Artificial Intelligence Algorithm for Heart Disease Diagnosis using Phonocardiogram Signals

Ibrahim Abdel-Motaleb and Rohit Akula
Department of Electrical Engineering, Northern Illinois University, DeKalb, IL 60115
ibrahim@niu.edu

***Abstract*-An artificial intelligence system has been developed using Artificial Neural Networks (ANN) algorithms to diagnose heart disease from Phonocardiogram (PCG) signals. Four new featured characteristics of the signals, namely activity, complexity, mobility and the spectral peaks from the power spectral density plots are used as input to the neural network. Ninety-four PCG signals for three heart diseases were used in this study to test the accuracy of the neural networks. After the signals are filtered and the feature characteristics are extracted, the features are fed to the neural networks. Classification is carried using the Radial Basis Function (RBF) network and the Back Propagation Network (BPN) techniques. Receiver operating characteristic (ROC) is calculated to measure the accuracy for both structures. The results show that RBF provided 98% accuracy in predicting the disease compared with 90.8% for BPN. The developed artificial intelligence algorithm has been shown to be a powerful technique in automatic diagnosis of heart diseases using PCG signals.**

I. INTRODUCTION

Cardiac auscultation is the first basic analysis tool used to evaluate the functional state of the heart [1]. Electrocardiography (ECG) test is ordered if the heart sound from the Phonocardiogram (PCG) shows any abnormalities. Heart auscultation has the advantage of being heard and seen on the screen, which gives a higher level of assurance of the accuracy of the initial diagnosis. However, because of noise and human misinterpretation, diagnosis may not be as accurate as desired. Diagnosis accuracy can be substantially enhanced if an artificial intelligence machine is used to provide potential diagnoses using some characteristic features of the heart sound signals. Such process is expected to reduce mortality rate and cost of care.

There are a number of publications that propose different techniques for the extraction of features from the heart sounds and classify them using neural networks. In the late 80's Mohamed and Raafat developed a mathematical model to describe the heart sounds and murmurs by a finite number of parameters [2]. In this case, features were extracted based on fourth order linear prediction of the cardiac cycle frames, where classification was carried out based on the minimum distance between the features of the measured pattern and the reference patterns.

Patil and Kumaraswamy proposed an intelligent heart attack prediction system based on Data Mining and Artificial Neural Network [3]. In this method, the parameters vital to the heart attack are computed by using K-means clustering algorithm to the available data. These frequent patterns are mined from the data, with the aid of the Maximal Frequent Itemset Algorithm (MAFIA). The patterns are then selected based on the computed significant weight age. Although the above study reported that this method is capable of predicting the heart attack using MAFIA algorithm, the prediction accuracy was not reported for the work. Furthermore, this technique uses features corresponding to the behavioral habits of the subject, such as smoking and alcohol consumption, instead of feature characteristics of the heart sound signal itself.

A novel method of segmenting the heart sounds using homomorphic filtering and feature extraction from wavelet coefficients are classified using GAL (Grow and Learn) algorithm [4]. The accuracy of this method was reported to be 90.9%. Reed et al. [5] worked on analysis of the heart sounds for symptom detection, where heart sounds were segmented and transformed using wavelet decomposition. The transformed vectors were reduced to smaller vector sizes by discarding levels with shortest scales [5]. Finally, each vector was classified using a three layer neural network, which gave 100% accuracy for all heart sounds. The disadvantage of this technique is the need to use a large number of hidden layer neurons, which may reach 50 layers. Such large number will not only require a very long computation time, but it makes it very hard to train the neural network.

Heart rate variability (HRV) based classification of arrhythmia has been reported [6]. This method is based on both the General Discriminant Analysis (GDA) and the Multi Layer Perceptron (MLP) method. The results have shown that this method resulted in accuracy of 100% for the data acquired by the authors from MIT-BIH database. This method, however, uses ECG based HRV signals not PCG signals. It should be noted here that obtaining ECG signals are not considered routine tests for primary care physicians, since it requires laboratory set-ups, and consequently, it is time consuming and less cost effective.

In this work, a new algorithm is developed, where new features characteristics are extracted from the PCG signals and are used to develop artificial intelligence algorithms. Using these features, PCG signals from 94 human subjects, collected from Texas heart institute and Biosignetics Corporation, were classified using a neural network utilizing Back Propagation Network (BPN) and Radial Basis Function (RBF) network algorithms to assess the diagnosis predictability of the developed algorithm. Among these subjects, 32 are diagnosed

978-1-4673-0818-2/12/$31.00 ©2012 IEEE

with mitral regurgitation (disease-1), 31 with coarctation of the aorta (disease-2), and 31 with mitral stenosis (disease-3). Out of the 94 signals, 66 were used for training, 5 for validation and 23 for testing.

II. Feature extraction

The sampling rate of the signals was 440Hz. Filtering is done using Matlab Wavelet Tool Box. The parameters chosen for all signals in this work are Db5 wavelet with decomposition level 5 and soft thresholding with "rigrsure" selection scheme.

The objective of the feature extraction procedure is to find the features from the available preprocessed data and use them later for classification. The analysis of heart sound is difficult to perform in the time domain because of noise interference and overlapping of heart sound components. Therefore, it is recommended, instead, to process heart sound signals in the frequency domain. Since sound signals may have different ranges of amplitudes, normalization is used to make all the signals to be restricted to the same range. In our case all signals amplitudes are normalized to the range 0 to 1.

There are a number of feature extraction algorithms available; among these algorithms are the Linear Frequency Band Cepstral (LFBC), the Heart Sound Segmentation (HSS), the Mel Frequency Cepstrum Coefficients (MFCC) and the Discrete Wavelet Transform (DWT) coefficients. However, most of the classification algorithms use wavelet transform and the segmentation algorithms. The four features extracted in this work are (a) activity or variance, (b) mobility, (c) complexity or form factor, and (d) the number of peaks from the frequency domain plot of the Power Spectral Density (PSD). Features a, b and c are collectively known as Hjorth descriptors.

The features extracted are characteristic to the signal itself and they do not need to be changed to any other form, as normally done with features extracted using transform coefficients. The number of peaks is calculated from the PSD plot using 'findpeaks' function in MATLAB. PSD is computed using the FFT (Fast Fourier Transform) method. Since the number of peaks can be very large, the raw data may not be suitable for use as inputs to the neural networks. To solve this problem, a grading system depending on the number of peaks for each system is developed. The grading system shown in TABLE I is used in this study.

TABLE I
GRADING SYSTEM USED FOR THE PEAKS

Number of Peaks	Grading Value
0-20	0.1
20-40	0.2
40-60	0.3
60-80	0.4
80-100	0.5
100-120	0.6
120-140	0.7
140-160	0.8
160-180	0.9
180 and above	1.0

PSD plot corresponding to a sample belong to disease-2, coarctation of the aorta is shown in Fig. 1. Although the Figure identifies some peaks, the total number may exceed 200 peaks, and changes from one signal to the other. These computed graded values along with the Hjorth descriptors are saved in an MS EXCEL file to be fed as input parameters to the neural networks.

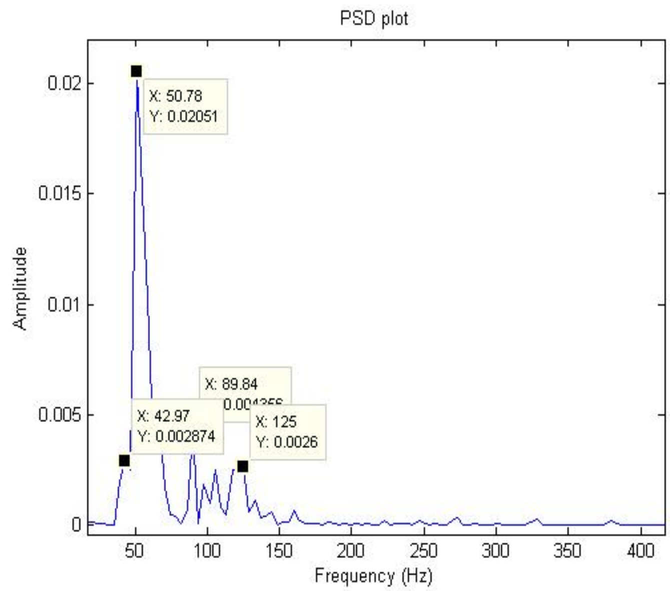

Fig.1. Power Spectral Density (PSD) plot showing the vectors of some of the peaks.

III. Features description

The three Hjorth descriptors are "Activity", "Mobility", and "Complexity" [7]. Activity is defined as the variance of the signal representing energy. Let's say we have a signal 'x' in the digital domain represented by $x(n)$ where 'n' represents the number of samples, the standard deviation of $x(n)$ is given by

$$\sigma_x = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i-\mu)^2}\ , \tag{1}$$

where N is the number of samples in the signal, μ is the mean of the signal. The variance is defined as the square of the standard deviation. Hence, the activity (A) is obtained from

$$A= \sigma_x^2. \tag{2}$$

Mobility is given by

$$M_x= \square_x{}'/\square_x\ , \tag{3}$$

where $\square_x{}'$ is the standard deviation of the first derivative of x, $x[n]'$, which can be obtained from

$$x[n]' = x[n]-x[n-1]. \tag{4}$$

Complexity (C) or the Form Factor (FF), which gives a computational value of the shape of the signal, is given by

$$FF = \frac{M_{x'}}{M_x} = \frac{\frac{\sigma_{x''}}{\sigma_{x'}}}{\frac{\sigma_{x'}}{\sigma_x}}, \tag{5}$$

where σ_x'' is the standard deviation of the second derivative of x, $x[n]''$, which is obtained from,

$$x[n]'' = x[n]-2x[n-1]+x[n-2]. \tag{6}$$

IV. Classification using Neural Networks

Classification is the process of assigning a label to an unknown pattern so that it is categorized into one of several known categories. The advantage of neural networks is that they provide a robust, general and practical method for learning real-valued, discrete-valued, or vector-valued functions from samples [1]. Two algorithms will be used, the back-propagation and the radial basis function networks.

A. Back-propagation Network (BPN)

BPN is a feed-forward network with three layers, namely input layer, hidden layer, and output layer, as shown in Fig. 2. The number of hidden layers can be more than one, depending on the complexity of the problem. In our study, we used one hidden layer to minimize the computational time and reduce complexity of training.

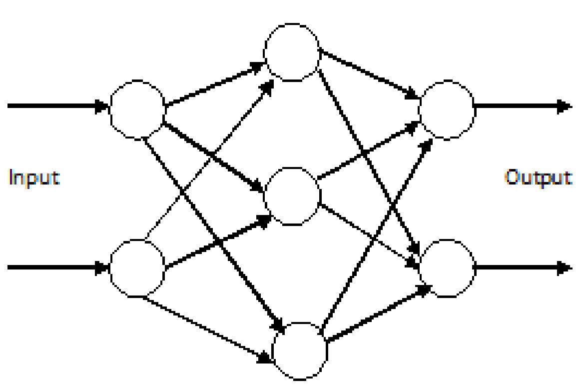

Fig.2.An example of a feed forward neural network

Determining the number of layers and the number of processing elements per layer are important decisions, which are made by the programmer while creating and training the network. Training inputs are applied to the input layer of the network, and the desired outputs are compared at the output layer. The difference between the output of the final layer and the desired output is back-propagated to the previous layer(s). The back-propagated signals are usually modified by the derivative of the transfer function and the connection weights, which are usually adjusted using the Delta Rule. The minimum mean square error between the actual output layer of the network and the desired output is minimized using the gradient descent algorithm. The performance of a neural network depends on the weights and the transfer function (input-output function) specified for the units. In this work, sigmoid function is used because of its similarities with the biological neuron. Since there are three classes of diseases, a state numerical code is assigned for each disease. The networks were trained to reproduce the related code at the outputs. The output can be left in decimal form, where the decimal numbers 1, 2, and 3 represent the classes, or codes in binary form are used instead. In our work, a binary classification system is used for the output layer, where the outputs are coded in binary form as: code 00, for disease 1; code 01 for disease 2; and code 10 for disease 3.

B. Radial Basis Function Network (RBF)

It is a three-layer network, namely the input, the output and the hidden layer, where each hidden unit in a hidden layer implements a radial activated function; see Fig. 3. The main advantages of RBF's over feed-forward networks are its accuracy and shorter computational time. As Venkatesan and Anitha [8] explained, the response of the jth-hidden unit can be mathematically expressed as

$$Z_J = \emptyset\left[\left|\frac{X-\mu_j}{\sigma_j^2}\right|\right], \tag{7}$$

where $\emptyset$ is a strictly positive, radially symmetrical function (kernel) with a unique maxima at its center, μ_j, and σ_j is the width of the receptive field. The error between the target and the desired output is minimized using gradient descent algorithm.

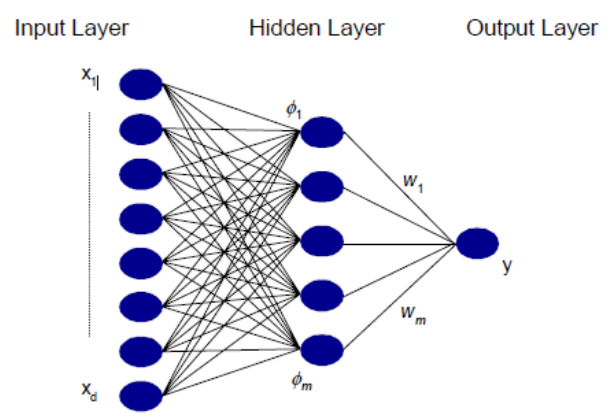

Fig.3. RBF network architecture.

V. RESULTS

The samples for each disease are denoted as S1, S2,.. Sn, where n is the number of samples for the disease. The features are referred to as F1 for Activity, F2 for Complexity, F3 for Mobility, and F4 for the graded number of peaks. These features are shown in TABLE II, TABLE III, and TABLE IV, for disease 1, disease 2, and disease 3, respectively.

TABLE II
FEATURE VALUES FOR DISEASE 1.

SAMPLE	F1	F2	F3	F4
S1	0.15	0.44	0.22	0.7
S2	0.1	0.48	0.23	0.8
S3	0.135	0.7	0.225	0.4
S4	0.148	0.73	0.24	0.2
S5	0.09	0.74	0.257	0.2
S6	0.2	0.68	0.18	0.8
S7	0.19	0.67	0.158	0.9
S8	0.18	0.74	0.19	0.1
S9	0.17	0.75	0.16	0.5
S11	0.165	0.76	0.17	0.6
S12	0.168	0.74	0.148	0.6
S13	0.2	0.78	0.35	0.7
S14	0.158	0.77	0.48	0.8
S15	0.165	0.72	0.37	0.8
S16	0.17	0.16	0.34	0.2
S17	0.19	0.17	0.4	0.3
S18	0.18	0.18	0.41	0.1
S19	0.4	0.19	0.79	0.4
S20	0.39	0.54	0.717	0.3
S21	0.43	0.219	0.67	0.2
S22	0.45	0.21	0.676	0.2
S23	0.274	0.67	0.42	0.4
S24	0.42	0.72	0.31	0.7
S25	0.23	0.57	0.18	0.6
S26	0.07	0.23	0.32	0.3
S27	0.12	0.42	0.47	0.8
S28	0.34	0.78	0.71	0.6
S29	0.61	0.62	0.64	0.5
S30	0.52	0.28	0.51	0.4
S31	0.081	0.17	0.18	0.2
S32	0.31	0.2	0.39	0.3

Back-propagation neural network is implemented using neural network toolbox in MATLAB. The training state parameters of the network are shown in Fig. 4. The Figure shows the gradient value at the end of the training and the validation checks plot of the Back-propagation networks. In this plot, the gradient value starts at 1, decreases and stays around 0.01, and finally reaches 0.04 when the training is stopped. Low value of gradient plot indicates that the network is learning and only finer adjustments in the weights are taking place. In other words, the network is avoiding chances of false predictions because it is becoming more accurate and reliable. Validation plot shows the point where the network learned sufficiently and passed validation.

The performance plot of the Back-propagation network is shown in Fig. 5. The performance of a network is measured in terms of mean square error (MSE).

TABLE III
FEATURE VALUES FOR DISEASE 2.

SAMPLE	F1	F2	F3	F4
S1	0.2	0.89	0.24	0.4
S2	0.21	0.65	0.25	0.3
S3	0.29	0.66	0.6	0.1
S4	0.276	0.68	0.59	0.3
S5	0.25	0.64	0.58	0.2
S6	0.17	0.6	0.56	0.2
S7	0.18	0.23	0.9	0.8
S8	0.75	0.25	0.95	0.8
S9	0.17	0.24	0.94	0.8
S11	0.18	0.23	0.91	0.8
S12	0.24	0.22	0.65	0.3
S13	0.21	0.2	0.59	0.1
S14	0.6	0.224	0.2	0.5
S15	0.62	0.2356	0.17	0.6
S16	0.63	0.19	0.22	0.3
S17	0.64	0.17	0.19	0.4
S18	0.59	0.16	0.18	0.2
S19	0.9	0.2	0.64	0.7
S20	0.81	0.17	0.62	0.4
S21	0.86	0.16	0.67	0.2
S22	0.612	0.49	0.52	0.3
S23	0.72	0.23	0.39	0.5
S24	0.82	0.19	0.47	0.7
S25	0.21	0.31	0.2	0.4
S26	0.37	0.42	0.01	0.1
S27	0.2	0.57	0.92	0.9
S28	0.16	0.82	0.82	0.7
S29	0.42	0.1	0.78	0.5
S30	0.31	0.32	0.48	0.2
S31	0.87	0.44	0.62	0.3

TABLE IV
FEATURE VALUES FOR DISEASE 3.

SAMPLE	F1	F2	F3	F4
S1	0.7	0.1	0.62	0.3
S2	0.71	0.16	0.83	0.3
S3	0.73	0.15	0.84	0.2
S4	0.75	0.9	0.48	0.4
S5	0.48	0.8	0.49	0.6
S6	0.49	0.12	0.51	0.1
S7	0.53	0.7	0.52	0.4
S8	0.51	0.79	0.84	0.5
S9	0.58	0.8	0.54	0.2
S11	0.6	0.68	0.12	0.8
S12	0.9	0.66	0.1	0.7
S13	0.89	0.45	0.11	0.7
S14	0.94	0.48	0.13	0.4
S15	0.92	0.52	0.2	0.3
S16	0.91	0.51	0.87	0.2
S17	0.55	0.54	0.88	0.4
S18	0.43	0.47	0.91	0.3
S19	0.49	0.51	0.32	0.2
S20	0.59	0.53	0.43	0.5
S21	0.52	0.6	0.63	0.4
S22	0.49	0.61	0.58	0.8
S23	0.72	0.62	0.27	0.4
S24	0.58	0.72	0.29	0.4
S25	0.47	0.47	0.35	0.3
S26	0.56	0.58	0.41	0.6
S27	0.81	0.39	0.46	0.8
S28	0.74	0.6	0.81	0.2
S29	0.92	0.83	0.69	0.3
S30	0.84	0.57	0.16	0.4
S31	0.52	0.61	0.73	0.5

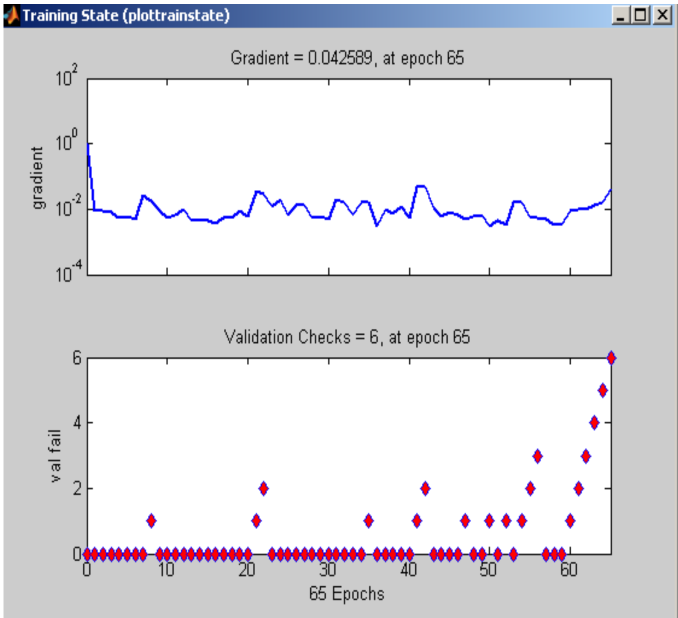

Fig.4. Training state parameters of the network

From Fig. 5, we can observe that the MSE was rapidly decreasing as the network was being trained at the epoch number 59. MSE is the average squared error between targets and actual outputs, which should be kept to a minimum. A value of zero MSE means there is no error. The mean square error value is found to be 0.041. The figure shows that the blue curve (training), the green curve (validation) and red curve (testing) are almost the same, which means the network was trained well with optimum performance.

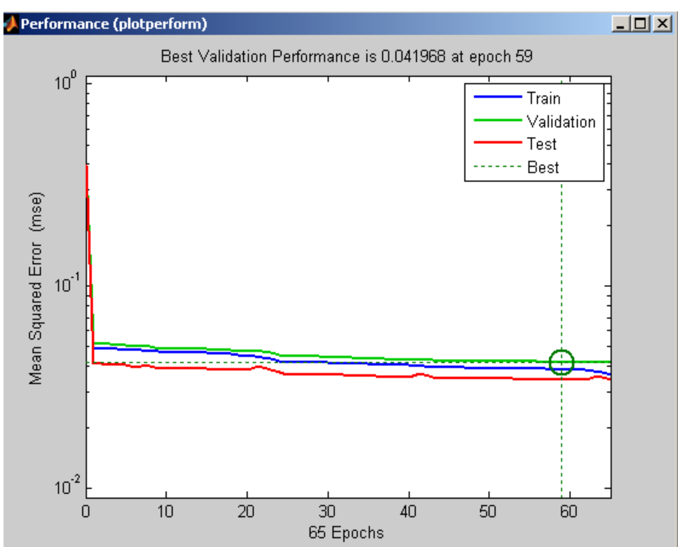

Fig.5. Performance plot of the network after training

Receiver operating characteristic (ROC) curve is a common metric of measuring accuracy of the classifier. The ROC curve for the overall data using BPN algorithm is shown below in Fig.6. For RBF, ROC curves are plotted in Fig.7. In figures 6 and 7, the blue, red and green curves represent the training, testing and validation data. The percentage accuracy is found by calculating the area under curve (AUC) using a trapezoidal rule. From Fig.6, the area under the curve for training data is computed and found to be 0.908. This means that the algorithm has 90.8 % accuracy.

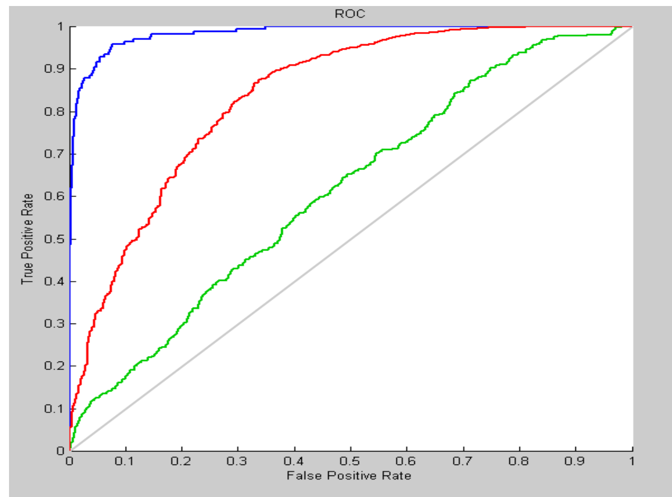

Fig.6. ROC plot for overall data using BPN networks.

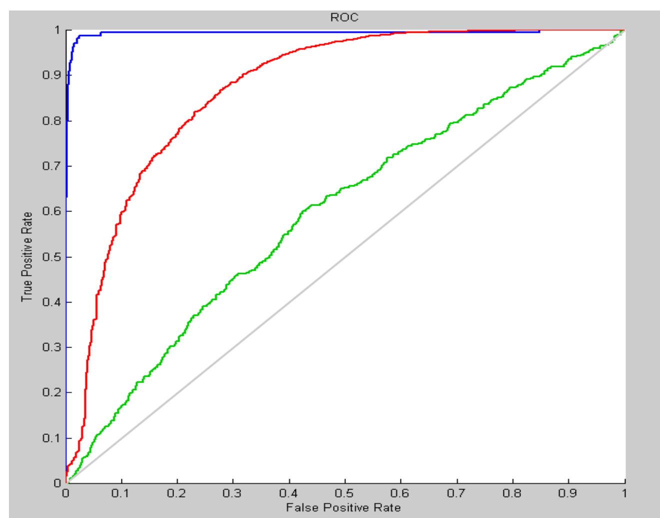

Fig.7. ROC plot for overall data using RBF networks.

On the other hand, the area under ROC training curve for RBF network is found to be 0.98, Fig. 7. This means that this network can provide 98% accuracy. Therefore, in this case of classification problem, RBF networks outperform BPN algorithm networks.

CONCLUSIONS

This work preprocesses diseased heart sound signal from 94 human subjects. Among these subjects, 32 are diagnosed with mitral regurgitation (disease-1), 31 with coarctation of the aorta (disease-2), and 31 with mitral stenosis (disease-3). Out of the 94 signals, 66 were used for training, 5 for validation and 23 for testing. The prepossessing was performed using Wavelet transform. Four independent feature characteristics related to the PCG signals are extracted. These features are fed as inputs to two neural networks: the traditional back-propagation network algorithm and the radial basis functions network algorithm. The two networks were trained using the 66 samples. The networks were tested using 23 samples for the three different diseases. The testing results show that the performance of RBF networks is superior when compared to the traditional BPN networks with 98% accuracy compared with 90.8% for the BPN.

AKNOWLEDGMENT

This work was supported in part by the College of Engineering and Engineering Technology at Northern Illinois University.

REFRENCES

[1] S. S. Tripathy. “System for diagnosing valvular heart disease using heart sounds”, Master’s Thesis, India, 2005

[2] A. S. A. Mohamed and H. M. Raafat. “Recognition of heart sounds and murmurs for cardiac diagnosis”. In Proceedings of 9th International Conference on Pattern Recognition, 1988.

[3] S. B .Patil and Y . S. Kumaraswamy. “ Intelligent and Effective Heart Attack Prediction System Using Data Mining and Artificial Neural Network”, European Journal of Scientific Research, 2009, ISSN 1450-216X Vol.31 No.4. pp.642-656.

[4] C. N. Gupta, Ramaswamy Palaniappan ,Sundaram Swaminathan , Shankar M. Krishnan. “Neural network classification of homomorphic segmented heart sounds “, Applied Soft Computing (7), 2007, pp. 286–297.

[5] T. R. Reed., N. E Reed. and P. Fritzson, “The analysis of heart sounds for symptom detection and machine-aided diagnosis”. In Proceedings of the 4th International EUROSIM conference, 2001,pp. 1-6.

[6] F. Yaghouby, A. Ayatollahi and R. Soleimani, Classification of Cardiac Abnormalities Using Reduced Features of Heart Rate Variability Signal”. World Applied Sciences Journal 6 (11), 2009, pp. 1547-1554.

[7] Ramaswamy Palaniappan,” Biological Signal Analysis”, Ramaswamy Palaniappan and Ventus Publishing Aps, U.K, 2010.

[8] P. Venkatesan* and S. Anitha. Application of a radial basis function neural network for diagnosis of diabetes mellitus. Current science, vol. 91, no. 9, 10 November, 2006, pp. 1195-1199.