ECG classification and abnormality detection using cascade forward neural network

Article *in* International Journal of Engineering Science and Technology · July 2011

DOI: 10.4314/ijest.v3i3.68420

2 authors:

Shahanaz Ayub
Bundelkhand Institute of Engineering and Technology

Prabhjot Singh Saini
Narsee Monjee Institute of Management Studies

MultiCraft

International Journal of Engineering, Science and Technology
Vol. 3, No. 3, 2011, pp. 41-46

INTERNATIONAL JOURNAL OF ENGINEERING, SCIENCE AND TECHNOLOGY
www.ijest-ng.com

ECG classification and abnormality detection using cascade forward neural network

Shahanaz Ayub[1], J.P. Saini[2]

[1]Department of Electronics and Communication Engineering, BIET, Jhansi, INDIA*
[2] MMMEC, Gorakhpur, UP, INDIA
[*]Corresponding Author: e-mail: shahanaz_ayub@rediffmail.com Tel +91-9415587596

Abstract

Electrical activity of the heart is called as electrocardiogram i.e. ECG. Arrhythmias are among the most common ECG abnormalities. ECGs provide lots f information about heart abnormalities. The diagnosis depends upon the physician and it varies from physician to physician and also depends upon the experience of the physician. Previously many techniques were tried for analysis and automisation of the analysis. This paper describes the use of MATLAB based artificial neural network tools for ECG analysis for finding out whether the ECG is normal or abnormal and if it is abnormal, what is the abnormality. There are various arrhythmia like Ventricular premature beats, asystole, couplet, bigeminy, fusion beats etc. To classify this, various weighted neural networks were tried with different algorithms. They were provided training inputs from the standard MIT-BIH Arrhythmia database and tested by providing unknown patient data from the same database. The results obtained with different networks and different algorithms are compared, it is found that to identify whether the ECG beat is normal or abnormal, cascade forward back network algorithm has shown 99.9 % correct classification. These results are compared with previous neural network techniques and found that method proposed in this paper gives best results.

Keywords: Arrhythmia, MATLAB, Artificial Neural Networks, Back propagation, Cascade- Forward Network, MIT-BIH arrhythmia data base.

1. Introduction

Cardiac problems are increasing day by day. ECG is one of the most commonly used tests to diagnose the heart problem. Detection and treatment of arrhythmias has become one of the cardiac care unit's major functions. Few of the arrhythmias are Ventricular Premature Beats, asystole, Couplet, Bigeminy, Fusion beats. In Ventricular Premature Beats (VPB), there is premature ventricular contraction arising in diastolic period of the preceding sinus beat followed by compensatory pause. Couplet is the case where pair of VPBs are observed. In Asystole, there is a lack of conduction observed for an extended duration. Bigeminy is the presence of VPB between alternate normal beat. Fusion beat is a parasystolic condition in which two pacemakers in heart discharge at their own inherent rate, occasionally causing simultaneous invasion of ventricular musculature, each activating part of ventricles. The resulting QRS complex has a configuration intermediate below 'pure' sinus beat and pure ventricular beat. More than 3 million ECGs are taken (*Kerala SCERT,2006*) worldwide each year for the patients with different cases. All the samples taken have one thing in common and that is, they are analysed by the experienced doctors who depending upon their knowledge predict out the problem(s) associated with the patient. If this morphological disturbance in ECG becomes somewhat complex (such as the case of fusion beats) then it is analysed by them depending upon their experience. This experience based analysis gives different interpretations. Hence there is a need of a system that could analyse the ECG signals properly and with a great accuracy so that there is a less chance of mistake as well as the problem is spotted in time so that an early treatment could be started.

So to achieve this objective many works have been done in this field based on image processing, Digital Signal Processing etc and prominent among them is the use of Artificial Neural Networks (*Zurada, 1999*) which has given promising results to such complex problems. Neural network based analyses made were either weight based or weightless. This work is based on weighted neurons with bias adjustments but with the application of MATLAB based algorithms and neural network structure.

The excellent features of the MATLAB (Gilat., 2005) such as wide range of tools for network structure development and adjustment according to requirements as well as tools to analyse the results, makes it a good option to solve this complex problem in a simple way, especially the case of fusion beats. . In this paper the case of ECG beat whether it is normal or abnormal is discussed so as to have an insight into the concept of identification of normal beats using cascade forward neural networks (MATLAB based) with back propagation algorithm.. The data base used in this paper to train and test the neural network, is the standard MIT-BIH arrhythmia database (*Brown.,2006*).

The objective of this work is to make the analysis of ECG beats as normal or abnormal and if it is abnormal what is the abnormality. The analysis is done so that the patient could be diagnosed for the heart problems in less time as well more accurately so that the medical practitioners have primary information about the ailment and could start a treatment early. Apart from this the project has been targeted towards the rural community and so we are also considering hardware implementation of this work but in low cost and greater efficiency. In this paper MATLAB based neural network tools are used to detect the abnormality. The network needs to train with sufficient number of samples and also needs to test with sufficient number of samples.

2. Methodology

The database provided by the MIT-BIH arrhythmia database (*Brown.,2006*) regarding different kinds of heart rhythm abnormalities for different class of patients, is the source of data used for training, testing and validation of the neural networks. The most of the data is taken from the patient number 208. But as it was not sufficient to train the network, other patients data were also taken so as to enhance the prediction capability of the trained neural network and make it more accurate .Also various patients data from this database has different arrhythmia cases.

The data taken was used to make training inputs which represented the whole ECG cycle as well as for making test inputs. The MATLAB based Perceptron and back propagation (*Demuth et al.,2008*) networks were developed and training parameters were fixed for certain quantities and varied for others. The network trained were analysed using the test inputs first for all unknowns which were not used for training and then for all inputs which included both training inputs and test inputs. Analysis plot tools (Demuth et al.,2008,2004) were also used to understand the network capability and other properties such as Mean Squared Error (MSE) value and learning capability.

3. Database

The MIT-BIH Arrhythmia Database contains 48 half-hour excerpts of two-channel ambulatory ECG recordings, obtained from 47 subjects studied by the BIH Arrhythmia Laboratory between 1975 and 1979. The recordings were digitized at 360 samples per second per channel with 11-bit resolution over a 10 mV range. Two or more cardiologists independently annotated each record; disagreements were resolved to obtain the computer-readable reference annotations for each beat (approximately 110,000 annotations in all) included with the database. For our study purpose the record chosen was the MIT-BIH patient 208. This was chosen for two reasons:

1) The patient was without any kind of medication and had sufficient beats for normal and ventricular as well as maximum number of beats for fusion analysis among all patient data available.
2) Earlier weightless neural network analyses were done by other scholars using this patient so we wanted to compare our network and its result with their one.

The number of samples needed to train the network for other arrhythmia cases like Ventricular premature beats, fusion beats etc., were taken from other patients data as only 208 patient's data was not sufficient for analysis using neural networks. Figure 1. shows the sample record of data of patient showing different abnormality beats.

Figure 1. A sample record of data of patient showing different abnormal beats as V & F along with normal beats as N.

V stands for ventricular premature beat and F stands for fusion beat.

4. Inputs

In this paper one ECG beat corresponds to one sample of 301 inputs, which covers the whole ECG beat. The inputs for the networks were selected considering two important points:

a) The inputs must be of a standard size such that it is neither too small to cover up one ECG cycle and nor two high to increase the number of beats required to analyse the signal, thus increasing the hardware requirements.

b) Secondly the input must be so arranged that the R peak in the QRS complex must be at the centre of the signal cycle under considerations.

The first condition was achieved by setting up an arbitrary value of 301 samples of MLII lead data (*Schreck 2005*) obtained from the database in which the 150 samples were on the left side and 150 samples on the right side of the 151st sample value. From the database when the ventricular beats were derived we got the R peak values in the table form. This R peak was taken as centre and from the samples of the same patient 150 sample before this R peak value and 150 samples after it was taken to make a 301 sample input, where the ventricular beat was in the centre. Thus the input becomes a matrix of 301x1 and ready to be used in MATLAB. The same process was repeated to make all the inputs of all the kinds of beats that are normal, fusion and ventricular premature.

The second condition was achieved by allowing the 151^{st} sample to be the beat value of MLII lead signal obtained from the database for particular conditions. E.g. If a umber (2250-150=) 2100 to sample number (2251+150=) 2400 will be the input data. Table 1 shows the organization of the data.

Table 1. Organisation of the data

Rank of sample in a beat	No.of sample in a beat	ML II lead value of Sample in that beat
1st	2100	930
2nd	2101	970
3rd	2102	978
| | |	| | |	| | |
151st	2250	1264
| | |	| | |	| | |
300th	2399	955
301st	2400	920

Table 2 shows the input matrix structure. Total 301 input points are there. Total samples for normal beats are 1591. Out of which 1285 are taken for training, 300 for testing and 6 for validation.

Table 2. Input Matrix Structure

Kind	Inputs	Matrix Dimension	Training	Testing	Validation
Normal	1591	301 X 1591	301 X 1285	301 X 300	301 X 6
Fusion Beats	740	301 X 740	301 X 584	301 X 150	301 X 6
Ventricular Premature Beats	996	301 X 996	301 X 796	301 X 120	301 X 6
Unclassified (Abnormal)	260	301 X 260	301 X 180	301 X 80	-

5. Analysis

The following table provides the results related to various beats analysed using the best case of cascade- forward back propagation network, among them the results for the normal beats should be observed carefully.

Table 3. Cascade Forward Network Design Analysis Results, ‘ TRAINBFG’ : Training Algorithm

Inputs	1285 N / 3587 T	1285 N / 3587 T
HN	5	10
Time in Sec	438	3033
P(cc) Unknown	99.3 V, 99.9 N, 94F, 100U	99.8V, 99.9N, 94F, 100 U
P(cc) All	99.7 V, 99.9N, 94F, 99.2 U	99.7 V, 99.9N, 94.3F, 97 U
Ephochs (Max=1000)	34	34
MSE 0.0001	0.00642	0.00621
Result	Passed	Passed

TRAINBFG - BGFS Quasi Newton Back propagation training algorithm
P (CC) - Percentage of correct classifications
INPUTS- total number of samples is suffixed by T
HN- Hidden neurons, representing the number of neurons in the Hidden layer
MSE- Mean Squared Error; the error goal was fixed at 0.0001 and hence here the difference MSE-0.0001 is being tabulated
SUFFIXES V, N, F, AND U- in the P (cc) columns indicate respectively Ventricular premature, normal, fusion and unclassified beats and their percentage of correct classifications.

The following images are the analysis plots for this work where each of them interprets different properties about the network

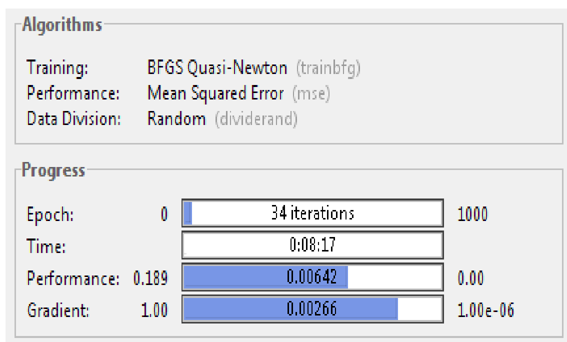

Figure 2. Training Process Results

Fewer epochs mean network learns in small repetitions. Less time means network achieved goal easily and shortly. Performance indicates the final MSE achieved. Lower value is associated with higher network accuracy.

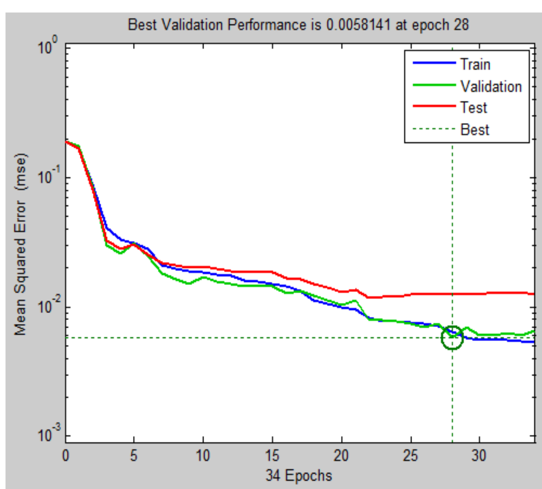

Figure 3. Mean Squared Error (MSE) Plot

Mean squared error plot shows the achieved error value. Lower value means the less probability of false predictions. Here the network has achieved quite low error probability.

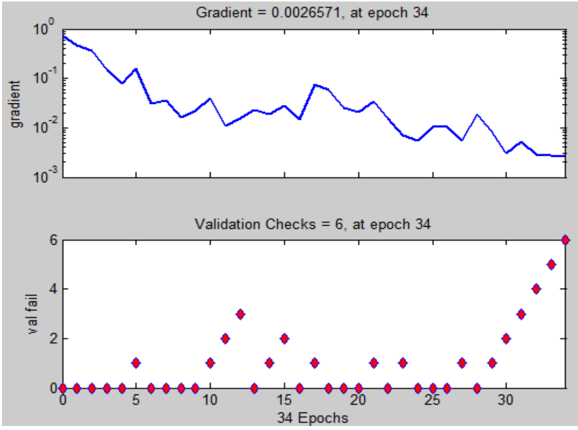

Figure 4. Gradient & Validation Check Plots

Low value of gradient plot indicates that the network is learning up to a large extent which means finer adjustments in the weights and bias. This in turn makes network more accurate and reliable, avoiding chances of false predictions. Validation plot shows the point where the network learned sufficiently and passed validation without. The point where the failures cross the defined limit is the stoppage point of training and indicates the starting of the over fitting data.

6. Inferences

i. This Network gives quite low value of MSE and it is near 0.00642 in just 34 epochs. Time taken for training may be large but it is giving good results.
ii. Though it was also tried that the trainlm case also gives good result but trainbfg case is used because of higher accuracy .

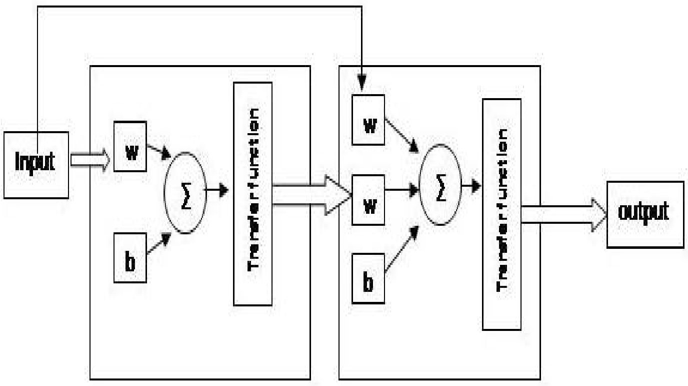

Figure 5. Cascade Forward Network General Structure

7. Conclusion

The same network is tried with different number of training input sizes, but the time taken to train the network differs. The network based on Cascade-Forward network algorithm with trainbfg training algorithm was best for the case of normal beat analysis because it has an accuracy of about 99.9% as well as the memory requirements were also low. Hence we preferred this network for the normal beat analysis. The conclusion derived from this work is that, by using the MATLAB based neural network design (*Demuth et al.,2008*); such networks can be made which have capability to understand different class of inputs when they are fed to be analysed. Such networks can be very reliable as MATLAB provides a good set of tools so that the network parameters can be adjusted easily and precisely by just adjusting values for them and change in full length code, as was done previously, is not required. Though the objective of this research was not to use MATLAB or Neural Networks, these were used to get higher accuracy in analysis of ECG which is more useful for the mankind.

The results obtained with other methods like weightless neural networks, MLP (*Chickh et al.,2002, Thomson et al., 1993.,Gao et al., 2003., Chow et al., 1993.,Nadal et al.,1993*) etc are compared with our results. Table 4 shows the comparison of the results.

Table 4. Comparative Results with Other Methods

Methods	% OF CORRECT CLASSIFICATION Normal Beats	Remark
MLP	98.85	Comparatively Lowest accuracy
HFNS	98.78	2 ECG leads are used as inputs
PCA	98.80	Same samples are used for training and Testing
Weightless	99.69	Comparatively low accuracy
Our Method	99.9	Best Accuracy

The networks are not tested with the current real patients data. But it will give same higher accuracy, as network is trained and tested with sufficient number of inputs. Again all the arrhythmias are not classified with the name, it is still mentioned as unclassified which is next step of this research.

References

Brown G.,2006, *MIT-BIH Arrhythmia database*,MIT

Chickh M. A.N. Belgacem, F. Bereksi-Reguig ,2002,Neural classifier to classify ectopic beats. Acte des IX emes rencontre de la *Societe Francophone de Classification*, Toulouse, le 16-18.

Chow H.S,. Moody G. B., and Mark R.G.1993, Detection of ventricular ectopic beats using neural networks, *Computers in Cardiology*, pp 659-662.

Demuth Horward, Beale Mark, Hagan Martin, 2008,MATLAB Neural Network Toolbox, MATHWORKS INC., *MATLAB* Version R2008b, October .

Demuth Horward, Beale Mark, Hagan Martin, 2004, MATLAB Simulink, MATHWORKS INC., *MATLAB* Version 6, October.

ECG & AMT Teacher's Source Book,2006, Government of Kerala, Department of Education, *State Council of* Educational *Research & Training (SCERT)*, pp.23.

Gao D, Kinouchi Y, Ito K, Zhao X, 2003,Time series identification and modeling with neural network. In. Pro. *IEEE Int. Joint conference on Neural Networks*, Portland, USA, 2454-59.

Gao D, Kinouchi Y, Ito K, Zhao X, 2003, Neural Networks for Event extraction from Time Series, A back propagation Algorithm Approach, Future generation , *Computer System*

Gilat Amos,2005, MATLAB- An Introduction with Applications, *Wiley & Sons publications INC*.

Nadal J. and Bosson M.C.,1993, Classification of cardiac arrhythmias based on principal component analysis and feedforward neural networks, *Computers m Cardiology*,pp 341- 344.

Schreck D. M.,2005, System and method for synthesizing leads of an Electrocardiogram *US patent*, 6, 901, 285.

Thomson D. C., Soraghan J. J, and Durrani T.S, 1993,An artificial neural-network based SVT/VT classification system, *Computers in Cardiology*,(pp 333-336.

Zurada, 1999, Introduction to Artificial Neural Systems, *West Publishing House*.

Biographical notes

Dr. Shahanaz Ayub is an Associate Professor in Electronics & communication Engineering department of Bundelkhand Institute of Engineering & technology (BIET) Jhansi, UP, India, currently holding the position of Head of the department. She has twelve years of work experience as Industrial, Research & Teaching. Her area of specialization is biomedical Signal Simulation and analysis, Antenna, Optical & photonic waveguide. She has a wide experience in microcontrollers and uses the artificial intelligence techniques in her research. She has published thirteen research papers in international journals and presented papers in twenty three national & international journals.

Dr.J. P. Saini is a Professor in the Department of Electronics and Communication Engineering, Bundelkhand Institute of Engineering and Technology Jhansi, India. He has more than 25 years of experience in teaching, research and administrative work. His current area of research includes Multi-criteria Decision-Making, Neural Networks, and Optical & Photonic waveguides. He has published more than twenty five papers in referred international journals. He has also presented more than forty research articles in national and international conferences. He is currently working as a Principal of Madan Mohan Malaviya Engineering College, Gorakhpur, India on deputation.

Received January 2011
Accepted March 2011
Final acceptance in revised form April 2011