CLASSIFICATION AND COMPRESSION OF CARDIAC VASCULAR DISEASE TO ENHANCE RURAL HEALTH CARE SYSTEM USING SOFTCOMPUTING TECHNIQUES

C.Anushya Devi
PG Scholar, Department Computer Science and Engineering (PG), National Engineering College, Kovilpatti.

K.Vimala
Assistant Professor, Department Computer Science and Engineering (PG), National Engineering College, Kovilpatti.

N.Sathiya Rani
PG Scholar, Department Computer Science and Engineering (PG), National Engineering College, Kovilpatti

ABSTRACT

The detection of Cardiac Vascular Disease (CVD) is to save a life of a heart patients, with the help of Public Health Care Centers by transmitting ECG signals to nearby hospital server. In this paper we analyze the abnormalities found in the ECG signals by identifying the Normal, Bradycardia Arrhythmia, Tachycardia Arrhythmia and Ischemia signal using the method of Neuro Fuzzy Classifier. DWT coefficients are used to extract the relevant information from the ECG input data. The extracted features are analyzed and classified using Adaptive Neuro Fuzzy Inference System (ANFIS) as a Neuro Fuzzy classifier. The compression algorithm is performed by using Huffman coding. Unit blocks size optimization, adaptive threshold adjustment, and 4-bit-wise Huffman coding methods are applied to reduce the processing cost while maintaining the signal quality.

Key Words - ECG, DWT, ANFIS, Huffman coding

I.INTRODUCTION

Electrocardiogram (ECG) is the electrical manifestation of the contractile activity of the heart that can be recorded fast and automatically. It is a noninvasive diagnostic tool, meaning that ECG signal can be measured without entering the body at all. The analysis system can bring the possibility to record the heart condition at early stage; the problem is being hard interpretation for non-trained people. Therefore the importance in developing the system that
make this interpretation easier for non-trained people and the system could detect the disease with high levels of accuracy because many people who died cause of heart disease showed no outward symptoms. This research is used by the other health care professionals including physicians, nurses, therapists and technicians to bring together knowledge from many technical sources to develop new procedures, or to solve clinical problems. The compression algorithm is used to transmit the ECG signal to hospital server efficiently. The compression minimizes the data amount during the compression. The Cardiac abnormalities can be identified fastly and efficiently using ECG signals. In this project, it is proposed to perform real-time classification of Cardiac Vascular Disease (CVD). Here the features selection for the ECG is done using feature extraction technique. The methods presented here are divided into three pieces of work. Firstly, procedures to identify and annotate of ECG signal for Normal, Bradycardia Arrhythmia, Tachycardia Arrhythmia and Ischemia characteristic. Secondly, a strategy is presented for extracting the features vector for each sample of selected heart disease using an algorithm that exploits the coefficient derived from Discrete Wavelet Transform. Lastly, this part presented the procedures of classification process using Adaptive Neural Fuzzy Inference System modeling.

II.RELATED WORKS

There are many techniques introduced by the researches to overcome the heart problems in day-today-life. The holter monitor system which records ECG signal continuously in ambulatory condition for a sizable time and system transmit the recorded data to hospital when the recorded period is completed.

The ECG analysis technique required the feature extraction and classifier stage. Before feature extraction the noise is removed from the ECG signal. For noise reduction many reduction techniques are proposed by different researchers. An approach to Cardiac Arrhythmia Analysis using Hidden Markov Models. This technique classified by detecting and analyzing QRS complex and determining the R-R intervals to determine the ventricular arrhythmias.

DWT-based feature extraction technique yields superior performance. ECG analysis using wavelet transform method can distinguish the between the QRS wave and P, T wave. This technique also can distinguish noise, baseline drift and artifacts. The two morphological feature extraction methods which are higher-order statistics and hermite basis functions.

A new approach to feature extraction which is Karhunen Lo’eve Transform (KLT) which is an attractive and powerful approach to the feature-extraction and shape representation process. It has the solution if the probability densities of population of pattern vectors of a problem domain are unknown. In designing an ECG classifier based on Neural Network (NN) , the normal procedure is to firstly train the network by presenting it with training data that is representative of the unknown data it is likely to experience during the classification process.

The Neuro-Fuzzy approach to the recognition and classification of the heart rhythms on the basis of ECG waveforms. It uses the new approach of heart beat recognition. This project is the resolution for the problem of less sensitivity to the morphological variation of the ECG.
III. METHODOLOGY

This system provides useful technique for fast identification and treatment of the heart disease. This paper provides the detection of cardiovascular disease in ECG signal by the help of health care centers in rural areas. In figure 1 the signal is recorded from the patient’s body and it is monitor by the physicians. The signal is then classified by the physicians by applying Neuro-Fuzzy classifier technique. This shows the result as whether the patient’s is in the normal, mild or severe condition. If the patient’s condition is critical, then the signal is transmitted to the hospital server by compressing the ECG signals using Huffman coding. For the severe cases physician will provide a tablet that could maintain the heart beat as normal for 6 hours which would help him to reach the Cardiac hospital. Then the patient can undergo for medical treatment by the cardiologist. If the patient’s condition is mild then they are advised to contact the cardiologist.

![Diagram](image.png)

**Figure 1:** Cardiovascular Abnormality Diagnosis Process.

A. DWT Technique for Feature Extraction

In this scope of this thesis, feature extraction was conducted by applying wavelet analysis techniques to patient data, thus providing ECG characteristic point detection capabilities. Since most recently published detectors are based on standard database libraries and limited wave detection, this application is an attempt to expand the horizons of current research efforts. The procedures of DWT implementation is describe as follow in Figure 2.
The general wavelet decomposition of DWT procedure involves three steps. The result of decomposed signal will shows the important details and approximation coefficients. Which represent the original signal. The basic version of the procedure follows the steps described below:

- Choose a wavelet types.
- Choose a wavelet name.
- Choose a level N which will compute the wavelet decomposition of the signals at level N.

The DWT wavelet types have been chosen in this features extraction method and the ECG signals were decomposed into time-frequency representations using single-level one-dimensional wavelet. The wavelet names of Daubechies wavelet filters db4 have been choosing and the number of decomposition levels was chosen to be 5. Thus, the ECG signals were decomposed into the details coefficients D1-D5 and one final approximation coefficient, A5.

The results of applying the Daubechies wavelet of order 4(db4) which is more suitable to detect changes of ECG signal is evaluated. The wavelet filter with scaling function more closely similar to the shape of the ECG signal achieved better detection. Db wavelet family is similar in shape to ECG signal and their energy spectrums are concentrated around low frequencies the signal is approximated by omitting the signals high frequency components. In this study, from the original intervals of ECG signal, five standard measures parameters used are used. A signal of 75 discrete data was selected as considered ECG signals data. For each ECG signals, the detail wavelet coefficients of fourth level (75 coefficients) were computed. In order to reduce the dimensionality of feature vectors, statistics over the set of the wavelet coefficients were used. The following statistical features were used to represent the time-frequency distribution of the ECG signals:
1. Energy of the wavelet of each ECG signals sample
2. Maximum of the wavelet coefficients of each ECG signals sample.
3. Minimum of the wavelet coefficients of each ECG signals sample.
4. Mean of the wavelet coefficients of each ECG signals sample.
5. Standard deviation of the wavelet coefficients of each ECG signal sample.

B. ANFIS implementation in classifying

**Heart Disease**

The classification was performed using the ANFIS in Fuzzy Logic Toolbox. ANFIS were trained with the back propagation gradient descent method in combination with the least squares method. The block of featured processed in ANFIS were shown in Figure 4. ANFIS required a predefined network structure and its membership function as well as other parameters can be trained during the learning process. The system is first designed using Sugeno Fuzzy Inference System (FIS).

---

**Figure 3**: Feature Extraction using DWT Technique
There are two types of FIS namely Grid Partition and Subtractive Clustering. The ANFIS Grid Partition was adopted in this study because this system required the number of membership functions for each input. This system uses the gbell shaped membership function to characterize the fuzzy set input and Sugeno output membership functions are linear types. In the layer 1, there are five nodes have been used for each input dimension $X_i$ where $i=1,2,...,d$ and $d$ is the member of input dimensions. The ANFIS which constructs a FIS, whose membership function parameters are tuned using a backpropagation algorithm in combination with a least squares type of method, will allows fuzzy systems to learn from the data that they are modeling. The FIS of heart disease classification. Based on five input-one output systems, the five variables were used which are Energy, Maximum, Minimum, Mean and Standard Deviation of DWT coefficients and the output class either Normal, Bradycardia Arrhythmia, Tachycardia Arrhythmia or Ischemia is taken as the output variable. The input parameters are represented by fuzzy set or linguistic Variables.

C. Lossless Compression: Huffman Coding

The delta coding and the lossless compression algorithm are adapted after the skeleton method. The delta coding is calculated with difference with the previous number ($y_i = x_i - x_{i-1}$). stores the changes instead of the absolute values to reduce the number of bit. The Huffman coding is selected because it provides minimum encoding cost when the original data has the unique distribution [5]. According to the Huffman Coding scheme, the most frequently occurring values have the shortest bit code length, and the rarely occurring data have the longest bit code length. After the skeleton step, the input data have Gaussian distribution, which is more than 50% of the data are located near the zero. Thus, these high frequently occurring data can be transformed with the short length of code by the Huffman coding.

The modified Huffman coding method transforms the sample-oriented format into the 4-bitwordoriented stream. So, it can obtain the unified data format for the efficient memory access, although the variable sample resolution is provided. When decoding the Huffman code, the bit stream is decoded into the 4-bit-word. The first bit is picked up and compared
with the Huffman table, and the original value is reconstructed from the remaining encoded data. The average CR of the Huffman coding is approximately 2:1 without the compression error rate.

IV. RESULT AND PERFORMANCE ANALYSIS

In this study, training and test sets were formed by 52 data train the samples. The 27 data samples were used for training and 25 data samples were for testing. The training dataset was used to train the ANFIS, whereas the testing dataset was used to verify the accuracy and the effectiveness of the trained ANFIS model for the detection of heart disease patients.

Table 1: Statistic of correct and incorrect Heart Disease Classification

<table>
<thead>
<tr>
<th>Heart disease type</th>
<th>Class</th>
<th>Correct classified</th>
<th>Misclassified</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>1</td>
<td>43</td>
<td>1</td>
</tr>
<tr>
<td>Bradycardia Arrhythmia</td>
<td>2</td>
<td>40</td>
<td>2</td>
</tr>
<tr>
<td>Tachycadia Arrhythmia</td>
<td>3</td>
<td>44</td>
<td>0</td>
</tr>
<tr>
<td>Ischemia</td>
<td>4</td>
<td>41</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>168</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 1 and Figure 5 show the correct classified and misclassified data samples of heart disease for each class. The MIT-BIH Arrhythmia Database is used to evaluate the performance. The sampling rate and the resolution are 360 samples/s and 12 bits, respectively. In addition, the Gaussian white noise source is injected from −10 to 20 dB SNR for the noise stress test. The test vector is produced by injection of the noise source with MIT-BIH record 100. The performances of compression coding are evaluated by the CR and the compression error rate of the percentage root mean square difference (PRD).

![Figure 5: Statistic of Heart Disease Classification](image-url)
V. CONCLUSION

This paper indicates that by using DWT and ANFIS, the classification of Normal, Bradycardia Arrhythmia, Tachycardia Arrhythmia and Ischemia signals can be classified; therefore the primary objective of this study is achieved. The simulation results show that the classification of heart disease is well predicted using DWT and ANFIS system and the system working well since it achieves the 97.68% of classification accuracy rate. This result indicates that it has some potential and had been found to be successful in heart disease detection. The proposed QLV delineate the ECG signal, and it supports both the flows to achieve better performance with low computation complexity. By applying the QLV, the overall CR improves while maintaining the small compression error rate, and high-accuracy performance can achieve at the segmentation and the R-peak detection stage. The performance is evaluated using the MIT/BIH Arrhythmia Database, and the noise robust test is also performed for the reliability of the algorithm. The average CR is 16.9:1 with 0.641% PRD value, and the encoding rate is 6.4 kbps. The accuracy performance of the R-peak detection is 100% without noise and 97.5% at the worst case with −10 dB SNR noise.

VI. REFERENCES