EARLY DIAGNOSIS OF ACUTE CORONARY SYNDROMES AUTOMATICALLY BY USING FEATURES OF ECG RECORDINGS

A THESIS

SUBMITTED TO THE DEPARTMENT OF ELECTRICAL AND

ELECTRONICS ENGINEERING

AND THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE

OF BILKENT UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF SCIENCE

By

Merve Begüm Terzi

September, 2014

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Prof. Dr. Orhan Arıkan(Advisor)

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Assoc. Prof. Dr. Sinan Gezici

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Assoc. Prof. Dr. Selim Aksoy

Approved for the Graduate School of Engineering and Science:

Prof. Dr. Levent Onural
Director of the Graduate School

ABSTRACT

EARLY DIAGNOSIS OF ACUTE CORONARY SYNDROMES AUTOMATICALLY BY USING FEATURES OF ECG RECORDINGS

Merve Begüm Terzi
M.S. in Electrical and Electronics Engineering
Supervisor: Prof. Dr. Orhan Arıkan
September, 2014

In patients with acute coronary syndrome, temporary chest pains together with changes in the ST/T sections of ECG occur shortly before the start of myocardial infarction. In order to diagnose acute coronary syndromes early, we propose a new technique which detects changes in the ST/T sections of ECG. For this purpose, by using real ECG recordings, we identify ECG features that are critical in the detection of acute coronary syndromes. By using support vector machines (SVM) operating with linear and radial basis function (RBF) kernels, we obtain classifiers that use 2 or 3 most discriminating features of the ST/T sections. To improve performance, classification results on multiple segments are fused. The obtained results over a considerable number of patients indicate that the proposed classification technique provides highly reliable detection of acute coronary syndromes. To develop a detection technique that can be used in the absence of unhealthy ECGs, we also investigate the detection of acute coronary syndromes based on ECG recordings of a patient obtained during healthy stage only. For this purpose, a Gaussian mixture model is used to represent the joint pdf of the selected features. Then, a Neyman-Pearson type of approach is developed to provide detection of outliers that would correspond to acute coronary syndromes.

Keywords: Electrocardiogram (ECG) signal classification, feature extraction, support vector machine (SVM), kernel method, acute coronary syndrome, acute myocardial infarction.

ÖZET

EKG KAYITLARININ ÖZNİTELİKLERİ KULLANILARAK AKUT KORONER SENDROMLARIN OTOMATİK OLARAK ERKEN TEŞHİSİ

Merve Begüm Terzi
Elektrik ve Elektronik Mühendisliği, Yüksek Lisans
Tez Yöneticisi: Prof. Dr. Orhan Arıkan
Eylül, 2014

Akut koroner sendrom hastalarında, miyokard infarktüsü başlangıcından kısa bir süre önce, geçici göğüs ağrıları ile birlikte EKG'nin ST/T kısmında değişiklikler meydana gelmektedir. Akut koroner sendromları erken teşhis etmek amacıyla, EKG sinyalinin ST/T kısmındaki değişiklikleri saptayan yeni bir teknik öneriyoruz. Bu amaçla, gerçek EKG kayıtları kullanarak, akut koroner sendrom teşhisi için kritik olan EKG özniteliklerini belirledik. Lineer ve radyal tabanlı çekirdek fonksiyonuna sahip destek vektör makineleri (DVM) kullanarak, ST/T kısmının ayırıcılığı en yüksek 2 veya 3 özniteliğini kullanan sınıflandırıcılar elde ettik. Performansı arttırmak amacıyla, çoklu segmentlerin sınıflandırma sonuçlarını birleştirdik. Kayda değer sayıda hastadan elde edilen sonuçlar, önerilen sınıflandırma tekniğinin oldukça güvenilir akut koroner sendrom tespiti sağladığını göstermektedir. Ayrıca, sağlıksız EKG yokluğunda kullanılabilecek bir tespit tekniği geliştirmek için, bir hastanın sağlıklı durumdayken elde edilen EKG kayıtlarına dayanarak akut koroner sendrom tespitini araştırdık. Bu amaçla, seçilen özniteliklerin ortak olasılık yoğunluk fonksiyonunu göstermek için Gauss karışım modeli kullanılmıştır. Ayrıca, akut koroner sendromları temsil eden aykırı değerlerin tespitini sağlamak amacıyla Neyman-Pearson tipi yaklaşım geliştirilmiştir.

Anahtar sözcükler: Elektrokardiyogram (EKG) sinyal sınıflandırma, öznitelik çıkarımı, destek vektör makinesi (DVM), kernel metodu, akut koroner sendrom, akut miyokard infarktüsü.

Acknowledgement

First and foremost, I would like to thank to my family for giving me a great deal of support in every respect throughout my life. Having said that, I would also like to express my deepest gratitude to:

- My thesis advisor Prof. Dr. Orhan Arıkan and my thesis committee: Assoc. Prof. Sinan Gezici and Assoc. Prof. Selim Aksoy.
- Prof. Dr. Adnan Abacı for his leadership in the TÜBITAK project, his expert guidance on ST/T analysis and invaluable effort in the real ECG recordings of the patients.
- Mehmet Dedeoğlu for his valuable assistance and contributions to the feature detection part of this thesis.
- TÜBİTAK for the scholarship. This thesis is funded by TÜBİTAK as a part of The Support Program for Scientific and Technological Research Projects (1001).
- Dr. Mehmet Köseoğlu, Dr. Elif Aydoğdu and Ismail Uyanık for guiding me about different aspects of academic degree.
- Ömer Faruk Oran, Necip Gürler, Fatih Hafalır, Rıdvan Cantaş and Haşim Meriç for giving me support and guidance when I was a new member of Bilkent EE department.
- My fellow graduate student friends: Dr. Kağan Oğuz, Dr. Aslı Ünlügedik, Gökçe Kuralay, Cemre Arıyürek, Maryam Salim, Burak Şahinbaş, Caner Odabaş and Yasin Kumru.
- To past and present members of signal processing group: Dr. Kıvanç Köse, Alican Bozkurt, Ali Alp Akyol, Onur Yorulmaz, Ahmet Yazar, Oğuzhan Teke, Mohammad Tofighi, Serkan Sarıtaş, Semih Çaycı.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Any set of clinical symptoms consistent with acute myocardial ischemia containing clinical cases such as non-ST segment elevation myocardial infarction (NSTEMI), ST segment elevation myocardial infarction (STEMI) and unstable angina (UA) are called as acute coronary syndrome (ACS) [1].

According to the World Health Organization (WHO), ischaemic heart disease is the leading cause of death for both men and women worldwide. Each year, only in the United States, approximately 1.36 million hospitalizations are required for ACS, of which 0.81 million are for myocardial infarction (MI) and the remainder are for UA [2].

The symptoms of UA/NSTEMI and STEMI are similar and differentiating the two requires medical evaluation and 12-lead electrocardiography (ECG). The guidelines for managing UA/NSTEMI, released by the American College of Cardiology (ACC) and the American Heart Association (AHA), state that patients with symptoms suggestive of ACS should be referred to a facility that has capabilities for 12-lead ECG recording and evaluation by a physician [3]. The literature shows that monitoring of heart electrical activity provides early recognition of symptoms of myocardial infarction and can reduce the morbidity and mortality of patients [4].

1.1 Anatomy of the Heart

Heart is a muscular organ that supplies blood through the body, which is located between lungs in the left side of the sternum. It consists of four chambers which are right atrium, right ventricle, left atrium and left ventricle. Figure 1.1 illustrates the structure diagram of a human heart.

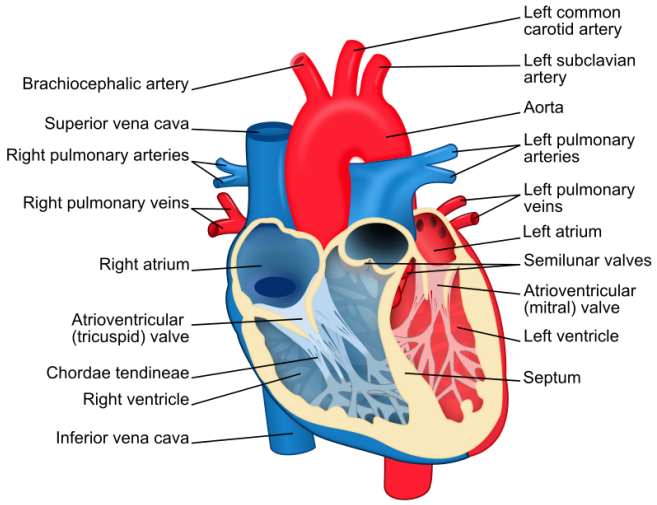


Figure 1.1: Structure diagram of the human heart from an anterior view. Red components indicate oxygenated pathways and blue components indicate deoxygenated blood pathways [5].

The components of the cardiovascular system that are composed of different chambers, valves, arteries and veins are defined as below.

Right Atrium: The deoxygenated blood returning from the body enters to the right atrium by superior vena cava and is then passed on to the right ventricle through the tricuspid valve.

Right Ventricle: It consists of deoxygenated blood which passes through the pulmonary valve and then through lungs for oxygenation.

Pulmonary Valve: It is a one way valve which controls the flow of blood from right ventricle to pulmonary arteries.

Pulmonary Arteries: Deoxygenated blood is supplied to the lungs via the pulmonary arteries, where the blood gets oxygenated.

Pulmonary Veins: The oxygenated blood flows from the lungs to the pulmonary veins. These veins supply the oxygenated blood to left atrium.

Left Atrium: The oxygenated blood coming from pulmonary veins enters to left atrium. The blood is then forced into the left ventricle via the mitral valve.

Mitral Valve: It is a one way valve which controls the flow of blood from the left atrium to the left ventricle.

Left Ventricle: The oxygenated blood entering the left ventricle through the mitral valve is forced from the left ventricle into the aorta through the aortic valve.

Aortic Valve: It is a one way valve that controls the flow of blood from the left ventricle to the aorta.

Aorta: It is the largest artery in the body and by branching into smaller arteries it carries the oxygenated blood from the heart to the other parts of the body.

The cardiovascular system which contains the pulmonary and systemic circulation consists of the components that are described above.

1.2 Physiology of the Heart

The heart is made of cardiac muscle tissue that contracts and relaxes throughout the lifetime of a person and this contraction and relaxation of the muscles drives the blood from the heart. The contraction of the cardiac muscles of the heart's ventricles is called *systole*, while the relaxation of the cardiac muscles of the

heart's ventricles is called *diastole*. A network of nerve fibers coordinates the contraction and relaxation of the cardiac muscle tissue to obtain an efficient, wave-like pumping action of the heart. Figure 1.2 demonstrates the schematic representation of the sinoatrial (SA) node, atrioventricular (AV) node, bundle of His and Purkinje fibers.

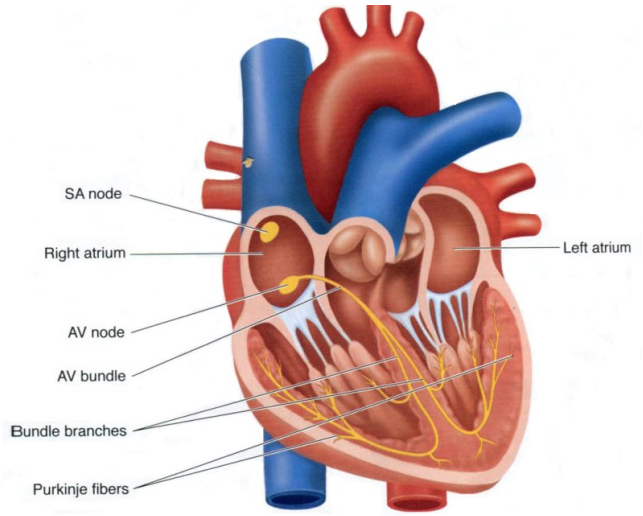


Figure 1.2: Schematic representation of the sinoatrial (SA) node, atrioventricular (AV) node, bundle of His and Purkinje fibers [5].

Sinoatrial node (SA node) which is the natural peacemaker of the heart, triggers an electrical impulse that produces a heartbeat. The impulse trigger that passes through the atria causes the atrium muscles to contract. After the contraction of the atrium muscles, the impulse travelling from the sinoatrial node reaches to the atrioventricular node (AV node). Another impulse is triggered by the atrioventricular node which causes the ventricles to contract. The triggering impulse from the AV node reaches to bundle of His which divides the impulse into the left and right bundle branches, resulting in contraction of the left and right ventricles.

1.3 Myocardial Infarction

Myocardial infarction (MI), which is the medical term for a heart attack, is defined as the damaging or death of an area of the heart muscle (myocardium) resulting from blocked blood supply to the area.

Coronary heart diseases form as a result of atherosclerosis (narrowing and hardening of heart muscles) which happens in the arteries that supply blood to the heart. As a result of accumulation of cholesterol and other fatty substances, plaque or atheroma formation occur on the arterial wall. The deposit of plaque hamper blood flow to the cardiac muscles. Figure 1.3 shows schematic representation of myocardial infarction caused by the blockage of a coronary artery.

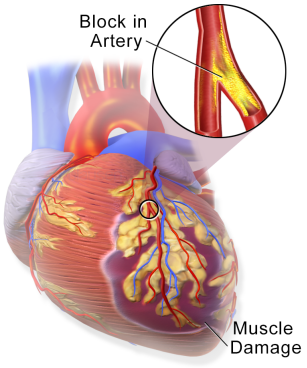


Figure 1.3: Schematic representation of myocardial infarction caused by the blockage of a coronary artery [6].

When the plaque formed on the anterior wall disperses, the blood clots spread and blocks the coronary artery, which results in necrosis. As the blockage of the coronary artery causes a shortage in the oxygen supply to the artery, the heart muscle cells start to die. The region of injury corresponds to the area of the heart muscle cells which encompasses the region of infarction. The heart muscle cells in the region of injury are not completely dead, but they are rather partially functional. The region of infarction and injury represents a permanent damage in that area of the heart, in other words a heart attack or myocardial infarction.

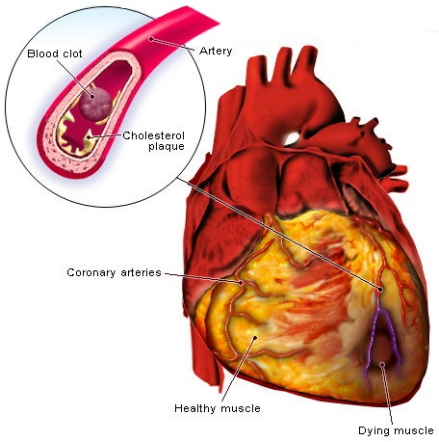


Figure 1.4: Schematic representation of the region of infarction and permanent muscle damage [6].

The decrease in the amount of blood supply flowing through the coronary arteries causes chest pain, which is called angina pectoris. Chest pain is the most common symptom of ACS and is often described as a sensation of tightness, pressure, or squeezing. Pain radiates most often to the left arm, but may also radiate to the lower jaw, neck, right arm, back, and upper abdomen, where it may mimic heartburn [7].

1.4 Electrocardiography

Electrocardiography (ECG) is the most reliable tool for measuring electrical activity of the heart. In electrocardiography, the electrical current transmitted by the heart to all over the body is picked up by an electrical sensing device which is attached to an ECG machine. The electrical sensing devices which detect heart impulses are called electrodes and they are non-invasively placed on the surface of the body. Figure 1.5 demonstrates an example of disposable ECG electrodes.

Figure 1.5: An example of disposable ECG electrodes [8].

The recorded trace acquired from the ECG device is called an electrocardiogram. Electrocardiogram represents the depolarization and repolarization of the heart. By examining the specific features of electrocardiogram, the diagnosis of patients with different kind of heart diseases can be realized. In the evaluation of acute chest pain, ECG is the first diagnostic tool which is considered. Figure 1.6 shows an example of a 12-lead ECG.

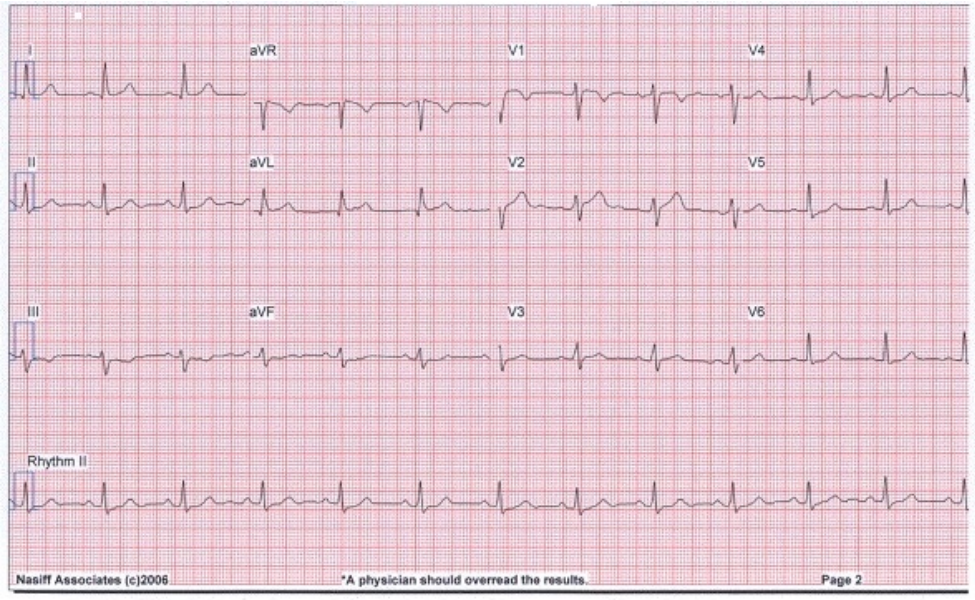


Figure 1.6: An example of a 12-lead ECG [8].

There are mainly three deviations in an ECG; the P wave, the QRS complex and the T wave. Figure 1.7 demonstrates schematic diagram of a normal sinus rhythm.

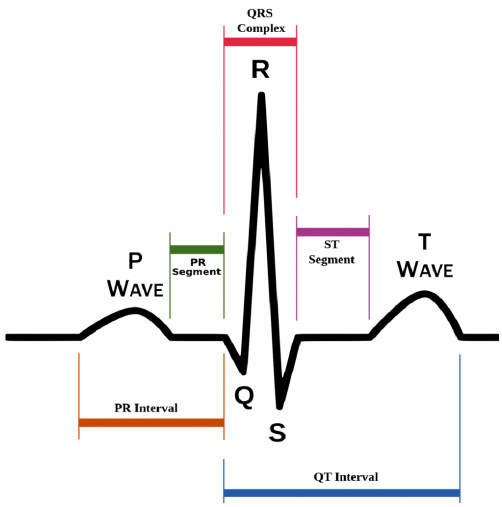


Figure 1.7: Schematic diagram of a normal sinus rhythm [9].

1.4.1 ECG Waves and Intervals

The different components of ECG signal which consist of waves, intervals and segments are described below.

P Wave: The impulse generated by the sinoatrial (SA) node triggers an excitation to the atria before reaching to the atrioventricular (AV) node. The P wave represents electrical activity of the atria in electrocardiogram. The duration of the P wave is generally 0.08 to 0.1 seconds. (80 ms to 100 ms.)

PR Interval: The impulse vector traveling from the AV node to the bundle of His creates an isoelectric pattern in electrocardiogram. The PR interval corresponds to the trace between the inception of the P wave to the inception of QRS complex. It is the indicator of the trace between the onset of atrial depolarization and the onset of ventricular depolarization. The length of PR interval is usually between 0.12 to 0.20 seconds (120 to 200 ms).

PR Segment: The impulse vector traveling from the AV node to the bundle of His, to the bundle branches and then to the Purkinje fibers represents the PR segment. Since this electrical activity does not produce a contraction, it shows up flat between the P wave and the QRS complex in electrocardiogram. The duration of PR interval is largely between 0.05 to 0.12 seconds. (50 to 120 ms.)

QRS complex: The travel of the impulse from the bundle of His to the ventricular walls throughout the right and left bundle branches results in the contraction of the ventricular walls. The rapid depolarization of the right and left ventricles is represented in ECG by the QRS complex. Since the ventricles have a large muscle mass compared to the atria, the QRS complex generally has a much larger amplitude than the P wave. The length of the QRS complex changes between 0.06 to 0.1 seconds (60 to 100 ms), indicating the relatively shorter period of the ventricular depolarization.

J Point: It is the point which corresponds to the end of the QRS complex and the beginning of the ST segment. The degree of ST elevation or depression is measured by using the J point.

ST Segment: After QRS complex, ST segment takes place, forming the isoelectric line in ECG signal. It represents the period when the ventricles are depolarized. The duration of the ST segment changes between 0.08 to 0.12 seconds. (80 to 120 ms.)

T Wave: The repolarization of the ventricles constitutes the T wave in ECG. The duration of the T wave is usually around 0.16 seconds. (160 ms.)

QT Interval: The QT interval is measured from the beginning of the QRS complex to the end of the T wave. It indicates the time between ventricular depolarization and ventricular repolarization. The period of the QT interval is generally between 0.2 and 0.4 seconds. (200 to 400 ms.)

1.4.2 ECG Measurements

In order to record the electrical activity of the heart and to identify different kind of heart anomalies, ECG electrodes are strategically established onto the body surface. Ten electrodes that are detailed below are used for a 12-lead ECG.

Bipolar Leads: Bipolar leads are leads which have one positive and one negative pole. In a 12-lead ECG, the limb leads (I, II and III) are bipolar leads [10].

Limb Leads: Leads I, II and III are called limb leads, because the electrodes that form these signals are located on the limbs, one on each arm and one on the left leg. The combination of lead I, II and III form the Einthoven's triangle which is shown in Figure 1.8.

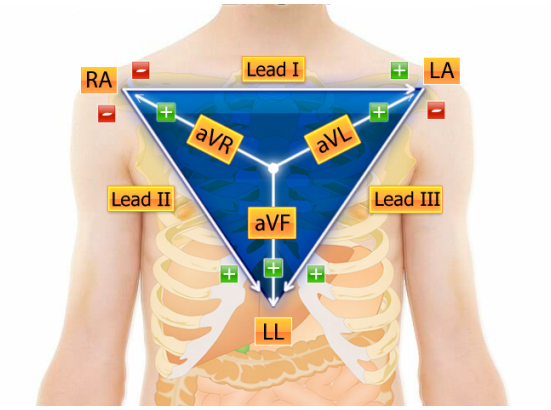


Figure 1.8: Graphical representation of Einthoven's triangle [11].

- **Lead I:** It is the voltage between the (positive) left arm (LA) electrode and right arm (RA) electrode.

$$I = LA - RA \tag{1.1}$$

- **Lead II:** It is the voltage between the (positive) left leg (LL) electrode and the right arm (RA) electrode.

$$II = LL - RA \tag{1.2}$$

- **Lead III:** It is the voltage between the (positive) left leg (LL) electrode and the left arm (LA) electrode.

$$III = LL - LA \tag{1.3}$$

Unipolar Leads: These leads also have two poles, however, the negative pole is a composite pole made up of signals from multiple other electrodes. All leads except the limb leads are unipolar (aVR, aVL, aVF, V1, V2, V3, V4, V5, and V6) in a 12-lead ECG.

Augmented Limb Leads: From Leads I, II and III, additional three leads which are called the augmented limb leads, aVR, aVL and aVF are generated. Augmented limb leads are the recordings between one limb and two other limbs [10].

- **aVR (Augmented Vector Right):** The negative electrode is a combination of the left arm electrode and the left leg electrode. It has the positive electrode on the right arm.

$$aVR = RA - \frac{1}{2}(LA + LL) \tag{1.4}$$

  or by using the limb leads

$$aVR = -\frac{I + II}{2} \tag{1.5}$$

- **aVL (Augmented Vector Left):** The negative electrode is a combination of the right arm electrode and the left leg electrode. It has the positive electrode on the left arm.

$$aVL = LA - \frac{1}{2}(RA + LL) \tag{1.6}$$

  or by using the limb leads

$$aVL = I - \frac{II}{2} \tag{1.7}$$

- **aVF (Augmented Vector Foot):** The negative electrode is a combination of the right arm electrode and the left arm electrode. It has the positive electrode on the left leg.

$$aVF = LL - \frac{1}{2}(RA + LA) \tag{1.8}$$

or by using the limb leads

$$aVF = II - \frac{I}{2} \tag{1.9}$$

Precordial Leads: The electrodes for the precordial leads (V1, V2, V3, V4, V5 and V6) are considered to be unipolar and Wilson's central terminal is used for the negative electrode. Electrodes are placed directly on the chest. They don't require augmentation because of their close proximity to the heart.

Wilson's central terminal V_W is produced by connecting the electrodes RA, LA, and LL together to give an average potential across the body.

$$V_W = \frac{1}{3}(RA + LA + LL) \tag{1.10}$$

Chapter 2

ECG Feature Extraction for the Detection of Acute Coronary Syndrome

In this section, we present the signal analysis techniques which are developed for the detection of critical heartaches in patients who suffer from coronary artery disease. Early diagnosis of acute coronary syndromes is possible by analyzing the changes in ST/T segment of the ECG signal [12].

In order to investigate the performance of the proposed detection techniques, a database of real ECG recordings including basal and ischaemic states of the patients which is constructed in the Cardiology Department of Gazi University Faculty of Medicine, is used.

During ECG recordings, the signal can be distorted because of muscular activity, respiration of the patient, electrode artifacts due to perspiration and electrode movements. In order to minimize the distortion of the signal, the electrodes are placed on the shoulders and on the hip instead of the arms and the leg. This change in the electrode placement is called the Mason-Likar modification [13]. Figure 2.1 illustrates Mason-Likar modification of the 12-lead ECG system.

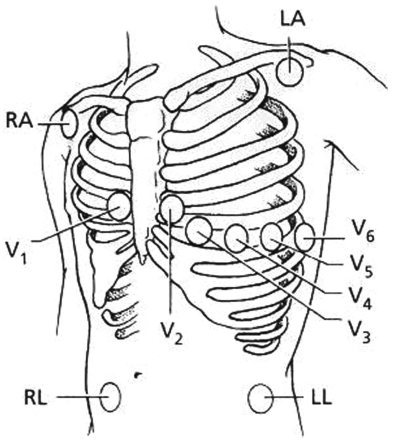


Figure 2.1: Mason-Likar modification of the 12-lead ECG system [8].

2.1 ECG Data Acquisition

The ECG recordings in the database have been obtained by using a simple pc-based, portable 12-lead Microcor ECG device shown in Figure 2.2. This device amplifies the ECG signals acquired from the patient, digitizes, processes and transmits them wirelessly to a computer via a USB adapter attached to it.

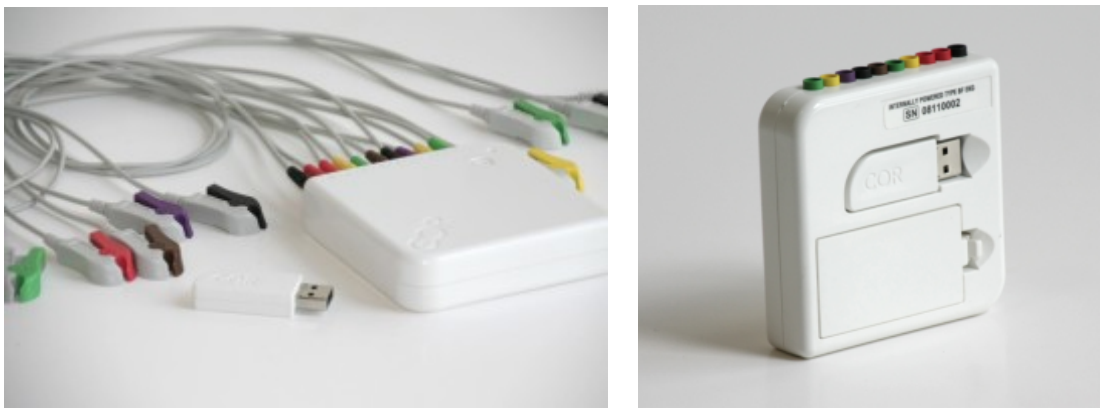


Figure 2.2: 12-lead portable Microcor ECG device.

2.1.1 Electrode Quality Check and Calibration Control

At the beginning of ECG recording, the skin contact quality of all the electrodes are displayed, so that faulty and low-quality electrode connections can be detected and corrected. In addition, calibration control is realized for detecting channel gain and ECG cable continuity faults before starting to record ECG signals.

2.1.2 Microcor PC Software

After the completion of the ECG recording, data is sent from Microcor ECG device to Microcor PC software via a USB adapter. A specific patient folder is then formed for each patient and the 12 lead basal and ischaemic ECG data is saved to that folder, including information about the patient's name, age, gender and the date of the recording. Figures 2.3 and 2.4 show Microcor ECG device visual display of 12-lead ECG recording.

Figure 2.3: Microcor ECG device visual display of 12-lead ECG recording in patient window.

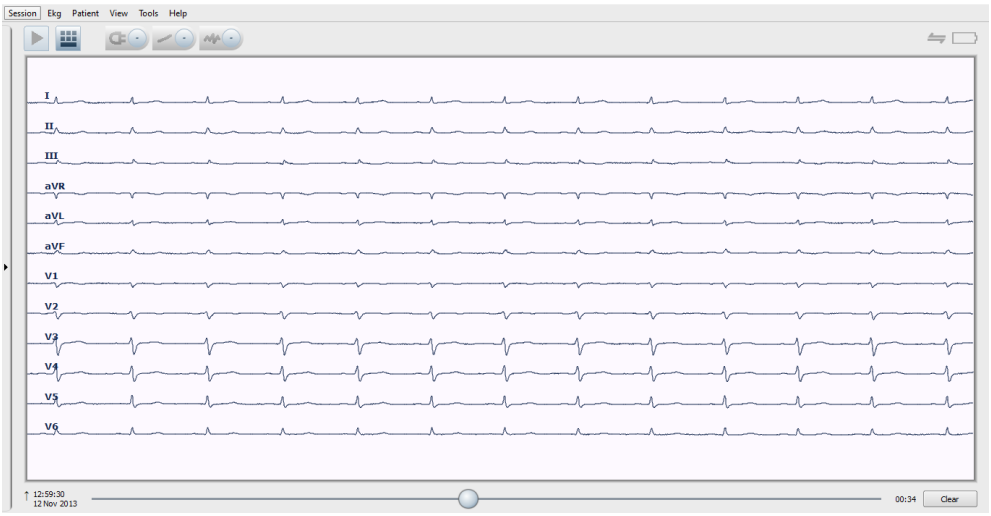


Figure 2.4: Microcor ECG device visual display of 12-lead ECG recording in data window.

2.2 ECG Signal Analysis Techniques

The following standard signal analysis techniques are implemented on the real ECG recordings in the database.

2.2.1 Preprocessing of ECG Signal

ECG signals typically include high frequency muscle artifacts and low frequency baseline wander, which may occur due to the movement of the patient during the ECG recording session [14]. Since the baseline wander and muscle artifact do not contain any information about ECG data, they are removed from the ECG signal by using a band-pass filter. Additionally, notch filter is used to eliminate 50 Hz line voltage [15].

In order to avoid Gibbs phenomenon, which is caused by the abrupt changes at the onset and at the end of the signal, ECG signal is first multiplied by a raised cosine window, then it is filtered by a cascade of band-pass and notch filter

combinations with cut-off frequencies f_L=0.05 Hz and f_H=150 Hz for the band-pass filter and f_L=49 Hz and f_H=51 Hz for the notch filter [16]. Figure 2.7 shows the raised cosine window and 2.8 shows the cascade of band-pass and notch filter combinations.

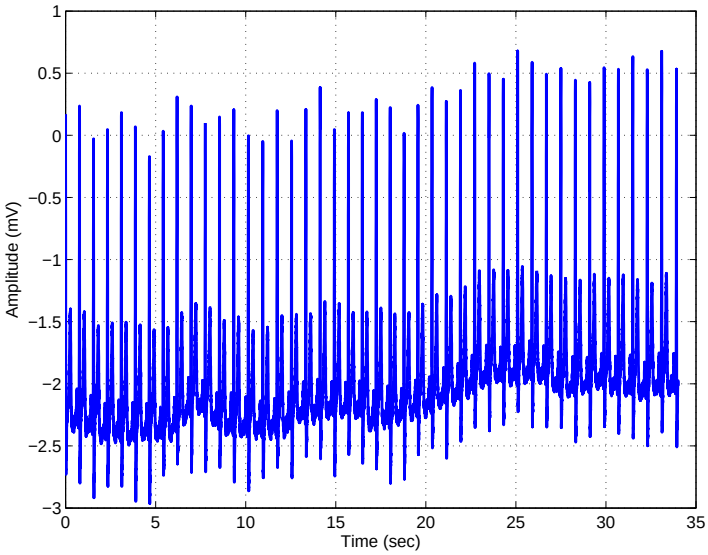


Figure 2.5: Basal ECG raw data taken from a patient.

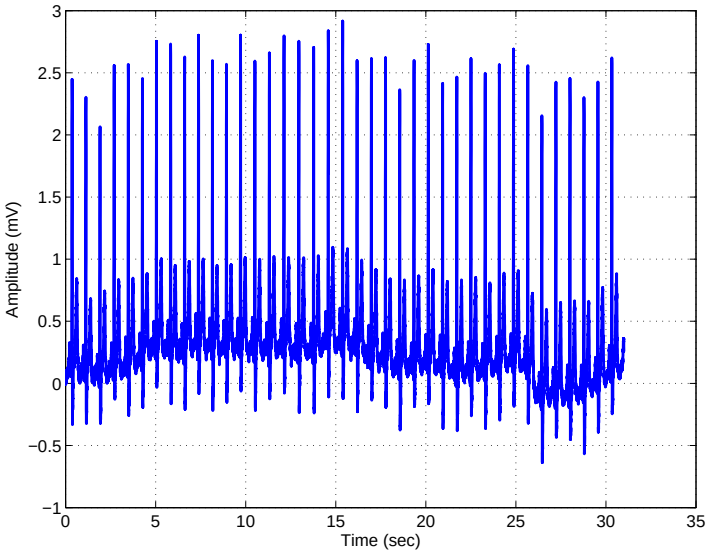


Figure 2.6: Ischaemic ECG raw data taken from a patient.

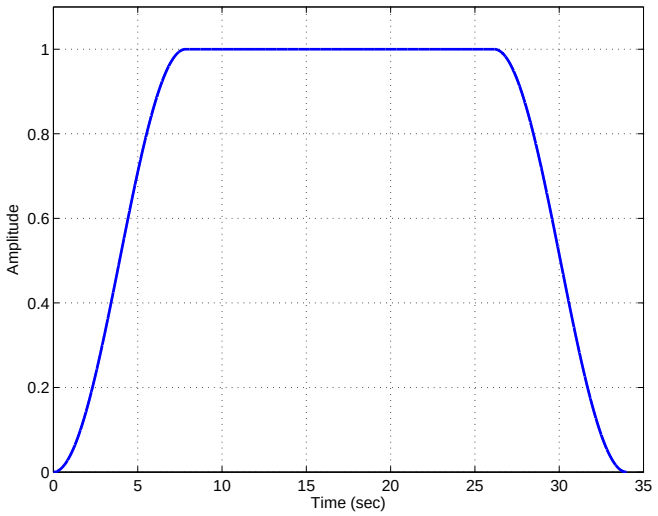


Figure 2.7: Raised cosine window.

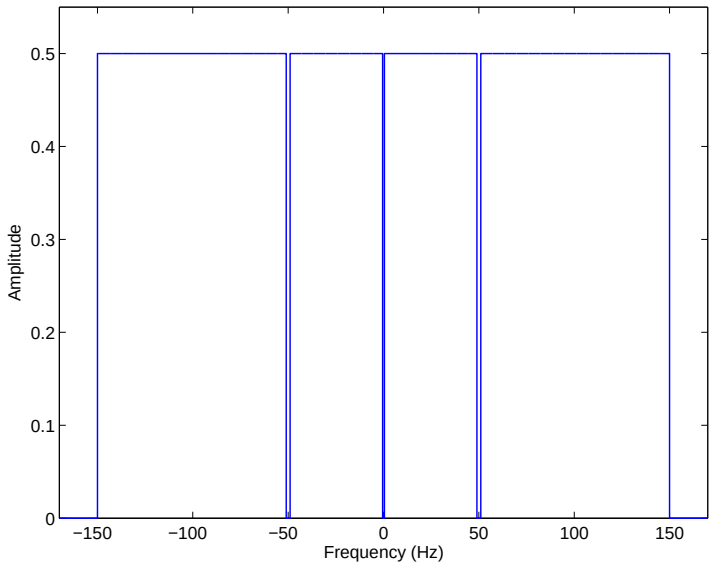


Figure 2.8: Cascade of band-pass and notch filter combinations with cut-off frequencies f_L=0.05 Hz and f_H=150 Hz for the band-pass filter and f_L=49 Hz and f_H=51 Hz for the notch filter.

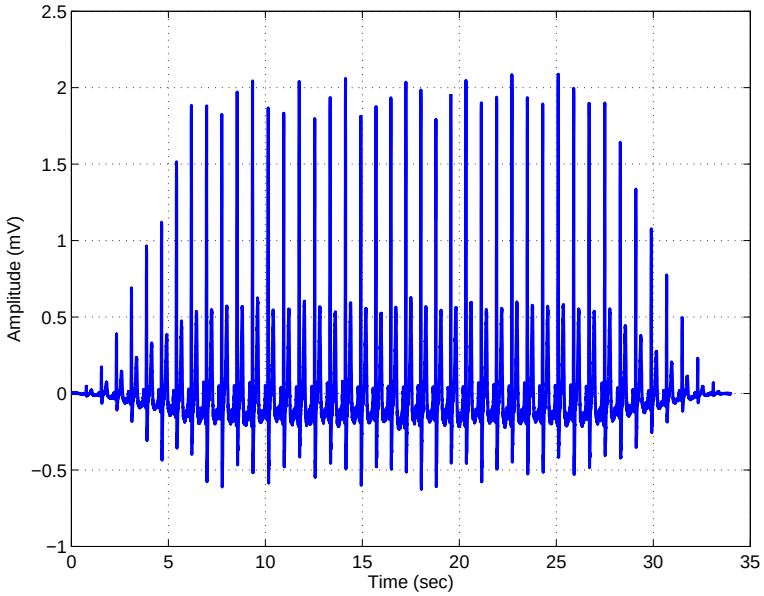


Figure 2.9: Filtered basal ECG recording.

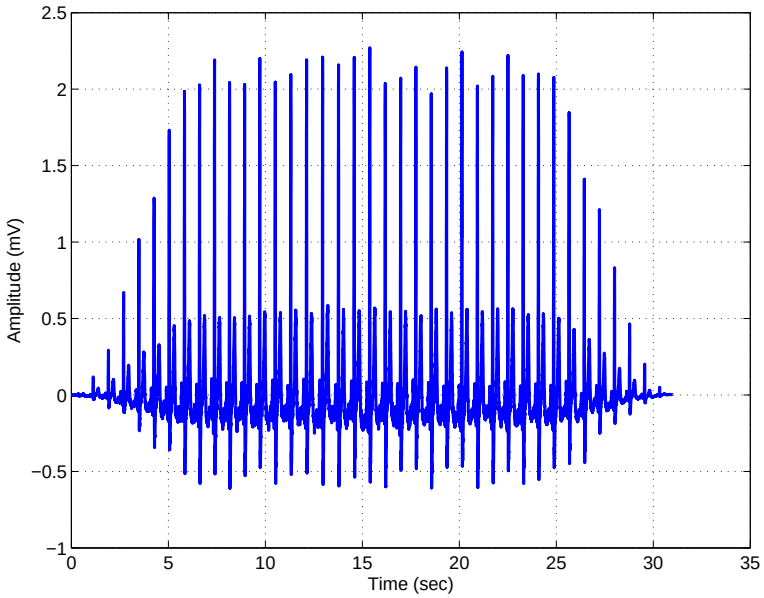


Figure 2.10: Filtered ischaemic ECG recording.

2.2.2 Detection of the QRS Complex

Since the QRS complex is the most distinctive part of the ECG signal, it serves well in the classification of each period in the recording. For the detection of QRS complex, we first set a threshold level which is higher than P and T waves and lower than the QRS complex in amplitude. By comparing the signal values where the ECG signal amplitude is higher than the set threshold, the time domain localization of the QRS complexes are obtained. Figures 2.11 and 2.12 show QRS detection of basal and ischaemic ECG recordings.

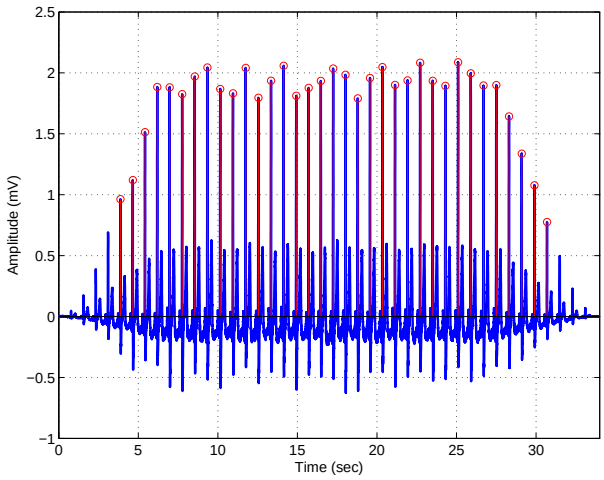


Figure 2.11: QRS detection of basal ECG recording.

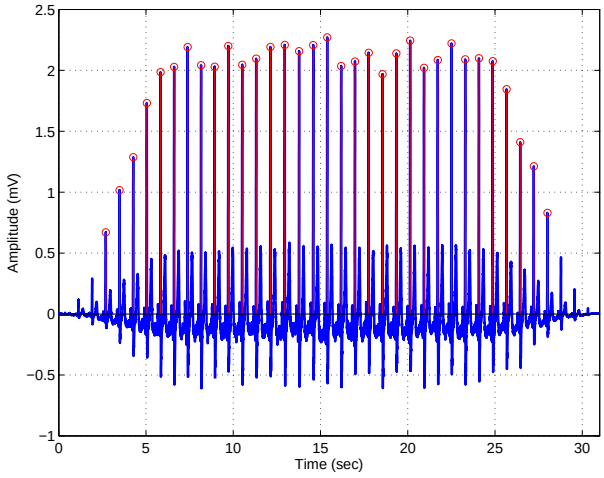


Figure 2.12: QRS detection of ischaemic ECG recording.

2.2.3 Segmentation of the ECG Signal

According to the previously detected QRS positions, ECG signals are segmented into individual periods each corresponding to a single heartbeat. Each period starts 230 ms before and ends 500 ms following the peak position of the QRS complex. Figures 2.13 and 2.14 demonstrate a single period of basal and ischaemic ECG recordings.

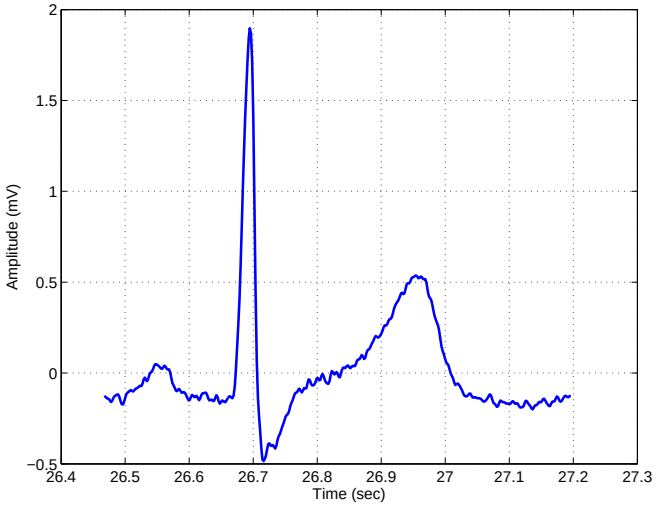


Figure 2.13: A single period of basal ECG recording.

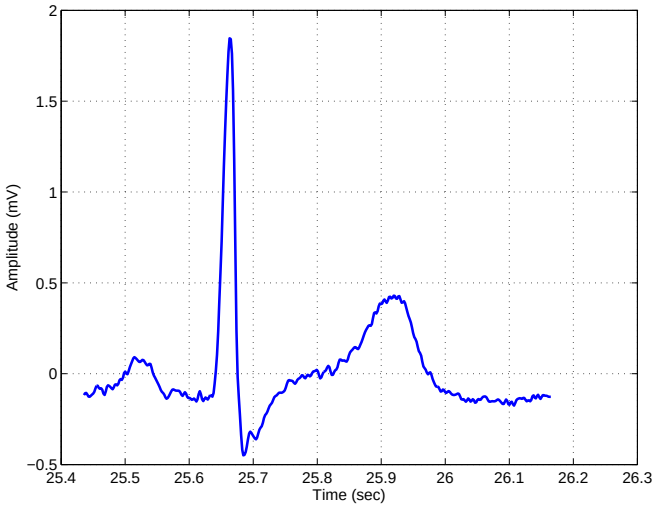


Figure 2.14: A single period of ischaemic ECG recording.

2.2.4 Estimation of the Isoelectric Line

Isoelectric line represents the reference potential level of the measured heart activity for each recorded heartbeat. In ECG signals, isoelectric line is defined as the level of the recorded signal during the PR segment. Our purpose in defining the isoelectric line is to determine the depression or elevation amount in the ST segment. After the removal of isoelectric line, the amplitude of the signal in PR segment oscillates around zero. Figures 2.15 and 2.16 illustrate a single period of basal and ischaemic ECG recordings before and after removal of the isoelectric line.

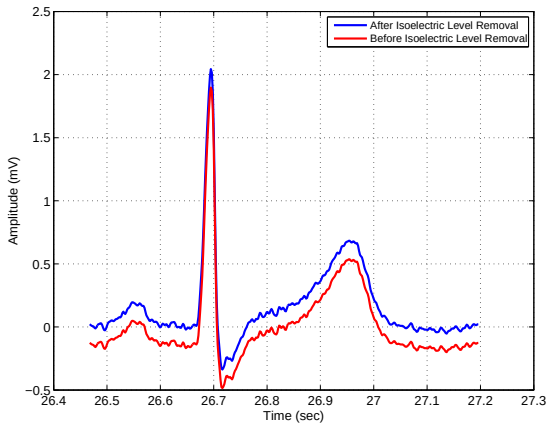


Figure 2.15: A single period of basal ECG recording before and after removal of the isoelectric line.

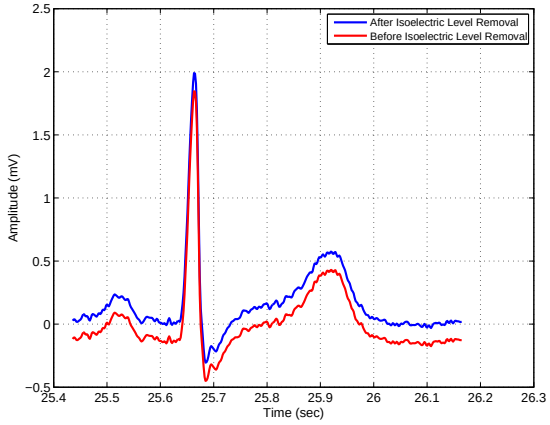


Figure 2.16: A single period of ischaemic ECG recording before and after removal of the isoelectric line.

Figures 2.17 and 2.18 show a single period of basal and ischaemic ECG recordings after removal of the isoelectric line.

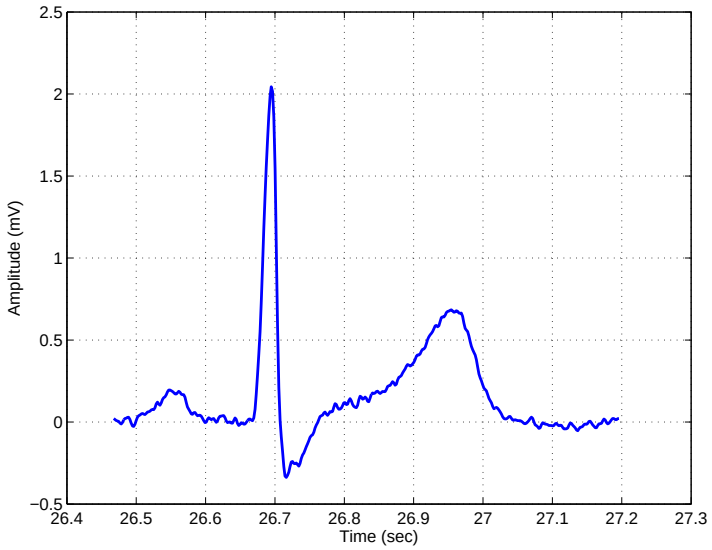


Figure 2.17: A single period of basal ECG recording after removal of the isoelectric line.

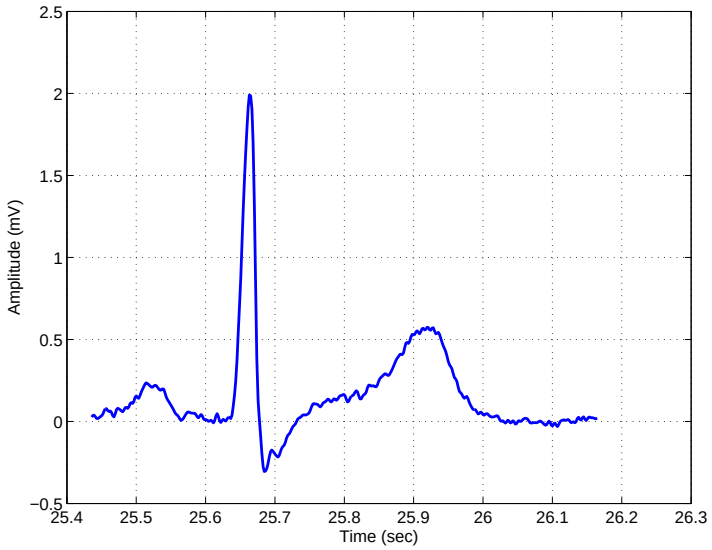


Figure 2.18: A single period of ischaemic ECG recording after removal of the isoelectric line.

2.2.5 Detection of the J Point

We find the location of the J point in a ECG signal in order to locate the corresponding ST segment and T wave. In a period of ECG recording, J point roughly corresponds to the end of the QRS complex and the beginning of the ST segment. For the accurate location of the J point, we find the first value of the signal which is lower than zero within 20 sample points following the QRS peak in time. Alternatively, we also identify the point where the signal reaches down to its minimum value within 20 sample points following the QRS peak. Results show that both of these methods give almost the same detection results. The point corresponding to 15 ms after this detected point is defined as the J point.

2.2.6 Identification of the ST Segment

The location of ST segment in ECG signal is precisely defined by using the J point. The onset of the ST segment corresponds to the J point, while the end of the ST segment is defined as 160 ms after the J point.

2.2.7 Detection of the T Wave

The changes in the T wave may indicate symptoms of acute coronary syndrome. Therefore, we define the location of the T wave in ECG signal by taking the J point as a reference. The onset of the T wave corresponds to 150 ms after the J point, while the end of the T wave corresponds to 350 ms after the J point. Figure 2.19 shows the detected T wave of the ECG signal.

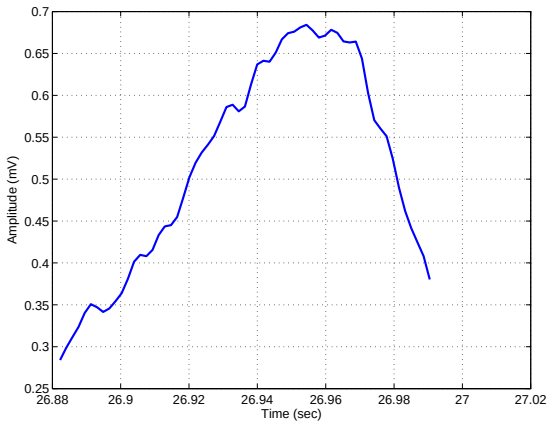


Figure 2.19: The T wave.

2.3 Extraction of ECG Features from the ST Segment

2.3.1 Depression or Elevation in the ST Segment

One of the symptoms of acute coronary syndrome is the depression or elevation in the ST segment. ST level is identified by taking the isoelectric line as a reference. According to the isoelectric line, if the instant ECG ST level is lower than the basal ECG ST level, we conclude that a depression has taken place in ST segment. However, if the instant ECG ST level is higher than the basal ECG ST level, we conclude that an elevation has taken place in ST segment [17].

2.3.2 Slope of the ST Segment

Slope of the ST segment constitutes one of the tools for analyzing the effects of acute coronary syndrome on ECG signals. The slope of ST segment, is estimated as the slope of the best fitting line in terms of least squares to the samples of the ST segment. Figures 2.20 and 2.21 illustrate a single period of basal and ischaemic ECG recordings after implementation of the signal analysis techniques.

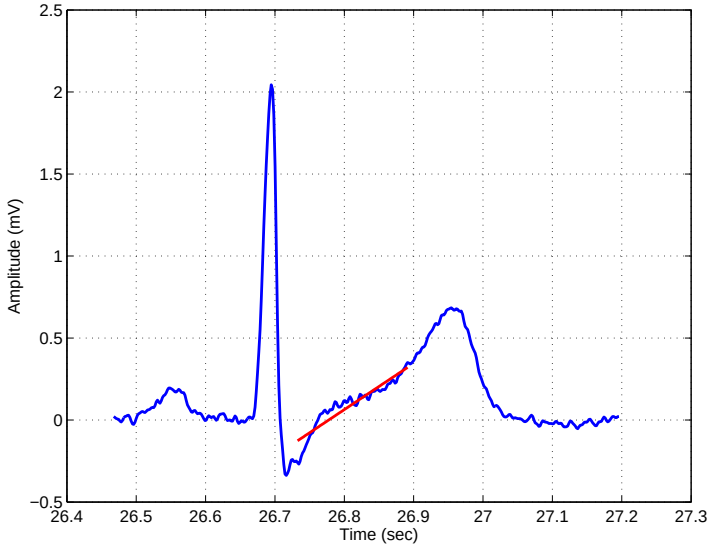


Figure 2.20: A single period of basal ECG recording after implementation of the signal analysis techniques.

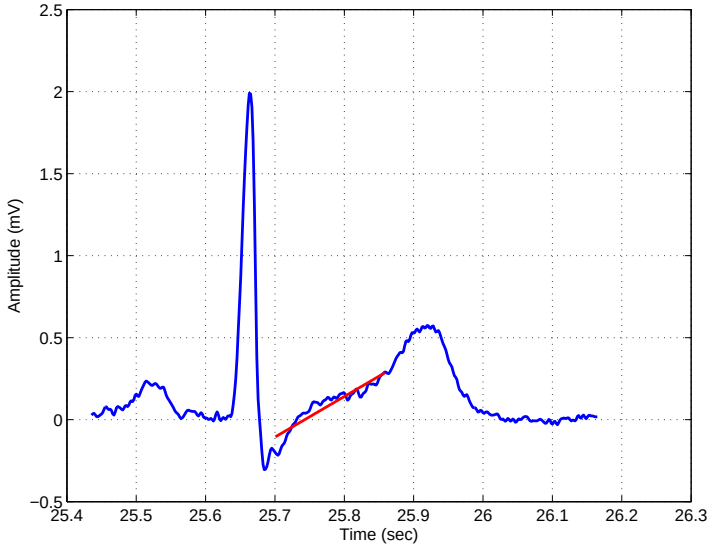


Figure 2.21: A single period of ischaemic ECG recording after implementation of the signal analysis techniques.

2.4 Extraction of ECG Features from the T Wave

2.4.1 Maximum/Minimum Point of the T Wave

In some cases, due to the lengthening of ST segment, the onset and the end of the T wave can be delayed. Therefore, for accurately locating the T wave in ECG signals, the extremum point of the T wave can be used. The extremum point of the T wave is found between the time interval corresponding to 150 ms and 350 ms following the J point for more precise identification of the T wave.

2.4.2 Area Under the T Wave

The area under the T wave is considered as a tool for the diagnosis of acute coronary syndromes. It is approximated by using the trapezoidal rule and all the values of the signal between the onset and the end of the T wave.

2.5 Statistical Variation of Computed ECG Features

To investigate their statistical variation, the ECG features of the recordings in the database have been computed. After obtaining the ECG features, in order to analyze the differences between basal and ischaemic ECG signals, the mean, maximum, minimum and standard deviation values of three different ECG features are calculated for all 12 leads of the ECG signals. Table 2.1 and 2.2 shows an example of computed ECG feature values for single lead ECG signal that are obtained from a single patient.

Table 2.1: Mean, standard deviation, maximum and minimum values of basal ECG features.

	ST Level (mV)	**ST Slope** (mV/sec)	**T Area** (mV.sec)
Mean Value	0.24	5.42	0.12
Standard Deviation	0.05	0.72	0.01
Maximum Value	0.49	7.12	0.16
Minimum Value	0.14	3.24	0.05

Table 2.2: Mean, standard deviation, maximum and minimum values of ischaemic ECG features.

	ST Level (mV)	**ST Slope** (mV/sec)	**T Area** (mV.sec)
Mean Value	0.11	2.31	0.09
Standard Deviation	0.03	0.41	0.02
Maximum Value	0.28	4.91	0.15
Minimum Value	0.09	1.36	0.05

In clinical studies, when basal ECG recording is used as a reference, an ST segment elevation or depression of 0.1 mV in at least one of the leads is considered as an indication of acute coronary syndrome. The comparison between ST segment feature results of basal and ischaemic ECG recordings showed that there is an ST segment elevation or depression which is greater than 0.1 mV in at least one of the leads in ECG recordings of each patient. Table 2.1 and 2.2 shows an example of such a case where there is an ST segment depression of 0.13 mV.

In previous studies, the contribution of T wave analysis to the sensitivity of ischaemia detection has not been investigated [3]. Therefore, in this thesis, the analysis of the effect of T wave changes on ACS detection has important contributions to the related studies. The results of the analysis show that, the changes in the area of T wave are coherent with the changes in the slope of ST segment in each lead. Table 2.1 and 2.2 demonstrates an example of this result where a decrease in the T wave area occurs simultaneously with a decrease in the

ST segment slope. For this reason, the results indicate that using the T wave feature in addition to the ST segment features has a significant contribution to the analysis on increasing the sensitivity and specificity of ischaemia detection.

For the purpose of analyzing the differences between basal and ischaemic ECG signal, bar graphs of three different ECG features which are ST level, ST slope and T wave area are plotted for all 12 leads of ECG recordings for each patient. Figures 2.22, 2.23 and 2.24 demonstrate the mean and standard deviation values of 12 leads of ECG recordings from a single patient for three different features. As a result of the bar graph analysis, it is observed that in different patients the discriminative strength of the ST level, ST slope and T wave area features are different in a specific lead. In addition, the results showed that the discriminative strength of a specific feature in 12 different leads is different in each patient.

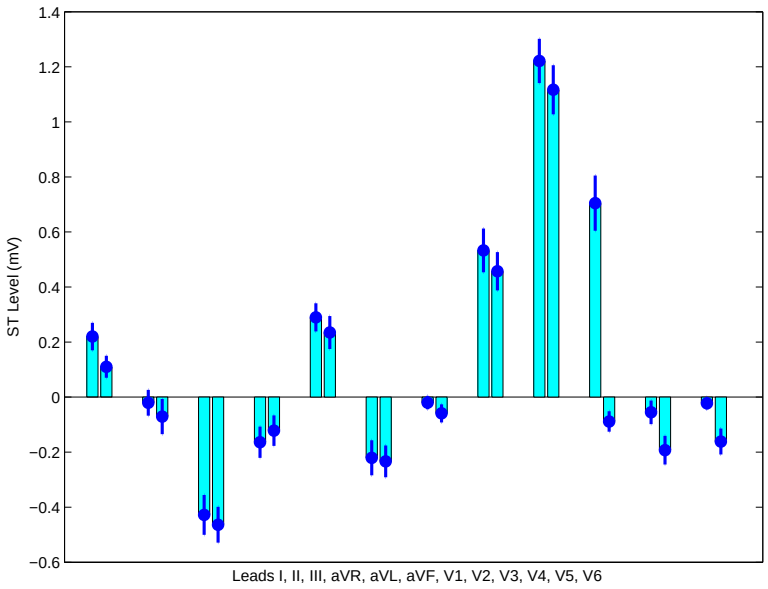


Figure 2.22: ST level bar graph showing mean and standard deviation values of basal and ischaemic ECG for 12 lead.

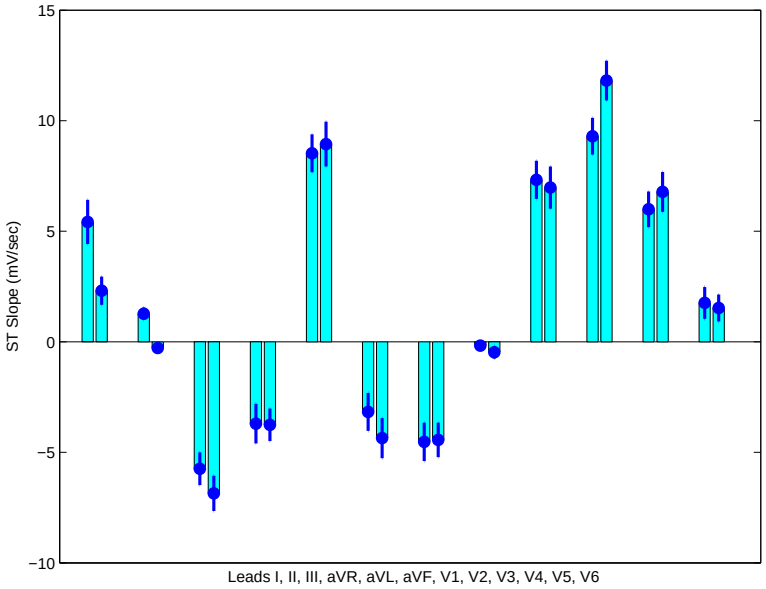


Figure 2.23: ST slope bar graph showing mean and standard deviation values of basal and ischaemic ECG for 12 lead.

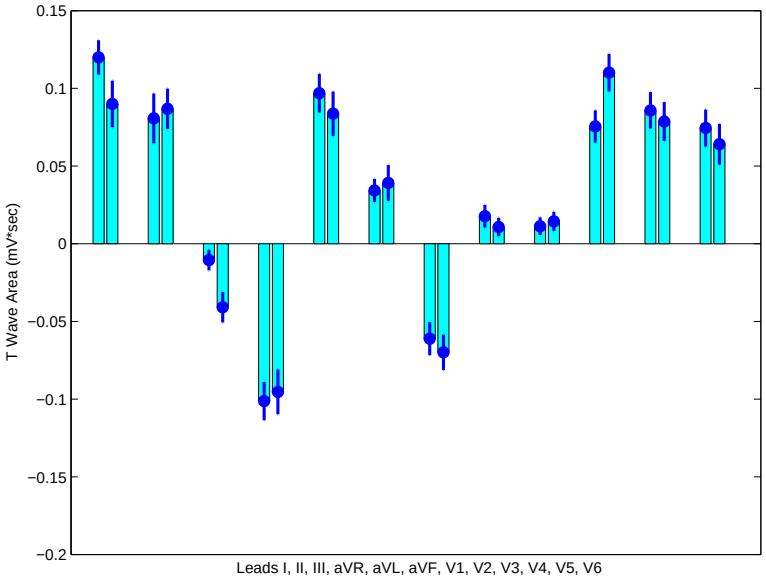


Figure 2.24: T wave area bar graph showing mean and standard deviation values of basal and ischaemic ECG for 12 lead.

Chapter 3

Proposed Detection Technique of Acute Coronary Syndrome

3.1 Formulation of the Detection Framework Using ECG Features

After realizing the feature extraction step, the most discriminating ECG features which are required for the detection of acute coronary syndromes are obtained. These three features are acquired from basal and ischaemic ECG recordings of the patients. By using support vector machines (SVM) operating with linear and radial basis function (RBF) kernels, we obtained classifiers that use two and three most discriminating features of the ST/T sections. The set of ECG features constitute the training data for the support vector machines. The classification problem can be expressed as follows.

Given a set of n points which is the training data of the form

$$\mathcal{D} = \{(\mathbf{x}_i, y_i) \mid \mathbf{x}_i \in \mathbb{R}^p,\ y_i \in \{-1, 1\}\}_{i=1}^{n} \tag{3.1}$$

where each $\mathbf{x}_i$ is a p-dimensional feature vector and y_i shows to which class the point $\mathbf{x}_i$ belongs with a value either 1 or -1.

The aim of the classification task is to decide in which class a new data point, which is called the test data, will be, when we have given training data points belonging to one of the two classes. For this purpose, we want to learn if we can distinguish points by using a (p-1)-dimensional hyperplane where a data point is seen as a p-dimensional vector [18]. Among many hyperplanes which may classify the data, the hyperplane that provides the widest margin between two classes is identified for each case by using support vector machines. In other words, the hyperplane which provides maximum space between itself and the closest data points on both side, which are named as the support vectors, is determined.

3.2 Review of Classification Techniques Used in the Proposed Framework

In the following sections a review of support vector machines (SVM), kernel methods and performance metrics of binary classifiers are introduced.

3.2.1 Support Vector Machine

Support vector machines (SVMs) are supervised learning models used for classification in machine learning, with associated learning algorithms that analyze data and recognize patterns. An SVM training algorithm creates a model that allocates new samples into one class or the other when a set of training samples each marked as belonging to one of the two classes are given. The aim of an SVM model is to represent the samples as points in space, mapped so that the samples of the separate classes are distinguished by an apparent space that is as large as possible [19]. After the mapping of the new samples, they are predicted to be a part of a class according to the side of the space that they correspond. Nonlinear classification can also be realized by support vector machines by carrying out the kernel trick, which refers to mapping their inputs to high-dimensional

feature spaces where the classes become separable. In other words, for the realization of classification tasks, a support vector machine creates a hyperplane in a high dimensional space. Since the generalization error of the classifier decreases as the margin increases, the hyperplane having the greatest distance to the closest training data of any class provides a good separation [20]. In some cases, although the problem is defined in a finite dimensional space, the samples to separate may not be linearly separable in the same space. Therefore, in order to make the separation easier, it was suggested that the finite-dimensional space be mapped to a higher-dimensional space. By describing the mappings in terms of a kernel function $K(x, y)$ that is suitable for the task, the mappings are formed to guarantee that the dot products are easily computed in respect to the original space variables [21]. The definition of higher-dimensional space hyperplanes can be made as the set of points which have a constant dot product with a vector in the same space. Linear combinations of feature vector images with parameters α_i constitutes the vectors that define the hyperplanes [22]. When the hyperplane is chosen in this way, the following relation defines the points x in the feature space that are mapped into the hyperplane.

$$\sum_i \alpha_i K(x_i, x) = \text{constant}$$

The proximity level of the test point x to the corresponding data point x_i is measured by each term in the sum when $K(x, y)$ becomes smaller as y grows further away from x. In this way, the sum of kernels above can be used to measure the relative nearness of each test point to the data points originating in one or the other of the sets to be discriminated. Thus, relative closeness of each test point to the data points which belong to one of the sets to be separated can be measured by the sum of kernels stated above [23].

3.2.2 Linear SVM

Given a set of n points which is the training data of the form

$$\mathcal{D} = \{(\mathbf{x}_i, y_i) \mid \mathbf{x}_i \in \mathbb{R}^p,\, y_i \in \{-1, 1\}\}_{i=1}^{n} \tag{3.2}$$

where each $\mathbf{x}_i$ is a p-dimensional feature vector and y_i shows to which class the point $\mathbf{x}_i$ belongs with a value either 1 or -1. The aim is to determine maximum-margin hyperplane which separates points having $y_i = 1$ from those having $y_i = -1$. Every hyperplane can be identified by using the set of points $\mathbf{x}$ which satisfies

$$\mathbf{w} \cdot \mathbf{x} - b = 0, \tag{3.3}$$

where $\cdot$ states the dot product and $\mathbf{w}$ states the vector which is normal to the hyperplane. The offset of the hyperplane from the origin through the normal vector $\mathbf{w}$ is identified by the parameter $\frac{b}{\|\mathbf{w}\|}$. For the cases where the training data is linearly separable, two hyperplanes which separate the data can be chosen and the space between these two hyperplanes can be maximized. The region between hyperplanes is called "the margin". The following equations define these hyperplanes.

$$\mathbf{w} \cdot \mathbf{x} - b = 1 \tag{3.4}$$

and

$$\mathbf{w} \cdot \mathbf{x} - b = -1 \tag{3.5}$$

Since the interval between these two hyperplanes is $\frac{2}{\|\mathbf{w}\|}$, we will minimize $\|\mathbf{w}\|$. The following constraint have to be added in order to avoid the existence of the data points in the margin: for each i either

$$\mathbf{w} \cdot \mathbf{x}_i - b \geq 1 \qquad \text{for } \mathbf{x}_i\text{of the first class} \tag{3.6}$$

or

$$\mathbf{w} \cdot \mathbf{x}_i - b \leq -1 \qquad \text{for } \mathbf{x}_i\text{of the second class} \tag{3.7}$$

We can rewrite this as:

$$y_i(\mathbf{w} \cdot \mathbf{x}_i - b) \geq 1, \quad \text{for all } 1 \leq i \leq n. \tag{3.8}$$

The optimization problem can be obtained by putting these together:
Minimize

$$\|\mathbf{w}\| \tag{3.9}$$

subject to

$$y_i(\mathbf{w} \cdot \mathbf{x_i} - b) \geq 1, \quad \text{for any i} = 1, \ldots, \text{n} \tag{3.10}$$

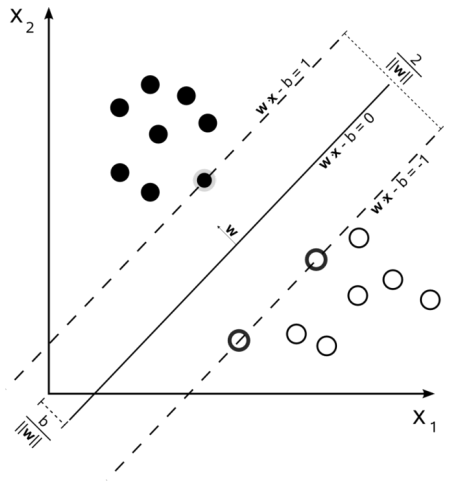


Figure 3.1: Margins and maximum-margin hyperplane for an SVM trained with samples from two classes. Support vectors are the samples on the margin [24].

3.2.2.1 Soft margin

A modified maximum margin idea that permits to mislabeled samples is suggested by Corinna Cortes and Vladimir N. Vapnik [23]. In cases where the two samples

cannot be separated by a hyperplane, a hyperplane which separates the samples as cleanly as possible and which has the maximum distance to the closest cleanly separated samples will be chosen by the soft margin method. Non-negative slack variables, ξ_i, that measure the misclassification level of the data x_i is presented by this method [25].

$$y_i(\mathbf{w} \cdot \mathbf{x_i} - b) \geq 1 - \xi_i \quad 1 \leq i \leq n. \tag{3.11}$$

The objective function is then increased by a function which penalizes non-zero ξ_i, and the optimization becomes a trade off between a small error penalty and a large margin. The optimization problem in case of a linear penalty function:

$$\arg\min_{\mathbf{w},\xi,b} \left\{ \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{n} \xi_i \right\} \tag{3.12}$$

subject to

$$y_i(\mathbf{w} \cdot \mathbf{x_i} - b) \geq 1 - \xi_i, \quad \xi_i \geq 0, \quad \text{for any i=1,. . . , n} \tag{3.13}$$

By using Lagrange multipliers, constraint in (2.10) together with the objective of minimizing $\|\mathbf{w}\|$ can be done by solving the following problem.

$$\arg\min_{\mathbf{w},\xi,b} \max_{\boldsymbol{\alpha},\boldsymbol{\beta}} \left\{ \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{n} \xi_i - \sum_{i=1}^{n} \alpha_i[y_i(\mathbf{w} \cdot \mathbf{x_i} - b) - 1 + \xi_i] - \sum_{i=1}^{n} \beta_i \xi_i \right\} \tag{3.14}$$

with $\alpha_i, \beta_i \geq 0$.

3.2.3 Nonlinear Classification

By implementing the kernel trick to maximum-margin hyperplanes, Vladimir N. Vapnik, Isabelle M. Guyon and Bernhard E. Boser proposed a way to create nonlinear classifiers [26]. The resulting algorithm is formally similar, except that every dot product is replaced by a nonlinear kernel function. The difference of the final algorithm from the linear case is that a nonlinear kernel function replaces

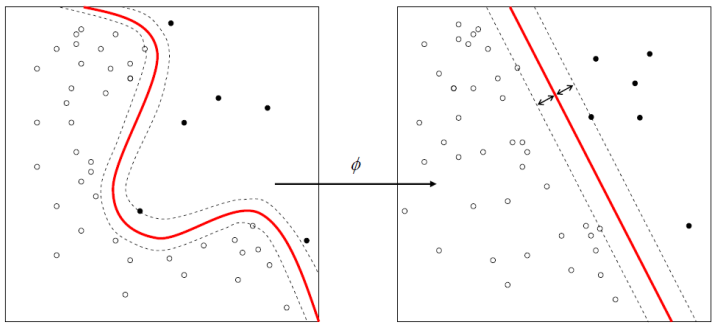


Figure 3.2: Schematic representation of a kernel machine [28].

the dot product. That way, the maximum-margin hyperplane can be fitted by the algorithm in a transformed feature space. The transformed space can be high dimensional and the transformation can be nonlinear. For this reason, the classifier may be a hyperplane in the high-dimensional feature space, although it is nonlinear in original input space [27].

3.2.3.1 Kernel Method

Different types of relations in data sets are found and studied as a pattern analysis task. In order to realize these tasks, data is usually represented as feature vectors. In pattern analysis concept, kernel methods are algorithms that are popular with support vector machine as a well known member [29]. Kernel functions allow kernel methods to work in high-dimensional feature spaces without computing data coordinates in that space [30]. The inner products between all pairs of data images in the feature space are computed by kernel methods. This method is computationally cheaper compared to the apparent coordinate computation and it is named as the kernel trick [31].

3.2.3.2 Kernel Trick

In any algorithm that is only affiliated with the dot product between two vectors, kernel trick can be implemented. In the context of kernel trick, kernel function

takes place of dot product. Linear algorithms are converted to non-linear algorithms with the implementation of the kernel method [32]. In the range space of a feature space φ, non-linear algorithms are equivalent to their linear ones. Besides, the function φ doesn't need to be apparently calculated because of the usage of kernels. Since the mentioned high-dimensional feature space can be infinite-dimensional and therefore impossible to compute, this situation is very desirable. The apparent mapping which is needed for the linear algorithms to learn the non-linear function is prevented by the kernel trick [32]. Some functions $K(x, y)$ may be stated as an inner product for x, y on S. In general, K is mentioned as a kernel function. In a specific machine learning problem, one may constitute $\varphi : S \to V$ such that

$$K(x, y) = \langle \varphi(x), \varphi(y) \rangle_V \tag{3.15}$$

and confirm that $\langle \cdot, \cdot \rangle_V$ is an inner product. Moreover, it is enough to know that V is an inner product space, so an apparent presentation of φ is not needed.

3.2.3.3 Linear Kernel

The linear kernel is given by the inner product $< x, y >$ plus a constant c.

$$k(\mathbf{x_i}, \mathbf{x_j}) = (\mathbf{x_i} \cdot \mathbf{x_j} + c) \tag{3.16}$$

3.2.3.4 Gaussian Radial Basis Function Kernel

The radial basis function kernel, in other words RBF kernel, is a well known kernel function used for support vector machine classification [33].

$$k(\mathbf{x_i}, \mathbf{x_j}) = \exp(-\gamma \|\mathbf{x_i} - \mathbf{x_j}\|^2), \quad \text{for } \gamma > 0. \tag{3.17}$$

It is sometimes parameterized using $\gamma = 1/2\sigma^2$.

3.2.4 Parameter Selection

The effectiveness of SVM depends on the selection of kernel, the kernel's parameters, and soft margin parameter C [34]. In order for the SVM algorithm to be efficient, the choice of the kernel function, kernel function parameters and soft margin parameter C must be chosen properly. By using geometrically growing sequences of C and γ, the ideal combination of C and γ is frequently selected by making a grid search. In general, after checking every combination of parameters by using the cross validation method, the parameters that give the higher cross-validation accuracy values are selected. Eventually, the whole training data is trained with the selected ideal parameters and final model that was used to test and classify new data.

3.2.5 Performance Evaluation of Binary Classifiers

3.2.5.1 Confusion Matrix

Confusion matrix is also known as an error matrix or a contingency table and it provides performance visualization of a supervised learning algorithm [35]. Rows of the table represents the cases in an actual class and columns of the table represents the cases in a predicted class. There are two columns and two rows in a confusion matrix which illustrates the number of true positives, true negatives, false positives and false negatives. Therefore, confusion matrix provides a more elaborate performance analysis compared to only accuracy results [36].

Table 3.1: Confusion matrix for a binary classification model.

		Predicted Class	
		Negative	**Positive**
Actual Class	**Negative**	True Negative (TN)	False Positive (FP)
	Positive	False Negative (FN)	True Positive (TP)

The entries in the confusion matrix have the following definitions:

True Negative (TN): The number of correct predictions that an instance is negative.

False Positive (FP): The number of incorrect predictions that an instance is positive. It is also referred as type I error.

False Negative (FN): The number of incorrect predictions that an instance negative. It is also referred as type II error.

True Positive (TP): The number of correct predictions that an instance is positive.

3.2.5.2 Derivations and Terminology from the Confusion Matrix

By using the four different entries of the confusion matrix, the following derivations that are significant for the measurement of a classifier's performance can be obtained.

Accuracy: Accuracy (ACC) is the proportion of the total number of predictions that were correct.

$$ACC = \frac{TP + TN}{TN + FP + FN + TP} \tag{3.18}$$

Sensitivity: Sensitivity, which is also called as the probability of detection, true positive rate, hit rate and recall, is the proportion of positive cases that were correctly identified.

$$TPR = \frac{TP}{TP + FN} \tag{3.19}$$

False Positive Rate: False positive rate (FPR), which is also called as the false alarm rate and fall-out, is the proportion of negative cases that were incorrectly classified as positive.

$$FPR = \frac{FP}{FP + TN} \tag{3.20}$$

Specificity: Specificity, which is also called as the true negative rate, is defined as the proportion of negative cases that were classified correctly.

$$SPC = \frac{TN}{TN + FP} \tag{3.21}$$

False Negative Rate: False negative rate (FNR), which is also called as the miss rate, is the proportion of positive cases that were incorrectly classified as negative.

$$FNR = \frac{FN}{FN + TP} \tag{3.22}$$

Positive Predictive Value: Positive predictive value (PPV), which is also called as the precision, is the proportion of the predicted positive cases that were correct.

$$PPV = \frac{TP}{TP + FP} \tag{3.23}$$

Negative Predictive Value: Negative predictive value (NPV) is the proportion of the predicted negative cases that were correct.

$$NPV = \frac{TN}{TN + FN} \tag{3.24}$$

Classified Rate: Classified rate is the proportion of the classified samples.

$$\text{Classified Rate} = \frac{\text{Classified Samples}}{\text{Total Number of Samples}} \tag{3.25}$$

Inconclusive Rate: Inconclusive rate is the proportion of the non-classified samples.

$$\text{Inconclusive Rate} = \frac{\text{Non-classified Samples}}{\text{Total Number of Samples}} \tag{3.26}$$

3.2.5.3 Receiver Operating Characteristic

A receiver operating characteristic (ROC) curve demonstrates the performance of a binary classifier for different discrimination thresholds of the classifier [37]. It can be formed by plotting false positive rate (FPR) vs. true positive rate (TPR) for several discrimination thresholds.

ROC space shows trade-offs between false positive and true positive values and it is identified by TPR and FPR as y and x axes respectively. Different points in the ROC space indicate different prediction results of a confusion matrix [38].

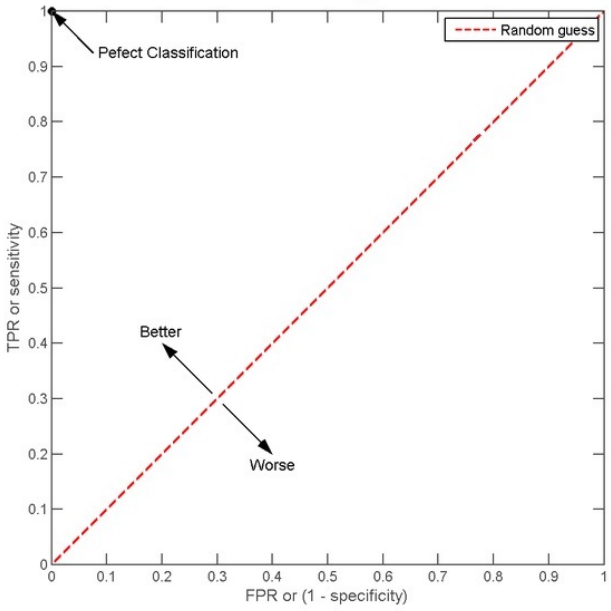


Figure 3.3: Schematic representation of the ROC space [39].

The point that indicates the perfect classification result on the ROC space has the coordinate (0,1) and it stands for 100% sensitivity and 100% specificity. A classifier which predicts every sample as negative has the coordinate (0,0), on the other hand a classifier which predicts all samples as positive has the coordinate (1,1). A classifier which is wrong for every classification has the coordinate (1,0) in the ROC space. The diagonal line on the ROC space stands for a random prediction, while the points under the diagonal line stands for poor classification results and the points above the diagonal line stands for good classification results [39].

Chapter 4

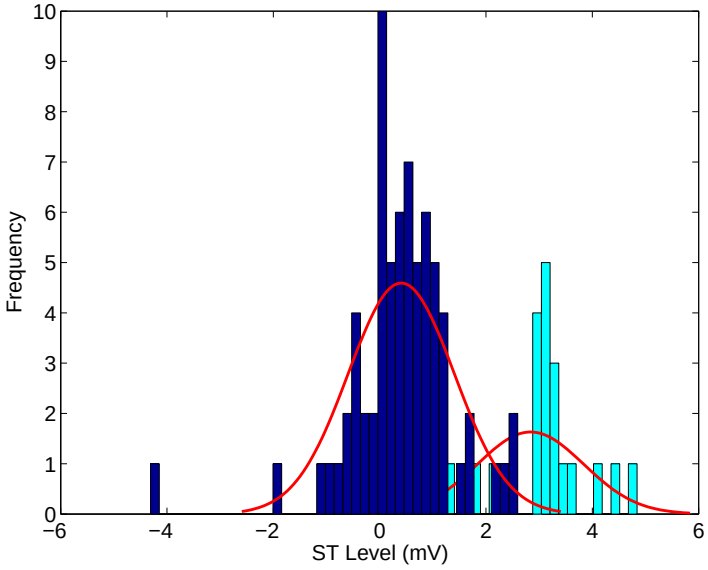
Proposed Detection Techniques Using Individual and Multiple Features of ECG Data

4.1 ECG Data Classification Using Individual Features

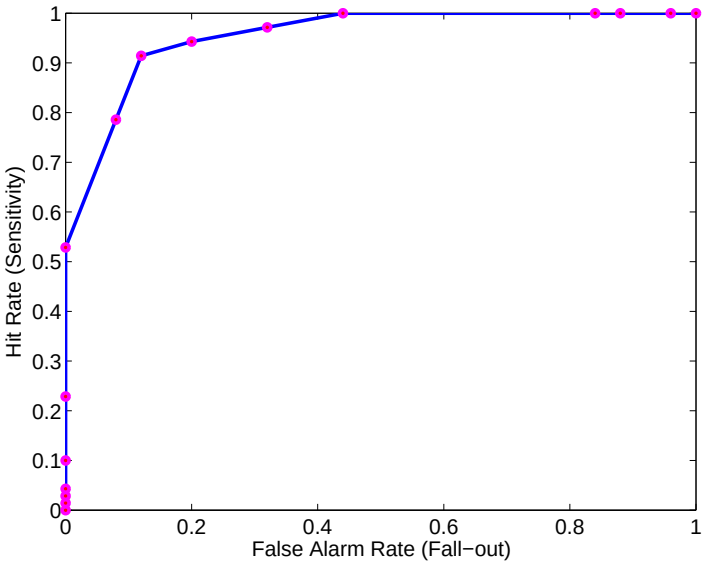
The graphical representation of the data distribution for basal and ischaemic ECG signals are plotted as a histogram for three different ECG features. While plotting histogram graphs, quantization of histogram bin values is realized for the purpose of comparing three different feature histograms. In order to visually identify the estimate of the underlying probability distribution, a normal density function is fitted to the histograms of both ECG data [40].

With the intent of illustrating the classification performance by using histogram, hit rates and false alarm rates are calculated for different discrimination threshold values and histogram ROC curves are plotted for each ECG feature. Figures 4.1, 4.2 and 4.3 show histograms of three different ECG features and their corresponding ROC curves. For this patient, it can be seen that ST level, ST

slope and T wave area histograms do not overlap and the corresponding ROC curves illustrate good classification results. For this reason, these non-overlapping histograms are a sign of features which have a significant role in discriminating the two classes from each other.

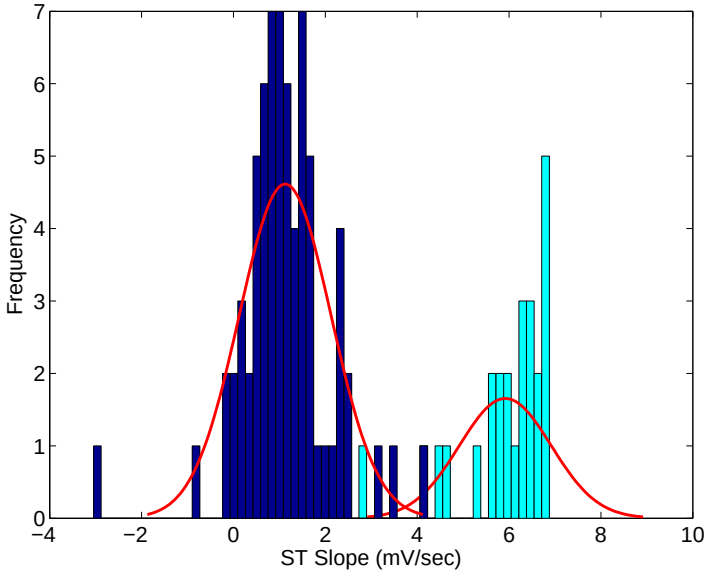


4.1.1 ST level histogram.

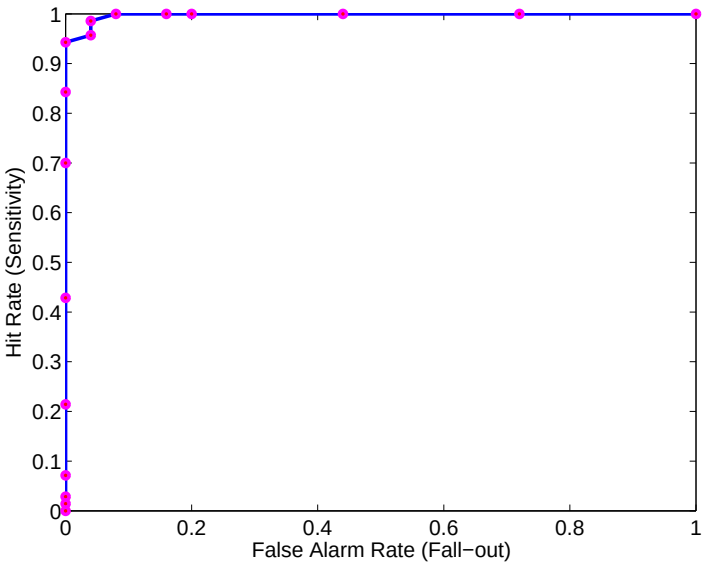


4.1.2 ROC curve of ST level histogram.

Figure 4.1: ST level histogram of basal *(cyan)* and ischaemic *(blue)* ECG and ROC curve of ST level histogram.

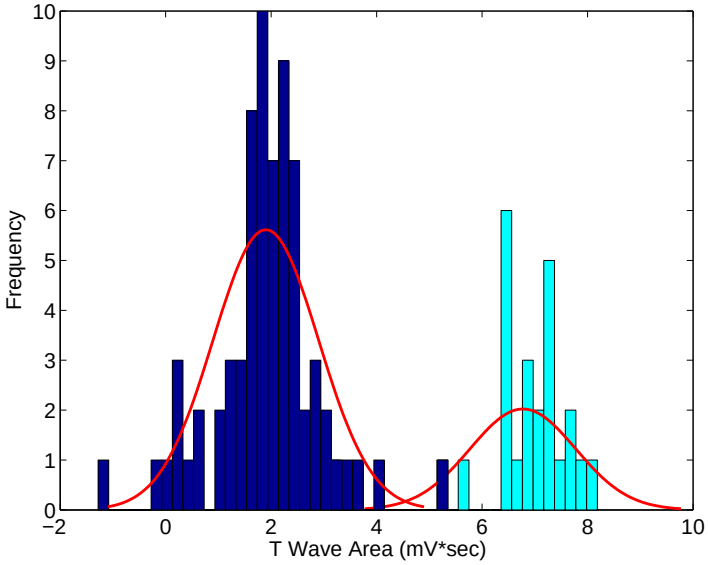


4.2.1 ST slope histogram.

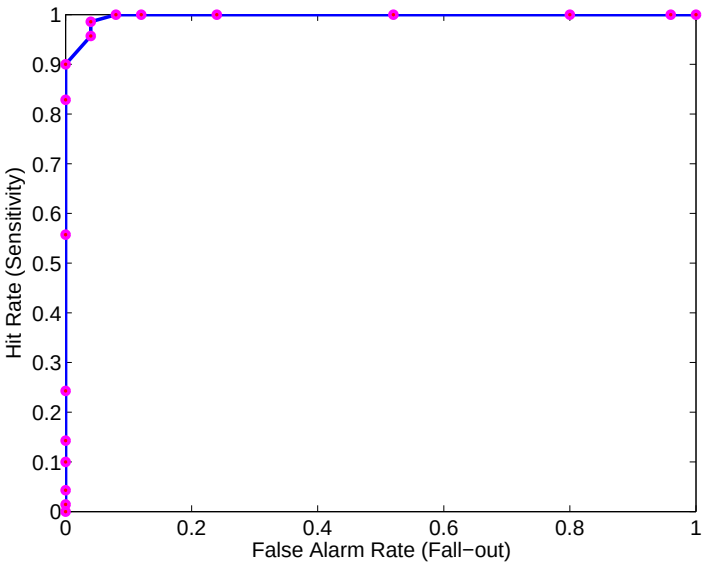


4.2.2 ROC curve of ST slope histogram.

Figure 4.2: ST slope histogram of basal *(cyan)* and ischaemic *(blue)* ECG and ROC curve of ST slope histogram.



4.3.1 T wave area histogram.

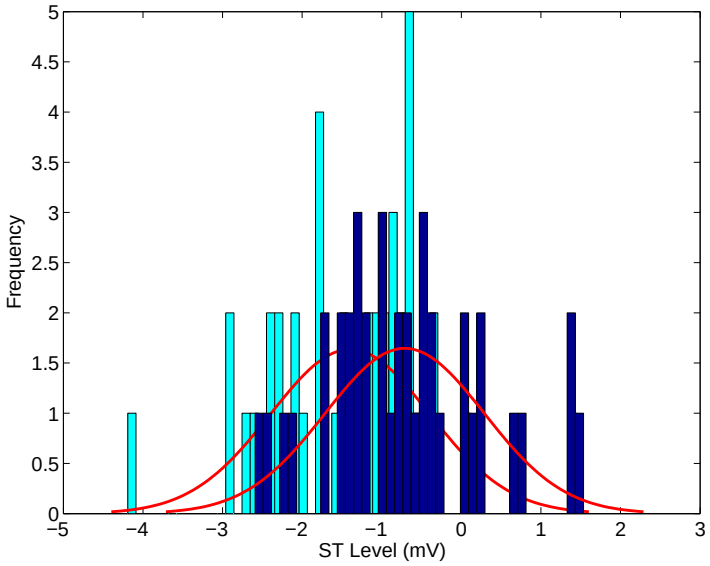


4.3.2 ROC curve of T wave area histogram.

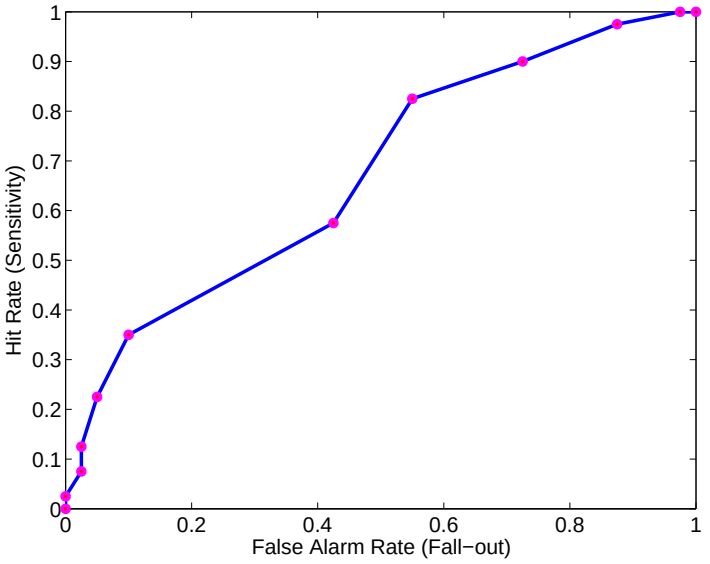
Figure 4.3: T wave area histogram of basal *(cyan)* and ischaemic *(blue)* ECG and ROC curve of T wave area histogram.

However, in Figures 4.4, 4.5 and 4.6 which show histograms of three different ECG features and their corresponding ROC curves for another patient, there is an overlap in ST level, ST slope and T wave area histograms. Since, the overlapping histograms are a sign of features which don't have a significant role in discriminating the two classes from each other, the corresponding ROC curves illustrate a lower classification performance.

Since there are cases where the discriminative strength of an individual feature is low and for the purpose of obtaining higher classification performance results, support vector machines are used to make classification by using joint features.

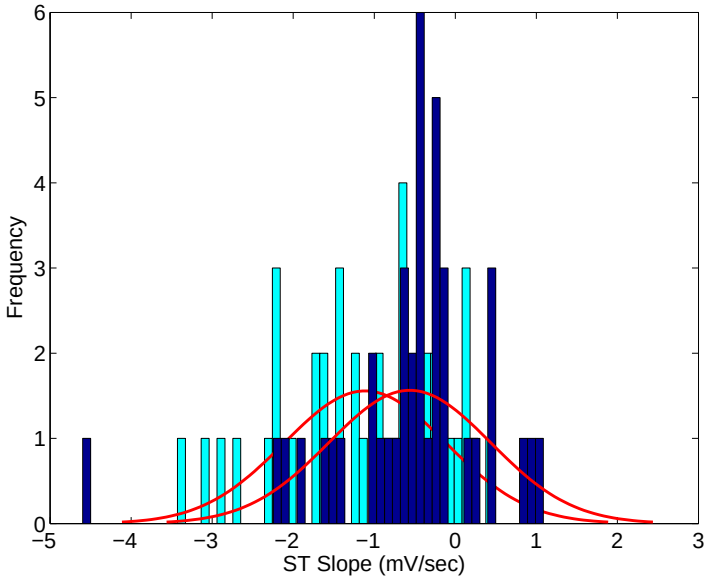


4.4.1 ST level histogram.

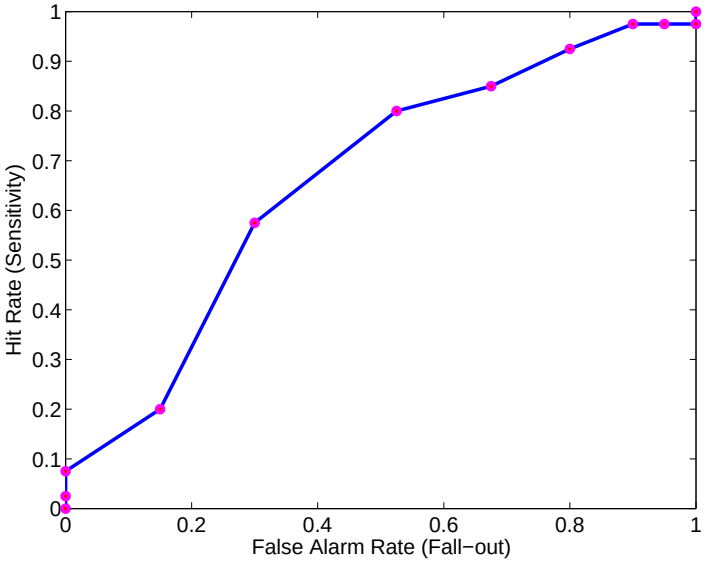


4.4.2 ROC curve of ST level histogram.

Figure 4.4: ST level histogram of basal *(cyan)* and ischaemic *(blue)* ECG and ROC curve of ST level histogram.

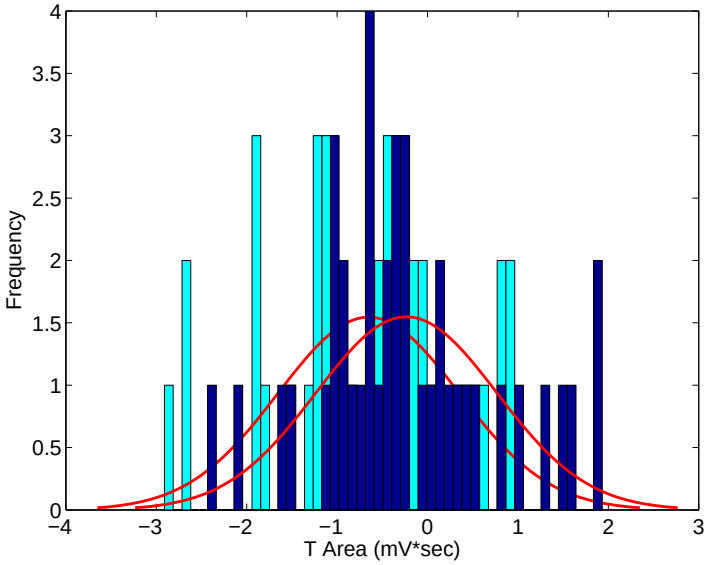


4.5.1 ST slope histogram.

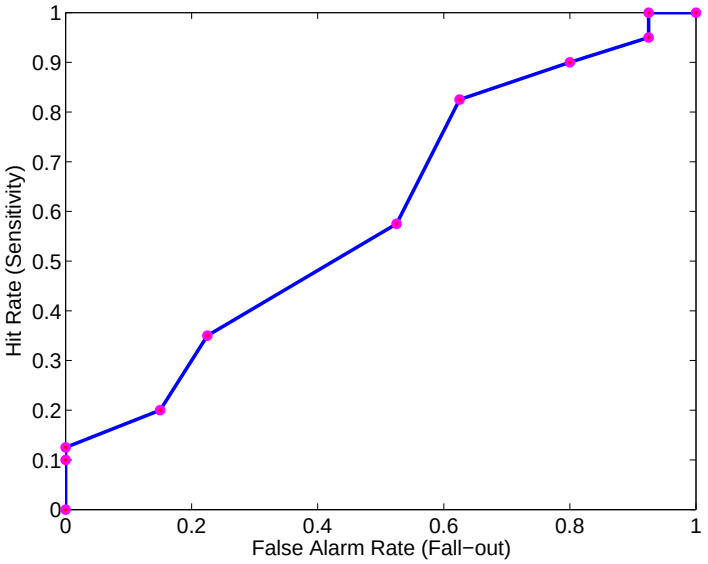


4.5.2 ROC curve of ST slope histogram.

Figure 4.5: ST slope histogram of basal *(cyan)* and ischaemic *(blue)* ECG and ROC curve of ST slope histogram.



4.6.1 T wave area histogram.



4.6.2 ROC curve of T wave area histogram.

Figure 4.6: T wave area histogram of basal *(cyan)* and ischaemic *(blue)* ECG and ROC curve of T wave area histogram.

4.2 ECG Data Classification Using Two and Three Joint Features

In the literature, various techniques have been used for the classification of ECG signals. These techniques include hybrid-fuzzy neural networks [41], high-order spectral analysis methods [42], rule based rough-set decision systems [43] and linear discriminant classifiers [44]. In previous studies, it is shown that SVM approach has higher generalization capability as compared to other classification techniques for the classification of ECG signals. Moreover, it is indicated that due to the high generalization capability of the SVMs, they are less sensitive to the curse of dimensionality and they provide higher classification accuracies [45].

In this thesis, in order to classify two distinct ECG signal by achieving high classification performance results, we developed and implemented algorithms which use support vector machine approach in combination with kernel methods. In cases where the data is linearly separable, we used support vector machines operating with linear kernel functions to map the training data into kernel space. On the other hand, for the data which are not linearly separable, we employed support vector machines operating with radial basis function (RBF) kernel.

The support vector machine is first trained with training the data and then the trained machine is used to classify new data which is the test data. During training process of the SVM classifier, quadratic programming method is employed to find the separating hyperplane or hypersurface.

To obtain satisfactory predictive accuracy, the parameters of linear and radial basis function kernels are tuned. While tuning kernel parameters cross-validation technique is realized, in other words, the data is partitioned into complementary subsets in order to perform the analysis on one subset which is called training set and validate the analysis on the other subset which is called testing set.

In the course of tuning the SVM classifier operating with linear kernel function, a geometric sequence of box constraint parameters within a range are

used for training and the best parameter values are found by using 5-fold cross-validation. Likewise, while tuning the SVM classifier operating with radial basis function (RBF) kernel, a geometric sequence of box constraint and RBF sigma parameters are used for training and the best parameter values are found by using 5-fold cross-validation. The resulting SVM classifiers are specifically designed for each patient.

In order to perform 5-fold cross validation, ECG data is randomly partitioned into 5 equal size subsamples. Of the 5 subsamples, a single subsample is retained as the validation data for testing the model and the remaining 4 subsamples are used as the training data. To reduce variability, the cross-validation process is repeated 5 times and each of the 5 subsamples are used exactly once as the validation data. The advantage of this method is that all observations are used for both training and validation and each observation is used for validation exactly once.

In each cross validation fold, statistical measures of the performance of SVM classifier, such as detection rate (sensitivity), false alarm rate (fall-out), specificity , accuracy, precision (positive predictive value), negative predictive value, classified rate and inconclusive rate are computed. The performance results from the folds are then averaged to produce a single estimation which enables to choose the classifier that gives the best performance results.

Figures 4.7 and 4.9 illustrates an example of the SVM classifier which classifies data by finding the best hyperplane (which has the largest margin between the two classes) that separates all data points of one class represented by 'Basal', from those of the other class represented by 'Ischaemic'. In Figure 4.8, the ROC and precision-recall curves of the SVM classifier operating with linear kernel are illustrated.

On the other hand, Figures 4.10 and 4.12 show an example of a binary classification problem which don't have a simple hyperplane as a useful separating criterion. For these cases, nonlinear transformation is realized by using support vector machines operating with radial basis function kernels (RBF). In Figure 4.11, the ROC and precision-recall curves of the SVM classifier operating with RBF kernel

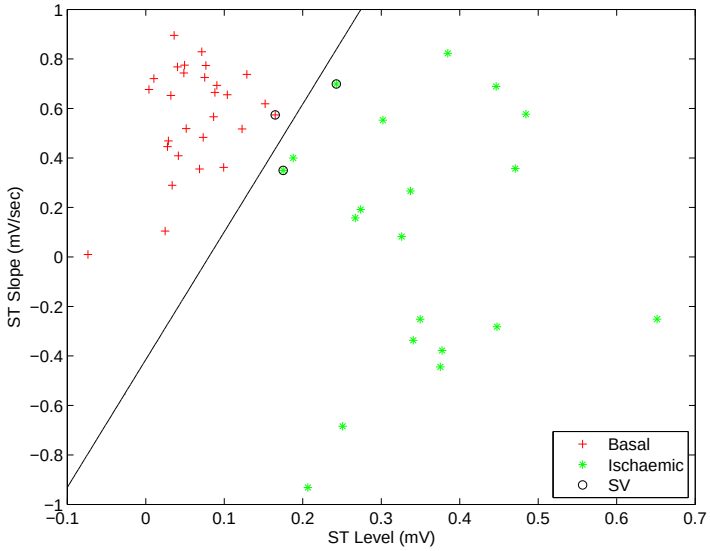
are illustrated.

In Figures 4.7, 4.9, 4.10 and 4.12 the red points which are labeled as 'Basal' stand for basal ECG data, the green points which are labeled as 'Ischaemic' stand for ischaemic ECG data. In the same figures, the pink points represent basal ECG data that are correctly classified and the blue points represent ischaemic ECG data that are correctly classified.

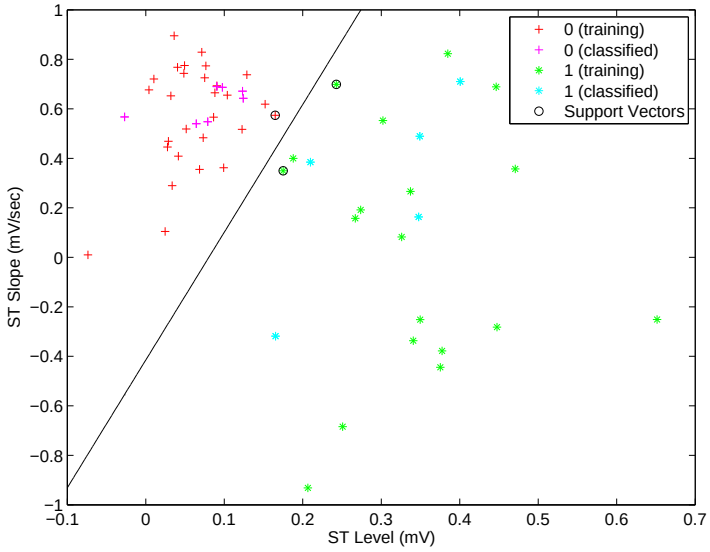
Table 4.1 demonstrates the statistical performance measure results of SVM classifier operating with linear kernel when the optimum kernel parameters and joint features ST level and ST slope are used. Likewise, Table 4.2 shows the statistical performance measure results of the SVM classifier operating with RBF kernel when the optimum kernel parameters and joint features ST level and ST slope are used.

Moreover, the SVM classifiers which use radial basis function kernel and three joint features, which are ST level, ST slope and T wave area, are designed and their statistical performance measures are computed to compare with performance results of the SVM classifiers which employ 2 joint features. When the classification performance results of the SVM classifiers which use two and three joint features are compared, it is observed that SVMs which employ three joint features have higher classification performance results than SVMs which employ two joint features. Figure 4.13 illustrates an example of such a SVM classifier which uses RBF kernel with optimum kernel parameters and three joint features and Table 4.3 shows the increased classification performance results of the classifier.

Table 4.4 and 4.5 show the SVM classifier performance results (%) for optimum linear kernel parameters and two joint features for different patients. Tables 4.6, 4.8, 4.10, 4.12 and 4.7, 4.9, 4.11, 4.13 demonstrate the SVM classifier performance results (%) for optimum radial basis function (RBF) kernel parameters and two and three joint features for different patients, respectively.

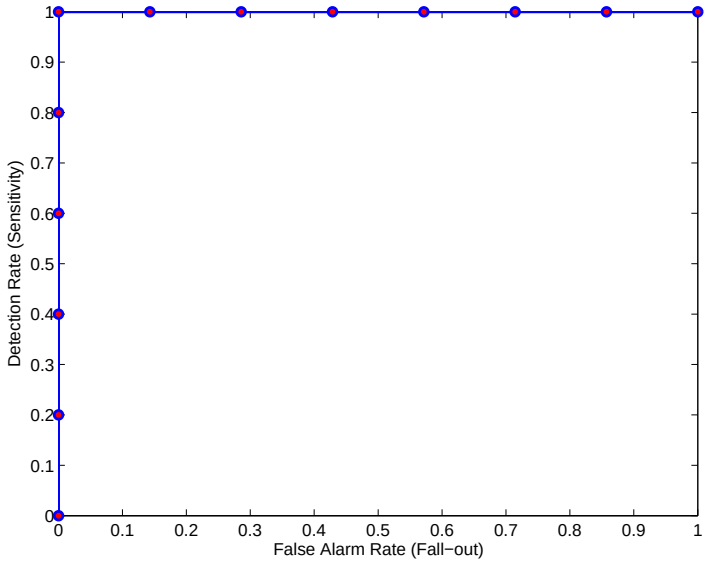


4.7.1 SVM training.

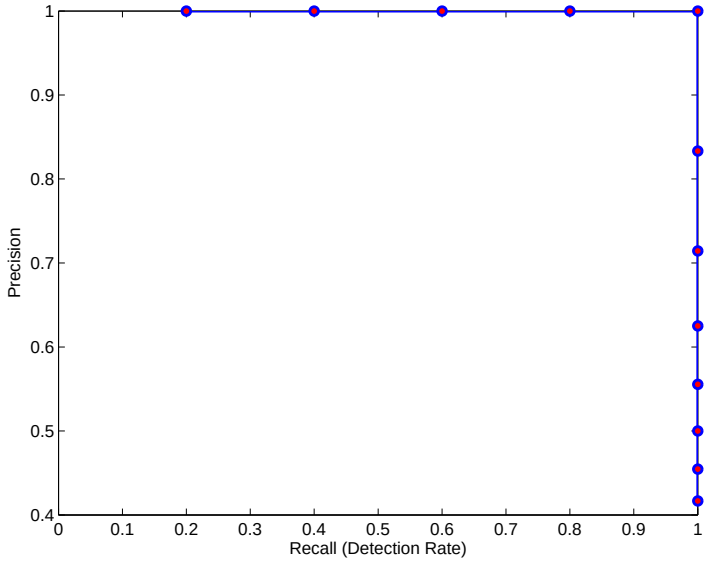


4.7.2 SVM classification.

Figure 4.7: SVM training and classification with linear kernel and joint features ST level and ST slope.



4.8.1 ROC curve.



4.8.2 Precision-recall curve.

Figure 4.8: ROC and precision-recall curves for SVM classification with linear kernel and joint features ST level and ST slope.

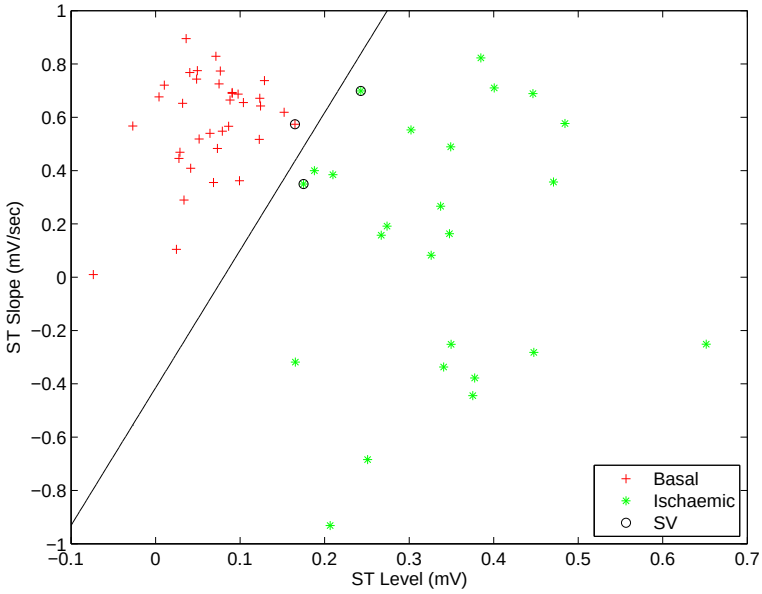
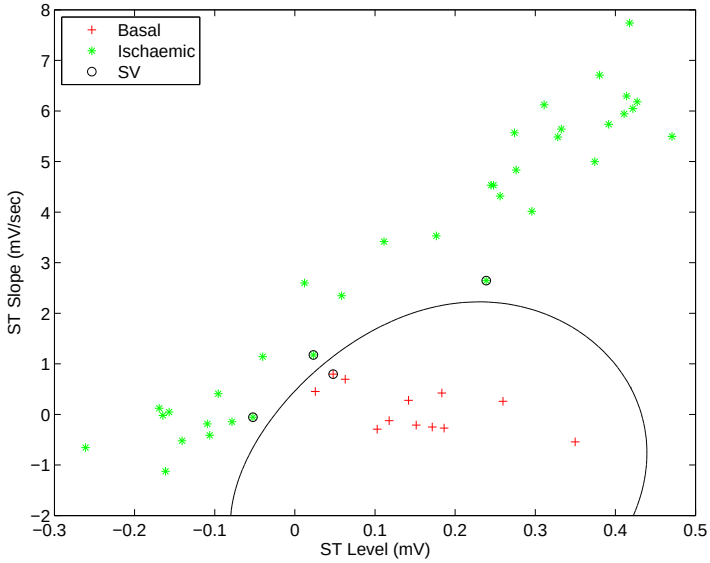


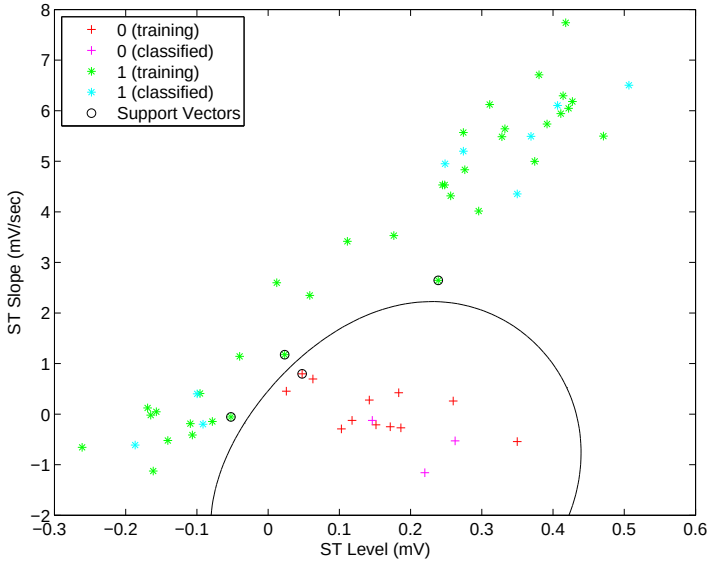
Figure 4.9: SVM training of whole data set with linear kernel and optimum kernel parameters for joint features ST level and ST slope.

Table 4.1: SVM classifier performance results (%) for optimum linear kernel parameters and joint features ST level and ST slope.

	SVM Classifier Performance Results (%)
Detection Rate	100.00
False Alarm Rate	0
Specificity	100.00
Accuracy	100.00
Precision	100.00
Negative Predictive Value	100.00
Classified Rate	100.00
Inconclusive Rate	0
Average Cross-Validation Error	0

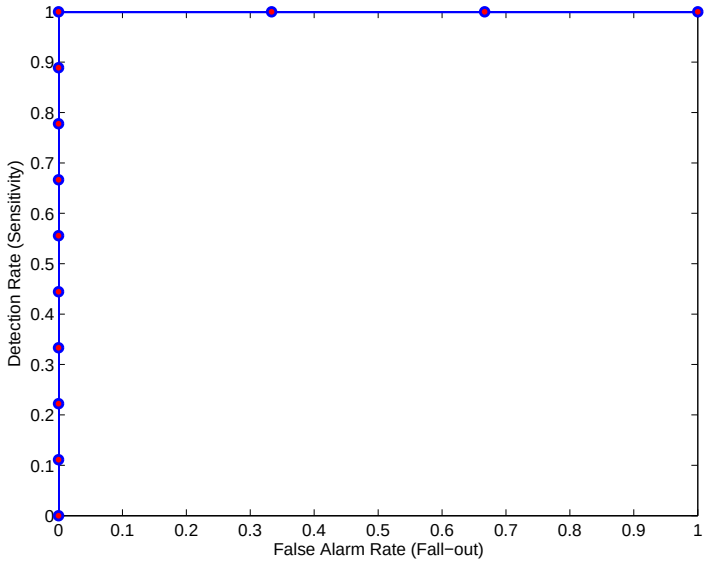


4.10.1 SVM training.

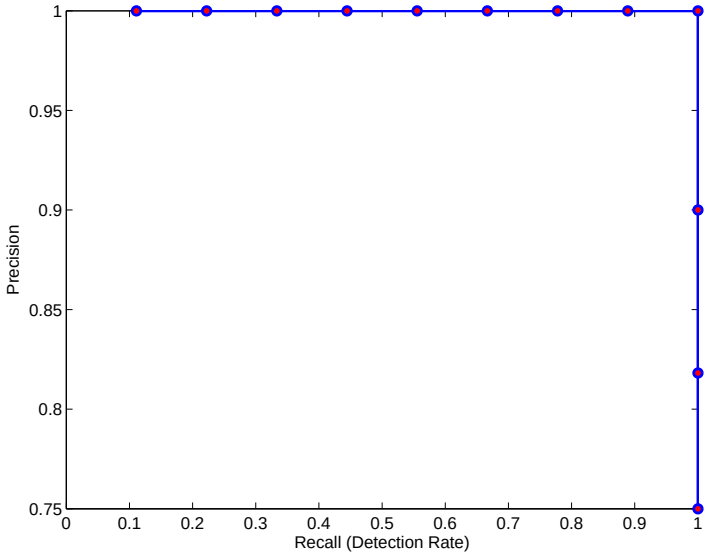


4.10.2 SVM classification.

Figure 4.10: SVM training and classification with radial basis function (RBF) kernel and joint features ST level and ST slope.



4.11.1 ROC curve.



4.11.2 Precision-recall curve.

Figure 4.11: ROC and precision-recall curves for SVM classification with radial basis function (RBF) kernel and joint features ST level and ST slope.

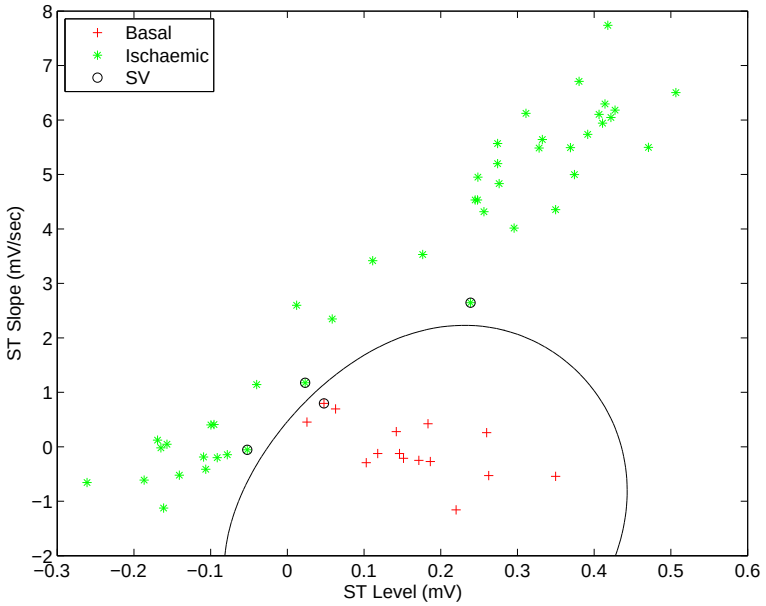


Figure 4.12: SVM training of whole data set with radial basis function (RBF) kernel and optimum kernel parameters for joint features ST level and ST slope.

Table 4.2: SVM classifier performance results (%) for optimum radial basis function (RBF) kernel parameters and joint features ST level and ST slope.

	SVM Classifier Performance Results (%)
Detection Rate	98.55
False Alarm Rate	0
Specificity	100.00
Accuracy	99.14
Precision	100.00
Negative Predictive Value	98.00
Classified Rate	100.00
Inconclusive Rate	0
Average Cross-Validation Error	0.94

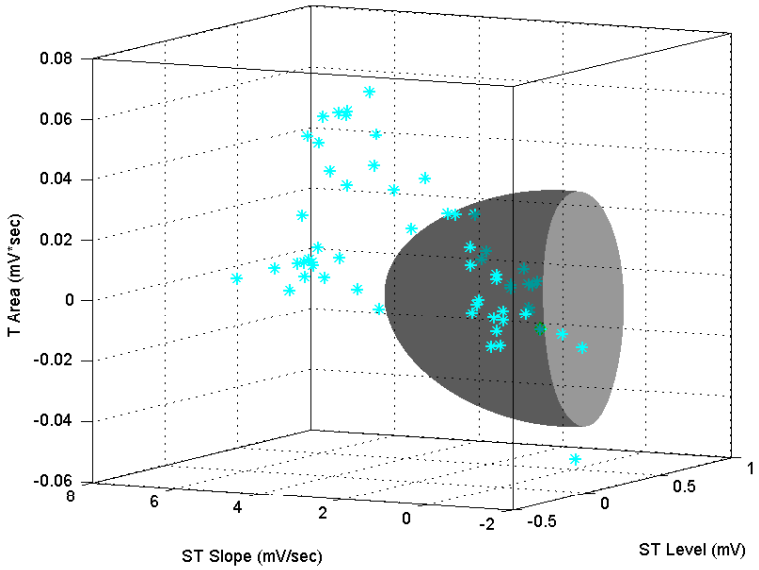


Figure 4.13: SVM training of whole data set with radial basis function (RBF) kernel and optimum kernel parameters for joint features ST level, ST slope and T wave area.

Table 4.3: SVM classifier performance results (%) for optimum radial basis function (RBF) kernel parameters and joint features ST level, ST slope and T wave area.

	SVM Classifier Performance Results (%)
Detection Rate	100.00
False Alarm Rate	0
Specificity	100.00
Accuracy	100.00
Positive Predictive Value	100.00
Negative Predictive Value	100.00
Classified Rate	100.00
Inconclusive Rate	0
Average Cross-Validation Error	0

Table 4.4: SVM classifier performance results (%) for optimum linear kernel parameters and joint features ST level and ST slope.

	Patients							
Performance Measures	**P1**	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**
Detection Rate	**100.00**	99.7	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
False Alarm Rate	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
Specificity	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
Accuracy	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
Positive Predictive Value	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
Negative Predictive Value	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	99.4	**100.00**	**100.00**
Classified Rate	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
Inconclusive Rate	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
Average Cross-Validation Error	**0**	0.07	**0**	**0**	**0**	0.03	**0**	**0**

Table 4.5: SVM classifier performance results (%) for optimum linear kernel parameters and joint features ST level and ST slope.

	Patients							
Performance Measures	**P9**	**P10**	**P11**	**P12**	**P13**	**P14**	**P15**	**P16**
Detection Rate	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
False Alarm Rate	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
Specificity	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
Accuracy	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	99.50	**100.00**
Positive Predictive Value	**100.00**	**100.00**	98.9	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
Negative Predictive Value	**100.00**	**100.00**	96.2	**100.00**	**100.00**	98.00	**100.00**	**100.00**
Classified Rate	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
Inconclusive Rate	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
Average Cross-Validation Error	**0**	**0**	0.08	**0**	**0**	**0**	0.05	**0**

Table 4.6: SVM classifier performance results (%) for optimum radial basis function (RBF) kernel parameters and joint features ST level and ST slope.

	Patients							
Performance Measures	**P17**	**P18**	**P19**	**P20**	**P21**	**P22**	**P23**	**P24**
Detection Rate	**100.00**	**100.00**	95.61	**100.00**	**100.00**	98.55	**100.00**	93.47
False Alarm Rate	**0**	**0**	2.10	**0**	**0**	**0**	**0**	**0**
Specificity	97.48	**100.00**	94.64	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
Accuracy	**100.00**	**100.00**	93.72	**100.00**	**100.00**	99.14	**100.00**	94.18
Positive Predictive Value	**100.00**	**100.00**	97.80	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
Negative Predictive Value	95.21	**100.00**	**100.00**	**100.00**	**100.00**	98.00	**100.00**	90.14
Classified Rate	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
Inconclusive Rate	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
Average Cross-Validation Error	0.16	**0**	4.13	**0**	**0**	0.94	0.33	3.26

Table 4.7: SVM classifier performance results (%) for optimum radial basis function (RBF) kernel parameters and joint features ST level, ST slope and T wave area.

	Patients							
Performance Measures	**P17**	**P18**	**P19**	**P20**	**P21**	**P22**	**P23**	**P24**
Detection Rate	**100.00**	**100.00**	98.56	**100.00**	**100.00**	99.81	**100.00**	96.28
False Alarm Rate	**0**	**0**	0.24	**0**	**0**	**0**	**0**	**0**
Specificity	**100.00**	**100.00**	97.85	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
Accuracy	**100.00**	**100.00**	97.74	**100.00**	**100.00**	**100.00**	**100.00**	97.41
Positive Predictive Value	**100.00**	**100.00**	99.82	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
Negative Predictive Value	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	99.75	**100.00**	94.58
Classified Rate	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
Inconclusive Rate	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
Average Cross-Validation Error	**0**	**0**	0.68	**0**	**0**	0.21	**0**	0.92

Table 4.8: SVM classifier performance results (%) for optimum radial basis function (RBF) kernel parameters and joint features ST level and ST slope.

	Patients							
Performance Measures	**P25**	**P26**	**P27**	**P28**	**P29**	**P30**	**P31**	**P32**
Detection Rate	**100.00**	**100.00**	94.30	**100.00**	88.71	**100.00**	**100.00**	91.74
False Alarm Rate	2.17	**0**	**0**	**0**	5.42	**0**	**0**	3.30
Specificity	97.82	**100.00**	**100.00**	**100.00**	90.18	96.75	**100.00**	92.34
Accuracy	98.04	**100.00**	96.93	**100.00**	92.00	98.60	**100.00**	95.12
Positive Predictive Value	92.00	**100.00**	**100.00**	**100.00**	89.27	**100.00**	**100.00**	98.21
Negative Predictive Value	**100.00**	**100.00**	94.29	**100.00**	92.10	99.10	**100.00**	96.16
Classified Rate	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
Inconclusive Rate	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
Average Cross-Validation Error	3.71	**0**	4.27	**0**	6.21	0.68	**0**	2.43

Table 4.9: SVM classifier performance results (%) for optimum radial basis function (RBF) kernel parameters and joint features ST level, ST slope and T wave area.

	Patients							
Performance Measures	**P25**	**P26**	**P27**	**P28**	**P29**	**P30**	**P31**	**P32**
Detection Rate	**100.00**	**100.00**	97.66	**100.00**	94.75	**100.00**	**100.00**	97.36
False Alarm Rate	**0**	**0**	**0**	**0**	1.91	**0**	**0**	1.10
Specificity	99.76	**100.00**	**100.00**	**100.00**	90.18	98.50	**100.00**	97.34
Accuracy	99.81	**100.00**	98.94	**100.00**	92.00	99.87	**100.00**	98.71
Positive Predictive Value	95.38	**100.00**	**100.00**	**100.00**	89.27	**100.00**	**100.00**	**100.00**
Negative Predictive Value	**100.00**	**100.00**	97.65	**100.00**	92.10	**100.00**	**100.00**	98.15
Classified Rate	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
Inconclusive Rate	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
Average Cross-Validation Error	0.86	**0**	1.52	**0**	2.57	**0**	**0**	0.82

Table 4.10: SVM classifier performance results (%) for optimum radial basis function (RBF) kernel parameters and joint features ST level and ST slope.

	Patients							
Performance Measures	**P33**	**P34**	**P35**	**P36**	**P37**	**P38**	**P39**	**P40**
Detection Rate	**100.00**	94.10	96.85	**100.00**	91.88	**100.00**	**100.00**	89.61
False Alarm Rate	**0**	4.13	4.35	**0**	3.51	**0**	**0**	6.13
Specificity	93.25	**100.00**	**100.00**	97.13	89.74	97.56	**100.00**	94.36
Accuracy	97.86	90.28	98.12	95.28	96.57	**100.00**	**100.00**	93.28
Positive Predictive Value	**100.00**	88.25	**100.00**	97.29	**100.00**	95.13	**100.00**	97.45
Negative Predictive Value	94.18	90.21	92.16	**100.00**	90.84	**100.00**	**100.00**	90.14
Classified Rate	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
Inconclusive Rate	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
Average Cross-Validation Error	2.43	4.36	5.10	1.56	4.62	**0**	**0**	6.87

Table 4.11: SVM classifier performance results (%) for optimum radial basis function (RBF) kernel parameters and joint features ST level, ST slope and T wave area.

	Patients							
Performance Measures	**P33**	**P34**	**P35**	**P36**	**P37**	**P38**	**P39**	**P40**
Detection Rate	**100.00**	98.12	99.10	**100.00**	94.73	**100.00**	**100.00**	92.55
False Alarm Rate	**0**	0.57	1.14	**0**	1.22	**0**	**0**	6.13
Specificity	96.48	**100.00**	**100.00**	**100.00**	93.54	**100.00**	**100.00**	96.87
Accuracy	**100.00**	95.68	**100.00**	97.13	98.76	**100.00**	**100.00**	97.31
Positive Predictive Value	**100.00**	93.85	**100.00**	98.76	**100.00**	**100.00**	**100.00**	**100.00**
Negative Predictive Value	96.75	94.36	94.88	**100.00**	93.72	**100.00**	**100.00**	92.81
Classified Rate	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
Inconclusive Rate	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
Average Cross-Validation Error	**0**	1.42	1.68	0.61	1.74	**0**	**0**	2.10

Table 4.12: SVM classifier performance results (%) for optimum radial basis function (RBF) kernel parameters and joint features ST level and ST slope.

	Patients							
Performance Measures	**P41**	**P42**	**P43**	**P44**	**P45**	**P46**	**P47**	**P48**
Detection Rate	**100.00**	89.53	**100.00**	95.48	92.41	97.11	**100.00**	94.45
False Alarm Rate	**0**	7.13	1.59	2.47	5.36	2.67	**0**	4.12
Specificity	96.21	86.55	92.40	98.41	88.63	**100.00**	**100.00**	90.78
Accuracy	**100.00**	88.24	94.17	99.12	93.10	91.24	**100.00**	96.29
Positive Predictive Value	98.15	84.72	**100.00**	92.34	**100.00**	93.78	97.24	98.17
Negative Predictive Value	95.46	96.35	90.55	**100.00**	88.74	89.86	**100.00**	91.43
Classified Rate	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
Inconclusive Rate	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
Average Cross-Validation Error	**0**	7.58	1.81	3.25	5.74	2.89	**0**	4.80

Table 4.13: SVM classifier performance results (%) for optimum radial basis function (RBF) kernel parameters and joint features ST level, ST slope and and T wave area.

	Patients							
Performance Measures	**P41**	**P42**	**P43**	**P44**	**P45**	**P46**	**P47**	**P48**
Detection Rate	**100.00**	94.56	**100.00**	98.16	96.79	**100.00**	86.17	97.28
False Alarm Rate	**0**	2.04	**0**	0.18	2.26	**0**	8.13	1.13
Specificity	99.10	90.76	95.74	**100.00**	93.51	**100.00**	91.87	93.71
Accuracy	**100.00**	92.75	97.15	**100.00**	98.00	97.84	90.31	99.00
Positive Predictive Value	**100.00**	89.42	**100.00**	97.82	**100.00**	95.72	92.18	**100.00**
Negative Predictive Value	97.48	**100.00**	93.55	**100.00**	92.78	93.61	88.71	94.68
Classified Rate	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
Inconclusive Rate	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
Average Cross-Validation Error	**0**	3.12	0.37	0.21	2.43	0.42	**0**	1.42

Table 4.14 shows the classification performance results of the hardest, easiest and average cases for optimum RBF kernel parameters and three joint features. The performance of the hardest classification case represents the lowest classification result, whereas the performance of the easiest classification case represents the highest classification result. The result of the analysis shows that the performance of the average classification results of all patients is between the easiest and the hardest classification performance results for each performance measure.

Table 4.14: SVM classier performance results (%) of the hardest, easiest and average cases for optimum RBF kernel parameters and three joint features.

Performance Measures	P47	P38	P_{AVG}
Detection Rate	86.17	100.00	94.54
False Alarm Rate	8.13	0	3.11
Specificity	91.87	100.00	95.43
Accuracy	90.31	100.00	94.65
Positive Predictive Value	92.18	100.00	96.74
Negative Predictive Value	88.71	100.00	93.38
Classified Rate	100.00	100.00	100.00
Inconclusive Rate	0	0	0
Average Cross-Validation Error	0	0	1.37

4.3 Decision Fusion of Sequential Segments

Each ECG recording is divided into equal size segments where each segment consists of three different feature vectors of an ECG recording which is 5 seconds long. In order to analyze the change in probability of false alarm as the number of ECG segment increases, probability of false alarm of sequentially increasing segments are theoretically calculated by assuming that each ECG segment is independent and identically distributed. In other words, it is assumed that the performance of the classifier stays the same for sequential segments and the decisions between the segments are statistically independent from each other [46]. The probability of false alarm values of sequentially increasing segments are calculated by using the following equation.

$$P_{FA_{final}} = \sum_{k=N_s/2}^{N_s} \binom{N_s}{k} (P_{FA})^k (1 - P_{FA})^{k-1} \tag{4.1}$$

where N_s is the number of segments and P_{FA} is the probability of false alarm of a single segment.

The probability of false alarm values that are obtained by using SVM classifiers for the same increasing segments are compared with the theoretically calculated performance results. The result of the analysis indicates that probability of false alarm decreases as the number of ECG segment increases. Additionally, it is observed that the observed probability of false alarm values demonstrate very close results to the theoretically calculated probability of false alarm values. Figures 4.15, 4.17 and 4.19 are three different cases which demonstrate the change in probability of false alarm values as the number of ECG segments increases.

Furthermore, probability of detection versus number of ECG segment graphs are plotted in order to analyze the change in probability of detection values as the number of ECG segment increases. The result of the analysis indicates that as the number of ECG segment increases, the discriminative strength of the data also increases, leading to an increase in probability of detection results. Figures 4.14, 4.16 and 4.18 are three different cases which demonstrate the change

in probability of detection values as the number of ECG segments increases.

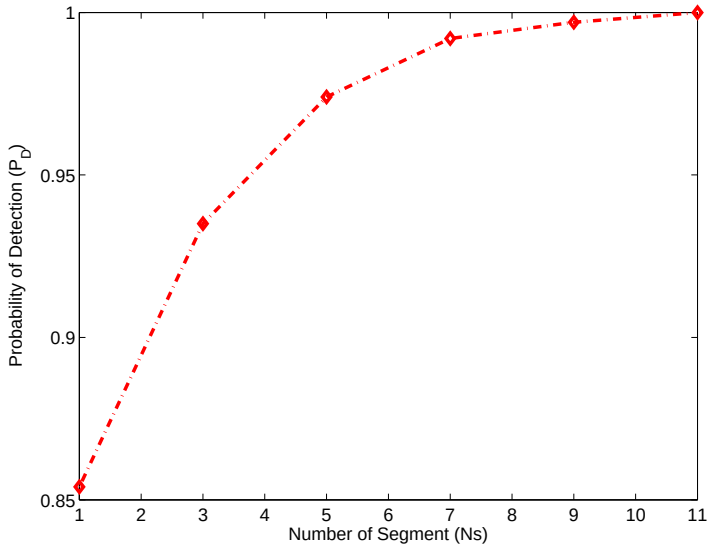


Figure 4.14: Graphical representation of the relation between probability of detection and number of segments.

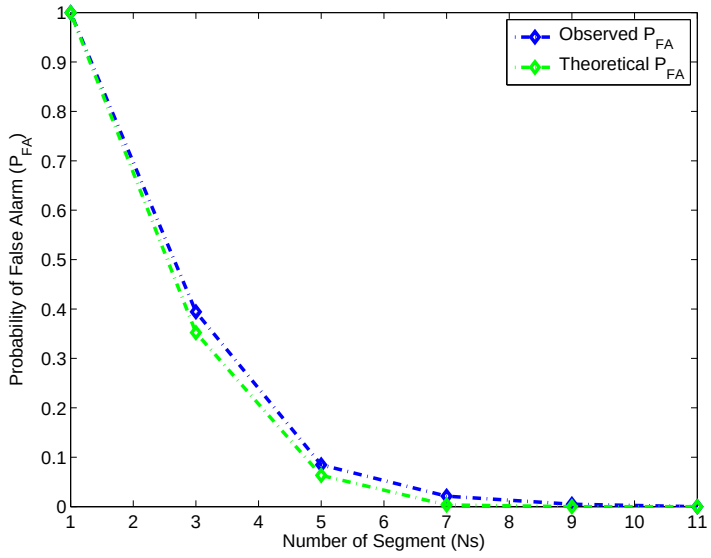


Figure 4.15: Graphical representation of the relation between probability of false alarm and number of segments.

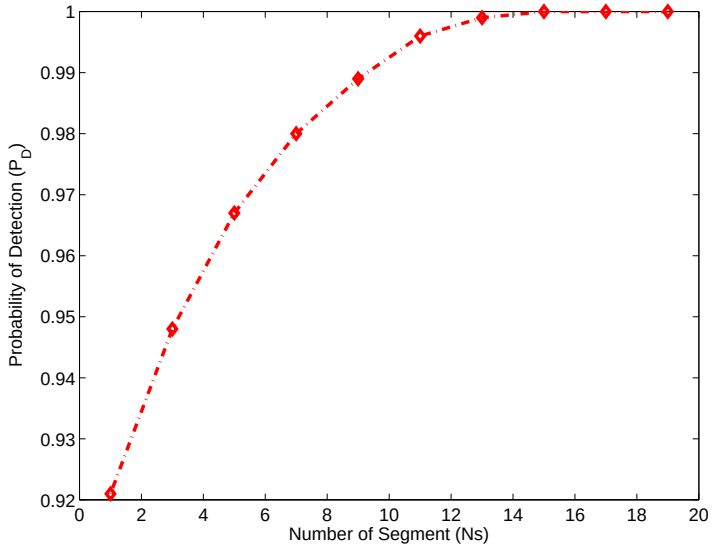


Figure 4.16: Graphical representation of the relation between probability of detection and number of segments.

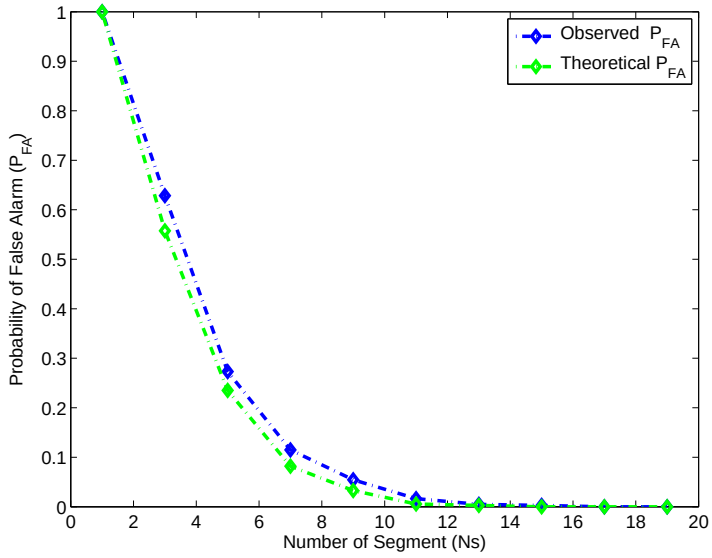


Figure 4.17: Graphical representation of the relation between probability of false alarm and number of segments.

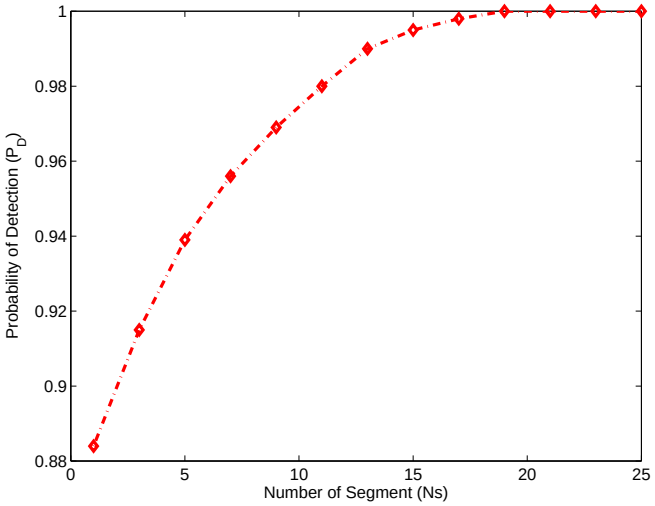


Figure 4.18: Graphical representation of the relation between probability of detection and number of segments.

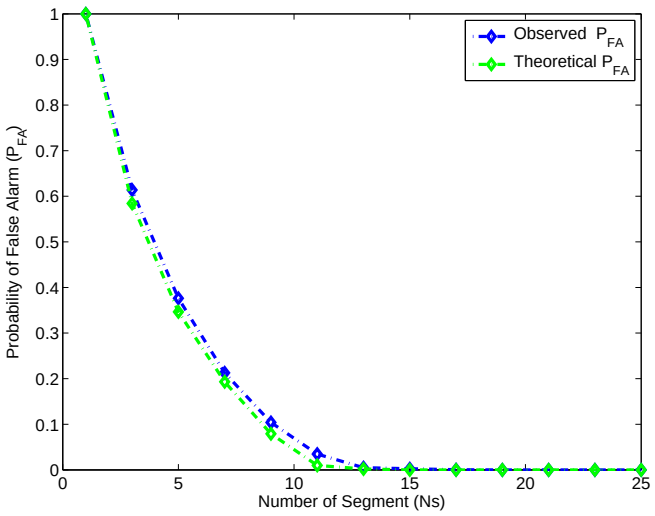


Figure 4.19: Graphical representation of the relation between probability of false alarm and number of segments.

The classification performance results (%) of the hardest classification case for single and multiple segments is shown in Table 4.15. The comparison between the classification performance results of single and multiple segments ($N_s = 11$) show that majority decision fusion on multiple segments improves the classification performance results significantly.

Table 4.15: The classification performance results (%) of the hardest classification case for single and multiple segments.

Performance Measures	Single Segment	Multiple Segment
Detection Rate	86.17	98.65
False Alarm Rate	8.13	0
Specificity	91.87	97.84
Accuracy	90.31	97.43
Positive Predictive Value	92.18	98.12
Negative Predictive Value	88.71	96.53
Classified Rate	100.00	100.00
Inconclusive Rate	0	0

Finally, the decisions which are obtained from three different ECG leads, which are lead I, II and III, are fused by using majority decision fusion rule in order to obtain the overall decision of the patients' ECG recording.

4.4 Gaussian Mixture Model Representation of the Joint Features

The detection of acute coronary syndromes based on ECG recordings of a patient obtained only during healthy stage is also investigated in order to develop a detection technique which can be used in the lack of ischaemic ECGs. For this purpose, the joint probability density function (pdf) of the features ST level and ST slope is represented by a Gaussian mixture model [47].

Figures 4.20 and 4.22 illustrate the joint probability density function of the basal ECG features ST level and ST slope represented by a Gaussian mixture model for two different patients.

Additionally, the detection of the outliers which correspond to acute coronary syndromes is realized by using the developed Neyman-Pearson type of approach [48]. Neyman-Pearson decision strategy is used by computing the average log-likelihood values of the segments and comparing them with a range of different

threshold values as shown in equations 4.2, 4.3 and 4.4.

$$g(\underline{x}) = \frac{1}{(2\Pi)^{\frac{K}{2}}|detR|^{\frac{1}{2}}} exp(-\frac{1}{2}(\underline{x} - \underline{x}_i)^T R^{-1}(\underline{x} - \underline{x}_i)) \tag{4.2}$$

$$P(\underline{x}|H_o) = \frac{1}{N_B}\sum_{i=1}^{N_B} g(\underline{x} - \underline{x}_i) \tag{4.3}$$

$$D = \begin{cases} H_0 & \text{if} \quad P(\underline{x}|H_o) > \Gamma \\ H_1 & \text{if} \quad P(\underline{x}|H_o)) < \Gamma \end{cases} \tag{4.4}$$

For different discrimination threshold values and number of ECG segments (N_s), probability of detection and probability of false alarm values are computed. Figures 4.21 and 4.23 show ROC curves of the Gaussian mixture model based classification for joint features ST level and ST slope and three different ECG segment numbers. As a result of the analysis, it is observed that increasing number of ECG segments provides higher performance results for Gaussian mixture model based classification. Figures 4.21 and 4.23 illustrate the increase in the performance results of the Gaussian mixture model based classification for increasing number of ECG segments.

Table 4.16 illustrates the performance results (%) for SVM and Gaussian mixture model based classification of the hardest and easiest classification cases. The comparison between support vector machine and Gaussian mixture model based classification shows that support vector machines provide higher classification performance results for both the easiest and hardest classification problem. This result can be explained by the fact that GMM based classification uses only basal ECG data, whereas support vector machines use both basal and ischaemic ECG data for classification.

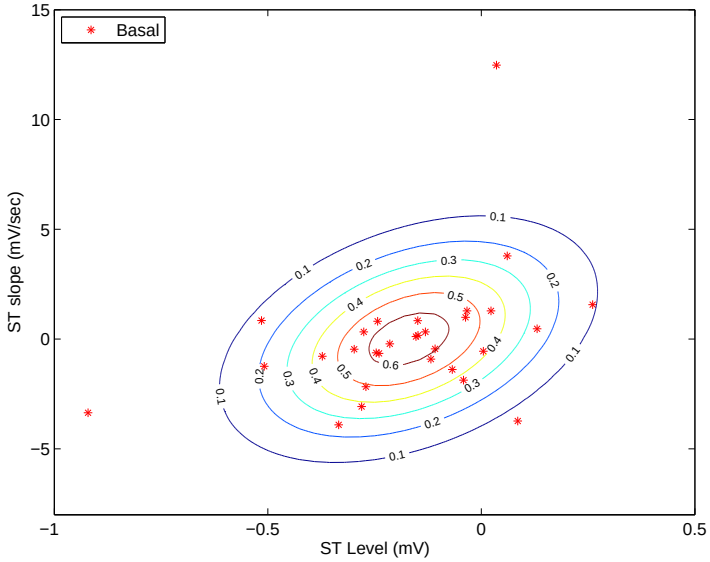


Figure 4.20: Joint probability density function of the basal ECG features ST level and ST slope represented by a Gaussian mixture model.

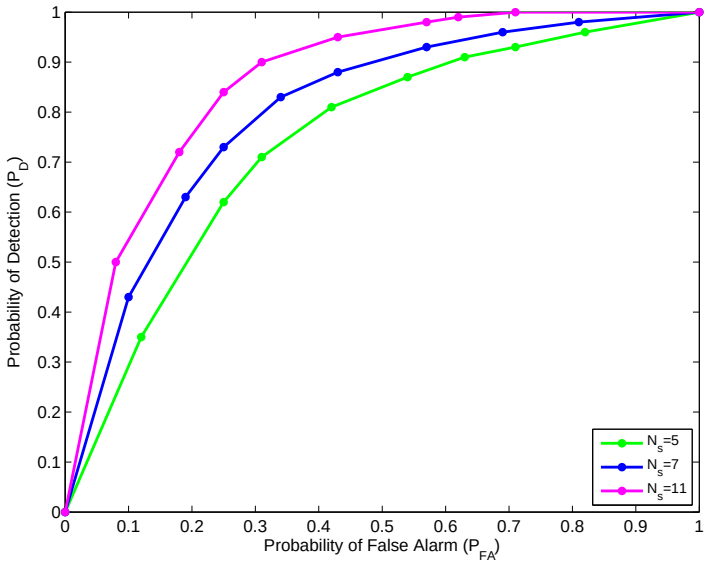


Figure 4.21: ROC curve for Gaussian mixture model based classification with joint features ST level and ST slope for three different segment numbers.

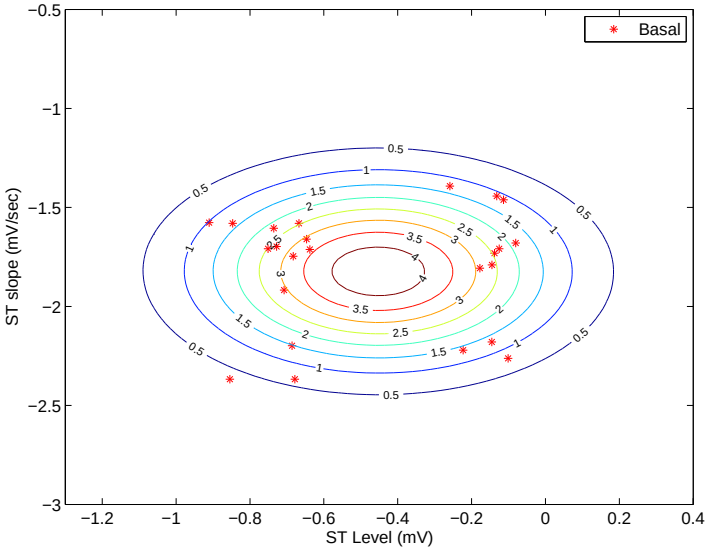


Figure 4.22: Joint probability density function of the basal ECG features ST level and ST slope represented by a Gaussian mixture model.

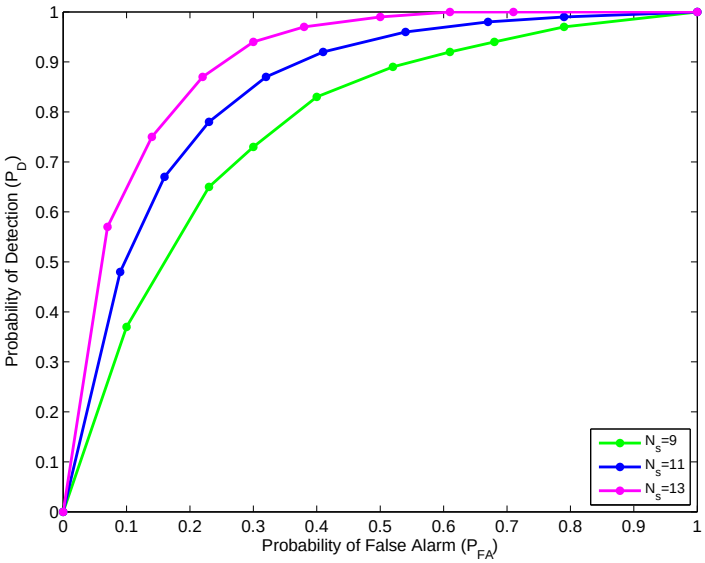


Figure 4.23: ROC curve for Gaussian mixture model based classification with joint features ST level and ST slope for three different segment numbers.

Table 4.16: Classification performance results (%) for SVM operating with RBF kernel and Gaussian mixture model based classification of the hardest and easiest cases.

Performance Measures	SVM_{P47}	GM_{P47}	SVM_{P38}	GM_{P38}
Detection Rate	86.17	83.47	100.00	95.71
False Alarm Rate	8.13	10.23	0	0
Specificity	91.87	88.32	100.00	91.67
Accuracy	90.31	87.94	100.00	94.38
Positive Predictive Value	92.18	90.82	100.00	92.29
Negative Predictive Value	88.71	85.73	100.00	90.85
Classified Rate	100.00	100.00	100.00	100.00
Inconclusive Rate	0	0	0	0

In conclusion, to obtain the highest classification performance results, the proposed detection block uses the SVM classification technique, three joint features which are ST level, ST slope and T wave area and performs majority decision fusion. In Figure 4.24 which illustrates the proposed detection block diagram, the different ECG segments (S_k) come as an input to the previously trained SVM. The classification results (D_k) which are obtained by SVM are then fused by using majority decision fusion rule and the overall decision of the ECG recording is acquired.

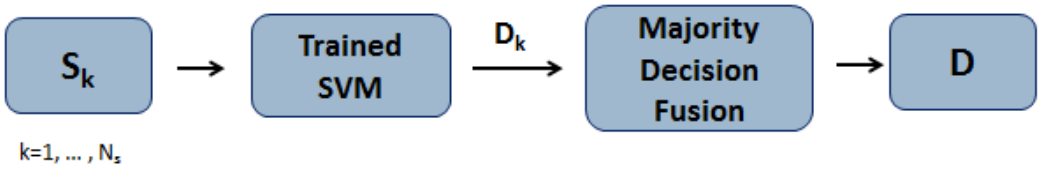


Figure 4.24: The proposed detection block diagram.

Chapter 5

Conclusions

By using a 12-lead ECG device, real ECG recordings are obtained from Gazi University Faculty of Medicine, Cardiology Department. A database which contains 12-lead ECG data of 48 patient is constituted. For each patient, there exists two different types of ECG data, which correspond to basal and ischaemic ECG signals.

Several signal analysis techniques are developed and implemented to 12-lead ECG data in order to realize the preprocessing step of ECG signal. During the preprocessing step, low frequency baseline wander, high frequency motion artifact and 50 Hz line voltage are removed from ECG signal by using a cascade of band-pass and notch filter combinations [16]. Before the application of band-pass filter, ECG signal is first multiplied by a raised cosine window in order to avoid Gibbs phenomenon, which is caused by the abrupt changes at the onset and at the end of the signal.

After preprocessing step, several important properties of ECG signals such as the QRS peaks, isoelectric line, J point, ST segment and T wave are detected. By using the detected properties, the extraction of ECG features which are critical for the detection of acute coronary syndromes, such as ST segment level, ST segment slope and T wave area, is realized.

For the one dimensional classification of two different ECG data, histogram graphs illustrating the graphical representation of the data distribution for basal and ischaemic ECG signals are obtained individually for three ECG features. By fitting a normal density function to the histograms of both ECG data, the estimate of underlying probability distribution is identified. To illustrate the performance of the classification by using histograms, hit rate and false alarm rate performance measures are calculated for different range of discrimination threshold values.

As a result of histogram classification analysis, it is seen that, in some cases, ST level, ST slope and T wave area histograms do not overlap and the their corresponding ROC curves illustrate good classification performance results. This result indicates that these non-overlapping histograms represent features which have a significant role in discriminating the two classes from each other. On the other hand, in most of the cases, it is observed that there are overlapping regions in ST level, ST slope or T wave area histograms. Since the overlapping histograms are a sign of features which don't have a significant role in discriminating the two classes from each other, the corresponding ROC curves showed lower classification performance results.

Since there are cases where the discriminative strength of an individual feature is low and for the purpose of obtaining higher classification performance results, support vector machines operating with kernels are designed to make classification by using joint features.

Support vector machines are designed specifically for each patient by tuning the kernel parameters in order to obtain the optimal classification results. In cases where the data is linearly separable, we used support vector machines operating with linear kernel functions to map the training data into the kernel space. On the other hand, for the data which are not linearly separable, we employed support vector machines operating with radial basis function (RBF) kernel [49]. To obtain satisfactory performance results, the parameters of linear and radial basis function kernels are tuned by implementing 5-fold cross validation.

Statistical measures of performance of SVM classifiers, such as detection rate

(sensitivity), false alarm rate (fall-out), specificity , accuracy, precision (positive predictive value), negative predictive value, classified rate and inconclusive rate are computed for each case.

For each case, statistical performance measure results of SVM classifier operating with linear kernel, optimum kernel parameters and joint features of ST level and ST slope showed almost %100 sensitivity, specificity, accuracy, positive and negative predictive values.

Statistical performance measure results of SVM classifier operating with RBF kernel, optimum kernel parameters and joint features of ST level and ST slope showed sensitivity, specificity, accuracy, positive and negative predictive values between %86 and %100.

Moreover, SVM classifiers which use radial basis function kernel and three joint features, which are ST level, ST slope and T wave area, demonstrated performance results between %90 and %100. When the classification performance results of SVM classifiers which use two and three joint features are compared, it is observed that SVMs which employ three joint features have higher classification performance results than SVMs which employ two joint features.

Furthermore, probability of detection versus number of ECG segment graphs are plotted in order to analyze the change in probability of detection values as the number of ECG data segment increases. Consequently, it is observed that as the number of ECG segment increases, the discriminative strength of the data also increases, leading to an increase in probability of detection results.

Likewise, probability of false alarm versus number of ECG segment graphs are plotted in order to analyze the change in probability of false alarm values as the number of ECG segment increases. Furthermore, probability of false alarm of sequentially increasing segments are theoretically calculated by assuming that each ECG segment is independent and identically distributed. The results show that increasing number of ECG segment results in a decrease in the probability of false alarm values. In addition, the comparison between observed and theoretically calculated probability of false alarm values demonstrated very approximate

results.

Additionally, the detection of acute coronary syndromes based on ECG recordings of a patient obtained only during healthy stage is also investigated in order to develop a detection technique which can be used in the absence of ischaemic ECGs. For this purpose, the joint probability density function (pdf) of the features of ST level and ST slope is represented by a Gaussian mixture model [47].

Moreover, the detection of the outliers which correspond to acute coronary syndromes is realized by using the developed Neyman-Pearson type of approach [48]. Neyman-Pearson decision strategy is used by computing the average log-likelihood values of the segments and comparing them with a range of different threshold values. For different discrimination threshold values and number of ECG segments (N_s), probability of detection and probability of false alarm values are computed and ROC curves are obtained. As a result of the analysis, it is observed that increasing number of ECG segments provide higher performance results for Gaussian mixture model based classification.

In addition, the comparison between support vector machine and Gaussian mixture model based classification showed that support vector machines provide higher classification performance results for ECG signal. This result can be explained by the fact that support vector machines use both basal and ischaemic ECG data, whereas GMM based classification uses only basal ECG data.

In conclusion, the proposed detection block diagram uses the SVM classification technique, three joint features and performs majority decision fusion to improve performance of the classification.

Bibliography

[1] B. Hamilton, E. Kwakyi, A. Koyfman, and M. Foran, "Diagnosis and management of acute coronary syndrome: Diagnostic et prise en charge du syndrome coronarien aigu," *African Journal of Emergency Medicine*, vol. 3, no. 3, pp. 124 – 133, 2013.

[2] A. S. Go, D. Mozaffarian, V. L. Roger, E. J. Benjamin, J. D. Berry, M. J. Blaha, S. Dai, E. S. Ford, C. S. Fox, S. Franco, H. J. Fullerton, C. Gillespie, S. M. Hailpern, J. A. Heit, V. J. Howard, M. D. Huffman, S. E. Judd, B. M. Kissela, S. J. Kittner, D. T. Lackland, J. H. Lichtman, L. D. Lisabeth, R. H. Mackey, D. J. Magid, G. M. Marcus, A. Marelli, D. B. Matchar, D. K. McGuire, E. R. Mohler, C. S. Moy, M. E. Mussolino, R. W. Neumar, G. Nichol, D. K. Pandey, N. P. Paynter, M. J. Reeves, P. D. Sorlie, J. Stein, A. Towfighi, T. N. Turan, S. S. Virani, N. D. Wong, D. Woo, and M. B. Turner, "Heart disease and stroke statistics2014 update: A report from the american heart association," *Circulation*, vol. 129, no. 3, pp. e28–e292, 2014.

[3] H. H. Ting, H. M. Krumholz, E. H. Bradley, D. C. Cone, J. P. Curtis, B. J. Drew, J. M. Field, W. J. French, W. B. Gibler, D. C. Goff, A. K. Jacobs, B. K. Nallamothu, R. E. O'Connor, and J. D. Schuur, "Implementation and integration of prehospital ecgs into systems of care for acute coronary syndrome: A scientific statement from the american heart association interdisciplinary council on quality of care and outcomes research, emergency cardiovascular care committee, council on cardiovascular nursing, and council on clinical cardiology," *Circulation*, vol. 118, no. 10, pp. 1066–1079, 2008.

[4] J. L. Anderson, C. D. Adams, E. M. Antman, C. R. Bridges, R. M. Califf, D. E. Casey, W. E. Chavey, F. M. Fesmire, J. S. Hochman, T. N. Levin, A. M. Lincoff, E. D. Peterson, P. Theroux, N. K. Wenger, and R. S. Wright, "2011 accf/aha focused update incorporated into the acc/aha 2007 guidelines for the management of patients with unstable angina/nonst-elevation myocardial infarction: A report of the american college of cardiology foundation/american heart association task force on practice guidelines," *Circulation*, vol. 123, no. 18, pp. e426–e579, 2011.

[5] P. Iaizzo, *Handbook of Cardiac Anatomy, Physiology, and Devices*. Current clinical oncology, Humana Press, 2005.

[6] W. Gentry and R. Williams, *Psychological aspects of myocardial infarction and coronary care*. Mosby, 1979.

[7] J. P. Berger, T. Buclin, E. Haller, G. Van Melle, and B. Yersin, "Right arm involvement and pain extension can help to differentiate coronary diseases from chest pain of other origin: a prospective emergency ward study of 278 consecutive patients admitted for chest pain," *Journal of Internal Medicine*, vol. 227, no. 3, pp. 165–172, 1990.

[8] D. Foster, *Twelve-Lead Electrocardiography: Theory and Interpretation*. Springer, 2007.

[9] G. Wagner, *Marriott's Practical Electrocardiography*. LWW medical book collection, Wolters Kluwer Health/Lippincott Williams & Wilkins, 2008.

[10] C. Physiology, "Electrocardiogram leads," November 2009.

[11] U. of Maryland School of Medicine Emergency Medicine Interest Group, "Lead placement," August 2009.

[12] P. J. Leisy, R. R. Coeytaux, G. S. Wagner, E. H. Chung, A. J. McBroom, C. L. Green, J. W. W. Jr., and G. D. Sanders, "Ecg-based signal analysis technologies for evaluating patients with acute coronary syndrome: A systematic review," *Journal of Electrocardiology*, vol. 46, no. 2, pp. 92 – 97, 2013.

[13] M. Papouchado, P. R. Walker, M. A. James, and L. M. Clarke, "Fundamental differences between the standard 12 lead electrocardiograph and the modified mason likar exercise lead system," *European Heart Journal*, vol. 8, no. 7, pp. 725–733, 1987.

[14] S. S, A. D, G. CB, and et al, "Prognostic value of the admission electrocardiogram in acute coronary syndromes," *JAMA*, vol. 281, no. 8, pp. 707–713, 1999.

[15] G. D. Clifford, F. Azuaje, and P. McSharry, *Advanced Methods And Tools for ECG Data Analysis*. Norwood, MA, USA: Artech House, Inc., 2006.

[16] F. Buenda-Fuentes, M. A. Arnau-Vives, and A. Arnau-Vives, "High-bandpass filters in electrocardiography: Source of error in the interpretation of the st segment," *ISRN Cardiology*, vol. 2012, no. 12, pp. 725–733, 2012.

[17] C. Ye, B. Kumar, and M. Coimbra, "Heartbeat classification using morphological and dynamic features of ecg signals," *IEEE Trans. Biomed. Eng.*, vol. 59, pp. 2930–2941, Oct 2012.

[18] N. Acir, "A support vector machine classifier algorithm based on a perturbation method and its application to ecg beat recognition systems," *Expert Systems with Applications*, vol. 31, no. 1, pp. 150 – 158, 2006.

[19] E. Zanaty, "Support vector machines (svms) versus multilayer perception (mlp) in data classification," *Egyptian Informatics Journal*, vol. 13, no. 3, pp. 177 – 183, 2012.

[20] B. M. Asl, S. K. Setarehdan, and M. Mohebbi, "Support vector machine-based arrhythmia classification using reduced features of heart rate variability signal," *Artificial Intelligence in Medicine*, vol. 44, no. 1, pp. 51 – 64, 2008.

[21] Q. Zhao and L. Zhang, "Ecg feature extraction and classification using wavelet transform and support vector machines," in *Neural Networks and Brain, 2005. ICNN B '05. International Conference on*, vol. 2, pp. 1089–1092, Oct 2005.

[22] C.-W. Hsu and C.-J. Lin, "A comparison of methods for multiclass support vector machines," *IEEE Trans. Neural Netw.*, vol. 13, pp. 415–425, Mar 2002.

[23] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.

[24] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press, 2000.

[25] S. Abe, *Support Vector Machines for Pattern Classification*. Advances in Computer Vision and Pattern Recognition, Springer, 2010.

[26] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, COLT '92, (New York, NY, USA), pp. 144–152, ACM, 1992.

[27] S. Hongzong, W. Tao, Y. Xiaojun, L. Huanxiang, H. Zhide, L. Mancang, and F. BoTao, "Support vector machines classification for discriminating coronary heart disease patients from non-coronary heart disease," *West Indian Medical Journal*, vol. 56, pp. 451 – 457, 10 2007.

[28] C. W. Hsu, C. C. Chang, and C. J. Lin, *A practical guide to support vector classification*. Department of Computer Science and Information Engineering, National Taiwan University, Taipei, Taiwan, 2003.

[29] S. Amari and S. Wu, "Improving support vector machine classifiers by modifying kernel functions," *Neural Networks*, vol. 12, no. 6, pp. 783 – 789, 1999.

[30] K. Muller, S. Mika, G. Ratsch, K. Tsuda, and B. Scholkopf, "An introduction to kernel-based learning algorithms," *IEEE Trans. Neural Netw.*, vol. 12, pp. 181–201, Mar 2001.

[31] B. Baesens, S. Viaene, T. Van Gestel, J. A. K. Suykens, G. Dedene, B. De Moor, and J. Vanthienen, "An empirical assessment of kernel type performance for least squares support vector machine classifiers," in *Knowledge-Based Intelligent Engineering Systems and Allied Technologies, 2000. Proceedings. Fourth International Conference on*, vol. 1, pp. 313–316 vol.1, 2000.

[32] L. Dioşan, A. Rogozan, and J.-P. Pecuchet, "Improving classification performance of support vector machine by genetically optimising kernel shape and hyper-parameters," *Applied Intelligence*, vol. 36, pp. 280–294, Mar. 2012.

[33] Y.-W. Chang, C.-J. Hsieh, K.-W. Chang, M. Ringgaard, and C.-J. Lin, "Training and testing low-degree polynomial data mappings via linear svm," *J. Mach. Learn. Res.*, vol. 11, pp. 1471–1490, Aug. 2010.

[34] I. Maglogiannis, *Emerging Artificial Intelligence Applications in Computer Engineering: Real Word AI Systems with Applications in EHealth, HCI, Information Retrieval and Pervasive Technologies*. Frontiers in artificial intelligence and applications, IOS Press, 2007.

[35] S. V. Stehman, "Selecting and interpreting measures of thematic classification accuracy," *Remote Sensing of Environment*, vol. 62, no. 1, pp. 77 – 89, 1997.

[36] T. Fawcett, "An introduction to roc analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861 – 874, 2006. ROC Analysis in Pattern Recognition.

[37] G. Forman, "An extensive empirical study of feature selection metrics for text classification," *J. Mach. Learn. Res.*, vol. 3, pp. 1289–1305, Mar. 2003.

[38] A. Khazaee and A. Ebrahimzadeh, "Classification of electrocardiogram signals with support vector machines and genetic algorithms using power spectral features," *Biomedical Signal Processing and Control*, vol. 5, no. 4, pp. 252 – 263, 2010.

[39] J. Fogarty, R. S. Baker, and S. E. Hudson, "Case studies in the use of roc curve analysis for sensor-based estimates in human computer interaction," in *Proceedings of Graphics Interface 2005*, GI '05, (School of Computer Science,

University of Waterloo, Waterloo, Ontario, Canada), pp. 129–136, Canadian Human-Computer Communications Society, 2005.

[40] O. Chapelle, P. Haffner, and V. Vapnik, "Support vector machines for histogram-based image classification," *IEEE Trans. Neural Netw.*, vol. 10, pp. 1055–1064, Sep 1999.

[41] S. Osowski and T. H. Linh, "Ecg beat recognition using fuzzy hybrid neural network," *IEEE Trans. Biomed. Eng.*, vol. 48, pp. 1265–1271, Nov 2001.

[42] L. Khadra, A. Al-Fahoum, and S. Binajjaj, "A quantitative analysis approach for cardiac arrhythmia classification using higher order spectral techniques," *IEEE Trans. Biomed. Eng.*, vol. 52, pp. 1840–1845, Nov 2005.

[43] S. Mitra, M. Mitra, and B. Chaudhuri, "A rough-set-based inference engine for ecg classification," *IEEE Trans. Instrum. Meas.*, vol. 55, pp. 2198–2206, Dec 2006.

[44] P. de Chazal, M. O'Dwyer, and R. Reilly, "Automatic classification of heartbeats using ecg morphology and heartbeat interval features," *IEEE Trans. Biomed. Eng.*, vol. 51, pp. 1196–1206, July 2004.

[45] F. Melgani and Y. Bazi, "Classification of electrocardiogram signals with support vector machines and particle swarm optimization," *IEEE Trans. on Inf. Technol. Biomed.*, vol. 12, pp. 667–677, Sept 2008.

[46] D. Li, K. Wong, Y. H. Hu, and A. Sayeed, "Detection, classification, and tracking of targets," *IEEE Signal Process. Mag.*, vol. 19, pp. 17–29, Mar 2002.

[47] P.-C. Chang, J.-J. Lin, J.-C. Hsieh, and J. Weng, "Myocardial infarction classification with multi-lead ecg using hidden markov models and gaussian mixture models," *Applied Soft Computing*, vol. 12, no. 10, pp. 3165 – 3175, 2012.

[48] Y. Leschyshyn, B. Yavorskyy, T. Rafa, and G. Shadrina, "Probability integral parameters identification at heart rate variability type detection by neyman - pearson criterion," in *Modern Problems of Radio Engineering,*

Telecommunications, and Computer Science, 2006. TCSET 2006. International Conference, pp. 628–629, Feb 2006.

[49] T. Ince, S. Kiranyaz, and M. Gabbouj, "A generic and robust system for automated patient-specific classification of ecg signals," *IEEE Trans. Biomed. Eng.*, vol. 56, pp. 1415–1426, May 2009.