International Journal of Advance Research in Science and Engineering
Volume No.07, Special Issue No.01, March 2018
www.ijarse.com
IJARSE
ISSN: 2319-8354

SCREENING OF ABNORMAL ECG USING MACHINE LEARNING TECHNIQUE

Diksha Santosh Gaonkar [1], Purti Savardekar [2]

[1] *Dept. of Electronics and Telecommunication Engineering, Goa College ofEngineering*

[2] *Dept. of Electronics and Telecommunication Engineering, Goa College ofEngineering*

ABSTRACT

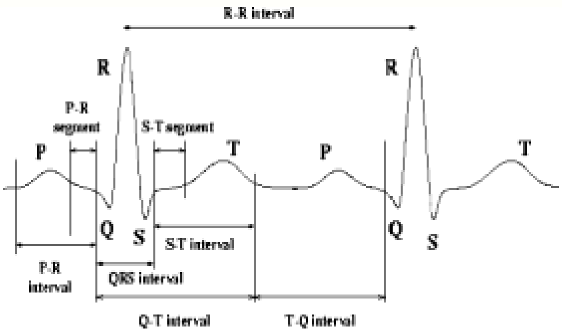
Meteoric advancements in technology and telemedicine sector have eased prior diagnosis of various diseases. One of the most recognized medical conditions that need prior diagnoses is cardiac arrhythmia. Since ECG provides the essential graphical trace of bio potential activity of heart, it has become a primary diagnosis tool for the assessment of various heart diseases. In this paper we propose a feed-forward back propagation neural network classification system for arrhythmia detection using bandpass filtering strategy. Proposed method uses BPNN to classify the ECG signals as normal and abnormal. The features are divided into two classes that are morphological features and DWT based features which are provided as input to the classifier. The performance of system was evaluated based on the percentage of accuracy. Our experimental result on MIT-BIH arrhythmia database showed overall system accuracy of 91.3% with the use of BPNN classifier.

Keywords: *ECG, Arrhythmia, MIT-BIH, BPNN, DWT*

1. Introduction

An electrocardiogram (ECG) is a bioelectrical signal which records the heart’s electrical activity versus time and thus is an important diagnostic tool for evaluation of heart functions. The signals that make the heart’s muscle to contract come from the sino atrial node, the normal pacemaker of heart which is located at the top of right atrium[5].

The ECG waveform generally consists of P, Q, R, S, T and U wave. The muscle contraction of the atria produces the P wave which represents atria depolarization. Ventricular depolarization is characterized by Q, R and S wave generally termed together as QRS complex which is the most vital feature of the ECG waveform. The ending of the atrial contraction and the beginning of the ventricular contraction is marked by the R wave peaks whose magnitude ranges from 0.1mV and 1.5mV. At last the ending of ventricular contraction is marked by the T-peak. The beginning of normal electrical sequence is always marked by atrial depolarization followed by ventricular depolarization. But in case of cardiac arrhythmias this sequence gets disrupted and rhythms tend to be irregular [1]. Over the years several techniques have been developed for ECG signal analysis aiming to achieve good percentage of accuracy, specificity and sensitivity. These techniques include classification methods such as support vector machine, RBF neural networks, fuzzy logic, wavelet coefficients and self-organizing maps. The figure 1 shows the typical ECGwaveform.



2. DatabaseCollection

Fig 1: Typical ECG waveform

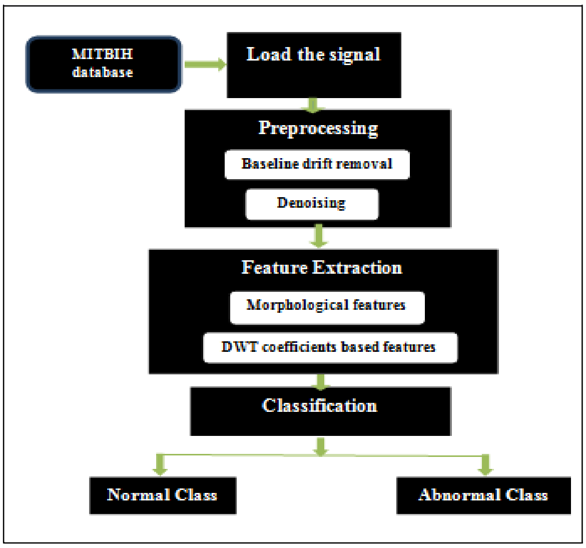
For this paper, Massachusetts Institute of Technology Beth Israel Hospital (MIT-BIH) arrhythmia database from physionet is used [4]. This database contains total 48 files of 30 minutes recordings but we have selected 46 files of 10 sec duration where 24 are considered as normal class and 22 as abnormal class based on maximum number of beats present in each record. Database includes division of recordings into two classes where first part is of 23 records (numbered from 100 to 124 inclusive with some numbers missing) chosen at random from this set, and second part is of 25 records (numbered from 200 to 234 inclusive, again with some numbers missing). The ECG waveform from MIT-BIH Database contains – a text header file, a binary file and a binary annotationfile.

3. Methodology

The block diagram of proposed technique is depicted in figure 2. The entire methodology is sectioned into three main divisions namely:

- Preprocessing: It includes baseline wander removal from raw ECG signal followed bydenoising.
- Featureextraction:Itincludesextractingandconvertingtheinputdatainformationintosetfeatures called feature vectors, by reducing data representationpattern.
- Classification: It includes the use of BPNN as a classifier to classify signal into two classes i.e.normal and abnormal.

From the figure 2 it can be seen that the raw ECG signal is loaded from the database and provided for preprocessing followed by feature extraction which involves preparing the precise input which best resembles the original signal and finally the classification stage where the processed signal is classified into abnormal and normal class.



4. ECG SignalPreprocessing

Fig 2 : Block diagram of arrhythmia detection system

Preprocessing is the initial stage in ECG signal analysis [3]. ECG being a non-stationary signal is often contaminated with various sought of noise artifacts which corrupt the raw ECG. It is therefore necessary to eliminate these noise artifacts for further analysis. This preprocess of ECG involves removal of various trends affecting the ECG and denoising using band pass filtering strategy.

4.1 Removal of baseline wander

The most common noise artifact responsible for contaminating the ECG signal is the Baseline Drift. It generally results from respiration, motion and changes in electrode impedance and lies between 0.15Hz and 0.3Hz. It can actually mask the necessary information from the ECG and cause irregularities in beat morphology. In this paper, the baseline drift of the ECG signal is eliminated by using the moving average filter to achieve the smoothed signal. Then the smoothed signal is subtracted from the original signal to get rid of the baseline drift from the ECG signal. Hence the signal thus obtained is free from baseline

4.2 Denoising

With the eradication of baseline drift the ECG signal is free from DC offset but still contains some noise. In this paper we try to implement band-pass filtering approach for noise rejection as it improves SNR and reduces influence of these noise sources Since the desired pass band 5-15Hz is unable to achieved using bandpass filters directly for our chosen sample rate we use the cascade of low pass and high pass filter to achieve 3dB pass band from 5-12Hz.

It is the cascade of low pass filter and high pass filter. The low pass filter eliminates 50Hz power line noise and electromyogram noise, having cutoff of 11Hz and with the gain 36 followed by processing delay of 6 samples.

The difference equation for low pass filter is:

$y(nT) = 2y(nT-T)-y(nT-2T)+x(nT)-2x(nT-6T)+x(nT-12T)$ ……………………………………………(1) The high pass filter eliminates lower frequency components which are motion artifacts, P wave and T wave.

The difference equation for high pass filter is:

y(nT) = 32x(nT-16T)-[y(nT-T)+x(nT)-x(nT-32T)]… ……………………………………………(2)

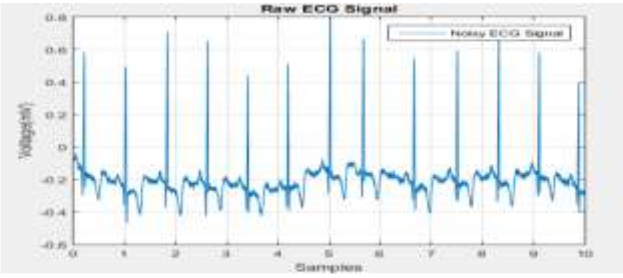


Fig 3: Raw ECG Signal

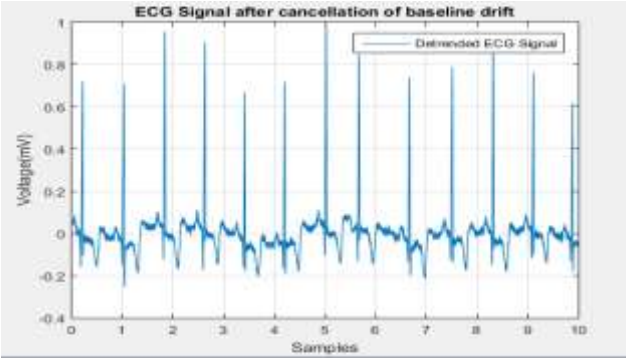


Fig 4: ECG signal with baseline drift removal

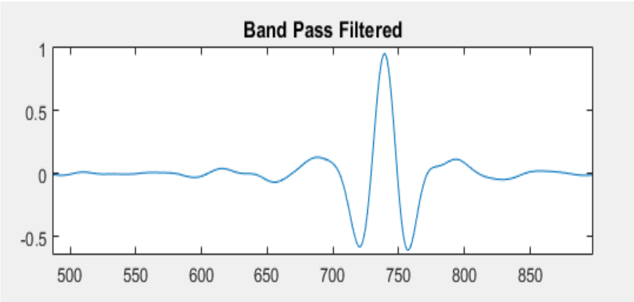


Fig 5: Denoised ECG signal

5. FeatureExtraction

Followed by removal of baseline wanders, noise reduction and peak detection, the feature extraction of ECG waveform is essential so as to use it in the preceding stage of ECG signal analysis. The feature set extracts the appropriate information from the input data in order to perform the classification task. It is thus the prevalent step in pattern recognition. In this method, two types of features are considered of ECGwaveform.

1. MorphologicalFeatures
2. Wavelet Coefficients based StatisticalFeatures

Selecting appropriate features plays a very crucial role in pattern recognition task.

In this work , the considered morphological features include QRS complex, maximum R interval , minimum R interval , QT interval , ST interval , heart rate (HR), maximum amplitude of R wave, maximum amplitude of T wave, maximum amplitude of S wave and maximum amplitude of Q wave.

Apart from morphological features wavelet coefficients based statistical features are also obtained. Since DWT coefficients represent the energy distribution of the signal they can thus be provided as feature vectors representing the signal to the classifier. Considering the dimensional reduction issue the statistics of wavelet coefficients are used.

The wavelet coefficient based features utilized are as follows:

a. Mean values of the detail and approximate coefficients at eachlevel.

b. Variance values of the detail and approximate coefficients at eachlevel.

c. Standard deviation values of the detail and approximate coefficients at eachlevel.

So at the end of feature extraction stage total 48 wavelet coefficients based features and 10 morphological features for each ECG signal was obtained which comprised of 58 features in total.

6. Neural Network

Artificial neural network (ANN) is a massively parallel-distributed processor resembling the human brain that has a natural propensity for storing experimental knowledge and making it available for use.

In this paper neural network is used in pattern recognition providing input units as feature vectors and output units as the class to be classified. Each corresponding feature vector is served to the input layer whose output is considered to be corresponding element in the vectors. Hidden layer is responsible for calculation of weighted sum of its input thus providing net activation of its scalar.

1.1 Feed-forward back propagation neural network

The term back propagation refers to the manner in which the gradient is computed for nonlinear multilayer networks. Standard back propagation is a gradient decent algorithm which is used to find the weights.Backpropagation algorithm uses feedfordward architecture containing hidden layers.

7. ExperimentalResults

In this work, the entire system was evaluated on MIT-BIH arrhythmia database sectioned into two classes namely abnormal and normal. Among the total 48 files of 30 min long only 46 files of 10 sec duration were utilized. The total 58 number of features were divided into two separate sections. These include 48 DWT coefficients based statistical features and 10 morphological features of ECG signal. Simulation and training of the network involved 32 samples from the database with 7 samples provided for testing and 7 samples for validation set. The classifier was designed with 3 hidden layers containing 20 neurons in first hidden layer, 25 neurons in second hidden layer and 31 in third hidden layer respectively. The output layer is responsible for deciding the targets matrix as (1, 0) and (0, 1) referring to normal and abnormal class.

Figure 6 depicts the simulation results of the neural network in terms of confusion matrix.

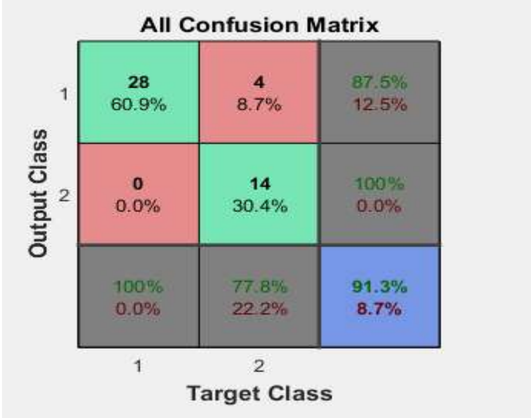


Fig 6: Confusion matrix of BPNN

With the reference to the confusion matrix there is no misclassification carried out for normal ECG samples. That means out of 28 samples provided to the network all 28 samples were classified correctly by BPNN .Whereas 4 abnormal samples were classified wrongly by BPNN as normal samples, that means 14 samples are classified correctly out of total 18 samples. Therefore the overall classification accuracy of 91.3% was achieved using back propagation neural network.

8. Conclusion

This paper addresses an effective feedfordwad back propagation neural network based system for classification of cardiac arrhythmia into two separate classes as normal and abnormal. This work was carried out on 46 files of MIT-BIH arrhythmia database consisting of 48 files 30 mins long. With total 58 features comprising of 10 morphological features and 48 DWT coefficients based features the overall system accuracy of 91.3% was achieved using 3 hidden layers with 20,25 and 31 number of neurons in first , second and third hidden layer respectively

Acknowledgment

I am highly indebted to my guide Miss Purti Savardekar for her guidance and constant supervision regarding the project.

I would also like to express my sincere gratitude towards the HOD and staff of electronics and telecommunication department Goa College of Engineering Farmagudi, where the entire project was carried out.

References

[1] K Amtul Salam, G. Srilakshmi, “An algorithm for ECGanalysisof arrhythmia detection”,Electrical,

Computers and Communication Technologies (ICECCT), IEEE International Conference, March 2015

[2] Ali Baharami Rad,Trygve Etestol,Kjersti Engan, Unai Irusta,Jan Johansen,Lars Wik,Aggelos K. Katsaggelos, “ ECG based classification of Resuscitation Cardiac Rhythms for Retrospective Data Analysis”, IEEE Transaction on Biomedical Engineering ,volume:64,Issue:10,October2017.

[3] K. K. Makesheva, Y. A. Altay, "Noise-resistant processing of electrocardiosignal in MATLAB", *European Journal of Scientific Researc*, vol. 13, no. 1, pp. 898-903,2016.

[4] “TheMIT-BIHArrhythmiaDatabase,http://physionet.ph.biu.ac.il/physiobank/database/mitdb/

[5] Hari Mohan Rai,Anurag Trivedi, “ECG signal Classification using Wavelet Trasform and Backpropagation neural network”,Computers and Devices for communication (CODEC), 5th IEEE International Conference, December 2012.

[6] Hilmy Assodiky, Iwan Syarif, Tessy badriyah, “Deep learning algorithm for arrhythmia detection”,Knowledge Creation and Intelligent Computing (IES-KCIC) ,IEEE International Electronics Symposium,September 2017.

[7] Miad Faezipour,Adnan Saeed,Hlaing Minn,lakshman Tamil, “A Patient Adaptive profiling Scheme for ECG Beat Classification”,IEEE Trans. On Information Technology in Biomedicine , volume:14, Issue:5,September 2010.