition, it inevitably doping many kinds of noise, on the premise of guarantee the quality of the signal, how to get rid of these noise is the key point of research. For the electrocardiogram of different noises, the influence of the diagnosis accuracy is very large. Therefore, choose a perfect method of de-noising. According to the study, there are three main types of noise: Frequency interference (50 Hz and its frequency), baseline drift (below 0.7 Hz) and myoelectric interference, therefore, the three kinds of ECG noise need to be filtered.

A. Frequency interference filtration

The frequency interference is an inevitable disturbance in the electrocardiogram. It is caused by the alternating

current of 50 Hz, which is caused by many harmonics, which is shown as some small burrs in the electrocardiogram. Filters can be removed for such interference, such as many traditional filters or adaptive filters. While the traditional digital filter technology is relatively mature and simple, and the noise effect is very good, so digital filter is a good choice. There are a lot of frequency-frequency classical filters, such as a smooth filter and a narrow band trap with fixed center frequency. For these cases, a Butterworth band resistance filter is used to remove the interference. Set the cut-off frequency of the resistance band to 49, and the upper limit frequency of the resistance band is set to 51, so that the frequency noise of the 50 Hz and its vicinity can be filtered out. The principle of Butterworth Formula is (1) .

$$|\mathrm{H}(\omega)|^2 = \frac{1}{1+ \left(\frac{\omega}{\omega_c}\right)^{2n}} \qquad (1)$$

n——The order number of the filter;

ω_c——Represents the cut-off frequency.

B. Myoelectric interference filtration

In the ECG acquisition process, P wave, T wave and U wave, it is easily affected by the electromyography interference, and many kinds of disease diagnosis depend on P wave and T wave. Main method with adaptive filtering electromyography interference filter and wavelet filter method, the method based on wavelet transform is mainly through MATLAB db5 wavelet ECG signal wavelet transform, then the wavelet coefficients of different scales. Here is the use of 5 wavelet transform. Then, the approximate wavelet coefficients of scale 3 are extracted, and the threshold values of each order are selected by using stein's unbiased likelihood estimation principle for the three-order detail wavelet coefficients. Finally, the electrocardiogram can be reconstructed according to the following wavelet coefficients, which can not only remove the noise caused by myoelectric interference, but also remove some of the baseline drift noise.

C. Baseline drift correction

The baseline drift is caused by slight movement of the human body, contact with electrodes, breathing etc., and its frequency is less than 1 Hz, and in the heart electrical signal is a superimposed low-frequency noise. In the ECG signal cardiac cycle, the end of the QRS composite wave is the dipole state of the electric dipole, and there is no potential change in normal conditions. The frequency band of baseline drift is 0.05 Hz~ 1.1.5 Hz, while the real signal and baseline drift noise are mostly coincident, which will inevitably have a significant impact on the real signal. For a higher accuracy of diagnosis of disease, must remove the baseline drift current noise filtering method and the baseline fitting method is the most commonly used two kinds of baseline drift suppression method, because at the time of electromyography interference filter has been removed part of the noise and baseline drift here use of moving window median filter/line drift to correction. The specific approach is to take some simple treatment of the original signal first, and

then use the median filter method to fit the baseline drift function. Finally, the actual signal function is used to subtract the proposed baseline drift function, so that the actual signal can be obtained.

D. Filter and its effect contrast

Through a large number of experimental comparison, the above set of more effective ECG pretreatment method is obtained, namely, the Butterworth filter is used to filter the interference. Then through wavelet decomposition for ECG signal decomposition, in view of the noise in the wavelet decomposed layers, and have little impact on the real ECG signal wavelet coefficient, then reconstruct signal and get new ECG signal, so that we can remove noise part ECG interference and baseline drift. Since the baseline drift has a great influence on the later feature extraction, especially for the extraction of waveform characteristics, the baseline drift is corrected by the value filtering method in the mobile window. By means of the above method, it can obtain the stable and reliable electrocardiogram signal and provide the data base for the later operation.

V. ECG DIAGNOSIS BASED ON BPNN MULTI-ORDER CLASSIFICATION ALGORITHM

Artificial neural Networks (ANN) is a simple network model built through abstract brain neurons and used to deal with complex practical problems. BP neural network is especially widely used in various fields in the diagnosis process of ECG is also a common method. The BP neural network model consists of four parts: input and output model, Action function model, error calculation model and self-learning model. The parts complement each other and work together to complete the whole training process.

A. Multi-order classification algorithm based on BPNN

The traditional BP neural network, when the sample size reaches a certain extent, the discrimination of the output target vectors in the identification process will be reduced, leading to the misjudgment. For ECG classification recognition research is need a lot of samples to train and test, therefore, on the basis of BP neural network proposed a multi-level classification algorithm based on BPNN.

In this algorithm, firstly the sample data were randomly divided into multiple data sets: a training set, multiple false test set and a test set, random purpose is to ensure that within the sample set of independent there is no different from other collections of data. After the establishment of multiple data sets, the training set was classified by the BPNN classifier, resulting in a network 1, which is our first-order network. Then use 1 false test set for simulation network, the network will output a result vector. In the result vector, some of the results are very distinct, but there are a lot of differences that are not obvious, so we set a threshold here, so that the fake test set is divided into two parts, and the difference between the obvious samples and the non-obvious samples. Set$\{Y_1, Y_2 Y_3 \dots Y_n\}$For the output vector set, the classification function is shown as Formula (2):

$$Class = \begin{cases} 1, & Y_i \geq thre \\ x, & Y_i \leq thre \end{cases} \qquad (2)$$

Type in the thre——The threshold value

Depending on the formula above, you can get Class =x Output vector of the sample, the sample as a second order network training set, through the BPNN classifier for the second order training set to train a network 2. Repeat the process over and over, and get the training network of different order Numbers. The classification identification of the test set first test set 1 through the network simulation, it get a set of output vector, by threshold function can get the two parts as a result, have been identified is a part of the sample, the other part is not classified samples. The first part is the first order output, and then the second part gets the output vector set, so that the second order classification results are obtained, and then the multi-order classification results can be obtained by analogy.

B. Multi-order classification algorithm

It experiment adopts an internationally recognized database of MIT-BIH arrhythmia, which has a rich sample of abnormal ECGand is annotated with each electrocardiogram. According to the annotation, and in the premise of guaranteeing the number of samples, classification of 5 types of heart beats, including: Normal (N), left bundle branch block (L), right bundle branch block (R), room premature beat (A) and ventricular premature beat (V), it including all kinds of cardiac number and number of samples taken from the data in Table 1:

TABLE I. SAMPLING CARDIAC TYPE AND QUANTITY

Cardiac type	Record number	Numbers of samples
Normal (N)	100, 101, 103, 115	900
left bundle branch block(L)	109, 111, 214	900
right bundle branch block(R)	118, 124, 212, 231900	
Ventricular premature beat(V)	119, 203, 205, 223 900	
room premature beat(A)	207, 209, 232	900

Will these five kinds of heart beat is divided into three groups, from 5 400 samples from each type of cardiac form a training set containing 2000 samples, 400 samples from each of a false test set containing 2000 samples, the remaining 500 samples as test set. According to the above methods, pretreatment and feature extraction are carried out, and multi-order classification algorithm is applied to identify. The recognition rate of experimental results is shown in table 2,N stands for normal heart beat, L is the left bundle branch block, R is the right bundle branch block, A is the room sex early beat, V is the ventricular premature beat, the Ave represents the average recognition rate:

TABLE II. MULTI-ORDER CLASSIFICATION ALGORITHM BASED ON WAVEFORM CHARACTERISTICS

N	L	R	V	A	Ave
1.000	0.985	0.980	0.930	0.945	0.968

VI. RELATED FEATURE EXTRACTION ALGORITHM

The three feature extraction methods are: small baud algorithm, improved RBP algorithm and waveform feature.

A. Small baud levy algorithm

The specific method of small baud levy is to detect the QRS complex wave group of the original signal first, then determine the position of the R point, then extract the first 43 points of R wave and then 84 points after the detection of the R spot sequence, plus the R point to form a signal containing 128 points. So take 4 R spread to the nearby signals, these four signals constitute a 512, 128, if the original signal in the R wave number less than 4, is replaced with the first. The specific practice is shown in Fig.2 below:

a) 128 point signal

b) 512 point combination signal

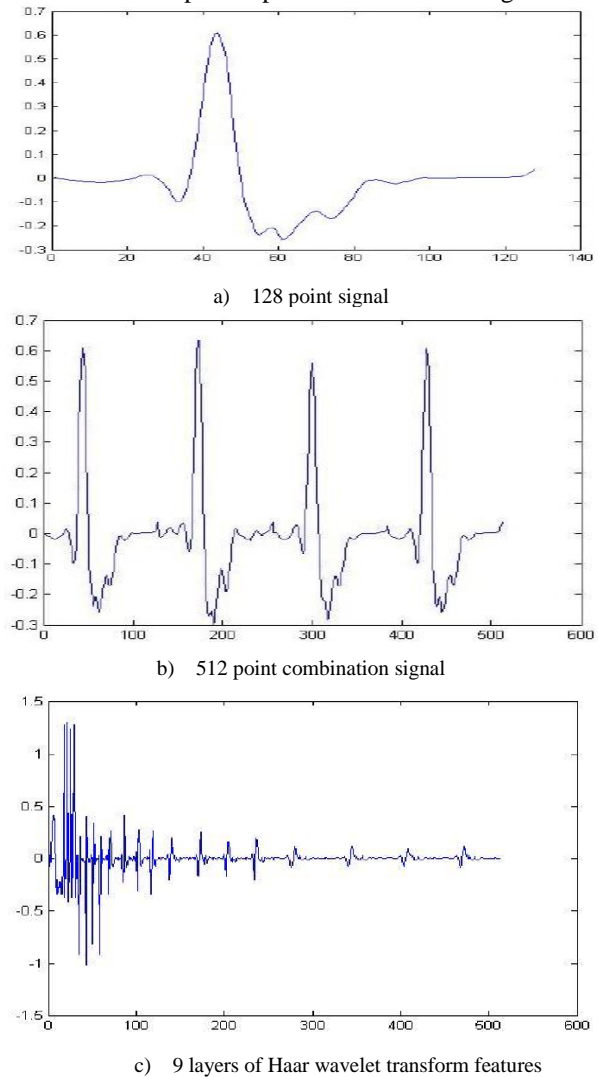c) 9 layers of Haar wavelet transform features

Figure 2. Wavelet transform feature extraction process

An original signals by the above method to extract, Fig2a) shows a 128-dimensional signal being intercepted. The four extracted signals are then merged into a 512 signal, which represents the original signal, as shown in Fig2b). Using Haar wavelet transform method 9 layers wavelet decomposition, the wavelet coefficients in each layer

combined into a 512 characteristics of wavelet coefficient, As shown in Fig2c), then classify and identify.

B. The improved RBP algorithm

All of the ECG signal is smooth, it can be as a vector. To each pair of consecutive input data do some comparison and classification, there are increased or reduced in both cases, the growth of 1, will reduce down to 0.For example, there is a string of input cardio data $\{x_1, x_2, \ldots x_i\}$, In this x_i represents the it input strength value of the signal, and the specific transformation is shown in Formula (3)

$$y_i = \begin{cases} 1, & x_i < x_{i+1} \\ 0, & x_i \geq x_{i+1} \end{cases} \quad (3)$$

ECG will be converted from the original data vector to a 2-bit sequence with only 0 and 1. Then the whole length of 2 m a bit sequence as the window of a size of m, and will be m a window into decimal number, each completed a transfer window move back one continue to convert, until the remaining m bit is converted to a decimal number finally, get the integer is also called the m - bit word.

Sometimes the basic algorithm of RBP in disease diagnosis system can't obtain good effect, because of the above algorithm in the generated 2 bits of the 0 s and 1 s, compare the two adjacent sampling point data values, this approach has limitations, easily affected by the interference of noise or local mutation. Here you can choose to increase the sampling room To modify the algorithm, the algorithm will be more flexible. Although noise interference and local mutation can't be completely avoided, the risk can be greatly reduced. The formula (3)above will become formula

$$5y_i = \begin{cases} 1, & x_i < x_{i+interval} \\ 0, & x_i \geq x_{i+interval} \end{cases} \quad (3)$$

interval——Sampling interval.

The change of sampling interval not only reduces the above two risks, but also decreases the dimension of the feature set as the interval increases, thus reducing the computational complexity.

C. Waveform character

For waveform feature extraction methods, we need to detect first before feature extraction, Q, R, S, Tposition of each of the waveform points, Whether it is correct or not is the key to the formation of characteristic properties. Adopted here is based on wavelet transform binary spline 4 layer, the layer 3 wavelet detail coefficients of maximum and minimum detection of R wave. Choose 3 layer wavelet detail coefficients to detect R wave, because three layer wavelet transformation R wave and other noise difference is the largest. In order to get the exact R wave position, we also

include the right R wave adjacent to the two QRS waves when we take the data, so that we get the P,Q,R,S,T position information, which is a sequence of position, We choose the middle of a value as "rumet" point, determine which in turn determine the Q and S point after "rumet" point, these two points is the most close to "rumet" point, one on the left side of the "rumet" point, the other one on the right side of the "rumet" point. And then we have P and T, which are closest to Q and S. You can get the exact P,Q,R,S,T position through the above steps，Then according to the location information, we calculate the four types of waveform character attributes, a total of 19, they including distance (RQ, RS ,RP, RT, PQ, TS, QS), the amplitude (amp RQ, amp RS, amp PQ, amp ST, amp PT), the slope (slope RQ, slope RS, slope RP, slope RT, slope PQ, slope TS) and the area (area QRS).

VII. CONCLUSIONS

By comparison with theoretical research, the simulation results and experimental results, found that the multistage classification algorithm based on BP neural network, the denoising and the pretreatment of the sampling frequency conversion method, combining with waveform feature extraction methods, ECG classification recognition rate is improved obviously, greatly improve the accuracy of ECG diagnosis.

ACKNOWLEDGMENT

REFERENCES

[1] Maeda M, Suenaga M and Miyajima H. Qubit neuron according to quantum circuit for XOR problem [J]. Applied Mathematics and Computation, 2007, 185(2): 1015-1025.

[2] Noh Y H, Hwang G H, Jeong D U. Implementation of Real- time Abnormal ECG Detection Algorithm for Wearable Healthcare [C] // 6th International Conference on Computer Sciences and Convergence Information Technology, IEEE, 2011, 136(5): 111-114.

[3] Lee Y G, Jeong W S, Yoon G. Smartphone-based Mobile Health Monitoring [J]. Telemedicine Journal and E- health: the Official Journal of the American Telemedicine Association, 2012, 18(8): 585-590.

[4] Gholam - Hosseini H, Nazeran H, Reynolds K J. ECG Noise C ancellation Using Digital Filters [C] // 2nd International Conference on Bioelectromagnetism, IEEE, 1998: 151-152.

[5] Lagaris IE, Likas A and Fotiadis D I. Artificial neural network methods in quantum mechanics [J]. Computer Physics Communications, 1997, 114 (1-3): 1-14.