Application of Prototype Based Fuzzy Classifiers for ECG based Cardiac Arrhythmia Recognition

Muhammad Usman Akram

(Thesis Submitted in partial fulfillment of the requirements for the degree of BS in Computer and Information Science)

Pakistan Institute of Engineering and Applied Sciences

Nilore 45650, Islamabad

July, 2008

He [Allah] grants wisdom to whom He pleases; and he to whom wisdom is granted indeed receives a benefit overflowing,

But none will grasp the Message except men of understanding.

Declaration

I hereby declare that the work contained in this thesis and the intellectual content of this thesis are the product of my own work. This thesis has not been previously published in any form nor does it contain any verbatim of the published resources which could be treated as infringement of the international copyright law.

I also declare that I do understand the terms ‘copyright’ and ‘plagiarism,’ and that in case of any copyright violation or plagiarism found in this work, I will be held fully responsible of the consequences of any such violation.

Signature:_____

Author’s Name: *Muhammad Usman Akram*

Certificate of Approval

Certified that work contained in this thesis entitled

Application of Prototype Base Fuzzy Classifiers for ECG based Cardiac Arrhythmia Recognition

was carried out by Muhammad Usman Akram under our supervision and that in our opinion, it is fully adequate, in scope and quality, for the degree of BS in Computer and Information Sciences.

Approved by:

Signature: _____

Supervisor's Name: Dr. Muhammad Arif

Signature: _____

Co-Supervisor' Name: Fayyaz ul Amir Afsar Minhas

August, 2008

Dedicated to My Beloved Family,

Teachers and all Caring Friends

Acknowledgements

Praise be to Allah Almighty for He bestowed upon us intelligence, gave us the strength to unveil the secrets of the universe, bestowed us with motivation and potential to explore and the power to aim for and achieve our goals.

My greatest debt of thanks is to my supervisor, Dr. Syed Muhammad Arif, DCIS, PIEAS has been an invaluable source of support, guidance throughout this project, and I have benefited much from his insight during my time project. I would also like to thank my examiner, Fayyaz ul Amir Afsar Minhas, DCIS, PIEAS who provided much support and kept me under strict surveillance, which helped a lot in all stages of this work. I am also grate full to my examination panel members, Dr. Abdul Majid, DCIS, PIEAS and Dr. Abdul Jalil, DCIS, PIEAS for their interest and suggestions.

I must also thank Dr. Anila Usman, HOD, DCIS, PIEAS and Dr. Javed Khusrshid, DCIS, PIEAS for the provision of Special Computational Lab Facilities and workplace for the project.

My parents have always encouraged and supported my academic pursuits and I am extremely grateful for everything they done for me. Most of all, special thanks to my teacher *Miss Sobia* and my friends *Bushra Sadia*, Numan Ali, Ali Imran and all other for their sincere suggestions and constant support.

_____

Muhammad Usman Akram

PIEAS, Nilore, Islamabad

Table of Contents

List of Tables

List of Figures

Abstract

"As many as 1.4 million children are suffering from heart related diseases in Pakistan" The NEWS (Sunday 18th December, 2006). All over the world major causes of death are heart diseases. Thus, there is need of computer aided reliable system, which can share load of cardiac experts, in monitoring and detecting arrhythmias.

The project is aimed at the exploration of the various approaches involved in the Fuzzy learning of a Classification Systems. The objective is to study different techniques involved in effective learning of Fuzzy Classifier from data, and applying the built system on ECG based arrhythmia recognition

Our methodology is based on learning rules from data using prototype based Fuzzy System. DWT was used for feature extraction from segmented QRS complex. The Pruned Weighted Fuzzy K-NN (using fast search) system was used for Beat Classification. The system was also tested after adding different levels of noise in data and for data reduction, giving accuracy of about ~ 97.6% for 6 classes and ~97.0% for 9 classes. Data pruning was used to reduce training samples thus, increasing computational efficiency.

Chapter 1. Introduction

Heart diseases are leading cause of deaths in Pakistan and all over the world. In Pakistan, on average 1 out of 4 persons is suffering from cardio vascular diseases.
Early detection of such diseases is necessary, where symptoms can be seen when observed for long period of time, at initial stages. As the chip size is getting smaller, computers and algorithms getting faster, techniques for biomedical signal processing and analysis are getting more reliable. Computer Aided System for automatic Cardiac Disease Diagnosis using ECG is becoming essential for helping experts and reducing their load.

Fuzzy classification system for detection of cardiac diseases using ECG was built. ECG analysis was done to extract useful features from annotated ECG signals of normal persons and patients of various cardiac disorders.
Most of the techniques used for beat classification are too rigid and crisp to deal with automated ECG analysis. So, Lotfi A. Zadeh (University of California, Berkeley) purposed Fuzzy Sets, they were able to represent and work on natural language variable, natural language vagueness and noise robust.

The purposed methodology uses Fuzzy k- Nearest Neighbor for the classification purpose, which gives robust and reliable results. The feature extraction was done by wavelet domain analysis of ECG data. 11 features from wavelet domain analysis and RR interval were used for classification giving accuracy of about ~ 97.6% for 6 classes and ~97.0% for 9 classes.

1.1 Project Objectives

Project is aimed at development of an initially offline ECG analysis and Arrhythmia Classification System using ECG. ECG is used because it is most widely used technique for cardiac disease detection and diagnosis. The reason for widespread use of ECG is because; it is noninvasive and reliable technique for getting information about activity of heart. It is an effective way to analyze heart's electrical and mechanical activities. Any abnormality in hearts activity can be seen in ECG.

1.2 System Architecture

The steps involved in technique used are shown in Figure 1.1:

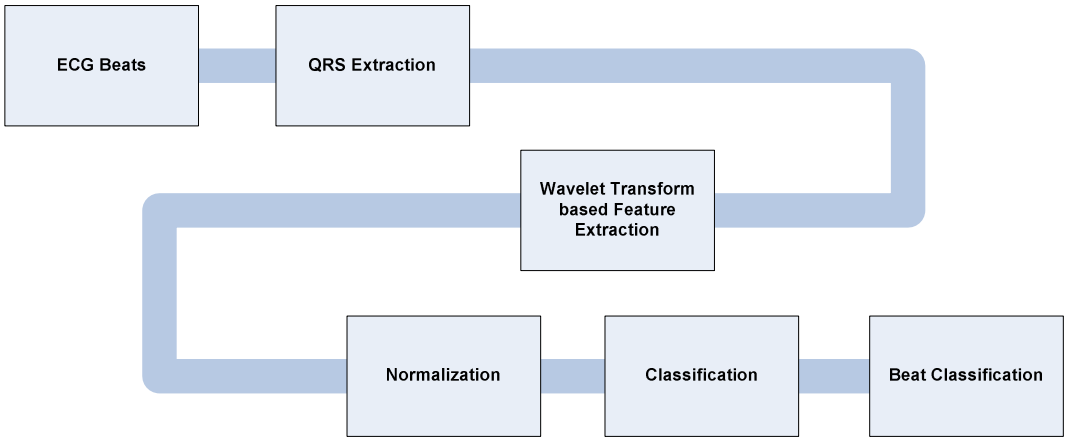


Figure 1.1 Steps of DWT Based Beat Classification

The techniques used involve signal preprocessing, QRS Detection & segmentation, feature Extraction and classification using prototype based fuzzy classifiers, and MATLAB environment was used for purpose of all implementation. MATLAB distributed computing environment was used to for feature extraction techniques to reduce time taken. Database for ECG came from the Massachusetts Institute of Technology, Beth Israel Hospital (MIT-BIH) arrhythmia database. The architecture of the system designed will be as shown in Figure 1.2.

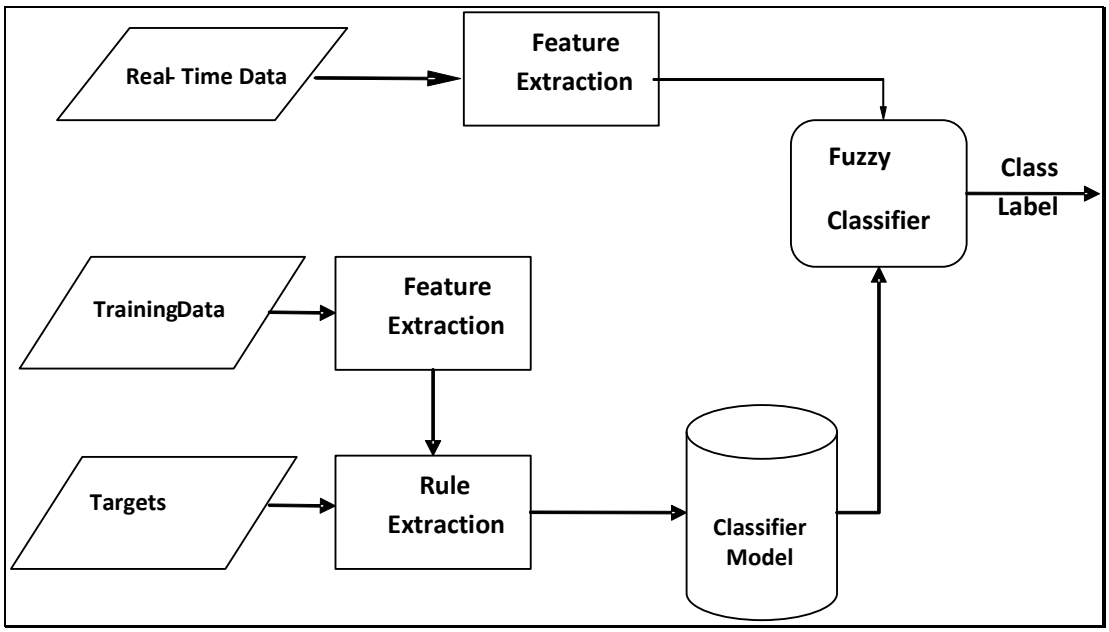


Figure 1.2 Fuzzy Classification System

1.3 Organization of thesis

Thesis draft serves as a detailed description of the project. It presents steps involved in Feature extraction and Cardiac arrhythmia Classification. Chapter 1 gives the introduction to project. Chapter 2 gives an introduction to Heart and ECG. Chapter 3 gives the detail about ECG and Cardiac Diseases. Chapter 4 gives a brief literature survey. Chapter 5 gives details about ECG Processing for Feature extraction. Chapter 6 gives description about dataset used. Chapter 7 presents Beat classification in detail. Chapter 8 presents Results & Discussion and gives details about performance evaluations. Chapter 9 gives the conclusion of work and future work proposals.

Chapter 2. Heart & ECG

Heart is biological blood pumping device. It is an electromechanical pump like any other electric pump. It also needs electric current (rather impulse) to perform its mechanical activity (pumping). Heart has special node which generate required electric impulse. Heart has even a backup battery (Backup node which provides electric impulse if primary nodes cannot provide an impulse).

2.1 Working of Heart

Heart pumps oxygenized blood from lungs to body and from deoxygenized body to lungs. Heart is an electromechanical pump which consists of myocardium (heart muscles) which contract and expand on electric impulse. All cells of heart have ability to polarize and depolarize. Heart requires a source of energy and oxygen in order to function. Energy needed for heart's pumping action comes from its built-in electrical conduction system.

Heart consists of four chambers and is located in left half of chest behind the rib cage and is surrounded by pericardium. The pericardium is a two layer structure with a adequate quantity of lubricating fluid between them. When it becomes inflamed, quantity of fluid between layers increases. This compresses heart and restricts its action. The pericardium is fluid filled membrane that surrounds heart and roots of blood vessels and acts as shock absorber by reducing friction between both layers. The pericardium has several functions. It keeps the heart contained in the chest cavity. The pericardium also prevents the heart from over expanding when blood volume increases.

Heart has two main functions gathering deoxygenized blood from body and sending it to lungs and getting oxygenized blood from lungs and sending it to body and it’ self. Upper tow chambers of heart are called atria, they are connected with veins and they are responsible for getting blood (right atrium receives blood from body via Vena Cava and left atrium receives blood from lungs via pulmonary veins). Atria’s then contract supplying stored blood to lower chambers which are called Ventricles through one way valves. On ventricular contraction blood is pushed to

body and lungs from left and right ventricles respectively. Heart also needs blood supply, which it gets from coronary arteries originating from aorta.

2.2 Electrical Activity of Heart

Being an electromechanical pump, heart uses its internal electric supply generated within heart. Cardiac contraction and expansion occur as a result of cardiac action potential.

2.3 Cardiac Action Potential and Myocardium

The significant purpose of heart cells is to contract and expand with a certain pattern for making heart pump the blood. The contractions of heart cells are triggered by an electrical impulse (spike) known as the action potential [23]. Before explaining the cardiac action potential, we must first know what an action potential is.

A cellular membrane surrounds every cell in our body. Different ions (charged particles) can move across membrane through the special ion channels. The channels only allow one type of ions to pass through at one time and block other type of ions. Due to this partial passing of ions, an electrical gradient is established across the membrane, this potential generates a contraction gradient across the membrane. Cells of our body are normally negative from inside with respect to its surroundings; this is known as resting potential. Some cells are specialized in rapidly reversing their rest potential to go positive from inside; these are known as excitable cells. The potential generated by this is known as action potential. This is caused by opening of voltage gate of ion channels which allowing to entry of positive ions through them.

Cardiac cells have a negative membrane potential when at rest. A variation between cardiac myocytes and remaining is how they raise Ca2+ (positive ions) to stimulate contraction. In cardiac myocytes, release of Ca2+ from the sarcoplasmic reticulum is induced by Ca2+ invasion into cell from voltage-gated calcium channels (ion channel only allowing calcium) causing muscle contraction. After a delay, (the absolute refractory period), Potassium channels reopen and the resulting outward flow of K+ causes repolarization to the resting state. Cells at pacemaker nodes have calcium channels which allow fast depolarization, while ventricular cells have channels which open slowly.

The action potential of the ventricular cells is used to understand cardiac action potential, also shown in Figure 2.1. The action potential has 5 phases (numbered 0-4). Phase 4 is the resting potential, when the cell is not being stimulated.

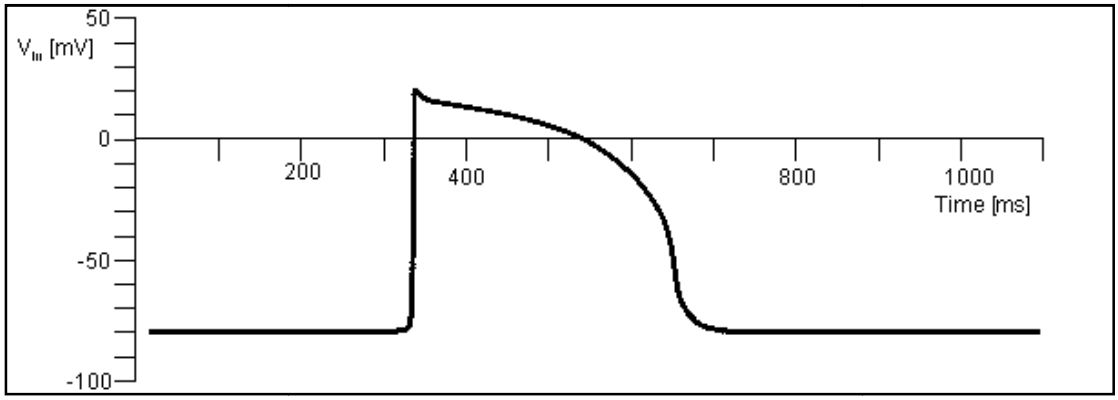


Figure 2.1 Cardiac Action Potential [21]

After being eclectically simulated the cell, it begins series of actions involving influx and efflux of different ions which produces the action potential of cell and simulating the adjacent cells. This conducts the stimulation to all cells of heart.

2.3.1 Phase 4

Phase 4 is the resting potential. Cell remains in this phase until it is stimulated by an external source. This phase of the action potential is associated with diastole.

2.3.2 Phase 0

Phase 0 is the swift depolarization phase. The slope of phase 0 represents the quick depolarization of cell. Due to the opening of fast Na+ channels causing rapid influx of Na+ ions into the cell. Ability of cell to open the fast Na+ channels during phase 0 is related to resting potential at moment of excitation.

If resting potential is at its baseline (about -85 mV), all fast Na+ channels are closed, and excitation will open them all, causing a large influx of Na+ ions.

However, if resting potential is less negative, some of fast Na+ channels will remain closed, thus resulting in lesser response to excitation of cell. For this reason, if the resting potential becomes too positive, cell may not be excitable, and conduction through the heart may be delayed, increasing the risk for arrhythmias.

2.3.3 Phase 1

This phase of action potential occurs with closing of fast Na+ channels. The transient net outward current causes small downward deflection of action potential, due to movement of K+ and Cl- ions across ion channels.

2.3.4 Phase 2

This phase of the cardiac action potential is sustained by balance between inward movement of Ca2+ through calcium channels (L-type which open slower and remain open longer) and outward movement of K+ through the slow delayed rectifier potassium channels.

2.3.5 Phase 3

During phase 3 of action potential, the L-type Ca2+ channels close, while the slow delayed rectifier K+ channels are still open. This guarantees a net outward current, corresponding to negative change in potential, allowing more types of K+ channels to open.

These are rapid delayed rectifier K+ channels and the inwardly rectifying K+ current. This net outward, positive flow causes cell repolarization. These channels close when potential is restored to about -80 to -85 mV.

2.3.6 Pacemaker Potential

The heart's cells possess ability to generate action potentials, causing cardiac contraction; sinoatrial node (Primary pacemaker node) is primarily responsible for initiation of process. SA node can generate impulses faster than other pacemaker nodes. Cardiac cells have refractory periods following contraction, during this period no other contraction can be initiated, their pacemaker potential is overridden by the sinoatrial node. Cells in SA node naturally discharge (create action potentials) at about 60-100 times/minute [20].

In case SA node stop functioning, then another cluster of cells with much similar properties becomes the pacemaker and is known as atrioventricular node (AV node), which is an area between the atria and ventricles. If even this backup node does

not work then, many small pacemaker sites start working. Conduction network with pacemakers is shown in Figure 2.2:

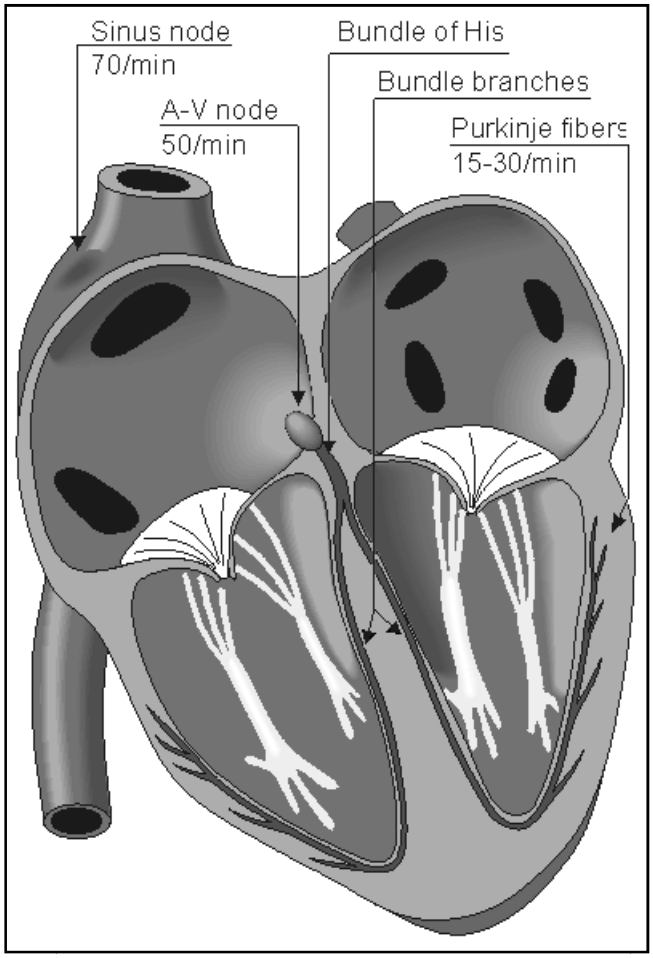


Figure 2.2 Cardiac Conduction System

2.3.7 Conduction of Electrical Activity

The electrical conduction system of heart allows propagation of impulse generated by SA node. The myocardium (heart muscle) contracts after stimulation. It is a rhythmic flow allowing efficient contraction of the heart.

The electrical potential is originated in SA node; it depolarizes and moves quickly to AV node. Conduction through AV node is very slow after which impulse again moves quickly until it reaches his bundle from where depolarization moves to whole through Purkinje fiber. The generated potential difference is moves to surface of body through tissues in contact with the heart and is used to monitor electrical activity in the heart, non-invasively.

2.4 Measurement of Cardiac Electrical Activity

Measurement of electric activity of any electrical equipment is necessary for maintaining check and balance, or to detect any abnormalities in function. Same is right in case of heart, as heart is in side body needing a noninvasive measurement technique. In 1872, Alexander Muirhead recorded of heartbeat of a fever patient by attaching wires to his wrist and using Lippmann electrometer. If these are the requirements then ECG (Electrocardiogram) is the best choice available. ECG is a noninvasive tool to observe cardiac electric activity and is in use from almost a century. ECG is widely used for analyzing heart rhythm and conduction network of heart. ECG records electrical activity of heart versus time. ECG was invented a century ago by Willem Einthoven (1901) and got Noble Prize in 1924 for his invention (Figure 2.5 Einthoven's ECG Machine). He used a string galvanometer, very thin and long filament made up of glass with silver coating to allow conduction and delicate optical equipment, as shown in Figure 2.3. Initially published ECG is shown in Figure 2.4.

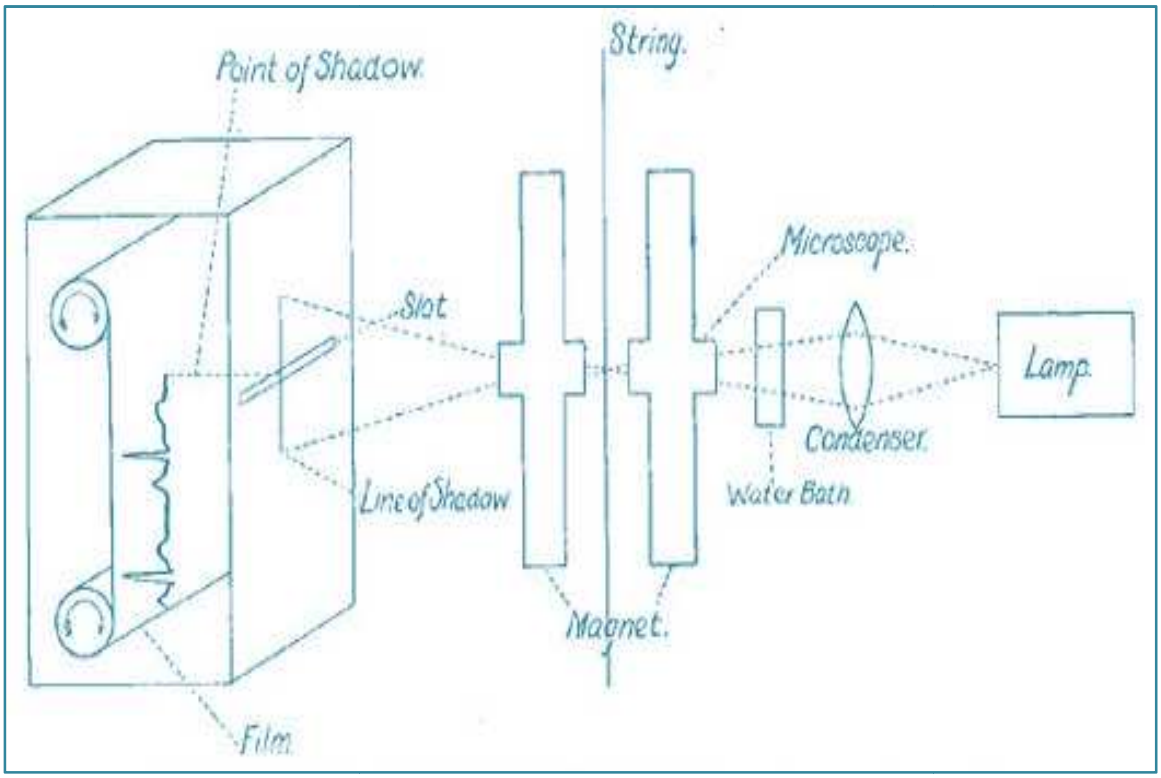


Figure 2.3 Architecture of Einthoven's ECG machine

ECG is calculated by observing electrical potential difference between electrodes connected to different parts of body using biomedical instrumentation amplifier. A specific combination of these electrodes makes up a lead.

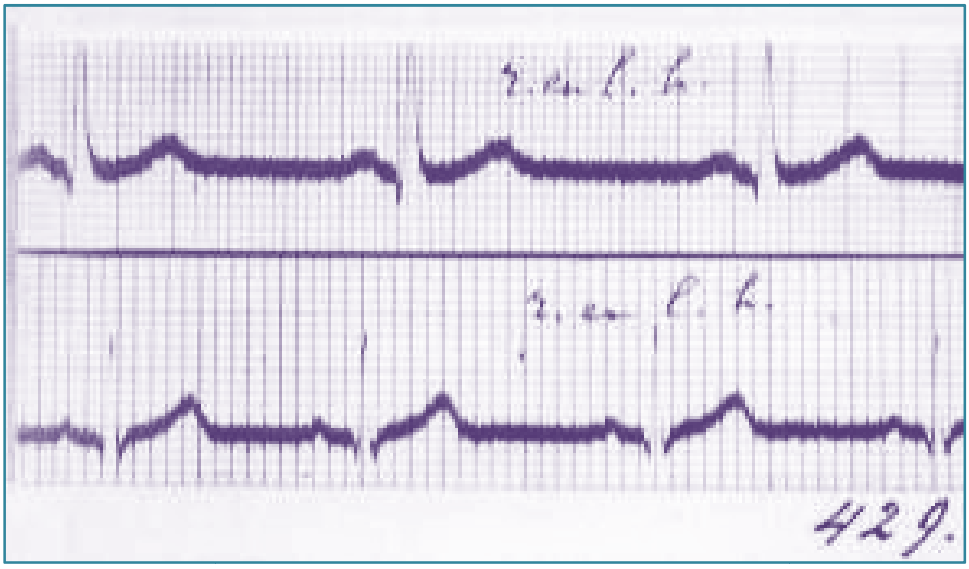
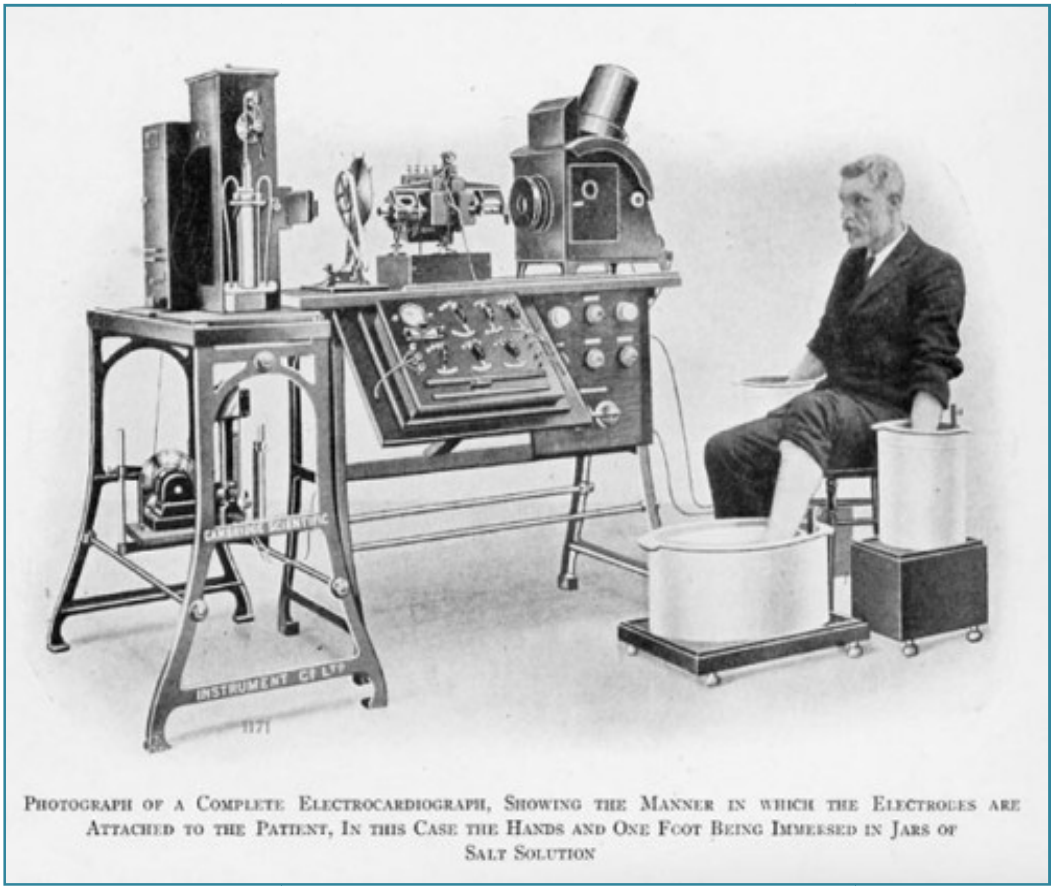


Figure 2.4 Early Einthoven ECG



PHOTOGRAPH OF A COMPLETE ELECTROCARDIOGRAPH, SHOWING THE MANNER IN WHICH THE ELECTRODES ARE ATTACHED TO THE PATIENT, IN THIS CASE THE HANDS AND ONE FOOT BEING IMMERSED IN JARS OF SALT SOLUTION

Figure 2.5 Einthoven's ECG Machine

2.5 Components of ECG

There are seven main components of an ECG wave shown in Figure 2.6.

2.5.1 Baseline

It is also called isoelectric line. It represents absence of any electrical activity in cardiac tissues. Typically it is measured as a portion of tracing following the T wave and preceding the next P wave.

2.5.2 P wave

P wave represents the atria depolarization and movement of cardiac impales from SA node to AV node.

2.5.3 PR interval

PR interval is the time gap between atrial and ventricular depolarization, it is due to the slow conduction of impulse through AV node.

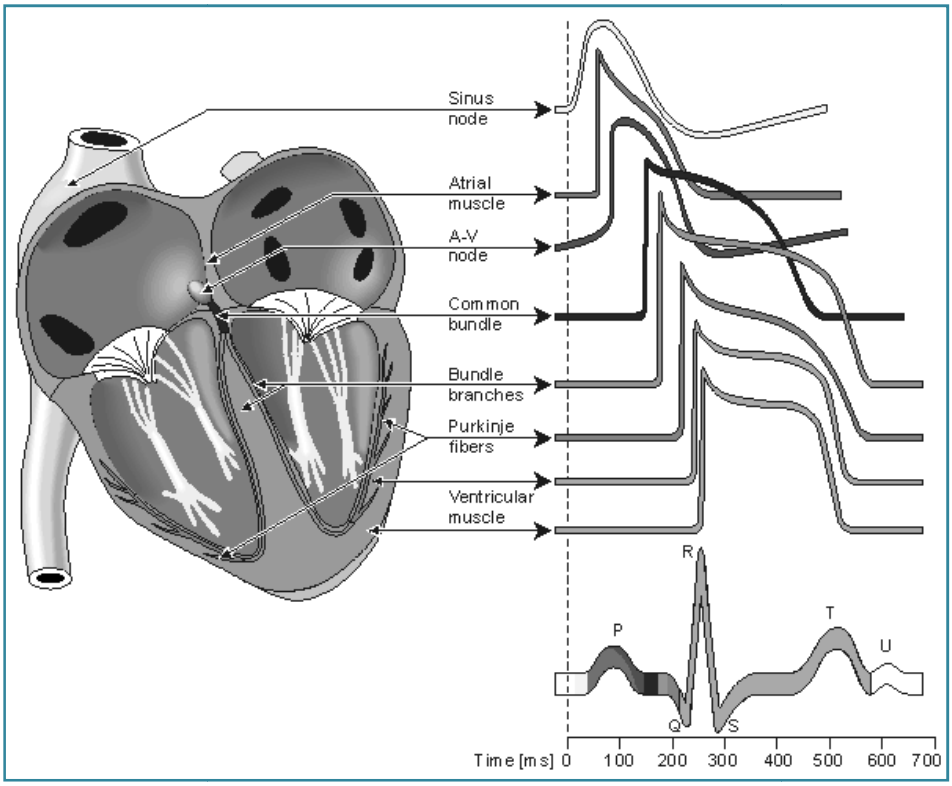


Figure 2.6 Components of ECG

2.5.4 QRS complex

This complex represents the contraction / depolarization of ventricles, when impulse is conducted through Purkinje fibers.

2.5.5 ST segment

It is isoelectric segment between S and T waves.

2.5.6 T wave

It represents ventricular repolarization.

2.5.7 U wave

U wave is normally visible in 50 to 75% of ECGs.

Chapter 3. Cardiac Diseases

ECG is a non-invasive method to monitor cardiac activities. Being a non-invasive method it is majorly used to watch out of heat activities. Pacing of heart is due to variation in bioelectric potential in heart. This potential is generated by pacemaker sites which regulates rate of heart beats. This bioelectric impulse / potential spread through special conduction network.

The normal/regular sinus rhythm shows the stability in electric conduction system and pacemaker nodes. Any abnormality in electric system may cause abnormal rhythm (also known as arrhythmias) reflecting their presence in ECG. Automated classification of arrhythmias is of great importance because some of arrhythmia (like Premature Ventricular Contraction) indicates life threatening conditions.

ECG presents electric activity in heart generated by pacemaker sites. Any incongruity in this conduction system or pacemaker sites results in change of normal ECG pattern which may be harmful.

3.1 Types of Diseases

Cardiac disease is a root term for a number of different diseases which affect the heart. Cardiac diseases are a very significant cause of death all over the globe. The most common cardiac diseases are:

- Coronary Artery Disease is a condition where deposits, cell- increase or blood clot in the arteries, resulting in a failure or reduction of blood supply to the myocardium (heart muscle) and the coronaries, which may cause infarction or heart attack.
- Ischemia or Ischemic Heart Disease (IHD) is a condition in which blood supply to the myocardium (heart muscle) is reduced, usually due to coronary artery disease (atherosclerosis of the coronary arteries). It is also known as or Myocardial Ischemia (lack of blood to heart muscles).
- Cor Pulmonale, is disorder in right ventricle, due to an increased pressure in right ventricle, causing Ventricular Hypertrophy.

- Hypertensive heart disease, caused by high blood pressure.
- Inflammatory heart disease causes inflammation of the myocardium and/or the tissue surrounding it.
- Valvular heart disease causes abnormalities in valves of the heart (on left, aortic and mitral valves and pulmonary and tricuspid valves on right).
- Arrhythmia is a term for a large number of different conditions in which there is abnormal electrical activity in the heart. The heart beat may be too fast or too slow, and may be irregular.

As the subject at hand is Cardiac Arrhythmias. Thus, we will be discussing this topic in detail now onwards.

3.2 Cardiac Arrhythmia

These abnormalities in ECG are may be result of and/or represent a cardiac disease. Some of these abnormalities are life-threatening, needing urgent medical care else can result in heart attack. While, some cause infuriating symptoms such as abnormal consciousness of heart beat, and may be merely irritating. While, others possibly not related to any symptoms, but prompt toward potentially life threatening stroke.
Arrhythmia may be classified with respect to the site of origin:

3.2.1 Atrial Arrhythmias

- Premature Atrial Contraction (PAC) / Atrial Premature Beats (APB)
- Wandering Atrial Pacemaker
- Atrial Flutter (AF)
- Atrial Fibrillation

3.2.2 Junctional Arrhythmias

- Super Ventricular Tachycardia (SVT)
- Paroxysmal Supra-ventricular Tachycardia (PSVT)
- Jynctional Rhythm
- Junctional Tachycardia
- Premature Junctional Complex

3.2.3 Atrio-ventricular

- AV reentrant tachycardia
    - Wolff-Parkinson-White syndrome
    - Lown-Ganong-Levine syndrome

3.2.4 Ventricular

- Premature Ventricular Contractions (PVC)
- Accelerated Idioventricular Rhythm
- Monomorphic Ventricular Tachycardia
- Polymorphic Ventricular Tachycardia
- Ventricular Flutter
- Ventricular Fibrillation

3.2.5 Heart Blocks

- First Degree Heart Block
- Second Degree Heart Block
- Third Degree Heart Block

3.3 Arrhythmias Used for Beat Classification

Classifying arrhythmias involves the recognition of characteristic patterns of the electrocardiogram (ECG). Beat classification is an important step in designing an arrhythmia classifier as many arrhythmias simply consist of a single abnormal beat as opposed to a constant rhythm disturbance. Type of beats (arrhythmias) used in classification are:

- Normal Sinus Rhythm (NSR)
- Bundle Branch Block (BBB)
- Premature Ventricular Contraction (PVC)
- Atrial Premature Beat (APB/PAC)
- Paced Beat/Rhythm (PB/PR)
- Ventricular Flutter (VF)

- Fusion of VF and NSR (FV&N)
- Fusion of NSR and PVC (FN&P)

3.3.1 Normal Sinus Rhythm (NSR)

This is the normal pattern of beats under a normal heart. The SA node acts as the primary source of electric signal to heart via his bundle branch. It is recognized by beat rate ranging from 60 ~ 100 bpm and regular P-waves in leads II, III and aVF. The Extracted QRS Complex for such rhythm is shown in Figure 3.1.

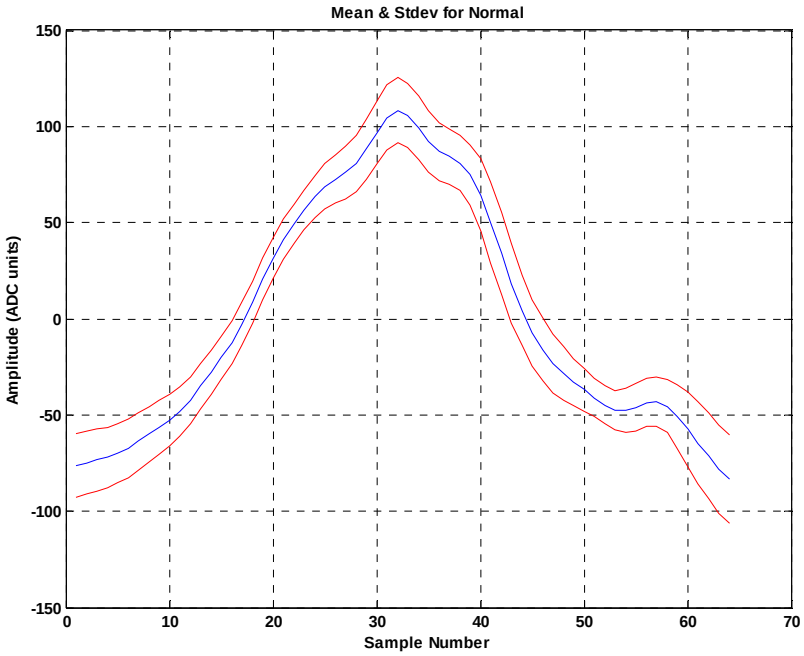


Figure 3.1 QRS Complex for NSR

3.3.2 Bundle Branch Blocks (BBB)

As the name suggests BBB is the result of a blockage in conduction of electric signal, when it passes through his Bundle Branch. There are two type of Bundle Branch Blockage:

3.3.2.1 Left BBB (LBBB)

It is caused by the blockage in left bundle branch. This blockage results in:

- ST depression in leads I, aVL, V5 and V6

- Widening of QRS complexes (>0.12s)

The Extracted QRS Complex for such rhythm is shown in Figure 3.2.

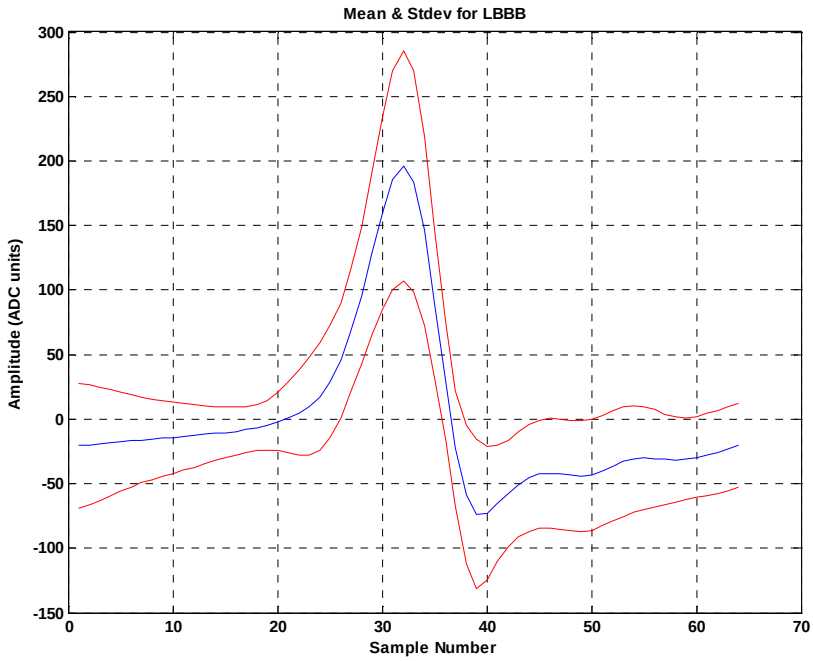


Figure 3.2 QRS Complex for LBBB

3.3.2.2 Right BBB (RBBB)

It is caused by the blockage in right bundle branch. This blockage results in:

- Terminal broad S wave in leads I
- RSR' complex in lead V1
- Widening of QRS complexes (>0.12s).

The Extracted QRS Complex for such rhythm is shown in Figure 3.3.

3.3.3 Premature Ventricular Contraction (PVC)

PVC is caused when an electrical signal is generated within ventricles before sinus beat is expected. PVC results in:

- No P-wave
- Widening of QRS complex

The Extracted QRS Complex for such rhythm is shown in Figure 3.4.

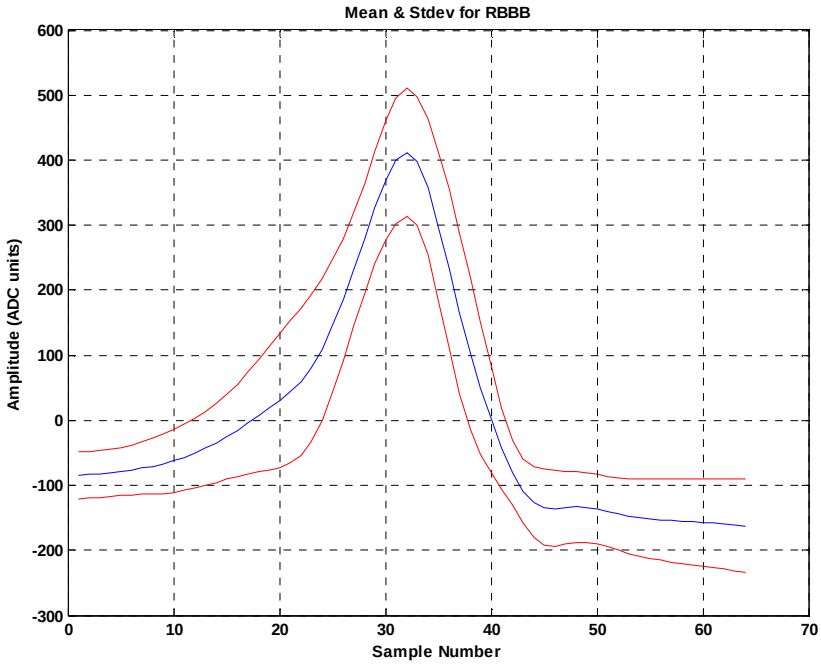


Figure 3.3 QRS Complex for RBBB

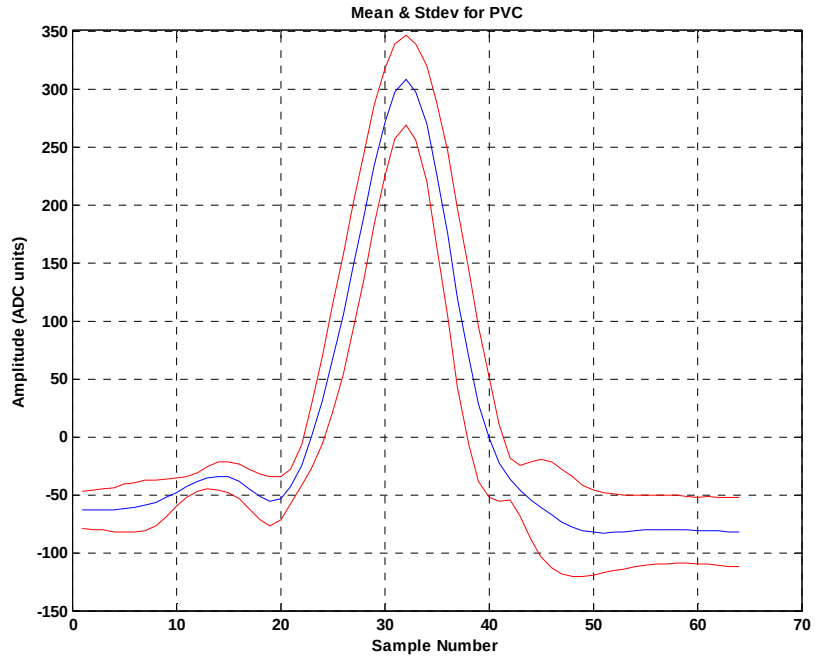


Figure 3.4 QRS Complex for PVC

3.3.4 Atrial Premature Beat (APB/PAC)

APB is caused when an electrical signal is generated within atria before sinus beat is expected and shown in Figure 3.5. APB results in:

- Contour P-wave
- Widening of PR interval
- Narrowing of QRS complex (<0.10s)

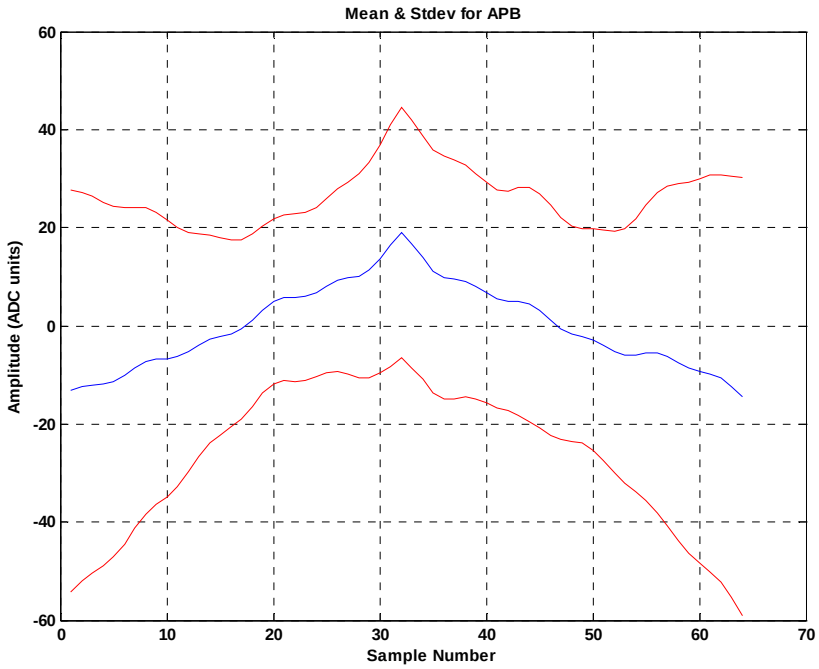


Figure 3.5 QRS Complex for Artial Premature Beats

3.3.5 Paced Beat/Rhythm (PB/PR)

It is caused by artificial pacemaker activates electric signals in to ventricles. This results in:

- No P-wave
- Widening of QRS complex

The Extracted QRS Complex for such rhythm is shown in Figure 3.6.

3.3.6 Ventricular Flutter (VF)

There is no P wave visible. A form of rapid ventricular tachycardia in which the electrocardiographic complexes assume a regular undulating pattern without distinct QRS and T waves with a frequency between 180 and 250 beats per minute. The extracted QRS complex for VF is shown in Figure 3.7.

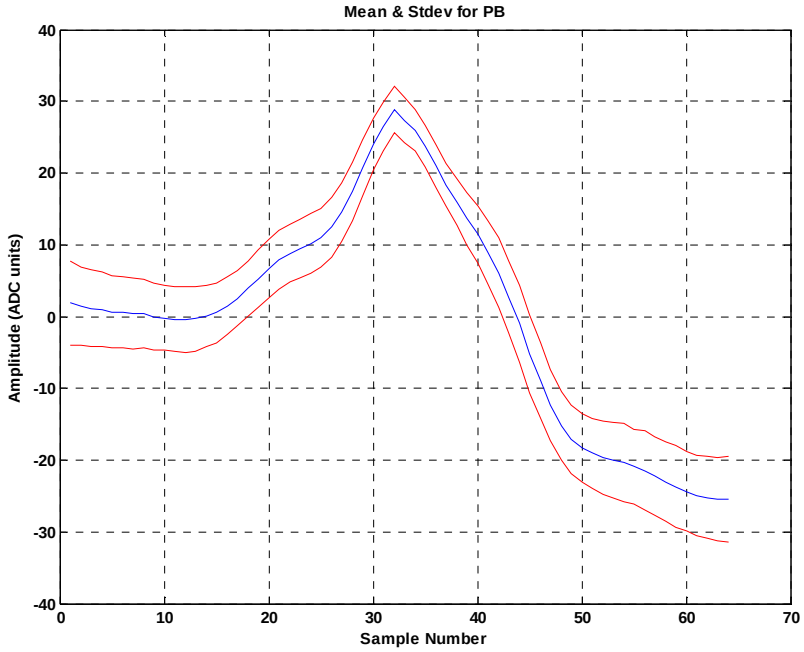


Figure 3.6 QRS Complex for Paced Beats

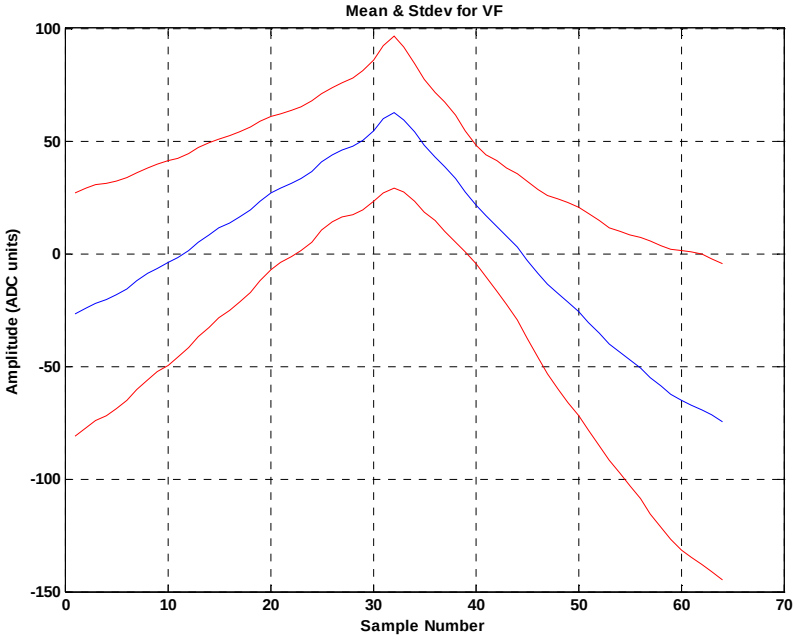


Figure 3.7 QRS Complex for Ventricular Flutter

3.3.7 Fusion Beats

Fusion beats occur when two opposing electrical currents meet and collide within the same chamber at the same time.

3.3.8 Fusion of VF and NSR (FV&N)

Fusion of VF and NSR (i.e. P wave is present). The Extracted QRS Complex for such rhythm is shown in Figure 3.8.

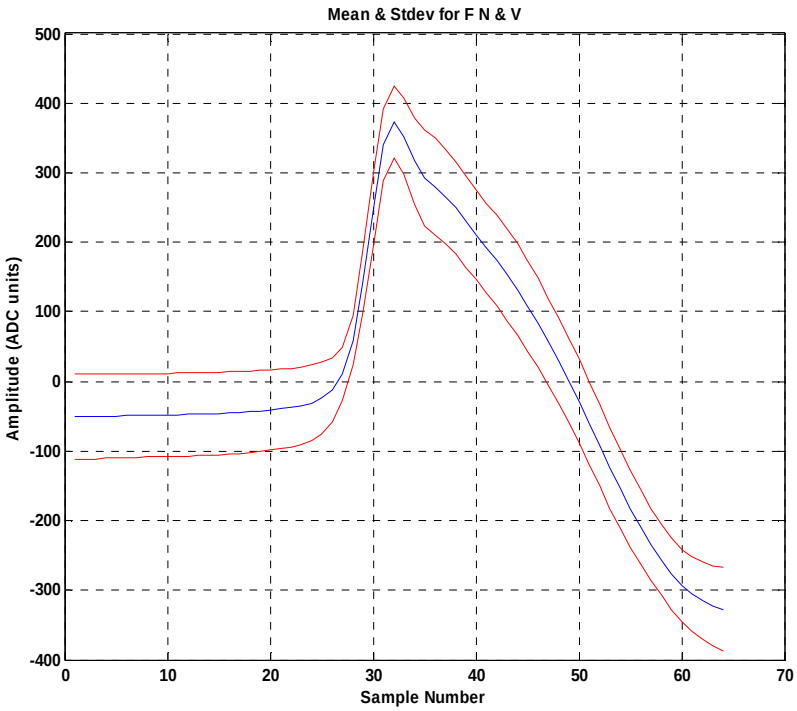


Figure 3.8 QRS Complex for Fusion of NSR & VF

3.3.9 Fusion of NSR and PVC (FN&P)

PVC occurs at in relation to the same time that a regular beat occurs and depolarizing the ventricles simultaneously in two different directions. It characterized by complexes that have both the characteristics of PVC and QRS complex of normal rhythm. The complex is usually narrower and of lesser amplitude than a PVC alone. The P to P, and R to R interval will remain constant. It cannot be recognized clinically, it can only be recognized as electrically. The criteria of fusion includes regular rhythm, and heart rate of normal rhythm, with QRS usually wider than 0.12, P wave same as normal rhythm and PR interval may be shorter than normal rhythm. Fusion of PVC and NSR (i.e. P wave is present).

The Extracted QRS Complex for such rhythm is shown in Figure 3.9.

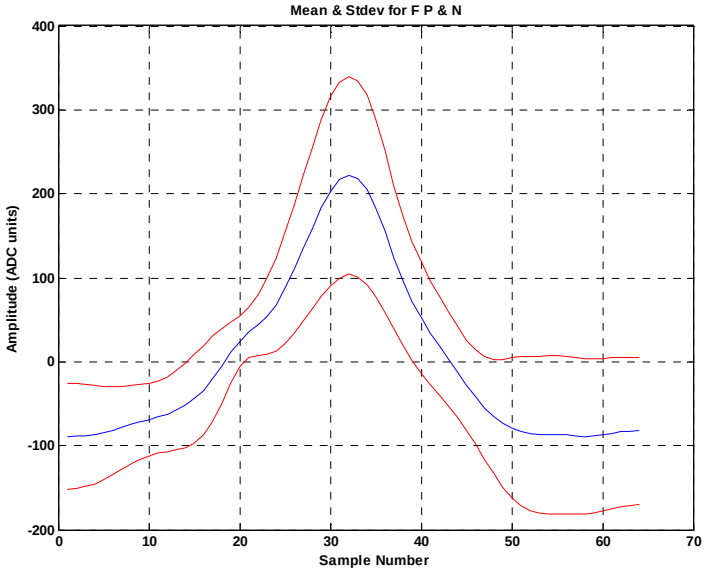


Figure 3.9 QRS Complex for NSR & PVC

Chapter 4. Literature Review

Electrocardiography deals with the electrical activity of the heart. The condition of cardiac health is given by ECG and heart rate. A study of the nonlinear dynamics of electrocardiogram (ECG) signals for arrhythmia characterization is considered. The statistical analysis of the calculated features indicates that they differ significantly between normal heart rhythm and the different arrhythmia types.

Many researches explored different techniques for classification of cardiac rhythms. There are two components of medical diagnostic system.

- Feature Extraction
- Classification

4.1 Feature Extraction

The statistical analysis of the calculated features indicates that they differ extensively between normal heart rhythm and different arrhythmia types and hence, can be rather useful in detection of ECG arrhythmia. The discrimination of ECG signals using non-linear dynamic parameters is of essential importance in the cardiac disease therapy and chaos control for arrhythmia defibrillation in the cardiac system.

Many feature extraction techniques used, exist in literature for arrhythmia classification. These can be classified in following domains:

- Time Domain
- Frequency Domain
- Time-Frequency (Wavelet)
- Filter Banks
- Blind Source Separation (BSS)
- Independent Component Analysis
- Higher Order Statistics
- Hermite Basis
- Phase Space Reconstruction
- nonlinear dynamical modeling

For the purpose of classification many classifiers build, such as; Linear Discrimination, Neural Networks, Support Vector Machines, Fuzzy and Neuro-Fuzzy Expert Systems and Ensemble based Techniques.

Many techniques exist in literature for the wandering baseline removal. These techniques can be classified as:

- Filtering Based Techniques
  A high pass filter is designed, such that it removes the slowly changing baseline. For cutoff frequency, the lowest frequency component in the ECG is selected to minimize the signal distortion.
- Polynomial Based Techniques
  Polynomial fitting based approaches aim at curve fitting on ECG or some sampled points from signal. Polynomial based approaches do not cause any distortion in ECG as is caused by linear filtering. A popular approach is to use cubic spline fitting [24].

In [11] Five levels of discrete wavelet transformation (DWT) are applied to decompose the signal into six sub band signals with different frequency distribution. Higher order statistics proceeds to calculate valuable features from the three midland signals. Three RR interval-related features are added to build a feature vector of 30 features. The proposed method demonstrates a promising accuracy of 97.53 % using Set6.

Many researchers have used effective signal analysis techniques for the classification of cardiac rhythms. In this section we present a review of existing approaches for beat classification. Minami [12] have used Fourier Transform (FT) based Frequency Domain techniques for beat classification. This method achieves a Sensitivity/PPV of ~98%.

A technique using filter banks was given by Alfonso [8]. Frequency based techniques for rhythm classification offer more reliable prospects as they are more robust to noise in contrast to time domain methods and present a more effective representation of the QRS complex. Dokur [9] carried out a comparative study Fourier Transform and Wavelet Transform (WT) demonstrating the efficiency of the WT as it provides a higher classification accuracy for ten types of beats from the MIT-BIH Arrhythmia database [10] in contrast to Fourier Transform.

4.2 Beat Classification

Classifying arrhythmias involves the recognition of characteristic patterns of the electrocardiogram (ECG). Beat classification is an important step in designing an arrhythmia classifier as many arrhythmias simply consist of a single aberrant beat as opposed to a sustained rhythm disturbance. Previous studies have employed different features in most beat classifiers.

Senhadji [25] investigated features extracted from the wavelet coefficients using linear discriminates. Hu [26] used the amplitude of points surrounding the QRS complex as features and a neural network model as the classifier.

Yeap [3] also employed neural networks as the classifier model and used the QRS width and amplitude along with three other measurements made on the ECG as features. It is difficult to compare the results as these studies employed different data sets.

A method extracting 11 features from wavelet decomposition sub-bands of an input ECG signal and applies a probabilistic neural network for classification of 6 types of beats from MIT-BIH Arrhythmia database achieving accuracy greater than 99% was presented by Yu [13].

Method using features such as heart beat intervals, RR-intervals and spectral entropy of the ECG signal along with a NN classifier to achieve an accuracy of 99.02% over the MIT-BIH Arrhythmia database was used by Niwas [13].

ECG analysis by extracting 30 principal components from the ECG signal for classifying 4 types of heart beats from the MIT-BIH arrhythmia database with an accuracy of 99.17% was performed by Hao [15].

Exarchos [16] used a rule mining approach for the classification of arrhythmias using a fuzzy inference system achieving 96% accuracy in classifying 4 types of beats (Ventricular Fibrillation, PVC, 2nd degree heart block, and Normal beats) in the MIT-BIH database using time domain features for each beat.

Yu [17] based on the use of independent component analysis (ICA) has been proposed, using 27 features for the classification of 6 types of beats from MIT-BIH database with an accuracy of 99%.

Chen [18] used DWT along with higher order statistics and RR interval features to produce a 30 dimensional feature set for detection of 8 types of rhythms

with a feed-forward back propagation NN with 97.53% accuracy. Chen et al. also present the results of their method for the classification of the 6 types of heart rhythms ~99.7% accuracy.

Chapter 5. ECG Processing

The proposed method uses features extracted from ECG through the Dyadic Wavelet Transform as in [19] for ECG delineation. The application of the same wavelet transform for ECG delineation along beat classification reduces overall system complexity. Furthermore, the Dyadic Wavelet Transform makes the feature extraction process more robust to noise. The 11 features used are also very simple to compute from the sub-band decompositions generated from DWT and there is no significant computational load. Moreover we have used PCA for further feature reduction from 11 to 6, making the process more suitable for real-time application.

To develop an efficient beat classifier, features are to be established to distinguish between different beats. The first set of features is extracted from ECG after the beat detection process. This set consists of features based on the R-R interval, the amplitude of points in the beat template and amplitudes of points in QRS template. After the QRS (onset or offset) of each beat detected is determined, features based on the QRS width 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 [27].

5.1 QRS Segmentation

As QRS complex is the most important feature of ECG, as it is associated with ventricular activation. QRS complexes are extracted on basis of R-peak identification. Window size of 64 point is used for QRS extraction centered at R-peaks.

5.2 Wavelet Transform

Only two-level DWT is used because of short length of QRS segment. This gives three sets of wavelet coefficients. These are shown in Figure 5.1.

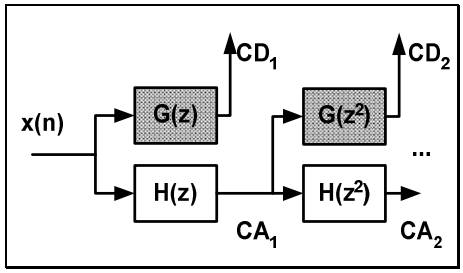


Figure 5.1 Two-level DWT Algorithm a'trous [6]

Primarily QRS complex are used which has zero crossing associated in the Wavelet Transform at scales 21 and 22 [6] therefore we utilize only these two levels.

5.3 Feature Extraction

Features are extracted from wavelet coefficients computed earlier. Details of features are:

- AC power of original signal σ_S^2. This feature measures the power in the original QRS complex signal.
- AC power of wavelet coefficients σ_{A2}^2, σ_{D2}^2 and σ_{D1}^2. This feature measures the power in each of the sub-bands.
- AC power of autocorrelation function of wavelet coefficients $\sigma_{R(A2)}^2$, $\sigma_{R(D2)}^2$ and $\sigma_{R(D1)}^2$. This is a measure of coherence in wavelet sub-bands.
- Ratio of minimum to maximum wavelet r_{A2}, r_{D2} and r_{D1}. These features represent the morphological characteristics of sub-band coefficients and the amount of change in frequency distribution of the ECG signal.

These features are combined with RR interval to get feature set given by $\left\{\sigma_S^2, \sigma_{D1}^2, \sigma_{R(D1)}^2, r_{D1}, \sigma_{D2}^2, \sigma_{R(D2)}^2, r_{D2}, \sigma_{A2}^2, \sigma_{R(A2)}^2, r_{A2}, RR\right\}$ for a single beat.

5.4 Normalization

As the features can be on different scales, normalization is necessary to homogenize all features to a same level. The relation used for normalization is given by:

$$x_{ij}^{'} = \text{tansig}\left(\frac{x_{ij} - \overline{x}_j}{\sigma_{x_j}}\right) \qquad \textbf{(4.1)}$$

Where

- x_{ij} is the j[th] component of the i^{th} feature vector, with
- $\overline{x}_j$ the mean of the j^{th} component of feature vector
- σ_{x_j} variance of the j^{th} component of feature vectors

The normalization function ranges features to [-1, 1].

5.5 Principle Component Analysis (PCA) Feature Reduction

Feature reduction is done to reduce classification time and to avoid curse of dimensionality. Feature reduction also increase performance in terms of training and classification times.

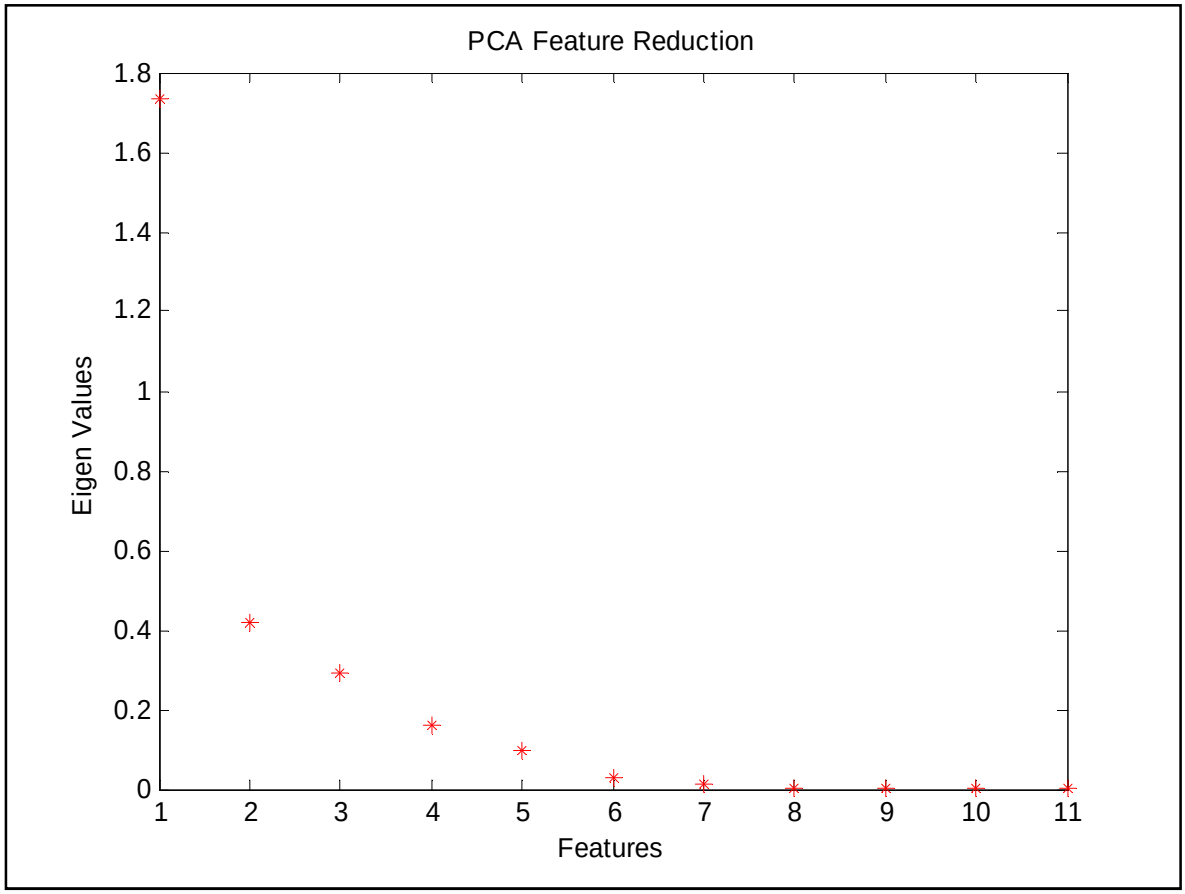


Figure 5.2 PCA Feature Reduction

Principal component analysis is a feature reduction method for achieving simplification. The method produces new set of variables, called principal

components. Each principal component is a linear combination of the original variables. All the principal components are orthogonal to each other, so there is no redundant information. The principal components as a whole form an orthogonal basis for the space of the data [7]. Figure 5.2 shows the PCA feature reduction with some features and eigen values.

Chapter 6. ECG Databases

In this work we have used MIT BIH Arrhythmia Databases from Physionet for assessing the quality of the algorithms developed. Physionet is a free resource for open access to everyone. It offers large collections of recorded biomedical signals and related open-source software. In this chapter we give a brief description of database used and its subsets used.

6.1 MIT BIH Arrhythmia Database

Database for ECG came from the Massachusetts Institute of Technology, Beth Israel Hospital (MIT-BIH) arrhythmia database between 1975 and 1979. This database has 48 subjects chosen at random from ~4000 long term Holter recordings. Each recording is 30 minutes long.

Table 6-1Different Annotations in Used Dataset

Beat Type	**No. Beats**
VF	470
PB	3616
APB	2495
FV&N	774
LBBB	8067
Normal	74716
RBBB	7247
PVC	7058
FN&P	258
Total	104690

Subjects were 25 mean aged from 32 to 89 years and 22 women aged from 23 to 89. The database contains approximately 109,000 annotated beats of different types.

6.2 Data Sets used

Following Data Sets were used:

1. Set6 (6 classes with 23200 Samples)
2. Set6 ALL (6 classes with 103199 Samples)
3. Set6 G (6 classes with 103199 Samples used for Leave One Patient Out)
4. Set9 ALL (9 classes with 104701Samples)

Chapter 7. Beat Classification

ECG presents electric activity in heart generated by pacemaker sites. Any incongruity in this conduction system or pacemaker sites results in change of normal ECG pattern which may be harmful. These abnormalities in ECG are may be result of and/or represent a cardiac disease.

7.1 Classification

For purpose of classification we used prototype based classifiers. While keeping our focus on prototype based fuzzy classifiers.

7.1.1 Prototype Based Classifiers

This class of classifiers uses exemplar data to classify unknown samples. They are easily adaptable and allow online learning. Thus, no prior knowledge about distribution of data is needed.

7.1.1.1 k - Nearest Neighbor

The Fuzzy k-nearest neighbor (k-NN) is a method was used for classification, a modified form of K- NN. It is a supervised / prototype based classification method. K-Nearest neighbors of unknown sample were selected initially and then class for that sample was selected by majority voting amongst them. The parameters in k-NN techniques are the number k which determines how many prototypes are to as the neighbors, and the distance function, generally the Euclidian distance is used.

$$\mu_i(x) = \begin{cases} 1 \; if \; x \in c_i \\ 0 \; \; if \; x \notin c_i \end{cases} \tag{7.1}$$

$$\mu_i(x) = \sum_{j=1}^{K} \mu_{ij} \tag{7.2}$$

$$C(x) = \arg\left(\max_{i=1}^{M} \left(\mu_i(x) \right) \right) \tag{7.3}$$

The problem with crisp K-NN is that how near or far a neighbor is does not matter until it is in k- nearest neighbors, it will have equal weight to other neighbors. Due to this even if a training sample lies just next to testing sample but all other majority nearest neighbors are at a longer distance compared, but the class of test sample will be determine on majority voting in which distant majority may win. Another problem with crisp K-NN is that if a tie situation occurs, the class is assigned arbitrarily to the lower class label. Moreover giving equal weights to all k- nearest neighbor prototypes can introduces error, if there is noise in prototype (relative to the chosen value of K).

7.1.1.2 Fuzzy k - Nearest Neighbor

This technique [5] is based on supervised learning using k- nearest neighbor. A sample is assigned degree of membership values of each class based on membership values of its *k* nearest neighbors in those classes. Fuzzy K-NN resolves ties by using degree of memberships of neighbors to get degree of membership of sample in each of the classes available. Classification method is also shown in Figure 7.1.

In Fuzzy K-NN an unknown sample is assigned membership to the class most represented by its K nearest neighbors, while giving a fuzzy weighting to the distance of neighbors. Fuzzy K-NN removes the crispness problem of Crisp K-NN, and generally produces more reliable & accurate results. Degree of membership of sample *x* in i^{th} class is given by:

$$\mu_i(x) = \begin{cases} 1 \; if \; x \in c_i \\ 0 \; \; if \; x \notin c_i \end{cases} \tag{7.4}$$

$$\mu_i(x) = \frac{\sum_{j=1}^{K} \mu_{ij} \left(\frac{1}{\left\| x - x_i \right\|^{\frac{2}{m-1}}} \right)}{\sum_{j=1}^{K} \left(\frac{1}{\left\| x - x_i \right\|^{\frac{2}{m-1}}} \right)} \tag{7.5}$$

$$C(x) = \arg\left(\max_{i=1}^{M} \left(\mu_i(x) \right) \right) \tag{7.6}$$

There are two significant limits of values for *m*.

- $m < 1$: For smaller values of m, distant samples have greater influence in classification of an unknown sample. For $-1 < m < 1$, as m decreases, the influence increases exponentially.
- $m > 1$: For larger the values of m, distant samples have the lesser influence in classification. For $1 < m < 3$ as m increases, the influence of distant samples decreases exponentially.
- As m approaches $\pm\infty$ the results of the classifications approach an estimation of crisp K-NN.

The values of *k* have an effect on noisy data, greater the value of *k* more robust the classification becomes. But that makes boundaries of classes fuzzy. The smaller values of *k*, makes distinct boundaries between classes but are less robust against noise. The cross validation was used to choose optimal values for *k* and *m*. Generally, heuristic techniques are used to choose optimal *k* and *m*, like cross validation or leave one out. Another popular approach is to use evolutionary algorithms to find optimal values for *k* and *m*.

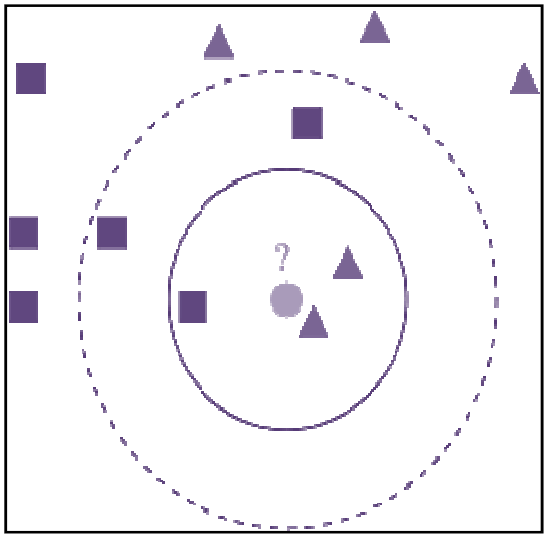


Figure 7.1 Showing nearest neighbors [4]

7.1.1.3 Pruned Weighted Fuzzy k - Nearest Neighbor (Purposed Methodology)

Reducing number of prototype sample to reduce the classification time. This is a proposed methodology for a systematic reduction in prototypes while maintaining the classification accuracy.

Proposed Algorithm:

- ***Classification***
  - *For any unknown point X*
  - *Find k-NN (X_j) using ATRIA*
  - *Evaluate membership values using*

$$\mu_i(x) = \frac{\sum_{j=1}^{K} \mu_{ij} \left(d_j^{\frac{-2}{m-1}} \right)}{\sum_{j=1}^{K} \left(d_j^{\frac{-2}{m-1}} \right)} \tag{7.7}$$

$$dj = \left\| x - x_i \right\| \tag{7.8}$$

  - *Evaluate weighted firing strengths*

$$c(x) = \arg\left(\max_{i=1}^{M} \left(w_i \mu_i(x) \right) \right) \tag{7.9}$$

- ***Training***
  - *Selection of border points*
  - *For every point in T,*
  - *Find k- nearest neighbors of other class*
  - *Add these points to prototype set P*
  - *Calculate Class Weights*

$$w_{c_i} = \left[\frac{1}{|c_i| \Big/ \min\left\{ |c_l| \,|\, l = 1...M \right\}} \right]^{1/Exp}$$

$$Exp > 1 \tag{7.10}$$

$$\mu_i(v_i) = \begin{cases} 0.517 + 0.49\,{}^{k_i}\!/_{k} & if\ class(v_i) = w_i \\ 0.49\,{}^{k_i}\!/_{k} & else \end{cases} \quad (7.11)$$

- *For every point in training set*
- *classify using prototype set P*
- *if any misclassified*
- *Add misclassified point to P (IB2)*
- *Update class weights again on any addition in P*

7.1.1.4 Large Margin Nearest Neighbor

This algorithm takes a transformation matrix L and transforms data such that samples from different classes push each other away and of same classes pull each other.

7.1.2 Parameter Optimization

Parameter optimization was done using following techniques:

7.1.2.1 Leave One Out Cross-Validation (LOOCV)

Leave 10% Out Cross Validations was used to optimize parameters (k and m) for nearest neighbor.

7.1.2.2 Convex Optimization

Convex optimization was used to find L (transformation matrix) for Large Margin Nearest Neighbor. For convex optimization problem, if a local minimum exists, then it is a global minimum and set of all (global) minima is convex. If function is strictly convex, then there exists at most one minimum. The theoretical construction for convex optimization uses the facts above in combination with notions from convex analysis such as the Hilbert projection theorem, the separating hyperplane theorem, and Farkas's lemma.

Chapter 8. Results & Discussion

This chapter presents the techniques used to classify different class of beats in the ECG. For classification purpose MATLAB distributed computing toolbox is used, to done parallel feature extraction.

This data has been previously been used for evaluation in [6].In order to provide stable analysis, mean and standard deviation of results of five runs of each experiment by randomly varying training and testing samples, i.e. training and testing datasets are not fixed as in [6]. Following parameters are used for evaluation purpose:

a. Positive Predictive Values (*PPV*) of each class. *PPV* is defined by,

$$PPV_c = \frac{TP_c}{TP_c + FP_c} \tag{8.1}$$

b. Sensitivity Values (*Se*) of each class. Sensitivity is defined by,

$$Se_c = \frac{TP_c}{TP_c + FN_c} \tag{8.2}$$

c. Total Accuracy (*A*) of each class. Total Accuracy is define by,

$$A = \frac{\sum_{k=1}^{M} TP_{c_k}}{N_{test}} \tag{8.3}$$

d. Geometric Mean of Sensitivity values (*G*), given by,

$$G = \left(\prod_{k=1}^{M} Se_{c_k} \right)^{1/M} \tag{8.4}$$

Where, TP_c is True Positives of class c, FP_c is False Positives in class c, FN_c is False Negatives in class *c*, *M* is number of classes and *Se* is sensitivity of k^{th} class.

Noise in ECG comes from a multitude of sources, like electrical interferences, muscular movement etc. Noise in ECG spans across the signal range (0 to ~40Hz) and beyond which is electrical interference (~50 to ~60Hz)[1 (34)]. The test of robustness of our system to noise, we tested the system at different levels of Gaussian white noise which effectively models majority of noise types in the ECG. We analyzed accuracy of the system at different Signal to Noise Ratios, defined by:

$$SNR = 10 * \log\left(\frac{\sigma_s^2}{\sigma_e^2}\right) \tag{8.5}$$

Where σ_s^2 and σ_e^2 are the power of the signal and noise respectively. The table below shows the effect of noise against positive productivity value, sensitivity and overall classification accuracy. Such robustness removes need of sophisticated signal processing techniques for noise removal thus lowering system complexity.

8.1 Parameter Optimization Results

Parameter optimization is one of the major tasks while applying any algorithm. We cannot get best results from any algorithm (that is inside boundary limitation of algorithm), without optimizing its parameters. Parameter optimization is a method to push any algorithm towards its best limits.

The parameters for Fuzzy and Crisp k-NN were selected using "Leave 10% Out Cross Validation". Parameters optimized for further using this method were *k* & *m*. The parameter k & m are important with respect to both Fuzzy and Crisp k-NN. They are used in evaluating membership of unknown samples.

For Crisp k-NN, *k* is the parameter which presents the number of neighbors used to evaluate an unknown sample. And for Fuzzy k-NN, *k* has the same job and has an additional parameter m which give selects that; how much factor should effect the classification. The effects of varying m are given in discussion of Fuzzy k-NN in 7.1.1.2. The equation for Fuzzy membership using *k* & *m* is given below.

$$\mu_i(x) = \frac{\sum_{j=1}^{K} \mu_{ij}\left(d_j^{\frac{-2}{m-1}}\right)}{\sum_{j=1}^{K}\left(d_j^{\frac{-2}{m-1}}\right)} \tag{8.6}$$

$m > 3$ Fuzzy k-NN starts to act like crisp k-NN. So, keeping $-1 < m < 3$ & $m \neq 1$. The favorable vale pair for *k*, *m* is 5, 1.5 respectively. Here on wards, all results will be taken on the optimized values of $k = 5$ and $m = 1.5$.

The cross validation results are presented in Table 8-1 below.

Table 8-1 Cross validation results for selection k & m

m	k = 1	k = 3	k = 5
-0.5	99.45	99.27	99.33
0	99.40	99.357	99.17
0.5	99.40	99.19	99.17
1.5	99.31	99.41	99.47
2	99.43	99.41	99.44
2.5	99.45	99.45	99.344
4	99.35	99.34	99.22

8.2 Results for Set6

For comparison of work with existing literature, set6 was used. This database has following distribution as shown in Table 8-2:

Table 8-2 Distribution of SET6

Class	Number of Samples
N	1200
LBBB	1200
RBBB	1200
APB	1500
PVC	1700
PB	1200
Total	23200

The results below present positive Geometric means for Sensitivity values and classification accuracy with their mean and standard deviations over 5 runs. For purpose of training *N/2* samples were selected each time at random from the dataset and the remaining samples were used as test set. The parameters for Fuzzy and crisp k-NN were optimized using Leave 10% Out Cross Validation. The optimized parameters *k = 5 & m = 1.5* were used for obtaining results. The sensitivity for all beats is approximately equal or greater than ~99, with total accuracy ~99.4%.

Using crisp k-NN [6] obtained total accuracy 99.49% with *k = 1* without any noise, which seems to be better than Fuzzy k-NN with 99.43% accuracy. Comparison for noise robustness is given in Table 8-3.

Table 8-3 Confusion Matrix for FKNN SET6

Data Dist.	N	LBBB	RBBB	PVC	PAC	PB	Total	
N_{train}	3604.6	2428	2387.2	1132.2	848.6	1199.4	11600	
N_{test}	3595.4	2372	2412.8	1167.8	851.4	1200.6	11600	
Class Labels	Confusion Matrix						SEN Mean	SEN Std
N	3546	0	4	0	1	0	99.82	0.09
LBBB	2	2388	9	14	1	0	98.94	0.15
RBBB	2	7	2412	1	3	1	99.50	0.14
PVC	0	10	1	1170	1	0	98.59	0.27
PAC	0	3	0	0	853	0	99.34	0.47
PB	0	1	1	0	0	1169	99.93	0.07
PPV Mean	99.82	99.00	99.24	99.16	99.11	99.95	99.35	
PPV Std	0.11	0.08	0.09	0.24	0.21	0.08	Total Acc.	99.43

Table 8-4 Noise Robustness Crisp & Fuzzy k-NN Comparison

SNR(dB)	**Crisp k-NN [6]**	**Fuzzy k-NN**
No Noise	99.4±0.1	99.4±0.03
40	99.4±0.1	99.5±0.01
35	99.3±0.1	99.5±0.02
30	99.2±0.1	99.5±0.02
25	98.9±0.0	99.4±0.06
20	98.2±0.2	99.2±0.04
15	95.8±0.1	98.8±0.12
10	86.3±0.2	97.4±0.09
5	99.4±0.1	89.6±0.36

Feature reduction was done using Principle Component Analysis (PCA). After feature reduction 6 out of 11 features were used, producing an accuracy of ~99.4%. Comparison for noise robustness is given in Table 8-5.

8.3 Results for Set6 ALL

Table 8-5presents positive Geometric Means for Sensitivity value, and overall classification accuracy with their mean and standard deviation for the results of 6 classes obtained with 11 features (without using PCA) with N/2 samples used for

training and remaining used for testing out of total of 104701 beats and Fuzzy k-NN with k = 5 was used as classifier. The sensitivity for all beats is approximately equal or greater than ~94 except APB, with total accuracy ~97.6%.

Table 8-5 Set6 ALL Noise Robustness Crisp & Fuzzy k-NN Comparison

SNR(dB)	Crisp k-NN [6]	Fuzzy k-NN
No Noise	99.4±0.1	99.4±0.03
40	99.4±0.1	99.5±0.01
35	99.3±0.1	99.5±0.02
30	99.2±0.1	99.5±0.02
25	98.9±0.0	99.4±0.06
20	98.2±0.2	99.2±0.04
15	95.8±0.1	98.8±0.12
10	86.3±0.2	97.4±0.09
5	99.4±0.1	89.6±0.36

While having ~97.6% accuracy, we can see in Table 8-6, where the beats are getting confused cause of this drop in accuracy is just by adding more samples to the training and testing sets

Table 8-6 Confusion Matrix for FKNN SET6 ALL with no noise

Data Dist.	PB	APB	LBBB	N	RBBB	PVC	Total	
N_{train}	1759	1237	4028	37339	3622	3615	51600	
N_{test}	1857	1258	4039	37377	3625	3443	51599	
Class Labels	**Confusion Matrix**				**5**	**runs**	**SEN Mean**	**SEN std**
PB	1820	0	0	1	0	0	99.94	0.00
APB	0	1044	2	173	9	13	83.47	1.15
LBBB	0	0	3747	192	3	26	94.53	0.47
Normal	0	182	178	36797	52	134	98.54	0.05
RBBB	1	6	3	33	3654	10	98.59	0.19
PVC	0	11	33	139	6	3330	94.36	0.28
PPV Mean	99.93	83.11	94.64	98.54	98.30	94.58		
PPV Std	0.05	1.11	0.35	0.05	0.14	0.26	Total Acc	97.8

The value of parameter '*k*' was chosen to be 5 through cross validation (leave 10% out). The results show a total accuracy of ~97.6% with a geometric mean of ~94.7%. Detailed results are shown below in Table 8-7.

Table 8-7 Noise Rubustness for Crisp KNN

SNR(dB)	Total	G. Mean
No Noise	97.58±0.04	94.69±0.16
40	97.57±0.09	94.68±0.24
35	97.49±0.04	94.47±0.20
30	97.40±0.06	94.43±0.17
25	97.10±0.05	93.59±0.25
20	96.42±0.07	92.18±0.27
15	95.31±0.30	89.89±0.53
10	92.52±0.09	84.47±0.37
5	84.72±0.24	68.40±0.68

We now analyze the effects of using Fuzzy KNN (FKNN) classifier. The parameter values of *k=5* and *m=1.5* were chosen through leave 10% out cross-validation. The results for Fuzzy k-NN based classification are shown in Table 8-8.

Table 8-8 Noise Robustness with FKNN

SNR(dB)	Total	G. Mean
No Noise	97.63±0.02	94.74±0.11
40	97.59±0.04	94.56±0.21
35	97.52±0.06	94.45±0.17
30	97.31±0.04	94.18±0.19
25	97.11±0.07	93.65±0.17
20	96.45±0.10	92.27±0.28
15	95.35±0.26	91.94±0.49
10	92.43±0.08	83.9±0.49
5	84.84±0.18	68.75±0.72

These results show comparable accuracy to crisp k-NN. However Fuzzy k-NN classifier gives us the membership values of the unknown sample for all possible classes, using these membership values we can calculate a confidence metric showing

the distance between winning and runner-up classes. Confidence Metric of i[th] sample is given by:

$$Conf_i = \left(\frac{\mu_{ic_{winning}} - \mu_{ic_{runner-up}}}{\sum_{k=1}^{M} \mu_{i_{c_k}}}\right) 100\% \qquad (8.9)$$

Where, $\mu_{i_{c_{winning}}}$ is the membership of i^{th} sample in winning class and $\mu_{i_{c_{runner-up}}}$ is membership in runner-up class.

Fuzzy Confidence Metric gives flexibility of choosing label from one of the confusing (i.e. winning and runner-up class labels). We attained an accuracy of ~98% by choosing appropriate label from winning and runner-up, if confidence is less than 25%, as shown in Table 8-9.

Table 8-9 Noise Robustness for FKNN using confidence Metric

SNR(dB)	Total	G. Mean
No Noise	97.93±0.03	95.05±0.21
40	97.90±0.03	95.03±0.24
35	97.84±0.07	94.90±0.18
30	97.74±0.07	94.69±0.14
25	97.41±0.05	94.07±0.23
20	96.92±0.02	92.87±0.16
15	95.94±0.27	90.48±0.86
10	93.44±0.05	84.95±0.20
5	86.87±0.17	70.07±0.32

Weighted Fuzzy k - NN gives flexibility of handling imbalance dataset problem. We attained an accuracy of ~97% by class weighting, as shown in Table 8-10.

Table 8-10 Effect of noise on Set6 ALL using FWKNN (Without PCA)

Noise	Total	G. Mean
No Noise	97.30	95.07
40	97.26	95.05
35	97.30	95.04
30	97.10	94.77
25	96.75	94.18
20	96.00	92.96
15	94.96	91.30
10	91.64	85.74
5	83.14	71.13

Weighted Fuzzy k - NN gives flexibility of handling imbalance dataset problem. We attained an accuracy of ~97% shown in Table 8-11, by class weighting along with PCA feature reduction.

Table 8-11 Effect of noise on Set6 ALL using FWKNN (With PCA)

Noise	**Total**	**G. Mean**
No Noise	97.35	95.03
40	97.33	94.86
35	97.21	94.94
30	97.13	94.77
25	96.77	94.23
20	96.12	93.05
15	94.84	90.91
10	91.60	85.34
5	83.05	71.13

8.4 Results for Set9

Table below presents positive Geometric Means for Sensitivity value, and overall classification accuracy with their mean and standard deviation for the results of 9 classes obtained with 11 features (without using PCA) with N/2 samples used for training and remaining used for testing out of total of 107647 beats and Fuzzy k-NN with k = 5 was used as classifier. The sensitivity for all beats is approximately equal or greater than ~94 except APB, with total accuracy ~97.6%.

Confusion matrix for PKNN Set6 is Table 8-12.

Table 8-12 Confusion Matrix for PKNN SET6

Data Dist	VF	PB	APB	FV&N	LBBB	Normal	RBBB	PVC	FN&P	Total
N_{train}	233	1784	1265	379	4016	37350	3687	3513	124	52351
N_{proto}	203	121	838	324	1727	5150	1265	1305	112	11045
N_{test}	237	1832	1230	395	4051	37366	3560	3545	134	52350
Class Labels	Confusion Matrix									SEN Mean
VF	195	0	6	1	0	9	0	26	0	82.28
PB	0	1817	0	0	0	1	0	1	13	99.18
APB	2	0	1014	0	2	189	9	14	0	82.44
FV&N	0	0	1	291	1	59	1	42	0	73.67
LBBB	1	1	1	1	3807	205	5	29	1	93.98
Normal	16	0	259	69	206	36558	74	173	11	97.84
RBBB	0	0	6	0	1	31	3515	6	1	98.74
PVC	41	2	14	43	32	166	9	3237	1	91.31
FN&P	0	9	0	1	8	21	4	1	90	67.16
PPV Mean	76.47	99.34	77.94	71.67	93.84	98.17	97.18	91.73	76.92	

For prototype based classifiers classification time highly depend on number of prototypes. Thus, we used pruning for reducing number of prototypes, as shown in Table 8-13.

Table 8-13 Effect of Noise on PKNN SET9 without PCA

Noise	Total Acc.	G. Mean	REDUC
40	96.78	88.15	0.21
35	96.71	88.63	0.21
30	96.36	86.48	0.22
25	96.22	85.72	0.99
20	95.39	84.58	0.26
15	93.53	77.14	0.31
10	90.36	69.86	0.38
5	81.42	51.22	0.52

With feature reduction the same classifier gave a promising total accuracy of ~96.5%, as shown in Table 8-14.

Table 8-14 Effect of Noise on PKNN SET9 with PCA

Noise	Total Acc.	G. Mean	REDUC
No Noise	96.51	86.67	0.21
40	96.45	87.12	0.21
35	96.45	87.81	0.22
30	96.20	85.17	0.22
25	95.89	84.90	0.23
20	94.94	81.13	0.25
15	92.96	74.22	0.30
10	89.94	69.15	0.38
5	80.53	47.62	0.52

Crisp k-NN resolves ties by giving arbitrary label from one of the tie holders, to avoid these ties and to get Noise robust system Fuzzy Theory was used along with k-NN. Using fuzzy for its robustness against problems of crisp k-NN gives us accuracy of ~97%, as shown in Table 8-15.

Table 8-15 Effect of Noise on FKNN

SNR(dB)	Total	G. Mean
No Noise	97.04	87.90
40	97.02	87.83
35	96.98	88.11
30	96.84	87.64
25	96.49	85.80
20	95.71	82.91
15	94.60	79.10
10	91.29	68.57
5	83.44	49.40

Having large number of samples in prototype set, increases time and space complexity of classification, to resolve these complexities pruning of prototype set was done.

Using pruning with Fuzzy k-NN gives us accuracy of ~97%, as shown in Table 8-16.

Table 8-16 Effect of noise and reduction in prototype samples with PWFKNN

SNR (dB)	Total Acc.	G Mean of Sensitivity	Reduction Factor
No Noise	96.74	89.59	0.21
40	96.77	89.35	0.21
35	96.60	89.92	1.01
30	96.51	89.36	0.22
25	96.15	88.40	0.23
20	95.26	85.94	0.26
15	94.27	81.66	0.30
10	90.20	72.19	0.38
5	81.20	57.82	0.52

Chapter 9. Conclusion & Future Proposals

This project was aimed at design of an arrhythmia classification system for cardiac arrhythmia with a particular focus on an offline classification system using ECG.

9.1 Conclusion

For purpose of classification prototype based techniques were used, which are Crisp k-NN, Fuzzy k-NN, Pruned Weighted Fuzzy k-NN. Following Table 9-1 shows a comparison of different techniques used during ECG based Arrhythmia recognition.

Table 9-1 Comparison of techniques

Data Set	Total Accuracy	G. Mean	Method
Set6	99.50%	99.40%	k-NN
Set6	99.43%	99.35%	F k-NN
Set6 ALL	97.60%	94.81%	K-NN
Set6 ALL	97.12%	93.77%	Drop3 k-NN
Set 6 ALL	97.50%	95.24%	W k-NN
Set6 ALL	97.52%	95.05%	P k-NN
Set6 ALL	97.33%	94.98%	PW k-NN
Set6 ALL	97.63%	94.73%	F k-NN
Set6 ALL	97.31%	94.74%	PF k-NN
Set6 ALL	97.30%	95.07%	WF k-NN
Set6 ALL	97.52%	95.05%	P LMNN
Set6 ALL	97.51%	95.03%	WLMNN
Set6 ALL	96.86%	92.81%	LFDA

Data Set	Total Accuracy	G. Mean	Method
Set 9	97.30%	86.40%	P k-NN
Set 9	97. 04%	87.90%	F k-NN
Set9	96.83%	86.20%	WF k-NN
Set9	96.74%	89.59%	PWF k-NN
Set6 G (LOPO)	89.24%	80.95%	PWF k-NN

9.2 Future Proposals

For future work an online or real-time system can be designed for ECG analysis of single or multiple patients at hospital or even can be used for telemedicine services. It can be interfaced with portable ECG devices (such as Holter meter) with laptops for purpose of portability. It can be modified into an adaptive system if needed by adding or removing prototypes from Prototype Set.

References

[1] Xiaoguang Chang and John H. Lilly, "Evolutionary Design of a Fuzzy Classifier From Data," *IEEE Trans. on Systems, Man and Cybernetics – part: B Cybernetics*, vol. 34, pp. 1894-1905, (2004)

[2] Plamen Angelov, Xiaowei Zhou, Dimitar Filev and Edwin Lughofer, "Architectures for Evolving Fuzzy Rule-based Classifiers," *IEEE Trans. on Systems, Man and Cybernetics, (2007)*

[3] Ethem Alpaydin, *"Introduction to Machine Learning"* , Prentice-Hall of India: New Delhi ,pp. 175-190, (2005)

[4] "k-nearest neighbor algorithm", Wikipedia, (2007), [Online] Available: http://en.wikipedia.org/wiki/Nearest_neighbor_(pattern_recognition)

[5] Matthew Kerwin, "A Fuzzy K -Nearest Neighbor Algorithm", (2005) [Online] Available: http://mattyk.nqitx.net/Honours/CP5090/FuzzyK-NN.pdf

[6] Fayyaz ul Amir Afsar Minhas and Dr. M. Arif, "Robust Classification of ECG Beats using Discrete Wavelet Transform & Principal Component Analysis", ICET, (2007)

[7] Matlab 7.6, The MathWorks-Inc, Manual, (2008)

[8] Afonso, V.X., et al. Filter bank-based ECG beat classification. in Engineering in Medicine and Biology society, 1997. Proceedings of the 19th Annual International Conference of the IEEE. (1997)

[9] Dokur, Z., T. Olmez, and E. Yazgan, *"Comparison of discrete wavelet and Fourier transforms for ECG beat classification"*. ELECTRONICS LETTERS, pp. 1502-1504, (1999)

[10] Mark, R. and G. Moody, *MIT-BIH Arrhythmia Database Directory*. 1988: Cambridge, MA: MIT Press.

[11] Ying-Hsiang Chen; Sung-Nien Yu, "Subband Features Based on Higher Order Statistics for ECG Beat Classification," *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE* , pp.1859-1862, (2007[Online] Available: http://ieeexplore. ieee.org/iel5/4352184/4352185/04352677.pdf

[12] Minami, K., H. Nakajima, and T. Toyoshima, *Real-time discrimination of ventricular tachyarrhythmia with Fourier-transform neural network.* Biomedical Engineering, IEEE Transactions on, pp. 179-185. (1999)

[13] Yu, S.-N. and Y.-H. Chen, *Electrocardiogram beat classification based on wavelet transformation and probabilistic neural network.* Pattern Recogn. Lett., p p. 1142-1150. (2007)

[14] Issac Niwas, S., R. Shantha Selva Kumari, and V. Sadasivam, *Artificial neural network based automatic cardiac abnormalities classification,* in *Computational Intelligence and Multimedia Applications, 2005. Sixth International Conference*. (2005)

[15] Hao, Z. and Z. Li-Qing, *ECG analysis based on PCA and Support Vector Machines,* in *Neural Networks and Brain, 2005. ICNN&B '05. International Conference*. (2005)

[16] Exarchos, T.P., et al., *A methodology for the automated creation of fuzzy expert systems for ischaemic and arrhythmic beat classification based on a set of rules obtained by a decision tree.* Artificial Intelligence in Medicine, pp. 187-200. (2007)

[17] Sung-Nien, Y. and C. Kuan-To, *A switchable scheme for ECG beat classification based on independent component analysis.* Expert Syst. Appl., pp. 824-829. (2007)

[18] Chen, Y.-H. and S.N. Yu. *Subband Features Based on Higher Order Statistics for ECG Beat Classification*. in *29th Annual International Conference of IEEE Engineering in Medicine and Biology Society*. 2007. Lyon, France.

[19] Martinez, J.P., et al., *A wavelet-based ECG delineator: evaluation on standard databases.* Biomedical Engineering, IEEE Transactions on, pp. 570-581. (2007)

[20] "Basics of Cardiac Arrythmias", McGill. (2000) , [Online] Available: http://sprojects.mmi.mcgill.ca/cardiophysio/AnatomySAnode.htm

[21] "The Heart", Jaakko Malmivuo, Tampere University of Technology, [Online] Available: http://butler.cc.tut.fi/~malmivuo/bem/bembook/06/06.htm

[22] S.C. Bera, B. Chakraborty, J.K. Ray, *"A mathematical model for analysis of ECG waves in a normal subject"*, MeasurementVolume 38, Issue 1, pp. 53-60. (2005)

[23] Cardiac Action Potential, Wikipedia, (2008), [Online] Available: http://en.wikipedia.org/wiki/Cardiac_action_potential

[24] Meyer, C.R. and H.N. Keiser, "*Electrocardiogram baseline noise estimation and removal using cubic splines and state-space computation techniques"*. Comput. Biomed. Res., pp. 459-470. (1977)

[25] Senhadji L, Carrault G, Bellanger JJ, Passariello G. *Comparing Wavelet Transforms for Recognizing Cardiac Patterns*. IEEE Eng in Med and Biol. (1995)

[26] Hu YH, Tompkins WJ, Urmsti JL, Afonso V. *Applications of artificial neural networks for ECG signal detection and classification.* J of Electrocardiology, pp66-73. (1993)

[27] M O'Dwyer, P de Chazal, RI3 Reilly, *Beat Classification for Use in Arrhythmia Analysis*, University College Dublin, Ireland

Appendix: Detailed Results

Following Data Sets were used:

1. Set6 (6 classes with 23200 Samples)
2. Set6 ALL (6 classes with 103199 Samples)
3. Set6 G (6 classes with 103199 Samples used for Leave One Patient Out)
4. Set9 ALL (9 classes with 104701Samples)
5. Set9 G (9 classes with 104701Samples used for Leave One Patient Out)

Appendix A. Set 6

Table A-1 Noise Effect on Fuzzy KNN (k=5,m=1.5), without PCA Feature Reduction

Noise		Effect of noise Without PCA Feature Reduction (N_{train} = 11600, N_{test} = 11600)						
		PB	**APB**	**LBBB**	**Normal**	**RBBB**	**PVC**	**Total**
No noise	PPV	99.82	99.00	99.24	99.16	99.11	99.95	99.43
		0.11	0.08	0.09	0.24	0.21	0.08	0.03
	Sensitivity	99.82	98.94	99.50	98.59	99.34	99.93	99.35
		0.09	0.15	0.14	0.27	0.47	0.07	
40	PPV	99.87	99.12	99.51	99.07	98.85	100.00	99.50
		0.06	0.18	0.08	0.30	0.57	0.00	0.01
	Sensitivity	99.76	99.22	99.41	99.02	99.42	100.00	99.47
		0.06	0.12	0.13	0.24	0.21	0.00	
35	PPV	99.89	99.13	99.52	99.05	99.00	100.00	99.52
		0.05	0.22	0.12	0.21	0.31	0.00	0.02
	Sensitivity	99.83	99.21	99.55	98.67	99.49	99.97	99.45
		0.06	0.13	0.11	0.44	0.24	0.04	
30	PPV	99.84	99.20	99.34	99.16	99.05	99.97	99.49
		0.10	0.19	0.04	0.22	0.41	0.05	0.02
	Sensitivity	99.84	99.19	99.49	98.65	99.33	99.97	99.41
		0.06	0.11	0.06	0.37	0.14	0.05	
25	PPV	99.83	98.87	99.39	98.93	98.95	99.97	99.40
		0.07	0.14	0.14	0.53	0.47	0.05	0.06
	Sensitivity	99.87	98.95	99.38	98.34	99.42	99.95	99.32
		0.05	0.23	0.20	0.18	0.38	0.08	
20	PPV	99.80	98.77	98.97	98.54	98.47	99.97	99.21
		0.06	0.25	0.15	0.24	0.38	0.08	0.04
	Sensitivity	99.78	98.42	99.29	98.28	99.12	99.88	99.13
		0.04	0.21	0.20	0.38	0.20	0.05	
15	PPV	99.69	98.15	98.25	98.16	98.18	99.92	98.83
		0.10	0.49	0.37	0.08	0.44	0.00	0.12
	Sensitivity	99.64	97.69	98.72	97.91	98.78	99.88	98.77
		0.08	0.32	0.22	0.55	0.08	0.09	
10	PPV	98.39	96.96	95.20	97.22	97.62	99.64	97.39
		0.17	0.32	0.43	0.30	0.42	0.19	0.09
	Sensitivity	98.68	95.89	95.83	97.06	98.18	99.37	97.49
		0.22	0.17	0.29	0.58	0.61	0.22	
5	PPV	91.40	89.11	81.45	93.85	94.20	94.53	89.63
		0.55	0.62	1.14	0.56	0.80	0.55	0.36
	Sensitivity	91.91	89.48	81.39	92.75	93.71	93.73	90.39
		0.34	0.61	1.19	0.57	0.60	0.48	

Table A-2 Noise Effect on Fuzzy KNN, with PCA Feature Reduction

Noise		Effect of noise With PCA Feature Reduction (N_{train} = 11600, N_{test} = 11600)						
		PB	APB	LBBB	Normal	RBBB	PVC	Total
No noise	PPV	99.78	99.17	99.27	98.80	98.85	99.89	99.39
		0.04	0.13	0.18	0.40	0.40	0.09	0.05
	Sensitivity	99.84	98.80	99.39	98.95	99.03	99.93	99.32
		0.05	0.14	0.15	0.15	0.15	0.04	
40	PPV	99.83	99.16	99.44	98.85	99.05	100.00	99.47
		0.02	0.09	0.09	0.18	0.24	0.00	0.05
	Sensitivity	99.82	99.02	99.60	98.66	99.32	99.98	99.40
		0.10	0.11	0.12	0.09	0.17	0.04	
35	PPV	99.83	99.16	99.44	98.85	99.05	100.00	99.47
		0.02	0.09	0.09	0.18	0.24	0.00	0.05
	Sensitivity	99.82	99.02	99.60	98.66	99.32	99.98	99.40
		0.10	0.11	0.12	0.09	0.17	0.04	
30	PPV	99.82	99.21	99.28	98.85	98.94	99.97	99.44
		0.04	0.22	0.28	0.14	0.33	0.05	0.09
	Sensitivity	99.83	98.77	99.49	98.91	99.48	99.97	99.41
		0.09	0.31	0.23	0.20	0.17	0.07	
25	PPV	99.83	99.11	99.28	98.80	98.82	99.88	99.40
		0.08	0.21	0.12	0.27	0.28	0.08	0.07
	Sensitivity	99.81	98.85	99.29	98.90	99.41	99.95	99.37
		0.06	0.13	0.25	0.51	0.40	0.05	
20	PPV	99.75	99.02	98.99	98.36	98.42	99.91	99.22
		0.08	0.17	0.23	0.41	0.40	0.12	0.04
	Sensitivity	99.76	98.34	99.16	98.88	99.11	99.90	99.19
		0.09	0.11	0.13	0.26	0.20	0.07	
15	PPV	99.58	98.10	97.90	98.33	98.06	99.85	98.72
		0.11	0.30	0.35	0.32	0.26	0.16	0.09
	Sensitivity	99.62	97.39	98.49	97.87	98.89	99.83	98.68
		0.09	0.26	0.21	0.51	0.61	0.17	
10	PPV	98.34	96.71	95.17	97.26	97.16	99.29	97.26
		0.09	0.49	0.25	0.48	0.89	0.04	0.14
	Sensitivity	98.58	95.76	95.65	96.72	98.27	99.28	97.37
		0.15	0.30	0.47	0.68	0.42	0.11	
5	PPV	91.04	88.82	79.01	94.29	94.09	88.25	88.35
		0.32	0.88	1.05	0.53	1.15	0.82	0.21
	Sensitivity	91.70	88.36	78.92	92.61	93.65	89.18	88.93
		0.53	0.60	0.22	0.39	0.74	1.12	

Appendix B. Set 6 ALL

Table B-1 Effect of Noise on k-NN (k = 5) without PCA Feature Reduction

Noise (db)		Effect of noise Without PCA Feature Reduction (N_{train} = 51600, N_{test} = 51599)						
		PB	APB	LBBB	Normal	RBBB	PVC	Total
No Noise	PPV	99.89	82.44	94.01	98.56	97.83	95.05	97.57
	Sensitivity	100.00	83.83	94.88	98.41	98.78	93.96	
40	PPV	99.83	82.87	94.25	98.59	98.42	94.47	97.63
	Sensitivity	100.00	83.76	95.07	98.48	98.47	94.26	
35	PPV	100.00	82.97	94.27	98.43	97.85	94.56	97.48
	Sensitivity	99.89	83.43	93.90	98.52	98.30	93.52	
30	PPV	99.95	82.20	94.05	98.36	98.46	93.75	97.39
	Sensitivity	99.89	82.73	93.26	98.42	98.40	93.86	
25	PPV	99.95	80.79	92.96	98.05	97.44	93.88	96.99
	Sensitivity	99.79	81.12	92.21	98.17	97.91	93.03	
20	PPV	99.89	79.14	91.36	97.86	96.60	92.10	96.46
	Sensitivity	99.83	79.51	91.15	97.69	97.98	92.46	
15	PPV	99.83	76.27	88.38	97.22	95.89	91.23	95.62
	Sensitivity	99.78	75.73	86.79	97.35	97.06	90.92	
10	PPV	99.44	63.91	80.50	95.24	88.03	88.81	92.55
	Sensitivity	99.38	62.38	79.49	95.47	89.01	87.57	
5	PPV	91.68	43.87	65.15	90.31	62.90	81.02	84.81
	Sensitivity	90.20	42.16	63.43	91.11	62.30	78.52	

Table B-2 Effect of Noise on k-NN (k = 5) with PCA Feature Reduction

Noise (db)	Effect of noise With PCA Feature Reduction (N_{train} = 51600, N_{test} = 51599)							
		PB	APB	LBBB	Normal	RBBB	PVC	Total
No noise	PPV	99.94	81.33	94.04	98.46	98.02	94.75	97.46
	Sensitivity	99.83	82.85	94.14	98.43	98.43	93.98	
40	PPV	99.83	82.95	94.67	98.50	98.38	94.15	97.56
	Sensitivity	100.00	83.68	94.42	98.47	98.20	94.47	
35	PPV	99.95	82.68	94.06	98.33	98.45	94.67	97.44
	Sensitivity	99.89	82.28	93.59	98.48	98.50	93.78	
30	PPV	99.94	82.81	94.08	98.36	97.87	94.19	97.39
	Sensitivity	99.94	82.09	93.66	98.45	98.48	93.42	
25	PPV	99.89	78.68	93.62	98.23	97.63	93.57	97.10
	Sensitivity	100.00	81.43	92.74	98.25	97.87	92.98	
20	PPV	99.83	76.98	91.47	97.85	96.80	92.15	96.46
	Sensitivity	99.83	76.07	90.42	97.83	98.17	92.67	
15	PPV	99.60	70.83	86.72	96.85	94.47	90.95	94.97
	Sensitivity	99.16	70.31	85.83	96.92	96.39	89.91	
10	PPV	99.41	63.99	78.97	95.10	89.79	88.77	92.48
	Sensitivity	99.46	60.79	79.30	95.51	89.00	86.70	
5	PPV	91.59	41.51	63.44	90.29	63.28	82.19	84.65
	Sensitivity	90.54	39.77	63.41	90.97	62.48	78.59	

Table B-3 Effect of Noise with FKNN with PCA Feature Reduction

Noise (db)		Effect of noise With PCA Feature Reduction (N_{train} = 51600, N_{test} = 51599)						
		PB	APB	LBBB	Normal	RBBB	PVC	Total
No Noise	PPV	99.89±0.07	82.56±0.26	94.04±0.22	98.40±0.07	97.94±0.37	94.14±0.58	97.41±0.07
	Sensitivity	99.85±0.09	83.06±0.52	93.53±0.42	98.40±0.05	98.72±0.15	93.77±0.22	
40	PPV	99.96±0.02	83.11±1.06	93.76±0.17	98.36±0.05	98.02±0.19	93.98±0.28	97.37±0.04
	Sensitivity	99.89±0.08	82.79±0.88	93.77±0.31	98.38±0.08	98.57±0.25	93.44±0.52	
35	PPV	99.89±0.08	82.80±0.92	94.05±0.39	98.38±0.05	98.07±0.31	93.72±0.31	97.37±0.07
	Sensitivity	99.94±0.06	83.54±0.59	93.33±0.29	98.36±0.06	98.56±0.09	93.94±0.28	
30	PPV	99.82±0.05	82.67±1.05	93.61±0.35	98.24±0.07	97.94±0.35	93.48±0.50	97.22±0.07
	Sensitivity	99.89±0.13	81.76±2.13	93.08±0.52	98.32±0.03	98.35±0.26	93.18±0.35	
25	PPV	99.82±0.05	81.54±0.23	92.45±0.57	98.04±0.06	97.64±0.24	93.00±0.74	96.90±0.06
	Sensitivity	99.88±0.09	80.79±0.93	91.72±0.23	98.13±0.11	98.09±0.12	92.79±0.52	
20	PPV	99.76±0.10	77.49±1.13	90.50±0.66	97.55±0.04	96.72±0.34	91.67±0.60	96.14±0.12
	Sensitivity	99.76±0.13	76.76±0.54	89.54±0.64	97.70±0.16	97.47±0.17	90.89±0.45	
15	PPV	99.39±0.55	72.17±3.26	84.93±1.54	96.80±0.08	94.55±0.37	90.27±0.69	94.79±0.16
	Sensitivity	99.49±0.43	69.92±1.86	85.10±0.76	96.84±0.17	95.44±0.34	89.78±0.84	
10	PPV	99.12±0.15	64.82±1.76	78.18±0.90	95.15±0.11	88.65±0.29	87.57±0.29	92.30±0.09
	Sensitivity	98.95±0.21	61.25±1.74	77.54±0.57	95.35±0.14	89.45±0.59	87.44±0.71	
5	PPV	84.69±0.29	41.96±1.46	62.38±0.58	90.24±0.12	61.32±0.67	79.62±0.61	84.01±0.10
	Sensitivity	85.08±0.72	40.83±0.91	62.22±1.30	90.62±0.08	60.31±0.92	78.25±0.79	

Table B-4 Effect of Noise with FKNN without PCA Feature Reduction

Noise (db)		Effect of noise Without PCA Feature Reduction (N_{train} = 51600, N_{test} = 51599)						
		PB	APB	LBBB	Normal	RBBB	PVC	Total
No Noise	PPV	99.93±0.05	83.11±1.11	94.65±0.35	98.55±0.05	98.30±0.14	94.58±0.26	97.63±0.02
	Sensitivity	99.95±0.00	83.47±1.15	94.53±0.47	98.54±0.05	98.59±0.19	94.36±0.28	
40	PPV	99.91±0.06	83.51±0.68	94.45±0.42	98.48±0.06	98.20±0.04	94.96±0.44	97.59±0.04
	Sensitivity	99.91±0.10	82.83±1.30	94.48±0.21	98.54±0.04	98.63±0.16	94.09±0.49	
35	PPV	99.89±0.07	82.38±1.18	94.60±0.31	98.47±0.05	98.11±0.13	94.38±0.43	97.52±0.06
	Sensitivity	99.97±0.05	82.42±0.55	94.14±0.20	98.48±0.08	98.59±0.20	94.26±0.31	
30	PPV	99.82±0.09	81.56±1.43	93.65±0.40	98.37±0.08	97.99±0.23	93.99±0.21	97.31±0.04
	Sensitivity	99.95±0.05	82.36±0.74	93.57±0.24	98.34±0.09	98.47±0.16	93.54±0.36	
25	PPV	99.91±0.06	81.38±0.89	93.00±0.70	98.21±0.08	97.74±0.22	93.65±0.71	97.11±0.07
	Sensitivity	99.90±0.05	80.79±1.27	93.06±0.27	98.23±0.03	98.44±0.18	92.90±0.27	
20	PPV	99.85±0.05	78.53±1.23	91.55±0.16	97.74±0.04	96.92±0.33	92.49±0.37	96.45±0.10
	Sensitivity	99.81±0.15	77.78±0.64	90.12±0.46	97.88±0.08	97.77±0.29	92.19±0.21	
15	PPV	99.69±0.39	73.98±1.88	87.08±0.69	97.06±0.12	95.22±0.79	91.87±0.51	95.35±0.26
	Sensitivity	99.67±0.32	72.67±2.59	86.84±0.35	97.21±0.17	96.20±0.60	90.25±0.49	
10	PPV	99.18±0.12	63.91±1.31	79.23±0.62	95.16±0.09	88.66±0.33	88.22±0.38	92.43±0.08
	Sensitivity	99.25±0.14	60.57±0.83	78.20±0.81	95.51±0.07	89.35±0.41	86.98±0.62	
5	PPV	91.46±0.62	41.23±1.30	64.02±0.86	90.53±0.07	63.81±1.37	80.61±0.86	84.84±0.18
	Sensitivity	91.15±0.72	40.57±1.83	63.42±0.65	91.02±0.17	63.01±1.00	78.53±0.94	

Table B-5 Effect of variation of number of training examples for FKNN without PCA

N_{train}	Effect of variation of number of training examples (N_{train}) Without PCA Feature Reduction (N_{test} = 103199 - N_{train})							
		PB	APB	LBBB	Normal	RBBB	PVC	Total
51600	PPV	99.93±0.05	83.11±1.11	94.65±0.35	98.55±0.05	98.30±0.14	94.58±0.26	97.63±0.02
	Sensitivity	99.95±0.00	83.47±1.15	94.53±0.47	98.54±0.05	98.59±0.19	94.36±0.28	
25800	PPV	99.88±0.08	81.85±0.66	93.76±0.31	98.31±0.08	97.80±0.33	93.82±0.43	97.27±0.08
	Sensitivity	99.93±0.05	81.58±0.63	93.46±0.33	98.35±0.05	98.32±0.23	93.38±0.49	
12900	PPV	99.89±0.04	79.18±1.18	92.63±0.53	98.07±0.13	97.63±0.17	93.71±0.57	96.92±0.09
	Sensitivity	99.88±0.03	80.33±0.80	92.13±0.62	98.17±0.05	98.26±0.21	92.13±0.42	
6450	PPV	99.76±0.13	79.73±1.23	91.11±0.73	97.91±0.07	96.85±0.46	92.26±0.96	96.55±0.11
	Sensitivity	99.57±0.39	77.81±1.05	91.36±0.78	97.90±0.08	97.97±0.49	91.81±0.57	
3225	PPV	99.67±0.21	75.25±3.73	89.60±1.36	97.39±0.12	96.10±0.94	92.04±1.07	95.88±0.11
	Sensitivity	99.57±0.26	74.54±2.79	88.31±0.75	97.63±0.13	97.55±0.70	89.98±0.66	
1613	PPV	99.63±0.10	71.87±3.60	86.91±2.23	96.84±0.18	94.76±1.01	90.48±1.41	95.01±0.26
	Sensitivity	99.57±0.24	69.99±3.21	85.56±1.58	97.21±0.18	96.37±0.33	87.52±2.31	

Table B-6 Effect of variation of number of training examples for FKNN with PCA

N_{train}	Effect of variation of number of training examples (N_{train}) With PCA Feature Reduction (N_{test} = 103199 - N_{train})							
		PB	APB	LBBB	Normal	RBBB	PVC	Total
51600	PPV	99.89±0.07	82.56±0.26	94.04±0.22	98.40±0.07	97.94±0.37	94.14±0.58	97.41±0.07
	Sensitivity	99.85±0.09	83.06±0.52	93.53±0.42	98.40±0.05	98.72±0.15	93.77±0.22	
25800	PPV	99.90±0.03	80.93±1.47	93.22±0.29	98.24±0.10	97.89±0.14	93.28±0.55	97.13±0.07
	Sensitivity	99.88±0.05	81.44±1.14	92.63±0.41	98.22±0.07	98.39±0.16	93.48±0.02	
12900	PPV	99.80±0.12	80.59±1.76	92.11±0.52	97.98±0.05	97.34±0.38	93.05±0.50	96.79±0.10
	Sensitivity	99.61±0.41	79.27±1.44	91.57±0.37	98.10±0.16	98.11±0.18	92.26±0.30	
6450	PPV	99.82±0.05	77.44±2.31	90.55±0.97	97.78±0.05	97.23±0.51	91.96±0.80	96.35±0.09
	Sensitivity	99.65±0.40	78.06±3.15	90.44±0.69	97.75±0.11	97.55±0.21	91.84±0.80	
3225	PPV	99.57±0.28	77.88±3.03	88.63±1.33	97.37±0.16	96.38±0.51	90.73±1.64	95.78±0.16
	Sensitivity	99.31±0.37	74.99±3.25	87.43±0.85	97.54±0.19	97.48±0.49	90.52±0.81	
1613	PPV	99.39±0.06	75.24±3.57	85.26±1.42	96.96±0.21	95.19±0.74	89.88±2.30	95.01±0.18
	Sensitivity	99.29±0.32	70.24±2.24	86.38±1.19	97.07±0.18	96.34±0.65	88.27±1.98	

Table B-7 Effect of Noise on Pruning FKNN with PCA Feature Reduction

Noise (dB)		Effect of Noise With PCA Feature Reduction (N_{train} = 51600, N_{test} = 51599)							
		PB	APB	LBBB	Normal	RBBB	PVC	Accuracy/ GM	Red. Fact.
No Noise	PPV	99.92	78.88	93.56	98.54	97.62	93.92	97.31	0.19
		0.049	0.85	0.30	0.08	0.21	0.32	0.10	
	Sensitivity	99.91	84.46	94.24	98.13	98.78	93.78	94.74	
		0.083	1.14	0.40	0.09	0.14	0.20		
40	PPV	99.96	78.88	93.39	98.52	97.57	94.16	97.29	0.19
		0.025	1.98	0.22	0.07	0.27	0.28	0.03	
	Sensitivity	99.85	83.28	94.70	98.14	98.69	93.55	94.53	
		0.14	1.25	0.22	0.09	0.18	0.56		
35	PPV	99.92	78.75	93.33	98.52	97.39	93.70	97.23	0.19
		0.06	1.39	0.23	0.05	0.09	0.36	0.05	
	Sensitivity	99.88	83.98	94.08	98.06	98.61	94.05	94.62	
		0.07	0.62	0.25	0.05	0.19	0.42		
30	PPV	99.91	78.14	92.92	98.41	97.49	93.62	97.12	0.19
		0.05	1.52	0.52	0.07	0.32	0.43	0.20	
	Sensitivity	99.91	82.80	93.94	98.02	98.51	93.39	94.25	
		0.12	1.01	0.46	0.08	0.33	0.55		
25	PPV	99.86	77.02	91.76	98.26	97.01	93.11	96.82	0.20
		0.06	0.80	0.21	0.05	0.32	0.13	0.06	
	Sensitivity	99.89	81.49	93.04	97.80	98.39	92.90	93.71	
		0.07	1.14	0.28	0.05	0.21	0.37		
20	PPV	99.89	72.67	89.89	97.78	96.11	91.80	96.04	0.23
		0.10	0.63	0.48	0.09	0.22	0.65	0.08	
	Sensitivity	99.73	78.65	90.36	97.31	97.83	91.53	92.28	
		0.116	0.86	0.18	0.08	0.26	0.75		
15	PPV	99.80	68.37	85.12	97.14	94.05	90.61	94.87	0.28
		0.11	3.00	0.88	0.12	0.66	0.59	0.31	
	Sensitivity	99.53	73.40	87.47	96.50	96.32	89.64	90.03	
		0.33	1.98	0.24	0.28	0.54	0.62		
10	PPV	99.33	57.45	75.66	95.24	85.26	87.21	91.52	0.36
		0.17	1.38	0.41	0.09	0.31	0.84	0.07	
	Sensitivity	99.14	61.65	78.70	94.16	90.21	86.28	84.05	
		0.17	1.90	0.61	0.07	0.56	0.58		
5	PPV	91.60	36.99	59.34	90.34	56.53	79.26	82.91	0.49
		0.60	2.32	0.99	0.17	0.59	0.83	0.13	
	Sensitivity	90.24	41.45	64.04	88.43	62.94	77.70	68.53	
		0.90	2.20	0.90	0.14	0.58	0.62		

Table B-8 Effect of Noise on Pruning FKNN without PCA Feature Reduction

Noise (dB)		Effect of Noise Without PCA Feature Reduction (N_{train} = 51600, N_{test} = 51599)							
		PB	APB	LBBB	Normal	RBBB	PVC	Acc	Red.
No Noise	PPV	99.93	78.61	93.66	98.51	97.51	94.46	97.32	0.19
		0.05	0.68	0.23	0.06	0.33	0.52	0.05	
	Sensitivity	99.86	83.49	94.60	98.17	98.65	93.72	94.58	
		0.076	0.98	0.32	0.03	0.14	0.50		
40	PPV	99.88	78.77	93.65	98.54	97.55	94.07	97.32	0.19
		0.059	0.94	0.52	0.04	0.25	0.52	0.06	
	Sensitivity	99.89	83.50	94.62	98.12	98.70	94.03	94.64	
		0.10	0.35	0.27	0.08	0.18	0.31		
35	PPV	99.91	77.90	93.21	98.48	97.51	93.63	97.17	0.19
		0.03	1.48	0.29	0.04	0.13	0.26	0.08	
	Sensitivity	99.94	83.42	94.30	98.02	98.49	93.67	94.47	
		0.04	0.74	0.17	0.08	0.23	0.49		
30	PPV	99.93	78.12	93.18	98.39	97.42	93.63	97.12	0.19
		0.06	1.63	0.56	0.043	0.14	0.46	0.04	
	Sensitivity	99.82	83.35	93.89	98.04	98.53	93.07	94.28	
		0.08	0.94	0.48	0.06	0.07	0.45		
25	PPV	99.88	77.31	91.83	98.20	97.06	92.75	96.76	0.20
		0.06	0.95	0.31	0.03	0.29	0.38	0.04	
	Sensitivity	99.88	82.57	92.44	97.77	98.31	92.86	93.79	
		0.09	0.55	0.21	0.04	0.17	0.21		
20	PPV	99.82	74.05	89.83	97.79	96.34	91.87	96.12	0.23
		0.088	0.75	0.23	0.12	0.37	0.26	0.09	
	Sensitivity	99.83	78.29	90.77	97.40	97.76	91.40	92.28	
		0.08	1.24	0.55	0.06	0.34	0.73		
15	PPV	99.68	68.52	85.57	97.10	94.49	90.43	94.91	0.28
		0.28	0.74	0.47	0.14	0.59	0.81	0.17	
	Sensitivity	99.68	73.58	86.53	96.58	96.45	90.14	90.04	
		0.19	1.90	0.70	0.04	0.44	1.14		
10	PPV	99.38	56.74	76.41	95.25	86.13	87.35	91.66	0.36
		0.18	1.14	0.28	0.12	0.35	0.72	0.10	
	Sensitivity	98.87	62.85	78.77	94.34	89.69	86.32	84.24	
		0.23	1.49	0.56	0.12	0.73	0.45		
5	PPV	91.36	37.02	59.78	90.52	56.68	79.38	83.02	0.50
		0.60	1.04	0.54	0.13	1.20	0.64	0.10	
	Sensitivity	90.16	43.16	64.58	88.36	63.73	77.80	69.23	
		0.99	1.02	1.32	0.17	0.50	0.82		

Table B-9 Effect of Noise Pruned Weighted FKNN with PCA Feature Reduction

Noise (db)		Effect of Noise With PCA Feature Reduction (N_{train} = 51600, N_{test} = 51599)							
		PB	APB	LBBB	Normal	RBBB	PVC	Total	Red.
No Noise	PPV	99.76	77.34	93.01	98.73	97.33	93.84	97.30	0.19
		0.06	1.36	0.41	0.06	0.40	0.36	0.04	
	Sensitivity	99.92	84.97	95.36	97.94	98.76	94.29	95.07	
		0.05	0.30	0.15	0.06	0.25	0.43		
40	PPV	99.76	77.12	93.05	98.69	97.49	93.53	97.26	0.19
		0.12	1.44	0.18	0.08	0.21	0.56	0.10	
	Sensitivity	99.91	85.16	94.71	97.92	98.90	94.50	95.05	
		0.09	1.31	0.31	0.07	0.11	0.32		
35	PPV	99.84	77.08	93.03	98.66	97.28	93.21	97.21	0.19
		0.05	0.94	0.35	0.07	0.18	0.32	0.10	
	Sensitivity	99.87	84.97	94.60	97.88	98.85	94.26	94.94	
		0.06	1.03	0.28	0.06	0.13	0.48		
30	PPV	99.77	76.26	92.58	98.63	97.16	93.17	97.10	0.20
		0.13	1.08	0.36	0.06	0.35	0.29	0.08	
	Sensitivity	99.94	84.51	94.60	97.80	98.64	94.01	94.77	
		0.04	0.63	0.27	0.06	0.23	0.70		
25	PPV	99.80	73.94	91.61	98.40	97.27	92.51	96.75	0.21
		0.09	0.68	0.40	0.03	0.39	0.20	0.06	
	Sensitivity	99.87	83.50	93.20	97.57	98.51	93.45	94.18	
		0.14	0.89	0.31	0.09	0.23	0.29		
20	PPV	99.87	70.91	88.95	98.02	95.91	91.37	96.00	0.23
		0.10	1.26	0.61	0.01	0.14	0.27	0.08	
	Sensitivity	99.79	80.80	91.34	96.99	98.01	92.20	92.96	
		0.10	1.41	0.49	0.13	0.26	0.22		
15	PPV	99.83	67.43	85.05	97.46	94.29	90.04	94.96	0.28
		0.09	1.90	0.52	0.10	0.45	0.32	0.17	
	Sensitivity	99.61	78.03	87.88	96.26	96.65	91.12	91.30	
		0.08	0.76	0.54	0.16	0.40	0.27		
10	PPV	99.20	56.18	74.88	95.73	86.09	86.88	91.64	0.36
		0.33	1.01	0.44	0.13	1.20	0.16	0.12	
	Sensitivity	99.14	67.15	80.53	93.81	90.30	87.49	85.74	
		0.15	0.38	0.77	0.15	0.50	0.62		
5	PPV	90.23	36.82	59.82	91.18	56.90	78.63	83.14	0.50
		1.11	0.73	0.33	0.17	0.51	0.30	0.12	
	Sensitivity	90.59	47.23	65.72	87.85	65.50	80.06	71.13	
		1.05	0.97	1.12	0.16	0.71	0.61		

Table B-10 Effect of Noise Pruned Weighted FKNN without PCA Feature Reduction

Noise (dB)	Effect of Noise Without PCA Feature Reduction (N_{train} = 51600, N_{test} = 51599)

		PB	APB	LBBB	Normal	RBBB	PVC	Acc	Reduc.
No Noise	PPV	99.85	77.95	93.41	98.70	97.46	93.73	97.35	0.19
		0.07	0.84	0.49	0.05	0.17	0.33	0.04	
	Sensitivity	99.90	84.89	95.07	98.03	98.64	94.48	95.03	
		0.04	0.50	0.33	0.09	0.29	0.18		
40	PPV	99.83	77.73	93.39	98.66	97.46	94.03	97.33	0.19
		0.08	0.72	0.36	0.07	0.27	0.51	0.05	
	Sensitivity	99.87	84.12	95.09	98.05	98.69	94.29	94.86	
		0.12	0.49	0.20	0.06	0.20	0.72		
35	PPV	99.84	77.08	93.03	98.66	97.28	93.21	97.21	0.19
		0.05	0.94	0.35	0.07	0.18	0.32	0.10	
	Sensitivity	99.87	84.97	94.60	97.88	98.85	94.26	94.94	
		0.06	1.03	0.28	0.06	0.13	0.48		
30	PPV	99.85	76.84	92.74	98.60	97.39	92.95	97.13	0.20
		0.11	1.40	0.64	0.07	0.30	0.31	0.09	
	Sensitivity	99.91	84.73	94.32	97.85	98.55	94.11	94.77	
		0.06	1.12	0.51	0.13	0.31	0.30		
25	PPV	99.87	73.78	91.71	98.41	96.95	92.81	96.77	0.21
		0.03	1.01	0.37	0.07	0.41	0.56	0.06	
	Sensitivity	99.88	83.48	93.58	97.56	98.52	93.34	94.23	
		0.13	1.26	0.21	0.06	0.15	0.36		
20	PPV	99.70	72.03	89.14	98.05	96.29	91.67	96.12	0.23
		0.17	1.41	0.65	0.04	0.23	0.28	0.08	
	Sensitivity	99.81	80.78	91.45	97.12	97.87	92.60	93.05	
		0.13	0.71	0.32	0.11	0.23	0.32		
15	PPV	99.56	66.68	84.72	97.36	94.12	90.29	94.84	0.28
		0.43	2.34	0.38	0.12	0.84	0.65	0.25	
	Sensitivity	99.65	76.27	87.75	96.18	96.71	91.00	90.91	
		0.25	1.43	0.57	0.19	0.62	0.95		
10	PPV	99.14	55.18	75.66	95.62	85.34	87.30	91.60	0.36
		0.13	0.98	1.12	0.13	0.97	0.83	0.19	
	Sensitivity	99.02	65.82	80.53	93.87	90.12	86.99	85.34	
		0.17	1.20	0.55	0.21	0.52	0.22		
5	PPV	89.94	37.03	60.05	91.11	56.69	78.34	83.05	0.50
		0.85	2.00	0.75	0.16	0.59	0.26	0.22	
	Sensitivity	90.91	47.06	65.41	87.73	65.79	80.22	71.13	
		0.46	1.77	0.79	0.21	0.99	0.83		

Table B-11 Effect of Noise on Pruned KNN without PCA Feature Reduction

Noise (dB)	Effect of Noise Without PCA Feature Reduction (N_{train} = 51600, N_{test} = 51599)								
		PB	APB	LBBB	Normal	RBBB	PVC	Total	Reduc.
No noise	PPV	99.96	78.38	93.47	98.56	97.53	94.16	97.31	0.32
		0.02	2.09	0.17	0.05	0.24	0.17	0.05	0.23
	Sensitivity	99.88	83.32	94.97	98.11	98.69	93.74	94.61	
		0.10	0.90	0.17	0.11	0.11	0.61		
40	PPV	99.94	78.54	93.78	98.53	97.52	93.94	97.30	0.33
		0.04	1.19	0.37	0.03	0.33	0.48	0.07	0.23
	Sensitivity	99.84	83.92	94.35	98.11	98.63	94.21	94.69	
		0.11	0.63	0.23	0.08	0.18	0.35		
35	PPV	99.91	78.64	93.39	98.50	97.62	93.90	97.27	0.33
		0.06	1.52	0.26	0.06	0.21	0.57	0.05	0.22
	Sensitivity	99.86	83.74	94.35	98.11	98.58	93.67	94.56	
		0.18	1.73	0.31	0.03	0.34	0.22		
30	PPV	99.89	78.38	92.81	98.40	97.37	93.41	97.08	0.33
		0.08	0.29	0.45	0.07	0.12	0.32	0.05	0.23
	Sensitivity	99.87	82.43	93.86	98.00	98.56	93.34	94.15	
		0.09	0.43	0.39	0.03	0.24	0.51		
25	PPV	99.91	76.37	92.31	98.23	97.02	92.98	96.80	
		0.03	1.17	0.38	0.04	0.45	0.37	0.07	0.35
	Sensitivity	99.82	81.64	93.00	97.83	98.45	92.48	93.66	0.23
		0.07	1.09	0.25	0.10	0.10	0.44		
20	PPV	99.85	71.68	89.78	97.80	96.12	91.96	96.05	0.39
		0.10	1.49	0.51	0.07	0.22	0.23	0.12	0.24
	Sensitivity	99.80	77.63	90.91	97.28	97.62	91.75	92.18	
		0.10	0.78	0.40	0.16	0.11	0.38		
15	PPV	99.74	69.67	85.23	97.13	94.49	90.76	94.94	0.45
		0.28	1.98	0.63	0.15	0.57	0.59	0.24	0.27
	Sensitivity	99.63	73.74	87.10	96.62	96.34	89.92	90.12	
		0.15	1.13	0.61	0.17	0.30	0.52		
10	PPV	99.08	58.41	75.86	95.37	85.86	86.41	91.61	0.36
		99.18	65.86	78.83	94.11	89.99	86.46	85.00	
5	Sensitivity	91.33	35.66	60.81	90.60	56.21	78.95	83.03	0.50
		90.12	40.87	65.20	88.28	64.25	78.65	68.92	

Table B-12 Effect of Noise on Pruned KNN with PCA Feature Reduction

Noise	Effect of Noise With PCA Feature Reduction (N_{train} = 51600, N_{test} = 51599)								
		PB	APB	LBBB	Normal	RBBB	PVC	Total	Reduc.
No noise	PPV	99.89	78.28	93.90	98.55	97.57	94.08	97.33	0.32
		0.04	0.93	0.59	0.06	0.18	0.22	0.05	0.23
	Sensitivity	99.89	83.61	94.60	98.15	98.64	93.97	94.65	
		0.10	0.53	0.31	0.08	0.21	0.41		
40	PPV	99.91	77.44	93.28	98.49	97.60	93.85	97.20	0.32
		0.03	0.59	0.11	0.09	0.27	0.40	0.09	0.22
	Sensitivity	99.94	83.44	94.45	98.05	98.50	93.51	94.48	
		0.00	0.67	0.41	0.06	0.12	0.33		
35	PPV	99.96	78.58	93.48	98.51	97.42	94.16	97.28	0.32
		0.05	0.89	0.43	0.05	0.23	0.45	0.06	0.22
	Sensitivity	99.87	83.58	94.52	98.11	98.69	93.73	94.58	
		0.05	1.08	0.17	0.05	0.18	0.37		
30	PPV	99.91	77.58	92.86	98.37	97.53	93.29	97.04	0.33
		0.08	0.84	0.54	0.08	0.15	0.37	0.10	0.23
	Sensitivity	99.88	83.21	93.73	97.96	98.40	93.13	94.21	
		0.07	0.88	0.50	0.05	0.33	0.48		
25	PPV	99.91	76.15	91.79	98.27	97.16	92.82	96.78	0.35
		0.10	1.24	0.20	0.05	0.31	0.14	0.07	0.23
	Sensitivity	99.89	82.60	93.03	97.74	98.37	92.71	93.87	
		0.05	0.66	0.22	0.10	0.27	0.50		
20	PPV	99.82	72.18	90.07	97.74	96.43	91.61	96.04	0.39
		0.10	0.53	0.66	0.05	0.19	0.81	0.05	0.25
	Sensitivity	99.81	77.95	90.35	97.34	97.59	91.63	92.14	
		0.11	1.44	0.29	0.11	0.12	0.63		
15	PPV	99.62	68.00	85.35	97.12	94.16	90.26	94.84	0.45
		0.42	2.46	1.42	0.23	0.63	0.82	0.42	0.27
	Sensitivity	99.53	73.35	86.99	96.48	96.24	90.00	89.98	
		0.26	3.52	0.78	0.28	0.64	0.95		
10	PPV	99.44	57.13	76.83	95.25	86.20	87.48	91.77	0.36
	Sensitivity	99.11	60.18	79.48	94.34	90.02	87.38	84.02	
5	PPV	90.40	37.02	59.21	90.64	57.49	79.13	83.03	0.50
	Sensitivity	90.50	42.97	65.35	88.26	64.89	76.81	69.41	

Table B-13 Effect of Noise on Pruned Weighted KNN without PCA Feature Reduction

Noise	Effect of Noise Without PCA Feature Reduction (N_{train} = 51600, N_{test} = 51599)								
		PB	APB	LBBB	Normal	RBBB	PVC	Total	Redc Ratio
No Noise	PPV	99.81	77.68	93.34	98.68	97.47	93.94	97.33	0.19
		0.08	0.42	0.26	0.06	0.45	0.28	0.03	
	Sensitivity	99.81	84.69	95.31	98.01	98.68	94.20	94.98	
		0.05	1.23	0.24	0.04	0.14	0.39		
40	PPV	99.90	78.07	92.90	98.64	97.21	93.31	97.21	0.19
		0.04	1.31	0.15	0.07	0.15	0.25	0.07	
	Sensitivity	99.88	83.98	94.87	97.91	98.77	94.31	94.80	
		0.04	1.51	0.40	0.08	0.21	0.38		
35	PPV	99.83	77.01	92.83	98.66	97.28	93.46	97.21	0.19
		0.04	0.75	0.43	0.02	0.24	0.16	0.03	
	Sensitivity	99.86	84.76	94.81	97.91	98.60	94.13	94.87	
		0.03	1.25	0.29	0.04	0.25	0.29		
30	PPV	99.87	77.28	92.75	98.60	97.34	93.02	97.12	0.20
		0.05	1.48	0.36	0.09	0.11	0.22	0.08	
	Sensitivity	99.88	84.99	94.53	97.83	98.68	94.01	94.85	
		0.12	0.74	0.47	0.07	0.35	0.32		
20	PPV	99.67	71.87	89.24	98.00	96.25	91.66	96.09	0.23
	Sensitivity	99.67	80.14	91.67	97.10	97.66	92.61	92.90	
15	PPV	99.78	67.48	84.46	97.39	94.18	90.38	94.88	0.28
	Sensitivity	99.72	76.37	87.91	96.24	96.51	90.89	90.93	
10	PPV	99.12	55.65	76.64	95.71	85.54	87.13	91.70	0.36
	Sensitivity	99.23	68.00	80.91	93.91	89.98	87.16	85.91	
5	PPV	89.63	39.41	58.92	91.21	56.41	78.00	83.04	0.49
	Sensitivity	91.30	49.65	64.97	87.80	64.81	79.78	71.51	

Table B-14 Effect of Noise on Pruned Weighted KNN with PCA Feature Reduction

Noise (dB)	Effect of Noise With PCA Feature Reduction (N_{train} = 51600, N_{test} = 51599)								
		PB	APB	LBBB	Normal	RBBB	PVC	Total	Ratio
No noise	PPV	99.89	78.77	93.52	98.64	97.45	93.95	97.35	0.19
		0.13	1.28	0.24	0.05	0.29	0.25	0.03	
	Sensitivity	99.86	84.37	95.00	98.08	98.78	94.22	94.90	
		0.12	0.71	0.31	0.06	0.18	0.31		
40	PPV	99.82	78.03	93.39	98.67	97.26	93.53	97.29	0.19
		0.08	0.79	0.52	0.05	0.52	0.41	0.06	
	Sensitivity	99.88	84.40	95.11	97.99	98.82	94.03	94.89	
		0.08	0.79	0.31	0.07	0.12	0.41		
35	PPV	99.80	75.81	93.16	98.63	97.46	93.05	97.15	0.19
		0.11	1.29	0.57	0.06	0.29	0.37	0.05	
	Sensitivity	99.85	84.49	94.59	97.85	98.62	94.27	94.80	
		0.07	0.64	0.35	0.06	0.09	0.26		
30	PPV	99.89	75.74	92.32	98.61	97.25	93.51	97.09	0.20
		0.07	1.60	0.34	0.09	0.25	0.24	0.04	
	Sensitivity	99.90	85.01	94.60	97.81	98.68	93.55	94.79	
		0.08	1.38	0.47	0.08	0.06	0.26		
25	PPV	99.77	74.88	90.89	98.31	96.93	91.85	96.58	0.21
		0.08	1.55	0.39	0.04	0.26	0.35	0.05	
	Sensitivity	99.84	83.57	92.45	97.46	98.32	93.24	93.99	
		0.11	0.89	0.12	0.10	0.12	0.51		
20	PPV	99.73	70.41	88.64	97.93	96.30	90.02	95.83	0.23
	Sensitivity	99.89	81.33	90.81	96.90	97.38	91.51	92.76	
15	PPV	99.72	67.27	81.42	96.92	94.41	90.13	94.27	0.28
	Sensitivity	99.77	77.60	84.18	95.92	95.53	90.41	90.24	
10	PPV	98.88	58.04	73.66	95.62	86.39	85.55	91.43	0.35
	Sensitivity	98.88	66.87	79.55	93.78	89.53	87.20	85.30	
5	PPV	82.87	35.73	58.50	90.72	55.98	76.66	82.32	0.50
	Sensitivity	86.32	46.73	62.98	87.54	61.98	80.09	69.26	

Table B-15 Effect of Noise on Pruned KNN using DROP3 without PCA Feature Reduction

Noise (dB)	Effect of Noise Without PCA Feature Reduction (N_{train} = 51600, N_{test} = 51599)								
		PB	APB	LBBB	Normal	RBBB	PVC	Total	Reduc.
No noise	PPV	99.89	80.82	93.20	98.25	97.75	93.36	97.12	0.03
	Sensitivity	99.84	81.86	94.05	98.23	97.89	99.99	93.77	

Appendix C. Set 9 ALL

Table C-1 Effect of Noise on Pruned KNN without PCA Feature Reduction

Noise	Effect of Noise Without PCA Feature Reduction (N_{train} = 52351, N_{test} = 52350)											
		VF	PB	APB	FV&N	LBBB	Normal	RBBB	PVC	FN&P	Total	Reduc
40	PPV	78.02	99.72	78.95	72.47	93.19	98.38	97.80	92.68	78.51	96.78	0.21
	Sensitivity	80.09	99.49	83.00	74.01	94.83	97.98	98.38	92.25	77.87	88.15	
35	PPV	82.64	99.78	76.34	74.22	93.51	98.38	97.38	92.37	80.00	96.71	0.21
	Sensitivity	82.99	99.67	84.98	71.97	94.82	97.83	98.34	92.11	79.37	88.63	
30	PPV	75.52	99.45	77.35	66.09	92.59	98.17	97.23	91.37	88.68	96.36	0.22
	Sensitivity	81.25	99.61	84.12	71.01	93.70	97.69	98.34	90.61	68.61	86.48	
25	PPV	78.28	99.29	77.28	64.22	92.29	98.00	97.43	91.90	77.23	96.22	0.99
	Sensitivity	78.28	99.78	81.19	72.46	93.25	97.60	98.36	91.15	66.67	85.72	
20	PPV	76.39	99.62	74.48	64.47	88.89	97.53	96.13	90.38	78.76	95.39	0.26
	Sensitivity	79.11	99.19	78.04	64.81	91.06	97.09	97.59	89.39	72.36	84.58	
15	PPV	74.49	98.41	66.91	49.53	84.30	96.62	92.71	87.73	58.59	93.53	0.31
	Sensitivity	75.10	99.28	71.20	57.73	85.22	95.86	96.71	86.77	46.03	77.14	
10	PPV	63.71	99.04	55.74	43.13	75.97	94.73	86.23	84.27	46.73	90.36	0.38
	Sensitivity	65.83	98.54	63.28	46.25	77.80	93.81	89.79	82.77	38.46	69.86	
5	PPV	58.55	89.84	35.49	20.09	59.59	89.95	55.96	74.70	18.46	81.42	0.52
	Sensitivity	55.92	90.38	42.46	22.47	63.24	87.73	63.36	72.77	19.67	51.22	

Table C-2 Effect of Noise on Pruned KNN with PCA Feature Reduction

Noise	Effect of Noise With PCA Feature Reduction (N_{train} = 52351, N_{test} = 52350)											
		VF	PB	APB	FV&N	LBBB	Normal	RBBB	PVC	FN&P	Total	Reduc.
No Noise	PPV	76.47	99.34	77.94	71.67	93.84	98.17	97.18	91.73	76.92	96.51	0.21
	SEN	82.28	99.18	82.44	73.67	93.98	97.84	98.74	91.31	67.16	86.67	
40	PPV	80.83	99.28	78.23	69.89	92.94	98.19	97.51	91.36	77.78	96.45	0.21
	SEN	81.86	99.39	82.37	75.25	93.98	97.75	98.63	91.13	69.42	87.12	
35	PPV	78.46	99.37	77.94	69.95	93.07	98.25	97.03	91.52	71.90	96.45	0.22
	SEN	81.78	99.32	83.97	76.61	93.76	97.72	98.20	91.07	72.50	87.81	
30	PPV	78.48	99.17	77.81	67.90	91.80	98.01	97.26	91.43	78.15	96.20	0.22
	SEN	78.48	99.33	82.78	67.24	92.81	97.66	98.38	91.12	66.91	85.17	
25	PPV	79.06	98.86	75.57	71.29	91.56	97.94	96.68	89.85	66.94	95.89	0.23
	SEN	79.40	99.13	80.87	74.13	92.19	97.39	97.99	90.67	60.90	84.90	
20	PPV	75.95	98.66	73.53	57.31	88.99	97.36	95.39	89.01	68.57	94.94	0.25
	SEN	73.47	99.10	77.82	68.21	89.66	96.83	97.98	88.15	52.55	81.13	
15	PPV	71.31	97.97	65.52	53.35	83.51	96.19	92.41	85.75	40.52	92.96	0.30
	SEN	74.90	97.86	68.90	52.65	83.62	95.76	95.28	85.31	39.50	74.22	
10	PPV	62.50	98.50	55.47	39.80	73.87	94.53	85.43	83.78	42.50	89.94	0.38
	SEN	65.61	97.53	62.33	42.62	75.95	93.50	89.78	82.41	40.48	69.15	
5	PPV	48.98	82.89	35.03	18.26	61.19	89.53	52.95	72.54	12.78	80.53	0.52
	SEN	52.63	83.89	41.47	22.11	63.05	87.32	60.00	72.02	13.08	47.62	

Table C-3 Effect of Noise on FKNN without PCA Feature Reduction

Noise		Effect of Noise Without PCA Feature Reduction (N_{train} = 52351, N_{test} = 52350)									
		VF	PB	APB	FV&N	LBBB	Normal	RBBB	PVC	FN&P	Total
No Noise	PPV	80.48	99.66	83.54	77.68	94.46	98.27	98.16	92.60	88.07	97.04
		1.94	0.07	0.82	1.72	0.26	0.06	0.19	0.36	3.79	0.06
	Sensitivity	80.87	99.74	82.07	74.19	94.43	98.38	98.44	92.78	75.20	87.90
		3.14	0.13	1.33	2.37	0.13	0.05	0.09	0.26	1.83	
40	PPV	80.04	99.59	82.84	76.46	94.06	98.30	98.08	93.14	83.80	97.02
		2.87	0.15	1.36	1.51	0.37	0.08	0.07	0.58	2.80	0.02
	Sensitivity	82.18	99.77	83.55	73.79	94.49	98.33	98.47	92.26	72.87	87.83
		1.87	0.10	0.98	2.78	0.42	0.09	0.12	0.22	4.44	
35	PPV	77.42	99.61	82.72	77.82	94.21	98.24	98.09	93.08	83.78	96.98
		2.04	0.09	1.36	1.62	0.44	0.04	0.24	0.26	3.06	0.03
	Sensitivity	83.88	99.71	82.95	73.61	94.08	98.33	98.43	92.33	74.51	88.11
		1.70	0.05	0.86	2.99	0.31	0.06	0.12	0.25	2.17	
30	PPV	79.81	99.49	82.16	73.75	94.35	98.16	97.84	92.46	86.25	96.84
		2.30	0.12	0.70	2.89	0.40	0.10	0.12	0.26	2.37	0.02
	Sensitivity	82.48	99.71	82.09	72.98	93.61	98.24	98.38	92.21	74.22	87.64
		2.26	0.10	0.97	2.26	0.43	0.09	0.05	0.69	1.75	
25	PPV	79.46	99.41	80.45	73.72	92.70	98.00	97.69	91.96	80.80	96.49
		2.83	0.15	2.01	2.74	0.54	0.06	0.27	0.53	3.45	0.06
	Sensitivity	80.69	99.69	81.74	70.29	92.95	98.01	98.25	91.29	66.69	85.80
		1.67	0.03	1.51	0.58	0.31	0.10	0.19	0.55	2.34	
20	PPV	77.92	99.29	77.84	66.65	90.72	97.49	96.95	90.69	79.52	95.71
		2.17	0.14	1.23	2.75	0.44	0.12	0.36	0.56	4.52	0.06
	Sensitivity	76.83	99.59	77.73	68.49	90.30	97.62	97.56	89.85	58.80	82.91
		1.63	0.14	0.88	1.36	0.65	0.04	0.26	0.61	2.81	
15	PPV	75.67	99.29	74.09	62.14	87.62	96.75	95.07	89.07	69.90	94.60
		3.09	0.20	2.08	3.39	0.78	0.21	0.65	0.54	5.64	0.29
	Sensitivity	75.42	99.34	72.05	60.82	86.64	97.00	96.34	88.40	51.58	79.10
		1.82	0.26	2.43	2.85	0.60	0.20	0.47	0.79	5.20	
10	PPV	66.24	98.76	63.47	47.30	78.31	94.68	88.78	85.03	54.22	91.29
		2.32	0.33	0.99	3.53	0.89	0.15	0.71	0.65	4.19	0.07
	Sensitivity	65.41	99.06	60.32	43.58	78.43	95.09	89.33	83.70	35.26	68.57
		3.67	0.19	1.76	3.10	1.01	0.13	0.72	0.49	5.90	
5	PPV	56.65	91.04	41.48	24.35	63.32	90.05	63.34	77.15	16.75	83.44
		3.50	0.54	0.98	1.80	0.66	0.12	0.70	0.90	2.42	0.14
	Sensitivity	54.62	90.99	41.18	23.76	63.56	90.39	63.74	74.43	13.23	49.40
		5.00	0.43	1.41	1.74	0.67	0.14	1.10	0.86	2.02	

Table C-4 Effect of Noise on FKNN with PCA Feature Reduction

Noise	Effect of Noise With PCA Feature Reduction (N_{train} = 52351, N_{test} = 52350)										
		VF	PB	APB	FV&N	LBBB	Normal	RBBB	PVC	FN&P	Total
No Noise	PPV	79.51	99.43	82.51	77.08	93.95	98.15	98.00	92.30	78.48	96.82
		2.52	0.23	0.14	1.48	0.60	0.04	0.25	0.44	3.51	0.05
	Sensitivity	79.75	99.36	82.34	73.66	93.78	98.25	98.43	92.03	70.11	86.85
		1.89	0.19	0.70	1.47	0.17	0.03	0.21	0.50	4.25	
40	PPV	80.23	99.22	82.07	74.19	93.68	98.16	97.68	92.57	79.24	96.76
		1.61	0.12	1.16	1.65	0.45	0.02	0.18	0.47	2.32	0.03
	Sensitivity	78.98	99.32	82.63	74.30	93.65	98.19	98.69	91.73	66.13	86.28
		2.84	0.23	0.78	2.31	0.35	0.02	0.25	0.65	2.72	
35	PPV	80.34	99.31	82.23	75.25	93.65	98.18	97.94	92.09	76.67	96.77
		1.89	0.12	0.93	3.47	0.09	0.02	0.16	0.22	4.60	0.04
	Sensitivity	79.92	99.34	82.60	74.62	93.33	98.19	98.50	92.16	68.60	86.79
		3.52	0.24	0.61	2.09	0.45	0.07	0.13	0.29	4.68	
30	PPV	80.41	99.20	81.54	74.01	93.49	98.04	97.87	91.73	79.03	96.60
		2.07	0.10	1.22	0.72	0.29	0.07	0.30	0.64	4.78	0.05
	Sensitivity	80.04	99.39	82.73	72.01	92.90	98.10	98.31	91.70	63.93	85.68
		0.68	0.07	1.16	1.50	0.16	0.10	0.15	0.53	3.84	
25	PPV	77.07	99.03	81.94	73.56	92.12	97.83	97.54	91.51	76.93	96.32
		3.77	0.31	1.32	2.71	0.51	0.11	0.30	0.46	2.51	0.06
	Sensitivity	77.33	99.35	80.77	69.95	91.99	97.98	98.30	91.03	59.11	83.94
		1.92	0.21	1.38	2.30	0.56	0.13	0.15	0.32	5.68	
20	PPV	74.79	98.72	76.64	67.54	90.08	97.32	96.77	89.20	69.11	95.37
		4.60	0.09	1.02	2.78	0.31	0.03	0.27	0.58	5.93	0.04
	Sensitivity	76.41	99.23	76.67	64.50	89.57	97.37	97.41	89.30	55.17	81.40
		1.32	0.21	0.73	2.65	0.29	0.06	0.24	0.36	3.85	
15	PPV	73.60	98.53	71.71	58.94	85.62	96.46	94.34	87.66	61.85	93.95
		3.59	0.82	2.82	1.07	1.58	0.23	0.90	0.77	12.05	0.35
	Sensitivity	72.56	98.74	70.66	56.86	84.78	96.63	95.56	87.43	45.48	76.49
		1.38	0.81	2.33	3.07	1.21	0.15	0.44	1.00	5.19	
10	PPV	63.53	98.12	63.15	47.20	77.90	94.59	88.89	84.45	47.15	91.07
		2.06	0.40	1.20	2.47	0.44	0.08	0.09	0.42	5.27	0.12
	Sensitivity	62.15	98.33	61.13	46.20	77.63	94.98	89.01	82.98	35.38	68.51
		3.14	0.29	1.45	1.97	0.35	0.10	0.30	0.70	1.98	
5	PPV	57.28	83.98	42.37	20.54	61.68	89.75	61.30	75.38	14.15	82.58
		4.42	0.90	1.33	1.38	0.64	0.15	0.61	0.38	2.35	0.18
	Sensitivity	52.88	84.04	40.43	21.03	63.07	90.03	60.35	73.99	12.04	47.16
		2.28	0.96	1.76	2.16	0.82	0.24	0.88	0.54	3.27	

Table C-5 Effect of Noise on Pruned Weighted KNN without PCA Feature Reduction

Noise	Effect of Noise Without PCA Feature Reduction (N_{train} = 52351, N_{test} = 52350)							
		PB	APB	LBBB	Normal	RBBB	PVC	Total
No Noise	PPV	99.76	77.34	93.01	98.73	97.33	93.84	97.30
		0.06	1.36	0.41	0.06	0.40	0.36	0.04
	Sensitivity	99.92	84.97	95.36	97.94	98.76	94.29	95.07
		0.05	0.30	0.15	0.06	0.25	0.43	
40	PPV	99.76	77.12	93.05	98.69	97.49	93.53	97.26
		0.12	1.44	0.18	0.08	0.21	0.56	0.10
	Sensitivity	99.91	85.16	94.71	97.92	98.90	94.50	95.05
		0.09	1.31	0.31	0.07	0.11	0.32	
35	PPV	99.88	77.90	93.24	98.68	97.52	93.42	97.30
		0.06	1.22	0.52	0.07	0.14	0.31	0.07
	Sensitivity	99.89	85.25	94.88	97.98	98.69	94.33	95.04
		0.07	0.63	0.44	0.08	0.29	0.30	
30	PPV	99.77	76.26	92.58	98.63	97.16	93.17	97.10
		0.13	1.08	0.36	0.06	0.35	0.29	0.08
	Sensitivity	99.94	84.51	94.60	97.80	98.64	94.01	94.77
		0.04	0.63	0.27	0.06	0.23	0.70	
25	PPV	99.80	73.94	91.61	98.40	97.27	92.51	96.75
		0.09	0.68	0.40	0.03	0.39	0.20	0.06
	Sensitivity	99.87	83.50	93.20	97.57	98.51	93.45	94.18
		0.14	0.89	0.31	0.09	0.23	0.29	
20	PPV	99.87	70.91	88.95	98.02	95.91	91.37	96.00
		0.10	1.26	0.61	0.01	0.14	0.27	0.08
	Sensitivity	99.79	80.80	91.34	96.99	98.01	92.20	92.96
		0.10	1.41	0.49	0.13	0.26	0.22	
15	PPV	99.83	67.43	85.05	97.46	94.29	90.04	94.96
		0.09	1.90	0.52	0.10	0.45	0.32	0.17
	Sensitivity	99.61	78.03	87.88	96.26	96.65	91.12	91.30
		0.08	0.76	0.54	0.16	0.40	0.27	
10	PPV	99.20	56.18	74.88	95.73	86.09	86.88	91.64
		0.33	1.01	0.44	0.13	1.20	0.16	0.12
	Sensitivity	99.14	67.15	80.53	93.81	90.30	87.49	85.74
		0.15	0.38	0.77	0.15	0.50	0.62	
5	PPV	90.23	36.82	59.82	91.18	56.90	78.63	83.14
		1.11	0.73	0.33	0.17	0.51	0.30	0.12
	Sensitivity	90.59	47.23	65.72	87.85	65.50	80.06	71.13
		1.05	0.97	1.12	0.16	0.71	0.61	

Table C-6 Effect of Noise on Pruned Weighted KNN with PCA Feature Reduction

Noise		Effect of Noise With PCA Feature Reduction (N_{train} = 52351, N_{test} = 52350)						
		PB	APB	LBBB	Normal	RBBB	PVC	Total
No Noise	PPV	99.85	77.95	93.41	98.70	97.46	93.73	97.35
		0.07	0.84	0.49	0.05	0.17	0.33	0.04
	Sensitivity	99.90	84.89	95.07	98.03	98.64	94.48	95.03
		0.04	0.50	0.33	0.09	0.29	0.18	
40	PPV	99.83	77.73	93.39	98.66	97.46	94.03	97.33
		0.08	0.72	0.36	0.07	0.27	0.51	0.05
	Sensitivity	99.87	84.12	95.09	98.05	98.69	94.29	94.86
		0.12	0.49	0.20	0.06	0.20	0.72	
35	PPV	99.84	77.08	93.03	98.66	97.28	93.21	97.21
		0.05	0.94	0.35	0.07	0.18	0.32	0.10
	Sensitivity	99.87	84.97	94.60	97.88	98.85	94.26	94.94
		0.06	1.03	0.28	0.06	0.13	0.48	
30	PPV	99.85	76.84	92.74	98.60	97.39	92.95	97.13
		0.11	1.40	0.64	0.07	0.30	0.31	0.09
	Sensitivity	99.91	84.73	94.32	97.85	98.55	94.11	94.77
		0.06	1.12	0.51	0.13	0.31	0.30	
25	PPV	99.87	73.78	91.71	98.41	96.95	92.81	96.77
		0.03	1.01	0.37	0.07	0.41	0.56	0.06
	Sensitivity	99.88	83.48	93.58	97.56	98.52	93.34	94.23
		0.13	1.26	0.21	0.06	0.15	0.36	
20	PPV	99.70	72.03	89.14	98.05	96.29	91.67	96.12
		0.17	1.41	0.65	0.04	0.23	0.28	0.08
	Sensitivity	99.81	80.78	91.45	97.12	97.87	92.60	93.05
		0.13	0.71	0.32	0.11	0.23	0.32	
15	PPV	99.56	66.68	84.72	97.36	94.12	90.29	94.84
		0.43	2.34	0.38	0.12	0.84	0.65	0.25
	Sensitivity	99.65	76.27	87.75	96.18	96.71	91.00	90.91
		0.25	1.43	0.57	0.19	0.62	0.95	
10	PPV	99.14	55.18	75.66	95.62	85.34	87.30	91.60
		0.13	0.98	1.12	0.13	0.97	0.83	0.19
	Sensitivity	99.02	65.82	80.53	93.87	90.12	86.99	85.34
		0.17	1.20	0.55	0.21	0.52	0.22	
5	PPV	89.94	37.03	60.05	91.11	56.69	78.34	83.05
		0.85	2.00	0.75	0.16	0.59	0.26	0.22
	Sensitivity	90.91	47.06	65.41	87.73	65.79	80.22	71.13
		0.46	1.77	0.79	0.21	0.99	0.83	

Table C-7 Effect of Noise on Pruned Weighted FKNN without PCA Feature Reduction

Noise	Effect of Noise Without PCA Feature Reduction (N_{train} = 52351, N_{test} = 52350)											
		VF	PB	APB	FV&N	LBBB	Normal	RBBB	PVC	FN&P	Total	Reduc.
No Noise	PPV	68.87	99.55	79.77	74.77	92.48	98.52	97.27	93.07	80.47	96.74	0.21
	Sensitivity	90.04	99.49	84.34	76.56	94.79	97.84	98.38	92.34	76.30	89.59	
40	PPV	71.84	99.62	76.61	67.94	93.73	98.59	97.57	92.58	87.76	96.77	0.21
	Sensitivity	85.78	99.84	85.43	82.79	95.24	97.73	98.68	91.85	71.07	89.35	
35	PPV	72.32	99.72	76.54	69.35	93.10	98.53	97.32	92.36	78.86	96.60	1.01
	Sensitivity	87.08	99.50	85.11	78.96	94.90	97.55	98.78	92.23	78.23	89.92	
30	PPV	70.74	99.56	75.02	70.12	92.69	98.48	97.54	91.96	80.33	96.51	0.22
	Sensitivity	84.51	99.72	86.14	78.44	94.56	97.48	98.35	91.90	76.56	89.36	
25	PPV	70.93	99.52	74.59	63.51	91.66	98.33	97.01	90.75	83.62	96.15	0.23
	Sensitivity	88.36	99.73	83.13	74.44	93.74	97.27	98.18	91.22	74.05	88.40	
20	PPV	67.71	99.35	70.47	56.98	88.89	97.87	96.31	90.33	69.52	95.26	0.26
	Sensitivity	90.76	99.62	80.78	70.99	90.85	96.64	97.58	89.93	64.04	85.94	
15	PPV	68.20	99.45	71.01	49.89	84.68	97.38	94.32	88.28	61.74	94.27	0.30
	Sensitivity	75.74	99.18	78.04	63.19	88.24	95.98	96.97	88.96	59.66	81.66	
10	PPV	63.08	98.40	56.89	39.21	75.11	95.23	86.23	83.77	27.64	90.20	0.38
	Sensitivity	73.33	98.51	68.29	48.89	79.69	93.12	89.86	83.11	39.86	72.19	
5	PPV	48.37	89.70	37.40	16.24	61.53	91.13	56.23	73.45	9.40	81.20	0.52
	Sensitivity	64.35	89.10	48.09	26.03	65.32	86.44	66.48	74.65	35.90	57.82	

Table C-8 Effect of Noise on Pruned Weighted FKNN with PCA Feature Reduction

Noise	Effect of Noise Without PCA Feature Reduction (N_{train} = 52351, N_{test} = 52350)										
		VF	PB	APB	FV&N	LBBB	Normal	RBBB	PVC	FN&P	Total	Reduc.
No Noise	PPV	71.58	99.67	78.02	69.61	93.27	98.62	97.51	93.20	81.15	96.83	0.21
	Sensitivity	88.70	99.73	85.96	76.56	94.96	97.81	98.61	92.81	73.88	89.42	
40	PPV	73.09	99.38	76.34	72.67	93.44	98.58	97.45	92.38	78.69	96.72	1.00
	Sensitivity	90.16	99.44	84.13	78.80	94.63	97.75	98.59	92.56	75.59	89.79	
35	PPV	72.01	99.51	76.06	69.05	92.82	98.58	97.41	92.10	79.31	96.63	0.22
	Sensitivity	88.13	99.40	85.02	78.68	94.97	97.56	98.69	92.47	72.44	89.24	
30	PPV	73.81	99.46	75.98	70.09	93.42	98.18	97.51	91.61	82.76	96.38	0.21
	Sensitivity	88.57	99.73	82.55	75.57	93.42	97.64	98.37	91.19	78.69	89.12	
25	PPV	68.38	99.21	75.02	66.17	92.32	98.19	97.22	90.82	73.11	96.11	0.23
	Sensitivity	86.92	99.33	82.55	70.74	93.33	97.33	98.29	91.90	65.41	86.48	
20	PPV	68.11	99.27	72.04	57.24	89.26	97.97	96.21	89.98	72.73	95.38	0.26
	Sensitivity	86.86	99.27	81.40	70.92	91.50	96.73	97.98	90.21	61.54	85.32	
15	PPV	65.14	99.77	66.88	51.62	85.63	97.06	94.05	88.41	68.97	93.99	0.30
	Sensitivity	82.22	99.49	77.27	58.87	86.59	95.96	97.01	88.05	56.74	80.98	
10	PPV	59.79	98.70	54.60	41.29	75.80	95.31	85.92	84.92	31.48	90.37	0.38
	Sensitivity	75.00	98.81	66.61	50.00	79.51	93.27	90.17	83.93	41.80	72.87	
5	PPV	47.56	89.86	36.34	19.81	59.22	90.77	56.56	74.73	7.17	80.91	0.52
	Sensitivity	61.09	89.66	46.68	27.58	64.39	86.35	64.63	76.03	24.43	55.13	

Table C-9 Effect of Random Data Reduction on FKNN without PCA

N_{train}	Effect of Data Reduction Without PCA Feature Reduction (N_{test} = 104701 - N_{train})									
		VF	PB	APB	FV&N	LBBB	Normal	RBBB	PVC	FN&P	Total
52351	PPV	80.48	99.66	83.54	77.68	94.46	98.27	98.16	92.60	88.07	97.04
		1.94	0.07	0.82	1.72	0.26	0.06	0.19	0.36	3.79	0.06
	Sensitivity	80.87	99.74	82.07	74.19	94.43	98.38	98.44	92.78	75.20	87.90
		3.14	0.13	1.33	2.37	0.13	0.05	0.09	0.26	1.83	
26176	PPV	77.80	99.61	82.37	76.39	93.46	98.06	97.85	92.15	82.61	96.71
		2.57	0.15	1.24	0.32	0.36	0.05	0.09	0.33	3.09	0.03
	Sensitivity	79.32	99.63	81.31	69.93	93.50	98.23	98.32	91.69	68.30	85.88
		2.03	0.11	1.18	1.29	0.35	0.03	0.09	0.20	4.12	
13088	PPV	73.55	99.43	79.39	71.16	92.51	97.92	97.36	91.58	84.42	96.31
		2.37	0.15	0.83	2.87	0.51	0.07	0.40	0.30	4.65	0.05
	Sensitivity	76.64	99.68	80.79	70.62	92.37	97.96	98.36	90.38	62.66	84.50
		2.61	0.10	1.03	2.21	0.18	0.09	0.14	0.73	3.91	
6544	PPV	70.14	99.09	79.23	67.57	90.81	97.50	97.17	90.63	79.62	95.75
		7.17	0.30	2.47	4.91	1.06	0.10	0.31	0.88	6.53	0.10
	Sensitivity	72.11	99.60	77.57	67.24	90.39	97.80	98.20	88.39	52.87	81.11
		8.93	0.21	1.24	6.25	0.63	0.24	0.26	0.85	4.98	
3272	PPV	67.31	98.81	76.67	65.57	88.93	97.09	96.09	89.34	81.49	95.08
		6.16	0.48	2.70	2.02	1.01	0.12	0.53	1.22	8.27	0.14
	Sensitivity	63.46	99.12	74.16	64.08	88.69	97.46	97.69	87.23	44.77	77.30
		7.31	0.40	1.69	6.74	0.74	0.31	0.23	1.00	10.22	
1636	PPV	60.72	98.36	71.00	57.88	85.30	96.66	95.32	85.92	81.18	93.99
		6.65	0.22	2.40	7.82	0.93	0.26	0.52	2.18	8.67	0.09
	Sensitivity	50.44	99.31	72.85	50.09	86.81	96.78	96.98	84.23	35.05	70.67
		10.98	0.30	3.74	8.67	0.73	0.26	0.49	2.42	9.92	

Table C-10 Effect of Random Data Reduction on FKNN with PCA

N train	Effect of Data Reduction With PCA Feature Reduction (N_{test} = 104701 - N_{train})									
		VF	PB	APB	FV&N	LBBB	Normal	RBBB	PVC	FN&P	Total
52351	PPV	79.51	99.43	82.51	77.08	93.95	98.15	98.00	92.30	78.48	96.82
		2.52	0.23	0.14	1.48	0.60	0.04	0.25	0.44	3.51	0.05
	Sensitivity	79.75	99.36	82.34	73.66	93.78	98.25	98.43	92.03	70.11	86.85
		1.89	0.19	0.70	1.47	0.17	0.03	0.21	0.50	4.25	
26176	PPV	76.21	99.09	81.02	76.26	92.84	97.93	97.60	91.91	79.95	96.47
		2.67	0.23	1.22	2.57	0.23	0.03	0.29	0.26	6.22	0.05
	Sensitivity	78.99	99.44	81.00	71.02	92.40	98.08	98.27	91.37	60.73	84.66
		2.12	0.36	0.77	1.38	0.42	0.03	0.21	0.50	4.20	
13088	PPV	72.52	99.09	80.48	70.14	91.94	97.72	97.07	90.94	74.65	96.06
		2.22	0.11	1.61	1.83	0.50	0.11	0.12	0.56	6.49	0.10
	Sensitivity	74.40	99.36	79.34	68.72	91.09	97.88	98.21	90.36	57.18	82.77
		3.81	0.38	1.13	3.67	0.77	0.08	0.22	0.71	5.36	
6544	PPV	69.97	98.86	78.48	68.39	90.12	97.41	96.91	89.72	80.49	95.53
		2.37	0.24	1.96	7.24	0.60	0.10	0.44	0.38	2.90	0.11
	Sensitivity	68.31	99.32	78.97	67.29	89.31	97.61	97.82	88.98	46.00	79.43
		4.25	0.40	0.89	5.11	0.75	0.23	0.45	0.76	7.54	
3272	PPV	64.04	98.25	75.94	65.13	88.73	97.05	95.32	88.41	80.31	94.86
		6.75	0.47	1.51	4.42	0.93	0.21	0.51	0.60	11.19	0.07
	Sensitivity	65.28	98.97	75.37	60.50	87.60	97.31	97.50	87.10	39.31	75.92
		4.82	0.94	2.95	6.76	1.38	0.19	0.39	0.86	5.70	
1636	PPV	57.34	98.65	76.21	54.29	86.45	96.53	94.07	85.74	53.08	94.00
		4.56	0.48	1.49	9.63	1.91	0.31	0.76	0.81	31.46	0.25
	Sensitivity	56.38	98.98	70.41	44.50	85.89	96.95	96.22	85.92	28.65	68.79
		7.28	0.71	3.68	6.32	1.49	0.23	1.28	1.75	17.41	

Appendix D. Set 6 G

Table D-1 Leave One Patient Out with Pruned Weighted FKNN without PCA

Noise		Patient Leave One Out – without PCA Feature Reduction							
		PB	APB	LBBB	Normal	RBBB	PVC	Total	Reduction
No noise	PPV	26.96	82.20	95.72	99.33	98.27	96.51	89.18	0.19
	Sensitivity	100	58.48	76.22	92.51	81.51	81.96	80.66	

Table D-2 Leave One Patient Out with Pruned Weighted FKNN with PCA

Noise		Patient Leave One Out – with PCA Feature Reduction							
		PB	APB	LBBB	Normal	RBBB	PVC	Total	Reduc.
No noise	PPV	80.49	95.31	99.27	98.26	96.12	89.06	80.49	0.19
	Sensitivity	58.72	75.54	92.40	81.70	81.79	80.59	58.72	