CLASSIFICATION AND PREDICTION OF CARDIAC VASCULAR DISEASE

Dr. S.N. Shivappriya 1  Dr. R. Harikumar 2  A. Kalaiselvi 1
1 Assistant Professor II, ECE Department, Kumaraguru College of technology, Coimbatore, India-641 049
2 Professor, ECE Department, Bannari Amman Institute of Technology, Sathyamangalam, India-638 401
Email: shivapriyakvp@gmail.com

Abstract: Analysis of Cardiac Vascular Disease (CVD) is much important to diagnosis the severity of the heart problems and to identify Heart disease. This work compares the recent works in that how the disease can be detected, classified and predicted in advance with the Electro Cardio Gram. Because earlier prediction Cardiac diseases will save the human life and appropriate treatment can be given to the patient in appropriate time.

Key words: Cardiac Vascular Disease, Electro Cardio Gram, Classifier

1. Introduction

Cardiac Vascular Disease (CVD) causes the death of over 17 million people worldwide each year. The main causes of CVD includes heart attacks, strokes, heart valve problems and arrhythmia. With the help of the Electro cardio Gram (ECG) the functioning of heart can be analyzed with Computer Aided Diagonosis. Various methods of automated diagnosis have been adopted (Rosaria Silipo & Carlo Marchesi 1998), however the entire process can generally be subdivided into a number of disjoint processing modules: beat detection, feature extraction, feature selection, and classification. The initial pre-processing module of beat detection aims to locate each cardiac cycle in each of the recording leads and will insert reference markers indicating the beginning and end of each inter wave component. In order to classify different CVD the predominant features has to be extracted and then it is given to the classifier for the classification process.

2. Materials and Methods

Many commercial tools are available for automatic beat detection, but rather their performance is not satisfactory. In MATLAB, the possible enhancements in other earlier algorithms (Misiti 1996, Mikhled Alfaouri et al 2008) with the help of Wavelet Transforms is performed to reduce the complexity and to improve the Signal to Noise Ratio (SNR) of the ECG signal earlier to detection. The software application could take the ECG signal, denoise it, and perform the beat detection efficiently. Possible enhancements include reducing the number of fiducial marks (P, Q, R, S, and T waves) and thereby reducing the number of thresholds that could, to a large extent, reduce the complexity.
In the Feature Extraction process, the QRS frequency band has been used without identifying the optimum QRS frequency range for the detection of the QRS complexes. Different researchers used different passbands; for example, Thakor et al (1983) proposed an estimate of QRS complex spectra and suggested that the passband that maximizes the QRS energy is approximately 5–15 Hz. (Pan & Tompkins 1985) used cascaded low-pass and high-pass filters to achieve a passband of about 5–11 Hz. Omer T. Inan Laurent Giovangrandi & Gregory T. A. Kovacs (2006) used a quadratic spline wavelet with compact support and one vanishing moment. It was concluded by them that most QRS complex energies are at the scale of $2^2$; that is, the Fourier transform frequency range lies between 4 and 13.5 Hz.

Sahambi et al (1997) used the first derivative of a Gaussian smoothing wavelet and found that most QRS complex energies are at the scales of $2^3$ and $2^4$, with corresponding frequency ranges between 4.1 Hz and 33.1 Hz. Benitez et al (2000) developed a QRS detection algorithm using the properties of the Hilbert transform with band stop frequencies at 8 and 20 Hz in order to remove muscular noise and maximize the QRS complex, respectively. Moraes et al (2002) combined two improved QRS detectors using a band-pass filter between 9 and 30 Hz. Chen & Chen (2003) introduced a QRS detection algorithm based on real-time moving averages and assumed the QRS frequencies were concentrated at approximately 5–15 Hz. Mahmoodabadi et al (2005) used Daubechies 2 WT to detect QRS complex using scales of $2^3$–$2^5$, which covers the frequency range 2.2–33.3 Hz.

After the pre-processing and beat detection, effective beat classification is essential for correct analysis of different types of CVD. The general trend is to develop automated systems to classify cardiac beats. This can significantly help to simplify the diagnosis of heart diseases. For instance, heart-rate variability or diagnosis of certain arrhythmia (abnormal cardiac signal) may take up to several hours, when done by visual inspection. Even then, some vital information may be missed between due to tedious manual procedure.

### 3. Literature Review

Recently many computer based diagnosis tools are running around the hospitals, this work summarizes most recent and highly successful methods for classification of Cardio Vascular Diseases (CVD) and how it can be predicted with various hard and soft computing tools. If we can able to predict the Heart diseases with new technologies in CAD, which will save the life of many patients. Therefore, computer-based beat classification is essential and becoming the norm in clinical applications (Özbay 2009). There are several techniques such as maximum likelihood, (artificial) neural networks (Jehan Zeb Shah 2006), and support vector machines (Osowski et al 2004, Moavenian et al 2010) have been introduced for the ECG beat classification. These machine learning techniques identifies and matches new data instances based on the information extracted from the annotated training data in the learning phase. Most techniques provide a global classifier, that may not be always accurate for patient-specific cardiac variations.

Farid Melgani & Yakoub Bazi (2008) proposed a two folded system. First, the experimental study to show the superiority of the generalization capability of the Support Vector Machine (SVM) approach in the automatic
classification of ECG beats. Second, to propose a novel classification system based on Particle Swarm Optimization (PSO) to improve the generalization performance of the Support Vector Machine (SVM) classifier. Optimized SVM classifier is created by searching for the best value of the parameters that tune its discriminant function, and upstream by looking for the best subset of features that feed the classifier. The experiments were conducted on the basis of ECG data from the MIT-BIH arrhythmia database to classify five kinds of abnormal waveforms and normal beats. In particular, they were organized, so as to test the sensitivity of the SVM classifier and that of two reference classifiers used for comparison, i.e., the k-Nearest Neighbor (kNN) classifier and the Radial Basis Function (RBF) NN classifier, with respect to the curse of dimensionality and the number of available training beats. The obtained results clearly confirm the superiority of the SVM approach as compared to traditional classifiers, and suggest that further substantial improvements in terms of classification accuracy can be achieved by the proposed PSO-SVM classification system. On an average, over three experiments making use of a different total number of training beats (250, 500, and 750, respectively), the PSO-SVM yielded an overall accuracy of 89.72% on 40438 test beats selected from 20 patient records against 85.98%, 83.70%, and 82.34% for the SVM, the kNN, and the RBF classifiers, respectively.

Eduardo Pasolli & Farid Melgani (2010) proposed Iterative procedure - SVM Classifier:

- Margin Sampling (MS) in which the samples of the learning set more close to the hyper planes between the different classes are chosen;
- Posterior Probability Sampling (PPS) in which posterior probabilities are estimated for each class. Then, the samples that maximize the entropy between the posterior probabilities are selected
- Query By Committee (QBC) in which a pool of classifiers is trained on different features to label the set of learning samples. More sophisticated training set initialization strategy could further improve the performance, have 79.07% accuracy

Automated Arrhythmia-diagnosis systems which can provide high-classification accuracy rates for inter and intra-patient variation cases, are still an active area of research. The accuracy of detection of each cardiac cycle is of great importance because it contributes significantly to the overall classification result. The markers are subsequently processed by the feature extraction module, where measurements are produced for wave amplitudes and durations. The collective term for the measurements produced is commonly referred to as the input feature vector, which is considered to describe the morphology of the current recorded signal. The module of feature selection is an optional stage, whereby the feature vector is reduced in size including only, from the classification viewpoint, what may be considered as the most relevant features required for discrimination. The classification module is the last stage of automated diagnosis. It examines the input feature vector, produces a suggestive hypothesis (Grant PM 1989).

The Wavelet Transform (WT) can be applied to extract the wavelet coefficients of
discrete time signals. This procedure utilizes multirate signal processing techniques. The multiresolution feature of the WT allows the decomposition of a signal into a number of scales, each scale representing a particular coarseness of the signal under study. The WT provides very general techniques (Al-Fahoum & Howitt 1999) which can be applied to many tasks in signal processing. One of the most important application is the ability to compute and manipulate data in compressed parameters which are often called features (Hyejung Kim et al 2010). Thus, the ECG signal, consisting of many data points, can be compressed into a few parameters. These parameters characterize the behaviour of the ECG signal. This feature of using a reduced number of parameters to represent the ECG signal is particularly important for recognition and diagnostic purposes (Barro et al 2002). So identifying quality ECG signal, which does not require filtering and relatively clean signals are desired. These signals are then analyzed using Naive Bayes classifier and SVM.

Abdelhamid Daamouche et al (2012) proposed a novel approach for generating the wavelet that best represents the ECG beats in terms of discrimination capability is proposed. It uses polyphase representation of the wavelet filter bank and formulates the design problem within a PSO framework. Experimental results conducted on the benchmark MIT/BIH Arrhythmia database with the state-of-the-art SVM classifier confirm the superiority in terms of classification accuracy and stability of the proposed method over standard wavelets (i.e., Daubechies and Symlet wavelets).

Llamedo Soria & Martínez (2007) developed a Step-wise, randomized method for feature subset selection, and Linear Discriminant Analysis (LDA) had been used for additional dimensional reduction. Three classifiers: linear, quadratic and Mahalanobis distance were evaluated, using a k-fold like cross validation scheme. Results in the training set showed that the best performance was obtained with a 28-feature subset, using LDA and a Mahalanobis distance classifier. This model was evaluated in the test dataset with the following performance measurements: global accuracy: 86%; for Supraventricular beats, Sensitivity (Se): 86%, Positive Predictivity (PP): 20%; for Ventricular beats Se: 71%, PP: 61%. This results show the feasibility of classification based on the multilead wavelet features, although further development is needed in subset selection and classification algorithms.

Anuradha & Veera Reddy (2008) proposed, four non-linear parameters considered for CA classification of the ECG signals are Spectral entropy, Poincaré plot geometry, Largest Lyapunov exponent and Detrended fluctuation analysis, which are extracted from heart rate signals. The inclusion of Artificial Neural Network (ANN) in the complex investigating algorithms yields very interesting recognition and...
classification capabilities across a broad spectrum of biomedical problem domains. ANN classifier was used for the classification, and an accuracy of 90.56% had been achieved.

de Chazal et al (2004) shows the method for the automatic processing of the ECG for the classification of heartbeats. The method allocates manually detected heartbeats to one of the five beat classes recommended by ANSI/AAMI EC57:1998 standard, i.e., Normal Beat, Ventricular Ectopic Beat (VEB), Supra Ventricular Ectopic Beat (SVEB), fusion of a Normal and a VEB, or unknown Beat type. Data was obtained from the 44 non-pacemaker recordings of the MIT-BIH arrhythmia database. The data was split into two datasets with each dataset containing approximately 50,000 beats from 22 recordings. The first dataset was used to select a classifier configuration from candidate configurations. Twelve configurations processing feature sets derived from two ECG leads were compared. Feature sets were based on ECG morphology, heartbeat intervals, and RR-intervals. All configurations used a statistical classifier model utilizing supervised learning. The second dataset was used to provide an independent performance assessment of the selected configuration. This assessment resulted in a Se of 75.9%, a PP of 38.5%, and a False Positive Rate (FPR) of 4.7% for the SVEB class. For the VEB class, the Se was 77.7%, the PP was 81.9%, and the FPR was 1.2%. These results shows an improvement on previously reported results for automated heartbeat classification systems.

Thakor et al (2002) proposed, an algorithm for detecting VF and Ventricular Tachycardia (VT) by the method of sequential hypothesis testing. The algorithm generates a binary sequence by comparing the signal to a threshold value. The probability distribution of the time intervals of the binary sequence is obtained, and the sequential hypothesis testing of 85 cases resulted in identification of (1) 97.64% VF and 97.65% VT episodes after 5 s and (2) 100% identification of both VF and VT after 7 s. The desired False Positive (FP) and False Negative (FN) error probabilities can be programmed into the algorithm. A key feature of the sequential method is that extra time for detection can be traded off for improved accuracy, and vice versa.

Costas Papaloukas et al (2002) developed a new rapid and robust computerized system which is examined in detecting changes in long duration ECG recordings. The system distinguishes these changes between ST-segment deviation and T-wave alterations and can support the produced diagnosis by providing explanations for the decisions made. The ESC ST-T database was utilized for evaluating the performance of the system. Se and PP accuracy were the performance measures used and the proposed system scored 92.02% and 93.77%, respectively, in detecting ST-segment episodes and 91.09% and 80.09% in detecting T-wave episodes.

Cardiac beat classification is a key process in the detection of Myocardial Ischemic episodes in the electrocardiographic signal. Goletsis et al (2004) proposed a multicriteria sorting method for classifying the cardiac beats as Ischemic or not. Through a supervised learning procedure, each beat is compared to preclassified category prototypes under five criteria. These criteria refer to ST segment changes, T wave alterations, and the patient's age. The difficulty in implementing the above criteria is the determination of the required method parameters,
namely the thresholds and weight values. To overcome this problem, Genetic Algorithm (GA) is employed, which, after proper training, automatically calculates the optimum values for the above parameters. A task-specific cardiac beat database was developed for training and testing the proposed method using data from the ESC ST-T database. Various experimental tests were carried out in order to adjust each module of the classification system. The obtained performance was 91% in terms of both Se and Sp and compares favourably to other beat classification approaches proposed in the literature.

Markus Zabel et al (2000) has dealt with the Total Cosine R-to-T (TCRT) reflects the spatial angle between depolarization and repolarization, akin to the venerable concept of the ventricular gradient. T-wave loop dispersion extends this concept, reflecting variability of the T-wave vector loop. The normalized T-wave loop area measures heterogeneity of principal components of the T wave within its loop, whereas T-wave morphology dispersion expresses morphological heterogeneity within the 12-lead ECG. This analysis resulted in a improved accuracy of the prediction of cardiac mortality from the 12-lead resting ECG. Analysis of the digital ECG recordings was performed in a fully automatic manner with a custom-developed software implemented on a personal computer. The analysis program performs a Singular Value Decomposition (SVD) of the ECG signal into a minimum dimensional space. Complexity Ratio (CR) was the ratio of the singular value of the second most significant component to the square root of the sum of the squares of all 8 singular values. Based on the decomposition, several descriptors were calculated of spatial and temporal variations of T-wave morphology and repolarization wave front direction.

Conclusion

Sudden Cardiac Death occurs due to ventricular malfunction and Myocardial Infarction. This can be predicted in advance with the ECG signals- P QRS T wave forms. This work compares all existing approaches to denoise the ECG signal, Extracts features and to classify different arrhythmia signals. The in-depth analysis of the ECG signal will clearly indicate how the heart function varies, whether it leads to sudden death or needs pacemaker for further functioning.

References


Acknowledgements

The authors thank the management and Principal of Kumaraguru College of Technology, Coimbatore and Bannari Amman Institute of Technology, Sathyamangalam for providing excellent computing facilities and encouragement.
Authors' information

S.N.Shivappriya received her B.E (ECE) degree from Periyar University in 2004. She obtained her M.E (Communication System) degree from Anna university Chennai in 2006. She has 11 years 6 months of teaching experience at college level. She is completed her doctoral programme as a research scholar in Bio-Medical Engineering under guidance of Dr. R.Harikumar at Department of ECE, Bannari Amman Institute of Technology Sathyamangalam. Currently she is working as Assistant Professor II in Department of ECE at Kumaraguru College of Technology, Coimbatore. Her area of interest is Bio signal Processing, Soft computing, Embedded and Communication systems. She is member of ISTE.

Dr. R. Harikumar received his B.E (ECE) degree from REC Trichy 1988.He obtained his M.E (Applied Electronics) degree from College of Engineering, Guindy, Anna university Chennai in 1990. He was awarded Ph.D. in I&C Engg from Anna university Chennai in 2009. He has 21 years of teaching experience at college level. He worked as faculty in the Department of ECE at PSNA College of Engineering & Technology, Dindigul. He was Assistant Professor in IT at PSG College of Technology, Coimbatore. He also worked as Assistant Professor in ECE at Amrita Institute of Technology, Coimbatore. Currently he is working as Professor ECE at Bannari Amman Institute of Technology, Sathyamangalam. He is guiding eight PhD theses in these areas. He has published Twenty Eight papers in International and National Journals and also published around sixty nine papers in International and National Conferences conducted both in India and abroad. His area of interest is Bio signal Processing, Soft computing, VLSI Design and Communication Engineering. He is life member of IETE, IEEE and ISTE.

A. Kalaiselvi received her B.E (ECE) degree in 2004. She obtained her M.E (VLSI) degree from Anna university Chennai in 2006. She has 12 years 6 months of teaching experience at college level. She is doing her doctoral programme as a research scholar in VLSI signal processing under guidance of Dr. A Vasuki at Department of ECE, Kumaraguru College of Technology Coimbatore. Currently she is working as Assistant Professor II in Department of ECE at Kumaraguru College of Technology, Coimbatore. Her area of interest is VLSI signal Processing, Soft computing, Embedded systems.