CARDIAC ARRHYTHMIA PREDICTION USING IMPROVED MULTILAYER PERCEPTRON NEURAL NETWORK

VSR KUMARI¹ & P. RAJESH KUMAR²

¹Research Scholar, Department of ECE, Andhra University, Visakhapatnam, Andhra Pradesh, India
²Professor, Department of ECE, Andhra University, Visakhapatnam, Andhra Pradesh, India

ABSTRACT

Electrocardiogram (ECG) has much diagnostic information to ensure proper clinical decisions in cardiac arrhythmia. Heart rate variations are signposts of current heart disease or imminent cardiac diseases. This study uses an ECG to determine bundle branch block (BBB), a form of heart block involving conduction delay/failure in the heart’s bundle branch. Machine learning and data mining methods are considered to improve ECG arrhythmia detection accuracy. Usually an automated classification of cardiac arrhythmias procedure is suggested on the basis of linear and non-linear HRV analysis. This paper presents an automated method for classification of cardiac arrhythmic based on ECG rhythm.

RR intervals are extracted from ECG data using Symlet, and symmetric uncertainty is used for feature reduction. Extracted RR data is the classification feature with beats being classified through an Improved Neural Network and finally being evaluated through the use of MIT-BIH arrhythmia database.

KEYWORDS: Arrhythmia Classification, Electrocardiogram (ECG), RR Interval, MIT-BIH ECG Data Neural Network

INTRODUCTION

A major cause of death globally is cardiovascular disease. But the provision of pre-monitoring and pre-diagnostic methods can prevent heart attacks. Specifically, clinicians will be really helped if they can detect abnormalities in heart function called arrhythmias early. Electrocardiogram (ECG) can ensure proper clinical decisions in cardiac arrhythmia. Heart rate variations can be pointers of current heart disease or those in the near future [1, 2]. ECG consists of three - P, QRS, and T – waves which correspond to electrical phenomena induced on the heart’s surface, i.e.; atrial depolarization, P, ventricular depolarization, QRS complex, and ventricular repolarization, T [3]. The last few decades witnessed the development of many methods for automated arrhythmia detection to simplify monitoring [4].

This study plans to determine - based on ECG analysis - bundle branch block (BBB) [5] - a heart block type involving conduction delay/failure in the heart’s bundle branch which could be complete/incomplete, transient/permanent, or intermittent. It is named on the basis of the involvement of left/right bundle branch. An abnormal QRS complex is caused by cardiac impulses inability to conduct down bundle branches. BBB is evident in acute, anterior wall myocardial infarction, due to ischemia or bundle branches necrosis, or a mechanical branch compressed by a tumour.

Bundle branch blocks are heart problems known as intraventricular conduction defects (IVCD). Left and right bundle branches exist. BBB involves nerve impulses slowing/interruption. Patients with right bundle branch block (RBBB) have limited, or no conducted nerve impulses. The right ventricle receives impulses through the muscle-to-muscle spread beyond regular nerve pathway, the slow impulse transmission causing delayed right ventricle contraction. Left bundle branch block (LBBB) is of many types with each having its specific failure mechanism [6].

Many computer based diagnostic systems used Artificial Neural Network (ANN) as a base to design structural...
designs. Improved ANN techniques including data reduction and feature extraction were presented in literature [7-10]. MLP can recognize and classify ECG signals better than other ANN methods.

Neural networks are classification tools with research proving that they provide a good alternative to many conventional classification methods. The following are their advantages. First, as neural networks are data driven and self-adaptive they adjust to data without any functional specification or distributional form. Next, as they are functional approximators they approximate functions accurately. As classification requires intra group member’s functional relationships and object attributes, accurate identification of functions gains importance. Also, NN are nonlinear and hence flexible in modelling real world complex relationships. Finally, NN estimate posterior probabilities providing a base for classification rule establishment and statistical analysis performance.

This paper presents an automated ECG rhythm based cardiac arrhythmic classification procedure. ECG data using Symlet provide RR intervals to extract data with symmetric uncertainty being used to reduce features. Extracted RR data is a classification feature and beats are classified through a improved Neural Network. The final evaluation is by using a MIT-BIH arrhythmia database.

RELATED WORKS

Ozbay et al [11] presented a new fuzzy clustering NN architecture (FCNN) for early cardiac arrhythmia diagnosis. ECG signals from MIT-BIH ECG database, classify 10 different arrhythmias for training and they include normal sinus rhythm, sinus arrhythmia, sinus bradycardia, ventricular tachycardia, right bundle branch block, left bundle branch block, paced beat, atrial fibrillation, atrial premature contraction, and atrial flutter. The suggested structures were trained by backpropagation algorithm for testing using 92 patients experimental ECG records. The results proved that the new FCNN architecture generalized better than ordinary MLP architecture learning better and faster. It’s advantage lay in the fact that the proposed structure is due to decreasing segment numbers by grouping similar segments in training data by using fuzzy c-means clustering.

Arrhythmia was classified into normal and abnormal classes through a model proposed by Jadhav et al [12] called Modular neural network (MNN). Experiments were conducted on an UCI Arrhythmia data set. Missing data set attribute values are replaced by the concerned class’s nearest column value. This model was constructed by varying hidden layers from one to three and by varying training data set partition percentage. The method focused on high confident arrhythmia classification results applicable in diagnostic decision support systems. This data set can test classifiers as it is incomplete and ambiguous bio-signal data from total 452 patients. Classification performance evaluation is through use of six measures; classification accuracy, mean squared error (MSE), receiver operating characteristics (ROC), sensitivity, specificity, and area under curve (AUC). Experiments demonstrate more than 82.22% classification accuracy.

Moavenian et al [13] introduced a Kernel–Adatron (K–A) learning algorithm to help SVM (Support Vector Machine) classify ECG arrhythmias. The proposed classifier is compared to MLP (multi-layered perceptron) through back propagation (BP) training algorithm use. ECG signals from MIT-BIH arrhythmia database train and classify 6 different arrhythmia, plus normal ECG. MLP and SVM training/testing stages were carried out twice, being trained first with only one ECG lead signal. A second ECG lead signal was added to the training/testing datasets with the aim of investigating its influence on training/testing performance (generalization ability) and training time plus training time for both classifiers. Implementation of three criteria for ECG signals classification can reduce structural comparison issues which have not received much attention in earlier research. Results indicate that SVM compared to MLP is faster in training and performs seven times better, but MLP’s generalization ability as regards mean square error is reduced by more than three times.
Ceylan et al [14] presented a wavelet neural network for bundle branch blocks classification which was implemented by using Morlet and Mexican hat wavelet as activation function in a hidden layer. ECG data was taken from MIT-BIH ECG Arrhythmia Database. Training/test data included three beat types belonging to normal, right bundle branch block and left bundle branch block ECG signal classes. Experimental studies proved that Mexican hat wavelet designed wavelet neural network was successful than other networks.

Ghorbanian et al [15] developed an algorithm for detection/classification of 6 ECG signal beat types like normal beats (N), atrial premature beats (A), right bundle branch block beats (R), left bundle branch block beats (L), paced beats (P), and premature ventricular contraction beats (PVC or V) through the use of a neural network classifier. Many preprocessing stages are applied to prepare an input vector for the neural classifiers. Continuous wavelet transform (CWT) was applied for feature extraction from an ECG signal. Also, principal component analysis (PCA) reduces data size and then MIT-BIH database evaluated the proposed algorithm leading to 99.5% sensitivity, 99.66% positive predictive accuracy and 99.17% total accuracy.

METHODOLOGY

R-R Interval

This study chose R-waves for arrhythmia classification representing as they do the highest signal-to-noise ratio information available; arrhythmias usually create disturbances in R-R interval pattern. An RR wave with characteristic peaks and valleys P, Q, R, S and T in ECG is shown in Figure 1.

![Figure 1: R-R Interval Pattern in ECG](image)

R-wave detector is a two-step process [16]: (1) low frequency baseline elimination and (2) detection/identification of R-waves on cleaned-up signal. Baseline removal ensured accurate R-wave amplitude measurements. Experiments proved that this was required to prevent missing an R-wave during baseline’s high slope portions. Baseline removal was undertaken through a low pass, zero phase shift, moving average filter. 800 ms window length with a 40 ms sampling interval yielded a 3 dB point of 0.3Hz. ARS complex’s high slope segments relative to the remaining waveform identifies this complex. To ensure a well-defined fiducial QRS complex point, maximum startup slope point is the fiducial point.

Feature Extraction

Wavelets are waveforms bound in time and latency and their analysis splits signals into the mother wavelet’s shifted and scaled versions. Continuous Wavelet Transform (CWT) is known by the wavelet function $\psi$ adding signal times multiplied by scaled and shifted versions. The continuous wavelet is defined mathematically by

$$C(scale, position) = \int_{-\infty}^{\infty} f(t)\psi(scale, position, t)dt$$

(1)
CWT is responsible for many wavelet coefficients C, a function of scale and position. Original signals constituent wavelets are got through multiplying every coefficient by applicable scaled and shifted wavelet. Daubechies proposed symlets, symmetrical wavelets and got modifications of db family [17]. Both wavelet families are similar, the difference being db wavelets with maximal phase while symlets have minimal phase. The latter is supported wavelets with slight asymmetry with its wavelet coefficient being any positive even number and the highest number of vanishing moments for a specific support width.

Feature Reduction

High dimensional data - data sets with hundreds or thousands of features – has much irrelevant/redundant information degrading learning algorithms performance. Hence, feature selection is necessary for machine learning tasks to face high dimensional data. But this enormity on both size/dimensionality is a challenge to feature selection algorithms. Usually, a feature is good when relevant to class concept and not redundant to any other relevant features. If the correlation between two variables as a goodness measure is adopted, the above definition proves a feature is good when highly correlated to class but not to any other features.

Information-theoretical based correlation measure concept of entropy gives the measure the uncertainty of a random variable. Entropy of a variable X is defined as

\[ H(X) = \sum_i P(x_i) \log_2 \left( \frac{1}{P(x_i)} \right) \]  

(2)

and entropy of X after observing values of a variable Y is defined as

\[ H(X|Y) = \sum_j \sum_i P(y_j|x_i) \log_2 \left( \frac{1}{P(x_i|y_j)} \right) \]  

(3)

where \( P(x_i) \) is the prior probability for all values of X, and \( P(x_i|y_j) \) is a posterior probability of X given values of Y. The amount by which X entropy decreases reflects X’s additional information provided by Y and is the information gain [18], given by

\[ IG(X|Y) = H(X) - H(X|Y) \]  

(4)

Symmetry is measures correlations between features. But information gain favors features with more values. Also, values should be normal so that they are comparable and produce the same effect. Hence symmetrical uncertainty [19] is chosen and defined as follows.

\[ SU(X,Y) = 2 \left[ \frac{IG(X|Y)}{H(X)+H(Y)} \right] \]  

(5)

It compensates information gain’s bias against features with many values, normalizing its values to the range [0;1] with value 1 indicating that knowledge of either one predicts the value of the other. Also, value 0 indicates that X and Y are independent. It also treats two features symmetrically. Though Entropy-based measures require nominal features, they are still measure correlations between continuous features also when values are discretized in advance. [20].

Proposed Neural Network

Computational units multiple layers make up neural networks, interconnected and based on network design.
Inputs are fed on input layer and propagated to the output through layers and is computed using weights, bias and activation functions. The network is trained by propagation rules by back propagating errors and changing nodes weights. The difference between obtained and desired outputs is the error. Most artificial neural networks are monolithic in structure with many working on fully connected networks/layers. Such networks perform well on very small input space. But when complexity increases performance decreases with growing input dimension.

The organisation of the brain is a layered structure. Subsequent neuron layers are arranged hierarchically to form complex representations. In addition to a primate brain’s layered, hierarchical structure, multiple parallel processing streams/pathways that are another major neural construction principle, also exist. The brain structuring into distinct streams ensures independent processing of information types/modalities. This principle is the base for the proposed neural network.

The problem with using ANN to solve large-scale real-world issues is to divide a problem into smaller/simpler sub problems; and in assigning a network module to learn each sub-problem; and combining them as solutions to the original problem. Recently, researchers studied modular neural network learning approaches to overcome this issue[21, 22]. Neural network architectures improve performance when constrained. The brain’s modular connectivity constraints provide guidelines for an effective ANN architectures design.

The architecture of the proposed neural network is shown in figure 2.

![Figure 2: Architecture of Modular Neural Network](image)

The proposed modular neural network has two sub-networks each being made up of two hidden layers with differing transfer functions. This study uses a sigmoid activation function as transfer function. Table 1 provides the proposed neural network’s features.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input Neuron</td>
<td>25</td>
</tr>
<tr>
<td>Output Neuron</td>
<td>4</td>
</tr>
<tr>
<td>Number of Hidden Layer</td>
<td>2</td>
</tr>
<tr>
<td>Number of processing elements upper</td>
<td>4</td>
</tr>
<tr>
<td>Number of processing elements lower</td>
<td>4</td>
</tr>
<tr>
<td>Transfer function of hidden layer</td>
<td>Sigmoid</td>
</tr>
<tr>
<td>Learning Rule of hidden layer</td>
<td>Momentum</td>
</tr>
<tr>
<td>Step size</td>
<td>0.1</td>
</tr>
<tr>
<td>Momentum</td>
<td>0.7</td>
</tr>
</tbody>
</table>

**RESULTS**

Classifier performance was evaluated through the use of a MIT-BIH arrhythmia database. The evaluation dataset had 165 instance; 55 events each of Right bunch bundle block, Left bunch bundle block and Normal RR interval.
Continuous wavelet transforms are applied by using symlet2 filters. Feature number is lowered through use of symmetric uncertainty. The proposed network classifies instances from MIT-BIH arrhythmia database. A total of 25 features are selected for the classification process.

Classification accuracy experiments results are provided in the following Tables/Figures. Table 2 tabulates the summary of results for various techniques used. Figure 3 shows the plotted classification accuracy and Figure 4 show the root mean square error of each method.

**Table 2: Summary of the Results**

<table>
<thead>
<tr>
<th>Method</th>
<th>Classification</th>
<th>RMSE</th>
<th>Precision</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLP</td>
<td>0.93</td>
<td>0.2046</td>
<td>0.928</td>
<td>0.927</td>
</tr>
<tr>
<td>Proposed MLP</td>
<td>0.94</td>
<td>0.1982</td>
<td>0.940</td>
<td>0.939</td>
</tr>
<tr>
<td>MLP with feature selection</td>
<td>0.94</td>
<td>0.1876</td>
<td>0.945</td>
<td>0.945</td>
</tr>
<tr>
<td>Proposed MLP with feature selection</td>
<td>0.95</td>
<td>0.1765</td>
<td>0.951</td>
<td>0.951</td>
</tr>
</tbody>
</table>

**Figure 3: Classification Accuracy**

**Figure 4: Root Mean Squared Error**

It is observed from the above table and figures that with the classification accuracy improves with the proposed MLP with feature selection. Similarly, the root mean squared error also reduces significantly.

**Figure 5: Precision and Recall**
For a good classification system, the values of precision and recall should be high which is achieved in the proposed Modular Neural network (Figure 5).

CONCLUSIONS

This paper based on RR intervals investigates ECG classification procedure for arrhythmic beat classification. The methodology included RR interval beat extraction through the use of Symlet conversion on ECG data. Feature reduction is through Symmetric uncertainty. The classification feature with beats is the extracted RR data, the beats being classified through a Modular Neural Network and being evaluated through the MIT-BIH arrhythmia database. The latter evaluates classification efficiency with instances being classified as Right bunch bundle block, Left bunch bundle block and Normal RR interval. Classification accuracy achieved is 95.01% when 25 features are used for instance classification. Equally satisfactory precision and recall is achieved.

REFERENCES


