http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 12 (18): 4549-4552, 2013
ISSN 1812-5638 / DOI: 10.3923/itj.2013.4549.4552

Arteriosclerosis Diagnosis Based on Probabilistic Neural Network

Qi Hui
Qingdao University, Qingdao, Shandong 266071, China

Abstract: The waveform characteristics of pulse signal can be used to diagnose early atherosclerosis. Based on the research achievement of predecessors, eight feature parameters including both time-domain and frequency-domain related to arteriosclerosis are chosen for research so that we can analyze the pulse signal in many aspects. Using probabilistic neural network to train model of arteriosclerosis, spread constant of neural network is optimized through the adaptive genetic algorithm. Through the simulation experiment, it shows that probability neural network can predict the arteriosclerosis well. A feasible method which can achieve the classification of arteriosclerosis is provided.

Key words: Pulse wave, arteriosclerosis, neural network, genetic algorithm

INTRODUCTION

Pulse wave is the beating of the heart propagation which flows to the peripheral along the arterial blood vessels and blood. Many studies have found that pulse wave propagation is related to numbers of physiological parameter of cardiovascular system closely (Safar *et al*., 2003; Matthys *et al*., 2007; Alastruey *et al*., 2008).

Atherosclerosis is the main cause of various cardiovascular diseases. There will be a great significance if we can detect and analyze the hemodynamic parameter during diagnose before the onset. Existing research and clinical statistics show that the change of cardiovascular in blood flow parameters reacts in the change of pulse wave waveform firstly. Therefore, according to the normal and abnormal conditions of pulse wave comparative judgment, we can diagnose the potential risk of cardiovascular disease as soon as possible. In this paper, we extracted the time-frequency domain features in the pulse signal analysis. On this basis, the introduction of probabilistic neural network to predict atherosclerosis can achieve very good results.

PROBABILISTIC NEURAL NETWORK

The mathematical model of the probabilistic neural network model is shown in Fig. 1.

where, R is the dimensions of the input vector, Q is the number of overlay layer Radial Basis Function (RBF) neurons and the training samples. K is the number of input data which can be divided into categories. Operator frame ||dist|| is used to calculate the Euclidean distance between the input vector and weight vector and operator *indicates the threshold vector b1 and ||dist|| output vector corresponding elements multiplication, among them:

$$a_i 1 = \text{radbas}(\|_i IW - P\| b_i 1)$$

$$a2 = \text{compet}(L W_{2,1} a1)$$

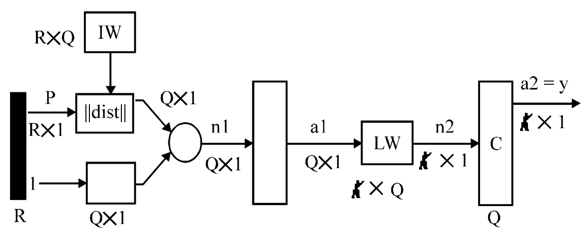

Fig. 1: Model of PNN

It should be noted that the threshold of radial basis function can adjust the sensitivity of function. However, from a practical engineering point of view, another parameter is used more widely, which is called spread. There are many ways to obtain the congruent relationship between b1 and spread in practical application (Ahmadlou and Adeli, 2010). The value of spread reflects the response range of the output to input. The greater the value, the larger the coverage area. At the same time, the smoothness between neurons is good and the response range of the hidden layer neurons for input is big. The smaller the value, the less the covered area, the sharper the curve and the cross coverage between neighbor classes can be reduced.

GENETIC ALGORITHM

Genetic Algorithm (GA) is a kind of search algorithm based on the natural selection and natural genetic mechanism. It simulates the mechanism of the evolution of creatures and realizes the optimization of specific objectives in the artificial system. GA is mostly used to solve the combinatorial optimization problems and the nonlinear optimization problems which have the non-differentiable functions or the complicated constraint conditions.

Fitness function: The standard error of convergence for classification problems is that there are just a bit errors in the classification of the training sample, then the network structure can be regarded as correct. So the error can be used to the objective function:

$$e = \left\| D - \text{compet}(HW') \right\|_F$$

where, D is the desired output and H is the hidden layer output matrix, W is the weight matrix which connect between the hidden layer and the output layer. $\|\bullet\|_F$ is Frobenius norm. The objective function is the smallest problem, so we choose the fitness function as follows:

$$\text{Fit}(x) = \begin{cases} c_{max} - e & e < c_{max} \\ 0 & \text{else} \end{cases}$$

where, c_{max} is the biggest estimate of e.

EXPERIMENT

Data acquisition: Pulse signal is collected from multi-channel physiological signal acquiring system based on RM6240. The maximum collection frequency is 800 Hz. Scanning speed is 250 ms/div. Sensitivity is 25 mV. Filtering parameters is 10 Hz.

The reason the age of subjects we choose is above 20 is that there is almost non-existent possibilities for the children and adolescents suffering from arteriosclerosis. One out of all 83 subjects is checked out that there is a great disturbance of the pulse signal which feature points can not be found with severe waveform distortion. So we remove this data. We use filter to extract the signal feature parameter.

Pretreatment: Predecessors have already done a lot of research (Rogers *et al*., 2000; Luo *et al*., 1996; Jiao and Fang, 2000) on the relationship between atherosclerosis and specific characteristics parameter of the pulse signal. It is impossible to be used to analyze entirely, which can only select mature parameter with significant and relatively to study. Among them, age of the subjects, blood pressure (mainly systolic blood pressure), pulse signal H1 = (h2-h1)/h1 (indicates the ratio of the front wave and the main wave of the gravity waves), T (indicates oscillation frequency of the radial artery within left ventricular ejection time), M1, M2, M3, M4 (indicate fourth harmonic amplitude) are significantly correlated with atherosclerosis. We can get a more comprehensive analysis of the pulse signal from many aspects with both time domain parameters and frequency domain parameters. Take three preparative schemes for eight characteristic parameter as the input of the PNN network:

Project 1: The original data without any pretreatment.

Project 2: The original data is normalized to a mean of 0 and standard deviation 1.

Project 3: The original data is normalized to [-1-1].

The output of the network is whether atherosclerosis or not.

Data packet: Due to the serial number of the subjects are random, 82 sets of data will be obtained, taking the first 54 subjects data as the training set, the last 28 subjects pulse data as a test set to verify the effectiveness of the network.

Genetic optimization spread constant: Spread constant is selected for the optimal parameter, ranges (0-10), using binary coding, accuracy of 0.01 and the chromosome length is chosen as 10. The population size is chosen as 20. The crossover rate and mutation rate can be selected by using the adaptive law (Duan and He, 1998).

Experimental results and analysis: Classification results are shown in Table 1.

The Table 1 shows that the second kind of pretreatment solution is not ideal, when the first solution spread constant is 0.05~0.6 and the third solution spread constant is 0.02~0.5. The accuracy of both the training set and the test set can reach 100%. When spread constant is

Table 1: The relation of misclassification rate and spread

Preparative regimen	Spread	Training set	Test set
3	0.6	1	0
	0.02~0.5	0	0
	0.01	0	1
2	0.01~0.6	4	1
1	0.05~0.6	0	0
	0.04	0	1

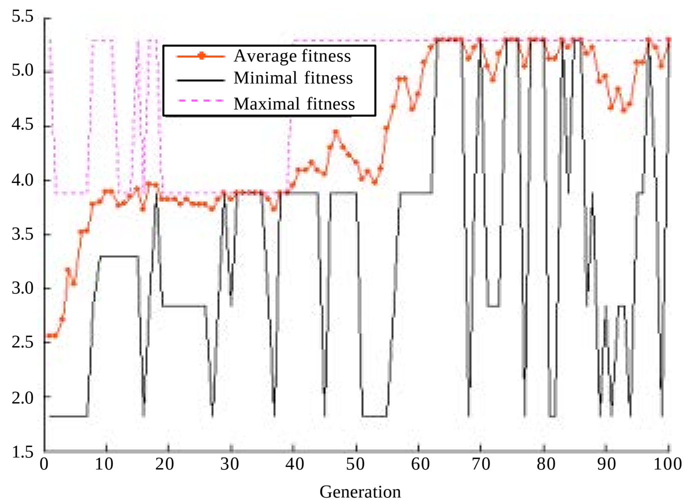

Fig. 2: The curve of fitness for generation

0.04, the wrong sample is 1. The subject who is error sample is 43, it indicates that the person who suffered from high blood pressure for many years is not clinically diagnosed atherosclerosis. It also looks quite good from the waveform, but the result of comprehensive experiment diagnose arteries. A miscalculation occurs. But such a mistake reflects the prediction of experimentally diagnose arteries from another sides. It is a similar situation when spread constant is 0.6 and 0.01 in the third solution.

From the experimental results, probabilistic neural network can predict arteries well. By choosing the appropriate spread constant, the accuracy can reach 100%. So the key problem is the determination of the optimum, that is, the spread constant. So we try to use genetic algorithm to simplify the process of determine the optimal spread constant. The result is shown in Fig. 2.

In the third pretreatment scheme, genetic training maximum fitness reaches maximum after the 40th generation, the minimum, maximum and average fitness will tend to smooth after the 60th generation. The fitness corresponds to spread constant is between 0~0.5. The data obtained from this test method match the data obtained previously, which means the range of training fitness is maximum, the error is minimum and the error rate is minimum. It proves the effectiveness of genetic algorithm.

DISCUSSION

Based on the research achievement of predecessors, eight characteristic parameters including both time-domain parameters and frequency-domain parameters related with atherosclerosis are chosen so that we can analyze the pulse signal in many aspects, which previous studies have not been reported. By learning the probabilistic neural network, we have mastered the knowledge of certain categories of atherosclerosis and the sample we have not learned before. Moreover, the learning process is relatively simple. Using genetic algorithms to optimize the spread constant of probabilistic neural network, which improves the prediction accuracy of the network, provides a feasible method to achieve the classification of atherosclerosis.

REFERENCE

Ahmadlou, M. and H. Adeli, 2010. Enhanced probabilistic neural network with local decision circles: A robust classifier. Integrated Comput. Aided Eng., 17: 197-210.

Alastruey, J., K.H. Parker, J. Peiro and S.J. Sherwin, 2008. Lumped parameter outflow models for 1-D blood flow simulations: Effect on pulse waves and parameter estimation. Commun. Comput. Phys., 4: 317-336.

Duan, Y.Q. and J.L. He, 1998. Genetic algorithm and its modification. Proc. EPSA, 10: 39-52.

Jiao, X.J. and X.Y. Fang, 2000. Research progress of methods of continuous measurement of blood pressure. Space Med. Med. Eng., 13: 148-151.

Luo, Z.C., S. Zhang, W.M. Yang and Z.B. Yang, 1996. A research on characteristic information of pulse wave. J. Beijing Polytechnic Univ., 22: 71-79.

Matthys, K.S., J. Alastruey, J. Peiro, A.W. Khir and P. Segers et al., 2007. Pulse wave propagation in a model human arterial network: Assessment of 1-D numerical simulations against in vitro measurements. J. Biomech., 40: 3476-3486.

Rogers, W.J., J.W. Prichard, Y.L. Hu, P.R. Olson and D.H. Benckart et al., 2000. Characterization of signal properties in atherosclerotic plaque components by intravascular MRI. Arterioscler. Thromb. Vasc. Biol., 20: 1824-1830.

Safar, M.E., B.I. Levy and H. Struijker-Boudier, 2003. Current perspectives on arterial stiffness and pulse pressure in hypertension and cardiovascular diseases. Circulation, 107: 2864-2869.