PERGAMON

Expert Systems with Applications 25 (2003) 1–13

Expert Systems with Applications

www.elsevier.com/locate/eswa

Neural network analysis of internal carotid arterial Doppler signals: predictions of stenosis and occlusion

Elif Derya Übeyli, İnan Güler*

Department of Electronics and Computer Education, Faculty of Technical Education, Gazi University, 06500 Teknikokullar, Ankara, Turkey

Abstract

Doppler ultrasound is a noninvasive technique that allows the examination of the direction, velocity, and volume of blood flow. Doppler ultrasound has proven to be a valuable technique for investigation of artery conditions. Therefore, Doppler ultrasonography is known as reliable technique, which demonstrates the flow characteristics and resistance of internal carotid arteries in stenosis and occlusion conditions. In this study, internal carotid arterial Doppler signals were obtained from 130 subjects, 45 of them had suffered from internal carotid artery stenosis, 44 of them had suffered from internal carotid artery occlusion and the rest of them had been healthy subjects. Multilayer perceptron neural network employing backpropagation training algorithm was used to predict the presence or absence of internal carotid artery stenosis and occlusion. Spectral analysis of internal carotid arterial Doppler signals was done by Burg autoregressive method for determining the neural network inputs. The network was trained, cross validated and tested with subject's internal carotid arterial Doppler signals. Performance indicators and statistical measures were used for evaluating the neural network. By using the network, the classifications of healthy subjects, subjects having internal carotid artery stenosis, and subjects having internal carotid artery occlusion were done with the accuracy of 95.2, 91.3, and 91.7%, respectively.

Keywords: Doppler ultrasound; Spectral analysis; Multilayer perceptron neural network; Backpropagation; Pattern classification; Internal carotid artery

1. Introduction

Doppler ultrasound is a noninvasive technique which is widely used in medicine for the assessment of blood flow in vessels. The technique has improved much since Satomura first demonstrated the application of the Doppler effect to the measurement of blood velocity in 1959 (Keeton & Schlindwein, 1997). Doppler ultrasonography works by emitting a focused ultrasound beam with a base frequency into the body via a piezoelectric transducer and detecting the change in frequency that occurs after the beam is reflected or scattered by moving targets. Since the velocity components are proportional to the frequency shifts, it is possible to track the velocity distribution by obtaining the power spectral density (PSD) estimates (Evans, McDicken, Skidmore, & Woodcock, 1989; Güler, Hardalaç, & Müldür, 2001; Güler, Kara, Güler, & Kıymık, 1996; Güler, Kıymık, & Güler, 1995a,b). Doppler PSD estimates of internal carotid arterial Doppler signals were obtained by using spectrum analysis techniques. By using spectrum analysis techniques, the variations in the shape of the Doppler PSDs were presented in order to obtain medical information.

Doppler ultrasound has proven to be a valuable technique for investigation of internal carotid artery disease. It demonstrates the flow characteristics and resistance of internal carotid arteries in stenosis and occlusion conditions. Internal carotid artery disease is a form of disease that affects the vessels leading to the head and brain (cerebrovascular disease). The results of the studies in the literature have shown that Doppler ultrasound evaluation can give reliable information on both systolic and diastolic blood velocities of the internal carotid arteries and have supported that Doppler ultrasound is useful in screening certain hemodynamic alterations in internal carotid arteries (Adams, 1993; Baker, 1985).

ANNs are computational tools for pattern classification that have been the subject of renewed research interest during the past 15 years. ANNs, employing several formats and learning algorithms, are being used in academic research, industrial, and medical applications (Basheer & Hajmeer, 2000; Fausett, 1994; Haykin, 1994; Miller, Blott, & Hames, 1992; Pham & Sagiroglu, 2001). A feedforward multilayered neural network is the popular model that has

* Corresponding author. Tel.: +90-312-212-3976; fax: +90-312-212-0059.

E-mail address: iguler@tef.gazi.edu.tr (İ. Güler).

0957-4174/03/$ - see front matter © 2003 Elsevier Science Ltd. All rights reserved.
doi:10.1016/S0957-4174(03)00002-2

been playing a central role in applications of neural networks (Basheer & Hajmeer, 2000). The backpropagation algorithm is a widely used training procedure that adjusts the connection weights of a multilayer perceptron (MLP) (Rumelhart, Hinton, & Williams, 1986). It is a gradient descent algorithm that minimizes the mean square error (MSE) between the perceptron output signals and desired response signals in an iterative manner (Basheer & Hajmeer, 2000; Fausett, 1994; Haykin, 1994; Pham & Sagiroglu, 2001).

Important tools in modern decision-making, in any field, include those that allow the decision-maker to assign an object to an appropriate group, or classification. Clinical decision-making is a challenging, multifaceted process. Its goals are precision in diagnosis and institution of efficacious treatment. Achieving these objectives involves access to pertinent data and application of previous knowledge to the analysis of new data in order to recognize patterns and relations. Practitioners apply various statistical techniques in processing data to assist in clinical decision-making and to facilitate the management of patients. As the volume and complexity of data have increased, use of digital computers to support data analysis has become a necessity. In addition to computerization of standard statistical analysis, several other techniques for computer-aided data classification and reduction, generally referred to as ANN, have evolved (Basheer & Hajmeer, 2000; Miller et al., 1992). On analyzing recent developments, it becomes clear that the trend is to develop new methods for computer decision-making in medicine and to evaluate critically these methods in clinical practice. Diagnosis of diseases may be considered as a pattern classification task (Basheer & Hajmeer, 2000; Baxt, 1995). Applications of ANNs in the medical field are numerous and include diagnosis of myocardial infarction (Baxt, 1991); electrocardiogram analysis (Edenbrant, Heden, & Pahlm, 1993; Hilera, Martinez, & Mazo, 1995); electrogastrogram analysis (Chen, Lin, Wu, & McCallum, 1995); electroencephalogram analysis (Prahadan, Sadasivan, & Arunodaya, 1996); coronary artery disease (Akay, 1992; Baxt, 1990; Mobley, Schechter, Moore, McKee, & Eichner, 2000); prognosis prediction for patients with heart failure (Ortiz, Ghefter, Silva, & Sabbatini, 1995); electromyogram analysis (Abel, Zacharia, Forster, & Farrow, 1996); and differentiation of assorted pathological data (Miller et al., 1992). However, neural network analysis of Doppler shift signals is a relatively new approach (Baykal, Reggia, Yalabık, Erkmen, & Beksaç, 1996; Beksaç, Başaran, Eskiizmirliler, Erkmen, & Yörükan, 1996; Ronco & Fernandez, 1999; Siebler, Rose, Sitzer, Bender, & Steinmetz, 1994; Smith, Graham, & Taylor, 1996; Wright & Gough, 1999; Wright, Gough, Rakebrandt, Wahab, & Woodcock, 1997). These numerous applications exhibit the suitability of ANNs in pattern classification including diagnosis of diseases. A potential application of neural networks is predicting medical outcomes such as internal carotid artery stenosis and occlusion.

In this study, MLP neural network employing backpropagation training algorithm was used for the interpretation of internal carotid artery Doppler waveforms. Analysis of internal carotid arterial Doppler signals involved several stages during the implementation of the neural network. Internal carotid arterial Doppler signals were obtained from 130 subjects that 45 of them had suffered from internal carotid artery stenosis, 44 of them had suffered from internal carotid artery occlusion and the rest of them had been healthy subjects and entered into a database. Then MLP neural network was used to predict the presence or absence of internal carotid artery stenosis and occlusion. Selection of network input parameters is important in classification systems. For determining the neural network inputs spectral analysis of internal carotid arterial Doppler signals was done by using Burg autoregressive (AR) method. The network was trained, cross-validated and tested with subject records from the database. Training and testing performance of the neural network was analyzed for determining whether the neural network is adequate to classify data or not.

2. Materials and method

2.1. Subjects

In the present study, internal carotid arterial Doppler signals were obtained from 130 subjects. The group was consisted of 62 females and 68 males with ages ranging from 18 to 65 years and a mean age of 34.5 ± 0.5 years. Toshiba 140A color Doppler ultrasonography was used during examinations and sonograms were taken into consideration. According to examination results, 45 of 130 subjects had suffered from internal carotid artery stenosis, 44 of them had suffered from internal carotid artery occlusion and the rest of them had been healthy subjects. The group having internal carotid artery stenosis was consisted of 22 females and 23 males with a mean age 38.5 ± 0.5 years (range 24–62), the group having internal carotid artery occlusion was consisted of 21 females and 23 males with a mean age 39.0 ± 0.5 (range 27–65) and the healthy subjects were consisted of 19 females and 22 males with a mean age 35.0 ± 0.5 (range 18–60).

2.2. The procedure of the classification system

Fig. 1 shows the procedure used in the development of the classification system. It consists of four parts: (a) measurement of internal carotid arterial Doppler signals, (b) spectral analysis (neural network inputs were selected), (c) classification using neural network, (d) classification results. These procedures are explained in the rest of this study.

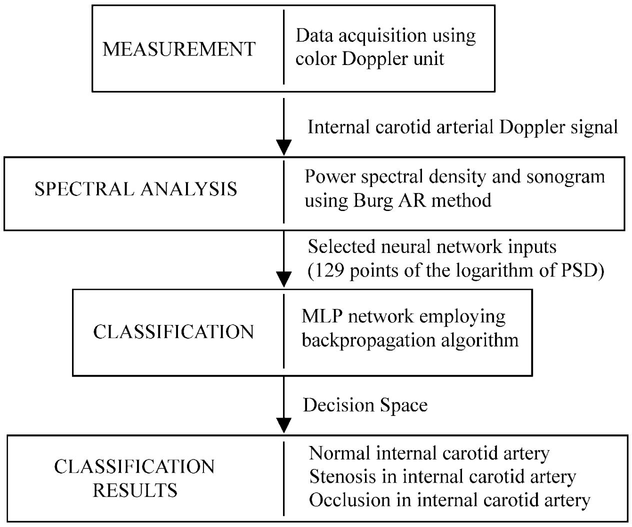

Fig. 1. The procedure used in the development of the classification system.

2.3. Measurements of internal carotid arterial Doppler signals

Doppler ultrasound system is used as a noninvasive method to observe the hemodynamics of internal carotid artery. Internal carotid artery examinations were performed with a Doppler unit using a 5 MHz ultrasonic transducer. The block diagram of the measurement system is shown in Fig. 2. The system consists of five units. These are 5 MHz ultrasonic transducer, analog Doppler unit (Toshiba 140Å Color Doppler Ultrasonography), recorder (Sony), analog/digital interface board (Sound Blaster Pro-16 bit), a personal computer with a printer (Güler et al., 2001; 1996; 1995a,b). The ultrasonic transducer was applied on a horizontal plane to the neck using water-soluble gel as a coupling gel. Care was taken not to apply pressure to the neck in order to avoid artifacts. The probe was most often placed at an angle of 60° towards the internal carotid artery.

2.4. Spectral analysis of internal carotid arterial Doppler signals

Doppler signal is conventionally interpreted by analyzing its spectral content. Diagnosis and disease monitoring are assessed by analysis of spectral shape and parameters. Doppler shift signal can be processed to achieve either a flow velocity waveform or a Doppler power spectrum. Clinically useful information can be extracted from these types of output. Interpretation of the shape of Doppler waveforms may be thought as a process of pattern recognition. Abnormal Doppler waveforms can be recognized by analyzing waveforms, even if physiological or pathological change gives rise to a particular change in waveform shape (Baykal et al., 1996; Beksaç et al., 1996; Evans et al., 1989; Smith et al., 1996; Wright & Gough, 1999; Wright et al., 1997). Burg AR method was used for spectrum analysis of internal carotid arterial Doppler signals in order to extract the features from Doppler signals for the purpose of recognizing normal, stenosis, and occlusion conditions in internal carotid arteries. Using this spectrum analysis technique, PSD estimates and sonograms of internal carotid arterial Doppler signals were obtained. MATLAB software package (MATLAB version 6.0 with signal toolbox) was used for spectral analysis of internal carotid arterial Doppler signals.

2.4.1. Power spectral density

The model-based (parametric) methods are based on modeling the data sequence $x(n)$ as the output of a linear system characterized by a rational system. In the model-based methods, the spectrum estimation procedure consists of two steps. Given the data sequence $x(n)$, $0 \le n \le N-1$, the parameters of the method are estimated. Then from these estimates, the PSD estimate is computed. AR method is most frequently used parametric method because estimation of AR parameters can be done easily by solving linear equations. Since Burg AR method is computationally efficient and yields stable estimates, PSD estimates of internal carotid arterial Doppler signals are obtained by using Burg AR method. The method is based on the minimization of the forward and backward prediction errors and on estimation of the reflection coefficient. The forward and backward prediction errors for a pth-order model are defined as:

$$\hat{e}_{\mathrm{f},p}(n) = x(n) + \sum_{i=1}^{p} \hat{a}_{p,i} x(n-i), \qquad n = p+1, \dots, N, \quad (1)$$

$$\hat{e}_{\mathrm{b},p}(n) = x(n-p) + \sum_{i=1}^{p} \hat{a}^{*}_{p,i} x(n-p+i), \quad (2)$$

$n = p+1, \dots, N.$

The AR parameters are related to the reflection coefficient $\hat{k}_p$ by

$$\hat{a}_{p,i} = \begin{cases} \hat{a}_{p-1,i} + \hat{k}_p \hat{a}^{*}_{p-1,p-i}, & i = 1, \dots, p-1 \\ \hat{k}_p, & i = p \end{cases}. \quad (3)$$

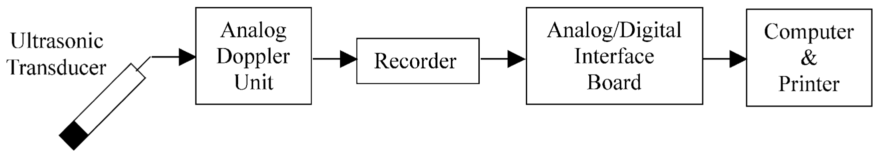

Fig. 2. Block diagram of measurement system.

Burg method considers the recursive-in-order estimation of $\hat{k}_p$ given that the AR coefficients for order $p-1$ have been computed. The reflection coefficient estimate is given by

$$\hat{k}_p = \frac{-2\sum_{n=p+1}^{N} \hat{e}_{\mathrm{f},p-1}(n)\hat{e}^{*}_{\mathrm{b},p-1}(n-1)}{\sum_{n=p+1}^{N} [|\hat{e}_{\mathrm{f},p-1}(n)|^2 + |\hat{e}_{\mathrm{b},p-1}(n-1)|^2]}. \tag{4}$$

From the estimates of the AR parameters, PSD estimation is formed as (Güler et al., 2001; 1996; 1995a,b)

$$\hat{P}_{\mathrm{BURG}}(f) = \frac{\hat{e}_p}{\left|1 + \sum_{k=1}^{p} \hat{a}_p(k)\mathrm{e}^{-\mathrm{j}2\pi fk}\right|^2}, \tag{5}$$

where $\hat{e}_p = \hat{e}_{\mathrm{f},p} + \hat{e}_{\mathrm{b},p}$ is the total least squares error.

Akaike information criterion (AIC) is used for selecting the model order which is based on minimizing Eq. (6)

$$\mathrm{AIC}(p) = \ln \hat{\sigma}^2 + 2p/N, \tag{6}$$

where $\hat{\sigma}^2$ is the estimated variance of the linear prediction error (Güler et al., 2001; 1996; 1995a,b). In this study, model order of AR method was taken as 10 by using Eq. (6).

2.4.2. Sonogram

Toshiba color Doppler ultrasonography displays Doppler frequencies as centimeters per second on the y-axis and plotted time in seconds on the x-axis called as sonograms. The variation in the shape of the Doppler power spectrum as a function of time is usually presented in the form of a sonogram. In sonograms, time is plotted along the horizontal axis, frequency along the vertical axis and the power at a particular frequency and time as the intensity of the corresponding pixel (Evans et al., 1989; Güler et al., 2001; 1996; 1995a,b).

A number of parameters related to the blood flow may be extracted from the sonogram and these are of high clinical value. Although such a display may be of great value for assessing the general quality of the signal and for making qualitative statements about disease state, it contains so much information that some feature extraction must take place before quantitative statements may be made (Evans et al., 1989; Güler et al. , 2001; 1996; 1995a,b). Feature extraction is a process of pattern recognition and consists of extracting and combining salient features of the pattern vector into a feature vector (for example, resistivity index, pulsatility index). Then decision is given whether such a feature vector is obtained from a normal or abnormal artery. Various indices derived from the waveforms defined as resistivity index (RI), pulsatility index (PI) and these can yield information that correlates closely to the disease present (Evans et al., 1989). Both RI and PI are reflections of the resistance to flow, downstream from the point of insonation. Great care must be taken with interpretation of RI and PI in the clinical setting as they are influenced by

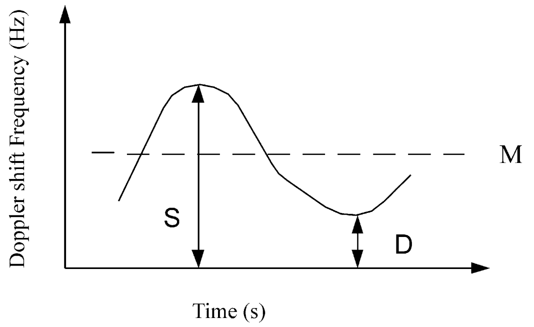

Fig. 3. Diagram illustrating the variables involved in the definitions of resistivity index and pulsatility index. S is maximum systolic height, D is end diastolic height and M is mean height of the waveform.

many factors including proximal stenosis, post stenosis and peripheral resistance. Specifically, maximum systolic height S, end diastolic height D, mean height of the waveform M (Fig. 3) are often used in the calculation of RI and PI (Baykal et al., 1996; Beksaç et al., 1996; Evans et al., 1989; Smith et al., 1996; Wright & Gough, 1999; Wright et al., 1997):

$$\mathrm{RI} = (S - D)/S, \tag{7}$$

$$\mathrm{PI} = (S - D)/M. \tag{8}$$

2.5. Artificial neural networks

ANNs consist of a great number of processing elements (neurons), which are connected with each other; the strengths of the connections are called weights. For the modeling of physical systems, a feedforward multilayered neural network is commonly used. It consists of a layer of input neurons, a layer of output neurons and one or more hidden layers. In order to cope with nonlinearly separable problems, additional layer(s) of neurons placed between the input layer (containing input nodes) and the output neuron are needed leading to the MLP architecture, as shown in Fig. 4. Since the intermediate layers do not interact with the external environment, they are called hidden layers and their nodes called hidden nodes. The addition of intermediate layers revived the perceptron by extending its ability to

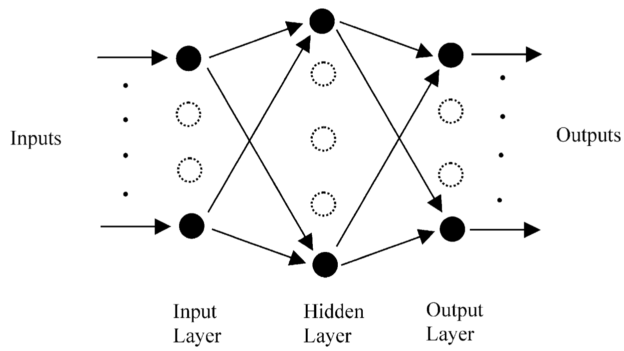

Fig. 4. A MLP neural network.

solve nonlinear classification problems. In ANNs, the knowledge lies in the interconnection weights between neurons. Therefore, training process is an important characteristic of the ANN methodology, whereby representative examples of the knowledge are iteratively presented to the network, so that it can integrate this knowledge within its structure. No assumption is needed about the underlying data probability distributions when ANN is used for pattern classification. Once trained, it can be configured to perform adaptively to improve its performance over time (Basheer & Hajmeer, 2000; Fausett, 1994; Haykin, 1994; Pham & Sagiroglu, 2001).

Although rules for neural network optimization are under development, neural network architectures are derived by trial and error. The determination of appropriate number of hidden layers is one of the most critical tasks in neural network design. Unlike the input and output layers, one starts with no prior knowledge as to the number of hidden layers. A network with too few hidden nodes would be incapable of differentiating between complex patterns leading to only a linear estimate of the actual trend. In contrast, if the network has too many hidden nodes it will follow the noise in the data due to overparameterization leading to poor generalization for untrained data. With increasing number of hidden layers, training becomes excessively time-consuming. The most popular approach to finding the optimal number of hidden layers is by trial and error (Basheer & Hajmeer, 2000; Fausett, 1994; Haykin, 1994; Pham & Sagiroglu, 2001). In the present study, MLP network consisted of one input layer, one hidden layer, and one output layer and the decision about the number of hidden layers in use was given empirically.

ANNs were implemented by using MATLAB software package (MATLAB version 6.0 with neural networks toolbox). MLP ANN was used. This choice is appropriate for solving pattern classification problems where supervised learning is implemented with backpropagation algorithm. The advantage of using this type of ANN is the rapid execution of the trained network, which is particularly advantageous in signal processing applications (Basheer & Hajmeer, 2000; Fausett, 1994; Haykin, 1994; Pham & Sagiroglu, 2001). The applications in the literature exhibit the suitability of ANNs in predicting medical outcomes of Doppler signals when ANNs have trained satisfactorily (Baykal et al., 1996; Beksaç et al., 1996; Miller et al., 1992; Ronco & Fernandez, 1999; Siebler et al., 1994; Smith et al., 1996; Wright & Gough, 1999; Wright et al., 1997).

2.5.1. MLP neural network employing backpropagation training algorithm

In most applications of MLP, the weights are determined by means of the backpropagation algorithm, which is based on searching an error surface (error as a function of ANN weights) using gradient descent for points with minimum error (Basheer & Hajmeer, 2000; Fausett, 1994; Haykin, 1994; Pham & Sagiroglu, 2001). During the training phase, the weights are successively adjusted based on a set of inputs and the corresponding set of desired output targets. Each iteration in backpropagation constitutes two sweeps: forward activation to produce a solution, and a backward propagation of the computed error to modify the weights. The forward and backward sweeps are performed repeatedly until the ANN solution agrees with the desired value within a prespecified tolerance. The backpropagation algorithm provides the needed weight adjustments in the backward sweep (Basheer & Hajmeer, 2000; Haykin, 1994). The backpropagation algorithm is a nonlinear procedure because of the nonlinear threshold element contained in each node, and its behavior is very complex because of the layered structure. However, this nonlinear behavior allows a perceptron to generate highly complex decision regions, which is a desirable property for pattern classification (Haykin, 1994). MLP neural networks employing backpropagation training algorithm are so versatile and can be used for signal processing, image compression, pattern recognition, medical diagnosis, prediction, classification, nonlinear system modeling, and control (Basheer & Hajmeer, 2000; Fausett, 1994). Since backpropagation training algorithm has rapid execution and has widely used in pattern classification problems, MLP neural network employing backpropagation training algorithm was used to predict the presence or absence of internal carotid artery stenosis and occlusion.

In MLP neural network, each neuron j in the hidden layer sums its input signals x_i after multiplying them by the strengths of the respective connection weights w_{ji} and computes its output y_j as a function of the sum

$$y_j = f\left(\sum w_{ji}x_i\right), \tag{9}$$

where f is activation function that is necessary to transform the weighted sum of all signals impinging onto a neuron. In this study, the activation function for hidden neurons was the conventional sigmoidal function with the range between zero and one, which introduces two important properties. First, the sigmoid is nonlinear, allowing the network to perform complex mappings of input to output vector spaces, and secondly it is continuous and differentiable, which allows the gradient of the error to be used in updating the weights. The sum of squared differences between the desired and actual values of the output neurons E is given in Eq. (10)

$$E = \frac{1}{2}\sum_j (y_{dj} - y_j)^2, \tag{10}$$

where y_{dj} is the desired value of output neuron j and y_j is the actual output of that neuron.

Each weight w_{ji} is adjusted by adding an increment Δw_{ji} to it. Δw_{ji} is selected to reduce E as rapidly as possible. How Δw_{ji} is computed depends on the training algorithm adopted. Backpropagation algorithm (Rumelhart et al., 1986) is then invoked to adjust all the weights in the network and gives the change $\Delta w_{ji}(k)$ in the weight of the connection between neurons i and j at iteration k as

$$\Delta w_{ji}(k) = -\alpha \frac{\partial E}{\partial w_{ji}(k)} + \mu \Delta w_{ji}(k-1), \quad (11)$$

where α is called the learning coefficient; μ, the momentum coefficient; $\Delta w_{ji}(k-1)$ is the weight change in the immediately preceding iteration. The learning coefficient is proportional to the adjustment which is made to individual weights during training and affects the rate at which the weights converge towards a solution. Momentum coefficient is proportional to the weight error derivative calculated at each training iteration, and appropriate values enable the backpropagation algorithm to converge towards the true minimum. It is possible to obtain a fast and stable approach towards a solution with appropriate values of learning coefficient and momentum coefficient (Fausett, 1994; Haykin, 1994). For this study, learning coefficient and momentum coefficient were determined empirically and α was taken as 0.1 and μ was taken as 0.2.

2.6. Performance indicators of the neural network

2.6.1. Measuring error

Given a random set of initial weights, the outputs of the network will be very different from the desired classifications. As the network is trained, the weights of the system are continually adjusted to reduce the difference between the output of the system and the desired response. The difference is referred to as the error and can be measured in different ways. The most common measurement is the MSE. The MSE is the average of the squares of the difference between each output and the desired output. In addition to MSE, normalized mean squared error (NMSE), mean absolute error (MAE), minimum absolute error and maximum absolute error can be used for the measuring error of the neural network (Basheer & Hajmeer, 2000; Fausett, 1994; Haykin, 1994). In this study, MSE, NMSE, MAE, minimum absolute error and maximum absolute error were used for measuring performance of the neural network.

2.6.2. Cross-validation

Cross-validation is a highly recommended criterion for stopping the training of a network. During performance analysis of network, cross-validation can be used for determining the final training. In general, it is known that a network with enough weights will always learn the training set better as the number of iterations is increased. However, neural network researchers have found that this decrease in the training set error was not always coupled to better performance in the test. When the network is trained too much, the network memorizes the training patterns and does not generalize well. The training holds the key to an accurate solution, so the criterion to stop training must be very well described. The aim of the stop criterion is to maximize the network's generalization (Basheer & Hajmeer, 2000; Mobley et al., 2000).

2.6.3. Classification and regression

Neural networks are used for both classification and regression. In classification, the aim is to assign the input patterns to one of several classes, usually represented by outputs restricted to lie in the range from 0 to 1, so that they represent the probability of class membership. While the classification is carried out, a specific pattern is assigned to a specific class according to the characteristic features selected for it. In regression, desired output and actual network output results can be shown on the same graph and performance of network can be evaluated in this way (Basheer & Hajmeer, 2000; Fausett, 1994; Haykin, 1994).

2.6.4. Confusion matrix

A confusion matrix is a simple methodology for displaying the classification results of a network. The confusion matrix is defined by labeling the desired classification on the rows and the actual network outputs on the columns. For each exemplar, 1 is added to the cell entry defined by desired classification and the actual network outputs. Since the actual network outputs and the desired classification wanted to be the same, the ideal situation is to have all the exemplars end up on the diagonal cells of the matrix (the diagonal that connects the upper-left corner to the lower right) (Beksaç et al., 1996).

2.6.5. Sensitivity, specificity and receiver operating characteristic curve analysis

The simplest classification problem is that of separating one-dimensional feature vectors into two groups. In this situation the only choice that needs to be made is where to locate the decision threshold. If there is no overlap between the magnitudes of the vectors obtained from patients belonging to the two classes, the threshold can simply be chosen to separate the classes completely. In general, the results from the two classes do overlap and so depending on where the threshold is placed some signals from normal subjects will be adjudged abnormal and/or some signals from abnormals will be adjudged normal. The best choice of threshold will then depend on a number of factors. There are two important measures of the performance of a diagnostic test; sensitivity (or true positive fraction) and specificity (or true negative fraction) which are defined as:

$$\text{Sensitivity } (TPF) = \text{Number of true positive decisions}/\text{Number of actually positive cases}, \quad (12)$$

and

$$\text{Specificity } (TNF) = \text{Number of true negative decisions}/\text{Number of actually negative cases}. \quad (13)$$

These measures are dependent since they are both affected by the position of the decision threshold and as the threshold is moved to increase sensitivity, so specificity decreases. The best method of assessing the value of a test and defining an appropriate decision threshold is to plot a receiver operating characteristic (ROC) curve for the test. Such a curve is derived by varying the decision threshold in small steps and determining the TPF and TNF for each new threshold value (Evans et al., 1989; Zweig & Campbell, 1993).

2.6.6. Correlation coefficient

The size of MSE can be used to determine how well the network output fits the desired output, but it may not reflect whether the two sets of data move in the same direction. The correlation coefficient (r) solves this problem. The correlation coefficient is limited with the range $[-1, 1]$. When $r = 1$ there is a perfect positive linear correlation between network output and desired output, which means that they vary by the same amount. When $r = -1$ there is a perfectly linear negative correlation between network output and desired output, that means they vary in opposite ways (when network output increases, desired output decreases by the same amount). When $r = 0$ there is no correlation between network output and desired output (the variables are called uncorrelated). Intermediate values describe partial correlations (Fausett, 1994; Haykin, 1994).

3. Results and discussion

In this study, for solving pattern classification problem MLP neural network employing backpropagation training algorithm was used. Effective training algorithm and better-understood system behaviour are the advantages of this type of neural network. Selection of network input parameters and performance of neural network are important for the prediction of internal carotid artery stenosis and occlusion by MLP neural network employing backpropagation training algorithm.

3.1. Selection of network input parameters

During training, the input and desired data will be repeatedly presented to the network. When using neural network, decision must be taken for how to divide data into a training set and a test set. In this study, 62 of 130 subjects were used for training and the rest of them were used for testing. For obtaining a better network's generalization 15 of training subjects were used as cross-validation set.

The outputs are represented by unit basis vectors:

[0 0 1] = normal
[0 1 0] = stenosis
[1 0 0] = occlusion

Selection of network input parameters plays an important role in classifying systems. During this study, two different sets of parameters were chosen for network inputs. In the first experiment, RI and PI values, which may be extracted from sonograms, were taken as network input parameters. Some sample sonograms of internal carotid arterial Doppler signals obtained from 33-year-old healthy subject, 35-year-old unhealthy subject suffering from internal carotid artery stenosis, and 36-year-old unhealthy subject

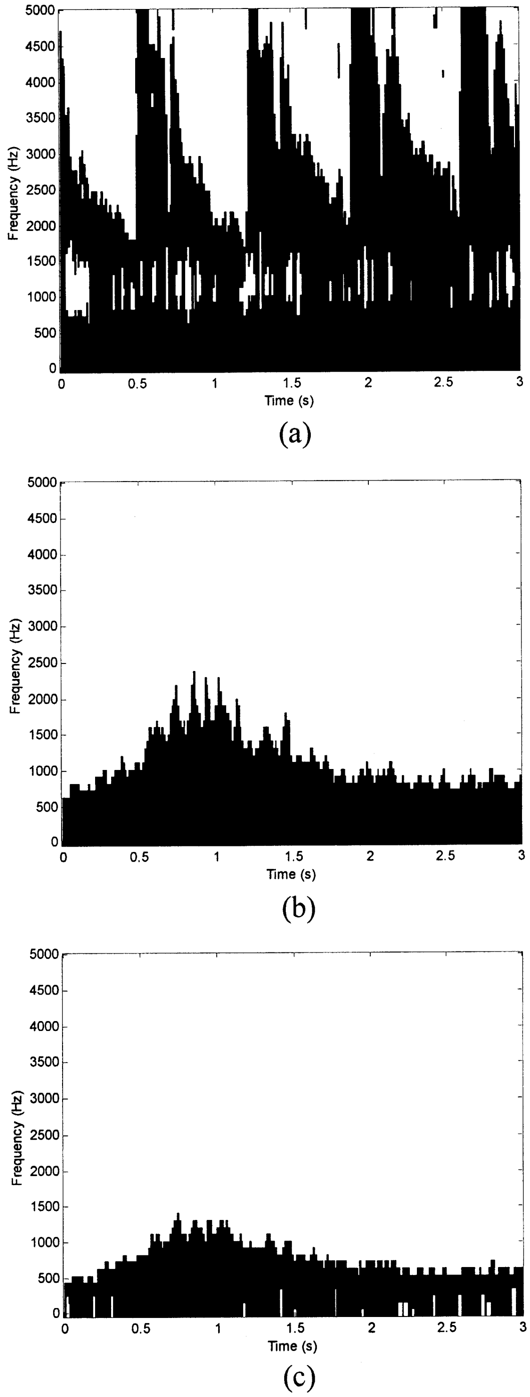

Fig. 5. Sonograms of internal carotid arterial Doppler signals: (a) 33-year-old healthy subject, (b) 35-year-old unhealthy subject suffering from internal carotid artery stenosis, (c) 36-year-old unhealthy subject suffering from internal carotid artery occlusion.

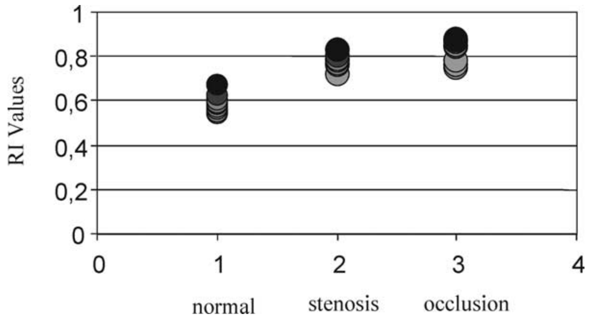

Fig. 6. Resistivity index values for normal, stenosis and occlusion groups.

suffering from internal carotid artery occlusion are shown in Fig. 5. Independent-samples t-tests were used to compare RI and PI values of normal subjects and subjects having stenosis and occlusion. t-Tests were performed with a statistical package (SPSS version 8.0). The level of statistical significance was taken as $p = 0.05$. The results of independent-samples t-tests showed that there was a statistically significant difference between RI and PI values of normal subjects and subjects having stenosis and occlusion ($p < 0.01$). However, Figs. 6 and 7 show the considerable overlap in RI and PI values of normal, stenosis and occlusion groups. Also, the previous studies have shown that standard waveform indices such as RI and PI are inadequate to evaluate Doppler waveforms (Baykal et al., 1996; Wright & Gough, 1999; Wright et al., 1997).

Since RI and PI values are not appropriate features for predictions of Doppler waveforms, internal carotid arterial Doppler PSD values were taken as network input parameters. Internal carotid arterial Doppler signal's PSD estimates were calculated by using the Burg AR method for each subject. PSD estimates of internal carotid arterial Doppler signals given in Fig. 5 are shown in Fig. 8. The correlation coefficients for internal carotid arterial Doppler PSD values of healthy subjects, unhealthy subjects suffering from artery stenosis, and unhealthy subjects suffering from artery occlusion were calculated with a statistical package (SPSS version 8.0). The calculated r values for healthy subjects are varying in the range $[0.905, 0.997]$ and show that there are perfect positive linear correlations among the PSD values of healthy subjects. The calculated r values for

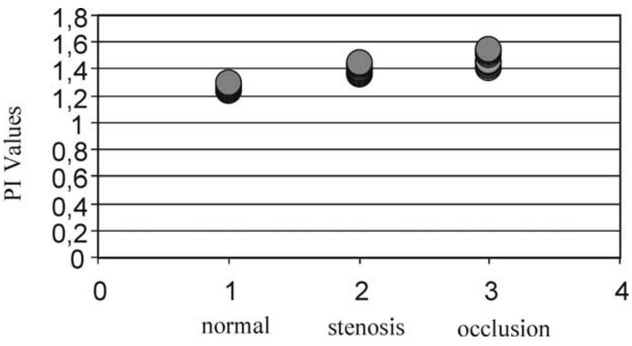

Fig. 7. Pulsatility index values for normal, stenosis and occlusion groups.

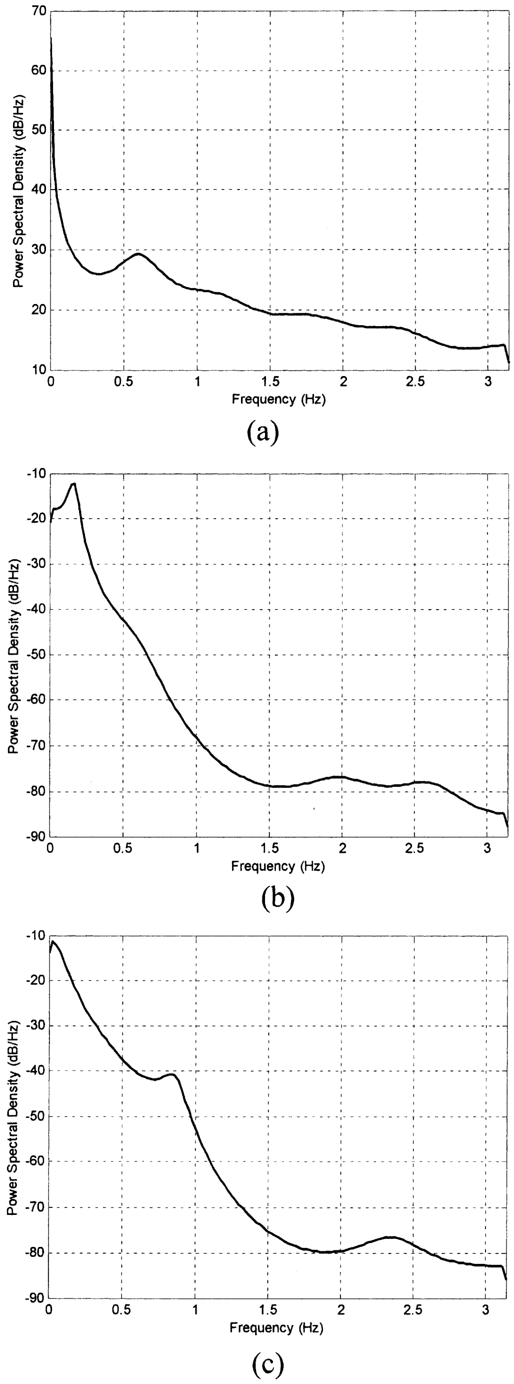

Fig. 8. PSD estimates of internal carotid arterial Doppler signals: (a) 33-year-old healthy subject, (b) 35-year-old unhealthy subject suffering from internal carotid artery stenosis, (c) 36-year-old unhealthy subject suffering from internal carotid artery occlusion.

unhealthy subjects suffering from artery stenosis are varying in the range $[0.924, 0.995]$ and indicate that there are perfect positive linear correlations among the PSD values of unhealthy subjects suffering from artery stenosis. The calculated r values for unhealthy subjects suffering from artery occlusion are varying in the range $[0.911, 0.992]$ and demonstrate that there are perfect positive linear

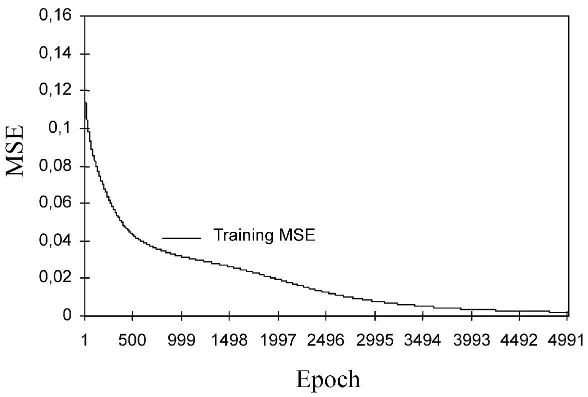

Fig. 9. Training MSE curve of MLP neural network.

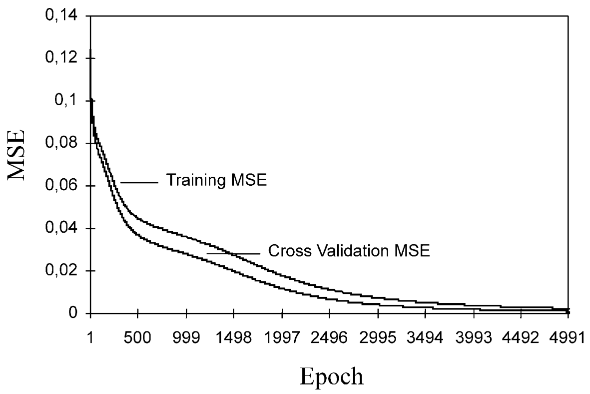

Fig. 10. Training and cross-validation MSE curves of MLP neural network.

correlations among the PSD values of unhealthy subjects suffering from artery occlusion. Then 129 points of the logarithm of the PSD values were used as network inputs in this study.

3.2. Performance analysis of MLP neural network

MLP neural network employing backpropagation was trained with the training set and checked with the test set. In this study, performance analysis of the neural network is examined in two parts as training performance and testing performance.

3.2.1. Training performance of MLP neural network

The neural network will find the input–output map by analyzing the training set repeatedly. This is called the network training phase. Most of the neural network design effort is spent in the training phase. Training is normally slow because the network's weights are being updated based on the error information. There is a need to monitor how well the network is learning. One of the simplest methods is to observe how the square difference between the network's output and the desired response changes over training iterations. The curve of the MSE versus iteration is called as the training curve. Training MSE curve of MLP neural network in 5000 epochs is shown in Fig. 9. As the network learns, the error is converging to zero. The values of minimum MSE and final MSE during training are given in Table 1.

A neural network is subject to what is known as the memorization of training data, otherwise known as the statistical phenomenon of overfitting when it is trained too much. If a network overfits or memorizes the training data, its generalized performance on other sample populations such as the test file, or on records for which prospective predictions are to be made, is likely to be severely compromised. Therefore, the most important criterion is choosing the number of iterations for training. Cross-validation is one of the most powerful methods to stop the training. In principle the training curve decreases exponentially to zero or a small constant. How small depends on the situation and a judgement must be used to find what error value is appropriate for the problem. In Fig. 10, the error in training set and the validation set is shown on the same graph. When the error in the cross-validation increases, the training should be stopped because the point of best generalization has been reached. In this study as it is seen from Table 2, training was done in 5000 epochs and number of epochs was determined according to cross-validation. The values of minimum MSE and final MSE during training and cross-validation are given in Table 2. Since MSE (Figs. 9 and 10) is converging to a small constant approximately zero in 5000 epochs, training of the neural network is determined as successful.

3.2.2. Testing performance of MLP neural network

After training phase, testing of the MLP neural network was done. The data, that the network had not seen before, was applied to the network for testing the network performance. Since training was successful and the network's topology was correct, it applied its past experience to test data and produced a good solution.

Table 1
The values of minimum and final MSE during training

Best network	Training
Epoch #	5000
Minimum MSE	0.001905056
Final MSE	0.001905056

Table 2
The values of minimum and final MSE during training and cross-validation

Best networks	Training	Cross-validation
Epoch #	5000	5000
Minimum MSE	0.002408279	0.001094893
Final MSE	0.002408279	0.001094893

Testing performance was analyzed in classification and regression form. Also, ROC curve analysis was done for the evaluation of testing performance of the neural network.

In this study, there were three classes as normal, stenosis and occlusion which were indicating situation of subject's internal carotid arteries. Classification results of the network was displayed by using a confusion matrix. In a confusion matrix, each cell contains the raw number of exemplars classified for the corresponding combination of desired and actual network outputs. The confusion matrix showing the classification results of this network is given below.

Confusion matrix:

Output/desired	Result (normal)	Result (stenosis)	Result (occlusion)
Result (normal)	20	0	0
Result (stenosis)	1	21	2
Result (occlusion)	0	2	22

When this confusion matrix is examined, it is seen that one normal subject classified incorrectly by the network as a subject having stenosis, two subjects having stenosis classified as subjects having occlusion and two subjects having occlusion classified as subjects having stenosis. The other option in confusion matrix, is to display each cell as a percentage of the exemplars for the desired class. In this format, each row of the matrix sums to 100.

Confusion matrix (percentage of correct):

	Result (normal)	Result (stenosis)	Result (occlusion)
Result (normal)	95.23809814	4.76190186	0
Result (stenosis)	0	91.30434418	8.69565582
Result (occlusion)	0	8.33333588	91.66666412

According to this confusion matrix, normal subjects were classified correctly with 95.23809814% and incorrectly with 4.76190186%. Subjects having internal carotid artery stenosis were classified correctly with 91.30434418% and incorrectly with 8.69565582%. Subjects having internal carotid artery occlusion were classified correctly with 91.66666412% and incorrectly with 8.33333588%. In this case, the classifications of normal subjects, subjects having stenosis and subjects having occlusion were done with the accuracy of 95.23809814, 91.30434418 and 91.66666412%, respectively.

Confusion matrix is presented for comparing similarities and dissimilarities between the predictions of the physician and the network. Paired-samples *t*-tests were used to compare predictions of physician and the network for normal subjects, subjects having stenosis and subjects having occlusion. The results of paired-samples *t*-tests showed that there was no statistically significant difference between predictions of physician and the network for normal subjects, subjects having stenosis and subjects having occlusion ($p > 0.05$).

For testing network's performance desired output and actual network output results can be shown on the same graph. This type of testing is known as regression. Desired outputs and actual network outputs for normal subjects, subjects having stenosis and subjects having occlusion are seen in Fig. 11(a)–(c), respectively.

Desired outputs and actual network outputs are restricted to vary in the range from 0 to 1. When Fig. 11(a) is examined, 21 normal subject's desired outputs are seen as 1. It is seen that actual network outputs are varying around 1 except one normal subject and actual network output of this one subject is seen as 0. Hence, it is understood that 1 of 21 normal subjects was classified incorrectly by the network. In the same manner, Fig. 11(b) is examined and desired outputs of 23 subjects having stenosis are seen as 1. Actual network outputs are varying around 1 except two subjects having stenosis and actual network outputs of these two subjects are seen as 0. In this situation, it is understood that two of 23 subjects having stenosis were classified incorrectly by the network. In the same way, Fig. 11(c) is examined and desired outputs of 24 subjects having occlusion are seen as 1. It is seen that actual network outputs are varying around 1 except two subjects having occlusion and actual network outputs of these two subjects are seen as 0. Thus, it is understood that two of 24 subjects having occlusion were classified incorrectly by the network.

ROC curve is one of the best method of evaluating the performance of a test and defining an appropriate decision threshold. The best choice of threshold will then depend on a number of factors including the consequences of making both types of false classification (false positive and false negative) and the prevalence of disease in the target population. For a given result obtained by a classifier system, four possible alternatives exist that describe the nature of the result: (i) true positive (TP), (ii) false positive (FP), (iii) true negative (TN), and (iv) false negative (FN). In this study, a TP decision occured when the positive diagnosis of the system coincided with a positive diagnosis according to the physician. A FP decision occured when the system made a positive diagnosis that did not agree with the physician. A TN decision occured when both the system and the physician suggested the absence of a positive diagnosis. A FN decision occured when the system made a negative diagnosis that did not agree with the physician. Also the best

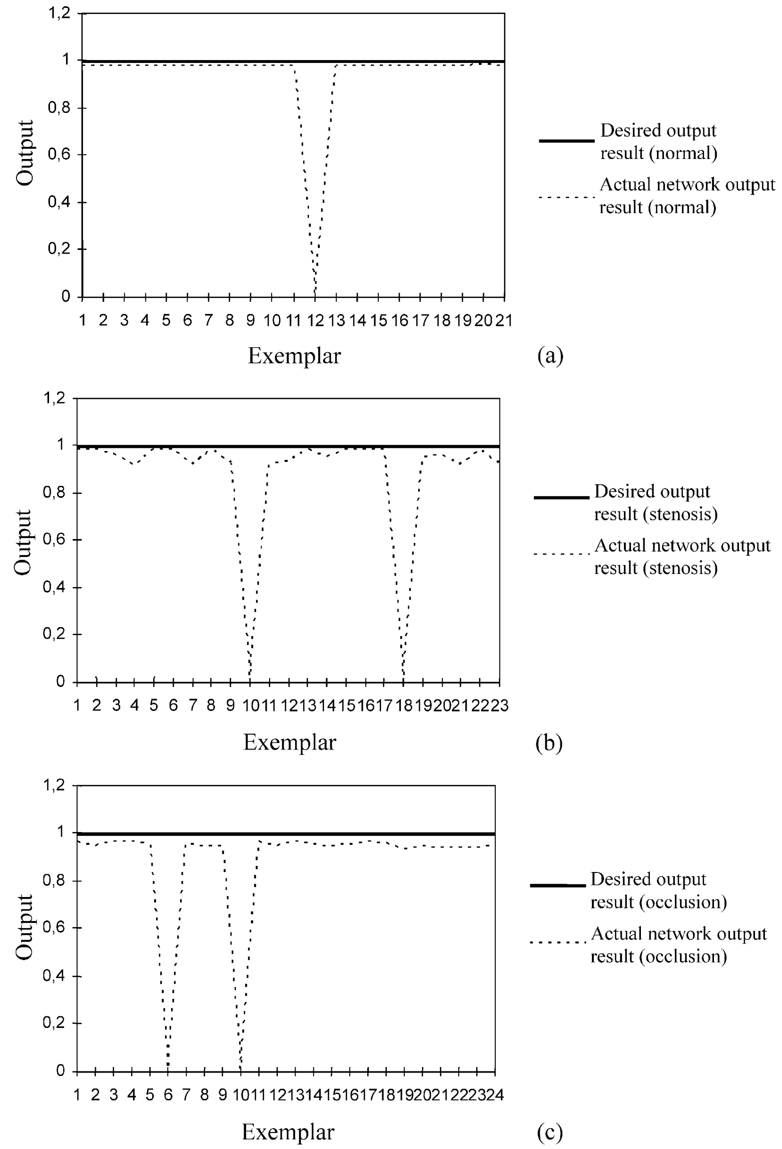

Fig. 11. Regression: (a) desired output result (normal) and actual network output result (normal), (b) desired output result (stenosis) and actual network output result (stenosis), (c) desired output result (occlusion) and actual network output result (occlusion).

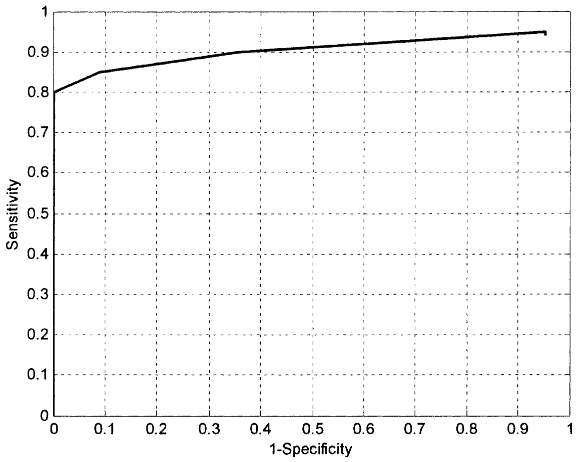

Fig. 12. ROC curve.

choice of threshold for any test depends on the shape of the ROC curve. A good test (curve in Fig. 12) is one for which sensitivity rises rapidly and one-specificity hardly increases at all until sensitivity becomes high. Hence, the closer the curve is to the upper left corner, the higher the overall

Table 3
Performance evaluation parameters and their values

Performance	Result (normal)	Result (stenosis)	Result (occlusion)
MSE	0.005049655	0.113833123	0.106269891
NMSE	0.430772984	0.475908485	0.447638604
MAE	0.201705291	0.208734353	0.208386512
Minimum absolute error	0.004068895	0.004531152	0.004310478
Maximum absolute error	0.793271264	0.925639853	0.867307127
r	0.947317005	0.907028173	0.908968343
Percent correct	95.23809814	91.30434418	91.66666412

Table 4
Applying nine subject records as test data

Subject no	Network inputs	Network's result (occlusion)	Network's result (stenosis)	Network's result (normal)	Physician's prediction
1	129 points of the logarithm of PSD	0	0	1	Normal
2	129 points of the logarithm of PSD	0	1	0	Stenosis
3	129 points of the logarithm of PSD	1	0	0	Occlusion
4	129 points of the logarithm of PSD	0	1	0	Stenosis
5	129 points of the logarithm of PSD	0	0	1	Normal
6	129 points of the logarithm of PSD	1	0	0	Occlusion
7	129 points of the logarithm of PSD	0	1	0	Stenosis
8	129 points of the logarithm of PSD	0	0	1	Normal
9	129 points of the logarithm of PSD	1	0	0	Occlusion

accuracy of the test. ROC curve which is seen in Fig. 12 represents the network performance on the test file of 130 subject records.

In this study, MSE, NMSE, MAE, minimum absolute error, maximum absolute error, correlation coefficient (r), and percent correct were used for measuring performance of the neural network. Performance evaluation parameters and their values in this neural network are given for normal subjects, subjects having stenosis, and subjects having occlusion in Table 3. From the results of performance evaluation and statistical measures, the neural network was found to be successful.

3.2.3. Applying test data

After training and testing this network, it was determined that the network adequately classified data. Then nine subject records were applied to this network. For this experiment network inputs of nine subject records were entered and desired outputs for these subjects were not given to the network. The results of network for each subject is compared with the physician's prediction. From Table 4 it is seen that nine subjects are classified correctly by this network.

4. Conclusion

ANN is a valuable tool in the medical field for the development of decision support systems. What makes neural networks a promising tool is their capacity to find near-optimum solutions from limited or incomplete data sets and the fact that learning is accomplished through training. The increased utilization of ANNs is linked to several features they possess, namely (i) the ability to recognize and learn the underlying relations between input and output without explicit physical consideration, regardless of the problem's dimensionality and system nonlinearity, and (ii) the high tolerance to data containing noise and measurement errors due to distributed processing within the network. ANNs also have limitations that should not be overlooked. These include (i) a lack of clear rules or fixed guidelines for optimal ANN architecture design, (ii) a lack of physical concepts and relations, and (iii) the inability to explain in a comprehensible form the process through which a given decision was made by the ANN. In addition to these characteristics, it has been shown that neural networks can combine data of a different nature in one system, such as data derived from clinical protocols and laboratory data obtained from measurements and features from signals and images, thus forming an integrated diagnostic system. Diagnosis of diseases may be considered as pattern classification task and the previous studies have shown that ANNs can offer a potentially superior method for classification and analysis of Doppler signals. Since testing and subsequent implementation is rapid, in this study, MLP neural network employing backpropagation training algorithm was used to predict the presence or absence of internal carotid artery stenosis and occlusion. The neural network was trained, cross-validated and tested with internal carotid arterial Doppler signals. Performance indicators and statistical measures were used for evaluating the neural network. For prediction purposes, it has been presented that MLP neural network employing backpropagation works reasonably well. The correct classification rate was 95.2% for healthy subjects, 91.3% for subjects having internal carotid artery stenosis, and 91.7% for subjects having internal carotid artery occlusion. Based on the accuracy of the neural network's predictions, it can be mentioned that classification of internal carotid artery waveforms is feasible by MLP neural network employing backpropagation training algorithm.

References

Abel, E. W., Zacharia, P. C., Forster, A., & Farrow, T. L. (1996). Neural network analysis of the EMG interference pattern. *Medical Engineering and Physics*, *18*(1), 12–17.

Adams, H. P. (1993). *Handbook of cerebrovascular diseases*. New York: Marcel Dekker.

Akay, M. (1992). Non-invasive diagnosis of coronary artery disease using a neural network algorithm. *Biological Cybernetics*, *67*, 361–367.

Baker, W. H. (1985). *Diagnosis and treatment of carotid artery disease*. New York: Futura Publishing Company.

Basheer, I. A., & Hajmeer, M. (2000). Artificial neural networks: Fundamentals, computing, design, and application. *Journal of Microbiological Methods*, *43*, 3–31.

Baxt, W. G. (1990). Use of an artificial neural network for data analysis in clinical decision making: The diagnosis of acute coronary occlusion. *Neural Computation*, *2*, 480–489.

Baxt, W. G. (1991). Use of an artificial neural network for the diagnosis of myocardial infarction. *Annals of Internal Medicine*, *115*, 843–848.

Baxt, W. G. (1995). Application of artificial neural networks to clinical medicine. *Lancet*, *346*, 1135–1138.

Baykal, N., Reggia, J. A., Yalabık, N., Erkmen, A., & Beksaç, M. S. (1996). Feature discovery and classification of Doppler umbilical artery blood flow velocity waveforms. *Computers in Biology and Medicine*, *26*(6), 451–462.

Beksaç, M. S., Başaran, F., Eskiizmirliler, S., Erkmen, A. M., & Yörükan, S. (1996). A computerized diagnostic system for the interpretation of umbilical artery blood flow velocity waveforms. *European Journal of Obstetrics and Gynecology and Reproductive Biology*, *64*, 37–42.

Chen, J. D. Z., Lin, Z., Wu, Q., & McCallum, R. W. (1995). Non-invasive identification of gastric contractions from surface electrogastrogram using backpropagation neural networks. *Medical Engineering and Physics*, *17*(3), 219–225.

Edenbrant, L., Heden, B., & Pahlm, O. (1993). Neural networks for analysis of ECG complexes. *Journal of Electrocardiology*, *26*, 66–73.

Evans, D. H., McDicken, W. N., Skidmore, R., & Woodcock, J. P. (1989). *Doppler ultrasound: physics instrumentation and clinical applications*. Wiley: Chicester.

Fausett, L. (1994). *Fundamentals of neural networks architectures, algorithms, and applications*. Englewood Cliffs, NJ: Prentice Hall.

Güler, İ., Hardalaç, F., & Müldür, S. (2001). Determination of aorta failure with the application of FFT, AR and wavelet methods to Doppler technique. *Computers in Biology and Medicine*, *31*, 229–238.

Güler, İ., Kara, S., Güler, N. F., & Kıymık, M. K. (1996). Application of autoregressive and fast Fourier transform spectral analysis to tricuspid and mitral valve stenosis. *Computer Methods and Programs in Biomedicine*, *49*, 29–36.

Güler, N. F., Kıymık, M. K., & Güler, İ. (1995a). Comparison of FFT and AR-based sonogram outputs of 20 MHz pulsed Doppler data in real time. *Computers in Biology and Medicine*, *25*, 383–391.

Güler, N. F., Kıymık, M. K., & Güler, İ. (1995b). Autoregressive-based sonogram output of 20 MHz pulsed Doppler data. *Medical Progress Through Technology*, *21*, 105–110.

Haykin, S. (1994). *Neural networks: A comprehensive foundation*. New York: Macmillan.

Hilera, J. R., Martinez, V. J., & Mazo, M. (1995). ECG signal processing with neural networks. *International Journal of Uncertainty Fuzziness and Knowledge-Based Systems*, *3*(4), 419–430.

Keeton, P. I. J., & Schlindwein, F. S. (1997). Application of wavelets in Doppler ultrasound. *Sensor Review*, *17*(1), 38–45.

Miller, A. S., Blott, B. H., & Hames, T. K. (1992). Review of neural network applications in medical imaging and signal processing. *Medical and Biological Engineering and Computing*, *30*, 449–464.

Mobley, B. A., Schechter, E., Moore, W. E., McKee, P. A., & Eichner, J. E. (2000). Predictions of coronary artery stenosis by artificial neural network. *Artificial Intelligence in Medicine*, *18*, 187–203.

Ortiz, J., Ghefter, C. G. M., Silva, C. E. S., & Sabbatini, R. E. M. (1995). One-year mortality prognosis in heart failure: A neural network approach based on echocardiographic data. *Journal of the American College of Cardiology*, *26*, 1586–1593.

Pham, D. T., & Sagiroglu, S. (2001). Training multilayered perceptrons for pattern recognition: A comparative study of four training algorithms. *International Journal of Machine Tools and Manufacture*, *41*, 419–430.

Prahadan, N., Sadasivan, P. K., & Arunodaya, G. R. (1996). Detection of seizure activity in EEG by an artificial neural network: A preliminary study. *Computers and Biomedical Research*, *29*(4), 303–313.

Ronco, A. L., & Fernandez, R. (1999). Improving ultrasonographic diagnosis of prostate cancer with neural networks. *Ultrasound in Medicine and Biology*, *25*(5), 729–733.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*, 533–536.

Siebler, M., Rose, G., Sitzer, M., Bender, A., & Steinmetz, R. (1994). Real-time identification of cerebral microemboli with ultrasound feature detection by a neural network. *Radiology*, *192*, 739–742.

Smith, J. H., Graham, J., & Taylor, R. J. (1996). The application of an artificial neural network to Doppler ultrasound waveforms for the classification of arterial disease. *International Journal of Clinical Monitoring and Computing*, *13*, 85–91.

Wright, I. A., & Gough, N. A. J. (1999). Artificial neural network analysis of common femoral artery Doppler shift signals: Classification of proximal disease. *Ultrasound in Medicine and Biology*, *24*(5), 735–743.

Wright, I. A., Gough, N. A. J., Rakebrandt, F., Wahab, M., & Woodcock, J. P. (1997). Neural network analysis of Doppler ultrasound blood flow signals: A pilot study. *Ultrasound in Medicine and Biology*, *23*(5), 683–690.

Zweig, M. H., & Campbell, G. (1993). Receiver-operating characteristic (ROC) plots: A fundamental evaluation tool in clinical medicine. *Clinical Chemistry*, *39*(4), 561–577.