

Classification of ECG Signals using ANN with Resilient Back Propagation Algorithm

G. Subramanya Nayak
Department of EC Engineering
MIT, Manipal

Dayananda Nayak
Department of EC Engineering
MIT, Manipal

ABSTRACT

Electrocardiogram is one important physiological signal, which is used in assessing cardiac health. The extraction of features used for identification of the state of ECG is discussed in this paper. Using MAT LAB programs/tools, different statistical features are extracted from both normal and arrhythmia spectra. These features include arithmetic mean, median, variance, residuals on curve fitting etc. The values of the feature vector reveal information regarding cardiac health state. Then a classical multilayer feed forward neural network with back propagation algorithm is employed to serve as a classifier of the feature vector, giving 100% successful results for the specific data set considered.

Keywords

Electrocardiogram, Back propagation algorithm, Neural Network

1. INTRODUCTION

There are many instances in which it is necessary to monitor and transmit physiological events from a distance for convenient Biotelemetry. Electrocardiogram is one important physiological signal, which is used in assessing cardiac health. In general the cardiologist looks critically at the various time intervals, polarities and amplitudes of the ECG to arrive at the diagnosis. Different features are extracted from these ECG spectra. This feature vector for both normal and arrhythmia spectra can be used for classification using ANN. PCA technique can also be implemented on the same data and a comparative study on performance of both ANN and PCA can be tabulated.

The proposed study is done on MIT arrhythmia ECG database, sampled at 360 samples per second. As the frame size is 400 samples, this method achieves a bit rate of 45bits per second, which is very much less in comparison of other methods. A large amount of data produced by ECG monitoring and recording facilities needs to be analyzed for classification of normal and arrhythmia conditions.

For further analysis using Matlab@6.1 algorithms, these data are recorded as sample points. Using MATLAB programs/tools, features are extracted for different spectra.

These features include curve fitting and residuals, arithmetic mean, median, variance etc. The features like 10th degree polynomial curve are fitted onto ECG spectra and residuals values are plotted as shown in Figure 1.

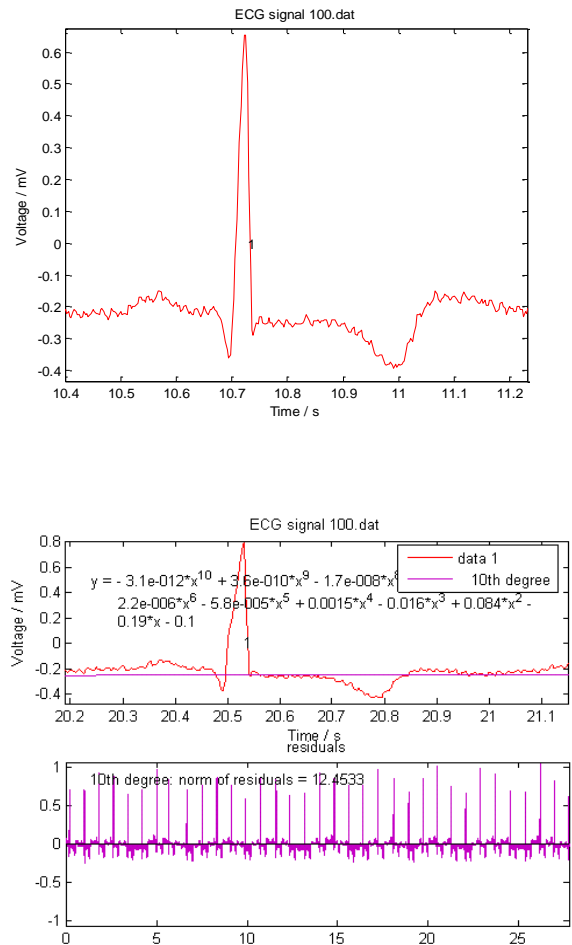


Figure 1: Curve fitting onto ECG Spectra

2. THE ANN DIAGNOSTIC SYSTEM ARCHITECTURE

An ANN structure is employed for the classification of spectra. It consists of four modules as shown in Figure 2.

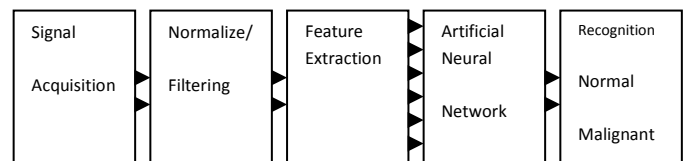


Figure 2. The model of ANN scheme used to perform diagnosis on human oral tissue samples.

The signals are filtered with 21 ordered median filters to eliminate undesirable spikes due to noise and other disturbances. The signal is then fed to the feature extraction module and different features were extracted.

3. FEATURE EXTRACTION

The statistical analysis and classification for discrimination among normal and pathology conditions were performed using MATLAB on the set of spectral data. Using MATLAB, the following features were extracted: residuals on curve fitting, arithmetic mean, median, variance, standard deviation, RMS (Root Mean Square), Summation, Skewness, Kurtosis. A brief description of each of these features is given below.

A. Arithmetic Mean

In simple terms, the arithmetic mean of a list of numbers is the sum of all the members of the list divided by the number of items in the list. If a set of data is denoted by $X = (x_1, x_2, \dots, x_n)$, then the sample mean is typically denoted with a horizontal bar over the variable

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{n} (x_1 + \dots + x_n)$$

B. Median

A median is described as the number separating the higher half of a sample, a population, or a probability distribution, from the lower half. The median of a finite list of numbers can be found by arranging all the observations from lowest value to highest value and picking the middle one. If there is an even number of observations, the median is not unique, so one often takes the mean of the two middle values. At most half the population has values less than the median and at most half have values greater than the median.

C. Variance

The variance of a sample is one measure of statistical dispersion, averaging the squared distance of its possible values from the expected value (mean). Whereas the mean is a way to describe the location of a distribution, the variance is a way to capture its scale or degree of being spread out. The unit of variance is the square of the unit of the original variable. The positive square root of the variance, called the standard deviation, has the same units as the original variable and can be easier to interpret for this reason. If random variable X has expected value (mean) $\mu = E(X)$, then the variance $\text{Var}(X)$ of X is given by:

$$\text{Var}(X) = E[(X - \mu)^2]$$

D. Standard deviation

The standard deviation of a multiple set of values is a measure of statistical dispersion of its values. The standard deviation is usually denoted with the letter σ . It is defined as the square root of the variance.

$$\sigma = \sqrt{E((X - E(X))^2)} = \sqrt{E(X^2) - (E(X))^2}$$

where $E(X)$ is the expected value of X . Standard deviation, being the square root of variance, measures the spread of data about the mean, measured in the same units as the data.

E. RMS

The root mean square is a statistical measure of the magnitude of a varying quantity. It is especially useful when variants are positive and negative, e.g. sinusoids. It can be calculated for a series of discrete values or for a continuously varying function. The RMS of a collection of n values is

The RMS of a periodic function is equal to the RMS of one period of the function. The RMS value of a continuous function or signal can be approximated by taking the RMS of a series of equally spaced samples.

F. Summation

Summation is the addition of a set of numbers; the result is their sum or total. The "numbers" to be summed may be

$$x_{\text{rms}} = \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2} = \sqrt{\frac{x_1^2 + x_2^2 + \dots + x_n^2}{n}}$$

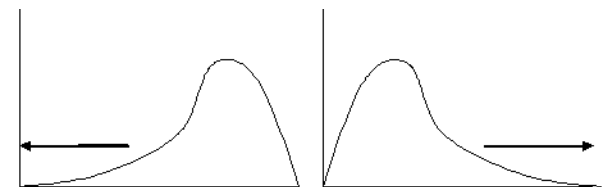
natural numbers, complex numbers, matrices, or signals. An infinite sum is a subtle procedure known as a series. Summation can also be represented as

G. Skewness

It is a measure of the asymmetry of the probability distribution of a real-valued random variable.

Considering the distribution in the above figure, the right side of the distribution tapers differently than the left one. These tapering sides are called tails, and they provide a visual means for determining which of the two kinds of skewness a

$$\sum_{i=m}^n x_i = x_m + x_{m+1} + x_{m+2} + \dots + x_{n-1} + x_n.$$



Negative Skew
Elongated tail at the **left**
More data in the left tail than would be expected in a normal distribution

Positive Skew
Elongated tail at the **right**
More data in the right tail than would be expected in a normal distribution

distribution has:

Negative skew (the left tail is longer): the mass of the distribution is concentrated on the right of the figure. The distribution is said to be left-skewed.

Positive skew (the right tail is longer): the mass of the distribution is concentrated on the left of the figure. The distribution is said to be right-skewed. Skewness, the third standard moment is written as γ_1 and defined as

$$\gamma_1 = \frac{\mu_3}{\sigma^3},$$

where μ_3 is the third moment about the mean and σ is the standard deviation. Equivalently, skewness can be defined as the ratio of the third cumulant κ_3 and the third power of the square root of the second cumulant κ_2 :

$$\gamma_1 = \frac{\kappa_3}{\kappa_2^{3/2}}$$

This is analogous to the definition of kurtosis, which is expressed as the fourth cumulant divided by the fourth power of the square root of the second cumulant.

For a sample of n values the skewness is

$$g_1 = \frac{m_3}{m_2^{3/2}} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3}{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2\right)^{3/2}},$$

where x_i is the i^{th} value, \bar{x} is the sample mean, m_3 is the sample third central moment, and m_2 is the sample variance..

H. Kurtosis

The fourth standardized moment is defined as

$$\frac{\mu_4}{\sigma^4},$$

where μ_4 is the fourth moment about the mean and σ is the standard deviation. This is sometimes used as the definition of kurtosis in older works, but is not the definition used here.

Kurtosis is more commonly defined as the fourth cumulant divided by the square of the second cumulant, which is equal to the fourth moment around the mean divided by the square of the variance of the probability distribution minus 3,

$$\gamma_2 = \frac{\kappa_4}{\kappa_2^2} = \frac{\mu_4}{\sigma^4} - 3,$$

which is known as “excess kurtosis”. The “minus 3” at the end of this formula is often explained as a correction to make the kurtosis of the normal distribution equal to zero.

For a sample of n values the sample kurtosis is

$$g_2 = \frac{m_4}{m_2^2} - 3 = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4}{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2\right)^2} - 3$$

where m_4 is the fourth sample moment about the mean, m_2 is the second sample moment about the mean (that is, the sample variance), x_i is the i^{th} value, and \bar{x} is the sample mean

4. ANN IMPLEMENTATION

A. Design of ANN with Back Propagation Algorithm:

Three layer networks are sufficient to design any nonlinear network. Input is a layer of nodes with three feature vectors as input. Single second and third layers have same activation function for all neurons. Activation function is a tan sigmoid function as shown in Figure 3.

Resilient back propagation algorithm, which is a standard Back propagation algorithm, is used for training.

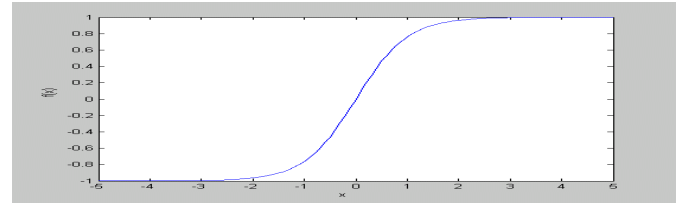


Figure 3. Activation function – Tan sigmoid

Single hidden layer is considered. An optimum number of seven neurons are considered in hidden layer and one neuron in output layer. The performance goal is set at 0.001 accuracy. Number of epochs is set for 5,000. The learning cycle updates the weights of output layer and progress backwards. Error function is MSE and updating in negative gradient.

B. Training of ANN

Data from normal and pathology were obtained from MIT. arrhythmia database of 80 normal and 80 arrhythmia are used to train the network.

Feature vector is constructed and training program is executed. The training of ANN stops once the performance goal is met, as shown in Figure 4. The convergence is achieved in 635 epochs.

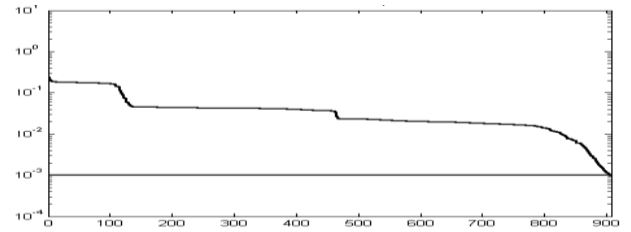


Figure 4. Training of ANN

5. RESULTS

Feature vector of 80 test data of Normal (N) and 80 test data of arrhythmia (P) are given to network to classify. The results are shown in Table 5.1.

TABLE 5.1
CLASSIFICATION RESULTS

Test data	Desired output	Classifier output	Output Characterization
N1	1	1.0000	Normal
N2	1	1.0000	Normal
N3	1	1.0000	Normal
N4	1	1.0000	Normal
N5-N15	1	1.0000	Normal
N16-N25	1	1.0000	Normal
N26-N35	1	1.0000	Normal
N36-N45	1	1.0000	Normal
N46-N55	1	1.0000	Normal
N56-N80	1	1.0000	Normal
P1	-1	-0.9025	Arrhythmia
P2 – P3	-1	-0.9008	Arrhythmia
P4	1	1.0000	Normal
P5 – P15	-1	-0.8070	Arrhythmia
P16-P25	-1	-0.1000	Arrhythmia
P26 –P35	-1	-0.8970	Arrhythmia
P36	1	1.0000	Normal
P37-P45	-1	-0.8033	Arrhythmia
P46-P54	-1	0.9776	Arrhythmia
P55	1	1.0000	Normal
P56-P80	-1	-0.9880	Arrhythmia

6. DISCUSSION

Our method achieves the discrimination between normal and arrhythmia ECG spectra. Furthermore, a small time needed to acquire and analyze the spectra together with the high rates of success, proves our method very attractive for real –time applications. The results of ANN has been verified with the results of PCA and found satisfactory. This gives us to firmly decide on the pathology. Further, performance parameters can be accurately judged if the tests are carried out on large number of spectra. As less number of features is used, computational delay in training the ANN is reduced. Gradient descent algorithm is used. But it is too slower and convergence depends on learning rate. Hence high performance algorithm (Levenberg – Marquardt algorithm) which operates in batch mode can be used for fast training. The performance parameters are shown in Table 6.1.

TABLE 6.1
PEFORMANCE PARAMETERS

Technique Used	Specificity %	Sensitivity %	Accuracy %
ANN	100	96.3	98

7. REFERENCES

- [1] Jacek M. Zurada , Introduction to Artificial Neural Systems,pp 25–89. Publisher-Pearson Education, 3rd Edition.
- [2] B.K.Manjunath, J.Kurien, C.Muralikrishna, Autofluorescence of Oral tissue for optical pathology in Oral malignancy, Journal of Photochemistry and Photobiology B. Biology 73 (2004) 49 -58.
- [3] George A.Rovithakis and Michail Maniadakis ,Artificial Neural Networks for discriminating Pathologic from Normal Peripheral Vascular Tissue,IEEE transactions on Biomedical Engineering ,vol 48,No 10, october 2001
- [4] Sigurdur Sigurdsson et ,al ,Detection of skin cancer by classification of Raman spectra, IEEE transactions on Biomedical Engineering ,2004.
- [5] Simon Hykin,Neural Networks-A Comprehensive foundation,section 4.1to4.10,page 156-234,Jaico Publishing House,5TH edition.
- [6] William Palm J, “Introduction to MATLAB 6 for Engineers”.
- [7] G.S.Nayak, Sudha Kamath et al. (2006) Principal Component Analysis and Artificial Neural Network Analysis of oral Tissue Fluorescence Spectra: Classification of Normal, Premalignant and Malignant Pathological Condition Biopolymers 82:152-166
- [8] G.Subramanya Nayak and Gopallakrishna Prabhu K (2004) Analysis of Electroglottographic Signal for Measurement of Vocal Fold Abduction / Larynx Movement ,using Amplitude Modulation Technique,ECCS-2004 Proc. National Conference on Electronic Circuits and Communication Systems, Patiala, India, 2004, pp 463 – 465.

AUTHORS PROFILE

Dr. G. Subramanya Nayak has B.E degree in E&C Engineering from Mysore University, M.Tech in Biomedical Engineering from Manipal University and Ph.D in the area of Bio signal recognition, under faculty of E & E Engineering, from VM University, Ariyanoor, India. Dr. Nayak has published more than 30 papers in National / International Journals and conferences.

Dr.Dayananda Nayak has B.E degree in E&C Engineering from Mysore University, M.Tech in Biomedical Engineering from Mangalore University and Ph.D in the area of Nano Electronics, under faculty of E & E Engineering, from VM University, Ariyanoor, India. Dr. Nayak has published more than 20 papers in National / International Journals and conferences.