بسم الله الرحمن الرحيم

BACHELOR OF SCIENCE IN ELECTRICAL AND ELECTRONIC ENGINEERING

AN APPROACH FOR CLASSIFYING ECG ARRHYTHMIAS BY FEATURE EXTRACTION METHOD

By

AHMED FARHAN (092477)

MD. SANAUL ISLAM (092478)

S.M. TOWHID HOSSAIN (092458)

Supervised by

Prof. Dr. Mohammad Rakibul Islam

Department of Electrical and Electronic Engineering

Islamic University of Technology (IUT)

October, 2013

Contents

List of Table

List of Figures

ACKNOWLEDGEMENTS

The first thank and honor go to the Almighty. He has given us the capability and opportunity to finish this work perfectly. We have tried our best through the whole year and this research is our most significant scientific accomplishment in our educational life. But, without His help, it would not be possible for us.

After that, we would like to thank our respective supervisor, Prof. Dr. Mohammad Rakibul Islam for his guidance, motivation and help during the thesis work. He has worked very hard and helped us a lot to finish and finalize this research work.

We would also like to thank other respective teachers, our friends and family members for their support and motivation which were also the key points behind our success.

Abstract

The irregularity of heart beat is known as arrhythmia. This disease is sometimes very dangerous for a human body. It can cause death. So, proper treatment is a must in this case. Before treatment, proper diagnosis is needed. There are different types of arrhythmias. Treatment is different for different types. So, efficiently and correctly classification of ECG arrhythmias is a great challenge for treatment of this disease. There are a lot of techniques to classify the ECG signal. Some of them are very efficient, but the process is complicated. Again, some of them are easy and simple, but not very efficient. So, it is a great challenge for the researchers to find out a simple but efficient process for the classification. Our proposed method is quite simple and it is a general method for classifying ECG data. This method is based on extracting the features of ECG signal and classifying according to them. The efficiency of this method is quite good. We have used the MIT-BIH database for testing our method. This database is considered as the standard database in the whole world. We have trained some data for different types of arrhythmias and then taking those data as reference data, we tested other data and found out different features of the ECG signal. According to these features, we have classified Normal, Right Bundle Branch Block and Left Bundle Branch Block ECG signal. There are many other types of abnormal signal. In the future we will try to classify them. We have shown a comparison at the last of this thesis paper. From the comparison, we see that the efficiency of our method is good, but not better than most other methods. There are many scopes to improve our proposed method. So, we will try to improve the efficiency of this proposed method.

CHAPTER 1

INTODUCTION

An arrhythmia is the abnormal rhythm of heart. It is also known as dysrhythmia. It causes the heart to pump less effectively. There are a lot of changes in the shape of the heart wave because of arrhythmia. ECG is a common term in the diagnosis of cardiac diseases. It provides information about the electrical activity of the heart. We can detect different kinds of heart diseases by analyzing the ECG signal. Higher efficiency in classifying ECG signal is very important nowadays. Detection of actual type of heart diseases is very important for further treatment. Heart disease is very dangerous for a human being. So, proper treatment is a must in this case. There are different kinds of arrhythmias. Different kinds of arrhythmias can be detected in different parts of the heart. Heart pumps blood in a regular way. But when it is affected by arrhythmia, it can't pump blood normally. Right bundle branch block (RBBB), left bundle branch block (LBBB), premature ventricular contractions (PVC), ventricular fibrillation (VF) are some serious arrhythmias. However ECG being a non-stationary signal, the irregularities may not be periodic and may not show up all the time, but would manifest at certain irregular intervals during the day. So, continuous ECG monitoring permits observation of cardiac variations over an extended period of time, either at the bad side or when patients are ambulatory, providing more information to physician. Thus, continuous monitoring increases the understanding of patients' circumstances and allows more reliable diagnosis of cardiac abnormalities. Detection of abnormal ECG signals is a critical step in administering aid to patients. Often, patients are hooked up to cardiac monitors in hospital continuously. It requires continuous monitoring by the physician. Modern era of medical science is facilitated by computer aided feature extraction a disease diagnostics in which various signal processing techniques have been used to extract various kinds of features from the ECG signals. Extracting the characteristics of each arrhythmia in details through signal processing techniques into a feature vector capable of correctly classifying among different types of ECG arrhythmia is a difficult job. It is very complex to classify the arrhythmias and also to implement in real life application. So, the ultimate overall goal of classification of arrhythmia is to find the simplest and most effective method in order to perform the classification task with greater sensitivity, specificity and accuracy.

1.1 Electrocardiogram

Electrocardiogram known shortly as ECG is a diagnosis tool which reported the electrical activity of the heart. It is recorded by skin electrode. The heart rate and the morphology reflect the cardiac health of human heart beat [1]. It is a noninvasive method which means this signal is measured on the surface of human body, which is used in identification of the heart diseases [2]. Any disorder of heart rate or rhythm, or change in the morphological pattern, is an indication of cardiac arrhythmia, which could be detected by analysis of the recorded ECG waveform. The amplitude and duration of the P-QRS-T wave contains useful information about the nature of disease afflicting the heart. The electrical wave is due to depolarization repolarization of Na+ and K+ ions in the blood [2]. The ECG signal provides the following information of a human heart

- Heart position and its relative chamber size
- Impulse origin and propagation
- Heart rhythm and conduction disturbances
- Extent and location of myocardial ischemia
- Changes in electrolyte concentrations
- Drug-effects on the heart.

ECG does not afford data on cardiac contraction or pumping function [3].

1.2 The Heart Anatomy

The heart contains four chambers that is right atrium, left atrium, right ventricle, left ventricle and several atrio-ventricular and sino-atrial node as shown in 'Figure-1.1' [1]. The two upper chambers are called the left and right ventricles. The atria are attached to the ventricles by fibrous, non-conductive tissue that keeps the ventricles electrically isolated from the atria. The heart conduction system is shown in 'Figure-1.2' [1]. The right atrium and right ventricle together form a pump to the circulate blood to the lungs. Oxygen-poor blood is received through large veins called the superior and inferior vena cava and flows into the right atrium. The right ventricle then pumps the blood to the lungs where the blood is oxygenated. Similarly, the left atrium and the left ventricle together form a pump circulate oxygen-enriched blood received from the lungs (via the pulmonary veins) to the rest of the body [4].

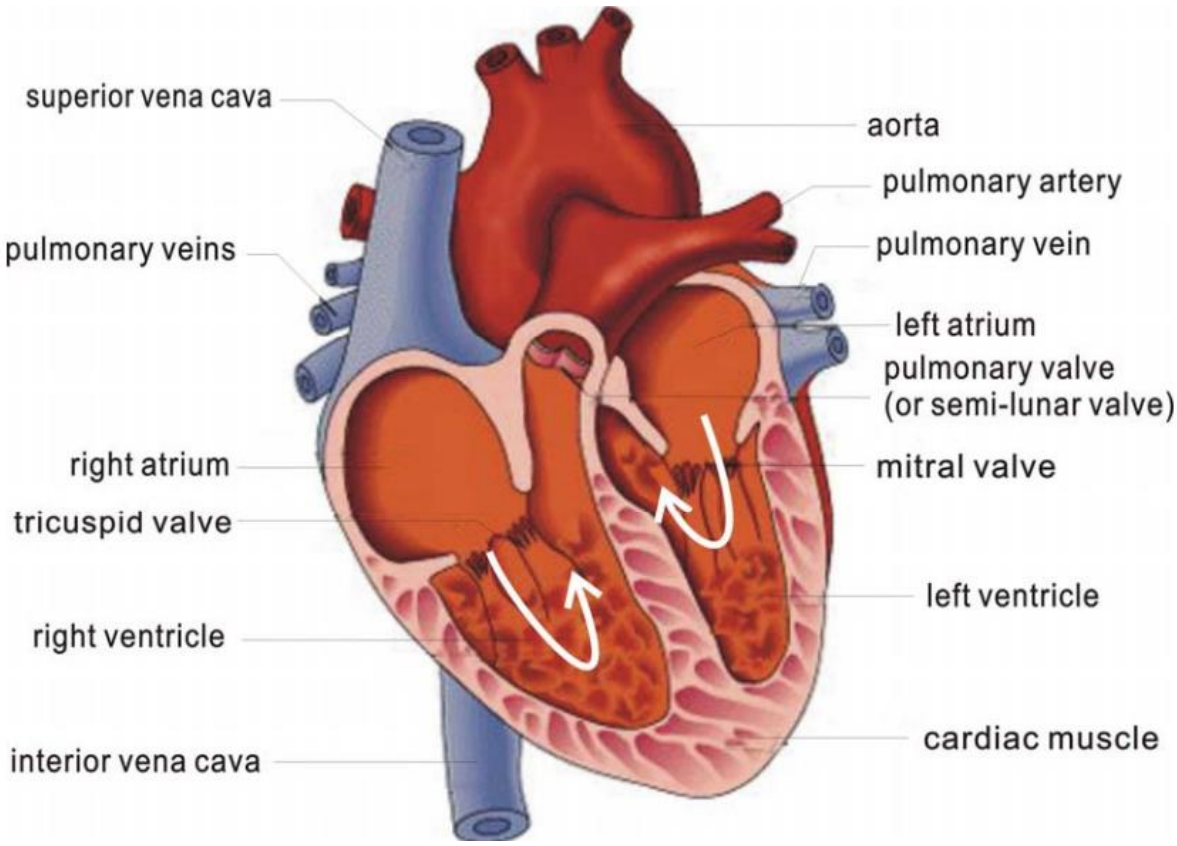

Figure 1.1: Basic Structure of Heart

In heart Sino atrial(S-A) node spontaneously generates regular electrical impulses, which then spread through the conduction system of the heart and initiate contraction of the myocardium. Propagation of an electrical impulse through excitable tissue is achieved through a process called depolarization. Depolarization of the heart muscles collectively generates a strong ionic current [1]. This current flows through the resistive body tissue generating a voltage drop. The magnitude of the voltage drop is sufficiently large to be detected by electrodes attached to the skin. ECGs are thus recordings of voltage drops across the skin caused by ionic current flow generated from myocardial depolarization [5]. Atrial depolarization results in the spreading of the electrical impulse through the atrial myocardium and appears as the P-wave. Similarly, ventricular depolarization results in the spreading of the electrical impulse throughout the ventricular myocardium.

1.3 Leads in ECG

The standard ECG has 12 leads which include 3 bipolar leads, 3 augmented unipolar leads and 6 chest (precordial) leads. In 'Figure-1.3' all leads' position are shown in human body. A lead is a pair of electrodes (+ve and –ve) placed on the body in designated anatomical locations and connected to an ECG record [3]. Bipolar leads record the potential difference between two points (+ve and –ve poles), whereas unipolar leads record the electrical potential at a particular point by means of a single exploring electrode.

Table 1.1: Types of leads used in ECG monitoring

Einthoven Leads	Limb Leads	Chest Leads
Bipolar leads	Unipolar leads	Unipolar leads
Lead I Lead II Lead III	aVR aVL aVF	V1 V2 V3 V4 V5 V6

1.3.1 Einthoven Leads

- Lead I : record potential between the left and right arm
- Lead II : between the right arm and left leg and
- Lead III : those between the left arm and the left leg

1.3.2 Unipolar limb leads

- aVR: When the +ve terminal is on the right arm
- aVL: When the +ve terminal is on the left arm and
- aVF: When the +ve terminal is on the left leg

One lead connected to +ve terminal acts as the different electrode, while the other two limbs are connected to the –ve terminal serve as the indifferent electrode [5]. Wilson leads (V1-V6) are unipolar chest leads positioned on the left side of the thorax in a nearly horizontal plane. The indifferent electrode is obtained by connecting the 3 standard limb leads. When used in combination with the unipolar limb leads in the frontal plane, they provide a tree-dimensional view of the integral vector.

1.3.3 Chest (periodical) leads

- V1: 4^{th} intercostal space, right sternal edge
- V2: 4^{th} intercostal space, left sternal edge
- V3: between the 2^{nd} and 4^{th} electrodes
- V4: 5^{th} intercostal space in the mid clavicular line
- V5: on 5^{th} rib, anterior auxiliary line
- V6: in the mid auxiliary line

To make recordings with the chest leads (different electrode), the three limb leads are connected to form an indifferent electrode with high resistances. The chest leads mainly detect potential vectors directed towards the back. These vectors are hardly detectable in the frontal plane [1]. Since the mean QRS vector is usually directed downwards and towards the left back region, the QRS vectors recorded by leads V1-V3 are usually negative, while those detected by V5 and V6 are positive [5]. In leads V1 and V2, QRS= -ve because, he chest electrode in these leads is nearer to the base of the heart, which is the direction of electronegativity during most of the ventricular depolarization process. In leads V4, V5, V6, QRS= +ve because the chest electrode in these leads is nearer the heart apex, which is the direction of electro-positivity during most of depolarization [3].

1.4 ECG Waves and Interval

ECG signal represents the condition of heart and the contraction of blood by different parts of hearts.

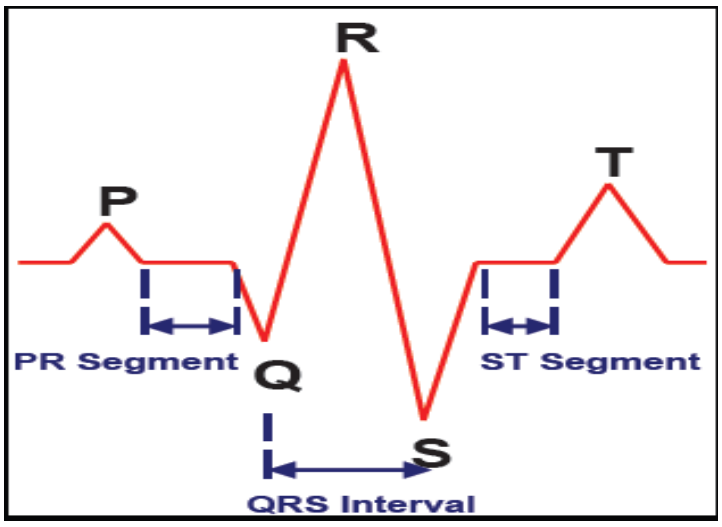

Figure 1.2: Basic ECG Signal

1.4.1 P Wave

P wave represents the sequential activation of the right and left atria and it is common to see notched or biphasic P waves of right and left activation. It is very difficult to analyze P waves with a high signal-to-noise ratio in ECG signal.

A clear P wave before the QRS complex represents sinus rhythm. Absence of P waves may suggest atrial fibrillation, sinus node arrest, junctional rhythm or ventricular rhythm.

1.4.2 PR Segment

The PR segment connects the P wave and the QRS complex. The impulse vector is from the AV node to the bundle of His to the bundle branches and then to the Purkinje Fibers. This electrical activity does not produce a contraction directly and is merely travelling down towards the ventricles and shows up flat on the ECG. The PR interval is more clinically relevant than PR segment to any abnormality.

1.4.3 PR Interval

The PR interval is measured from the beginning of the P wave to the first part of the QRS complex. It includes time for atrial depolarization, conduction through the AV node and conduction through the His- Purkinje system.

1.4.4 QRS Complex

The QRS Complex is the largest voltage deflection of approximately 10-20 mV but may vary in size depending on age and gender. The QRS represents the simultaneous activation of the right and left ventricles, although most of the QRS waveform is derived from the larger left ventricular musculature.

Duration of the QRS complex indicates the time for the ventricles to depolarize and may give information about conduction problems. Wide QRS complex indicates right and left bundle branch block, ventricular flutter or fibrillation etc.

1.4.5 T Wave

T wave represents ventricular repolarization. The interval from the beginning of the QRS complex to the apex of the T wave is referred to as the absolute refractory period. Tall T wave may suggest left bundle branch block, ventricular hyper trophy, hyperkalemia, stroke whereas small, flattened or inverted T wave implies myocardial ischemia, myocarditis, anxiety, certain drugs, right bundle branch block, hypokalemia etc.

1.4.6 ST Segment

The ST segment occurs after ventricular depolarization has ended and before repolarization has begun. The ST segment is usually isoelectric and has a slight upward concavity. It is always measured from the J point, where the QRS and ST segments meet. Elevation of ST segment implies myocardial ischemia, acute MI, LBBB, left ventricular hypertrophy, hyperkalemia, hypothermia whereas depression of ST segment may suggest myocardial ischemia, acute posterior MI, LBBB, RBBB etc.

1.4.7 ST Interval

The ST interval is measured from the J point to the end of the T wave.

1.4.8 QT Interval

The QT interval consists of the QRS complex (representing only a brief part of the interval), along with the ST segment and T wave, which constitutes the majority of the duration. The QT interval is used primary as a measure of membrane repolarization. Since the QT interval varies with heart rate, the QT interval is 'corrected' to make comparisons between ECG beats

1.4.9 U wave

The U wave is hypothesized to be caused by the repolarization of the inter-ventricular septum. Their amplitude is normally one third of the following T wave and even more often completely absent. It may be seen following the T wave and can make interpretation of the QT interval especially difficult. This is associated with metabolic disturbances, typically hypokalemia and ischemia.

Table 1.2: Amplitude and duration of waves, intervals and segments of ECG signal

Features	Amplitudes (mV)	Duration (mS)
P wave	0.1-0.2	60-80
PR Segment	_	50-120
PR interval	_	120-200
QRS complex	1	80-120
T wave	0.1-0.3	120-160
ST segment	_	80-120
ST interval	_	320
QT interval	_	300-420
RR interval	_	(0.4-1.2) s

1.5 Arrhythmias in ECG Signal

The normal rhythm of the heart where there is no disease or disorder in the morphology of ECG signal is called Normal sinus rhythm 'Figure-1.5(a)'. The heart rate of NSR is generally characterized by 60 to 100 beats per minute. The regularity of the R-R interval varies slightly with the breathing cycle. The source of the rhythm is the sino-atrial node, which is the normal

pacemaker of the heart. Hence another characteristic feature of NSR is a normal P-wave followed by a normal QRS-complex.

When the heart rate increases above 100 beats per minute, the rhythm is known as sinus tachycardia 'Figure-1.5(b)'. This is not an arrhythmia but a normal response of the heart which demand for higher blood circulation [1]. If the heart rate is too slow then this is known as bradycardia and this can adversely affect vital organs. When the heart rate is too fast, the ventricles are not completely filled before contraction for which pumping efficiency drops, adversely affecting perfusion. Rhythms that deviate from NSR are called arrhythmia since they are abnormal and dysfunctional. Arrhythmias may be easily understood by categorizing them in the following manner:

1. Sinus node Arrhythmias

2. Atrial Arrhythmias

3. Junctional Arrhythmias

4. Ventricular Arrhythmias

5. Atrio-ventricular Blocks

6. Bundle Branch and Fascicular Blocks

1.5.1 Sinus Node Arrhythmias

This type of arrhythmia arises from the S-A node of heart. As the electrical impulse is generated from the normal pacemaker, the characteristic feature of these arrhythmias is that P-wave morphology of the ECG is normal. These arrhythmias are the following types: Sinus arrhythmia, Sinus bradycardia and Sinus arrest etc.

1.5.1.1 Sinus Arrhythmia

This is not a disorder or a true arrhythmia, but a normal, physiologic variation in the sinus rate with the phases of respiration. The slowest instantaneous heart beat may less than 60 beats per minute, while the highest may exceed 100 beats per minute.

1.5.1.2 Sinus Bradycardia

In sinus bradycardia, the rhythm originates from the S-A node but at a rate of less than 60 beats per minute 'Figure-1.5(c)'. The ECG appears normal except for the slow heart rate. Mild sinus bradycardia is usually asymptomatic, while marked sinus bradycardia may lead to hypotension and result in insufficient perfusion of the brain and other vital organs. Treatment is indicated if the bradycardia is symptomatic.

1.5.2 Atrial Arrhythmias

Atrial arrhythmias originate outside the S-A node but within the atria in the form of electrical impulses. These arrhythmias types are given below:

1.5.2.1 Premature Atrial Contractions (PAC)

This type of arrhythmia results an abnormal P-wave morphology followed by a normal QRS complex and a T-wave. Because of an ectopic pacemaker firing before the S-A node, this happens. PACs may occur as a couplet where two PACs as generated consecutively. When three or more consecutive PACs occur, the rhythm is considered to be atrial tachycardia.

1.5.2.2 Atrial Tachycardia

The heart rate atrial tachycardia is fast and ranges from 160 to 240 beats per minute in atrial tachycardia. Frequently atrial tachycardia is accompanied by feelings of palpitations, nervousness, or anxiety.

1.5.2.3 Atrial Flutter

In atrial Flutter, the atrial rate is very fast, ranging from 240 to 360 per minute. The abnormal P-waves occur regularly and so quickly that they take morphology of saw tooth waveform which is called flutter waves.

1.5.2.4 Atrial Fibrillation

The atrial rate exceeds 350 beats per minute in this type of arrhythmia. This arrhythmia occurs because of uncoordinated activation and contraction of different parts of the atria. The higher atria

rate and uncoordinated contraction leads to ineffective pumping of blood into the ventricles. Atrial fibrillation may be intermittent, occurring in paroxysms or chronic [1].

1.5.3 Ventricular Arrhythmias

In this type of arrhythmia, the impulses originate from the ventricles and move outwards to the rest of the heart. In ventricular arrhythmia, the QRS complex is wide and bizarre in shape.

1.5.3.1 Premature Ventricular Contractions (PVC)

In PVC the abnormality is originated from ventricles. PVCs usually do not depolarize the atria or the S-A node and hence the morphology of P-waves maintains their underlying rhythm and occurs at the expected time. PVCs may occur anywhere in the heart beat cycle. PVCs are described as isolated if they occur singly and as couplets if two consecutive PVCs occur.

1.5.3.2 Ventricular Tachycardia (VT)

The heart rate of ventricular tachycardia is 110 to 250 beats per minute. In VT the QRS complex is abnormally wide, out of the ordinary in shape and of a different direction from the normal QRS complex. VT is considered life-threatening as the rapid rate may prevent effective ventricular filling and result in a drop in cardiac output.

1.5.3.3 Ventricular Fibrillation

It occurs when numerous ectopic pacemakers in the ventricles cause different parts of the myocardium to contract at different times in a non-synchronized fashion

1.5.4 Bundle Branch Block

Bundle branch block, cease in the conduction of the impulse from the AV-node to the whole conduction system. Due to this block there may occur myocardial infarction or cardiac surgery [1].

1.5.4.1 Right Bundle Branch Block (RBBB)

When the right bundle branch is blocked, the electrical impulse from the AV node is not able propagate to the Purkinje network to depolarize the right ventricular myocardium. Instead the impulse propagates in a convoluted manner through the left ventricle myocardium to reach the

right ventricular myocardium. Since propagation through the myocardium is much slower than through the specialized conducting tissue, the QRS complex becomes widened.

1.5.4.2 Left Bundle Branch Block (LBBB)

For LBBB, the right ventricle is depolarized primarily, generating an electrical wave front that eventually spreads to the left ventricular myocardium causing the myocardium to depolarize.

1.6 Objective of the thesis

The objectives of this thesis are:

- To develop a simple method for classifying ECG data efficiently.
- To investigate the performance of the proposed method for the classification of three types of arrhythmia using ECG signals available from the MIT-BIH arrhythmia database.
- To find out the sensitivity, specificity and accuracy of the proposed method.

1.7 Thesis Layout

This thesis comprises seven chapters.

- **Chapter 1** represents brief description on heart anatomy, basic ECG signal, different types of arrhythmias and the objective of the whole thesis work.
- **Chapter 2** represents a brief description on fuzzy logic, fuzzy classifier and different types of fuzzy rule based classifier.
- **Chapter 3** represents different methods for computing fractal dimension and ECG arrhythmia classification.
- **Chapter 4** represents brief description on Gaussian mixture model and classification process by this model.
- **Chapter 5** represents our proposed method that is feature extraction method for ECG arrhythmia classification.
- **Chapter 6** represents the simulation result and performance comparison.
- **Chapter 7** concludes the whole work with a brief summary and some future suggestion.

CHAPTER 2

REVIEW ON FUZZY CLASSIFIER METHODS

Classification of ECG arrhythmias is very important to provide treatment to the patients. Different classifiers and methods have been invented to analysis and classify ECG arrhythmias. This chapter represents the review on different methods based on fuzzy logic. Fuzzy logic is a multivalued logic. It deals with intermediate values. This method is very efficient. It is similar to human thinking. Zadeh introduced this logic.

Fuzzy logic can be used in many cases. It can be used for classifying ECG arrhythmias which is focused in this paper. This method can be used in modulation classifier in non-ideal environment which is very difficult or impossible by precise probabilistic methods [15]. Fuzzy if-then rules are used in many image processing applications [16]. This logic can be used in detecting H1N1 influenza. A hybrid method using Artificial Neural Network, Fuzzy Rule-base Classifier and Gaussian Mixture Model can be developed to detect H1N1 correctly [17].

Fuzzy rule-based classification system can be used in case of pattern classification problem. For this, error correction-based learning procedure and additional learning procedure is needed [18]. Type-2 Fuzzy sets also can be used for pattern recognition [19]. Genetic Fuzzy and Neuro-fuzzy classifier can be used for credit scoring. The Genetic Fuzzy classifier is more accurate in this case [20]. Again, Fuzzy classifier and fuzzy classified data are used in clinical proteomics. It is a very important tool for study at the protein and peptide level in medicine and health care [21]. Fuzzy rules are also applied in high-resolution multispectral satellite imagery for classification of urban and suburban areas [22].

Fuzzy logic has a lot of advantages. It does not require precise and noise free input. It is flexible and very easy to manage. It keeps the total system cost low. The system can be easily designed for many inputs and outputs because of rule-based operation. It can model non-linear functions of arbitrary complexity. It can model and control those types of non-linear systems which would be difficult or impossible to model mathematically [14]. Fuzzy logic also has many disadvantages. Development of fuzzy logic and the membership functions is quite difficult and tedious. The

analysis can be difficult with the interpretation of the fuzzy outputs in different ways. This process needs a lot of data [13]. Sometimes, it is difficult to collect too many data for this process.

2.1 Fuzzy logic

. Computer software understands only binary functions. Conventional Aristotelian and Boolean logic deal with only the absolute values of 0 and 1 or true or false. But, fuzzy logic is not like that. It expresses like medium, smaller and higher. Everything in this world cannot have absolute value and they cannot follow a linear function always. Fuzzy logic considers these phenomena. In fuzzy logic, everything is a matter of degree. A classical set of binary logic deals with crisp values while fuzzy sets have fuzzy values. It contains linguistic variables. It can be defined as low, medium and high. The values have fuzzy margins. They can overlap each other [13].

2.2 Fuzzy Inference System

Fuzzy inference system (FIS) is a framework. This system is based on fuzzy sets, fuzzy rules and fuzzy reasoning. Fuzzy reasoning is an approximate reasoning. This process draws conclusions from fuzzy sets and fuzzy rules. It has four components. They are fuzzifier, rule base, inference engine and defuzzifier. In fuzzifier, fuzzy sets are created from crisp value. Then, fuzzy rules are formed. The inference engine applies the fuzzy rules on the fuzzy sets. The fuzzy output is determined from this. The resultant is fuzzy value. So, in order to get crisp value as output, defuzzification process is needed [30].

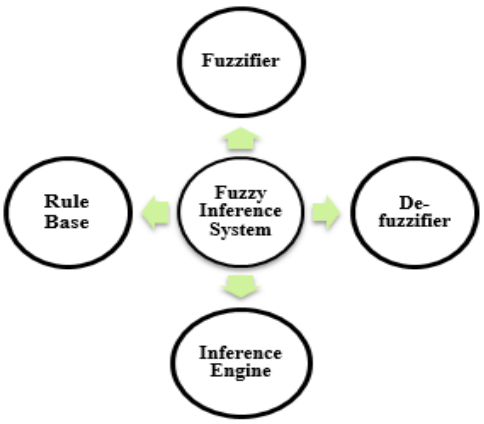

Figure 2.1**:** Fuzzy Inference System

2.3 Fuzzy Classifier for ECG arrhythmias detection

A classifier is a system which is based on a rule-based algorithm. It predicts the class. The classifiers use training algorithms and training data sets. It can be designed from prior knowledge when training data set is not available. After being trained, the classifier is ready for operation on unseen objects [11]. The classifier follows some rules and they are represented in the form of "If condition Then action" [12].

2.3.1 Simple Fuzzy Classifier

There are two major function blocks in this fuzzy classifier. They are ECG parameterizer and Fuzzy Classifier. In the ECG Parameterizer; initializing, pre-processing, fuzzification and defuzzification is done. Then, the data is sent to fuzzy classifier for classification [1]. In the fuzzy classifier, the fuzzy logic and if-then rule is applied.

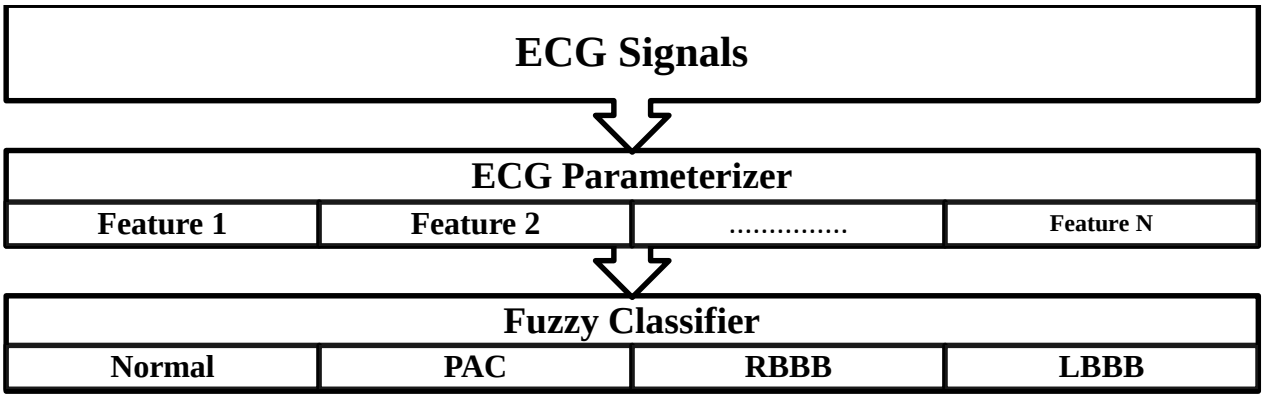

Figure 2.2**:** Structure of Fuzzy Classifier [23]

For the classification of ECG arrhythmias by fuzzy classifier, some ECG patterns and linguistic variables are needed. Some of them are given below.

Table 2.1: ECG patterns and linguistic variables [23]

ECG features	Medical linguistic variables
Prior-Heart Rate (RR_0)	{Short; Normal; Long}
Post-Heart Rate (RR_1)	{Short; Normal; Long}

P Wave	{Early; Normal; Disappear}
QRS Complex	{Upward; Downward}
R Wave Amplitude	{High; Normal; Low}
T Wave	{Upward; Downward; Disappear}

There are four steps through the whole classification system. They are initialization, preprocessing, fuzzification and defuzzification. In initialization, ECG data is read from the database that is used. In preprocessing step, ECG features are found out. The crisp values are converted into fuzzy or linguistic variables and membership functions are designed in the fuzzification step. The output comes as fuzzy value. So, in defuzzification step, fuzzy values are defuzzified to find crisp values as output [1].

Table 2.2: Definition of the membership function for Normal Beat [23]

Characteristic	**Feature**	**Function Type**	**Parameter (a)**	**Parameter (b)**
P up- ward	P peak value	S	0.10 mV	0.15 mV
QRS up-ward	R peak Value	**S**	0.70 mV	0.80 mV
T up-ward	T peak Value	**S**	0.10 mV	0.15 mV
RR_0	Prior-HR	Gaussian	80 bpm	20 bpm
RR_1	Post-HR	Gaussian	80 bpm	20 bpm

This definition is for normal beat. The ‘a’ and ‘b’ are based on the medical literature. After completing fuzzification, the incoming features are described by a membership value which is [0, 1]. The inference process uses if-then rule. For example:

IF (“Feature 1” is “Linguistic Variable 1”) AND

(“Feature 2” is “Linguistic Variable 2”) AND

…

("Feature N" is "Linguistic Variable N")

THEN (Class Name)

Now, for LBBB, the properties or statements are-

- The direction of T wave and QRS complex is opposite.
- Prior Heart Rate RR_0 is small.
- P wave is disappeared.

So, the LBBB classification can be defined as

IF ("T wave is upward") AND
("QRS wave is downward") AND
("P wave is disappeared") AND
("RR_0 is small")
THEN (Left Bundle Branch Block)

In this rule, the linguistic values are upward, downward, disappeared and small. The product of the membership grades [0, 1] is regarded as the hypothesis. If the product exceeds the limit, then the beat is in the class of LBBB.

2.3.2 Adaptive Fuzzy Classifier

For more accuracy, an adaptive fuzzy ECG classifier can be used. In the adaptive fuzzy ECG classifier, the membership functions come from general medical knowledge and datasets [23].
This method needs a learning stage in order to optimize the system parameters like threshold values and membership boundaries. The boundary values in the membership function are preset roughly during the learning stage. It is based on the medical knowledge. A specific beat is classified appropriately by using the predefined rule set. The boundary values in the membership function are modified using the significant values of these pre-classified beats. For each record, pre-classification and self-adaptation takes the first 10 minutes in the learning and testing purpose if the total time is 30 minutes. The rest 20 minutes is allocated for testing the updated AFC-ECG [23]. This process is quite perfect for classifying ECG signals.

Table 2.3: Performance Comparison of Fuzzy Classifier and Adaptive Fuzzy ECG Classifier [23]

Record	MIT-BIH Annotation		Fuzzy Classifier		Adaptive Fuzzy Classifier	
			Result	Accuracy	Result	Accuracy
Sig 106	Normal	1507	1208	80.2%	1258	83.5%
	PAC	0	7	-	1	-
	RBBB	0	30	-	27	-
	LBBB	0	0	-	1	-
	Unclassified	0	262	-	220	-

The average accuracy of the adaptive fuzzy ECG classifier is around 88.2%. It is about 77.0% in case of conventional fuzzy ECG classifier [23].

2.3.3 Fuzzy C-means Clustering method

Data clustering is used to find similarities in data. It puts the similar data into similar groups. Fuzzy C-means clustering (FCM) is a data clustering algorithm. In this process, each data point belongs to a cluster to a degree which is specified by a membership grade [25].

Table 2.4: Result of fuzzy classifier based on the FCM clustering [25]

Type	Sensitivity	Positive Predictivity	Accuracy
Normal	99.66%	96.76%	97.41%
LBBB	96.33%	98.29%	
RBBB	98.66%	95.79%	
PB	95.00%	98.95%	

In this method, the accuracy is 97.41%, which is a very good result for ECG arrhythmia classification.

2.3.4 Combination of Wavelet Transform and Fuzzy Neural Network for VPC Detection

This method is used to detect the ventricular premature contraction (VPC). The information that is used during QRS detection can be reused in this method. It is the main advantage of this method. The QRS duration is taken in scale three and the area under the QRS complex is taken in scale four. These are considered as the characteristic features. The R wave amplitude also has influence on the computation of the characteristic features. It is very necessary to normalize these features. The left bundle branch block beats are excluded here. The accuracy for VPC classification using FNN is 99.79%. Use of same wavelet has two advantages. There are less computation and less complexity in QRS detection and VPC classification [26].

2.3.5 Fuzzy Support Vector Machine

FSVM method is the concept of membership values into SVM. It is needed to define membership functions. For the calculation of degree of membership, classical membership functions are used. The data specific membership functions are also can be used. Four functions are used for assigning the degree of membership. They are one class weighing (OCW), distance to class mean (DTCM), Distance to One Class Mean (DTOCM), Cardinality (CAR), Fuzzy C-Means (FCM). The proposed method is applied to the UCI Arrhythmia Database. There are four techniques for dimension reduction. They are Principal Component Analysis (PCA), Factor Analysis (FA), Recursive Feature Elimination with Support Vector Machine (RFE-SVM) and Correlation Based Feature Selection (CFS). Using these membership functions, different accuracy was found [27].

Table 2.5: Accuracy Comparison

	PCA	**FA**	**RFE-SVM**	**CFS**
KNN15	67.14	76.43	78.33	74.29
MLP	72.86	79.52	79.05	80.48
SVM	78.57	82.62	82.14	81.43
FSVM-OCW	66.43	73.81	43.57	74.52

FSVM-DTCM	78.09	80.71	80.48	83.33
FSVM-DTOCM	78.33	80.95	80.71	82.86
FSVM-CAR	77.62	82.14	80.48	81.67
FSVM-FCM	78.33	82.14	81.19	81.90

2.3.6 Fuzzy-Genetic Based PCA and ICA

Principal Component Analysis (PCA) and Independent Component Analysis (ICA) are two methods for detecting different types of arrhythmia.

Fuzzy-Genetic based PCA (FGPCA) and Fuzzy-Genetic based ICA (FGICA) are the combination of Fuzzy C-Means (FCM) and Genetic Algorithm (GA) with PCA and ICA. They are very relevant methods [28].

Table 2.6: Comparison of accuracy between Fuzzy-Genetic Based PCA and ICA

Methods	**Accuracy**
PCA	86.7%
GPCA	90%
ICA	91%
GICA	93.3%
FGPCA	94.4%
FGICA	94.7%

So, the accuracy of FGPCA is 94.4% and FGICA is 94.7%.

2.3.7 Pruned fuzzy K-nearest neighbor classifier

The implementation of Fuzzy KNN is very easy. But, this process is very time consuming and needs large storage space. Pruned Fuzzy K-nearest neighbor is a good method for classifying ECG beats. It can classify six types of beats which are present in the MIT-BIH Arrhythmia database. The accuracy of this method is quite same as standard FKNN. But, this method reduces the computational complexity. The calculation time is very low comparison with FKNN. 11 features

are used in PFKNN. By using PCA, it can be reduced to 6 features. That means the calculation becomes easier. The accuracy of FKNN is 97.63% and PFKNN is 97.32% with 11 features and 97.31% with 6 features [30].

2.3.8 Adaptive neural fuzzy filter method

For early diagnosis of ECG arrhythmia, adaptive neural fuzzy filter method is a good technique. ANFF can learn by itself according to numerical training data or expert knowledge which is built up by the fuzzy if-then rules. In the structure of this method, there are five layers. Layer-1 nodes are input nodes. They represent input variables. Layer-2 and layer-4 nodes are called term nodes. They act as the membership functions. They represent the terms of respective input and output variables.

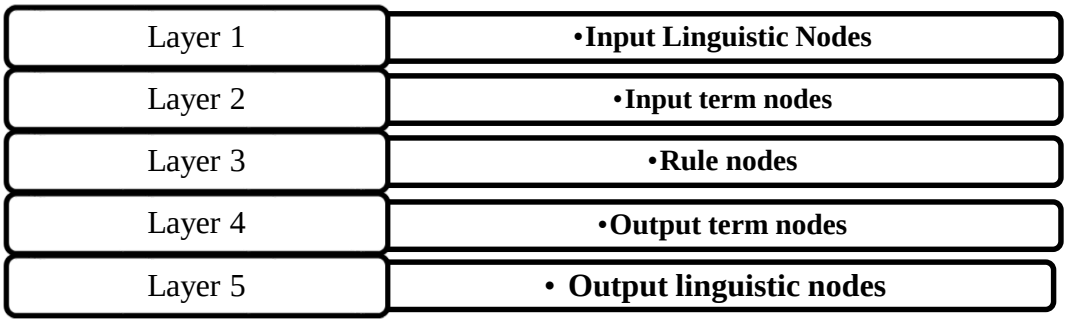

Figure 2.3: Structure of Adaptive Neural Fuzzy Filter (ANFF) [48].

The accuracy of this process is about 97.6% which is quite high. So, it is a very efficient method for classifying ECG beat.

2.3.9 Adaptive Neuro-Fuzzy Inference System

Adaptive Neuro-Fuzzy Inference System (ANFIS) is an efficient model for classification of Electrocardiogram (ECG) signals. Independent Component Analysis (ICA) is used for feature extraction. The feature extraction and Power spectrum with the RR interval serve as the input feature vector. This feature is used as input in the ANFIS classifier. Six types of ECG signals such as normal sinus rhythm (NSR), premature ventricular contraction (PVC), atrial premature contraction (APC), Ventricular Tachycardia (VT), Ventricular Fibrillation (VF) and Supraventricular Tachycardia (SVT) are classified by this method. It is the combination of

Neural Network Adaptive Capabilities and the Fuzzy Inference System. The accuracy is more than 97% [32].

2.3.10 Combination of fuzzy c-means clustering (FCMC) algorithm and neural networks

FCMCNN is the combination of fuzzy c-means clustering (FCMC) algorithm and neural networks. By this method, the correct classification rate is 99.99% [33].

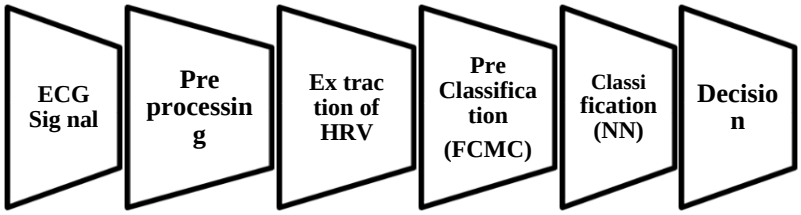

Figure 2.4: Block diagram of FCMCNN

2.3.11 Multi-class MLP ECG classifier using FCM

MLP system has inherent structure. This system is slow and bulky. Using FCM in the MLP system makes the total system fast and more efficient [34].

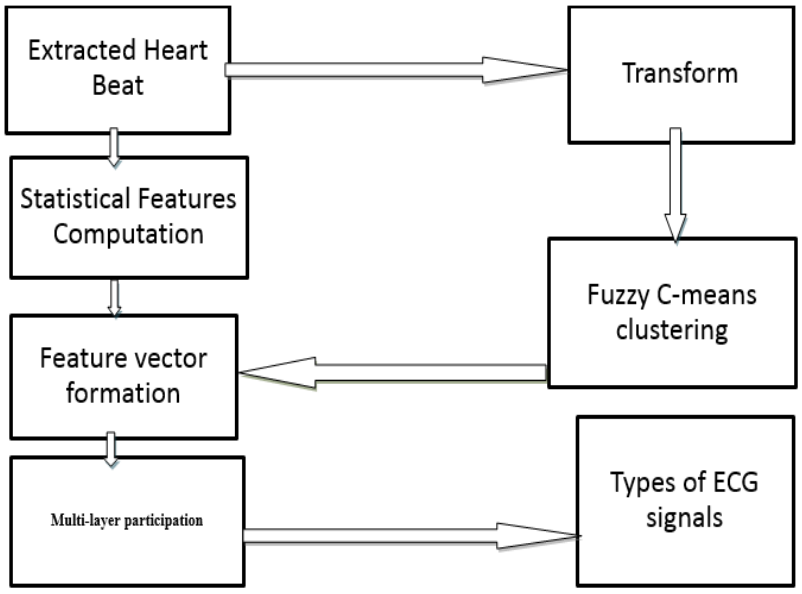

Figure 2.5: Block diagram of FCM based MLP ECG classifier

2.3.12 Type-2 fuzzy clustering neural network

Type-2 fuzzy c-means clustering is used to improve the performance of neural network. Different types of arrhythmias that are considered for this method are normal sinus rhythm (N), sinus bradycardia (Br), ventricular tachycardia (VT), sinus arrhythmia (SA), atrial pre-mature contraction (APC), paced beat (P), right bundle branch block (R), left bundle branch block (L), atrial fibrillation (A.Fib) and atrial flutter (A.Fl.) [35].

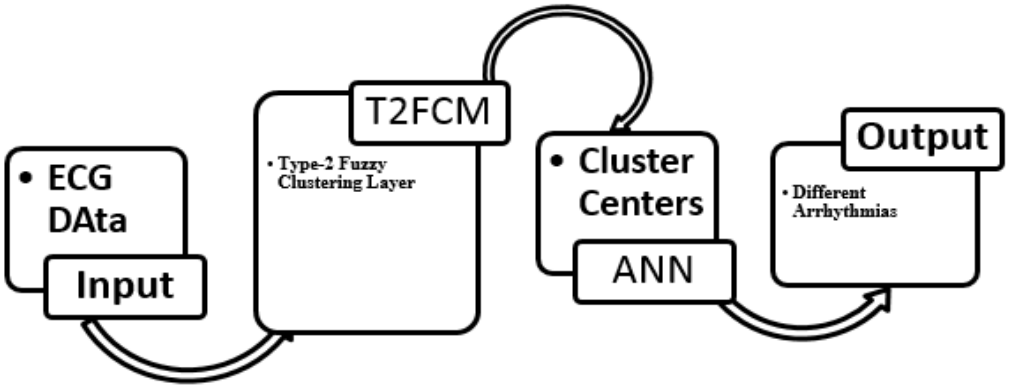

Figure 2.6: Optimum T2FCNN architecture [52].

The accuracy of this method is 99.99%. So, this technique is more efficient than other methods [35].

2.3.13 Fuzzy Gaussian Neural Network (FGNN)

FGNN is another method for heart disease diagnosis. Feature extraction from the QRST zone of ECG signals and Pattern classification for IHD diagnosis using the FGNN. This process has four layer structures. They are layer 1, layer 2, layer 3 and layer 4 [36].

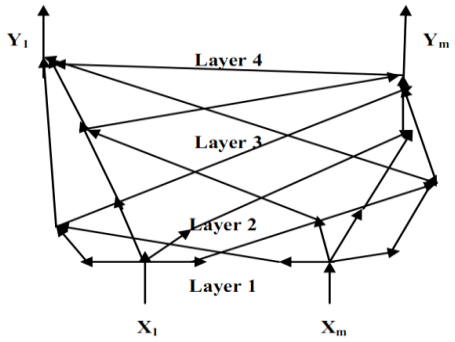

Figure 2.7: Structure of FGNN

2.4 Result Comparison of different methods using fuzzy logic

We have analyzed different methods, their methodology and their results. Some of the methods are very efficient for classifying ECG data. The comparison between the percentages of accuracy of the methods discussed above is given below in a tabular form.

Table 2.7: Comparison of accuracy between different Fuzzy logic methods

Method	Accuracy
Simple Fuzzy Classifier	93.13%
	93.78%
	77%
Adaptive Fuzzy ECG Classifier	88.2%
Fuzzy c-means clustering method	97.41%
Fuzzy Genetic based PCA method	94.4%
Fuzzy Genetic based ICA method	94.7%
Standard FKNN	97.63%
PFKNN (With 11 Features)	97.32%
PFKNN (With 6 Features)	97.31%
Adaptive Neural Fuzzy Filter	97.6%
Adaptive Neuro-Fuzzy Inference System	More than 97%
FCMCNN	99.99%
T2FCM	99.99%

2.5 Conclusion

We have focused an overview of different fuzzy logic methods and their accuracy in classifying ECG arrhythmias. We have seen that, combination of fuzzy logic with other methods is more efficient and helpful to classify ECG arrhythmias. Combination of other methods with fuzzy logic may be more efficient than the methods discussed above.

CHAPTER 3

REVIEW ON FRACTAL DIMENSION FEATURE EXTRACTION BASED ECG ARRHYTHMIA CLASSIFICATION

Fractals are of rough or fragmented geometric shapes. These shapes can be sub divided in parts. Our conventional geometry deals with lines, triangles, circles, sphere, cones but the fractal geometry is concerned with fractured shapes. Fractals are often mentioned by their fractal dimension [38]. If we consider a fracture object whose mass is function of its radius but the radius doesn't go like r^2 and neither does like r but rather it goes like a real power of r between 1 and 2. Such object with non-integer dimension is called fractals and the dimension of this type of object is called fractal dimension [39]. Fractal dimension measures the degree of fractal boundary fragmentation. It determines how the fractals differ from Euclidean objects. According to Mandelbrot, the relation between the data measuring scale δ and the length *L* can be expressed as [39]:

$$L(\delta) = k \cdot \delta^{1-D}$$

Where, *K* is a constant and *D* is known is the fractal dimension. Fractal Dimension D is very often a non-integer number [40].

3.1 Methods for computing Fractal dimension

Several methods have been devised to estimate fractal dimension. They are listed below [41]:

- Box counting method
- Higuchi method
- Katz's method
- Variance based
- Power spectrum density based (PSD)
- Empirical Mode Decomposition (EMD) based

- Wavelet Packet based

In the next section we will discuss the algorithm for finding fractal dimension by the listed (above) method. In the next section we will discuss ECG arrhythmia classification by computing fractal dimension using Power spectrum Density, EMD and wavelet packet method. We will also make a comparison among these results.

3.1.1 Box counting method

A box counting method is a very popular technique for computing fractal dimension. There are many applications for computing fractal dimension from a 2D image [42] [43] [44]. Now a day methods for computing fractal dimension of 3D objects has also been proposed [45] [46].

For computing fractal dimensions of objects self-similarity concepts is very important. The fractal dimension (FD) of a set of voxels is expressed by [47]:

$$FD = log\ (Nr)/\ log\ (1/r)$$

Here voxel is divided into N copies of a primitive shape, each one scaled up or down by a factor of r. There is a logarithmic relationship between N and r. When Nr is computed for various values of r, the fractal dimension can be estimated as the slope of the line. The line is estimated by the $log\ (Nr)$ plotted against $log(r)$ [47].

3.1.2 Higuchi method

This is another method for computing fractal dimension. This method has been proposed in 1988 [48]. The Higuchi algorithm computes the Fractal Dimension (FD) directly from the discrete time sequences [49].

3.1.3 Katz's method

This is another algorithm for calculating fractal dimension. By using Katz's method [52] the fractal dimension can be calculated of a sample as follows:

At first, the sum and average of the Euclidean distances between the successive points of the sample (L and a respectively) are calculated. Then we need to calculate the maximum distance

between the first point and any other point of the sample (d). The fractal dimension (FD) of the sample then calculated as follows [51]:

$$FD = \frac{\log(L/a)}{log(d/a)}$$

3.1.4 Variance Based Fractal Dimension Estimation

By this method we can also calculated fractal dimension. For this method if we consider a time series ECG signal x[n], then according to the Variance Based Fractal Estimation method, the variance of the amplitude increment is related to the time increment according to the following power law formula we can say that [54]:

$$\text{Var } \{(x [n_2]\text{-}x [n_1])\} \sim |n2 - n1|^{2H}$$

$$\text{Var } (\Delta x)\Delta n_m \sim \Delta n_m^{2H}$$

Where Δn=$|n2 - n1|$, (Δx) Δn=x[n_1]-x[n_2], $(\Delta n)_m$ is the discrete time increment at the m-th order scale, var(Δx) Δn_m is the variance of the amplitude increment over that time increment at the m-th order scale and here H is the Hurst component. H can be calculated from log-log plot of $(\Delta n)_m$ verses var(Δx) Δn_m. for calculating H the slop s of the line fitting the log-log plot is first estimated by a method named least square method. From there H is calculated as the half of the slop s. Then the fractal dimension (FD) is calculated as follows [54]:

$$FD=2-H$$

3.2 Fractal Dimension Estimation and ECG arrhythmia classification:

We have discussed some methods for computing fractal dimension in the above portion. Now we will discuss some methods of estimating fractal dimension and classifying ECG arrhythmias.

3.2.1 Power spectrum density based

Now we will discuss first how can we find fractal dimension using Power spectrum Density method and then we will discuss about ECG arrhythmia classification using the fractal dimension.

Power spectrum density is a well-known method for estimating fractal dimension from any biological signal like ECG signal [55]. The power spectrum of fractal process we can calculate the power law relationship of [53]

$$S(f_n) \sim pf_n^{-\beta}$$

We can write the equation as:

$$\text{Log}(S(f_n)) \sim \log(P) - \beta \log(f_n)$$

Here β is called the spectral index. β Is calculated as the slope of the line fitting the log−log plot of the power spectrum by a least square method. Fractal dimension FD is then calculated by the following equation [53]:

$$FD = \frac{5 - \beta}{2}$$

After calculating fractal dimension now we will see how they have classified ECG arrhythmias.

For ECG arrhythmia classification template matching method has been used. For matching the templates from the test dataset to those from the test dataset, two methods are used. The first method consists of calculating the Euclidean distance between both templates.

Euclidean distance (X_d) can be expressed as follows:

$$X_d = \sqrt{\sum_{i=1}^{W} (xi - yi)^2}$$

The second method consists of calculating the correlation coefficient. This gives the degree of closeness between the two templates [56]. The maximum correlation coefficient gives the arrhythmia class of the test ECG signal.

3.2.1.1 Performance and accuracy discussion:

In this method a local fractal feature based template matching technique has been used for ECG arrhythmia classification. The calculated fractal feature based ECG signal has been matched with the representative ECG signal for ECG classification. The matching is done using the Euclidean distance measuring method. Performance and accuracy has been given here as tabulated form [53]:

Table 3.1: Performance table for PSD

Different Arrhythmias	Detection Accuracy
Normal ECG	100%
RBBB	100%
Ventricular	95.4%
LBBB	98.6%
Misclassified	0%

Overall accuracy of this method is 99.498%

This performance is comparable to the best performances found in the open literature [57] and also some other works [58], [59].

3.2.2 Empirical Mode Decomposition (EMD) based fractal feature estimation for ECG arrhythmia classification

The EMD method is a necessary step to reduce any given data into a collection of intrinsic mode functions (IMFs). An IMF is defined as the function which satisfies the following requirements:

- In the whole data set, the number of extrema and the number of zero-crossings must either be equal or differ at most by one.
- At any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero.

3.2.2.1 Block diagram of this algorithm

Functional block diagram of the EMD based ECG arrhythmia classification has given below:

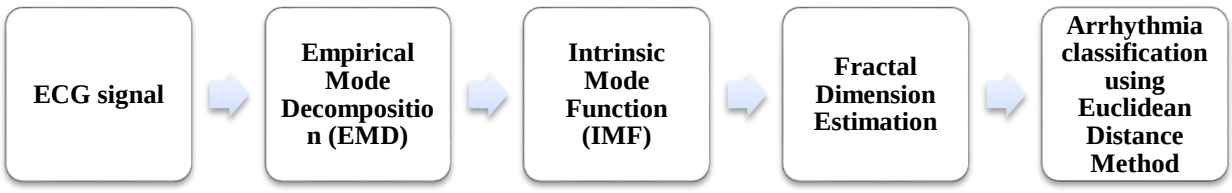

Figure 3.1: Block diagram of EMD

In this method MIT-BIH database has been used. From the ECG signal first IMF's has been computed using EMD method. By using this IMF's local fractal dimension has been estimated. From the calculated fractal dimension classification has been made by Euclidean Distance Method [54].

3.2.2.2 Performance and accuracy

The performance and the accuracy computed using this method has been tabulated below. Here performance accuracy for the normal beat and the five types of arrhythmias has been shown. At last the average accuracy has been given.

Table 3.2: Performance table for EMD

Arrhythmia types	Detection accuracy
Normal beat	98.26%
LBBB	97.73%
RBBB	99.09%
Ventricular	97.50%
Paced beat	98.03%
Atrial Premature	99.24%

Average performance accuracy of this method is 98.31%.

3.2.3 Wavelet packet based fractal feature estimation for ECG arrhythmia classification

In case of EMD for calculating local fractal dimension at least three IMF is required which is depended on the length of the segment of an ECG signal used for EMD operation. But it is seen that for some kind of the arrhythmias it is difficult to find three IMF's. In this case Wavelet Packet Decomposition (WPD) is very beneficial.

Wavelet Packet Decomposition is a popular method for fractal feature estimation as it satisfy the energy conversion law and using this method original signal can be reconstructed.

3.2.3.1 Block Diagram of this algorithm

Main function of this algorithm has been given below as a summarized form:

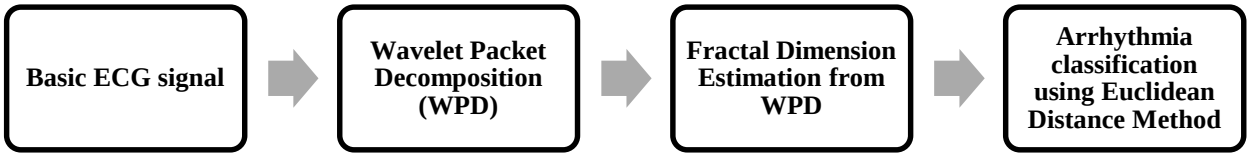

Figure 3.2: Block diagram of WPD

3.2.3.2 Performance and Accuracy

EMD has some limitation but using WPD those limits can be overcome. That's why the overall performance accuracy is also higher for WPD. Now the performance accuracy for the normal beat and the five classes of arrhythmias have been tabulated below for this method [54].

Table 3.3: Performance table for WPD

Arrhythmia types	Detection accuracy
Normal beat	99.44%
LBBB	98.72%
RBBB	99.56%
Ventricular	98.44%
Paced beat	98.83%
Atrial Premature	99.44%

Average accuracy for this algorithm is 99.07%

3.2.4 Other methods for ECG arrhythmia classification

Various works is going on in this field from a long since. Various methods have been proposed also for ECG arrhythmia classification. In this section some other method's algorithms and their performance accuracy has been discussed.

3.2.4.1 ECG arrhythmia classification based on Energy and Entropy of Detail Wavelet Packet Coefficients

Energy and the entropy based wavelet packet coefficient is simple but the most efficient method for multi-class ECG arrhythmia classification. After considering a detail analysis on entropy, energy and cross-correlation it is notified that only the 4^{th} level detail WPD coefficients are enough for ECG arrhythmia classification.

3.2.4.1.1 Block diagram of the algorithm

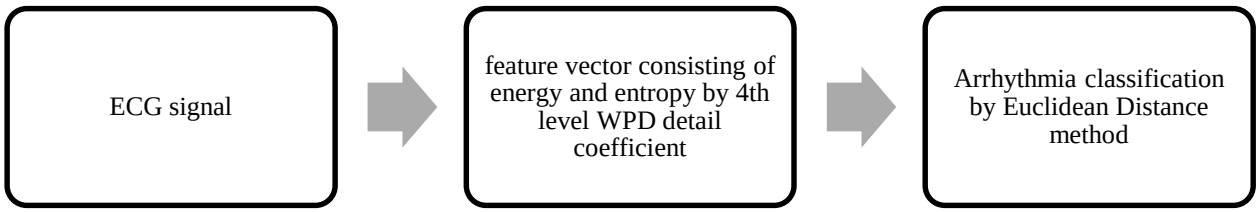

Figure 3.3: Block diagram of WPD detail coefficient

3.2.4.1.2 Performance and Accuracy

In this method MIT-BIH database has been used. The performance accuracy of this method for classifying ECG arrhythmias are given in tabulated form [54]:

Table 3.4: Performance table for WPD detail coefficient

Arrhythmia types	Detection accuracy
Normal beat	99.87%
LBBB	98.78%
RBBB	99.83%
Ventricular	99.69%
Paced beat	98.87%
Atrial Premature	99.84%

Average accuracy for this algorithm is 99.81%

3.2.4.2 EMD and Support Vector Machine (SVM) based ECG arrhythmia detection

Empirical mode decomposition (EMD) [59] is a new method which is pioneered by Huang et al. in 1998 for non-linear and non-stationary time series analysis. EMD method decomposes an ECG signals into (IMF's) and also extracts signal features adaptively [60].

Support Vector Machines (SVM) is a new learning method for ECG arrhythmia classification. It has shown many unique advantages in processing small samples case, nonlinear pattern recognition problems.

3.2.4.2.1 Block diagram of the algorithm

A summarized functional block has been given below. From this we will have some idea hoe this algorithm works.

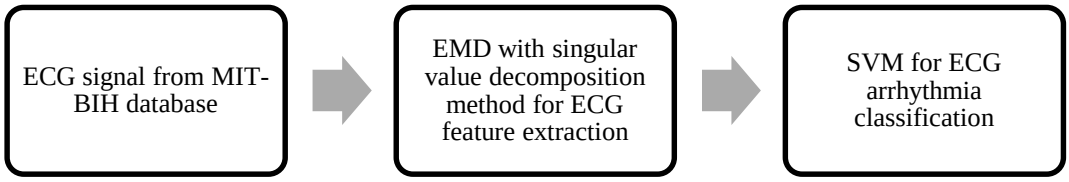

Figure 3.4: Block diagram of EMD+SVM

3.2.4.2.2 Performance and Accuracy

Performance of this method has been compared with BP neural network method. The comparison table has given below:

Table 3.5: Performance table for EMD+SVM

Classifier Type	**Accuracy (%)**		
	Normal	**PVC**	**BBB**
BP Neural Network	99.81	99.44	98.33
SVM	99.91	99.63	99.17

Average accuracy of this method is 99.72%.

3.3 Comparison of the results

In this section we will discuss about the performance result of different arrhythmia classification techniques. We will observe that how the results are varying and which one is the best method for ECG arrhythmia classification.

Table 3.6: Performance comparison table for different types of ECG arrhythmia classification method

ECG Arrhythmia Classifier	**Average Accuracy (%)**
Power Spectrum Density based	99.498
EMD	98.31
WPD	99.07
Energy and Entropy Based Detail Wavelet Packet Coefficients	99.81
EMD + SVM	99.72

3.4 Conclusion

In this paper we discussed about fractal feature based ECG arrhythmia classifier and some other classifier. We also make a comparison between these classifiers according to their performance accuracy. From the comparison table we see that the EMD with SVM technique makes a good performance. We hope that if we combine other methods that may also give some good accuracy.

CHAPTER 4

REVIEW ON CLASSIFICATION OF ECG ARRHYTHMIAS BY GAUSSIAN MIXTURE MODEL

In this chapter, a description on a classifier, based on Gaussian mixture model known as GMM to classify the ECG arrhythmias, is discussed. The procedure of classification of arrhythmias is summarized after that. Also it is focused here that how the Gaussian mixture model is used in other fields. The applications of GMM with most efficiency are reviewed. From the comparison of GMM with the other standard classifier, it can be seen that the efficiency of Gaussian mixture model is the highest value. So, there is a possibility of having greater or more efficiency in this field in the classification of arrhythmias, even in other application also. As GMM is far more efficient than the other classifiers, it can be made more effective in biomedical science.

In mathematics, a Gaussian function is a function which is named after Carl Friedrich Gauss. The graph of a Gaussian is a characteristic symmetric "bell curve" shape that quickly falls off towards zero. The parameter ‘a’ is the height of the curve's peak, *b* is the position of the center of the peak, and *c* (the standard deviation) controls the width of the "bell". Gaussian functions are widely used in statistics where they describe the normal distributions, in signal processing where they serve to define Gaussian filters, in image processing where two-dimensional Gaussians are used for Gaussian blurs, and in mathematics where they are used to solve heat equations and diffusion equations and to define the Weierstrass transform.

A model is basically a probability distribution that gets fit to some data that you have. Fitting a model to data means estimating parameters of the model so that it describes the data accurately. On the other hand, a mixture model is a probabilistic model for representing the presence of subpopulations within an overall population, without requiring that an observed data-set should identify the sub-population to which an individual observation belongs. Formally a mixture model corresponds to the mixture distribution that represents the probability distribution of observations in the overall population. Mixture models are used to make statistical inferences about the

properties of the sub-populations given only observations on the pooled population, without sub-population-identity information. Some ways of implementing mixture models involve steps that attribute postulated sub-population-identities to individual observations or weights towards such sub-populations, in which case these can be regarded as types of unsupervised learning or clustering procedures. However not all inference procedures involve such steps. Now, if it about Gaussian mixture model, then it will be that a Gaussian mixture model is nothing more than a mixture model of Gaussian distributions i.e. normal distributions. Gaussian Mixture Models (GMMs) are among the most statistically mature methods for clustering though they are also used intensively for density estimation.

4.1 Methodology

The methodology consists of a stage named pre-processing or this classification. But in real, pre-processing is a stage of several steps like re-sampling, QRS detection by middle point extraction using extended Pan Tompkin's method, segmentation on the basis of R point, linear pre- diction model for estimation, Principal component analysis (PCA) for RES computation and dimensionality reduction. Finally, the principal components (PCs) which are extracted are to be used for the classification using GMM classifier.

4.1.1 Pre-processing

This stage is to be carried out before the real classification of ECG arrhythmia. This stage consists of six steps such as, re-sampling, QRS detection, segmentation based on R point, Liner Prediction (LP) model estimation, residual signal estimation and Principal component analysis (PCA).

4.1.2 Re-sampling

At first ECG signal is sampled at 360 Hz for MIT - BIH Arrhythmia, whereas the same is sampled at 250 Hz for European ST-T database. In order to maintain the uniformity, the European ST-T database is re-sampled at 360 Hz by using FFT based interpolation method.

4.1.3 QRS detection

Pan Tompkins algorithm is the most well-known method for R peak (QRS complex) detection .It consists of taking the first difference of ECG signal, rectifying it and smoothing by passing through

a moving average filter of order 2. This procedure is repeated again and the two smoothed signals are summed up. An appropriate threshold is applied to this signal and rectangular pulses at the QRS complexes are obtained. The middle-point of each rectangular pulse is the R point in the ECG. The simplified Pan Tompkins algorithm follows following steps.

Step 1:

Compute the absolute value of first difference of the ECG signal, x (n).

Step 2:

Smooth this signal by passing through a moving average system of order 2 as follows:

$y(n) = 14\{x1(n) + 2x1(n-1) + x1(n-2)\}$

Step 3:

Take the derivative of the smoothened signal, y (n) from step 3 and compute its absolute value.

Step 4:

Smooth the signal obtained from step 3 through a moving average system.

Step 5:

Sum the two signals obtained from steps 2 and 4.

Step 6:

Threshold the signal with an adaptive threshold to get rectangular pulses.

4.1.4 Segmentation using R point

After detecting the R point in the ECG, 99 points to its left side and 100 points to its right side of R point along with R point itself are chosen as one segment, which is used for subsequent analysis and diagnostic classification.

4.1.5 Linear prediction (LP) model

In LP analysis of the ECG signal, x(n) is modeled as a linear combination of its past input signals x(n−k), k=1,2,...,p, where p de-note the order of prediction and a(k) is the kth LP coefficient. The error in prediction is given by

$e(n) = x(n) - \hat{x}(n)$

Though the ECG signal $x(n)$ can be predicted from the third order linear predictor using the LP coefficients, $a(k)$, the signal $e(n)$ represents the part of the ECG which cannot be predicted linearly, hence it is called RES. Here Levinson Durbin recursion is used to compute the LP coefficients. Third order prediction is used for 730 data segments. The output of the linear predictor e(n) is random in nature and obeys Gaussian distribution. The signal samples e(n) do not have any correlations among themselves. But the dimensionality of this data is large and this dimensionality is reduced by PCA technique.

4.1.6 Principal component analysis PCA

PCA is a linear dimensionality reduction technique that seeks a projection of the data that best represents the data in a least square sense. It includes computation of Eigen values and Eigen vectors of covariance matrix of the error signals obtained from ECG through LP model. Eigen vectors are orthogonal to each other and form as basis vectors for the system. The Eigen vectors are sorted in the descending order of the irrespective Eigen values. The Eigen vector corresponding to the highest Eigen value corresponds to the first PC, where the data variation is maximum. Let us consider the data matrix where *e(i)* and *e*(*j*) represent the *i*th and *j*th error signals ($1 \leq i, j \leq N$) derived from 730 ECG signals. Now the PCA algorithm is as:

Step 1:

Calculate covariance matrix of order $N \times N$ for the error signals. Find matrix V of Eigen vectors and diagonal elements of matrix D as Eigen values of covariance matrix.

Step 2:

Arrange the Eigen values from D in descending order and corresponding Eigen vector is chosen from V to construct PCs.

Step 3:

Compute the energy content based on eigenvalues (amount of variation). In this study, the above algorithm is applied to reduce the dimension of 200 error features based on 730 ECG samples obtained from arrhythmia and cardiac ischemia classes. The transformed features are called here PCs based on their energy content.

4.2 Classification using GMM

Euclidean distance (Orin general, Minkowski distance) is a commonly used distance metric, which works well when the dataset has common range of variations. The drawback of this Euclidean metric. Is that he largest scaled feature gets dominated. Solution to this problem could be normalization of the continuous features to a common range of variance. The RES *Rd* derived from electrocardiogram (ECG or EKG) signal is modeled statistically as Gaussian wide sense stationary process characterized by the model. It is easy to see that the number of parameter sin the covariance matrix grows quadratically with the dimensionality of data, which makes the model training or accurate parameter estimation for high- dimensional data a difficult task. Since standard PCA is a linear orthonormal transformation, the problem of data classification in the data space becomes data classification in the feature space. The bipartite graph provides a good understanding and visualization of this model based clustering algorithms, and emphasis on cluster centroids (in a model space), which is conceptually separated from the feature space. It assumes a set of N data objects, represented by $X = \{x1, x2, ..., xN\}$ in the data space which is mapped in to the feature space using PCA, i.e., $E = \{e1, e2, ..., ed\}$ and probabilistic generative in a model space .In fact ,there is many-to-one mapping from feature space to model space. Each cluster is represented by a model in the model space, which usually contains the models from a specific family of models. The model λk can be viewed as the generalized centroid of cluster k. Connection between object en and model λk indicates that the object en is being associated with cluster k, with connection weight (closeness) given by the log-likelihood $\log p(en|\lambda k)$. This idea of representing clusters by models is radically more general than the standard *K-means* algorithm, where both feature objects and cluster centroids are in the same feature space. The models also provide probabilistic interpretation of clusters, which is a desirable feature in this application. Next we observe how

these probabilistic models are derived using Expectation Maximization (EM) algorithm based on ML estimation.

4.3 GMM algorithm for classification

Here we have the two class problem of classification of two abnormalities in ECG using probabilistic model called as GMM. The database consisting of two abnormalities is assumed to be generated by the two Gaussian processes, of course each Gaussian could have generated by many biochemical processes inside the heart having arbitrary stochastic distribution. Therefore, the entropy of the resultant distribution tries to become highest. Hence, the resultant distribution is Gaussian which is having the maximum entropy among any other possible stochastic distributions. The conditional probability density is estimated. The objective function is formed by summing the class conditional density overall the classes for a feature in the feature vector; and again taking the product for all the features, assuming the features are linearly independent. Such likelihood based objective function for optimization is maximized by EM algorithm.

4.4 Comparison with other methods

Here is the comparison of GMM method with other method in case of accuracy.

Table 4.1: Comparison of GMM with other methods

Methods	**Classification accuracy (%)**
K-means algorithm	71.43
Fuzzy C-means	82.86
ANN	79
Hierarchical agglomerative	84
GMM based classifier	**94.29**

CHAPTER 5

FEATURE EXTRACTION METHOD FOR ECG ARRHYTHMIA CLASSIFICATION (OUR PROPOSED METHOD)

There exist a number of methods for ECG arrhythmia classification, such as Empirical Mode Decomposition (EMD) based, Wavelet Packet based, support Vector Machine based etc. Many other established methods are also available. Our proposed method is Feature Extraction Method based ECG arrhythmia classification.

In this chapter, we will discuss briefly about our method. We develop such an algorithm by which we extract different feature parts from an ECG signal like R peak, P point, S point, T point etc. We use these features for classifying normal and abnormal ECG signal and from the abnormal signal we correctly detect Right Bundle Branch Block (RBBB) and Left Bundle Branch Block (LBBB). In the next section we will give a brief overview of our whole process.

5.1 Feature Extraction

Feature extraction means extracting the main features from any object which will help us to analysis of that object. In our thesis we tried to extract different features from a basic ECG signal. For this we have built an algorithm which will give us different feature points from an ECG signal. Our proposed method has been discussed below.

5.1.1 Methodology

The electrocardiogram (ECG) is the non-invasive technique most used in heart disease diagnoses. It could be described as a record of the electrical phenomena originated from the cardiac activity. The ECG is frequently used to detect cardiac rhythm abnormalities, arrhythmia, and the utilization of pattern recognition techniques can help the physician to improve this detection and consequently make a more accurate diagnoses. An ECG signal has some very specific points like P, Q, R, S and

T. These are key points of an ECG signal. We can collect much information from these points for arrhythmia diagnosis. An ECG signal with its key features has been shown below:

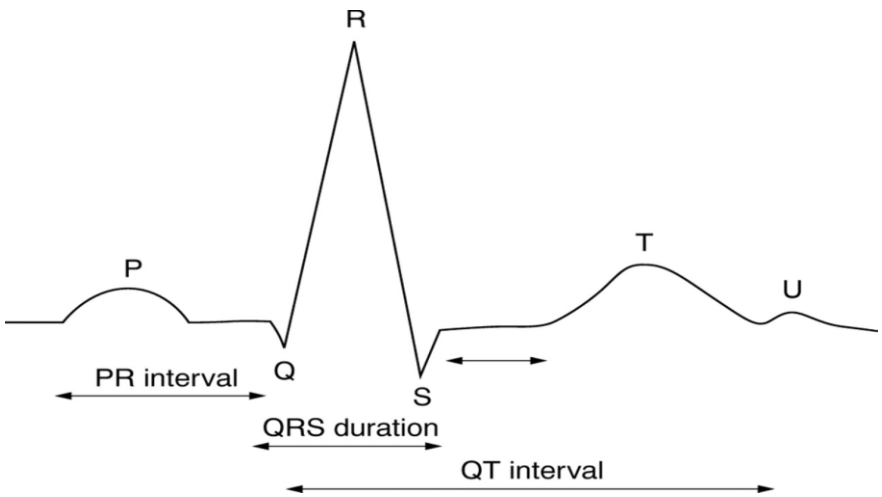

Figure 5.1: P, Q, R, S, T points in an ECG signal

5.1.2 Block Diagram

A block diagram of our methodology is given here.

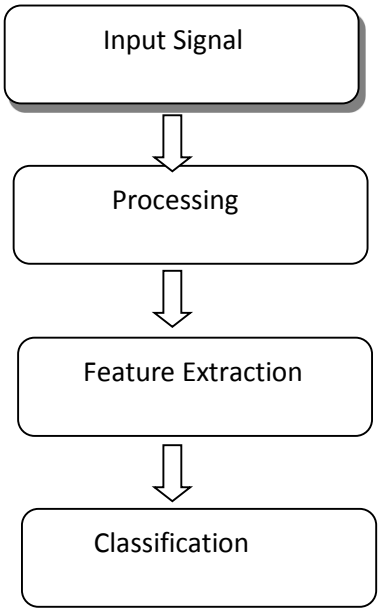

Figure 5.2: Block Diagram of our proposed method

5.1.3 Input Signal

At the very beginning of our work we find out the input signals for our algorithm. In our thesis we used MIT-BIH database for our input signals. It's a well-known database which is available in the internet. The MITBIH Arrhythmia database consists of ECG data band pass filtered at 0.1-100Hz and sampled at 360Hz. This data was used to evaluate the performance of the classifiers in this study. A data set was selected from the beginning of records 100, 101, 105, 109, 111, 115, 118, 121, 122, 124, 205, 207, 209, 212, 214, 221, 231, 234. We used these records for our input signal.

5.1.4 Processing of the signals

In order to extract features from the ECG signal certain processing tasks are necessary. The first of these tasks involves measurement of the R-R intervals, the P-R intervals and the QRS width. The detection of the P, Q, R, and S wave in the electrocardiogram is a complex problem since they have a time varying morphology and are subject to physiological variations due to individual patient variation and corruption from noise. Considerable research has been carried out in the development of QRS and P wave detectors; a review of these detectors can be found. However in this study the QRS width, the R-R and the P-R intervals were marked manually with the aid of the annotation files. The second of these processing tasks involves extracting templates from the ECG waveform. The ECG waveform is band-pass filtered (0.5-40 Hz) to remove baseline wander and power line interference. Then the waveforms are down-sampled at 80Hz. A window is taken around each beat, 400ms before the R wave and 600ms after. For each beat a beat template is formed using the amplitudes of the resulting 80 points in the windowed segment. For each beat a QRS template is formed using the amplitudes of 9 points extracted from a segment windowed 50ms before and after the R wave.

5.1.5 Feature Extraction

In order to develop an efficient beat classifier one needs to establish features to distinguish between the various beats. The first set of features is extracted from the ECG after the beat (QRS) detection process. This set consists of features based on the R-R interval (RR), the amplitude of points in the beat template (B amp) and the amplitudes of points in the QRS template (QRS amp). After the QRS onset/offset of each beat detected is determined, features based on the QRS width QRS duration are extracted. The remaining set of features is extracted after the P wave onset of each

detected P wave is determined. This set consists of features based on the P-R interval (PR). We use here three basic features. These are QRS duration, PR interval and RR interval.

5.1.6 Hilbert Transform based R peak Detection

In order to find out the QRS duration, PR interval and RR interval our main challenge was to find out the R peak of the basic ECG signal. We use Hilbert transform method for R peak detection. The detection of R-wave peak is achieved by comparing the envelope of an ECG signal against a fixed/adaptive amplitude- dependent and RR interval-dependent thresholds. Most R-peak detection methods use a similar approach to determine the threshold. The detection thresholds are adapted periodically based upon amplitudes, durations and RR-intervals of past detected R-peaks. In such a case, performance relies highly on the accurate estimation of initial parameters in the learning phase. Almost all methods use additional decision rules for the reduction of false-positive detections and introduce secondary thresholds to detect missed R-peaks. These heuristic rules may improve detections for regular rhythms but some rules may be in conflict with others. Furthermore, search- back mechanism with secondary threshold cannot be halted in case of irregular rhythms with varying QRS complexes and noise.

5.1.7 Other Points detection

After finding out the R peak our main goal was to find out the other features like P, Q, S and T wave. Then we make back and front calculation by our algorithm for finding those points. After getting those points we calculated the QRS duration, PR interval and RR interval by our algorithm.

5.2 Classification

Our classification was done by matching the features (QRS duration, PR interval and RR interval) of the training records and the testing records. By our algorithm we can identify normal and abnormal signal and among the abnormal signals we can detect RBBB and LBBB correctly. For this we at first trained some record for Normal, RBBB and LBBB by our algorithm. From there we collect the QRS duration, PR interval and RR interval. We took those data as standard. Then we tested some of the records from the MIT-BIH database. Based on our standard data we classify our tested sample as Normal ECG, RBBB and LBBB.

CHAPTER 6

SMULATION RESULT AND PERFORMANCE

Performance evaluation of the proposed method for classifying ECG arrhythmia plus normal ECG is a very important task. For this evaluation, confusion matrix criteria was taken into account. The popular MIT-BIH database is used for extracting ECG beats of different arrhythmia for simulation.

6.1 Simulation Database

In our proposed method, we have used MIT-BIH arrhythmia database. The ECG recordings are sampled at 360 samples per second per channel with eleven beat resolution over a 10 mV range. A list of table has been given below showing the different records that we have used for our simulation.

Table 6.1: Training and testing records

Type	MIT-BIH data file	No. of Training Sample	No. of Testing Sample
Normal	100, 101,115, 121, 122, 205, 221, 234	65	644
RBBB	118, 124, 212, 231	87	322
LBBB	109, 111, 207, 214	85	259

6.2 Simulation Results

In this section we will discuss briefly about our simulated results. We have done our simulation for Normal signal, RBBB and LBBB.

6.2.1 Result for Normal signal

In case of normal signal we first trained record #115(0-1 min) sample. From there we collected the range of QRS duration, PR interval and RR interval and took them as standard data for normal ECG signal. Then we tested others sets of records and compared them with the standard value.

6.2.1.1 Simulation result (Normal Signal)

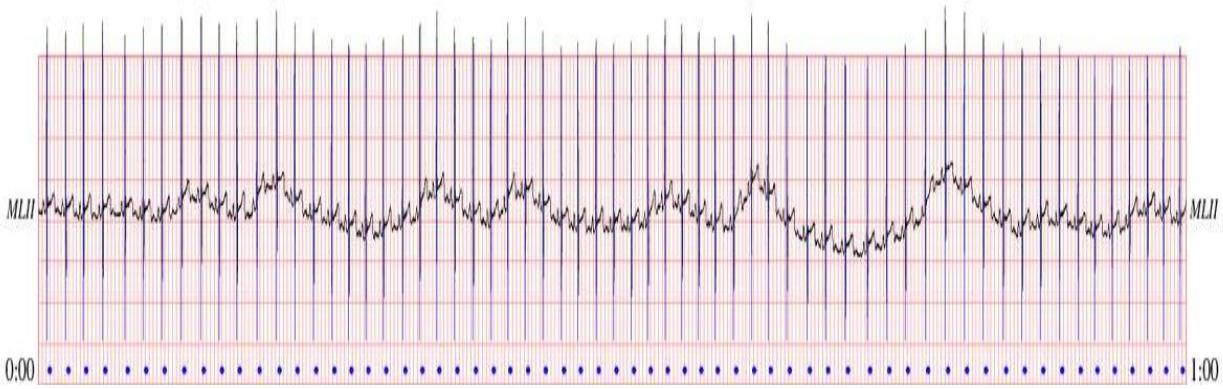

Figure 6.1: Training Signal #115 (0-1 min)

Range of QRS duration= (0.11 to 0.20) s

Range of PR interval= (0.03 to 0.24) s

Range of R-R interval= (.01 to 1.04) s

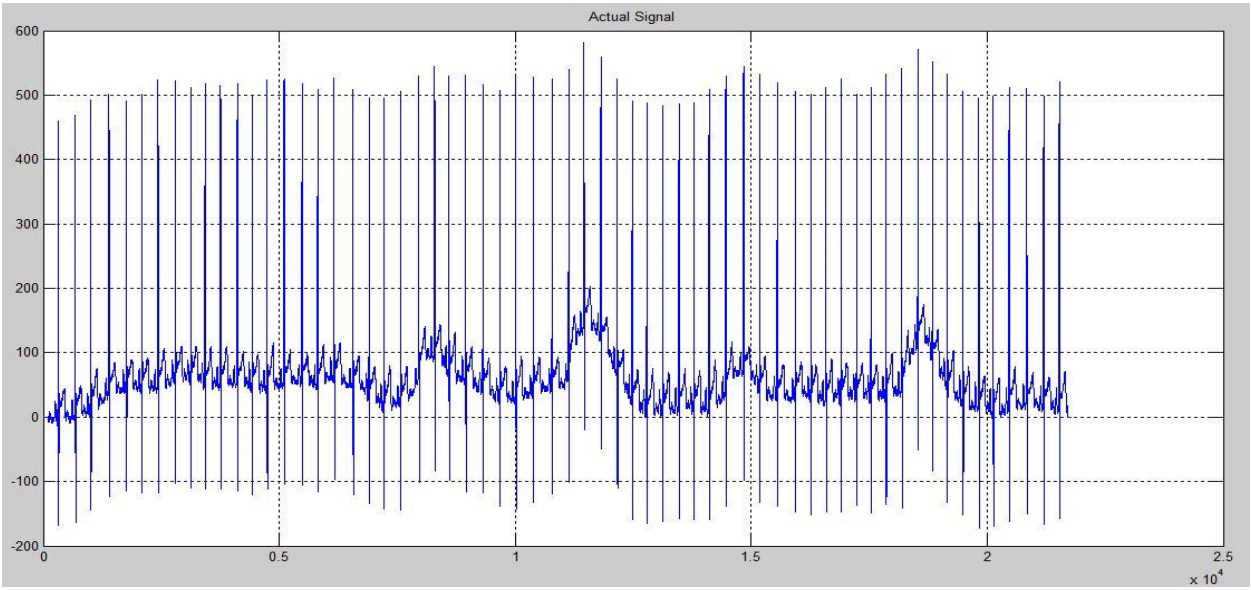

Figure 6.2: Testing Signal #115 (6-7 min)

The table below contains different features of normal signal.

Features	1	2	3	4	5
QRS Duration (sec)	0.17	0.18	0.21	0.20	0.21
PR interval (sec)	0.19	0.19	0.18	0.19	0.20

Table 6.2: Feature extraction for normal signal

6.2.2. Result for RBBB

In case of RBBB we did the same procedure as the normal signal discussed earlier.

6.2.2.1 Simulation result (RBBB)

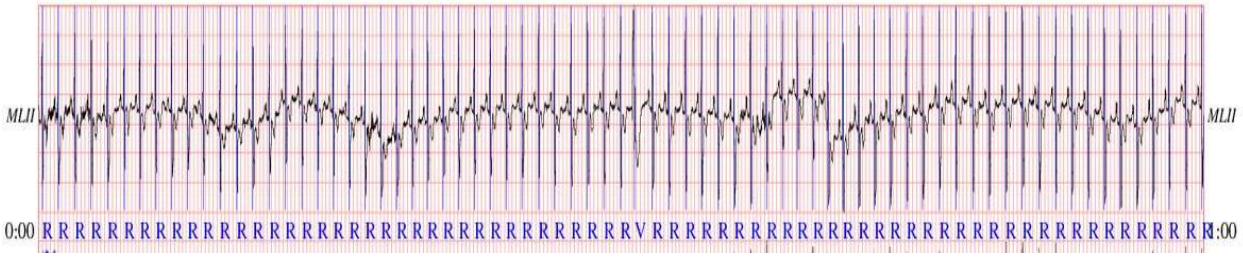

Figure 6.3: Training Signal #118 (6-7 min)

Range of QRS duration= (0.13 to 0.30) s

Range of PR interval= (0.03 to 0.27) s

Range of R-R interval= (.01 to 0.89) s

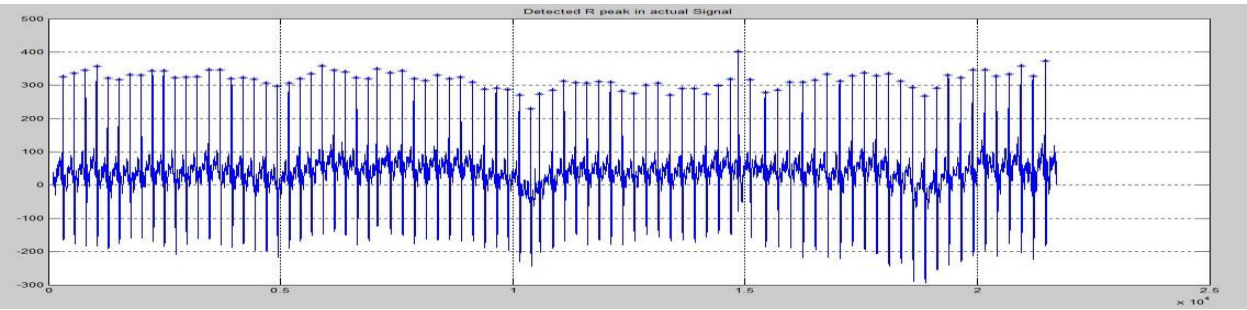

Figure 6.4: Testing Signal #118 (7-8 min)

The table below contains different features of RBBB signal.

Table 6.3: Feature extraction for RBBB signal

Features	1	2	3	4	5
QRS Duration (sec)	0.23	0.23	0.24	0.25	0.20
PR interval (sec)	0.20	0.22	0.19	0.20	0.21
R-R interval (sec)	0.67	0.66	0.69	0.70	0.72

6.2.3. Result for LBBB

In case of LBBB we did the same procedure as the normal signal discussed earlier.

6.2.3.1 Simulation result (LBBB)

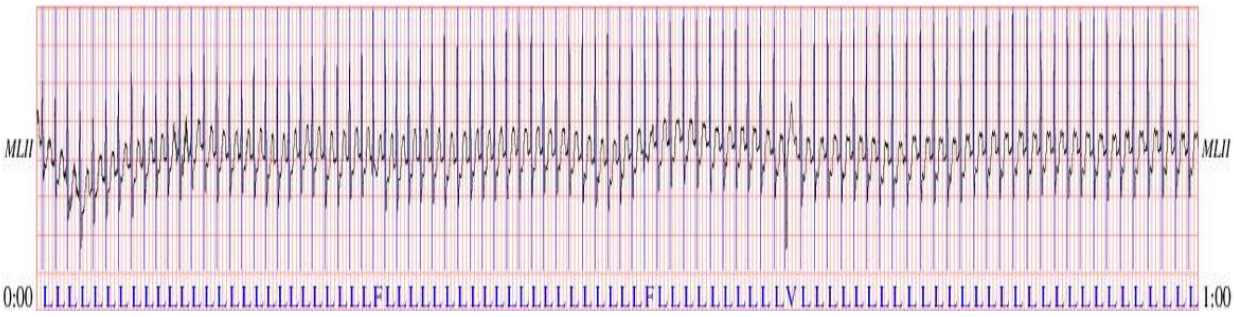

Figure 6.5: Training Signal #109 (0-1 min)

Range of QRS duration= (0.17 to 0.33) s

Range of PR interval= (0.03 to 0.26) s

Range of R-R interval= (.01 to 1.91) s

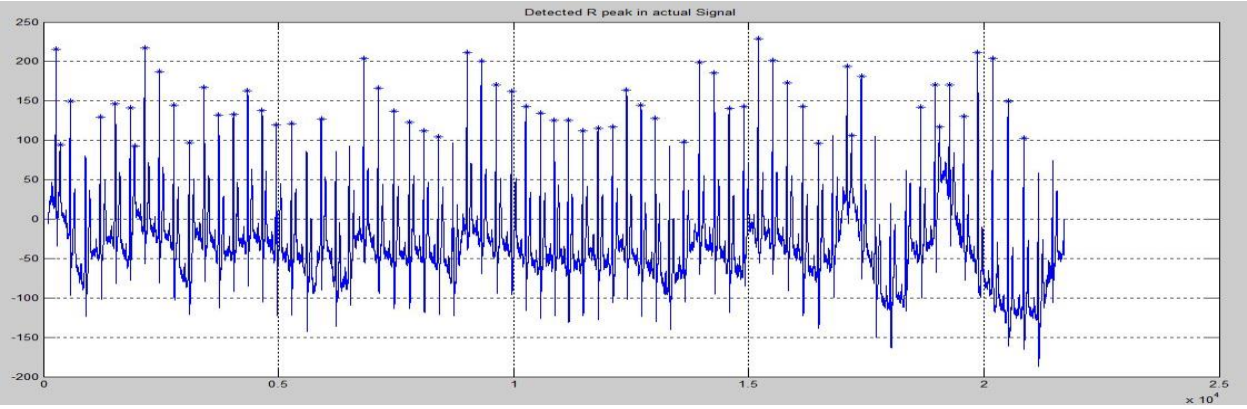

Figure 6.6: Testing Signal #111 (6-7 min)

The table below contains different features of LBBB signal.

Table 6.4: Feature extraction for LBBB signal

Features	1	2	3	4	5
QRS Duration (sec)	0.23	0.25	0.24	0.21	0.27
PR interval (sec)	0.18	0.16	0.22	0.20	0.23
R-R interval (sec)	0.92	0.85	0.88	0.89	0.91

6.3 Performance Evaluation

In this section we will discuss our performance parameters like efficiency, specificity, sensitivity etc.

6.3.1. Efficiency

Efficiency of our algorithm for detecting normal signal, RBBB and LBBB is given below:

Table 6.5: Efficiency for Normal Signal

#File	QRS detection	QRS efficiency	PR interval detection	PR interval efficiency	R-R interval detection	R-R interval efficiency
100	**59 of 75**	78.67%	**75 of 75**	100%	**24 of 24**	100 %
101	**50 of 61**	81.96%	**61 of 61**	100%	**17 of 24**	70.83%
121	**30 of 65**	46.15%	**63 of 65**	96.92%	**24 of 24**	100%
122	**91 of 99**	91.91%	**99 of 99**	100%	**24 of 24**	100%
205	**56 of 92**	60.87%	**92 of 92**	100%	**24 of 24**	100%
221	**28 of 95**	29.47%	**95 of 95**	100%	**22 of 24**	91.67%
234	**64 of 92**	69.56%	**85 of 92**	92.39%	**24 of 24**	100%

Avg. efficiency of QRS=65.51%

Avg. efficiency of PR interval=98.47%

Avg. efficiency of R-R interval= 94.64%

Table 6.6: Efficiency for RBBB Signal

#File	QRS detection	QRS efficiency	PR interval detection	PR interval efficiency	R-R interval detection	R-R interval efficiency
118	**85 of 86**	98.84%	**86 of 86**	100%	**24 of 24**	100%
124	**60 of 64**	93.75%	**64 of 64**	100%	**9 of 24**	37.5%

212	**106 of 114**	92.98%	**114 of 114**	100%	**24 of 24**	100%
231	**30 of 38**	78.95%	**38 of 38**	100%	**8 of 24**	33.33%

Avg. efficiency of QRS= 91.13%

Avg. efficiency of PR interval= 100%

Avg. efficiency of R-R interval= 67.71%

Figure 6.7: Efficiency for LBBB Signal

#File	**QRS detection**	**QRS efficiency**	**PR interval detection**	**PR interval efficiency**	**R-R interval detection**	**R-R interval efficiency**
109	**100 of 100**	100%	**100 of 100**	100%	**24 of 24**	100%
111	**69 of 70**	98.57%	**70 of 70**	100%	**23 of 24**	95.83%
207	**1 of 2**	50%	**2 of 2**	100%	**1 of 1**	100%
214	**82 of 87**	94.25%	**83 of 87**	95.4%	**24 of 24**	100%

Avg. efficiency of QRS= 85.71%

Avg. efficiency of PR interval= 98.85%

Avg. efficiency of R-R interval= 98.96%

Table 6.8: Overall Efficiency

Type of Signal	Individual Efficiency	Overall efficiency of this method
Normal	86.21%	
RBBB	86.28%	**89%**

LBBB	94.51%	

We will discuss about sensitivity and specificity which we have calculated using confusion matrix. A confusion matrix contains information about actual and predicted classifications done by a classification system. Performance of such systems is commonly evaluated using the data in the matrix. The following table shows the confusion matrix for a two class classifier.

The entries in the confusion matrix have the following meaning in the context of our study:

- a is the number of **correct** predictions that an instance is **negative**,
- *b* is the number of **incorrect** predictions that an instance is **positive**,
- *c* is the number of **incorrect** of predictions that an instance **negative**, and
- *d* is the number of **correct** predictions that an instance is **positive**.

Table 6.9: Confusion Matrix

		Predicted	
		Negative	Positive
Actual	Negative	**a**	**b**
	Positive	**c**	**d**

6.3.2 Sensitivity and Specificity Calculation

For calculating sensitivity and specificity some points need to be considered. They are listed below:

- True positive = correctly identified
- False positive = incorrectly identified
- True negative = correctly rejected
- False negative = incorrectly rejected
- $\text{Sensitivity} = \frac{\text{No.of true positive}}{\text{No.of true positive+No.of false negative}}$
- $\text{Specificity} = \frac{\text{No.of true negative}}{\text{No.of true negative+No.of false positive}}$

6.3.3 Our calculated data

Table 6.10: Confusion matrix table

Type	Normal	RBBB	LBBB
Normal	627	7	10
RBBB	3	319	0
LBBB	4	18	237

Table 6.11: Sensitivity calculation

Type	Sensitivity	Overall sensitivity
Normal	97.36%	96.08%
RBBB	99.37%	
LBBB	91.51%	

Table 6.12: Specificity calculation

Type	Specificity	Overall Specificity
Normal	96.95%	95.66%
RBBB	97.36%	
LBBB	92.67%	

So, our overall efficiency is 89%

Sensitivity is 96.08% and

Specificity is 95.66%

CHAPTER 7
CONCLUSION

7.1 Concluding Remarks

Feature extraction method that we have used is a very basic method. We have analyzed samples from MIT-BIH database and trained it. Then we have tested the other samples to get an efficient result. Our proposed method is eligible to classify normal, Right bundle brunch block and Left bundle branch block ECG signal. We have used the MIT-BIH database for our thesis work. We have faced a lot of problems during our thesis work. We have reviewed many methods and studied them. After that we have trialed many of them to classify ECG signal efficiently, but some of them were not efficient and some of them were quite hard for us to go further. So, at last we have chosen this method and we made our work successful. But, we have also faced some problems here. At first our efficiency was quite low. The reason was at first we tested very few data from the database. After that, we increased the number of samples and data and have found more efficient result. Randomly we have selected ECG beats. We have taken a sample from the MIT- BIH database and trained it. That was our reference for test. We have trained individually for each types. We matched our result with the MIT-BIH record of each beat. From there we have some data matched and very few mismatched. From the number of matched and mismatched data, we have calculated the sensitivity, specificity and accuracy. Our method was quite efficient.

7.2 Contribution of the thesis

The main contribution of our thesis is classifying the ECG signal efficiently in a simple way. We have classified normal and abnormal signals. Among abnormal signals, we have classified RBBB and LBBB. We have compared the results of our method with other methods. We also have calculated sensitivity, specificity and accuracy which were good.

7.3 Future Work

In this thesis, we have shown a good result but there are many scopes for developing the method and make the result more efficient. Here are some possible scopes.

- We have got a result by taking some samples and ECG beats from MIT-BIH database and our result was quite efficient. But, if we use more data, the result will be more efficient.
- Again, there are a lot of other methods which can be combined with our proposed method. By this combination, we hope our result will be more efficient.
- We have classified only normal, RBBB and LBBB ECG signals. In the future, we will try to classify other types of abnormal ECG signals.
- We have only used MIT-BIH database because it is considered as standard database. We can use other databases in the future and we will try to create our own database.

References:

1. R. Acharya, J. Suri, and J. Spaan, "Advances in cardiac signal processing", Springer Verlag, 2007.
2. C. Evans et al., "Principles of human physiology." Principles of human physiology, 9th Edition, 1945.
3. A. Moss, "Noninvasive electro-cardiology: Clinical aspects of holter monitoring", WB Saunders CO, 1996.
4. M. Khan, "Rapid ECG Interpretation", Humana Press, 2007.
5. F. Morris, W. Brady, A. Camm, and I. Ebrary, "ABC of clinical electrocardiography", BMJ Books, 2003.
6. B. Anuradha and V. C. Veera Reddy, "Cardiac arrhythmia classification using fuzzy classifiers", Journal of Theoretical and Applied Information Technology, 2008, pp. 353-359.
7. Muthuchudar, Lt. Dr. S. Santosh Baboo, "A Study of the Processes Involved in ECG Signal Analysis", International Journal of Scientific and Research Publications, Volume 3, Issue 3, March 2013, pp. 1-5.
8. Channappa Bhyri, Satish T. Hamde, Laxman M. Waghmare, "ECG Acquisition and Analysis System for Diagnosis of Heart Diseases", Sensors & Transducers Journal, Vol. 133, Issue 10, October 2011, pp. 18-29.
9. Introductory Guide to Identifying ECG Irregularities, DailyCare BioMedical Inc.
10. Miad Faezipour, Adnan Saeed, Suma Chandrika Bulusu, Mehrdad Nourani, Hlaing Minn & Lakshman Tamil, "A Patient-Adaptive Profiling Scheme for ECG Beat Classification," IEEE Transactions On Information Technology In Biomedicine, Vol. 14, No. 5, September 2010, pp. 1153-1165.
11. Ludmila I. Kuncheva (2008), Scholarpedia, 3(1): 2925.
12. Ryan J. Urbanowicz and Jason H. Moore, "Learning Classifier Systems: A Complete Introduction, Review, and Roadmap", Journal of Artificial Evolution and Applications, Volume 2009, Article ID 736398, 25 pages.

13. Saniya Siraj Godil, Muhammad Shahzad Shamim, Syed Ather Enam, Uvais Qidwai, “Fuzzy logic: A ‘simple’ solution for complexities in neurosciences?” Surgical Neurology International 2011, Vol-2, Issue-1, page 24.
14. Raj Kumar Bansal, Ashok Kumar Goel, Manoj Kumar Sharma, “MATLAB and Its Application in Engineering”, Pearson Publication, Fifth Impression, 2012.
15. Wen Wei and Jerry M. Mendel, “A Fuzzy Logic Method for Modulation Classification in Nonideal Environments”, IEEE Transactions on Fuzzy Systems, Vol. 7, No. 3, June 1999, pp. 333-344.
16. Tomoharu Nakashima, Gerald Schaefer, Yasuyuki Yokota, Hisao Ishibuchi, “A weighted fuzzy classifier and its application to image processing tasks”, Fuzzy Sets and Systems 158, 2007, pp. 284 – 294.
17. Reza Boostani, Mojtaba Rismanchib, Abbas Khosravani, Lida Rashidi, Samaneh Kouchaki, Payam Peymani, Seyed Taghi Heydari, B. Sabayan, K. B. Lankarani, “Presenting a hybrid method in order to predict the 2009 pandemic influenza A (H1N1)”, Advanced Computing: An International Journal (ACIJ), Vol.3, No.1, January 2012, pp. 31-43.
18. Ken Nozaki, Hisao Ishibuchi and Hideo Tanaka, “Adaptive Fuzzy Rule-Based Classification Systems”, IEEE Transactions on Fuzzy Systems, Vol. 4, No. 3, 1996, pp. 238-250.
19. Jia Zeng and Zhi-Qiang Liu, “Type-2 Fuzzy Sets for Pattern Recognition: The State-of-the-Art”, Journal of Uncertain Systems, Vol.1, No.3, 2007, pp. 163-177.
20. F. Hoffmann, B. Baesens, J. Martens, F. Put and J. Vanthienen, "Comparing a genetic fuzzy and a Neuro-fuzzy classifier for credit scoring", presented at Int. J. Intell. Syst., 2002, pp. 1067-1083.
21. F.M. Schleif, T. Villmann, B. Hammer, “Prototype based Fuzzy Classification in Clinical Proteomics”, International Journal of Approximate Reasoning, 2008, 47(1), pp. 4-16.
22. Aaron K. Shackelford and Curt H. Davis, “A Hierarchical Fuzzy Classification Approach for High-Resolution Multispectral Data Over Urban Areas”, IEEE Transactions on Geo-science And Remote Sensing, Vol. 41, No. 9, September 2003, pp. 1920-1932.
23. Wai Kei Lei, Bing Nan LI, Ming Chui Dong, Mang I. Vai, “AFC-ECG: An Intelligent Fuzzy ECG Classifier”, A. Saad et al. (Eds.): Soft Computing in Industrial Applications, ASC 39, 2007, pp. 189–199.
24. Yun-Chi Yeh, Wen-June Wang, and Che Wun Chiou, “Heartbeat Case Determination Using Fuzzy Logic Method on ECG Signals”, International Journal of Fuzzy Systems, Vol. 11, No. 4, December 2009, pp. 250-261.

25. Mohammad Reza Homaeinezhad , Ehsan Tavakkoli, Ali Ghaffari, “Discrete Wavelet-based Fuzzy Network Architecture for ECG Rhythm-Type Recognition: Feature Extraction and Clustering-Oriented Tuning of Fuzzy Inference System”, International Journal of Signal Processing, Image Processing and Pattern Recognition Vol. 4, No. 3, September, 2011, pp. 107-130.
26. Liang-Yu Shyu, Ying-Hsuan Wu, Weichih Hu, “Using Wavelet Transform and Fuzzy Neural Network for VPC Detection From the Holter ECG”, IEEE Transactions on Biomedical Engineering, Vol. 51, No. 7, July 2004, pp. 1269-1273.
27. N. Özlem Özcan, Fikret Gurgen, “Fuzzy Support Vector Machines for ECG Arrhythmia Detection”, International Conference on Pattern Recognition, 2010, pp. 2973-2976.
28. S. Murugan & Dr. S. Radhakrishnan, “Improving Ischemic Beat Classification Using Fuzzy-Genetic Based PCA and ICA”, International Journal on Computer Science and Engineering (IJCSE), Vol. 02, No. 05, 2010, pp. 1532-1538.
29. Eduardo Ramírez, Oscar Castillo, and José Soria, “Hybrid System for Cardiac Arrhythmia Classification with Fuzzy K-Nearest Neighbors and Neural Networks Combined by a Fuzzy Inference System”, P. Melin et al. (Eds.): Soft Comp. for Recogn. Based on Biometrics, SCI 312, 2010, pp- 37–55.
30. Muhammad Arif, Muhammad Usman Akram, Fayyaz-ul-Afsar Amir Minhas, “Pruned fuzzy K-nearest neighbor classifier for beat classification”, J. Biomedical Science and Engineering, 2010, 3, pp-380-389.
31. Glayol Nazari Golpayegani & Amir Homayoun Jafari, “A novel approach in ECG beat recognition using adaptive neural fuzzy filter”, J. Biomedical Science and Engineering, 2009, 2, pp. 80-85.
32. T.M. Nazmy, H. El-Messiry, B. Al-Bokhity, “Adaptive Neuro-Fuzzy Inference System for classification of ECG signals”, The 7th International Conference on Informatics and Systems (INFOS), Date of Conference: 28-30, March 2010, pp. 1-6.
33. A. Dallali, A. Kachouri and M. Samet, “Fuzzy C-Means Clustering, Neural Network, WT and HRV For Classification of Cardiac Arrhythmia”, ARPN Journal of Engineering and Applied Sciences, Vol. 6, No. 10, October 2011, pp. 112-118.
34. R. B. Ghongade and A. A. Ghatol, “Optimization of a multi-class MLP ECG classifier using FCM”, Indian Journal of Science and Technology Vol. 3, No. 9, Sep 2010, pp. 1102-1105.
35. Rahime Ceylan, Yuksel Ozbay, Bekir Karlik, “A novel approach for classification of ECG arrhythmias: Type-2 fuzzy clustering neural network”, Expert Systems with Applications, 30 August 2008, pp. 1-6.

36. Victor-Emil Neagoe, Iuliana Florentina Iatan and Sorin Grunwald, “A Neuro-Fuzzy Approach to Classification of ECG Signals for Ischemic Heart Disease Diagnosis”, AMIA Annu Symp Proc. 2003; pp. 494–498.
37. James Theiler, “Estimating fractal dimension”, Optical Society of America, 1990, pp: 1055-1073.
38. Benoit Mandelbrot, “How long is the coast of Britain?” Science, New Series, Vol. 156, No. 3775, May 5, 1967, pp. 636-638.
39. B. Mandelbrot, “The fractal geometry of nature”, Wh. Freeman, 1983.
40. S. Raghav and K. Misra, “Fractal Feature Based ECG Arrhythmia Classification”, IEEE, 2008, pp. 1-5.
41. B. B. Chaudhuri, Nirupam Sarkar, “Texture Segmentation Using Fractal Dimension”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 17, No. 1, pp.72–77, 1995.
42. A. Conci, C.B. Proenca,” A box-counting approach to color segmentation”, *International Conference on Image Processing*, Volume 1, pp.228–230, 1997.
43. N. Sarkar, B.B. Chaudhuri, An efficient differential box-counting approach to compute fractal dimension of image, *IEEE Transactions on Systems, Man and Cybernetics*, Volume 24, Issue 1, pp.115–120, 1994.
44. D. da Silva, F. Boudon, C. Godin, O. Puech, C. Smith, H. Sinoquet, A Critical Appraisal of the Box Counting Method to Assess the Fractal Dimension of Tree Crowns, Lecture Notes in Computer Sciences (Proceedings of ISVC 2006), Volume 4291, pp.751–760, 2006.
45. S. Kobayashi, S. Maruyama, H. Kawai, K. Kudo, Estimation of 3D fractal dimension of real electrical tree patterns, Proceedings of the 4th International Conference on Properties and Applications of Dielectric Materials, Vol.1, pp.359–362, 1994.
46. Higuchi T. Approach to an irregular time series on the basis of the fractal theory. Physica D 1988; 31:277–83.
47. Use of the Higuchi’s fractal dimension for the analysis of MEG recordings from Alzheimer’s disease patients.
48. T. Higuchi, “Approach to an irregular time series on the basis of the fractal theory,” Physica D, vol. 31, no. 2, pp. 277–283, 1988.
49. Chu K. Loo, A. Samraj, G. C. Lee, “Research Article Evaluation of Methods for Estimating Fractal Dimension in Motor Imagery-Based Brain Computer Interface”, Hindawi Publishing Corporation. Discrete Dynamics in nature and society, Volume 2011, Article ID 724697, doi:10.1155/2011/724697.

50. M. J. Katz, “Fractals and the analysis of waveforms,” Computers in Biology and Medicine, vol. 18, no. 3, pp. 145 -156, 1988.
51. S. Raghav and K. Misra, “Fractal Feature Based ECG Arrhythmia Classification”, IEEE, 2008, pp. 1-5.
52. S. Spasic, “Spectral and fractal analysis of bio-signals and colored noise”, in proc. 5th IEEE Int. Symposium Intelligent system and informatics, 2007, pp: 147-149.
53. Chunan-Chien, T.-H. Lin, and B. Y. Liau, “Using correlation coefficient in ecg waveform for arrhythmia detection,” *Biomed. Eng. Applications, basis and communication*, vol. 17, no. 3, 2005.
54. S.N.Yu and K.-T. Chou, “A switchable scheme for ECG beat classification based on independent component analysis,” Expert Systems with Applications, Elsevier, vol. 33, pp. 824–829, 2007.
55. G. K. Prasad and J. S. Sahambi, “Classification of ECG arrhythmias using multi-resolution analysis and neural network,” in IEEE Tencon 03, vol. 1, 2003, pp. 227–231.
56. K. Minami, H. Nakajima, and T. Toyoshima, “Real-time discrimination of ventricular tachyarrhythmia with Fourier transform neural network,” IEEE Tran. On Biomedical Engineering, vol. 46, no. 2, pp. 179–185, 1999.
57. S. Pal and M. Mitra, ”Empirical Mode decomposition based ECG enhancement and QRS detection ”, Computers in Biology and Medicine, vol. 542, no.1, 2
58. N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, and H. H. Shin, “The Empirical Mode Decomposition Method and the Hilbert Spectrum for Non-stationary Time Series Analysis,” Proc. Royal. Soc. London, vol. 454, pp. 903–995, April 1998.
59. L. X. Song, Y. J. Wang, and Q. Wang, “Heart Rate Variability Signal Analysis based on Hilbert-Huang Transformation,” Journal of Vibration and Shock, vol. 26, pp. 30–34, 2007.