Heart Arrhythmia Detection Using Continuous Wavelet Transform and Principal Component Analysis with Neural Network Classifier

Parham Ghorbanian[1], Ali Ghaffari[2], Ali Jalali[1], C Nataraj[1]

[1] Department of Mechanical Engineering, Villanova University, Villanova, PA, USA
[2] Department of Mechanical Engineering, K.N.Toosi University of Technology, Tehran, Iran

Abstract

The aim of this study is to develop an algorithm to detect and classify six types of electrocardiogram (ECG) signal beats including 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) using a neural network classifier. In order to prepare an appropriate input vector for the neural classifier several pre-processing stages have been applied. Continuous wavelet transform (CWT) has been applied in order to extract features from the ECG signal. Moreover, Principal component analysis (PCA) is used to reduce the size of the data. Finally, the MIT-BIH database is used to evaluate the proposed algorithm, resulting in 99.5% sensitivity (Se), 99.66% positive predictive accuracy (PPA) and 99.17% total accuracy (TA).

1. Introduction

Cardiovascular disease, including heart disease and stroke, remains the leading cause of death around the world. Yet most heart attacks and strokes could be prevented if some method of pre-monitoring and pre-diagnostic can be provided. In particular, early detection of abnormalities in the function of the heart, called arrhythmias, can be valuable for clinicians.

The electrocardiogram (ECG) plays an important role in the process of monitoring and preventing heart attacks. The ECG, as it is shown in figure 1, consists of three basic waves: P, QRS, and T. These waves correspond to the far field induced by specific electrical phenomena on the cardiac surface, namely, the atrial depolarization, P, the ventricular depolarization, QRS complex, and the ventricular repolarization, T.

The techniques, developed for automated electrocardiographic changes detection, work by transforming the mostly qualitative diagnostic criteria into a more objective quantitative signal feature classification problem. In order to address this problem, the analysis of ECG signals for detection of electrocardiographic changes have been carried out using techniques such as autocorrelation function, time frequency analysis, and wavelet transform (WT) ([1], [2], and [3]).

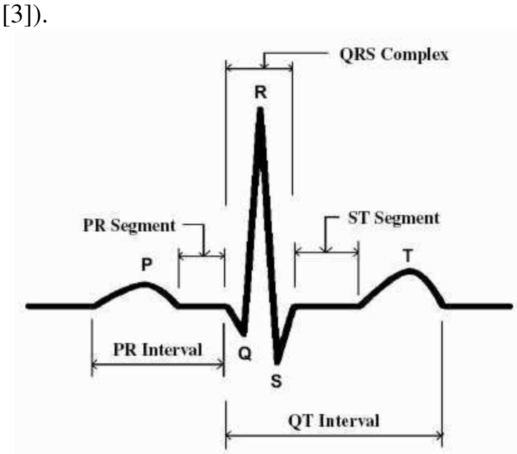

Figure 1. The components of ECG signal.

Results of studies in the literature have demonstrated that WT is the most promising method to extract features from ECG signals ([2] and [3]). These features characterize the behaviour of the ECG signals.

Application of wavelet transform, principal component analysis (PCA) and several types of neural network structures in order to detect and classify different kinds of heart arrhythmias have been presented by Silipo et al. [4]. In their research, results of different neural network structures in order to find the best neural network structure for the classification of specific types of arrhythmias have been compared. Papaloukas et al. [5] used a neural network classifier to detect and classify ischemic arrhythmia episodes in the ECG signal. They also used PCA as a method to select and extract features from the ECG signal. Foo et al. [6] used different types of multilayer neural network as a classifier to detect two types of ECG patterns. A comparison between different structures for heart arrhythmia detection algorithms based on neural network, fuzzy cluster, wavelet transform and principal component analysis, was carried out by Ceylan et al. [7]. Kutlu et al. [8] applied a K-nearest neighbourhood algorithm for the purpose of classification. They achieved an accuracy of 97.3% in classifying 5 types of heart arrhythmias. Cvikl et al. [9] designed a field-programmable gate array-based (FPGA) system for ECG signal processing.

The most difficult problem faced by today’s automatic ECG analysis is the large variation in the morphologies of ECG waveforms [10]. We address this problem of beat classifier performance by using a combination of continuous wavelet transform (CWT) and principal component analysis (PCA) in order to prepare a more effective input data for neural network classifier, leading to better classification results.

Continuous wavelet transform (CWT) is a time-frequency analysis method which differs from the more traditional short time Fourier transform (STFT) by having a variable window width, which is related to the scale of observation [11].

Principal components analysis (PCA) is a well-established technique for feature extraction and dimensionality reduction and has been used in a wide range of ECG signal analysis [3], [4], and [7].

2. Methodology

A schematic of the designed algorithm in this study for ECG beats classification is shown in Figure 2. The first stage is pre-processing stage including four levels of data processing which are signal filtering, sample selection, feature extraction, and finally dimensionality reduction. The other stages are main process and classification of ECG beats, which is the main goal of this study.

2.1. Signal filtering

In this stage a method of signal filtering presented by Ghaffari et al. [12] is applied to remove baseline wandering of the ECG signal. Figures 3.a shows raw ECG signal of records 232 from MIT-BIH database, which has baseline wandering. Figures 3.b shows the same range of ECG signals after applying the filtering method. It is obvious that the baseline wandering has been removed, leading to a better performance of the neural classifier.

2.2. Sample selection

The data of ECG signals used in this study are taken from the MIT-BIH ECG signals database, including normal beats and five types of different arrhythmia beats. For this paper, the selected types of arrhythmias are 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).

The appropriate and suitable range of samples from the raw ECG signal was found experimentally to be 150 samples after R wave for all types of signals, which are called a segment.

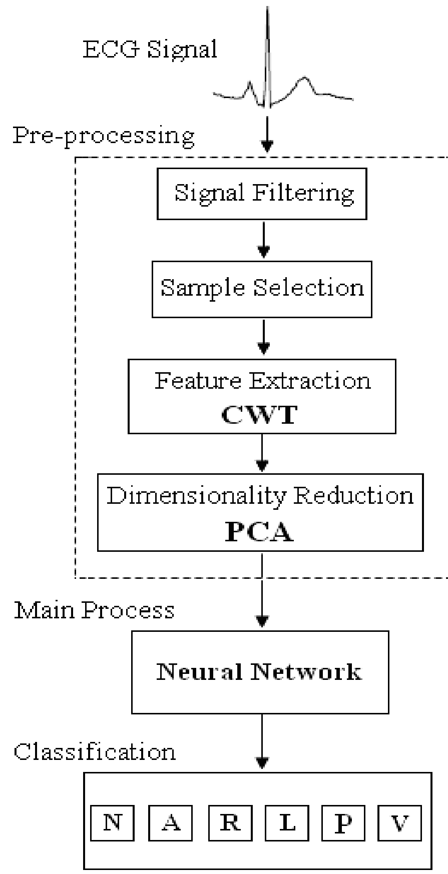

Figure 2. Schematic of the designed algorithm.

2.3. Feature extraction

The selection of the analyzing function in wavelet transform, which is called the mother wavelet, has a significant effect on the result of analysis and should be selected carefully based on the nature of the signal [3]. The most important factor to select mother wavelet in this study refers to the simplicity of the computed CWT coefficients, while having the appropriate and sufficient data about the ECG signal. In this study, ‘Haar’ mother wavelet has been selected for feature extraction. The analysis shows that extracted features from ECG signal by using the Haar mother wavelet would be suitable and appropriate for the data classification and the computed coefficients can represent morphological differences very well.

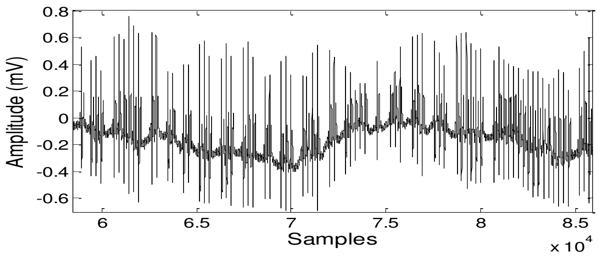

Figure 3.a. Filtered ECG signal from record 232

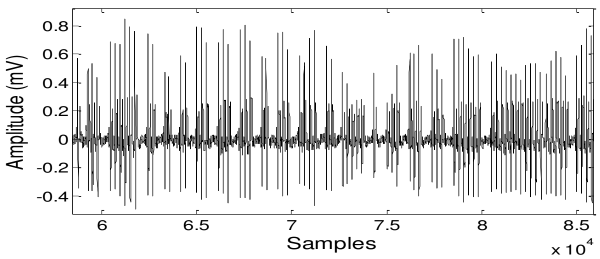

Fig. 3.b - Filtered ECG signal from record 232

In order to compute CWT of signals, a specific range of scales, which is suitable and appropriate for feature extraction, is needed. Computing CWT of signals in this range of scales should be sufficient for a complete feature extraction and should not lead to a huge volume of data. In this study, the wavelet coefficients are computed using the MATLAB software package.

In this study, the analysis shows that computing CWT of signals in the range of scales from a=5 through a=20, can lead to a complete and useful analysis. Since in this range, both noise of signals and the effect of morphological differences can be analyzed and the extracted features would be useful for classification of the signals under study. Using this range of scales has two advantages. First of all, by computing CWT in the range of a = 6 through a=9, the ECG signal can be analyzed in detail, and the second is that, by using range of a = 10 through a=15, the general morphology of the signal and its differences with other types of ECG signals can be highlighted.

2.4. Dimensionality reduction

A huge amount of data is not efficient to perform a pattern recognition process. In our algorithm and in the final level of pre-processing, PCA is applied on the computed matrices of wavelet coefficients, where each of them is a 10*150 matrix, resulting in 10 principal component (PC) vectors. In this study, the first three PC vectors have been selected and arranged as neural network classifier input vector. By selecting only three PC vectors, dimensionality reduction without significant loss of data information is achieved, leading to a better performance of the neural. The prepared vectors, which are the principal components, are used as the neural network classifier input vector. The analysis for providing the input vector structure is the same for both training and testing database.

2.5. Main process

Selection of the neural network inputs is the most important component of designing the neural network based pattern classification; since even the best classifier will perform poorly if the inputs are not selected well [13]. The inputs of neural network in this study are formed in the way which was described in previous section.

After finishing the pre-processing stages, data is ready as the input vector for the neural network classifier. In this study a classical multi-layer perceptron neural network (MLPNN) structure is used as the neural network classifier structure. This MLPNN has 2 hidden layers, with 60 nodes in the first hidden layer and 15 nodes in the second hidden layer for 160 iterations and is trained with the back propagation method of error.

From all 6 types of ECG beats under study and for neural network training data, two segments have been selected and processed in the way that was described in previous section.

2.6. Classification

When the neural network has been trained, it is ready as a classifier to detect and classify different types of ECG signals into one of six ECG beat groups under study. The classifier has been tested by 100 segments from each group of ECG signals. These testing segments are processed and prepared exactly like the input vector of the neural network; this means that all four levels of pre-processing stage have been applied to each segment in order to prepare it as a testing segment. These segments are used to test and evaluate the trained neural network classifier.

3. Results

The MIT-BIH arrhythmia database is used to evaluate the proposed algorithm. To assess the accuracy of the classifier, sensitivity, Se, positive predictive accuracy, PPA, and total accuracy, TA, has been calculated.

Table 1 shows the result of classification by neural network. It is shown in this table that from the whole testing data base, the classification fails only in 5 cases. According to this table, the algorithm achieves a good performance with 99.5 % Se, 99.66% PPA and 99.17% TA.

A comparison between results of this study and other studies in the field of heart arrhythmias detection has been presented in table 3. According to this table, the algorithms in this study show reasonably accurate results, and compare favourably to other studies.

Table 1 - Results of the algorithm on MIT- BIH database.

	N	A	R	L	P	V	Sum
All	100	100	100	100	100	100	600
TP	100	99	99	98	100	99	595
FN	0	0	0	2	0	1	3
FP	0	1	1	0	0	0	2
Se (%)	100	100	100	98	100	99	99.5
PPA (%)	100	99	99	100	100	100	99.66
TA (%)	100	99	99	98	100	99	99.17

Table 2 - Comparison of several classifier performances on MIT-BIH database.

	TA (%)	PPA (%)	Se (%)
Silipo et al. [4]	-	85 %	77 %
Papaloukas et al. [5]	-	89 %	90%
Foo et al. [6]	92 %	-	-
Kutlu et al. [8]	97.3 %	-	-
Cvikl et al. [9]	-	-	92.36 %
Guler et al. [13]	96.94 %	-	96.37 %
This Study	99.17 %	99.66 %	99.5 %

4. Conclusion

Automatic detection of heart arrhythmias could be very important in clinical usage and lead to early detection of a fairly common malady and could help contribute to reduced mortality. In this study, the use of neural networks for classification of the ECG beats is presented. Several stages of pre-processing have been used in order to prepare the most appropriate input vector for the neural classifier.

The main advantage of this study is that, by using 10 scales in computing CWT of signals, the morphological differences between several types of ECG signal are highlighted and the extracted features show the differences more clearly. Another advantage of this study is that the reduction of the dimension of data by applying PCA led to the most appropriate input vector for neural network classifier which improved the performance of neural network classifier significantly.

References

[1] Maglaveras N, Stamkopoulos T. ECG pattern recognition and classification using non-linear transformations and neural networks: A review. International Journal of Medical Informatics 1998; 52: 191-208.

[2] Kundu M, Nasipuri M. Knowledge-based ECG interpretation: a critical review. Pattern Recognition 2000; 33: 351-373.

[3] Addison PS. Wavelet transforms and the ECG: a review. Physiological Measurement 2005; 26: 155-199.

[4] Silipo R, Marchesi C. Artificial neural networks for automatic ECG analysis. Signal Processing 1998. 46; 1417-1425.

[5] Papaloukas C, Fotiadis DI. An ischemia detection method based on artificial neural network. Artificial Intelligence in Medicine 2002; 24: 167-178.

[6] Foo SY, Stuart G. Neural network-based ECG pattern recognition. Engineering Applications of Artificial Intelligence 2002; 15: 253-260.

[7] Ceylan R, Ozbay Y. Comparison of FCM, PCA and WT techniques for classification ECG arrhythmias using artificial neural network. Expert Systems with Applications 2007; 33: 286-295.

[8] Kutlu Y, Kuntalp D. Arrhythmia classification using higher order statistics. IEEE Signal Processing, Communication and Applications Conference 2008; IEEE Press: 1-4.

[9] Cvikl M, Zemva A. FPGA-oriented HW/SW implementation of ECG beat detection and classification algorithm. Digital Signal Processing 2010; 20: 238–248.

[10] Ozbay Y, Ceylan PR. Classification of heart rate data using artificial neural networks and fuzzy equivalence relation. Computers in biology and medicine 2005; 36:, 61-68.

[11] Osowski S, Linh T. ECG beat recognition using fuzzy hybrid neural network. IEEE Transaction on Biomedical Engineering 2001; 48: 1265-1271.

[12] Ghaffari A, SadAbadi H, Ghasemi M. A Mathematical algorithm for ECG Signals Denoising Using Window Analysis. Biomed Papers 2006; 151: 73-78.

[13] Guler I, Ubeyli E. ECG beat classifier designed by combined neural network model. Journal of the pattern recognition society 2005; 38: 199-208.

Address for correspondence.

Parham Ghorbanian
Department of Mechanical Engineering, Villanova University
800 E. Lancaster Ave. Villanova, PA. 19085
parham.ghorbanian@villanova.edu