AUTOMATED DIAGNOSIS METHODS FOR HEART DISEASES USING FLEXIBLE ANALYTIC WAVELET TRANSFORM

Ph.D. Thesis

By

MOHIT KUMAR

DISCIPLINE OF ELECTRICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY INDORE
FEBRUARY 2018

AUTOMATED DIAGNOSIS METHODS FOR HEART DISEASES USING FLEXIBLE ANALYTIC WAVELET TRANSFORM

A THESIS

Submitted in partial fulfillment of the

requirements for the award of the degree

of

DOCTOR OF PHILOSOPHY

by

MOHIT KUMAR

DISCIPLINE OF ELECTRICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY INDORE

FEBRUARY 2018

INDIAN INSTITUTE OF TECHNOLOGY INDORE

CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in the thesis entitled "**AUTOMATED DIAGNOSIS METHODS FOR HEART DISEASES USING FLEXIBLE ANALYTIC WAVELET TRANSFORM**" in the partial fulfillment of the requirements for the award of the degree of **DOCTOR OF PHILOSOPHY** and submitted in the **DISCIPLINE OF ELECTRICAL ENGINEERING, Indian Institute of Technology Indore**, is an authentic record of my own work carried out during the time period from December 2014 to February 2018 under the supervision of Prof. Ram Bilas Pachori, Indian Institute of Technology Indore, Indore, India.

The matter presented in this thesis has not been submitted by me for the award of any other degree of this or any other institute.

Signature of the student with date
(MOHIT KUMAR)

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

Signature of Thesis Supervisor with date
(PROF. RAM BILAS PACHORI)

MOHIT KUMAR has successfully given his Ph.D. Oral Examination held on

Signature of Chairperson (OEB) Date:	Signature External Examiner Date:	Signature of Thesis Supervisor Date:
Signature of PSPC Member:1 Date:	Signature of PSPC Member:2 Date:	Signature of Convener, DPGC Date:

Signature of Head of Discipline
Date:

ACKNOWLEDGEMENTS

I am immensely grateful to my supervisor Prof. Ram Bilas Pachori for providing inspiration, guidance, and support throughout my research work. It was my pleasure to work under the supervision of such an esteemed professor.

I would like to express my sincere gratitude to Prof. U. R. Acharya, Ngee Ann Polytechnic, Singapore, for providing valuable suggestions and the dataset of coronary artery disease for my research work. I am also grateful towards my PSPC committee members Dr. Santosh Kumar Vishvakarma and Dr. Ritunesh Kumar for their interesting discussions and suggestions during my research. I would like to thank my friends Rajeev Sharma, Abhay Upadhyay, Pankaj Sharma, Dipika Sharma, Rishi Raj Sharma, Shishir Maheshwari, Anurag Nishad, Abhijit Bhattacharyya, and Vipin Gupta who helped me in various ways during my thesis work.

Above all, I am indebted to my parents, my wife, my sister, and my son for their unconditional love and support. I cannot express my gratitude to my parents in words, without their blessings and inspiration, I would not have been able to accomplish the most significant milestones in my life.

Last but not the least, I thank God Almighty to enlighten my mind for carrying out this research work.

MOHIT KUMAR

Dedicated

to

My Family

ABSTRACT

Cardiovascular diseases (CVDs) are the major cause of death globally. The global contribution of the CVDs is 30% of all deaths every year. The electrocardiogram (ECG) is the most popular non-invasive tool for detecting the heart diseases. The cardiologists perform the manual inspection of the ECG recordings for the detection of heart diseases. The accuracy (Ar) of the diagnosis depends upon skills and experience of the cardiologist. Lack of expertise may result in an inaccurate diagnosis. Therefore, the automated decision making methods are needed to help the doctors during the diagnosis and also can reduce their workload. Hence, we have proposed computer-aided automated methods for the diagnosis of different types of heart disorders in this thesis work. These methodologies are described below:

Coronary artery disease (CAD) causes maximum death among all types of heart disorders. Therefore, we have proposed a new technique which can detect CAD automatically. In this method, the heart rate variability (HRV) signals are decomposed to sub-band signals using flexible analytic wavelet transform (FAWT). Two parameters namely; K-nearest neighbour (K-NN) entropy estimator and fuzzy entropy (FEnt) are computed from the decomposed sub-band signals. Various ranking methods are also used for optimising the features. The proposed methodology has shown better performance using entropy ranking technique. The least squares-support vector machine (LS-SVM) with Morlet wavelet and radial basis function (RBF) kernels yielded the highest classification Ar on the studied dataset.

An automated method for the detection of CAD using ECG signals is also proposed. First, the ECG signals of normal and CAD subjects are segmented into beats. The cross information potential (CIP) is computed from the detail coefficients obtained by FAWT based decomposition of ECG beats. For CAD subjects, mean values of CIP parameters are found higher than normal subjects. Further, the features are fed to LS-SVM classifier. We have observed significant improvement in the classification Ar up to the fourth level of decomposition. At the fifth level of decomposition, no significant improvement is noticed in the classification Ar as compared to the fourth level of decomposition. Hence, the ECG beats are analyzed up to the fifth level of decomposition. The Ar of classification is slightly higher for Morlet wavelet kernel (99.60%) than RBF kernel (99.56%).

Myocardial infarction (MI) is a condition which can cause the death of the heart muscles. Therefore, we have developed a method for automated identification of MI ECG signals using FAWT. First, the segmentation of ECG signals into beats is performed. Then, FAWT is applied to each ECG beat to decompose them into sub-band signals. Sample entropy (SEnt) is extracted from each sub-band signal. We have achieved the highest classification Ar of 99.31% when SEnt is fed to the LS-SVM classifier. We have also incorporated Wilcoxon and Bhattacharya ranking methods and observed no improvement in the performance.

A method for automated diagnosis of congestive heart failure (CHF) is proposed in this work. In this methodology, HRV signals of three different lengths (500, 1000, and 2000 samples) are analyzed using FAWT based decomposition. The accumulated fuzzy entropy (AFEnt) and accumulated permutation entropy (APEnt) are computed from the different combination of the obtained sub-band signals and ranked using the Bhattacharyya ranking method. Further, these ranked features are applied to the LS-SVM classifier. The proposed system has obtained the Ar of 98.21%, 98.01%, and 97.71%, for the 500, 1000, and 2000-sample length of HRV signals.

Atrial fibrillation (AF) represents a condition of abnormal heart rhythm. Hence, a new approach for the detection of AF using FAWT is developed. First, the small segments of 1000 ECG samples are extracted from the long duration ECG signals. Then, the permutation entropy (PEnt) and log energy entropy (LEE) features are computed from different sub-band signals which are obtained using FAWT. The classification Ar of LEE features is observed better as compared to PEnt features. We have obtained Ar, sensitivity (Ss), and specificity (Sc) of 96.84%, 95.8%, and 97.6% respectively using LEE features with random forest (RF) classifier.

Contents

List of Figures

List of Tables

List of Abbreviations and Symbols

AF:	Atrial fibrillation
AFEnt:	Accumulated fuzzy entropy
AFL:	Atrial flutter
ALLs:	Augmented limb leads
AM:	Amplitude modulation
ANN:	Artificial neural network
ANS:	Autonomic nervous system
APEnt:	Accumulated permutation entropy
Ar:	Accuracy
ASP:	Advanced signal processing
AV:	Atrioventricular
AVe:	Aortic valve
AVV:	Atrioventricular valve
BD:	Balanced dataset
BIDMC:	Beth Israel Deaconess Medical Center
BPNN:	Back propagation neural network
BPSO:	Binary particle swarm optimization
CAD:	Coronary artery disease
CART:	Classification and regression tree
CHF:	Congestive heart failure
CIP:	Cross information potential
CNN:	Convolutional neural network
CT:	Counterlet transform
CVDs:	Cardiovascular diseases
db6:	Daubechies 6

DCT: Discrete cosine transform

DFA: Detrended fluctuation analysis

DH: Density histograms

DT: Decision tree

DWT: Discrete wavelet transform

e: Embedding dimension/Sequence length

ECG: Electrocardiogram

ECS: Electrical conduction system

EEG: Electroencephalogram

EMD: Empirical mode decomposition

EMG: Electromyogram

EST: Exercise stress test

FAWT: Flexible analytic wavelet transform

FBE: Fourier-Bessel expansion

FEnt: Fuzzy entropy

FM: Frequency modulation

GA: Genetic algorithm

GMM: Gaussian mixture model

HF: High frequency

HFSBS: High frequency sub-band signals

HMM: Hidden Markov model

HOS: Higher order statistics

HR: Heart rate

HRP: High risk patients

HRV: Heart rate variability

HSS: Heart sound signals

ICA: Independent component analysis

IMFs: Intrinsic mode functions

J: Level of decomposition

K-NN: K-nearest neighbor

KS: Kolmogorov-Smirnov

KW: Kruskal-Wallis

LA:	Left arm electrode
LAm:	Left atrium
LDA:	Linear discriminant analysis
LEE:	Log energy entropy
LF:	Low frequency
LFSBS:	Low frequency sub-band signals
LL:	Left leg electrode
LLs:	Limb leads
LRP:	Low risk patients
LS-SVM:	Least squares-support vector machine
LVc:	Left ventricle
MCC:	Matthews correlation coefficient
MEES:	Multiscale energy and eigenspace
MI:	Myocardial infarcation
MIT-BIH:	Massachusetts Institute of Technology-Beth Israel Hospital
MLP:	Multilayer perceptron
MVe:	Mitral valve
NSR:	Normal sinus rhythm
p:	Probability
P:	Number of samples in the signal/signal length
PA:	Pulmonary artery
PCA:	Principal component analysis
PEnt:	Permutation entropy
PLs:	Precordial leads
PTB:	Physikalisch Technische Bundesanstalt
PV:	Pulmonary vein
PVe:	Pulmonary valve
QF:	Quality factor
RA:	Right arm electrode
RAm:	Right atrium
RBF:	Radial basis function
RF:	Random Forest

RL: Right leg electrode

RMSSD: Root mean square of successive RR differences

ROC: Receiver operating characteristic

RVc: Right ventricle

SA: Sino-atrial

Sc: Specificity

SDH: Standard density histograms

SEnt: Sample entropy

SLV: Semilunar valve

Ss: Sensitivity

STD: Standard deviation

STm: Shearlet transform

STFT: Short-time Fourier transform

SVM: Support vector machine

TEO: Teager energy operator

TFCV: Ten-fold cross-validation

TQWT: Tunable-Q wavelet transform

TVe: Tricuspid valve

UD: Unbalanced dataset

VLF: Very low frequency

WEKA: Waikato environment for knowledge analysis

WHO: World health organization

WPT: Wavelet packet transform

WT: Wavelet transform

XWT: Cross wavelet transform

μ: Mean

Chapter 1

Introduction

For a healthy human, proper circulation of the blood to the body tissues and organs is vitally important. The essential nutrients and oxygen are carried to the tissues and organs via the circulation of blood. It also carries carbon dioxide from the tissues to the lungs (Graaff et al., 2010). Blood circulation is controlled by the cardiovascular system. The heart is the main part of this system. It pumps the blood to suffice the demand of the human body. Hence, proper functioning of the heart is of prime importance for a living human. Its function may get disrupted due to the cardiovascular diseases (CVDs). The CVDs can be diagnosed using cardiac signals such as electrocardiogram (ECG) and heart rate variability (HRV).

The cardiac signals can be used to monitor the pathological conditions of the human heart. These signals have valuable information about a particular state of the human heart (Acharya et al., 2007). The clinical aspect of the acquisition of cardiac signals is to assess the information to distinguish the normal and pathological states of the heart. The meaningful information from cardiac signals can be assessed using signal processing methods and used for the diagnosis of CVDs. The obtained information can also be utilized to design the computer-aided decision-making system. These systems can be employed in the hospitals and remote locations for assisting the doctors in the diagnosis of various cardiac diseases. In the following section, a brief explanation of the human heart is provided.

1.1 Human heart and its functioning

Human heart is a muscular pump. The muscle wall consists inner, middle, and outermost layers. The inner layer which lines the heart chambers is named as endocardium. The middle layer is called as myocardium. Its function is to provide the sufficient force for blood pumping. The outermost layer, pericardium, which consists fluid to reduce the friction against the heart movements (Acharya et al., 2007). The heart is divided into the right side and the left side by a wall of tissues named as septum (Cromwell et al., 2012). These two sides are further divided into two parts. The left side of the heart is divided into the left atrium (LAm) and left ventricle (LVc), and the right side is divided into the right atrium (RAm) and right ventricle (RVc). Thus, the human heart can be considered as the collection of four different chambers namely, LAm, RAm, LVc, and RVc (Cromwell et al., 2012).

The heart receives blood through the atria (LAm and RAm) and pumps the blood out via the ventricles (LVc and RVc). The LAm is the smallest and the LVc is the largest among all the four chambers. Veins are those blood vessels through which blood comes to the heart, and the vessels that carried blood away from the heart are known as arteries. The oxygenated blood flows via the arteries and deoxygenated blood moves through the veins. The pulmonary artery (PA) and the pulmonary vein (PV) are the exceptions. The PA carries deoxygenated blood from the heart to lungs, and the PV transports the oxygenated blood from lungs to the heart.

The heart has two atrioventricular valves (AVVs) and two semilunar valves (SLVs). The AVV valve located between LAm and LVc is called as mitral valve (MVe). The other AVV is positioned between RAm and RVc which is named as tricuspid valve (TVe) (Cromwell et al., 2012; Graaff et al., 2010). The opening and closing of the AVVs depend on the pressure in the heart chamber. It opens when the pressure in the atria exceeds than the pressure in the ventricular part and vice-versa. One SLV is the pulmonary valve (PVe) which is located between the RVc and PA. The other SLV, aortic valve (AVe), is positioned between the LVc and aorta (Graaff et al., 2010). Functioning of these valves is also similar to the AVV. The SLVs open when the pressure in the LVc and RVc is higher in comparison to the pressure in aorta and PA.

The regulation of blood flow in the human body can be summarized in following

steps (Cromwell et al., 2012; Graaff et al., 2010):

Step 1 : First, the impure blood or the deoxygenated blood from the different organs and parts of the body comes to the RAm via the vene cava.

Step 2 : When the pressure of the blood in the RAm more than the RVc, TVe opens and impure blood starts passing to the RVc. Then, RAm contracts and passes all the impure blood to the RVc.

Step 3 : Now, the RVc contracts to pump the blood to the lungs. The TVe closes to make sure that no backward flow of the blood to the RAm. At this time PVe opens, and blood passes through PA to the lungs where the impure blood gets oxygenated. When the pressure in the RVc comes down to the pressure of the PA, then PVe closes and prevents the reverse flow of the blood to the RVc.

Step 4 : Then, the oxygenated blood comes to the LAm via PV.

Step 5 : From the LAm, it passes to the LVc through MVe in the same way as it passes RAm to RVc.

Step 6 : Finally, the LVc contracts and AVe opens to pump the oxygenated blood to the whole body via the aorta. At the time of LVc contraction, MVe closes to stop the blood flowing in the reverse direction. AVe closes when the pressure in the aorta increases as compared to the LVc to ensure the blood will not flow back to the LVc.

The above mentioned steps represent a cardiac cycle which starts with the receiving of the deoxygenated blood in the RAm from the organs of the body and completes when body organs and parts receive the oxygenated blood from the LVc. The cardiac cycle can be considered as a combination of two parts, namely, systole and diastole. Systole indicates the contraction of the heart muscles, and diastole represents the relaxation phase of the heart muscles (Cromwell et al., 2012; Graaff et al., 2010).

For the pumping of the blood in an efficient manner, all the heart chambers must contract in a synchronized way. It is achieved by the electrical conduction system (ECS) of the heart. The ECS is initiated by the sino-atrial (SA) node located in

the upper part of the RAm (Acharya et al., 2007). It is also termed as the natural pacemaker. It originates electrical impulse which propagates via the myocardium to the LAm and RAm and causes their contraction. The atria and the ventricles are electrically separated by the non-conducting tissues. The electrical impulse propagates to the ventricles from the atria via atrioventricular (AV) node. The AV node provides a short time delay to the electrical signal which gives time to the LVc and RVc to fill with the blood before contraction (Acharya et al., 2007). Then, electrical impulse propagates through the bundle of His to the myocardium of the LVc and RVc and leads their contraction (Cromwell et al., 2012). The working of the heart can be disturbed due to CVDs which are described in the following section.

1.2 Cardiac diseases

The CVDs are responsible for more deaths in the world than any other disease. Each year, these diseases contribute 30% of all deaths globally (Wong, 2014). The 80% deaths due to CVDs occur in low and middle-income countries. It is also a leading cause of death in India. The mortality rate in India due to CVDs is 272 per 100000 population. It is higher as compared to the global average CVDs death rate of 235 per 100000 population (Prabhakaran et al., 2016). The common CVDs are coronary artery diseases (CAD), myocardial infarction (MI), congestive heart failure (CHF), and atrial fibrillation (AF).

The CAD is a most common cause of mortality among all types of CVDs (Wong, 2014). In this disease, blood supplying arteries to the heart muscles become narrow due to the deposition of plaque inside them. This condition reduces the blood supply to the heart muscles (National Heart, Lung and Blood Institute, 2015a). Hence, heart muscles cannot receive the sufficient amount of oxygen and nutrients for proper functioning. This condition is becoming worse with the time and affects the metabolic activity of cardiac muscles. Finally, muscles become weak and lead to arrhythmias and heart failure (National Heart, Lung and Blood Institute, 2015a).

The MI is the indication of the death of the heart muscles (Thygesen & et al, 2012). This happens due to the disruption of the blood supply to a segment of the myocardium. It is the result of the plaque formation inside the coronary arteries.

This condition becomes worse when the deposited plaque ruptures and resulting in the build-up of blood clots and further restricts the flow of blood to the heart muscles (Thygesen & et al, 2012). Finally, the segment of the heart muscles start to die and may cause CHF or heart attack.

The CHF represents a condition in which the heart is unable to pump sufficient amount of blood to fulfill the requirement of the body. The CAD, diabetes, and high blood pressure are the common causes of the CHF (National Heart, Lung and Blood Institute, 2015b). Around the world, 17% to 45% of the CHF patients die within a year of their admission in the hospital (Ponikowski et al., 2014). The CHF is not a curable disease. Use of medicines and change in lifestyle may help the patient to live better and longer life (National Heart, Lung and Blood Institute, 2015b). Hence, timely diagnosis can save many of the lives.

The AF indicates irregularity and rapidity of heart rhythm. In AF, the electrical signals are not regulated by the SA node as in normal case (Cottrell, 2012). In this case, signals also generate from other parts of the atria. Hence, different parts of atria work in an uncoordinated manner, and atrial contraction rate increases to 350 per minute. In this arrhythmia, contraction rate of ventricle also increases to 180 beats/minute (Acharya et al., 2007). Therefore, the heart is unable to pump the blood properly.

The cardiac signals can be used to diagnose the CVDs. The cardiac signals are presented in the next section.

1.3 Cardiac signals

Cardiac signals are the noninvasive tool which can be used to assess the information about the health of the heart. The cardiac signals used in this work are described below:

1.3.1 Electrocardiogram

The ECG represents the electrical activity of the heart. A typical ECG signal taken from the Fantasia open-access database at physionet (Iyengar et al., 1996; Goldberger et al., 2000) is shown in Figure 1.1. An ECG beat comprises of P-wave,

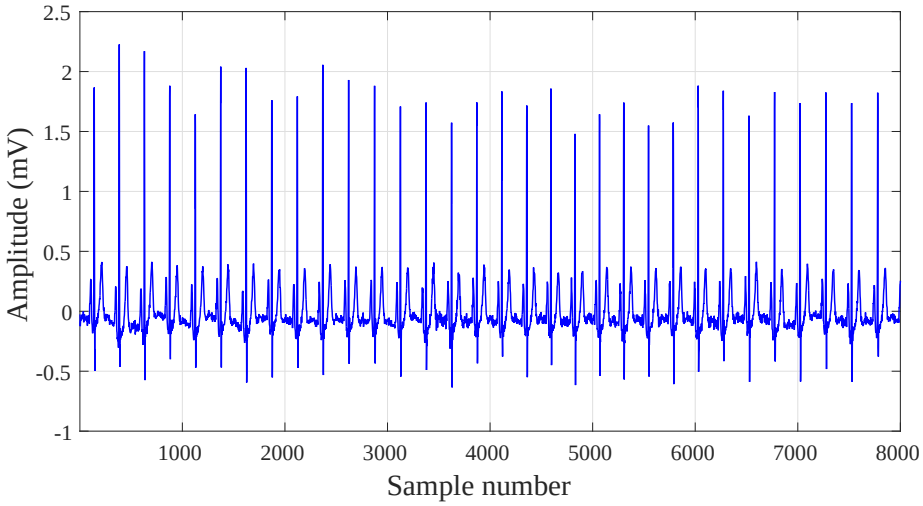


Figure 1.1: A plot of ECG signal.

QRS-complex, and T-wave (Cromwell et al., 2012). P-wave is the result of the depolarization of the atria. QRS-complex consists of three components, namely, Q, R, and S waves. This complex occurs in response to the ventricles depolarization. T-wave is the reflection of the repolarization of the ventricles (Graaff et al., 2010). Sometimes, one another wave after the T-wave also reflects in the ECG signal due to late depolarization of ventricles (Acharya et al., 2007).

The standard ECG recording system used by the physicians has twelve leads. It consists of 3-limb leads (LLs) (bipolar leads), 3-augmented limb leads (ALLs) (unipolar leads) and 6-precordial leads (PLs) (unipolar leads). The standard ECG system has ten surface electrodes (Cromwell et al., 2012). Out of 10, 6-electrodes (V1-V6) are placed on left side of chest called as precordial electrodes. In the four remaining electrodes, two are placed on the right and left arms which termed as right arm electrode (RA) and left arm electrode (LA), respectively. The other two electrodes are left leg electrode (LL), and right leg electrode (RL) placed on the left and right legs, respectively. The RL is used for the ground reference. A pair of electrodes (positive & negative) placed on the body is considered as lead (Acharya et al., 2007). In case of unipolar lead, the second electrode is virtual.

There are three bipolar leads, namely, lead-I, lead-II, and lead-III. A pair of LA and RA forms lead-I. The electrodes placed on the right arm and left leg form lead-II. Lead-III is the combination of the LL and LA. The three ALLs are aVL, aVR, and aVF. For aVL, the real terminal is LA, and the virtual terminal is the

combination of RA and LL. The aVR lead has the real electrode on the right arm, and LL and LA together work as the virtual electrode. For aVF, LL works as a real electrode, and the virtual terminal is formed by the LA and RA. The precordial electrodes (V1-V6) work as the real electrodes for the 6-PLs. The virtual electrode for the PLs is formed by the combination of LA, RA, and LL (Acharya et al., 2007). In the following section, the HRV signals are briefly explained.

1.3.2 HRV signals

The HRV signals also carry useful information about the state of the heart. These signals represent the variation in the time interval between the two consecutive R-peaks in the ECG signals (Acharya et al., 2007). A plot of typical HRV signal downloaded from Fantasia database is depicted in Figure 1.2. The analysis of HRV signals also provides the assess to the state of the heart and autonomic nervous system (ANS) (Acharya et al., 2007). A significant connection between ANS and cardiac mortality is observed in (Vaseghi & Shivkumar, 2008). Hence, HRV can also serve as an important tool for the detection of cardiac diseases.

The computer-based analysis of the cardiac signals (ECG, HRV) is discussed in the next section.

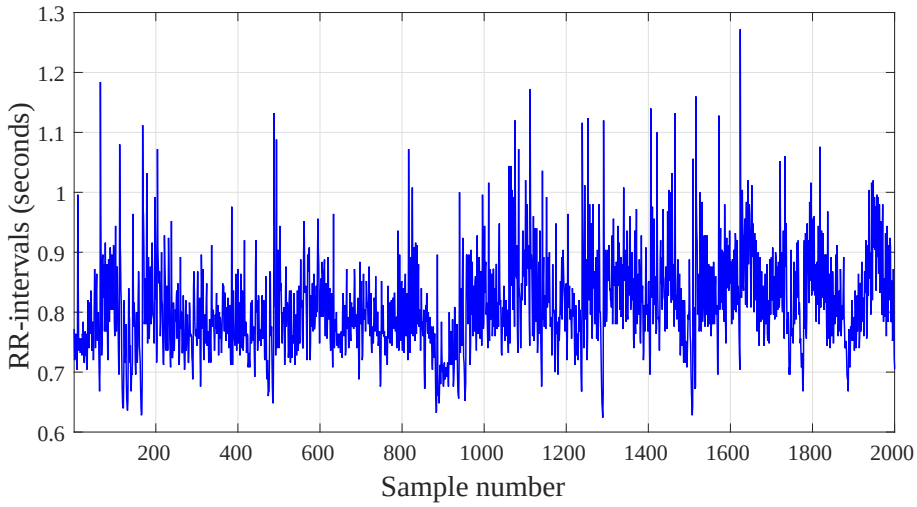


Figure 1.2: A plot of HRV signal.

1.3.3 Computer-based analysis of cardiac signals

The ECG signals possess information about the electrical function of the heart. Any change or alteration in the shape of P, QRS, and T waves may be the indication of any disease or pathological conditions (Acharya et al., 2007). It takes a lot of time for a cardiologist to diagnose the pathological conditions manually by viewing the ECG signals. Analysis of the ECG signal can also be performed using the computer-based methods. The implementation of such system is possible with the help of signal processing techniques and artificial intelligence methods (Patidar, 2014). These methods can provide a way for the better diagnosis of cardiac disorders by fetching more broader information from ECG signals. Development of these methods involve some necessary steps. These steps are discussed below.

First step is the acquisition and preprocessing of the cardiac signals. The aim of this stage is to provide the signal in a suitable format for the computer analysis. It is followed by the noise removal step. The ECG signals, during the time of recording, may be corrupted by noise due to the motion artifacts, respiration, power line interference, electrode contact noise, and muscle contraction (Acharya et al., 2007). It is an essential step, as the presence of noise can mislead the diagnostic result.

The second step involves feature extraction process. Numerous types of features for studying the ECG signals are suggested in the literature. Time domain features such as QRS width (Yeap et al., 1990; Chazal et al., 2004), T-wave duration (Chazal et al., 2004), and presence or absence of P-wave (Chazal et al., 2004) are suggested for ECG signal analysis. Other than these features, T-wave slope and hight of QRS are also suggested in (Yeap et al., 1990). Frequency domain features such as spectral power (Small et al., 2000) and predominant frequency (Chen, 2000) are used to analyze the ECG signals. The statistical features such as higher order cumulants are also explored in for studying the ECG beats (Osowski & Linh, 2001). The other way of ECG signals analysis is to convert them into the HRV signals. These signals also have wide application in assessing the information of heart states (Acharya et al., 2003). Various features such as average heart rate (HR), energy of the band, correlation dimension, and fractal dimension are studied to analyze the HRV signals (Acharya et al., 2003, 2005).

Next step is the selection of the suitable features for the classification. Feature selection can be performed using the ranking methods. These methods are able to reduce the complexity of the system without degrading the performance (Duda et al., 2000). Feature reduction methods such as principal component analysis (PCA) and independent component analysis (ICA) are also utilized for the optimization of the features (Giri et al., 2013).

Then, the last step is the decision making about the pathological condition of the heart. On the basis of this step further action may take place. Hence, it is an important clinical step. Various methods are proposed for decision making such as neural network based approach in (Chen & Yu, 2012) and support vector machine (SVM) based method in (Osowski et al., 2004; Ye et al., 2012). A neuro-fuzzy based approach in (Engin, 2004) and radial basis function (RBF) neural network in (Korurek & Dogan, 2010) are also utilized for ECG beat classification.

1.4 Motivation

The ECG is the most general tool used by the cardiologist for the diagnosis of the heart diseases. It is a non-invasive tool and represents the state of the heart. Minute changes in the ECG signals may be the indicator of a particular disease. The accuracy (Ar) of the diagnosis depends upon the correct identification of these changes. Visual detection of these minute changes is quite difficult. It requires a lot of skills and years of experience. Also, the manual analysis of ECG signals is a tedious and time consuming task as it requires the visual inspection of several hours of ECG recordings. Lack of skills and fatigue may cause inaccurate diagnosis. Therefore, a computer-aided system is required which can help the doctors to make their diagnosis faster and accurate.

Computer-based decision-making system involves two essential steps; feature extraction and classification for decision making. Advanced signal processing (ASP) techniques can be used for feature extraction, and machine learning methods can be utilized for decision making based on the extracted features. These techniques can be integrated with the portable ECG recording devices to detect the heart disorders. This integration can facilitate continuous monitoring of ECG signals for

a longer period of time without human intervention and diagnosis of heart disorders is possible in a more versatile way. These developed methodologies can also be installed in the hospitals. These systems can provide effective assistance to clinicians for improving their diagnostic Ar. These systems can also be upgraded for home-based telemonitoring applications. Home-based telemonitoring systems can provide better clinical assistance for the old age heart patients. These patients are more prone to severe complications and require continuous monitoring. These state of the art techniques for the diagnosis of heart patients can be useful in early prediction of complications and can play an important role in monitoring and controlling the heart disorders.

The cardiac signals (ECG, HRV) show nonlinear and non-stationary characteristics (Acharya et al., 2007, 2004). Therefore, the extraction of the meaningful information from the cardiac signals is a challenging task. Use of the linear methods may miss the subtle information. Hence, the state of the art method which can deal with the nonlinear and non-stationary nature of the cardiac signals will be more suitable for information extraction. Wavelet transform (WT) is a useful tool for analyzing the non-stationary signals. The performance of the WT based method depends upon the proper selection of the mother wavelet according to the application. Flexible analytic wavelet transform (FAWT) (Bayram & Selesnick, 2009) is an advanced WT method which provides the flexibility of adjusting the mother wavelet as per the requirement. Hence, it is a useful tool for analyzing the cardiac signals.

The FAWT decomposes the signals into sub-band signals. These sub-band signals can be used to capture the subtle information about the heart disease. The information present in the sub-band signals can be extracted using nonlinear parameters as cardiac signals have nonlinear characteristics. Further, these parameters can be used as an input of the classifiers for decision making. Therefore, we have investigated different nonlinear features in the FAWT framework for automated detection of cardiac signals related to different classes.

1.5 Objectives

The objectives of this dissertation are as follows:

- To develop FAWT based automated detection method for CAD disease using entropy based features extracted from HRV signals.

- Development of automated diagnosis method for CAD disease based on FAWT and cross information potential (CIP) using ECG signals.

- To develop an automated identification method for the MI ECG signals using sample entropy (SEnt) feature extracted in FAWT framework.

- To propose a computer-based approach for the automated diagnosis of CHF patients using FAWT based accumulated entropies extracted from HRV signals.

- Development of computer-aided identification method for AF ECG signals by computing the entropy based features in FAWT domain.

1.6 Contributions

The contributions of this thesis are summarized as follows:

- A FAWT based automated detection method for CAD disease is developed using HRV signals. The HRV signals are decomposed into the sub-band signals using FAWT. These sub-band signals are used to compute the K-nearest neighbor (K-NN) entropy estimator (Kraskov et al., 2004) and fuzzy entropy (FEnt) (Chen et al., 2007) features. These features yielded maximum classification Ar with the least squares-support vector machine (LS-SVM) (Suykens & Vandewalle, 1999) classifier on the studied dataset.

- An automated diagnosis method for CAD disease using ECG signals is developed. The ECG signals are segmented into the beats, and then, FAWT is applied to decompose the ECG beats. The CIP parameter (Xu & Erdogmuns, 2010) is extracted from the detail coefficients obtained using FAWT. Further, the features are used with LS-SVM classifier for decision making. The performance of the proposed method is found better than the other method studied on the same dataset.

- A method for automated identification of MI ECG signal using FAWT is developed. In this method, FAWT is applied to the ECG beats segmented from the ECG signals. Then, obtained sub-band signals are used to compute SEnt (Richman & Moorman, 2000). The suitable parameters to compute the SEnt and level of decomposition (J) in FAWT domain for the accurate detection of MI subjects are identified. The method achieved better classification performance than the compared works.

- A computer-based approach for the automated diagnosis of CHF patients is developed. In this approach, accumulated fuzzy entropy (AFEnt) and accumulated permutation entropy (APEnt) are computed over cumulative sums of the sub-band signals provided by FAWT based decomposition of HRV signals. Then, features are ranked and given as an input to the LS-SVM classifier. Our proposed system has obtained better performance than the other work studied on the same dataset with a smaller length of HRV signals.

- A computer-aided identification method for AF ECG signals using FAWT is developed. First, smaller ECG segments of 1000 samples are obtained from the long duration ECG signals. Further, FAWT is used to extract the sub-band signals from ECG segments. Then, log energy entropy (LEE) (Han et al., 2009) and permutation entropy (PEnt) (Bandt & Pompe, 2002) are computed from the sub-band signals. The performance of the proposed method is comparable to other methods. However, R-peak and P-wave detection are not required in this method. It is the advantage of this method than the other methods.

1.7 Organization of the thesis

The remaining portion of this thesis is organized in the following way:

- In chapter 2, the FAWT method is explained. Signal decomposition using FAWT is also presented in this chapter.

- The FAWT based automated detection method for CAD disease using HRV signals is explained in chapter 3. The extraction steps of K-NN entropy estimator and FEnt features are explained. Ranking methods are also presented

in this chapter. Explanation of LS-SVM classifier and various kernel functions are provided. Finally, results are presented and discussed.

- Development of automated diagnosis method for CAD disease using ECG signals based on FAWT is provided in chapter 4. The computation of CIP features from the FAWT coefficients is explained in this chapter. Finally, the obtained results using the LS-SVM classifier are discussed in detail.

- An automated identification method for the MI ECG signals in FAWT framework is presented in chapter 5. The SEnt computation from the decomposed sub-band signals is given in this chapter. The selection of suitable parameter for the computation of SEnt is explained. Selection of the decomposition level (J) is also discussed. The random forest (RF) classifier is also provided in this chapter. Finally, achieved results and discussion part is given.

- A computer-based approach for the automated diagnosis of CHF patients using HRV signals is described in chapter 6. The computation of PEnt is explained in this chapter. A way of obtaining the different frequency scales from the combination of decomposed sub-band signals is shown. The computation of AFEnt and APEnt from the accumulated sub-band signals is presented. The results obtained from the different length of the HRV signals are discussed.

- Development of computer-aided identification method for AF ECG signals using FAWT is demonstrated in chapter 7. The computation of LEE features from the sub-band signals is given in this chapter. The results achieved using this method are presented and discussed in the chapter.

- Finally, the whole work is concluded in chapter 8. The direction of future research work is also provided in this chapter.

Chapter 2

Flexible analytic wavelet transform

2.1 Introduction

The biomedical signals such as ECG, electroencephalogram (EEG), and electromyogram (EMG) exhibit non-stationary nature. The characteristics of these signals change rapidly. Analysis of these signals required the information in both time and frequency plane. Hence, the Fourier transform is not suitable for the analysis of this kind of signals, as it does not provide the information about the variation of frequency contents with time (Gao & Yan, 2011). The short-time Fourier transform (STFT) and WT-based methods are more suitable for the analysis of non-stationary signals.

In STFT, the width of the window remains constant during the analysis of the signal. Hence, the time and frequency resolution cannot be altered during analysis. The resolution in time and frequency domain depends on the width of the window function. The selection of a wider window in time-domain yields fine frequency resolution and poor time resolution, and vice-versa (Cohen, 1989). In many of the situations during the analysis of non-stationary signals, it requires good time resolution at the occurrence of high frequency and finer frequency resolution at the event of low-frequency (Chui, 1992). Therefore, the performance of the STFT is limited due to the fixed resolution. The WT has the property to provide multi-resolution analysis to deal this kind of situations. It has scaling parameter which

allows the analysis of a signal with different resolution in time and frequency domain. Hence, it finds suitable applications in the analysis of cardiac signals.

In (Inan et al., 2006), the ECG signals are analyzed using WT to detect the premature ventricular contractions. It is also used as a QRS detector in (Kadambe et al., 1999). The WT based morphological features are utilized for ECG signal classification in (Ince et al., 2009). In (Kaur et al., 2016), discrete wavelet transform (DWT) is employed for the detection of arrhythmia. The WT is also found useful in the study of myocardial ischaemic HRV signals (Gamero et al., 2002). A WT based HRV approach is also proposed for the analysis of ventricular tachycardia in (Chen, 2002). The performance of the WT based methods also depends on the proper selection of the mother wavelet according to the application.

The wavelet packet transform (WPT) based method is also used for the analysis of the HRV signals (German-Sallo, 2014). It is also explored for the analysis of AF ECG signals in (Qiao & Zhou, 2007). The dual tree complex wavelet transform is also employed for the denoising of the ECG signals in (Mishu et al., 2014). The tunable-Q wavelet transform (TQWT) method is also used for the analysis of the cardiac signals (Patidar et al., 2015a). The TQWT has an advantage than the other WT based method that it has the facility to adjust the Q-factor and the redundancy. Hence, a suitable wavelet for the application can be selected easily as compared to the other WT based method. Moreover, after selecting the Q-factor and the redundancy, it is not possible to adjust the dilation factor in the TQWT (Bayram, 2013).

The signal adaptive method such as empirical mode decomposition (EMD) (Huang et al., 1998) is also explored for the analysis of the cardiac signals. In this method, the signal is decomposed in to the intrinsic mode functions (IMFs). In (Ortiz et al., 2005), EMD is applied to HRV signals for studying the different conditions related to the fetal activity. It is also used for the investigation of CAD HRV signals in (Acharya et al., 2014; Sood et al., 2016). The EMD based method also incorporated for the analysis of ECG beats in (Shahnaz et al., 2015). The EMD technique suffers from the mode-mixing problem (Tang et al., 2012). The other limitation of the EMD is the lack of the mathematical background.

Recently, the FAWT has been proposed which allows to easily adjust the dilation

factor, quality factor (QF), and redundancy (R). The QF can be expressed as, QF=$\frac{2-\beta}{\beta}$ (Bayram, 2013). Redundancy is the ratio of the output samples and the input samples and can be derived as $R = (\frac{f}{g})\frac{1}{1-(b/c)}$ (Bayram, 2013). QF controls the frequency resolution of FAWT. The high QF provides finer filter banks for analyzing the signals in the frequency domain. For fixed dilation and QF, the redundancy controls the position of the wavelet in the time domain. The FAWT provides shift-invariance, tunable oscillatory bases, and flexible time-frequency covering (Zhang et al., 2015). These features make it more suitable for analyzing a signal as compared to the TQWT and other WT based methods.

2.2 Signal analysis based on FAWT

The FAWT (Bayram, 2013) has the facility to analyze the signal with easily adjustable parameters b, c, f, g and β. The parameters b and c are used to set the up and down sampling rates for low pass channel, respectively. The up and down sampling rates for high pass channels can be controlled by the parameters f and g, respectively. The J^{th} level decomposition of FAWT can be implemented using iterative filter bank (Bayram, 2013). Each level of implementation consists of two high pass channels and one low pass channel as depicted in Figure 2.1. The two high pass channels are used to separate the positive and negative frequencies and provide analytic bases. The wavelet bases of FAWT are obtained in Hilbert transform pairs (Bayram, 2013). These properties make it enable to analyze the transient and oscillatory parts of a signal ($x[n]$).

Frequency response of the low pass filter is defined as (Bayram, 2013):

$$H(w) = \begin{cases} (bc)^{1/2}, & |w| < w_p \\ (bc)^{1/2}\theta\left(\frac{w-w_p}{w_s-w_p}\right), & w_p \leq w \leq w_s \\ (bc)^{1/2}\theta\left(\frac{\pi-w+w_p}{w_s-w_p}\right), & -w_s \leq w \leq -w_p \\ 0, & |w| \geq w_s \end{cases} \tag{2.1}$$

where, The w_p and w_s are the cutoff frequencies of the pass-band and stop-band for the low pass filter, respectively, and can be computed as (Bayram, 2013):

$w_p = \frac{(1-\beta)\pi}{b} + \frac{\epsilon}{b}$, $w_s = \frac{\pi}{c}$,

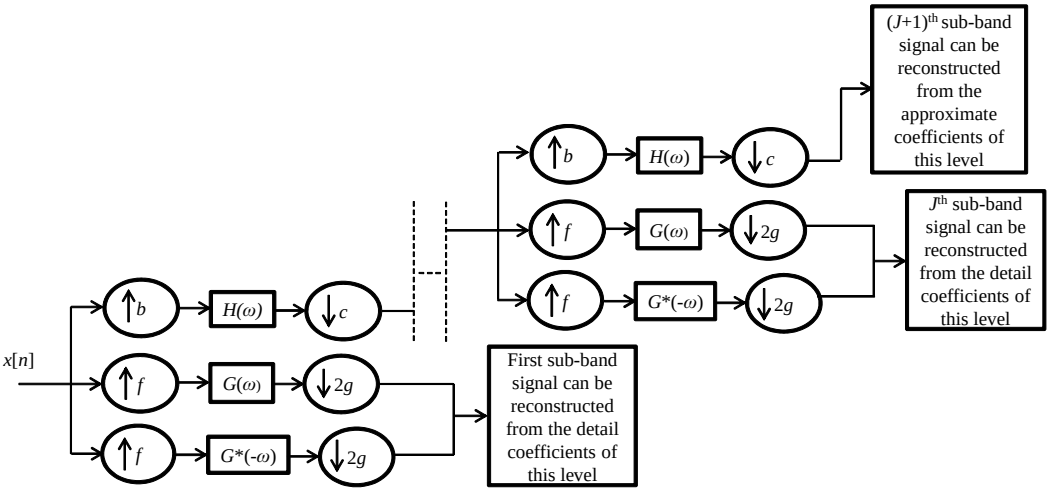


Figure 2.1: J^{th} level FAWT based decomposition of a signal.

Frequency response of the high pass filter is defined as (Bayram, 2013):

$$G(w) = \begin{cases} (2fg)^{1/2}\theta\left(\frac{\pi - w - w_0}{w_1 - w_0}\right), & w_0 \leq w < w_1 \\ (2fg)^{1/2}, & w_1 \leq w < w_2 \\ (2fg)^{1/2}\theta\left(\frac{w - w_2}{w_3 - w_2}\right), & w_2 \leq w \leq w_3 \\ 0, & w \in [0, w_0) \cup (w_3, 2\pi) \end{cases} \tag{2.2}$$

where, The w_0 and w_3 are the stop-band cutoff frequencies for the high pass filter, and w_1 and w_2 are the pass-band cutoff frequencies. These frequencies can be defined as (Bayram, 2013):

$w_0 = \frac{(1-\beta)\pi+\epsilon}{f}$, $w_1 = \frac{b\pi}{cf}$, and

$w_2 = \frac{\pi-\epsilon}{f}$, $w_3 = \frac{\pi+\epsilon}{f}$, $\epsilon \leq \frac{b-c+\beta c}{b+c}\pi$.

The $\theta(w)$ can be given as (Bayram, 2013; Bayram & Selesnick, 2009):

$$\theta(w) = \frac{[1+\cos(w)]\sqrt{(2-\cos(w))}}{2} \quad \text{for} \quad w \in [0, \pi] \tag{2.3}$$

In order to achieve perfect reconstruction filter bank, following conditions must be satisfied (Bayram, 2013):

$$|\theta(\pi - w)|^2 + |\theta(w)|^2 = 1 \tag{2.4}$$

$$\left(1 - \frac{b}{c}\right) \leq \beta \leq \left(\frac{f}{g}\right) \tag{2.5}$$

The FAWT is utilized for the faults detection in rotating machinery (Zhang et al., 2015) and for the study of EEG signals in (Gupta et al., 2017; Sharma et al., 2017b). Matlab toolbox of FAWT is available at http://web.itu.edu.tr/ibayram/AnDWT/.

The value of the factor c/b closer to 1 provides narrow sub-bands in the frequency domain. The factor g/f controls the oscillations of the bases. The ratio $g/f \geq 3$ provides high oscillatory bases. The parameter β can be used to further adjust the oscillations of the bases of FAWT (Zhang et al., 2015). Therefore, these parameters must be selected carefully according to the applications. In our work, we choose these parameters by trial and error experimentation. We varied the factor b/c between 0.5 to 1 and g/f between 1 to 3 as suggested in (Zhang et al., 2015). Finally, we selected the parameters values at which we achieved better results according to our applications. FAWT based decomposition of the ECG signal at $J = 4$ is shown in Figure 2.2, for parameter values, $b = 5$, $c = 6$, $f = 1$, $g = 2$, and β=0.8×(f/g).

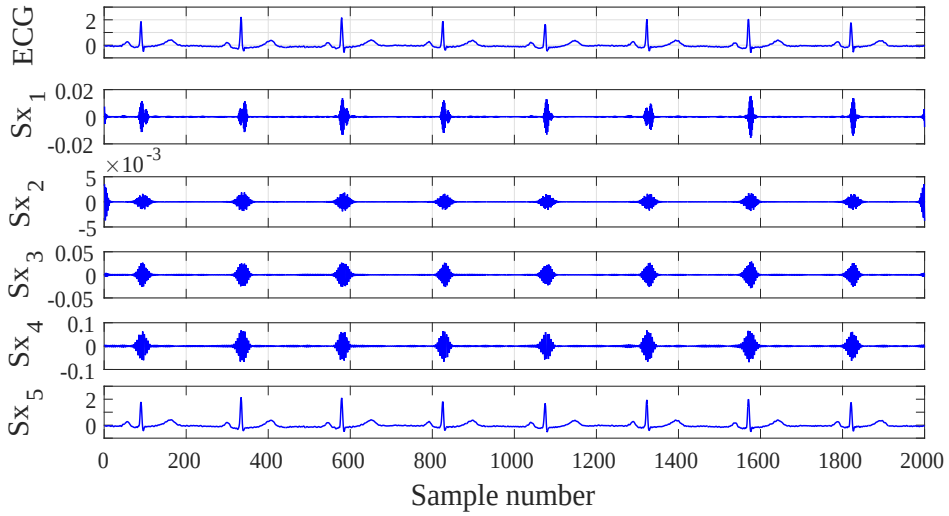


Figure 2.2: A plot of the sub-band signals obtained using FAWT based decomposition of the ECG signal.

In Figure 2.2., Sx_1 to Sx_4 are the sub-band signals which are reconstructed from the detail coefficients. The sub-band signal Sx_5 is reconstructed from the approximate coefficients. The 4^{th} level decomposition of the HRV signal is depicted in Figure 2.3.

The sub-band signal SH_5 is reconstructed from the approximate coefficients, and

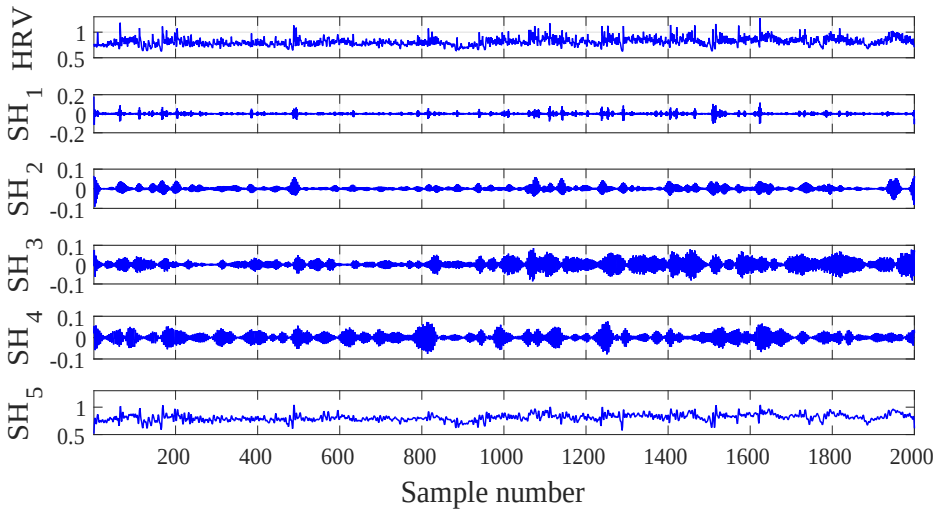


Figure 2.3: The plot of the decomposed components of the HRV signal extracted using FAWT based decomposition.

SH_1 to SH_4 are the sub-band signals which are reconstructed from the detail coefficients. The parameter values are kept same as for the signal shown in Figure 2.2. The ECG and HRV signals are taken from the Fantasia open-access database (Iyengar et al., 1996; Goldberger et al., 2000).

The advantage of using FAWT on the ECG and HRV signals is that it can decompose the signals in to the different subbands. The computation of the features from these subbands allow to assess broader information as compared to the computation of the features directly from the signals. In this work, the nonlinear features are computed from the subbands as the HRV and ECG signals possess the nonlinear characteristics.

2.3 Summary

The FAWT is a convincing technique for signal analysis as it consists adjustable parameters. The advantage of FAWT over WT is that by varying the parameters, the shape of the mother wavelet can be adjusted accordingly. Unlike the EMD method, it has a proper mathematical foundation. The main property of this method is that we can easily tune the QF, dilation factor, and redundancy as per the requirement.

Chapter 3

Automated detection method for normal and CAD HRV signals using FAWT

3.1 Introduction

The CVDs are the major causes of death in the world. Every year almost 17 million people die due to CVDs globally (Wong, 2014). Deaths due to CAD are higher than any other type of heart disease (Wong, 2014). The CAD is the result of the blockage of the coronary arteries due to the accumulation of cholesterol and fatty substances called plaque (National Heart, Lung and Blood Institute, 2015a). It reduces the blood supply to the heart muscles. Hence, there will be a reduction in oxygen and essential nutrients to the heart muscles (National Heart, Lung and Blood Institute, 2015a). It affects the strength of heart muscles, which disturbs heart pumping. Angina and shortening of breath are the common symptoms of the CAD. But in few people, CAD will be latent and does not show any symptoms (National Heart, Lung and Blood Institute, 2015a). Subsequently, it may lead to arrhythmia, heart attack and heart failure (National Heart, Lung and Blood Institute, 2015a). According to the report of world health organization (WHO), 7.4 million people died worldwide due to CAD in 2012 (World Health Organization: Cardiovascular diseases, 2015). Hence, CAD is a life-threatening disease and if detected at an early stage can save the life.

To detect the presence of CAD, physicians review the clinical history of the subject and perform diagnostic tests such as; exercise stress test (EST), ECG, echocardiogram, chest X-Ray, cardiac catheterization and coronary angiography (National Heart, Lung and Blood Institute, 2015a). All these techniques have a few limitations. In stress test, information about the status of the heart using ECG signals during the exercise is recorded. These recordings of heart information are termed as stress ECG. During EST, there is always a risk of cardiac arrest, and all CAD patients may not be able to achieve the targeted HR (Roman et al., 1998; Acharya et al., 2014). Resting ECG recordings are also used for the diagnosis of CAD (Schreck et al., 1988; Nowak et al., 1993).

It is difficult to interpret the minute changes in the ECG signals manually, due to small amplitude, presence of noise and baseline wander. In few cases, ECG recordings may not show significant difference between CAD and normal patients (Giri et al., 2013). Cardiac catheterization and coronary angiography are invasive techniques and can be performed in the presence of expert clinicians (National Heart, Lung and Blood Institute, 2015a; Giri et al., 2013). Hence, a computer-aided method with ASP techniques can help to capture the subtle information and minute changes in ECG signals. Therefore, an automated system can provide immense help to clinicians and doctors in the diagnosis of CAD during routine screening.

Advanced digital signal processing techniques can be used to diagnose the CAD patients. These techniques may be helpful to extract the information provided by the HR signals. These signals depict nonlinear and non-stationary characteristics and may have valuable information about the nature of heart disease (Acharya et al., 2004). The variations of HR signals are also termed as HRV. These signals have wide applications in biomedical engineering for detecting diseases related to the heart. In (Chua et al., 2008), HRV signals of normal and arrhythmia subjects are investigated using higher order spectral analysis. Moreover, HRV signals are used to predict the risk of CVDs (Acharya et al., 2008) and sudden cardiac death (Fujita et al., 2016). Other than heart diseases these signals are also used to detect diabetes (Pachori et al., 2015; Acharya et al., 2015b; Pachori et al., 2016). Significant variations are noticed in the range of nonlinear features for normal and CAD HR signals (Acharya et al., 2014). Power spectral analysis and time-domain based features are

obtained from HR signals (Bigger et al., 1995). All these parameters are found to be significantly lower for CAD patients. In (Hayano et al., 1990), it is shown that spectral components of HR signals have a relationship with the angiographic features of the CAD. In (Acharya et al., 2005), wavelet transform and fractal dimensions are used to analyse the HR signals. Fractal dimension showed decrease values for the diseased heart. Block entropies are estimated from healthy and CAD HR time series. They have shown higher values for healthy subjects as compared to CAD subjects (Karamanos et al., 2006).

HRV signals of normal and CAD subjects are studied for different sample lengths (Sood et al., 2016). Five features namely, amplitude modulation (AM) bandwidth, frequency modulation (FM) bandwidth, second-order difference plot, analytic signal representation area and mean frequency of Fourier-Bessel expansion (FBE) are extracted from IMFs. The FM-bandwidth, AM-bandwidth, and mean frequency of FBE showed more discriminating ability among five features. In (Ji et al., 2016), the ECG signals and photoplethysmography of CAD and normal subjects are studied. Further, heartbeat interval series, diastolic time interval series, and systolic time interval series are constructed from these signals. Thereafter, cross-correlation, mutual information, cross-conditional entropy, coherence function, cross FEnt and cross SEnt are studied on these constructed series for diagnosis of CAD.

The objective of present work is to develop a noninvasive methodology that can automatically diagnose the CAD using HR signals. To deal with the non-stationary nature of HR signals, FAWT (Bayram, 2013) is used to decompose the signals in terms of sub-band signals. Nonlinear features namely; K-NN entropy estimator and FEnt are used to extract the nonlinear dynamics corresponding to sub-band signals obtained from HR signals. The steps performed in the present work are shown in Figure 3.1.

Our proposed CAD diagnosis expert system is completely automated. It comprises two steps: (i) training and (ii) testing. During training step, classifier is trained using entropy features extracted from the subbands of FAWT. In the testing phase, unknown HR signal is fed to our system. Then same entropy features are extracted and fed to the pre-trained classifier for automated diagnosis (normal or CAD). Such an expert system will help the clinicians in their daily screening of car-

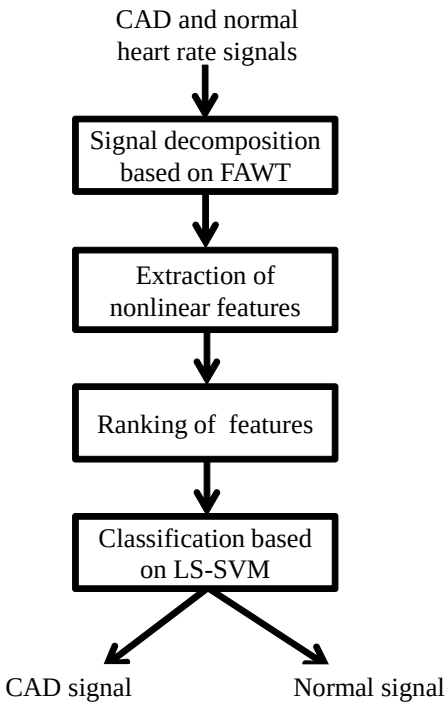


Figure 3.1: The proposed automated method for detecting CAD based on HRV signals.

diac patients. It can substantially reduce the possible human errors, screening time, and make the system robust. The organization of remaining part of this chapter is as follows:

Section 3.2 describes data acquisition, preprocessing, signal decomposition, nonlinear features studied, feature ranking methods, and LS-SVM classifier. The obtained results are provided in section 3.3. Section 3.4 presents the discussion part. Finally, section 3.5 provides the summary of the work.

3.2 Methodology

3.2.1 Process of data acquisition

The ECG signals were recorded from 10 healthy subjects and 10 CAD patients at Iqraa Hospital, Calicut, Kerala, India (Acharya et al., 2014). The BIOPAC^{TM} equipment was used to record ECG signals at a sampling rate of 500 Hz (BIOPAC Systems Canada, Inc., 2010). All the CAD subjects participated in the study were on similar medication. The age of all the subjects under the study ranges between 40

to 70 years. All the procedure of signal acquiring was performed with the assistance of an experienced cardiologist. The patients suffering from other heart diseases such as; myopathy, AF, right and left bundle branch block, and ventricular hypertrophy were not considered for the study. Finally, we have created 82 CAD files and 61 normal files with each file having 1000 samples from 10 normal and 10 CAD subjects.

3.2.2 Preprocessing of acquired signals

The following steps are performed to get the HR signal from the ECG signal (Acharya et al., 2014; Giri et al., 2013):

Step 1: Unwanted noise and baseline wander (Warlar & Eswaran, 1991) present in the ECG signals were eliminated by applying a band pass filter of lower and higher cut-off frequency of 0.3 Hz and 15 Hz, respectively.

Step 2: Power line interferences were eliminated, using a notch filter of 50 Hz cut off frequency.

Step 3: Pan-Tompkins algorithm (Pan & Tompkins, 1985) was used to identify the location of R peaks.

Step 4: Finally, the duration between the two consecutive R peaks (t_{RR}) was computed.

Finally, HR can be defined in terms of beat per minute as (Acharya et al., 2014):

$$\mathrm{HR} = \frac{60}{t_{RR}} \tag{3.1}$$

ECG data were recorded for 15 minutes. The plots of HR signals for CAD and normal subjects are shown in Figure 3.2.

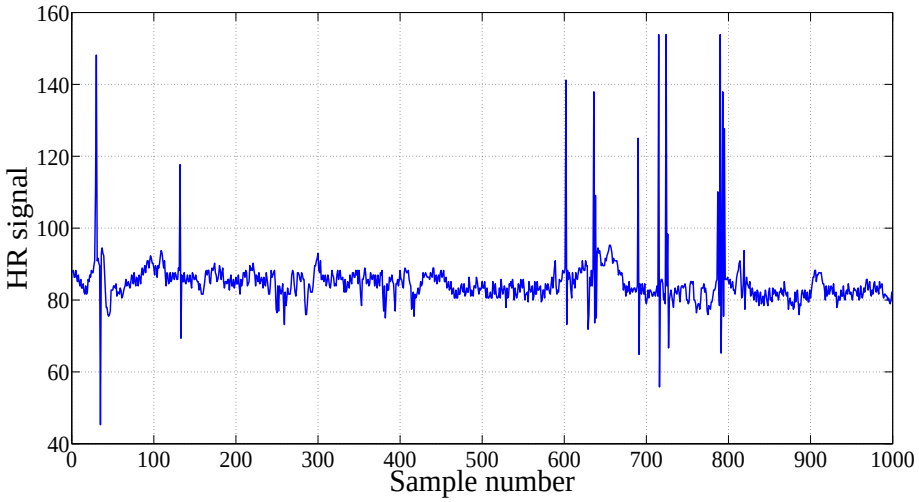
3.2.3 Signal decomposition based on FAWT

The plots of sub-band signals of CAD and normal classes obtained from FAWT are shown in Figure 3.3. In Figure 3.3, D_1, D_2 and D_3 are reconstructed from detail coefficients of levels 1, 2, and 3 respectively. A_3 is reconstructed from approximate

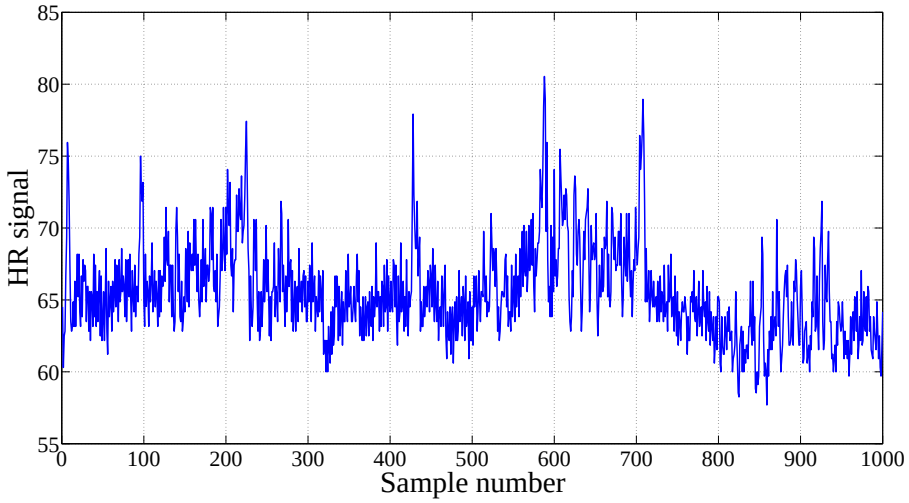
coefficients of level 3. In the present work, $b = 7, c = 8, f = 1, g = 2$, and $\beta = 0.8 \times (f/g)$ are selected (Zhang et al., 2015).

3.2.4 Nonlinear features extraction

Nonlinear features are widely used to analyze the HR signals (Acharya et al., 2004, 2013, 2014). These features unearth the hidden nonlinear nature of HR signals (Acharya et al., 2004; Giri et al., 2013). In this work, nonlinear parameters namely; K-NN entropy estimator and FEnt are used as features. These features are explained in the following sections.



(a)



(b)

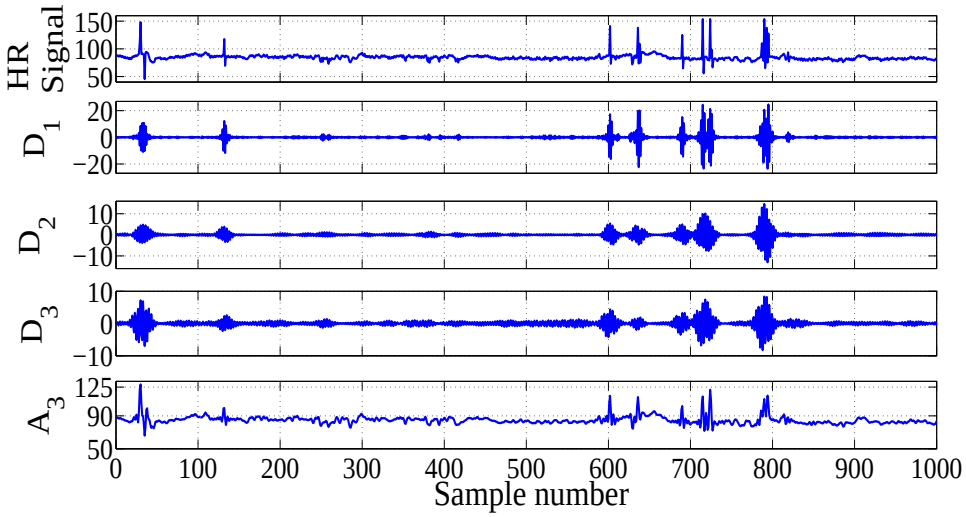
Figure 3.2: The plots of HR signals: (a) CAD subject, (b) Normal subject.

3.2.5 K-nearest neighbour entropy estimator

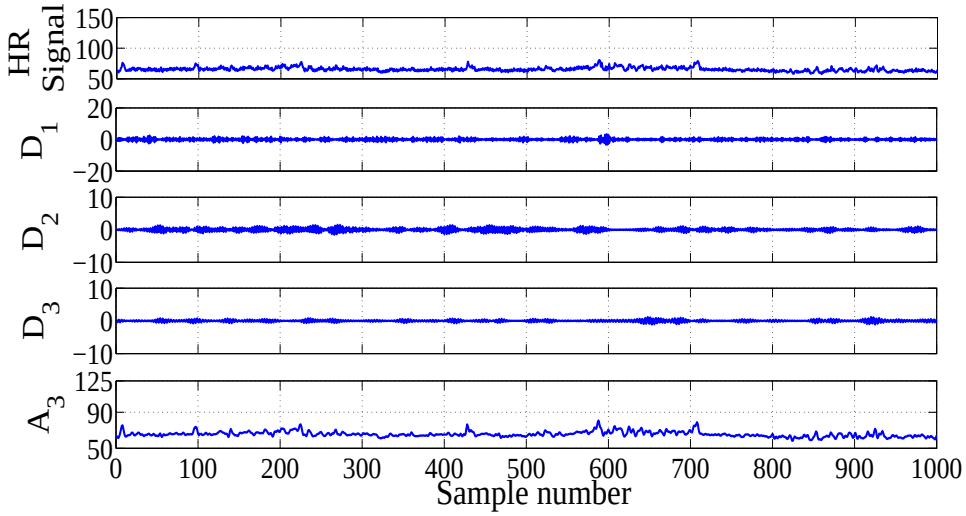
The K-NN entropy estimator $H(x)$ of differential entropy for a variable x can be defined as (Kraskov et al., 2004; Veselkov et al., 2010):

$$H(x) = -\psi(K) + \psi(M) + \log(C_D) + (D/M)\sum_{j=1}^{M}\log[\epsilon(j)] \tag{3.2}$$

where M denotes the total number of samples and D is used to represent the dimension of x. $\epsilon(j)$ is the distance between the j^{th} sample of x and its K nearest neighbours.



(a)



(b)

Figure 3.3: Typical level 3 FAWT based decomposition of HR signal: (a) CAD subject, (b) Normal subject.

In (7), $\psi(y)$ is a digamma function and defined as follows (Kraskov et al., 2004):

$$\psi(y) = \frac{1}{\Gamma(y)} \frac{d\Gamma(y)}{dy} \tag{3.3}$$

The C_D represents the volume of D-dimensional unit ball, and for Euclidean norm it is defined as follows (Veselkov et al., 2010):

$$C_D = \frac{\pi^{D/2}}{\Gamma(1 + \frac{D}{2})} \tag{3.4}$$

The K-NN entropy estimator is used to measure the systemic metabolic disruptions in patho-physiological states (Veselkov et al., 2010).

3.2.6 Fuzzy entropy

The computation of FEnt measures the similarity that occurs in a HRV signal. For the computation of the FEnt, the following steps have to be performed (Chen et al., 2007):

1. First, the sequences of length e are extracted from the HRV signal.

2. Compute the similarity degree S^e_{jk} between two sequences (j-th and k-th) using the fuzzy function (Chen et al., 2007) as follows:

$$S^e_{jk} = q(s^e_{jk}, u, v) \tag{3.5}$$

   where q, u, and v represent the fuzzy function, the gradient, and the width of the fuzzy similarity boundary, respectively and s^e_{jk} is the maximum absolute difference of the two sequence lengths.

3. Computation of $\varphi^e(u, v)$ as follows (Chen et al., 2007):

$$\varphi^e(u, v) = \frac{1}{P - e} \sum_{j\,=\,1}^{P-e} \frac{1}{P - e - 1} \left[\sum_{k\,=\,1, k\,\neq\,j}^{P-e} (S^e_{jk}) \right] \tag{3.6}$$

   where P denotes the total number of samples present in the HRV signal.

4. Finally, the FEnt can be computed as follows (Chen et al., 2007):

$$\mathrm{FEnt}(e, u, v, P) = \ln\left[\varphi^{e}(u, v)\right] - \ln\left[\varphi^{e+1}(u, v)\right] \tag{3.7}$$

The FEnt is used to characterize the surface EMG signals (Chen et al., 2007), epilepsy (Acharya et al., 2015a), and to discriminate non-focal and focal EEG signals (Sharma et al., 2015b).

3.2.7 Features ranking methods

Most significant features can be selected using features ranking methods. These methods assign ranks to the available features and arrange them according to their clinical significance. The lower ranked features can be neglected, and higher ranked features can be used for classification (Duda et al., 2000). These methods reduce the complexity of the system without affecting the classification performance. We have used four methods of features ranking: (1) Wilcoxon method, (2) Entropy method, (3) Receiver operating characteristic (ROC) method and (4) Bhattacharya space algorithm in this work. Wilcoxon method ranks based on non-parametric test (Derryberry et al., 2010). In ROC method, features are ranked according to the area under the ROC curve (Theodoridis & Koutroumbas, 2003). Entropy method utilizes the divergence method (Theodoridis & Koutroumbas, 2003) for measuring the separability between different classes. Bhattacharya space algorithm uses the Bhattacharya distance (Theodoridis & Koutroumbas, 2003) to measure the separability between the two classes.

3.2.8 Classification based on LS-SVM method

The SVM is a widely used machine learning technique in the field of pattern recognition. In SVM method, the features corresponding to different categories of data are classified by obtaining a hyperplane in the higher dimensional space (Suykens & Vandewalle, 1999). If SVM is employed with least square method, then it is termed as LS-SVM. It is prominently used in biomedical signal analysis applications. It is employed to analyze the EEG signals (Sharma et al., 2017a; Sharma & Pachori, 2017; Bhattacharyya et al., 2016, 2017a; Singh & Pachori, 2017) and analysis of

the heart sound signals (Patidar & Pachori, 2014; Patidar et al., 2015b). It is also used for automated diagnosis of diabetes in (Pachori et al., 2016) and glaucoma using fundus images in (Maheshwari et al., 2017a,b). LS-SVM is also utilized in the analysis of alcoholic EEG signals in (Patidar et al., 2017). Mathematically, it is represented as (Suykens & Vandewalle, 1999):

$$T = \text{sign}\left[\sum_{m=1}^{M} \alpha_m w_m F(y, y_m) + b\right] \tag{3.8}$$

where $F(y, y_m)$ represents a kernel function, α_m denotes the Lagrangian multiplier, y_m is the m^{th} input vector of D-dimension, b is used as bias term, and w_m is the target vector.

In the present study, RBF and Morlet wavelet are used as the kernel of LS-SVM. Kernel functions are used with LS-SVM to map the input space to the higher dimension space and the two classes can be separated using an optimal hyperplane (Suykens & Vandewalle, 1999).

RBF kernel can be expressed as (Khandoker et al., 2007):

$$F(y, y_m) = e^{\frac{-\|y-y_m\|^2}{2\gamma^2}} \tag{3.9}$$

In the RBF kernel, γ controls the width.

The expression of Morlet wavelet kernel is given as (Bajaj & Pachori, 2012; Zavar et al., 2011):

$$F(y, y_m) = \prod_{n=1}^{D} \cos\left[z_0 \frac{y^n - y_m^n}{l}\right] e^{\frac{-\|y^n - y_m^n\|^2}{2l^2}} \tag{3.10}$$

here l represents the scaling parameter of Morlet wavelet kernel and D denotes dimension of feature vectors set.

The other kernel functions that can also be used with LS-SVM are linear and polynomial kernels and mathematically defined as (Suykens & Vandewalle, 1999):

Linear kernel:

$$F(y, y_m) = y_m^T y \tag{3.11}$$

Polynomial kernel:

$$F(y, y_m) = (y_m^T y \ + \ 1)^r \tag{3.12}$$

where r is the order of the polynomial kernel.

In the present work, Ar, sensitivity (Ss), and specificity (Sc) (Azar & El-Said, 2014) are computed for performance evaluation.

3.3 Results

We have used FAWT to obtain the sub-band signals from the HRV signals of CAD and normal classes. In order to choose the level of decomposition, we computed probability (p) values for the features obtained from the sub-band signals at different levels of FAWT based decomposition using Kruskal-Wallis (KW) test. We observed lowest p-value at the third level of decomposition for FEnt computed from approximation signal (see Table 3.1). Lowest p-value indicates that this feature has highest discrimination ability. Therefore, we choose the third level of decomposition to analyze the CAD and normal HR signals. Furthermore, K-NN entropy estimator

Table 3.1: The p-values of features computed from the approximation signal at various levels of FAWT based decomposition for normal and CAD classes.

Features	Second level	Third level	Fourth level	Fifth level
FEnt	7.534×10^{-19}	1.11×10^{-19}	1.648×10^{-19}	4.738×10^{-18}
K-NN entropy estimator	9.407×10^{-7}	6.06×10^{-7}	2.99×10^{-7}	5.961×10^{-7}

and FEnt are computed from these sub-band signals. In our work, essential parameters required to compute K-NN entropy estimator and FEnt are selected using trial and error experimentation in such a way to maximize the classification Ar. For K-NN entropy estimator, the number of neighbours is selected to be 7. For FEnt, the value of e, u, and v are chosen to be 5, 2, and 0.3, respectively (Chen et al., 2007). The range of these features can be seen in Table 3.2.

The p-values computed using KW test (McKight & Najab, 2010) are also shown in Table 3.3. Significantly low p-values ($p < 0.05$) are observed for these features in Table 3.3, except K-NN entropy estimator computed from detail signal at the third level. Recently, KW test has been explored to test the statistical significance of the

Table 3.2: Mean (μ) and standard deviation (STD) of the features computed from sub-band signals of FAWT for normal and CAD classes.

Features	Sub-band signals	CAD class ($\mu\pm$ STD)	Normal class ($\mu\pm$ STD)
K-NN entropy estimator	D_1	-34.4144 ± 7.2939	-33.6003 ± 3.1676
	D_2	-38.2082 ± 7.8436	-38.1883 ± 2.4542
	D_3	-38.4634 ± 7.8750	-38.9568 ± 2.8859
	A_3	-25.1109 ± 4.1989	-22.5484 ± 1.7965
FEnt	D_1	0.1461 ± 0.0631	0.2471 ± 0.0653
	D_2	0.0724 ± 0.0698	0.0859 ± 0.0245
	D_3	0.0631 ± 0.0696	0.0640 ± 0.0254
	A_3	0.2767 ± 0.1171	0.5207 ± 0.1491

features in various biomedical signal analysis applications (Pachori, 2008; Pachori & Patidar, 2014; Sharma et al., 2015c; Bhati et al., 2017; Sharma et al., 2017e). The boxplots for K-NN entropy estimator and FEnt are depicted in Figures 3.4 and 3.5 respectively. N represents normal class, and C represents CAD class in Figures 3.4 and 3.5.

Table 3.3: The p-values of features computed from sub-band signals of FAWT for normal and CAD classes.

Features	D_1	D_2	D_3	A_3
K-NN entropy estimator	1.59×10^{-4}	2.73×10^{-2}	6.41×10^{-2}	6.06×10^{-7}
FEnt	6.14×10^{-18}	1.36×10^{-10}	9.10×10^{-7}	1.11×10^{-19}

Moreover, various ranking methods are used to achieve maximum Ar of classification using least number of feature vectors. Extracted features are arranged in

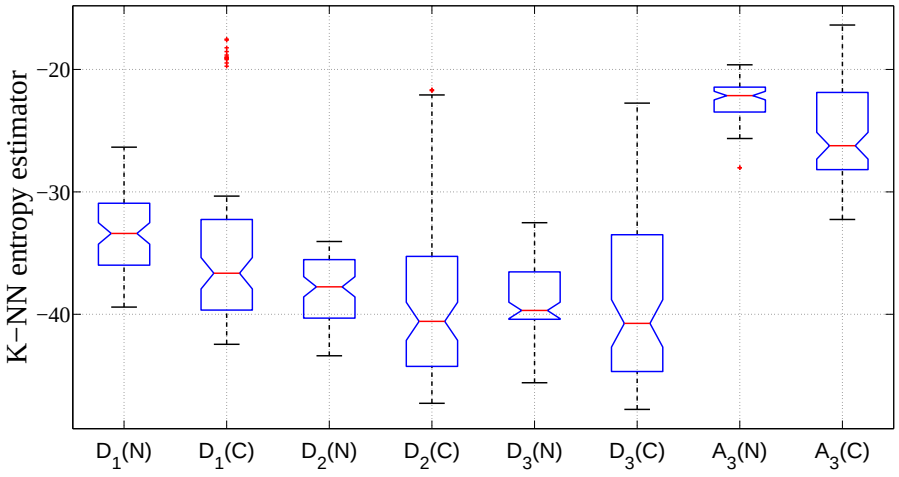


Figure 3.4: The boxplots showing ranges for K-NN entropy estimator.

the descending order according to their statistical significance. These ranked features are fed to the classifier one by one until the highest classification performance is reached. RBF and Morlet wavelet kernels are used with LS-SVM classifier to test the performance. For the performance validation of the LS-SVM, dataset is partitioned into training and testing sets using ten-fold cross-validation (TFCV) method (Kohavi, 1995). The TFCV method has widely used for the classification task (Sharma & Pachori, 2015; Sharma et al., 2015b, 2017c; Tiwari et al., 2017; Bhattacharyya et al., 2017b; Sharma & Pachori, 2018). A plot of the number of

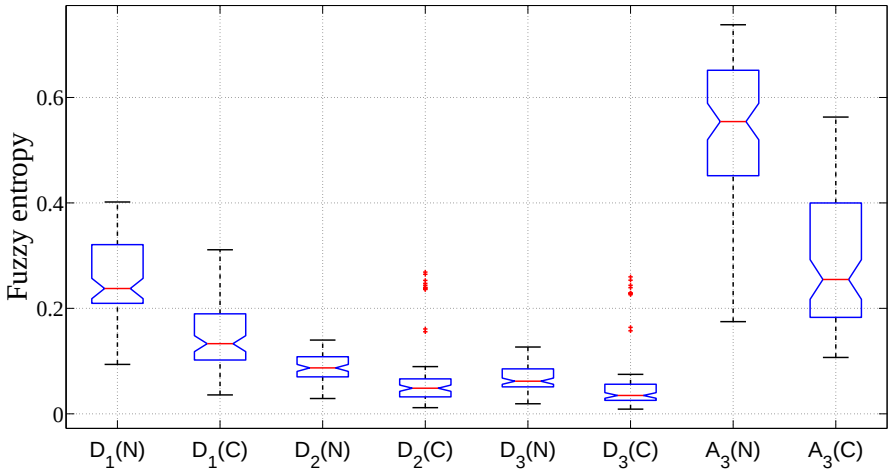
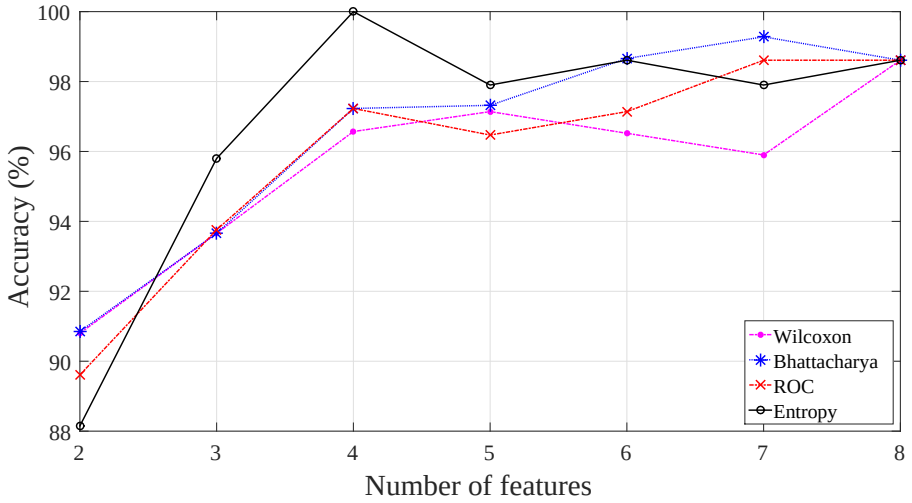


Figure 3.5: The boxplots for Fuzzy entropy.

features versus accuracies for the various ranking method is shown in Figures 3.6 and 3.7 for RBF and Morlet wavelet kernels, respectively. The LS-SVM attained



(a)

Figure 3.6: A plot showing accuracies for different number of features for RBF kernel.

maximum classification using entropy ranking method with four features, (see Figures 3.6 and 3.7). Maximum Ar of 100% is obtained for RBF and Morlet wavelet kernel functions. We have got 100% Ss and 100% Sc for RBF and Morlet wavelet kernels. These results are obtained using kernel parameter $\gamma = 0.4$ for RBF kernel, and $l = 1.05$ and $z_0 = 0.42$ for Morlet wavelet kernel.

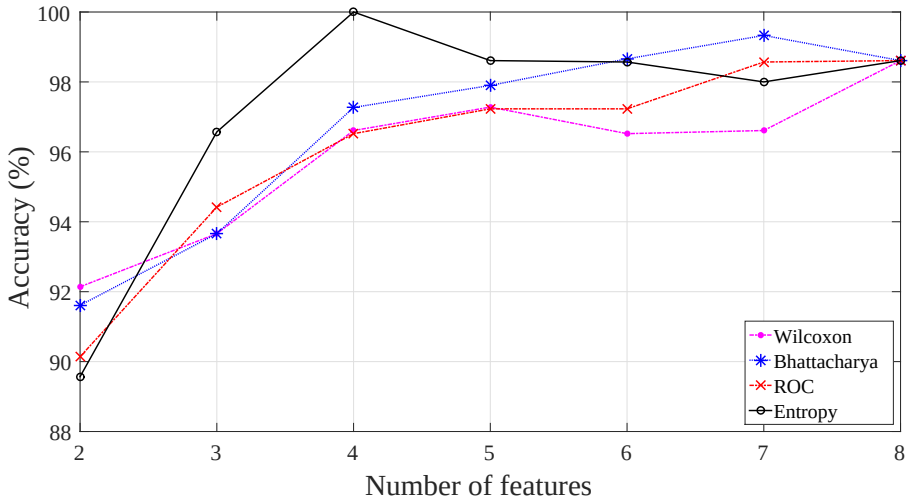


Figure 3.7: The variation of accuracies for different number of features with Morlet wavelet kernel.

3.4 Discussion

The performances of various methods used for automatic diagnosis of CAD subjects are summarised in Table 3.4. In (Karimi et al., 2005), heart sound signals (HSS) are used to diagnose CAD and normal subjects. The WPT and DWT are used to decompose the signals. Furthermore, variance, kurtosis, k-factor and skewness are computed and fed to the artificial neural network (ANN). The Ar of classification for abnormal cases is 85% and 90% for DWT and WPT respectively. In (Lee et al., 2007), nonlinear and linear features are used to detect the CAD subjects. The extracted features yielded the highest Ar of 90% using SVM classifier. In (Zhao & Ma, 2008), Teager energy operator (TEO) and EMD based methodology is used to diagnose CAD with HSS. Classification rate of 85% is obtained using back propagation neural network (BPNN).

In (Babaoglu et al., 2010), binary particle swarm optimization (BPSO) and genetic algorithm (GA) are used for feature selection from EST data. Thereafter, these features are tested with SVM classifier. Highest classification Ar of 81.46% is obtained using BPSO feature selection method. Furthermore, PCA is used to reduce the features extracted from EST data (Babaoğlu et al., 2010). These features in combination with SVM gave classification Ar of 79.17%. In (Dua et al., 2012), classification is performed using nonlinear features extracted from HRV signals. The PCA is applied to compute the principal components. These principal components are tested with eight different classifiers, and highest performance of 89.5% is achieved with multilayer perceptron (MLP). In (Giri et al., 2013), DWT is used to decompose the HRV signals and extracted features are reduced using ICA, PCA, and linear discriminant analysis (LDA). Highest classification Ar of 96.8% is attained using Gaussian mixture model (GMM).

Linear and nonlinear methods are applied to HRV signals to diagnose CAD patients (Poddar et al., 2015). Feature dimension is reduced using PCA and 91.67% Ar is reported using SVM classifier. In (Patidar et al., 2015a), HR signal decomposition is performed with TQWT, and correntropy features are extracted from the decomposed signals. Correntropies are further transformed using PCA to obtain more meaningful features. The 3-fold cross-validation method is used for the data partition. Finally, 99.72% average classification Ar is achieved using LS-SVM

classifier.

In this work, we have developed a new methodology to diagnose CAD using HRV signals based on FAWT decomposition method. These signals are decomposed upto third level using FAWT. K-NN entropy estimator and FEnt are used to fetch the information from detail and approximation signals at each level. K-NN entropy estimation is based on the distance of a sample from its K nearest neighbours, and it quantifies the degree of scattering of the time series (Veselkov et al., 2010). It should be noted that higher values of K-NN entropy estimator show a wider scattering of time series. On the other hand, low values of K-NN entropy estimator indicate the limited scattering of time series. From Table 3.2, we can observe higher values of K-NN entropy estimator for detail signals D_1, D_2, and approximate signal A_3 for normal subjects than CAD subjects. It shows slightly lower value for detail signal D_3 for normal subjects. However, for detail signal D_3, K-NN entropy estimator does not show discrimination ability as p-value is high ($p > 0.05$) in Table 3.3. For detail signals D_1, D_2, and approximate signal A_3 lower p-values ($p < 0.05$) are observed. Signal A_3 shows maximum discrimination ability for K-NN entropy estimator to separate the two classes as minimum p-value is observed for A_3 (see Table 3.3). Therefore, we are considering D_1, D_2, and A_3 sub-band signals to estimate K-NN entropy and study the dynamics of HRV signals. The lower values of K-NN entropy estimator for D_1, D_2, and A_3 for CAD subjects indicate that CAD HRV signals are less scattered compared to normal HRV signals.

FEnt is the measure of similarity in the time series. The measure of similarity is based on exponential function in (Chen et al., 2007). From Table 3.2, we can observe that, FEnt values for CAD subjects are smaller than normal subjects for all detail signals (D_1, D_2, and D_3) and approximation signal (A_3). All these signals (D_1, D_2, D_3, and A_3) show good discrimination ability for both classes ($p < 0.05$) in Table 3.3 for FEnt. For FEnt, approximation signal A_3 shows highest discrimination ability for both classes as p-value is found to be lowest (see Table 3.3). Moreover, the smaller value of FEnt shows more regularity of time series which means CAD HR signals have more regularity than normal HR signals. Finally, for RBF and Morlet wavelet kernels, maximum classification Ar of 100% is obtained using our proposed algorithm.

Table 3.4: Comparison of the classification performance of the proposed work with the existing work.

Authors	Methods	Features	Classifier	Ar (%) of classification
Karimi et al. (2005)	HSS, DWT, and WPT	Various statistical features	ANN	90
Lee et al. (2007)	HRV, linear & nonlinear methods	6 Linear & 5 nonlinear features	SVM	90
Zhao and Ma et al. (2008)	HSS, EMD, and TEO	Various statistical features	BPNN	85
Babaoglu et al. (2010)	EST, GA, and BPSO	11 Features	SVM	81.46
Babaŏglu et al. (2010)	EST and PCA	18 Features	SVM	79.17
Dua et al. (2012)	HRV and PCA	6 Nonlinear features	MLP	89.5
Giri et al. (2013)	HRV, DWT, and ICA	10 Features	GMM	96.8
Poddar et al. (2015)	HRV, linear & nonlinear methods	Various linear & nonlinear features	SVM	91.67
Patidar et al. (2015)	HRV, TQWT and PCA	2 Correntropy features	LS-SVM with Morlet wavelet kernel	99.72
Present work	HRV and FAWT	K-NN entropy estimator and FEnt	LS-SVM with RBF & Morlet wavelet kernel	100

The main advantage of our work as compared to other existing works is that the proposed method uses less number of features (only four) to obtain the highest Ar. However, (Patidar et al., 2015a) also achieved a classification Ar of 99.72% with two features. They used 3-fold cross-validation method, while we used TFCV method, which makes our method more robust and reliable as compared to the method proposed in (Patidar et al., 2015a). Therefore, the proposed system in this paper is suitable for clinicians and can be installed in the hospitals to detect CAD automatically using HR signals. As we are using four features, the diagnosis of CAD will be fast. In future, home telemonitoring can also be employed with this system to make it more dynamic. Therefore, the proposed methodology can help to save the life of the CAD patients. The limitation of the present work is small data set. Before the deployment of the proposed system for clinical purpose, it needs to be tested using large data set.

3.5 Summary

In the present work, a novel technique is proposed for the diagnosis of CAD patients based on FAWT decomposition method. In our work, HR signals are decomposed up to the third level using FAWT. The K-NN entropy estimator and FEnt are computed to fetch the nonlinear dynamics of HR signals. Several ranking methods are used to optimize the classification performance. Highest classification performance is observed with initial four significant parameters computed using entropy ranking method. Use of TFCV procedure makes the system more robust and reliable. Hence, it can be concluded by the observed performance that the proposed methodology is very efficient to diagnose the CAD subjects.

Chapter 4

An automated diagnosis method for CAD patients using ECG signals based on FAWT

4.1 Introduction

The commonly used methods for CAD detection by the doctors and their shortcomings are explained in chapter 3.

In literature, various computer-aided diagnostic methods are suggested based on ECG signals (Antanavičius et al., 2008; Arafat et al., 2005) and HRV signals (Acharya et al., 2014; Bigger et al., 1995; Karamanos et al., 2006; Sood et al., 2016). Nonlinear features such as, mutual information, embedding dimension error, fractal dimension, and recurrence percentage are found significantly different for ECG signals of CAD and non-CAD subjects (Antanavičius et al., 2008). In (Arafat et al., 2005), stress ECG signals are studied using fuzzy, probabilistic, and combined uncertainty models. Literature review on the analysis of CAD patients using HRV signals is provided in chapter 3.

The aim of this work is to develop an accurate, fast, and automated system to diagnose CAD non-invasively using ECG signals. In order to achieve this, first, we segmented the ECG signals into beats. Then, each beat is subjected to FAWT which decomposes it into detail and approximation coefficients. Further, CIP parameters are computed from the detail coefficients. In the end, classification is performed

using these CIP features. Flowchart of the proposed methodology is shown in Figure 4.1. The remaining sections of this chapter are arranged in the following way:

Section 4.2 presents the dataset, preprocessing steps, decomposition of the beats, and computation of CIP. The obtained results are described in section 4.3. Section 4.4 gives the discussion of the results. Finally, the summary of the work is presented in section 4.5.

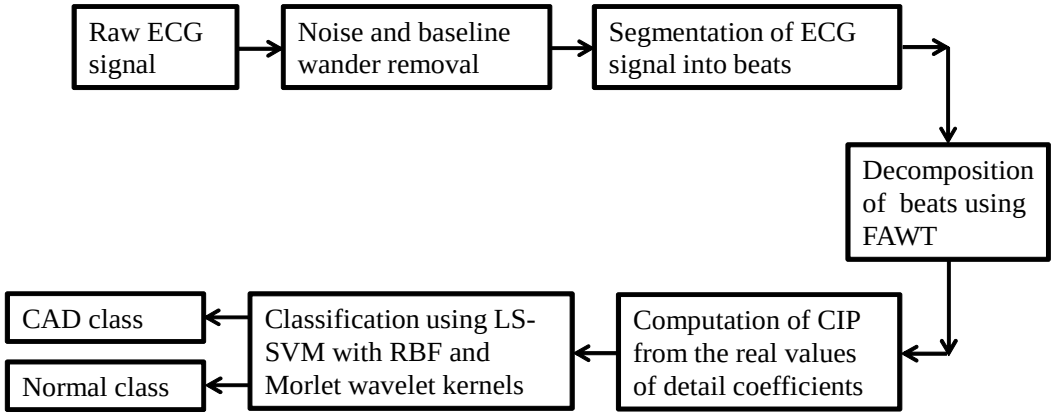


Figure 4.1: The automated diagnosis method for CAD disease using ECG signals.

4.2 Methodology

4.2.1 Dataset

In the present study, the ECG signals of CAD subjects are downloaded from St. Petersburg Institute of Cardiological Technics 12-lead Arrhythmia Database (Goldberger et al., 2000). The ECG signals of normal subjects are downloaded from Fantasia open-access database (Iyengar et al., 1996; Goldberger et al., 2000). In the present work, ECG signals of 40 normal subjects (20 young and 20 old subjects) and 7 CAD subjects are used. We have used ECG signals of lead-II in this work.

4.2.2 Preprocessing of downloaded ECG signals

The recording duration of the downloaded CAD and normal ECG are 0.5 hours and 2 hours respectively. Sampling frequency of CAD and normal ECG signals are 257 and 250 samples per second, respectively. In the present work, to make

the ECG dataset uniform, normal ECG signals are up-sampled to 257 samples per second. The baseline wander and noise exist in the ECG signals are eliminated using Daubechies 6 (db6) basis function of wavelet (Martis et al., 2013a,b).

4.2.3 Beats segmentation from preprocessed ECG signals

To segment the ECG signals into individual beats, first R-peak need to be detected. In this work, Pan-Tompkin's algorithm (Pan & Tompkins, 1985) is used to detect the R-peak from each ECG signal (Martis et al., 2013b). On the basis of detected R-peaks, ECG signals are segmented into beats. Each ECG beat consists of 64 samples before the R peak and 104 samples after R peak. Thus, one ECG beat is a segment of 169 samples and covers P, QRS, and T waves. In this work, 44,426 beats of CAD ECG signals and 137,587 beats of normal ECG signals are used.

4.2.4 Decomposition of the beats using FAWT

In this work, we have applied FAWT for the decomposition of ECG beats. The values of the parameters that are used in the present work are $b = 5$, $c = 6$, $f = 1$, $g = 2$, and $\beta = 0.8(f/g)$ (Zhang et al., 2015). Real values of the coefficients obtained after decomposing the ECG beats using FAWT are shown in Figures 4.2(a) and 4.2(b) for CAD and normal subjects, respectively. In Figures 4.2(a) and 4.2(b), CV is the magnitude of the coefficient, and CN represents the corresponding coefficient. These coefficients are the detail coefficients. CV_{1a} and CV_{1b} are representing detail coefficients of 1^{st} level. Likewise, CV_{2a} and CV_{2b}, CV_{3a} and CV_{3b}, CV_{4a} and CV_{4b}, and CV_{5a} and CV_{5b} represent 2^{nd} level, 3^{rd} level, 4^{th} and 5^{th} level detail coefficients, respectively.

4.2.5 Feature computation from the detail coefficients

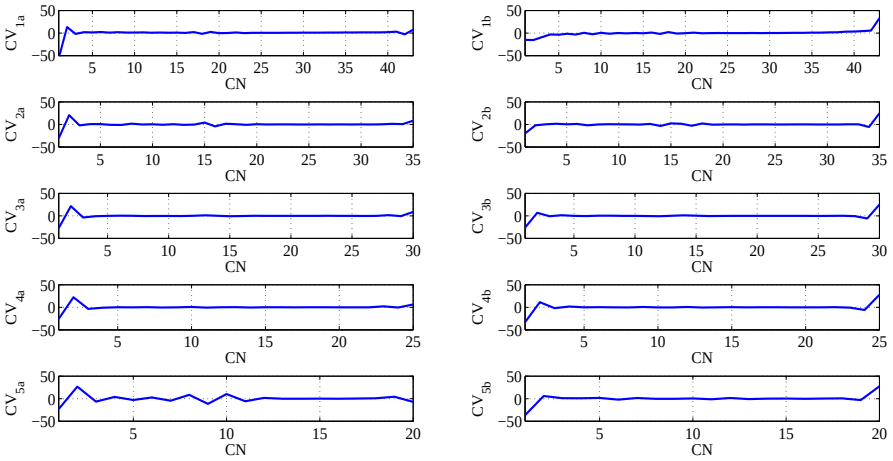
In the present work, CIP is computed between the real values of detail coefficients at each level.

Information potential is the non-parametric estimator of Renyi's quadratic entropy (Xu & Erdogmuns, 2010). It can be estimated using kernel-based approach

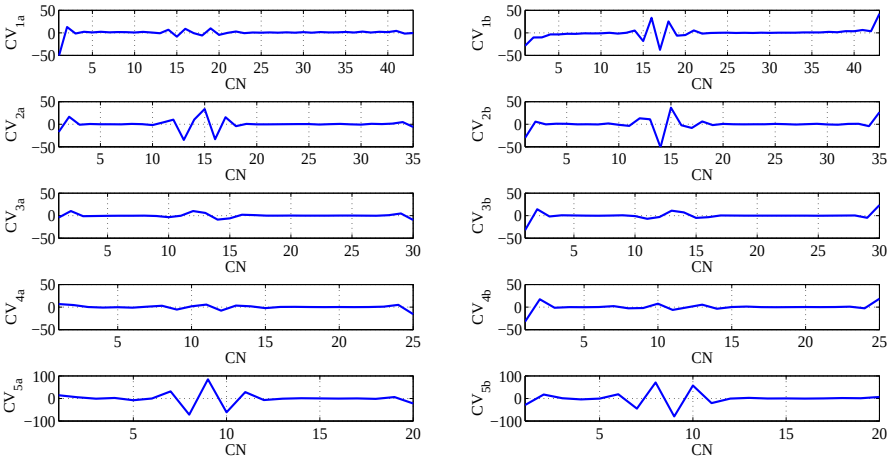
as (Xu & Erdogmuns, 2010):

$$\hat{V}(X) = \frac{1}{P^2} \sum_{a=1}^{P} \sum_{b=1}^{P} k(x_a - x_b) \tag{4.1}$$

where, $k(x_a - x_b)$ represents the kernel function, and P is the total number of samples. x_a and x_b are the a^{th} and b^{th} samples of the signal.



(a)



(b)

Figure 4.2: The plots of real coefficients obtained using FAWT based decomposition: (a) CAD subject and (b) Normal subject

CIP is the measure of similarity between the two probability density functions (Xu et al., 2008). CIP can be estimated as (Xu & Erdogmuns, 2010):

$$\hat{V}(X,Y) = \frac{1}{P^2}\sum_{a=1}^{P}\sum_{b=1}^{P} k(x_a - y_b) \tag{4.2}$$

x_a is the a^{th} sample of the signal X and y_b is the b^{th} sample of the signal Y. In the present work, we have used ITL toolbox (http://www.sohanseth.com/Home/codes) for computing the CIP, which uses the incomplete Cholesky decomposition to compute the CIP. Gaussian kernel is used to compute the CIP and kernel size (γ) of 1 is used in the present work.

4.2.6 Classification using LS-SVM

The LS-SVM method is briefly explained in chapter 3.

4.3 Results

In the present study, resting ECG signals are segmented into beats to diagnose CAD. The segmented beats are subjected to FAWT decomposition technique to get detail and approximation coefficients. Further, CIP parameters are computed from detail coefficients. In this work, we have analysed the ECG beats up to five decomposition levels. We have applied KW statistical test (McKight & Najab, 2010) and computed the p-values of features at different levels of decomposition. The p-value, μ and STD of CIP parameters at various levels of decomposition are provided in Table 4.1. We can observe from the table that, for normal subjects, CIP parameters have lower values at each decomposition level as compared to CAD beats. The box plots at each level of decomposition are computed using KW statistical test and shown in Figure 4.3.

Clinically significant p-values of CIP parameters at each level of decomposition can be seen in Table 4.1. At each level of decomposition, the effectiveness of CIP features is evaluated by computing Ar, Ss, Sc and Matthews correlation coefficient (MCC). The RBF and Morlet wavelet are used as kernel functions of LS-SVM classifier. The kernel parameters used in this work are shown in Table 4.2.

Table 4.1: The μ and STD values of CIP features computed from various levels of FAWT decomposition for normal and CAD ECG beats.

Level of decomposition	Features	CAD class ($\mu\pm$ STD)	Normal class ($\mu\pm$ STD)	p-values
1	CIP_{D1}	0.1987 ± 0.0734	0.1488 ± 0.1096	0
2	CIP_{D2}	0.1857 ± 0.0524	0.1595 ± 0.0768	0
3	CIP_{D3}	0.2281 ± 0.0546	0.1860 ± 0.0883	0
4	CIP_{D4}	0.2004 ± 0.0685	0.1593 ± 0.1045	0
5	CIP_{D5}	0.0830 ± 0.0417	0.0740 ± 0.0621	0

Table 4.2: Results of classification at different levels of decomposition using different kernel functions.

Level of decomposition	Kernel	Kernel parameter values	Ar(%)	Ss(%)	Sc(%)	MCC
Level 1	RBF	$\gamma = 1$	89.53	90.48	89.23	0.769
	Morlet wavelet	$l = 1.2$ $z_0 = 0.25$	89.58	90.16	89.39	0.769
Level 2	RBF	$\gamma = 1$	96.44	94.74	96.98	0.906
	Morlet wavelet	$l = 1.2$ $z_0 = 0.25$	96.51	94.93	97.02	0.908
Level 3	RBF	$\gamma = 1$	98.37	97.62	98.61	0.957
	Morlet wavelet	$l = 1.2$ $z_0 = 0.25$	98.53	97.87	98.74	0.961
Level 4	RBF	$\gamma = 1$	99.54	99.36	99.60	0.987
	Morlet wavelet	$l = 1.2$ $z_0 = 0.25$	99.58	99.45	99.62	0.988
Level 5	RBF	$\gamma = 1$	99.56	99.50	99.58	0.988
	Morlet wavelet	$l = 1.2$ $z_0 = 0.25$	99.60	99.57	99.61	0.989

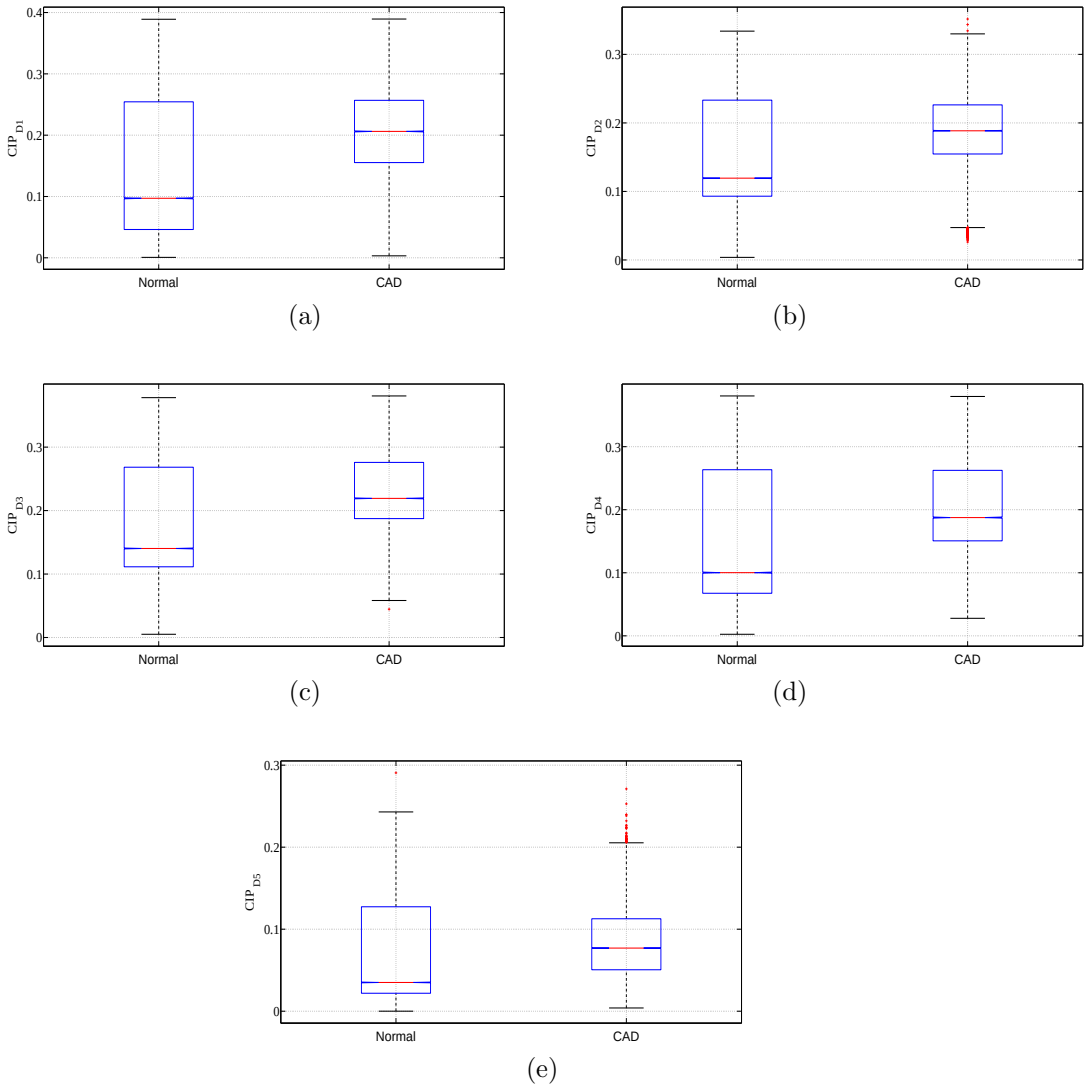


Figure 4.3: Box plots of CIP features at various decomposition levels of CAD and normal ECG beats: (a) Level-1, (b) Level-2, (c) Level-3, (d) Level-4 (e) Level-5.

Performance of the classifier is validated using TFCV method (Kohavi, 1995). At first level of decomposition, the classification accuracies are 89.53% and 89.58% with RBF and Morlet wavelet kernels, respectively. Remaining classification parameters can be observed in Table 4.2. Further, we have obtained classification accuracies of 96.44% for RBF kernel and 96.51% for Morlet wavelet kernel at the second level of decomposition. At this level, we have used CIP parameters (CIP_{D1} and CIP_{D2}) computed from both levels of detail coefficients. The values of Ar, Ss, Sc, and MCC are tabulated in Table 4.2. Significant improvement is observed in classification performance at the second level of decomposition as compared to the first level of decomposition.

At the third level of decomposition, the classification Ar of 98.37% and 98.53% are obtained using RBF and Morlet wavelet kernel functions, respectively. All four CIP features ($CIP_{D1}, CIP_{D2}, CIP_{D3}$, and CIP_{D4}) are fed to the LS-SVM classifier and observed an Ar of 99.54% and 99.58% using RBF and Morlet wavelet kernel functions at the fourth level of decomposition. The classification performance is improved at the fourth level of decomposition. At the fifth level, five CIP parameters ($CIP_{D1}, CIP_{D2}, CIP_{D3}, CIP_{D4}$, and CIP_{D5}) are applied to the input of LS-SVM. We have obtained classification performance parameters, Ar=99.56%, Ss=99.50%, Sc=99.58% and MCC=0.988, respectively for RBF kernel. The classification performance parameters, Ar=99.60%, Ss=99.57%, Sc=99.61% and MCC=0.989 are obtained for Morlet wavelet kernel. We did not observe significant improvement when we increased the decomposition from fourth to fifth level. Therefore, we did not increase the decomposition level further. The Ar, Ss, and Sc for various folds of TFCV is shown in Figures 4.4 and 4.5 for RBF and Morlet wavelet kernels, respectively.

4.4 Discussion

A summary of the performance of various diagnostic techniques used for CAD detection is presented in Table 4.3. In (Acharya et al., 2017e), higher order statistics (HOS) are used for diagnosis of CAD and normal subjects using ECG signals. HOS bispectrum and cumulant features are studied on the ECG beats. A classification Ar of 98.99% is achieved when 31 cumulant features are fed to the decision tree

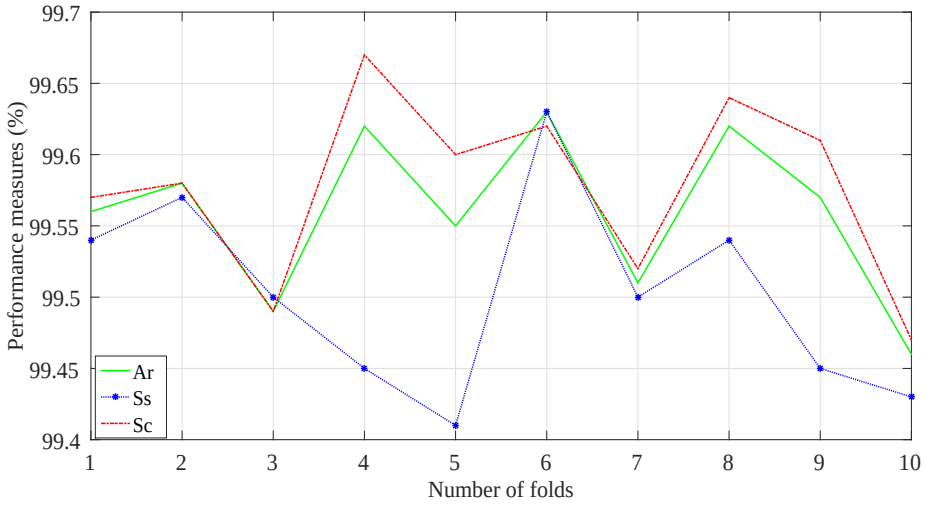


Figure 4.4: The plot of performance measures versus the number of folds for LS-SVM classifier for RBF kernel at 5^{th} level of decomposition.

(DT) classifier. The other methods for the detection of CAD are already discussed in chapter 3.

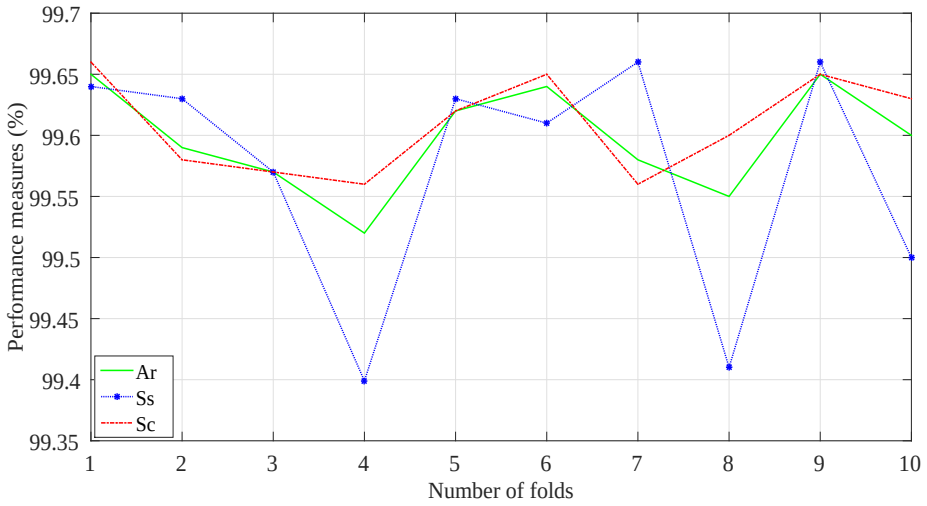


Figure 4.5: A plot of performance measures versus the number of folds when Morlet wavelet is used with LS-SVM classifier as a kernel (5^{th} level of decomposition).

In the present work, FAWT decomposition method and CIP are used to capture the hidden information from ECG beats of 40 normal and 7 CAD subjects. The advantage of using FAWT is that it separates the detail components from ECG beats by adjusting the Q factor, redundancy and dilation factor. Hence, more meaningful information can be extracted using FAWT. Furthermore, CIP is able

to extract information from detail coefficients provided by FAWT decomposition. CIP parameters are larger for CAD ECG beats. CIP parameter is the measure of

Table 4.3: Comparison of the proposed automated classification method of CAD and normal classes with the other existing method.

Authors	Data set used	Methodology	Classifier used	Ar %
Acharya et al. (2016)	ECG signals 40 normal and 7 CAD subjects	HOS 31 cumulant features	DT	98.99
In the present work	ECG signals 40 normal and 7 CAD subjects	FAWT and CIP feature	LS-SVM and Morlet wavelet kernel	99.60

similarity (Xu et al., 2008). Hence, larger values of CIP parameters indicate more similarity between the detail coefficients of FAWT based decomposition of CAD beats as compared to the normal beats. The total number of ECG beats used in our work is 182,013 (44,426 CAD ECG beats and 137,587 normal ECG beats). The size of the feature matrix is 182,013×5. This is a huge dataset of ECG beats. The present methodology is able to classify the 182,013 ECG beats with an Ar of 99.60% with LS-SVM classifier with TFCV method.

The advantage of the proposed algorithm is that we have used resting ECG signals. Hence, there is no risk of possible cardiac arrest during recording as in stress ECG signals. The developed algorithm is simple and fast. It is robust and repeatable as we employed TFCV. The limitation of the present work is that we have used only 7 CAD subjects.

4.5 Summary

In this work, a new methodology is proposed to diagnose the CAD automatically using ECG signals. First, ECG signals are segmented and decomposed using FAWT.

Five CIP parameters are computed from the five levels of FAWT detail coefficients. The performance of the method is examined at different levels of decomposition from $J = 1$ to $J = 5$. The CIP parameters with LS-SVM and RBF kernel yielded a classification Ar of 99.54% and 99.56% at the fourth and fifth level of decomposition respectively. The Ar of classification using Morlet wavelet kernel improved to 99.58% and 99.60% at the fourth and fifth level of decomposition respectively. The use of TFCV technique makes the proposed methodolpgy more robust. The present method can be extended for early detection of CAD.

Chapter 5

Automated identification method for normal and MI ECG signals using FAWT

5.1 Introduction

MI is a condition which indicates the injury of a heart cell due to the lack of oxygenated blood in the cardiac arteries (Thygesen & et al, 2012). The main cause of MI is the coronary heart disease which is responsible for nearly one-third of all deaths in the age group of above 35 years (Sanchis-Gomar et al., 2016; Acharya et al., 2016). MI is silent in nature and may lead to fast and non-recoverable damage to the muscles of heart (Acharya et al., 2016). If MI is not controlled timely then, myocardial structure and function of the LVe continue to damage further.

For the diagnosis of MI, the ECG is used due to its low operating cost and non-invasive nature (Lu et al., 2000). Vital information related to the functioning of the heart can be assessed by analyzing the ECG signals. Moreover, the MI results in the ST deviations and T wave abnormalities in the ECG signal (Lu et al., 2000). Manual identification of the changes in the ECG signals is a difficult task. Only 82% ST-segment elevation in MI subjects may be recognised by the experienced cardiologists (Liu et al., 2015). Therefore, an automated identification system for MI patients is needed to facilitate the clinicians in their accurate diagnosis. The classification of ECG signals and the extracted beats from ECG signals have been

studied in the literature for diagnosis of heart disorders (Crippa et al., 2015; Biagetti et al., 2014; Li et al., 2016).

In literature, various studies are performed for the detection of MI patients based on time-domain, frequency-domain, wavelet-domain, and nonlinear analysis. In (Safdarian et al., 2014), total integral of one ECG cycle and T-wave integral are suggested as features for the detection of MI. Time-domain features computed from 12-lead ECG signals are explored with fuzzy-MLP network to classify the MI ECG signals (Bozzola et al., 1996). In (Sun et al., 2012), a new multiple instance learning based approach is proposed for the detection of MI. The Hermite basis functions are used to decompose the multi-lead ECG signals, and the obtained coefficients are found effective for the detection of acute MI (Haraldsson et al., 2004). The phase space fractal dimension features and the ANN classifier are explored to detect the MI (Lahiri et al., 2009).

In (Lu et al., 2000), authors have applied neuro-fuzzy approach for the diagnosis of MI patients using multi-lead ECG signals. A hybrid approach based on hidden Markov models (HMMs) and GMMs is proposed to distinguish the MI ECG signal from the normal ECG signal in (Chang et al., 2009). In (McDarby et al., 1998), characterization of the QRS complex of normal and MI subjects is performed using DWT. Three different wavelets are used to decompose the ECG signals up to the fourth level of decomposition. The Daubechies wavelet performed best among the three chosen wavelets. The DWT technique is incorporated to extract the QRS complex of ECG signals, and it is found that identification of MI subject is possible by detecting the QRS complex (Banerjee & Mitra, 2010).

In (Tripathy & Dandapat, 2016), the phase of the complex wavelet coefficients obtained from the dual-tree complex wavelet transform of 12-lead ECG signals is computed. Then, multiscale phase alteration values are used as features to identify the normal, MI, and other abnormal ECG signals. In (Acharya et al., 2017a), the ECG signals of normal, MI, and CAD are applied to DWT, EMD, and discrete cosine transform (DCT) techniques. The authors in this study achieved the best performance when features obtained using DCT techniques are subjected to the K-NN classifier. Counterlet transform (CT) and shearlet transform (STm) based technique is proposed to distinguish normal, MI, CAD, and CHF subjects using

ECG beats in (Acharya et al., 2017d). The performance of the CT based technique is found to be better in comparison to the STm based method.

Our aim is to develop an automated diagnosis of MI patients in this work. We have analyzed normal and ECG beats using SEnt in FAWT (Bayram, 2013; Zhang et al., 2015) framework. First, preprocessing is performed to remove the baseline wandering and other noise present in the ECG signals. Then, ECG signals are segmented into the beats. Further, these beats are decomposed up to 24^{th} level of decomposition using FAWT. The SEnt is computed from each sub-band signal which is reconstructed from the corresponding coefficients of the FAWT based decomposition. The computed features are subjected to the RF (Breiman, 2001), J48 DT (Quinlan, 1986, 1993), BPNN (Jang et al., 1997), and LS-SVM classifiers (Suykens & Vandewalle, 1999) for separating the ECG beats of MI and normal classes. The steps performed in the present work are shown in Figure 5.1. The organization of the remaining sections of the chapter is as follows:

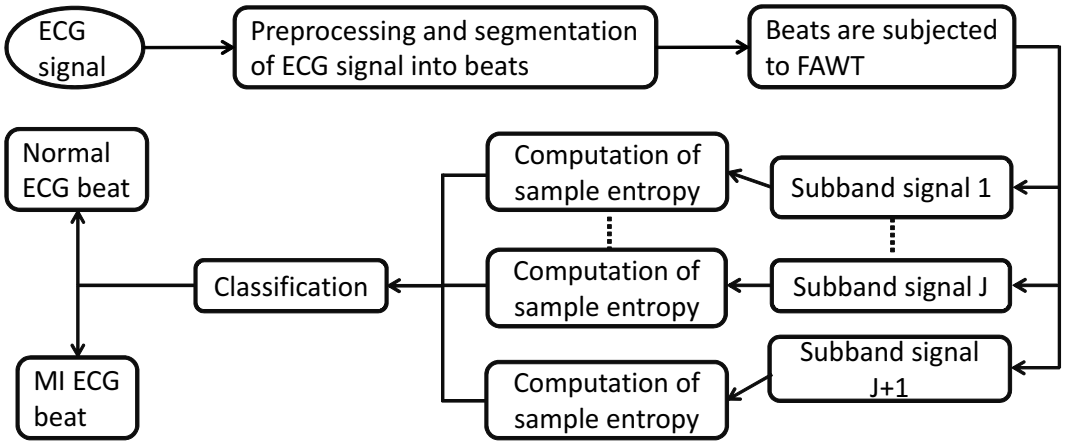


Figure 5.1: The proposed method for automated identification of the MI patients.

The dataset used, preprocessing, segmentation of the ECG signals into beats, SEnt, and classification methods are provided in section 5.2. The obtained results in this work are given in section 5.3 and discussed in section 5.4. Finally, section 5.5 presents the summary of the work.

5.2 Methodology

5.2.1 Dataset studied in this work

The dataset, containing normal and MI ECG signals, were downloaded from Physikalisch Technische Bundesanstalt (PTB) diagnostic ECG 12-lead database from the physiobank (Bousseljot et al., 1995; Goldberger et al., 2000). Each signal was acquired at the sampling rate of 1000 Hz. The dataset contains normal ECG recordings of 52 subjects and MI ECG recordings of 148 subjects. The ECG signals obtained from the lead-2 have been used in present work.

5.2.2 Preprocessing and segmentation of ECG signals

We have used db6 wavelet basis function to eliminate baseline wander and noise present in the ECG signals (Martis et al., 2013a,b). After preprocessing, each ECG signal is segmented into beats based on R-peak detection. The Pan–Tompkins algorithm is applied to identify the R-peaks (Pan & Tompkins, 1985). The 250 samples from the left and 400 samples from the right of the R-point are considered as one ECG beat (Acharya et al., 2016). Thus, each ECG beat contains 651 samples. Finally, we have 40182 MI ECG beats and 10546 normal ECG beats.

5.2.3 Sample entropy

SEnt (Richman & Moorman, 2000) measures the complexity of the time series. It improves the performance by excluding the bias due to the self-matches counted in the computation of approximate entropy. Higher values of SEnt indicate more complexity of the signal, on the other hand, the lower value of SEnt shows less complexity of the signal.

Let us consider a time-series $(y_1, y_2, ..., y_P)$ of length P for which the SEnt can be computed as (Sokunbi et al., 2013):

$$\mathrm{SEnt}(e, r, P) = -\ln\left[\frac{I^{e+1}(r)}{I^{e}(r)}\right] \tag{5.1}$$

where $I^e(r)$ is defined as follows (Sokunbi et al., 2013):

$$I^e(r) = \frac{1}{(P - e\tau)} \sum_{j=1}^{P-e\tau} C_j^e(r) \tag{5.2}$$

and

$$C_j^e(r) = \frac{S_j^r}{P - (e+1)\tau} \tag{5.3}$$

where S_j^r is the total count for which $L\left[Y(j), Y(k)\right] \leq r$ without considering the self-matches. The parameter $L\left[Y(j), Y(k)\right]$ is the distance between $Y(j)$ and $Y(k)$ vectors. $Y(j)$ and $Y(k)$ can be given as (Sokunbi et al., 2013):

$Y(j) = \{y_j, y_{j+\tau}, ..., y_{j+(e-1)\tau}\}$

$Y(k) = \{y_k, y_{k+\tau}, ..., y_{k+(e-1)\tau}\}$

where, j and k vary from 1 to $P - e\tau$ and $k \neq j$.

In this work, we have experimentally chosen threshold $(r) = 0.35$, delay $(\tau) = 1$, and embedding dimension $(e) = 5$. Parameter selection procedure has been explained in the section 5.3.

5.2.4 Studied classification techniques

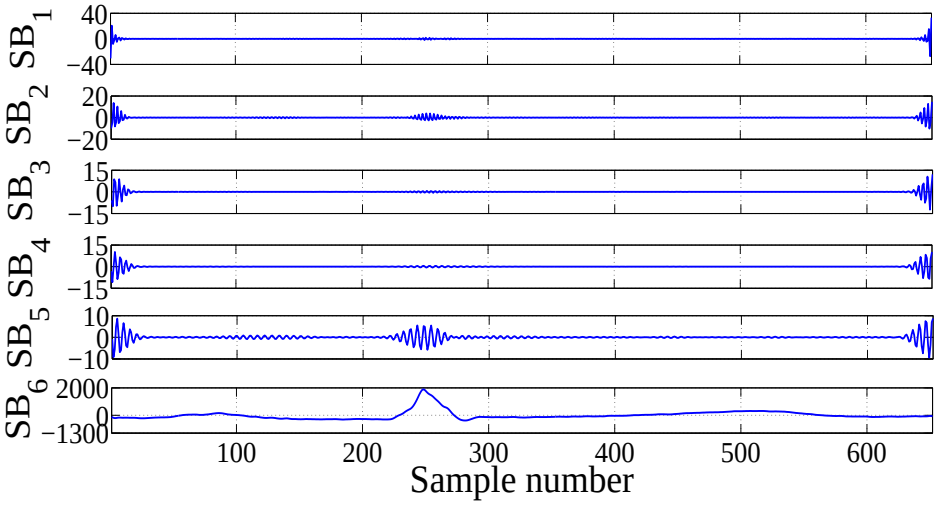
We have used RF, J48 DT, BPNN, and LS-SVM in this work to perform the classification of normal and MI ECG beats based on the extracted features. In this work, we have used Waikato environment for knowledge analysis (WEKA) toolbox for the implementation of RF, J48 DT, and BPNN classifiers (Hall et al., 2009). We have used default parameters provided in WEKA toolbox for RF, J48 DT, and BPNN classifiers. Recently, RF, J48 DT, and BPNN classifiers have been used to analyze the sleep stages in (Sharma et al., 2017d).

In the present work, we have evaluated the classification performance in terms of Ar, Ss, and Sc (Azar & El-Said, 2014).

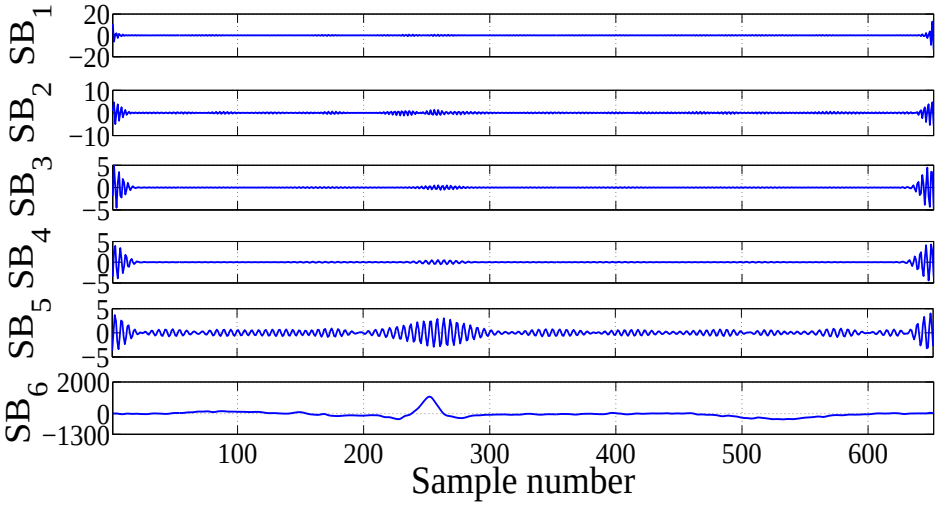
5.3 Results

In this work, we have computed SEnt in FAWT domain to classify MI and normal ECG beats. We have used $b = 5$, $c = 6$, $f = 1$, $g = 2$ and $\beta = (0.8 \times f)/g$ (Zhang et al., 2015) in the present work. First, we have segmented the MI and normal ECG

signals into the beats. Each ECG beat of both classes is decomposed into different sub-band signals using FAWT. The SEnt are computed from these different sub-band signals. We start performing the experiments with $J = 5$ and initial parameters for SEnt $e = 2$, $\tau = 1$, and $r = 0.15$ are chosen (Sokunbi et al., 2013). Typical sub-band signals extracted from the decomposition of normal and MI ECG beats at 5^{th} level of decomposition using FAWT are shown in Figures 5.2(a) and 5.2(b), respectively. In Figure 5.2, sub-band signals SB_1 to SB_5 are reconstructed from the detail coefficients from level 1 to level 5 and SB_6 is reconstructed from the approximate coefficients at level 5.



(a)



(b)

Figure 5.2: The plots of decomposed sub-band signals: (a) Normal ECG beat, (b) MI ECG beat.

We fed the features to the RF classifier for selecting the suitable parameters. Variation of classification accuracies for various values for e and τ is provided in Table 5.1. It can be inferred from the table that Ar of classification is highest for $e = 5$ and $\tau = 1$. Classification Ar for various values of r is shown in Figure 5.3. We have achieved the maximum classification Ar for $r = 0.35$. Hence, we have used the parameters $e = 5$, $\tau = 1$, and $r = 0.35$ to compute the SEnt in this work. Moreover, we have increased the decomposition level to $J = 6$, and observed that the classification Ar is increased to 91.95%. Hence, we further increased the decomposition level up to maximum possible decomposition level using FAWT with parameters values $b = 5$, $c = 6$, $f = 1$, $g = 2$ and $\beta = (0.8 \times f)/g$.

Table 5.1: Classification accuracies computed using RF classifier for different values of e and τ using SEnt with $r = 0.15$.

$e \longrightarrow$ $\tau\downarrow$	2	3	4	5
1	87.716%	89.353%	89.353%	89.629%
2	88.92%	89.128%	89.32%	89.075%
3	89.126%	88.84%	88.739%	88.84%

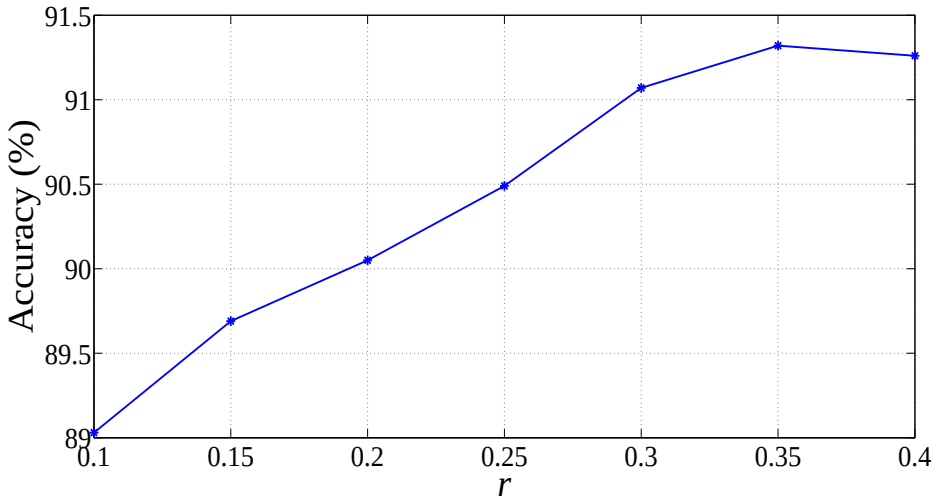


Figure 5.3: The plot of Ar (%) versus r of SEnt with RF classifier.

The plot of classification Ar versus decomposition levels is shown in Figure 5.4.

We can observe that Ar is increasing with increase in the decomposition level. The highest classification Ar of 97.10% is attained with RF classifier at $J = 24$. We have employed TFCV procedure for the training and testing of the classifier (Kohavi, 1995). The classification Ar achieved using J48 DT and BPNN classifiers are 93.97% and 92.85%, respectively.

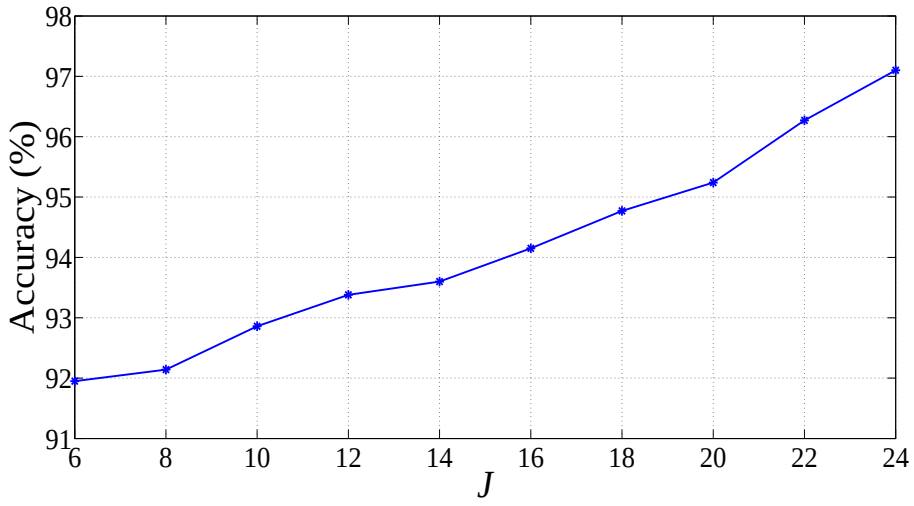


Figure 5.4: A plot of Ar (%) versus level of decomposition with RF classifier.

Further, we have also tested the features with LS-SVM classifier with different kernel functions namely, polynomial, linear, RBF, and Morlet wavelet kernels at $J = 24$. Variation of classification Ar (%) with the RBF kernel parameter γ is shown in Figure 5.5. We can observe from Figure 5.5 that classification Ar of LS-

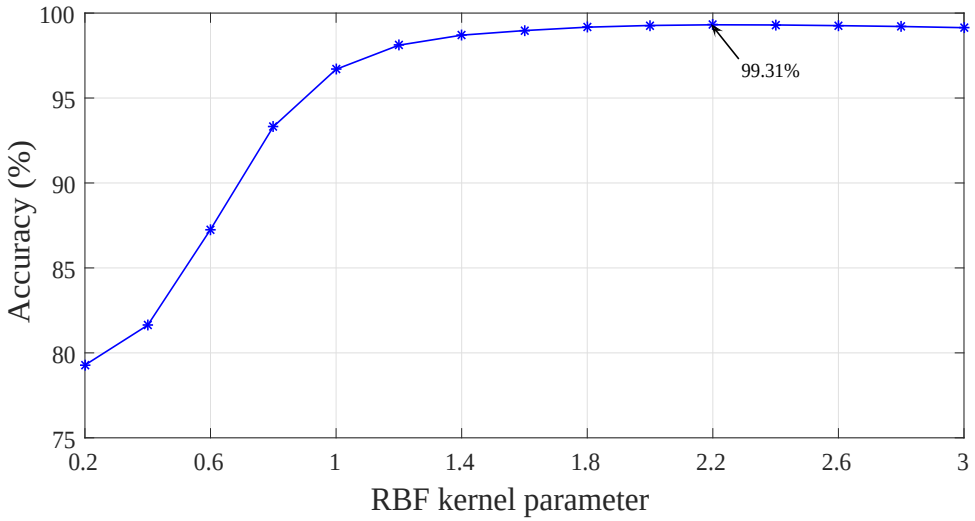


Figure 5.5: The plot of accuracies versus parameter γ of RBF kernel.

SVM is maximum for RBF kernel parameter $\gamma = 2.2$. A plot of changes in the value of Ar (%) for variation in the parameter l of Morlet wavelet kernel is shown in Figure 5.6. LS-SVM showed maximum Ar with Morlet wavelet kernel at $l =$

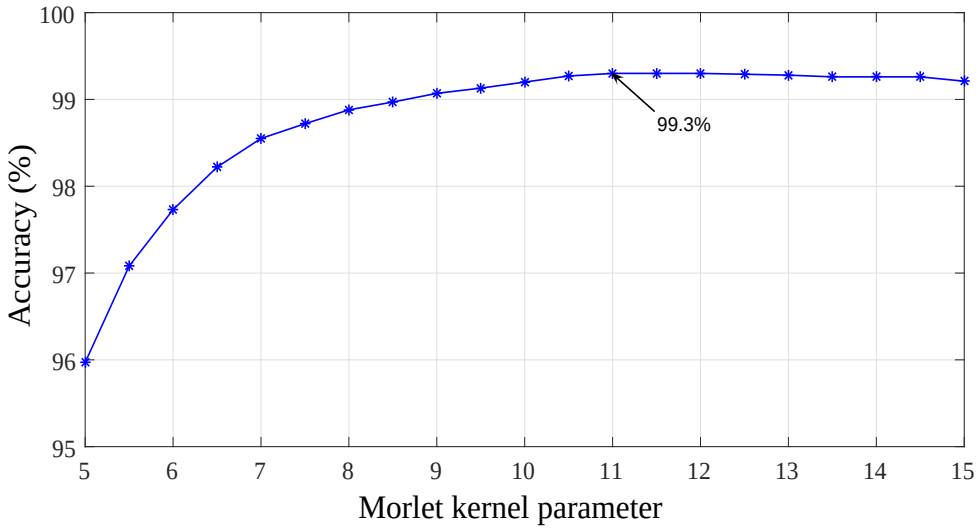


Figure 5.6: A plot of accuracies versus parameter l of Morlet wavelet kernel.

11. The performance of LS-SVM using four different kernels used in our work, is summarized in Table 5.2. LS-SVM yielded the highest classification performance with RBF kernel. The attained values of Ar, Sc, and Ss are 99.31%, 98.12%, and 99.62%, respectively.

Table 5.2: Classification performance (%) of LS-SVM for different kernel functions.

Kernel function	Parameters	Ar (%)	Ss (%)	Sc (%)
Linear		83.32	81.83	89.02
Polynomial	x = 2	96.30	96.01	97.43
	x = 3	96.74	96.44	97.92
RBF	$\gamma = 2.2$	99.31	99.62	98.12
Morlet wavelet	$l = 11$, $z_0 = 0.25$	99.30	99.64	97.92

We have also employed Wilcoxon and Bhattacharya ranking methods for improving the performance of the proposed system (Derryberry et al., 2010; Theodoridis & Koutroumbas, 2003). The plots of the classification Ar (%) for various ranked

features are shown in Figures 5.7 and 5.8 for RBF and Morlet wavelet kernels, respectively. It can be noted that the ranking methods are not able to improve the

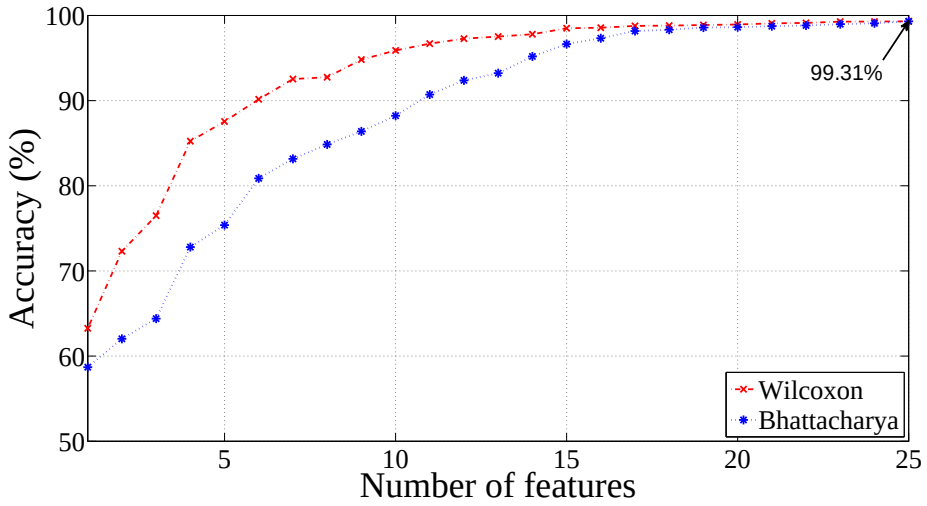


Figure 5.7: A plot of Ar (%) versus number of features using LS-SVM with RBF kernel.

classification performance. The discrimination ability of the features is determined by computing the p-values using KW test (McKight & Najab, 2010). The p-values

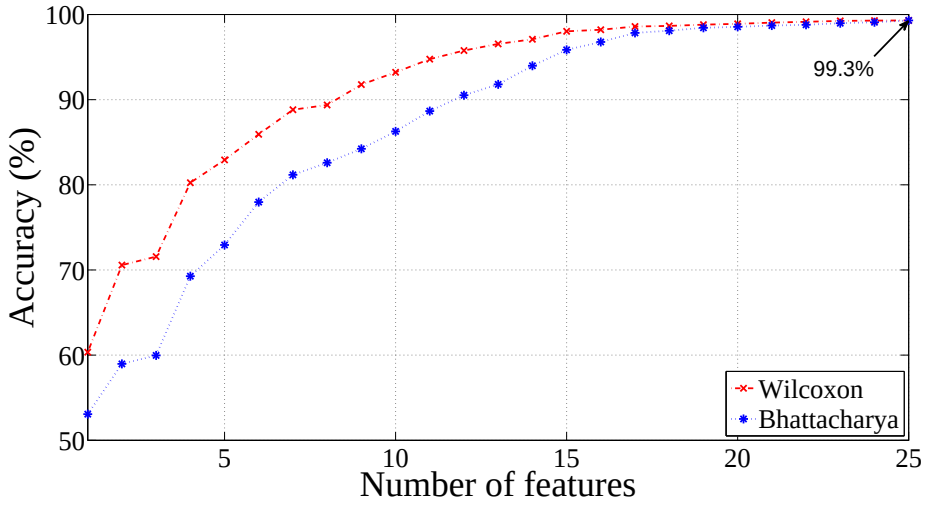


Figure 5.8: The plot of Ar (%) versus number of features using LS-SVM with Morlet wavelet kernel.

are found significantly low ($p < 0.0001$) for all the features (SEnt computed from 25 sub-band signals) which indicate good discrimination ability of all the computed features. The values of μ and STD for features are provided in Table 5.3. In Table 5.3, SEnt refers to the sample entropy and the subscript refers to the corresponding

sub-band signal from which SEnt is computed.

Furthermore, we test the classification performance with the equal number of beats (10,546 beats of each class) from both classes with same parameter values of SEnt and FAWT. Classification Ar for this case using LS-SVM classifier is presented in Table 5.4.

5.4 Discussion

In the present work, the ECG beats are decomposed upto 24^{th} level of decomposition using FAWT for obtaining sub-band signals. Further, SEnt is computed from each of the sub-band signals. We can observe from Table 5.3 that SEnt computed from the lower frequency sub-band signals (SB_{21}, SB_{22}, SB_{24}, and SB_{25}) showed higher values for normal ECG beats in comparison to the MI ECG beats. Therefore, lower frequency sub-band signals show higher complexity for normal ECG beats than MI ECG beats. However, lower values of SEnt are observed for higher frequency sub-band signals extracted from normal ECG beats. Hence, the complexity of higher frequency sub-band signals is lower for normal ECG beats. Finally, our method achieved 99.31% Ar using LS-SVM classifier with RBF kernel.

Summary of the comparison of the present work with the other existing work is provided in Table 5.5. In (Arif et al., 2010), time-domain features are computed from 12-lead ECG signals. The computed features are fed to the BPNN classifier which yielded Ss of 97.5%. In (Al-Kindi et al., 2011), a time-domain method is used for extracting the features from the ECG signals to diagnose the MI patients. The authors have used 12-lead ECG signals of 20 normal and 20 MI subjects. They achieved a Ss of 85% to detect the MI subjects. In (Banerjee & Mitra, 2014), the authors proposed a method based on the spectral differences of cross wavelet transform (XWT) of the ECG signals. Further, they proposed threshold based classifier and achieved 97.6% classification Ar. In (Liu et al., 2015), an algorithm based on the parametrization of ECG signal is developed. In this algorithm, a 20^{th} order polynomial is fitted with the ECG signal. Their method showed 94.4% classification Ar with J48 DT model for the diagnosis of MI. The approach presented in (Sharma et al., 2015a) utilized the evaluation of multiscale energy and eigenspace

Table 5.3: The μ, STD values for normal and MI classes.

Feature	Normal class ($\mu\pm$ STD)	MI class ($\mu\pm$ STD)
$SEnt_{SB1}$	0.0111 ± 0.0248	0.0448 ± 0.0651
$SEnt_{SB2}$	0.0250 ± 0.0385	0.0742 ± 0.0664
$SEnt_{SB3}$	0.0030 ± 0.0039	0.0071 ± 0.0112
$SEnt_{SB4}$	0.0032 ± 0.0026	0.0058 ± 0.0070
$SEnt_{SB5}$	0.0282 ± 0.0298	0.0660 ± 0.0491
$SEnt_{SB6}$	0.0625 ± 0.0431	0.0971 ± 0.0483
$SEnt_{SB7}$	0.0727 ± 0.0390	0.0973 ± 0.0420
$SEnt_{SB8}$	0.0696 ± 0.0388	0.0965 ± 0.0413
$SEnt_{SB9}$	0.0501 ± 0.0324	0.0722 ± 0.0338
$SEnt_{SB10}$	0.0493 ± 0.0246	0.0596 ± 0.0257
$SEnt_{SB11}$	0.0569 ± 0.0251	0.0680 ± 0.0244
$SEnt_{SB12}$	0.0674 ± 0.0305	0.0902 ± 0.0205
$SEnt_{SB13}$	0.0627 ± 0.0354	0.0928 ± 0.0288
$SEnt_{SB14}$	0.0599 ± 0.0340	0.0754 ± 0.0374
$SEnt_{SB15}$	0.0501 ± 0.0305	0.0663 ± 0.0380
$SEnt_{SB16}$	0.0480 ± 0.0221	0.0597 ± 0.0329
$SEnt_{SB17}$	0.0521 ± 0.0162	0.0607 ± 0.0247
$SEnt_{SB18}$	0.0894 ± 0.0151	0.0978 ± 0.0227
$SEnt_{SB19}$	0.1437 ± 0.0129	0.1442 ± 0.0157
$SEnt_{SB20}$	0.1491 ± 0.0056	0.1515 ± 0.0070
$SEnt_{SB21}$	0.1501 ± 0.0066	0.1475 ± 0.0087
$SEnt_{SB22}$	0.1230 ± 0.0100	0.1197 ± 0.0104
$SEnt_{SB23}$	0.0904 ± 0.0030	0.0911 ± 0.0038
$SEnt_{SB24}$	0.0665 ± 0.0010	0.0663 ± 0.0013
$SEnt_{SB25}$	0.0420 ± 0.0087	0.0363 ± 0.0107

Table 5.4: Classification performance (%) of LS-SVM with different kernel functions for balanced dataset.

Kernel function	Parameters	Ac (%)	Ss (%)	Sc (%)
Linear		85.74	84.64	86.83
Polynomial	x = 2	94.06	92.61	95.52
	x = 3	96.88	95.98	97.77
RBF	$\gamma = 2.2$	98.27	99.13	97.40
Morlet wavelet	$l = 11$, $z_0 = 0.25$	98.19	99.20	97.17

(MEES) features. The suggested method used SVM classifier with RBF kernel and achieved 96.15% classification Ar.

In (Acharya et al., 2016), ECG beats are decomposed up to the 4^{th} level of decomposition using DWT. From the DWT coefficients, 12 nonlinear parameters are extracted. The authors achieved 98.8% Ar using K-NN classifier. They also performed statistical tests for determining the significance levels of the studied features. A method to automatically detect the MI using ECG signals is also proposed in (Acharya et al., 2017b). The achieved accuracies were 93.53% and 95.22% using convolutional neural network (CNN) algorithms for the ECG beats with noise and without noise removal, respectively.

We have achieved highest Ar in comparison to the existing methods which are mentioned in Table 5.5. Moreover, the methods suggested in (Arif et al., 2010; Al-Kindi et al., 2011; Banerjee & Mitra, 2014; Liu et al., 2015; Sharma et al., 2015a) used ECG recording of the multiple leads. However, our method uses only lead-2 ECG recordings which makes our method less complex than multiple leads methods. The method suggested in (Acharya et al., 2016) also requires ECG records of one lead (lead-11) only. However, the method in (Acharya et al., 2016) achieved 98.8% classification Ar with 47 features. In comparison to the method in (Acharya et al., 2016), our method has achieved 99.31% Ar with 25 features. Our method showed better results than the method in (Acharya et al., 2016) with less number of features. The study proposed in (Acharya et al., 2017b) also used lead-2 ECG signals and achieved 95.22% Ar with 11-layer deep neural network. This method is more complex

Table 5.5: Summary of automated diagnosis of MI using ECG signals of PTB diagnostic ECG database.

Authors	Year	Analyzing method	Number of leads	Classification method used	10-fold Cross validation	Classification performance (%)
Arif et al.	2010	Time-domain method	12	BPNN	No	Ss = 97.5
Al-Kindi et al.	2011	Time-domain method	12	-	No	Ss= 85
Banerjee et al.	2014	XWT based method	3	Threshold based classifier	No	Ar = 97.6
Liu et al.	2015	ECG polynomial fitting	12	J48 DT	No	Ar = 94.4
Sharma et al.	2015	MEES based method	12	SVM	No	Ar = 96.15
Acharya et al.	2016	DWT, Nonlinear features	One (lead-11)	K-NN	Yes	Ar = 98.8
Acharya et al.	2017	No feature extraction and selection	One (lead-2)	CNN	Yes	Ar = 95.22
Present method		FAWT and SEnt	One (lead-2)	LS-SVM	Yes	Ar = 99.31

than our method and also time-consuming.

5.5 Summary

In this work, normal and MI ECG beats are analyzed using SEnt in FAWT framework. We have achieved the highest classification performance using lead-2 ECG signals than the reported works. We have identified the suitable parameters to compute the SEnt in FAWT domain for the detection of MI subjects accurately. Parameters for the computation of SEnt and decomposition level in FAWT domain are selected on the basis of classification Ar computed using RF classifier. Achieved classification accuracies with RF, J48 DT, BPNN, and LS-SVM classifiers are 97.10%, 93.97%, 92.85% and 99.31%, respectively using the entire dataset. Our method achieved a classification Ar of 98.27% with LS-SVM using balanced data set. Therefore, we can conclude that our methodology has performed well for the detection of MI patients using both balanced and unbalanced (entire) datasets. Our automated system can be used to assist cardiologists to cross-check their diagnosis.

Chapter 6

A computer-based method for the diagnosis of CHF HRV signals using FAWT

6.1 Introduction

Around the world, nearly 26 million people are living with CHF (Ponikowski et al., 2014). It is a pathophysiological condition in which heart is unable to provide sufficient blood supply to fulfill the requirements of the body (National Heart, Lung and Blood Institute, 2015b). It may be the result of structural or functional cardiac disorders. These disorders reduce the blood pumping ability of ventricles (Pazos-López et al., 2011). Dyspnea, edema, and fatigue are the common symptoms of CHF (National Heart, Lung and Blood Institute, 2015b; Pazos-López et al., 2011). It is the last stage of several cardiac diseases namely; heart valve disease, MI, and dilated cardiomyopathy (Jong et al., 2011). In many cases, CHF is not discernible. Moreover, CHF patients are more susceptible to sudden cardiac death (Khaled et al., 2006). Therefore, CHF must be diagnosed at an early stage. In this work, our aim is to develop such a system that can automatically distinguish the normal persons and CHF patients using HRV signals.

Visual detection of the variations present in the HRV signals can be a tedious work, and there is a possibility of the inaccurate classification of diseased and normal signals. Therefore, several studies have been carried out for automated detection

of normal and abnormal HRV signals. These signals have been widely used for diagnosing the heart diseases (Giri et al., 2013; Patidar et al., 2015a; Chua et al., 2008). HRV signals of post-MI patients are studied in (Stein et al., 2000). These signals are also studied to investigate hypertension patients (Mussalo et al., 2001) and patients who are at the risk of sudden cardiac death (Malliani et al., 1994).

In (Nolan et al., 1998), time-domain analysis of HRV signals is performed, and it is concluded that CHF has an association with autonomic dysfunction. In (Hadase et al., 2004), frequency-domain parameters, namely very low frequency (VLF), low frequency (LF), high frequency (HF), total power and the ratio of LF to HF powers, are computed from the HRV signals. The VLF power is found to be an independent risk predictor in CHF patients. Depressed HRV is observed in CHF patients as compared to healthy persons (Musialik-Łydka et al., 2003). In (Asyali, 2003), the STD of normal to normal beat intervals has shown the largest discrimination ability for normal and CHF subjects among nine long-term HRV measures. Significantly lower values of the normalized power of the LF component are observed in CHF patients (Guzzetti et al., 2000). Low risk patients (LRP) and high risk patients (HRP) of CHF are analyzed using HRV signals (Melillo et al., 2013). Frequency-domain parameters computed from HRV signals are found depressed in HRPs, except LF/HF. In (Arbolishvili et al., 2006), the authors have studied HRV in CHF subjects and observed lower values of standard HRV measures, except HF power. The low values of HRV parameters have a correlation with the functional severity of heart failure (Arbolishvili et al., 2006).

In (Maestri et al., 2007), the analysis of the nonlinear properties of HRV gives independent information in the risk stratification of patients with CHF. In (Thakre & Smith, 2006), complex Poincare plots are observed for the CHF patients. In (Shahbazi & Asl, 2015), linear and nonlinear features are used to discriminate the LRP and HRP of CHF disease. The discrimination ability of nonlinear features is found to be better than linear features. Various nonlinear features, such as detrended fluctuation analysis (DFA) (Jong et al., 2011), approximate entropy (Liu et al., 2014), and SEnt (Liu et al., 2014), are found to be effective in the analysis of CHF HRV signals.

In this work, we have analyzed the HRV signals of the CHF and normal subjects

for three different signal lengths at different frequency scales. The FAWT (Bayram, 2013; Zhang et al., 2015) is used to decompose the HRV signals up to the fifth level of decomposition. Decomposed sub-band signals are used to obtain different frequency scales of HRV signals. AFEnt and APEnt are computed from the signals at different frequency scales to measure the complexity of HRV signals at different frequency scales. A total of 10 different frequency scales are obtained in which five frequency scales are obtained by adding the lower frequency sub-band signals (LFSBSs) into the higher frequency sub-band signal (HFSBS), and the other five are obtained by adding the HFSBSs into the LFSBS. Finally, the obtained features are ranked using the Bhattacharyya ranking method (Theodoridis & Koutroumbas, 2003) and fed to the LS-SVM classifier (Suykens & Vandewalle, 1999). The steps followed in our work are dipicted in Figure 6.1. The remaining part of the paper is organized as follows:

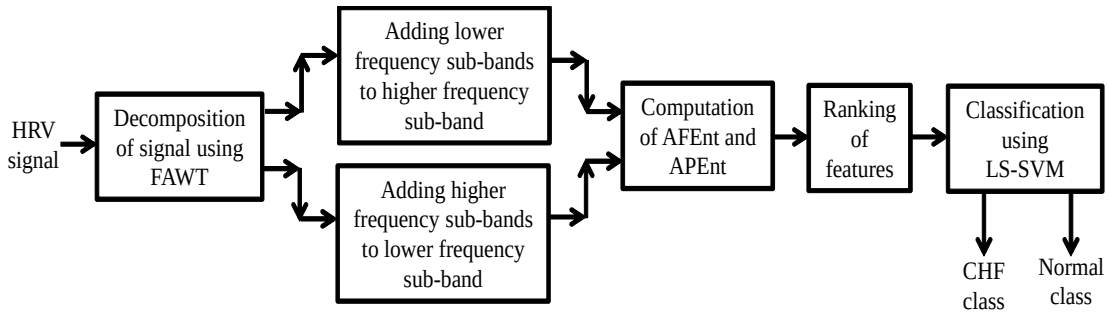


Figure 6.1: The proposed automated diagnosis method for CHF using HRV signals.

Section 6.2 describes the HRV dataset used in the present work, the segmentation of the HRV dataset, PEnt, FAWT-based accumulated entropies, the ranking method, and classification. Section 6.3 and section 6.4 present the obtained results and discussion part, respectively. Finally, the work is summarised in section 6.5.

6.2 Methodology

6.2.1 HRV dataset

In this work, we have used online available databases for CHF and normal subjects. For CHF patients, the Beth Israel Deaconess Medical Center (BIDMC) CHF dataset

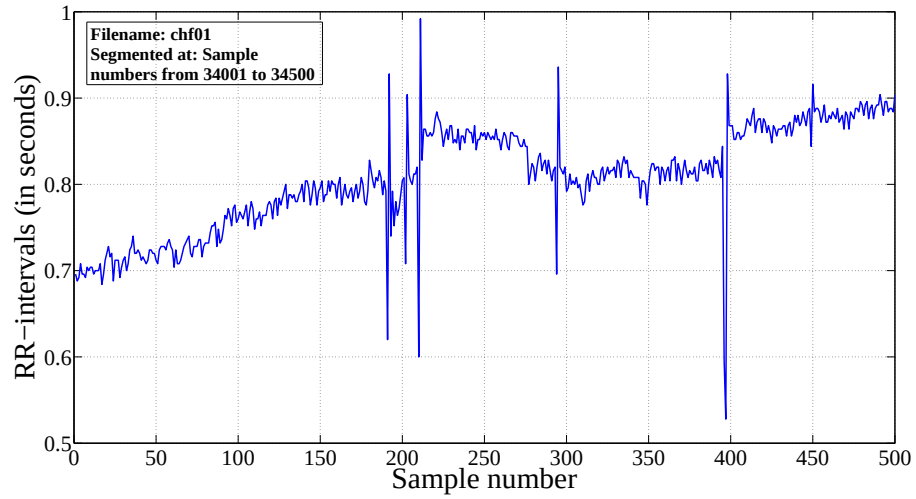
from physiobank (Baim et al., 1986; Goldberger et al., 2000) is used. The HRV signals of 15 CHF patients (11 males and 4 females) are used in this study. The age of the patients varies from 22 to 71 years for male subjects and from 54 to 63 years for female subjects. For normal subjects, two different datasets, Massachusetts Institute of Technology-Beth Israel Hospital (MIT-BIH) normal sinus rhythm (NSR) (Goldberger et al., 2000) and Fantasia (Iyengar et al., 1996; Goldberger et al., 2000) from physiobank, are used. From the MIT-BIH NSR dataset, we have obtained HRV signals of 18 subjects (5 males and 13 females). The age of male subjects ranges from 26 to 45 years, and for female subjects, age varies in between 20 and 50 years. From the Fantasia dataset, the HRV records of 40 subjects (20 young and 20 old) are obtained. The age variation of young subjects is 21 to 34 years, and for old subjects, it varies from 68 to 85 years. Information on RR-intervals is provided in text format in the physiobank. We have obtained RR-intervals directly from the physiobank.

6.2.2 Segmentation of HRV signals

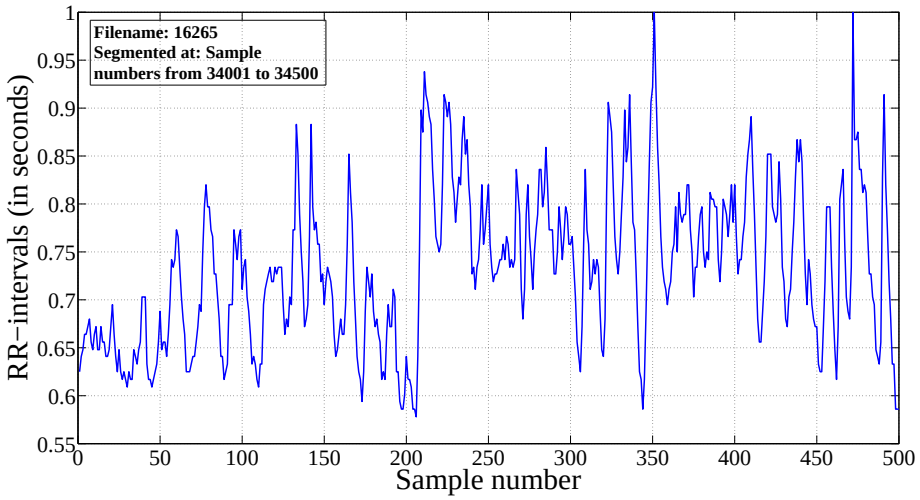
We have segmented the HRV signals into three different signal lengths: 500, 1000, and 2000 samples. Table 6.1 summarizes the number of HRV segments used for each signal length (500, 1000, and 2000 samples) of the two classes. The plots of normal (MIT-BIH NSR dataset) and CHF (BIDMC dataset) HRV signals for 500 samples length are shown in Figure 6.2.

Table 6.1: Total number of extracted segments for various signal lengths.

Database	Total segments		
	SL = 500 Samples	SL = 1000 Samples	SL = 2000 Samples
CHF/BIDMC	3212	1606	803
Normal/MIT-BIH	3420	1710	855
Normal/Fantasia	500	250	125



(a)



(b)

Figure 6.2: The plots of HRV signals of a 500-sample length: (a) CHF subject and (b) normal subject.

6.2.3 Features studied in this work

Nonlinear features are found suitable for analysis of the HRV signals (Acharya et al., 2004). Therefore, nonlinear parameters, namely AFEnt and APEnt, are proposed as features to classify the CHF and normal HRV signals in this work. A brief explanation of these parameters is provided in the following sections.

6.2.4 Permutation entropy

It measures the complexity of HRV signals, which is determined by comparing the permutation patterns present in the samples of the signal (Bandt & Pompe, 2002). Its computation depends on sequence length e and delay time τ. A total number of $e!$ permutations are possible for the selected value of e. If the probability of occurrence of the k-th permutation pattern is denoted by B_k, then the PEnt can be computed as follows (Bandt & Pompe, 2002; Zanin et al., 2012):

$$\text{PEnt} = -\sum_{k\,=\,1}^{e!} B_k \log(B_k) \tag{6.1}$$

where, B_k can be computed as follows (Bandt & Pompe, 2002):

$$B_k = \frac{\text{total occurance of } k^{\text{th}} \text{ pattern}}{P - e + 1} \tag{6.2}$$

where, P is the length of the signal, and $P - e + 1$ denotes the total number of vectors of e length.

6.2.5 Fuzzy entropy

The computation steps of FEnt are provided in chapter 3. In this work, the parameters $u = 2$, $v = 0.2$ (Chen et al., 2007) are selected to compute the FEnt.

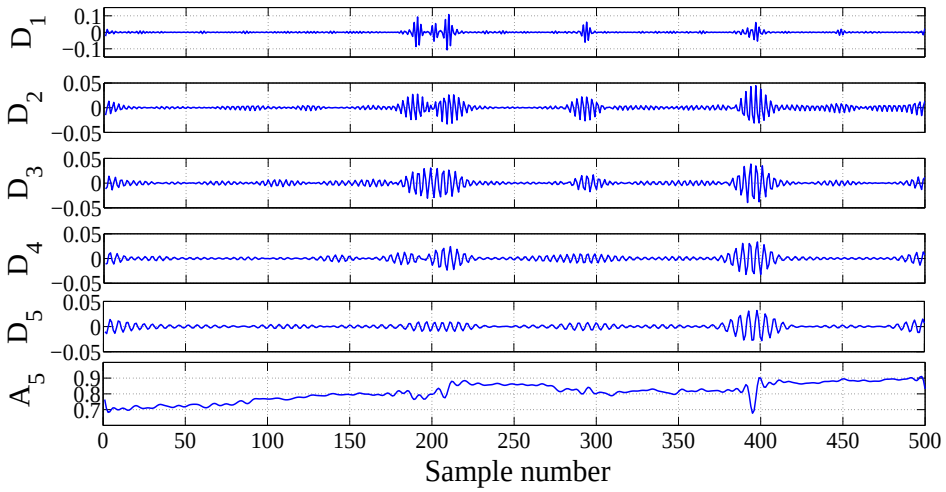
6.2.6 FAWT-based accumulated entropies

The values of the parameters used in the present work are $b = 5$, $c = 6$, $f = 1$, $g = 2$ and $\beta = (0.8f)/g$ (Zhang et al., 2015). The 5^{th} level FAWT-based decomposition of the signals in Figure 6.2 is shown in Figure 6.3. The sub-band signals are ordered from the highest frequency to the lowest one in Figure 6.3. A_5 represents the approximation signal with lower frequency components than D_1 to D_5.

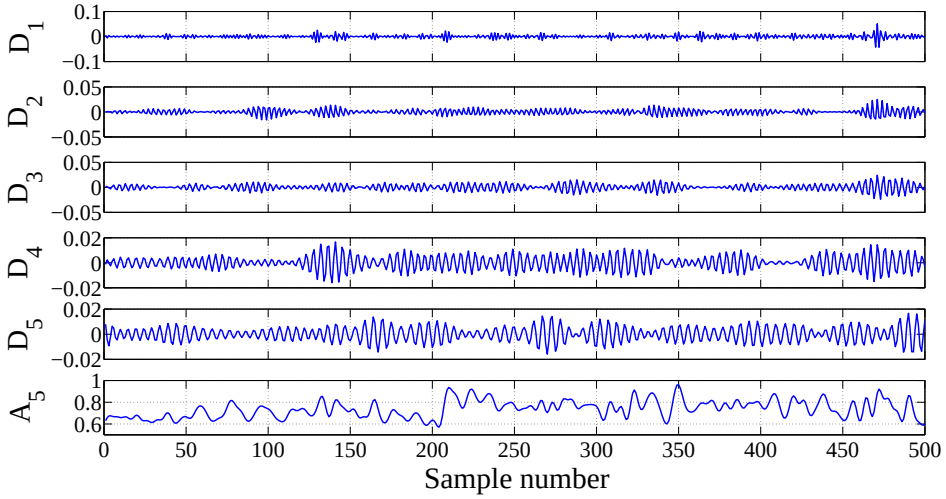
APEnt and AFEnt can be computed over cumulative sums of sub-band signals obtained by the decomposition of HRV signals using the FAWT method.

The APEnt can be defined as follows:

$$\text{APEnt}(i, e, \tau) = \text{PEnt}[S^i_{\text{sub-bands}}, e, \tau] \tag{6.3}$$



(a)



(b)

Figure 6.3: Decomposition of HRV signals using the FAWT method: (a) CHF subject and (b) normal subject.

The AFEnt can be defined as follows:

$$\text{AFEnt}(i, e, u, v, P) = \text{FEnt}[S^i_{\text{sub-bands}}, e, u, v, P] \tag{6.4}$$

where $S^i_{\text{sub-bands}}$ can be expressed as:

$$S^i_{\text{sub-bands}} = \text{SL}_i, \quad i = 1, 2, .., J \tag{6.5}$$

when sub-band signals are combined from the LF to the HF. For the combination

of sub-band signals from HF to LF, $S^i_{\text{sub-bands}}$ can be given as:

$$S^i_{\text{sub-bands}} = \text{SH}_i, \quad i = 1, 2, .., J \tag{6.6}$$

We have selected $\tau = 2$ in our work to compute the APEnt. In 6.5 and 6.6, J represents the decomposition level, which is selected as 5 in the present work. A total of 5 detail sub-band signals (D_1 to D_5) and the approximation signal (A_5) are used to obtain different frequency scales. The frequency-scaled sub-band signals are shown in Figures 6.4 and 6.5 for CHF and normal subjects corresponding to a 500-sample length of the HRV signal.

The combination of sub-band signals is provided in Table 6.2. In (Pachori et al., 2009), IMFs obtained from the EMD method are utilized in combination of LF to HF and HF to LF for analyzing the postural time series.

Table 6.2: Signals at different frequency scales and the combination of used sub-band signals.

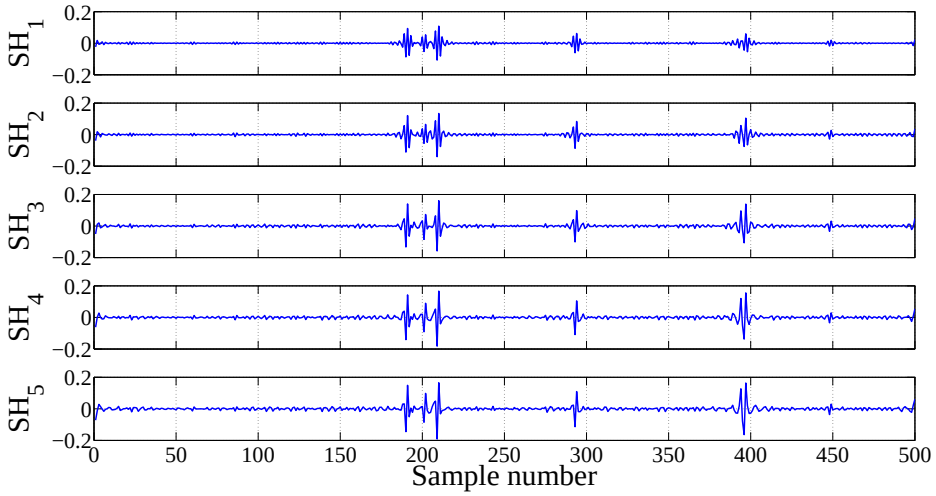
Signals at different frequency scales	Accumulation of sub-band signals	Signals at different frequency scales	Accumulation of sub-band signals
SL_1	A_5	SH_1	D_1
SL_2	$A_5 + D_5$	SH_2	$D_1 + D_2$
SL_3	$A_5 + D_5 + D_4$	SH_3	$D_1 + D_2 + D_3$
SL_4	$A_5 + D_5 + D_4 + D_3$	SH_4	$D_1 + D_2 + D_3 + D_4$
SL_5	$A_5 + D_5 + D_4 + D_3 + D_2$	SH_5	$D_1 + D_2 + D_3 + D_4 + D_5$

6.2.7 Ranking and classification

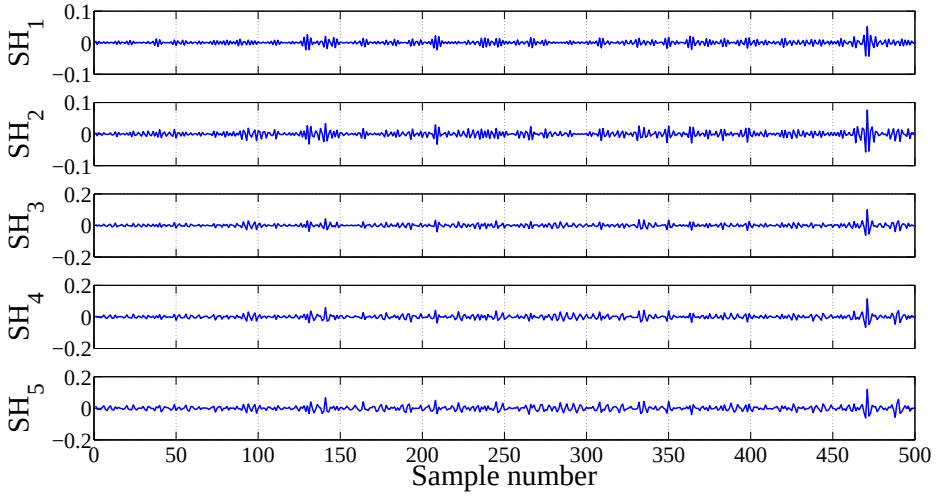
In this work, we have used Bhattacharyya ranking method and LS-SVM classifier which are already explained in chapter 3. In the present work, the classification performance measures, namely Sc, Ss and Ar (Azar & El-Said, 2014), are computed to determine the effectiveness of the classifier.

6.3 Results

In the present work, we have computed APEnt and AFEnt at different frequency scales. These different frequency scales of HRV signals are obtained by adding the sub-band signal from the HF component to the LF component and the LF component to the HF component. These sub-band signals are obtained from FAWT-based decomposition. We have analyzed HRV signals for three signal lengths, 500, 1000, and 2000 samples. The obtained results for these three different lengths of HRV signals are as follows:



(a)

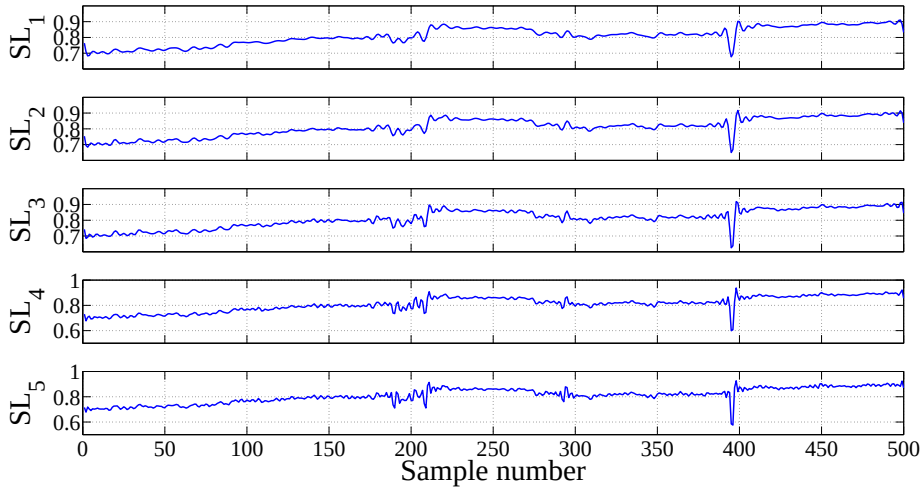


(b)

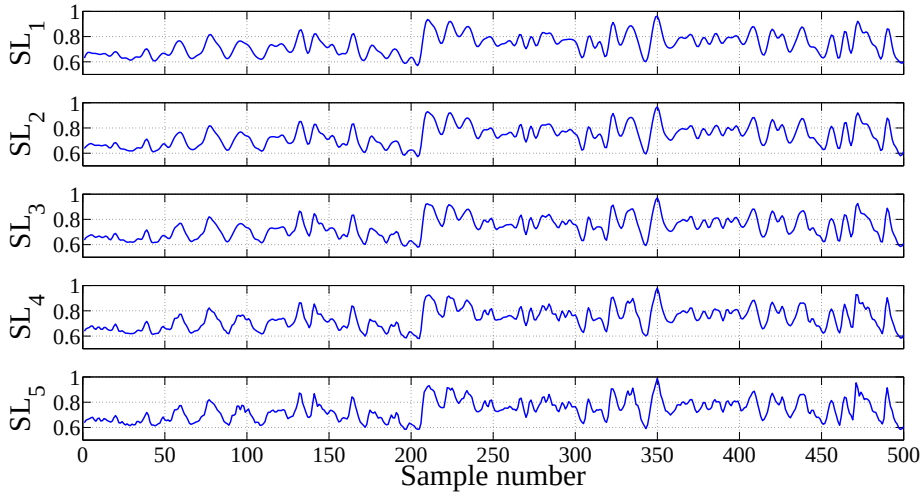
Figure 6.4: The plots of HRV signals at various frequency scales (the first subplot represents the highest frequency sub-band signal, and remaining subplots are the addition of the LFSBSs to the HFSBS): (a) CHF subject and (b) normal subject.

6.3.1 Results with a signal length of 500 samples

We have used one dataset for CHF patients (BIDMC) and two datasets of normal subjects (Fantasia dataset and MIT-BIH NSR dataset). The HRV signals of two classes (normal and CHF) are subjected to FAWT-based decomposition. The signals are decomposed up to the fifth level. We have made four different combinations of datasets, namely; unbalanced dataset 1 (UD1), unbalanced dataset 2 (UD2), Balanced dataset 1 (BD1) and balanced dataset 2 (BD2).



(a)



(b)

Figure 6.5: The plots of HRV signals at various frequency scales (the first subplot represents the lowest frequency sub-band signal, and remaining subplots are the addition of the HFSBSs to the LFSBS): (a) CHF subject and (b) normal subject.

Results for UD1

This dataset has the recordings of MIT-BIH NSR for normal subjects and the BIDMC dataset for CHF patients. This is the largest dataset among the four combinations of the dataset used in the present work. This selected dataset has 3420 normal segments and 3212 CHF segments of HRV signals. To find the suitable value of e for AFEnt and APEnt, we have computed the classification Ar at various value of e for AFEnt and APEnt using the LS-SVM classifier. Classification Ar for various values of e is shown in Table 6.3 for AFEnt and APEnt. From Table 6.3, we can observe that AFEnt and APEnt show better performance for the value of $e = 3$ and 4, respectively. For AFEnt and APEnt, the highest classification Ar of 96.29% and 91.39% using Morlet wavelet kernel is obtained respectively. For further analysis, we have used $e = 3$ for AFEnt and $e = 4$ for APEnt.

Table 6.3: Classification performance for UD1 (Length of the signal = 500 samples) using AFEnt and APEnt separately with LS-SVM for different values of e and different kernel functions and parameters (entropies computed from all of the frequency scales are used for classification).

Features	Sequence length	Linear kernel Ar (%)	RBF kernel Ar (%)	Morlet wavelet kernel Ar (%)	Polynomial kernel Ar (%)
AFEnt	3	90.66	95.67	96.29	95.02
	4	90.86	94.54	95.31	94.52
	5	89.77	94.90	95.98	94.43
	6	92.12	95.43	95.92	95.38
	7	91.36	94.84	95.55	94.79
APEnt	3	82.62	89.32	89.17	88.16
	4	84.80	91.26	91.39	90.89
	5	84.84	90.65	90.65	90.75
	6	83.05	89.24	89.2	89.02
	7	78.69	85.93	86.17	85.20

The μ and STD values of AFEnt for $e = 3$ and APEnt for $e = 4$ are given in

Table 6.4. We have also computed p-values using the KW test (McKight & Najab, 2010) to check the statistical significance of the features in order to discriminate normal and CHF HRV signals. We can observe in Table 6.4 that p-values are significantly low (p-value < 0.05) for all of the features. Further, we have applied the Bhattacharyya ranking method to rank the features. These ranked features are presented in Table 6.5. We can observe in Table 6.5 that AFEnt features achieve a higher ranking as compared to APEnt features. In Table 6.5, the indices (SH_1 to SH_5 and SL_1 to SL_5) represent the frequency-scaled signals from which AFEnt and APEnt are computed. These ranked features are fed to the LS-SVM classifier by appending the next lower ranking features to the highest ranking feature up to 20 features. The accuracies of the ranked features with LS-SVM for different kernels (linear, RBF, polynomial and Morlet wavelet) are shown in Figure 6.6. It can be observed from Figure 6.6 that the LS-SVM classifier achieved the highest Ar for Morlet wavelet kernel with the first 18 features, for the polynomial kernel with the first 17 features, and for the RBF kernel with the first 14 features. The classification Ar with the linear kernel is highest for 20 features. The classifier is trained and tested using the TFCV method (Kohavi, 1995). The best classification performance of LS-SVM for different kernels is provided in Table 6.6. The maximum value of Ar can be seen from Table 6.6, which is 98.21% for Morlet wavelet kernel.

Results for UD2

This dataset includes the HRV signals from Fantasia (normal subject) and the BIDMC dataset (CHF patients). This dataset has 500 segments of normal and 3212 segments of CHF HRV signals. The p-values, μ and STD values of AFEnt and APEnt are given in Table 6.7. We can observe from Table 6.7 that p-values are significantly low (p-value <0.05) for all of the features, except AFEnt for SH_1 and APEnt for SL_5 and SL_4.

The classification Ar of the LS-SVM classifier for the ranked features with the different kernels used in this work is shown in Figure 6.7. For RBF, polynomial, and Morlet wavelet kernels, the highest Ar of classification is 97.33% with 19 features, 96.20% with 17 features, and 97.31% with 20 features, respectively. For the linear kernel, maximum classification Ar is 90.38% from 20 features. The remaining

Table 6.4: The μ, STD and p-values of AFEnt and APEnt for UD1 (Length of the signal = 500 samples) at various frequency scales provided by FAWT-based decomposition applied to normal and CHF HRV signals.

AFEnt				APEnt			
Subband signals	CHF class $\mu\pm$ STD	Normal class $\mu\pm$ STD	p-values	Subband signals	CHF class $\mu\pm$ STD	Normal class $\mu\pm$ STD	p-values
SH_1	0.0132± 0.0222	0.0042± 0.0085	7.92×10^{-85}	SH_1	2.6932± 0.0717	2.6810± 0.0581	8.09×10^{-16}
SH_2	0.0313± 0.0420	0.0268± 0.0185	5.63×10^{-45}	SH_2	3.0413± 0.0870	3.0708± 0.0489	1.81×10^{-51}
SH_3	0.0447± 0.0554	0.0502± 0.0317	2.32×10^{-161}	SH_3	3.0829± 0.0955	3.0764± 0.1191	2.13×10^{-9}
SH_4	0.0555± 0.0654	0.0694± 0.0465	7.37×10^{-200}	SH_4	3.0418± 0.1221	2.9902± 0.1556	2.75×10^{-42}
SH_5	0.0610± 0.0716	0.0841± 0.0565	5.39×10^{-273}	SH_5	3.0393± 0.1289	2.9718± 0.1591	2.42×10^{-81}
SL_5	0.0361± 0.0467	0.0702± 0.0488	0	SL_5	3.0657± 0.0950	3.0192± 0.1019	5.52×10^{-120}
SL_4	0.0342± 0.0377	0.0689± 0.0476	0	SL_4	3.0442± 0.0916	2.9981± 0.1070	4.32×10^{-85}
SL_3	0.0394± 0.0397	0.0683± 0.0469	9.49×10^{-276}	SL_3	3.0739± 0.079	2.9524± 0.1242	0
SL_2	0.0361± 0.0369	0.0628± 0.0425	2.59×10^{-273}	SL_2	3.0360± 0.1210	2.8492± 0.1513	0
SL_1	0.0291± 0.0296	0.0568± 0.0352	0	SL_1	2.8731± 0.1484	2.6972± 0.1449	0

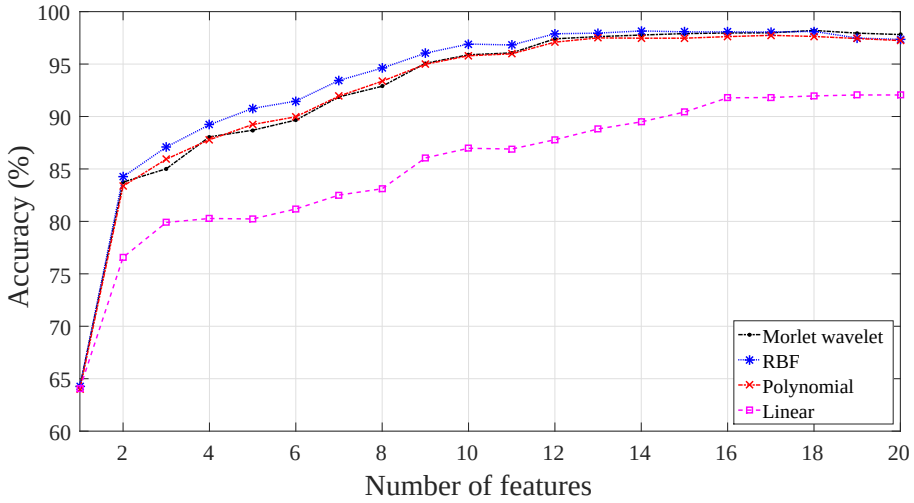


Figure 6.6: The plot of the classification accuracies versus the number of features for UD1 (Length of the signal = 500 samples).

classification parameters (Ss and Sc) are given in Table 6.6.

Results for BD1

In this dataset, we have used 500 HRV segments from the MIT-BIH NSR dataset of normal subjects and 500 HRV segments from the BIDMC dataset of CHF patients. Classification accuracies corresponding to the used number of features are presented in Figure 6.8. The highest classification accuracies for linear, Morlet wavelet, RBF, and polynomial kernels are 96.60% using 18 features, 99.4% using 15 features, 99.5% using 15 features, and 98.4% using 11 features, respectively. Other computed classification parameters, namely Ss and Sc, are provided in Table 6.6.

Table 6.5: Ranked features using the Bhattacharyya ranking method for UD 1 (Signal length = 500 samples).

Feature rank	1	2	3	4	5	6	7
Feature name	$AFEnt_{SH_1}$	$AFEnt_{SH_2}$	$APEnt_{SL_2}$	$APEnt_{SL_3}$	$APEnt_{SL_1}$	$APEnt_{SH_2}$	$AFEnt_{SH_3}$
Feature rank	8	9	10	11	12	13	14
Feature name	$AFEnt_{SH_4}$	$AFEnt_{SL_4}$	$APEnt_{SH_5}$	$APEnt_{SH_4}$	$AFEnt_{SL_1}$	$AFEnt_{SH_5}$	$APEnt_{SL_4}$
Feature rank	15	16	17	18	19	20	
Feature name	$AFEnt_{SL_3}$	$AFEnt_{SL_2}$	$AFEnt_{SL_5}$	$APEnt_{SH_3}$	$APEnt_{SH_1}$	$APEnt_{SL_5}$	

Table 6.6: Classification performance of the LS-SVM classifier for various combinations of datasets (Signal length = 500 samples) with different kernel functions.

Combinations of datasets	Kernels	Kernel parameters	Number of features	Ss (%)	Sc (%)	Ar (%)
UD1	Morlet wavelet	$l = 6.2,\ z_0 = 0.25$	18	98.07	98.33	98.21
	Polynomial	Order = 3	17	97.76	97.72	97.74
	RBF	$\gamma = 1$	14	97.98	98.33	98.16
	Linear		20	90.26	93.74	92.05
UD2	Morlet wavelet	$l = 7.9,\ z_0 = 0.25$	20	97.85	93.80	97.31
	Polynomial	Order = 3	17	96.51	94.20	96.20
	RBF	$\gamma = 1.2$	19	97.95	93.40	97.33
	Linear		20	91.38	84.00	90.38
BD1	Morlet wavelet	$l = 6.7,\ z_0 = 0.25$	15	99.60	99.20	99.40
	Polynomial	Order = 3	11	99.00	97.80	98.40
	RBF	$\gamma = 1$	15	99.80	99.20	99.50
	Linear		18	96.00	97.20	96.60
BD2	Morlet wavelet	$l = 7.7,\ z_0 = 0.25$	16	97.60	97.80	97.70
	Polynomial	Order = 3	12	97.60	96.40	97.00
	RBF	$\gamma = 1.2$	16	97.60	98.00	97.80
	Linear		20	93.60	90.20	91.90

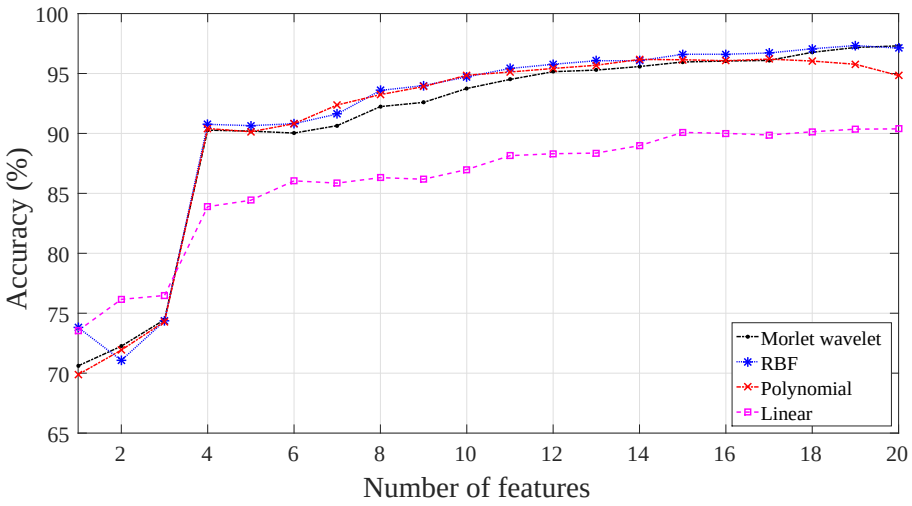


Figure 6.7: A plot of classification accuracies versus the number of features for UD2 (Length of the signal = 500 samples).

Results for BD2

This dataset has 500 normal HRV segments from the Fantasia dataset and 500 CHF HRV segments from the BIDMC dataset. Classification accuracies obtained for this dataset are shown in Figure 6.9. The highest values of Sc, Ss, and Ar obtained using LS-SVM for different kernels are summarized in Table 6.6.

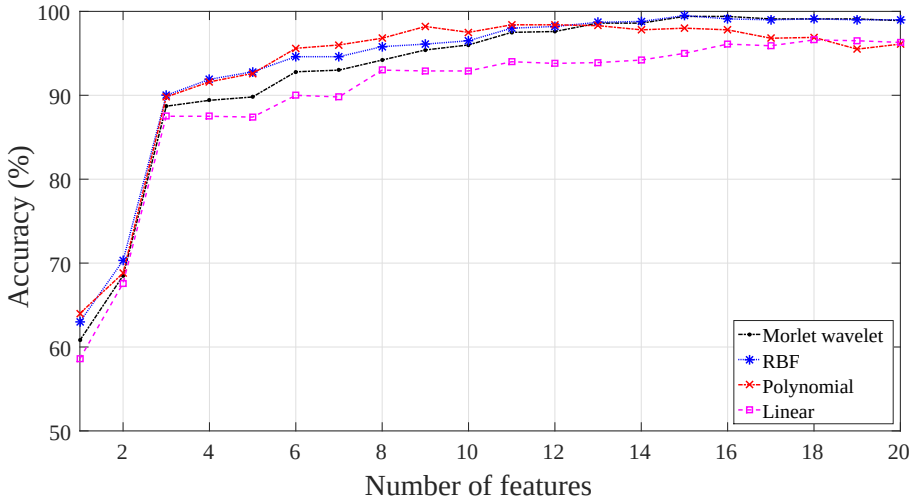


Figure 6.8: A plot of classification accuracies versus number of features for BD1 (Length of the signal = 500 samples).

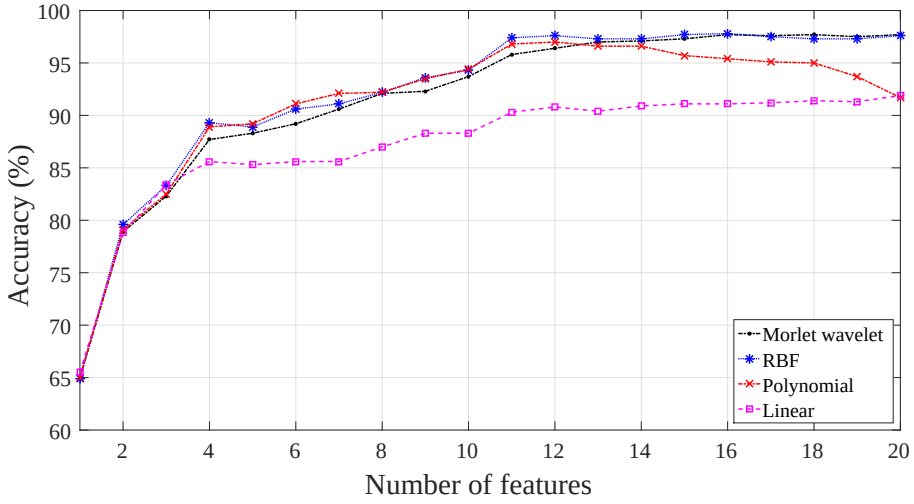


Figure 6.9: A plot of accuracies versus number of features for BD2 (Signal length = 500 samples).

Table 6.7: The μ, STD, and p-values of AFEnt and APEnt for UD2 (Length of the signal = 500 samples) at various frequency scales provided by FAWT-based decomposition applied to HRV signals of CHF patients and normal subjects.

AFEnt				APEnt			
Subband signals	CHF class $\mu\pm$ STD	Normal class $\mu\pm$ STD	p-values	Subband signals	CHF class $\mu\pm$ STD	Normal class $\mu\pm$ STD	p-values
SH_1	0.0132± 0.0222	0.0117 ± 0.0192	0.6866	SH_1	2.6932± 0.0717	2.7088 ± 0.0721	4.15×10^{-7}
SH_2	0.0313± 0.0420	0.0377 ± 0.0323	2.50×10^{-30}	SH_2	3.0413± 0.0870	3.0672 ± 0.0687	3.17×10^{-18}
SH_3	0.0447± 0.0554	0.0622 ± 0.0474	5.93×10^{-52}	SH_3	3.0829± 0.0955	3.0387 ± 0.0947	2.41×10^{-34}
SH_4	0.0555± 0.0654	0.0807 ± 0.0577	3.37×10^{-58}	SH_4	3.0418± 0.1221	2.9600 ± 0.1397	6.25×10^{-45}
SH_5	0.0610± 0.0716	0.0931 ± 0.0640	3.06×10^{-65}	SH_5	3.0393± 0.1289	2.9510 ± 0.1405	8.31×10^{-51}
SL_5	0.0361± 0.0467	0.0824 ± 0.0533	3.65×10^{-125}	SL_5	3.0657± 0.0950	3.0563 ± 0.1104	0.296
SL_4	0.0342± 0.0377	0.0796 ± 0.0509	5.82×10^{-116}	SL_4	3.0442± 0.0916	3.0304± 0.1205	0.255
SL_3	0.0394± 0.0397	0.0792 ± 0.0504	4.04×10^{-88}	SL_3	3.0739± 0.079	2.9953 ± 0.1397	6.43×10^{-42}
SL_2	0.0361± 0.0369	0.0740 ± 0.0470	9.86×10^{-91}	SL_2	3.0360± 0.1210	2.8926 ± 0.1602	5.69×10^{-99}
SL_1	0.0291± 0.0296	0.0659 ± 0.0404	1.41×10^{-106}	SL_1	2.8731± 0.1484	2.7450 ± 0.1437	1.21×10^{-74}

6.3.2 Results with a signal length of 1000 samples

For a signal length of 1000 samples, again, we have made four different combinations of the datasets as described in section 6.3.1. We have used the same parameters for AFEnt and APEnt and the same ranking method, classifier, and kernel functions as in section 6.3.1. The obtained highest classification performances for different combinations of datasets for a signal length of 1000 samples are provided in Table 6.8.

Table 6.8: Classification performance of the LS-SVM classifier for various combinations of datasets (Signal length = 1000 samples) with different kernel functions.

Combinations of datasets	Kernels	Kernel parameters	Number of features	Ss (%)	Sc (%)	Ar (%)
UD1	Morlet wavelet	$l = 7.7$, $z_0 = 0.25$	20	97.95	98.07	98.01
	Polynomial	Order = 3	17	96.70	97.54	97.13
	RBF	$\gamma = 1.8$	19	98.01	97.95	97.98
	Linear		20	91.22	94.44	92.88
UD2	Morlet wavelet	$l = 6.4$, $z_0 = 0.25$	20	98.75	92.80	97.95
	Polynomial	Order = 3	11	96.76	95.60	96.60
	RBF	$\gamma = 1.4$	20	98.57	94.00	97.95
	Linear		18	92.65	84.80	91.60
BD1	Morlet wavelet	$l = 6$, $z_0 = 0.25$	16	100.00	99.60	99.80
	Polynomial	Order = 3	10	99.20	98.40	98.80
	RBF	$\gamma = 1$	16	100.00	99.60	99.80
	Linear		18	97.20	99.60	98.40
BD2	Morlet wavelet	$l = 7.9$, $z_0 = 0.25$	20	98.40	98.40	98.40
	Polynomial	Order = 3	10	98.80	96.80	97.80
	RBF	$\gamma = 1.2$	20	97.20	98.40	97.80
	Linear		18	94.00	92.80	93.40

6.3.3 Results with a 2000-sample signal length

We have also formed four different combinations of datasets for a signal length of 2000 samples as explained in section 6.3.1. In this case, the same parameters are

selected for AFEnt and APEnt. For this case, we have also applied the same ranking method, classifier and kernel functions as described in section 6.3.1. For a signal length of 2000 samples, the highest obtained values of Ss, Sc, and Ar are provided in Table 6.9.

Table 6.9: Classification performance of the LS-SVM classifier for various combinations of datasets (Signal length = 2000 samples) with different kernel functions.

Combinations of datasets	Kernels	Kernel parameters	Number of features	Ss (%)	Sc (%)	Ar (%)
UD1	Morlet wavelet	$l = 7.1$, $z_0 = 0.25$ Order = 3 $\gamma = 1.4$	15	97.51	97.78	97.65
	Polynomial		15	95.39	96.84	96.14
	RBF		16	97.76	97.67	97.71
	Linear		20	92.41	95.32	93.91
UD2	Morlet wavelet	$l = 6$, $z_0 = 0.25$ Order = 3 $\gamma = 1.3$	19	98.38	91.92	97.52
	Polynomial		11	96.77	92.82	96.23
	RBF		19	98.38	91.92	97.52
	Linear		19	92.54	85.64	91.59
BD1	Morlet wavelet	$l = 7.6$, $z_0 = 0.25$ Order = 3 $\gamma = 1.8$	16	98.40	100.00	99.20
	Polynomial		9	99.17	98.33	98.80
	RBF		16	98.40	100.00	99.20
	Linear		19	98.46	99.17	98.78
BD2	Morlet wavelet	$l = 8$, $z_0 = 0.25$ Order = 3 $\gamma = 1.6$	20	98.40	99.23	98.82
	Polynomial		9	96.79	95.96	96.41
	RBF		20	98.40	99.23	98.82
	Linear		20	96.73	95.06	95.95

6.4 Discussion

In this work, we have analyzed the complexity of CHF and normal HRV signals at different frequency scales obtained using FAWT. The FAWT is applied to the CHF and normal HRV signals to decompose them into sub-band signals. We have combined these sub-band signals so that we can observe the complexity of HRV signals at different frequency scales. For complexity measurement, AFEnt and APEnt are

computed from different frequency-scaled signals. We observed that AFEnt is more suitable to analyze the complexity of CHF and normal HRV signals, as it showed higher Ar as compared to APEnt, and it can be seen in Table 6.3. Most of the frequency scaled signals (except SH_1 and SH_2 in Table 6.4) showed lower mean values of AFEnt for CHF HRV signals. These results show the lower complexity of CHF HRV signals as compared to the normal HRV signals for most of the frequency scales. The reduced complexity of CHF HRV signals may be the reflection of the reduced parasympathetic modulation of HR (Acharya et al., 2017c).

The comparison of our methodology with the other existing works is summarized in Table 6.10. In (Khaled et al., 2006), the normal and CHF HRV signals are analyzed using the Poincare plot and time-domain features. The positive predictive Ar of 98.19% is achieved for time-domain features with the BPNN classifier. Time and frequency-domain based features are computed to separate normal and CHF classes in (Pecchia et al., 2011). The classification and regression tree (CART) method provided 96.4% classification Ar. The DFA-based features with SVM yielded 96% classification Ar to discriminate normal and CHF HRV signals in (Jong et al., 2011). In (Yu & Lee, 2012), the authors have studied time-domain features, frequency-domain features and bispectrum features to analyze HRV signals of CHF and normal subjects. They incorporated GA and SVM classifier in their method and obtained 98.79% classification Ar. In (Narin et al., 2014), the standard HRV measures, nonlinear parameters, and wavelet-based measures are used to distinguish CHF and normal HRV signals. They achieved 91.56% Ar using 27 selected features with the SVM classifier.

Two balanced and two unbalanced datasets of HRV signals (CHF and normal subjects) are studied in (Acharya et al., 2017c). First, the signals are decomposed using the EMD method; then, 13 nonlinear parameters are computed from these decomposed IMFs. They achieved 97.64% Ar for UD1 with 22 features, 95.79% Ar for UD2 with 35 features, 96.7% Ar for BD1 with 12 features, and 94% Ar for BD2 with 11 features.

In this work, we have analyzed the HRV signals of CHF and normal subjects with 500, 1000, and 2000 sample lengths, while the authors in (Acharya et al., 2017c) have used a 2000 sample length in their study. We have achieved better classifica-

Table 6.10: Summary of automated diagnosis of CHF based on HRV signals.

Authors and year	Studied dataset Normal subject	CHF patient	Applied methods	Number of subjects/HRV signals	Classifier used	Total number of features	Results
Khaled et al. (2006)	MIT-BIH NSR and NSR RR interval databases	BIDMC CHF and CHF RR interval databases	Time-domain parameters and Poincare plots	Total 600 short-term HRV signals	BPNN	11	Positive predictive Ar = 98.19%
Pecchia et al. (2011)	NSR RR interval database	CHF RR interval database	Time and frequency-domain based parameters	54 Normal subjects and 29 CHF patients	CART	9	Ar = 96.4%
Jong et al. (2011)	NSR RR interval database	CHF RR interval database	DFA-based parameters	54 Normal subjects and 29 CHF patients	SVM	DFA-based features	Ar = 96%
Yu et al. (2012)	NSR RR interval database	CHF RR interval database	Time, frequency-domain parameters and bispectrum parameters	54 Normal subjects and 29 CHF patients	SVM	16	Ar = 98.79%
Narin et al. (2014)	NSR RR interval database	CHF RR interval database	Standard HRV measures, nonlinear parameters, and wavelet-based measures	54 Normal subjects and 29 CHF patients	SVM	27	Ar = 91.56%
Acharya et al. (2016)	MIT-BIH NSR and Fantasia databases	BIDMC CHF database	EMD, 13 nonlinear parameters HRV signal length = 2000 samples	58 Normal subjects and 15 CHF patients	SVM	22 35 12 11	UD 1, Ar = 97.64% UD 2, Ar = 95.79% BD 1, Ar = 96.7% BD 2, Ar= 94%
In the present work	MIT-BIH NSR and Fantasia databases	BIDMC CHF database	FAWT, AFEnt, and APEnt HRV signal length = 500 samples	58 Normal subjects and 15 CHF patients	LS-SVM	18 19 11 12	UD1, Ar = 98.21% UD2, Ar = 97.33% BD1, Ar = 98.40% BD2, Ar = 97%

tion Ar with fewer features as compared to the study presented in (Acharya et al., 2017c). The computational complexity of our method is lesser than the method proposed in (Acharya et al., 2017c), as we have computed only two parameters (AFEnt and APEnt), while (Acharya et al., 2017c) have computed 13 nonlinear parameters. Moreover, we have also achieved good classification accuracies for a 1000-sample signal length and a 500-sample signal length. Therefore, our methodology is suitable for short-term HRV signals, which can provide great help to the doctors and clinicians during their diagnosis of CHF patients. The proposed methodology requires HRV signals of a smaller length as compared to the method given in (Acharya et al., 2017c), which is desired in the biomedical signal recording.

The limitation of our work is that we have used only 15 CHF subjects in the study. Our developed algorithm needs to be tested with a huge dataset before clinical usage.

6.5 Summary

In this work, the CHF and normal subjects are analyzed using 500, 1000, and 2000 samples of HRV signals. The AFEnt and APEnt are computed at different frequency scales of HRV signals (CHF and normal classes). To get the different frequency scales of HRV signals, we have used the sub-band signals obtained by the decomposition of HRV signals using the FAWT method. The HRV signals are decomposed up to the fifth level. Our methodology performed well for all three signal lengths of HRV signals. We have tested our methodology with two combinations of unbalanced and balanced datasets. The presented method performed significantly well for all four combinations of datasets. Moreover, we have used the LS-SVM classifier with TFCV method and tested it with four different kernels (linear, RBF, polynomial, and Morlet wavelet), which increases the robustness of our method. This method may assist clinicians in the faster diagnosis of CHF patients and to provide timely treatment. Thus, it may be a useful tool in improving the quality of life of CHF-affected patients.

Chapter 7

A computer-aided methodology for the identification of normal and AF ECG signals

7.1 Introduction

AF is a type of cardiac arrhythmia which can be characterized by irregularity and rapidity of cardiac contraction (Asgari et al., 2015; Markides & Schilling, 2003). About 2.2 million people in the USA and 4.5 million people in Europian Union have this arrhythmia (Fuster & et. al., 2011). In the next 40 years, the number of people having AF is expected to double in north America (Gillis et al., 2013). AF may lead to stroke and CHF (Fuster & et. al., 2011). The death rate for AF patients is double as compared to the normal subjects (Fuster & et. al., 2011). The risk of death may be reduced if timely treatment is available for the AF patients. Hence, an early stage diagnosis of AF can avoid unnecessary complications and save a life. The AF patients may show dyspnea and fatigue (Asgari et al., 2015). These signs may not be present in all the AF patients. Hence, there is a possibility that many AF patients may be left undetected. In order to diagnose the AF, the ECG is widely used by the cardiologists. Manual detection of AF using ECG is a tedious and time-consuming task (Asgari et al., 2015). The Ar of the AF detection by many primary care physicians using ECG signals is not sufficiently high (Mant et al., 2007). Hence, a computer-aided detection system for AF patients is required to help the doctors

for accurate diagnosis.

Various approaches are suggested to diagnose the presence of AF. A method based on the absence of P-wave for AF detection over a short duration of ECG is proposed in (Ladavich & Ghoraani, 2015). An AF detection method based on time domain features is also presented in (Du et al., 2014). In (Slocum et al., 1992), a computer-based AF detection algorithm is suggested. This algorithm is based on power spectral analysis of remainder ECG. Percent power of the remainder ECGs is observed to be significantly different for the rhythms of AF and control groups. A method based on SEnt is presented in (Lake & Moorman, 2011) and can detect the AF using 12 ECG beats. The AF detection technique proposed in (Zhou et al., 2014) computes symbolic sequence from the RR-interval sequence. Further, Shannon entropy is computed from the symbolic-sequence to detect the presence of AF. In (Jiang et al., 2012), the RR-interval distribution difference curve is used to detect the transition between the AF and normal rhythms. It is the reflection of the variability of RR-interval. In (Lee et al., 2013), root mean square of successive RR differences (RMSSD), Shannon entropy, and SEnt are found to be useful in the assessment of AF. A four-step process is explored for the AF detection from RR-interval in (Huang et al., 2011). These steps include histogram, standard deviation analysis, numbering abbrent rhythm recognition, and Kolmogorov-Smirnov (KS) test.

Our objective is to propose a computer-based approach to diagnose the AF patients automatically. To achieve the objective, we have computed entropy based features in FAWT (Bayram, 2013; Zhang et al., 2015) domain to discriminate NR and AF ECG segments. Here, NR represents the rhythm other than AF, atrial flutter (AFL), and AV junctional rhythms in the MIT-BIH AF database (MIT-BIH AFDB) (Moody & Mark, 1983; Goldberger et al., 2000). The ECG segments are decomposed in to the sub-band signals using FAWT. We have computed LEE and PEnt from these sub-band signals. Then, the features are fed to the RF (Breiman, 2001) and J48 DT (Quinlan, 1986, 1993) classifiers to classify the ECG segments of NR and AF classes. The block diagram of the proposed system is shown in Figure 7.1. The remaining of the chapter is organized in the following way:

The dataset used, preprocessing, segmentation of the ECG signals, entropy based

features, and classification methods are provided in section 7.2. The obtained results in this work are given in section 7.3 and discussed in section 7.4. Finally, section 7.5 presents the summary of the work.

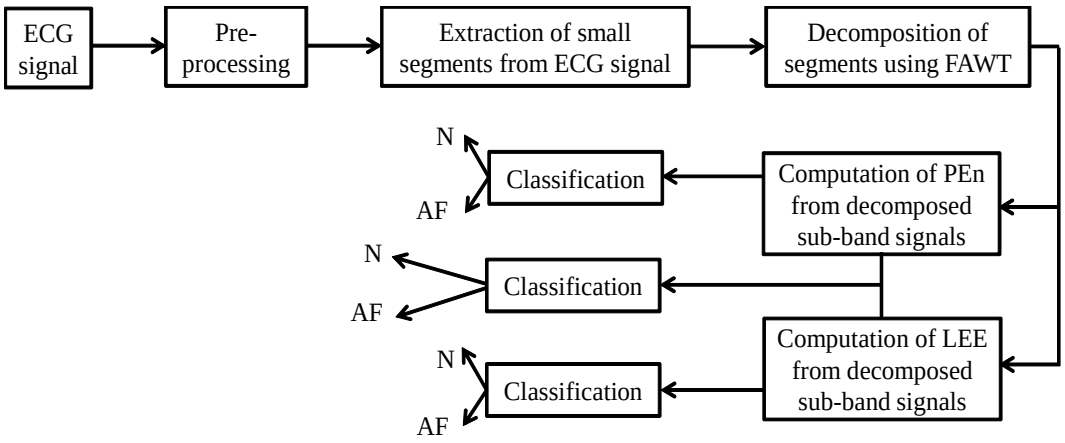


Figure 7.1: The proposed computer-aided identification method for AF ECG signals.

7.2 Methodology

7.2.1 Dataset used

In this work, we have used MIT-BIH AFDB which is publicly available at physionet. The dataset has ECG signals of 23 AF patients. The ECG signals present in the dataset are sampled at the rate of 250 samples per second. The duration of the recording is 10 hours. The dataset contains four types of rhythm AF, AFL, AV junctional, and NR. In this work, we have studied AF and NR rhythms. We have formed two classes namely, NR and AF from the database according to the annotations available in this dataset. We have excluded the two patients ("04936" and "05091") which are wrongly annotated (Zhou et al., 2014) in this study. Hence, we have analysed 243 NR and 247 AF rhythms of ECG signals corresponding to 21 AF patients. In the dataset, there are two ECG signals corresponding to each patient. We have used first ECG signal corresponding to each patient in this work.

7.2.2 Noise removal and segmentation of ECG signals

The db6 wavelet basis function for eliminating baseline wander and noise from the ECG signals is utilized (Martis et al., 2013a,b). After pre-processing, we have extracted 1000 samples segments from the NR and AF rhythms from 21 patients' ECG signals. The 1000 samples of the signal corresponding to the 4 second duration of the signal. Finally, we have 77,267 AF ECG segments and 114,214 NR ECG segments of length 1000 samples.

7.2.3 Features

We have used LEE and PEnt features to study the NR and AF ECG segments in this work. These features are described below:

7.2.4 Log energy entropy

The LEE parameter can be defined as follows (Han et al., 2009):

$$\text{LEE} = \sum_{j=1}^{P} \log(x_j^2)$$

where, x_j is the j^{th} sample of the signal, and P is the number of samples in the signal.

7.2.5 Permutation entropy

In this work, PEnt is used for complexity analysis of NR and AF ECG segment. Its explanation is provided in chapter 6. In this work, we have selected $e = 3$ (Bandt & Pompe, 2002).

7.2.6 Classification methods

The extracted features are applied to the input of J48 DT and RF classifiers to classify NR and AF ECG segments in this work. The WEKA toolbox is used to implement the classifiers with their default parameters (Hall et al., 2009).

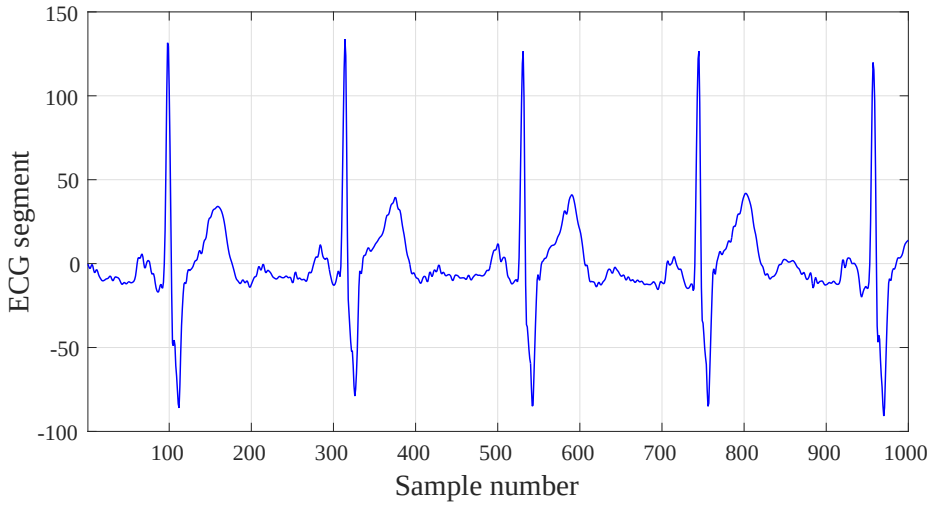
We have used the performance measure namely, Ar, Sc, and Ss in this work (Azar & El-Said, 2014). TFCV method is also employed to train and test the classifiers (Kohavi, 1995).

7.3 Results

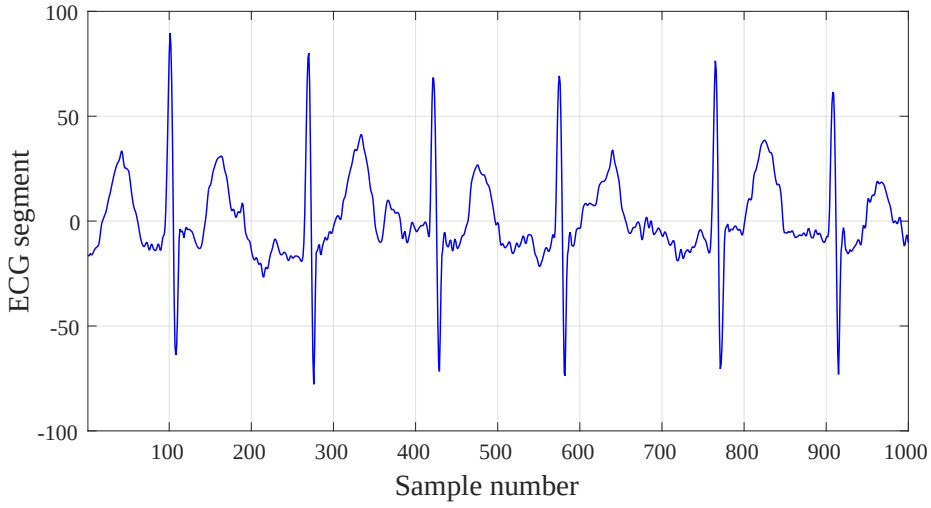
First, the long duration ECG signals are segmented into small segments of 1000 samples. The plots of ECG segments of both classes can be observed in Figure 7.2. Each segment is subjected to FAWT and decomposed into various sub-bands. The values of parameters b, c, f, and g are selected 5, 6, 1, and 2, respectively, and β is chosen as $(0.8 \times f)/g$ (Zhang et al., 2015) in this work. We have analysed the ECG segments at different levels of decomposition (J = 5, 10, 15, and 20). The obtained sub-band signals at $J = 5$ are shown in Figure 7.3. For $J = 5$, we have 5 detail sub-band signals and one approximate signal. In Figure 7.3, sub-band signals S_1 to S_6 are arranged in decreasing order of frequency. The sub-band signals S_1 and S_6 are corresponding to the highest and lowest frequency sub-band signals. The S_6 sub-band signal is reconstructed from approximate coefficients, and other sub-band signals are reconstructed from the corresponding detail coefficients obtained using FAWT. In the same way, for J = 10, 15, and 20, we have total 11 (10 detail and 1 approximate), 16 (15 detail and 1 approximate), and 21 (20 detail and 1 approximate) sub-band signals, respectively.

From the decomposed ECG segments, we have computed LEE and PEnt parameters. The STD and μ values for LEE and PEnt are summarised in Tables 7.1 and 7.2. In this work, we have employed Student's t-test (Box, 1987) to evaluate the discrimination ability of the computed features. The p-values corresponding to the computed features are also given in the Tables 7.1 and 7.2 respectively. We can observe that the p-values are less than 0.05 for all LEE and PEnt features except the LEE feature computed from S_2 sub-band signal.

Further, the features are applied to RF and J48 classifiers. We have observed the classification performance using both the features separately and also with their combination. The classification results for the individual feature are arranged in Tables 7.3 and 7.4. The classification performance of RF classifier is found to be



(a)

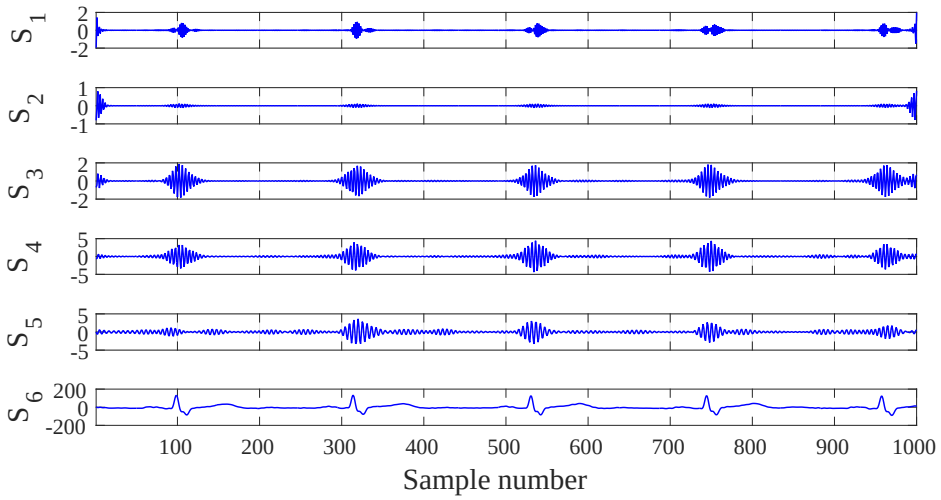


(b)

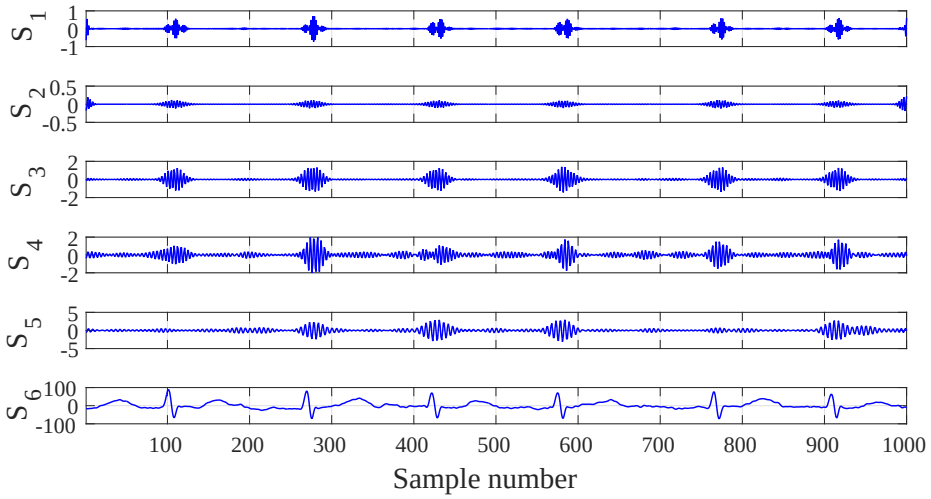
Figure 7.2: The plots of ECG segments for two classes: (a) NR, (b) AF.

Table 7.1: The μ and STD values and corresponding p-values for LEE features computed at $J = 5$ level.

Sub-band signal⟶		S_1	S_2	S_3	S_4	S_5	S_6
AF	Mean	-7019.06	-10099.54	-5065.82	-2982.59	-2492.53	4838.49
	STD	1604.63	1730.39	1700.76	1515.11	1296.66	1955.98
NR	Mean	-6966.14	-10087.94	-5110.26	-2930.00	-2181.84	5244.25
	STD	1276.95	1235.13	1294.48	1302.97	1264.18	1496.54
p-values		p <0.001	p=0.08	p <0.001	p <0.001	p <0.001	p <0.001



(a)



(b)

Figure 7.3: The plots of decomposed ECG segments using FAWT for two classes: (a) NR, (b) AF.

Table 7.2: The μ and STD values and corresponding p-values for PEnt features computed at $J = 5$ level.

Sub-band signal⟶		S_1	S_2	S_3	S_4	S_5	S_6
AF	Mean	1.59107	1.77694	1.74185	1.69385	1.62561	1.26887
	STD	0.0185	0.0018	0.0048	0.0062	0.0054	0.0863
NR	Mean	1.59416	1.77688	1.74266	1.69210	1.62449	1.27385
	STD	0.0196	0.0023	0.0053	0.0059	0.0052	0.0768
p-values		p <0.001	p <0.001	p <0.001	p <0.001	p <0.001	p <0.001

better than J48 classifier. The classification performances of both the classifiers are better for LEE features than PEnt features. The LEE features showed the highest Ar of 96.84% when subjected to the RF classifier with $J = 20$.

Table 7.3: Classification performance of RF classifier at various decomposition levels (J) of FAWT.

Feature	J	Total features	Ar (%)	Sc (%)	Ss (%)
LEE	5	6	86.12	90.01	80.2
	10	11	91.16	93.6	87.5
	15	16	95.12	96.3	93.5
	20	21	96.84	97.6	95.8
PEnt	5	6	70.50	80.6	55.6
	10	11	76.45	85.2	63.6
	15	16	81.07	85.6	74.3
	20	21	85.84	86.8	84.5

For PEnt, the maximum classification Ar of 85.84% is observed using RF classifier at $J = 20$. The LEE and PEnt features yielded 93.41% and 79.77% Ar with J48 classifier at $J = 20$ respectively. The classification results for the combination of the two features are provided in Table 7.5.

Table 7.4: Classification performance of J48 classifier for different values of J of FAWT.

Feature	J	Total features	Ar (%)	Sc (%)	Ss (%)
LEE	5	6	83.75	87.7	77.9
	10	11	87.6	90	84
	15	16	91.5	92.9	89.5
	20	21	93.41	94.5	91.8
PEnt	5	6	69.78	79.6	55.2
	10	11	71.93	79.1	61.3
	15	16	75.33	78.9	70
	20	21	79.77	81.1	77.8

From, Tables 7.3, 7.4, and 7.5, we can observe that LEE parameter has yielded

higher classification performance than PEnt, and the combination of PEnt and LEE for the same number of features. In Figure 7.4, classification Ar for each fold is shown for RF and J48 classifiers at $J = 20$.

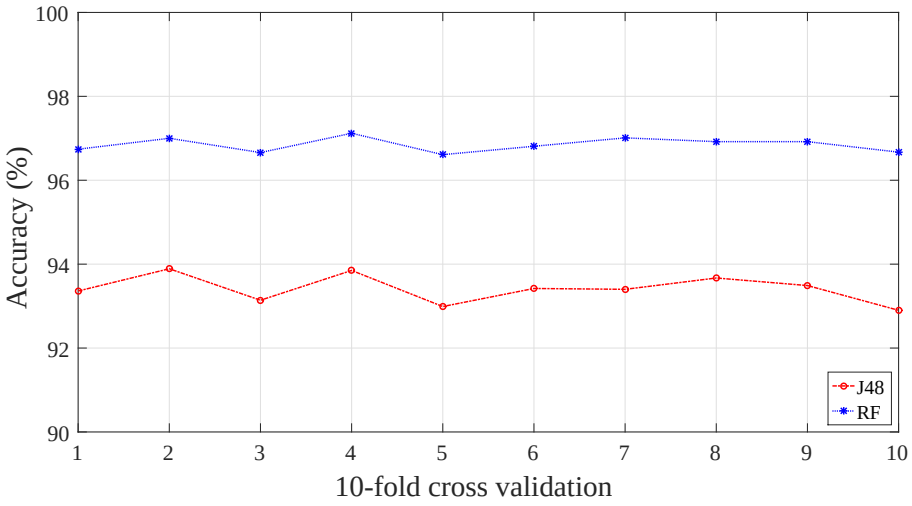


Figure 7.4: The plot of classification Ar at each fold for TFCV process.

Table 7.5: Classification performance of RF and J48 classifiers at various decomposition levels (J) of FAWT with the combination of LEE and PEnt features.

Classifier	J	Features	Ar (%)	Sc (%)	Ss (%)
RF	5	12	89.73	93.2	84.5
	10	22	92.24	94.7	88.5
J48	5	12	85.90	88.9	81.5
	10	22	88.06	90.2	85

7.4 Discussion

In this work, we have used FAWT to decompose the ECG segments of the two classes due to the non-stationary nature of ECG signals. The LEE and PEnt are nonlinear features which are employed to capture the nonlinear information present in the ECG segments. We have analysed the signal at different levels of decomposition. These decomposed sub-band signals are corresponding to the different frequency bands. The computation of entropy features from different frequency bands are performed.

We have noticed that the Ar of classification is increased with the increase in levels of decomposition which also increases the computation cost with the increase in the number of features.

From Table 7.1, we can observe that LEE features show lower value for AF in all the sub-band signals except S_3 sub-band signal. We have also observed negative values of LEE features for HFSBSs (S_1 to S_5) and positive values for LFSBS (S_6) in Table 7.1 which indicates that LFSBS (S_6) has higher energy as compared to the HFSBSs. From Table 7.2, we can observe that PEnt features showed higher value for S_1, S_3, and S_6 sub-band signals for NR class, and at S_2, S_4, and S_5 sub-band signals showed lower values for NR class as compared to AF class. Moreover, LEE is an energy-based feature and computation of PEnt is based on the occurrence of a pattern in the time series. The classification performance of LEE features is found much better than PEnt features, which means that energy based feature is more efficient as compared to the similarity based feature for the identification of AF ECG segment. Moreover, we have also used the combination of the two features with the classifiers. We found an increment in the Ar of classification. If we compare the classification performance of LEE features, PEnt features, and the combination of LEE and PEnt features, we can observe that for the same number of features performance of LEE is higher than PEnt, and the combination of LEE and PEnt features.

In Table 7.6, we have provided the summary of the automated diagnosis techniques used for the detection of AF. In (Tateno & Glass, 2001), the density histogram (DH) of the RR and $\triangle$ RR-intervals is determined and compared with the standard density histograms (SDH). Then, KS test is used to evaluate the difference between the DH of a subject and SDH. The $\triangle$ RR-interval based approach showed better results with a Ss of 94.4% and Sc of 97.2%. In (Babaeizadeh et al., 2009), an algorithm based on RR-intervals is provided which utilized Markov modeling for computing the R-R Markov score. The P wave morphology similarity and P-R interval variability measures are also used. The algorithm achieved 92% Ss in detecting the AF episodes.

In (Dash et al., 2009), RMSSD, turning points ratio, and Shannon entropy are computed from the RR-interval time series for the characterization of AF. They

Table 7.6: Summary of the computer-based automatic AF detection methods.

Author and years	Dataset	Methods	Results (%)
Tateno and Glass (2001)	MIT-BIH AFDB	RR and △ RR-intervals, KS-test	Ss = 94.40 Sc = 97.20
Babaeizadeh et al. (2009)	633 Holter ECG records MIT-BIH AFDB	P wave similarity and P-R interval, variation measures	Ss = 92
Dash et al. (2009)	MIT-BIH AFDB	RR-intervals analysis sample entropy	Ss = 94.4 Sc = 95.1
Rodenas et al. (2015)	MIT-BIH AFDB	Wavelet entropy analysis	Ar = 95.28 Ss = 96.47 Sc = 94.19
Present work	MIT-BIH AFDB	FAWT LEE features	Ar = 96.84 Ss = 95.8 Sc = 97.6

observed a Sc and Ss of 95.1% and 94.4%, respectively. A wavelet entropy-based approach is suggested for detecting AF episodes in (Rodenas et al., 2015). Their method yielded the Ar, Ss, and Sc of 95.28%, 96.47%, and 94.19% respectively.

In comparison to the methods (Tateno & Glass, 2001; Babaeizadeh et al., 2009; Dash et al., 2009; Rodenas et al., 2015), the main advantage of our method is that we do not need to detect the R-peak and P-wave. In (Tateno & Glass, 2001; Babaeizadeh et al., 2009; Dash et al., 2009), the Ar of AF detection depends on the correct detection of R-peaks as they are using RR-intervals in their analysis. If there is an error in the R-peak detection, then the Ar of AF detection is also affected. Due to the use of ECG signals, our system does not suffer from this limitation. Hence, our proposed method is more robust in comparison to these methods. In (Rodenas et al., 2015), TQ interval was computed after R-peak detection. They used an adaptive approach to make it insensitive to the possible error of R-peak detection. However, our results are comparable to their results. We have achieved better Ar as compared to (Rodenas et al., 2015).

7.5 Summary

In this work, we have computed LEE and PEnt in FAWT framework to analyze NR and AF ECG segments. Our results show that the LEE features performed better than the PEnt features in the automated diagnosis of AF using ECG signals. We have also observed that the performance of our system increased with decomposition levels. We have provided the results for 4 different levels of decomposition ($J = 5$, 10, 15, and 20). Classification Ar for $J = 20$ is higher as compared to $J = 5$, 10, and 15. This automated system can be implemented in hospitals to help the cardiologists to reduce their workload.

Chapter 8

Conclusion and future work

8.1 Conclusion

The cardiac signals (ECG, HRV) are able to serve as a potential tool for the diagnosis of different diseases related to the heart. The use of ASP methods together with the machine learning methods can effectively execute the task of decision making about the abnormal conditions of the heart.

CAD and normal HRV signals are investigated using FAWT. HRV signals are decomposed up to the third level using FAWT. Two computed parameters, K-NN entropy estimator and FEnt can capture the nonlinear dynamics of HRV signals. Several ranking methods are used to optimize the classification performance. Highest classification performance is observed with initial four significant parameters computed using entropy ranking method.

ECG beats of normal and CAD subjects are analyzed. Five CIP parameters are computed from the five levels of FAWT detail coefficients. These CIP parameters have shown higher values for CAD ECG beats as compared to normal ECG beats. It indicates more similarity between the detail coefficients of FAWT based decomposition of CAD beats than the normal beats. The present methodology can be extended for early detection of the CAD.

Normal and MI ECG beats are studied using SEnt in FAWT framework. Extracted features are applied to the input of different classifiers to achieve the highest performance. The classification Ar is computed for different values of J. The performance of the proposed method is highest for LS-SVM classifier at $J = 24$.

HRV signals of CHF and normal subjects are analyzed using AFEnt and APEnt at different frequency scales. To get the different frequency scales of HRV signals, the sub-band signals obtained using the FAWT method are accumulated in different combinations. We observed that AFEnt is more suitable to analyze the complexity of CHF and normal HRV signals. Most of the frequency-scaled signals showed lower mean values of AFEnt for CHF HRV signals.

ECG segments of NR and AF rhythms are studied using LEE and PEnt features in FAWT framework. The LEE and PEnt are nonlinear features which are employed to capture the nonlinear information present in the ECG segments. The results indicated that the LEE features performed better than the PEnt features in the automated diagnosis of AF.

8.2 Future work

The proposed automated identification methods based on ASP techniques and machine learning are found effective in decision making using cardiac signals. The proposed diagnosis framework for heart diseases consists of preprocessing, segmentation, feature extraction and selection, and classification of cardiac signals. However, further improvement of the proposed methodologies is possible in future by addressing the following issues.

Further work can be done for reducing the computation time of the used features in the proposed work. New features with lesser computation time can be proposed, and dedicated hardware may be implemented using the proposed methodologies. We have performed trial and error experimentation for the parameter selection of the studied features, kernels, and other classification parameters. Hence, it may possible that the selected values of these parameters are not the optimal values. Further work can be performed for the optimal selection of the parameters values and an automated parameter selection techniques may be developed. In future, we also intend to perform the classification using deep learning algorithms. Another limitation of the proposed methods is the use of the limited dataset. Hence, the testing of the developed methods with the huge dataset is needed before their implementation for the clinical purpose. Moreover, we have considered two class

classification in this thesis work. Further work can be performed in the direction of multi-class classification of the heart disorders.

The extension and modification of the suggested methods for the analysis of other biomedical signals such as EEG and EMG would be of great interest. With the incorporation of some new features, our proposed framework can also be utilized for the diagnosis of other abnormal conditions like as, arrhythmia, alzheimer's disease, ischemia, autism, septal defects, congenial heart diseases, and diseases related to the heart valves.

Bibliography

Acharya, U. R., Bhat, P. S., Iyengar, S. S., Rao, A., & Dua, S. (2003). Classification of heart rate data using artificial neural network and fuzzy equivalence relation. *Pattern Recognition*, *36*, 61–68.

Acharya, U. R., Bhat, P. S., Kannathal, N., Rao, A., & Lim, C. M. (2005). Analysis of cardiac health using fractal dimension and wavelet transformation. *ITBM-RBM*, *26*, 133–139.

Acharya, U. R., Faust, O., Kadri, N. A., Suri, J. S., & Yu, W. (2013). Automated identification of normal and diabetes heart rate signals using nonlinear measures. *Computers in Biology and Medicine*, *43*, 1523–1529.

Acharya, U. R., Faust, O., Sree, V., Swapna, G., Martis, R. J., Kadri, N. A., & Suri, J. S. (2014). Linear and nonlinear analysis of normal and CAD-affected heart rate signals. *Computer Methods and Programs in Biomedicine*, *113*, 55–68.

Acharya, U. R., Fujita, H., Adam, M., Lih, O. S., Sudarshan, V. K., Hong, T. J., Koh, J. E. W., Hagiwara, Y., Chua, C. K., Poo, C. K., & San, T. R. (2017a). Automated characterization and classification of coronary artery disease and myocardial infarction by decomposition of ECG signals: A comparative study. *Information Sciences*, *377*, 17–29.

Acharya, U. R., Fujita, H., Oh, S. L., Hagiwara, Y., Tan, J. H., & Adam, M. (2017b). Application of deep convolutional neural network for automated detection of myocardial infarction using ECG signals. *Information Sciences*, *415*, 190–198.

Acharya, U. R., Fujita, H., Sudarshan, V. K., Bhat, S., & Koh, J. E. W. (2015a). Application of entropies for automated diagnosis of epilepsy using EEG signals: A review. *Knowledge-Based Systems*, *88*, 85–96.

Acharya, U. R., Fujita, H., Sudarshan, V. K., Lih Oh, S., Muhammad, A., Koh, J. E. W., Hong Tan, J., Chua, C. K., Poo Chua, K., & San Tan, R. (2017c). Application of empirical mode decomposition (EMD) for automated identification of congestive heart failure using heart rate signals. *Neural Computing and Applications*, *28*, 3073–3094.

Acharya, U. R., Fujita, H., Sudarshan, V. K., Oh, S. L., Adam, M., Koh, J. E. W., Tan, J. H., Ghista, D. N., Martis, R. J., Chua, C. K., Poo, C. K., & Tan, R. S. (2016). Automated detection and localization of myocardial infarction using electrocardiogram: A comparative study of different leads. *Knowledge-Based Systems*, *99*, 146–156.

Acharya, U. R., Fujita, H., Sudarshan, V. K., Oh, S. L., Adam, M., Tan, J. H., Koo, J. H., Jain, A., Lim, C. M., & Chua, K. C. (2017d). Automated characterization of coronary artery disease, myocardial infarction, and congestive heart failure using contourlet and shearlet transforms of electrocardiogram signal. *Knowledge-Based Systems*, *132*, 156–166.

Acharya, U. R., Kannathal, N., & Krishnan, S. M. (2004). Comprehensive analysis of cardiac health using heart rate signals. *Physiological Measurement*, *25*, 1139–1151.

Acharya, U. R., Sankaranarayanan, M., Nayak, J., Xiang, C., & Tamura, T. (2008). Automatic identification of cardiac health using modeling techniques: A comparative study. *Information Sciences*, *178*, 4571–4582.

Acharya, U. R., Sudarshan, V. K., Koh, J. E. W., Martis, R. J., Tan, J. H., Oh, S. L., Muhammad, A., Hagiwara, Y., Mookiah, M. R. K., Chua, K. P., Chua, C. K., & Tan, R. S. (2017e). Application of higher-order spectra for the characterization of coronary artery disease using electrocardiogram signals. *Biomedical Signal Processing and Control*, *31*, 31–43.

Acharya, U. R., Suri, J. S., Spaan, J. A. E., & Krishnan, S. M. (2007). *Advances in Cardiac Signal Processing*. Springer-Verlag.

Acharya, U. R., Vidya, K. S., Ghista, D. N., Lim, W. J. E., Molinari, F., & Sankaranarayanan, M. (2015b). Computer-aided diagnosis of diabetic subjects by heart

rate variability signals using discrete wavelet transform method. *Knowledge-Based Systems*, *81*, 56–64.

Al-Kindi, S. G., Ali, F., Farghaly, A., Nathani, M., & Tafreshi, R. (2011). Towards real-time detection of myocardial infarction by digital analysis of electrocardiograms. In *1st Middle East Conference on Biomedical Engineering* (pp. 454–457). Sarjah, UAE.

Antanavičius, K., Bastys, A., Blužas, J., Gargasas, L., Kaminskienė, S., Urbonavičienė, G., & Vainoras, A. (2008). Nonlinear dynamics analysis of electrocardiograms for detection of coronary artery disease. *Computer Methods and Programs in Biomedicine*, *92*, 198–204.

Arafat, S., Dohrmann, M., & Skubic, M. (2005). Classification of coronary artery disease stress ECGs using uncertainty modeling. In *2005 ICSC Congress on Computational Intelligence Methods and Applications* (pp. 1–4). Istanbul, Turkey.

Arbolishvili, G., Mareev, V., Orlova, I., & Belenkov, I. (2006). Heart rate variability in chronic heart failure and its role in prognosis of the disease. *Kardiologiia*, *46*, 4–11.

Arif, M., Malagore, I. A., & Afsar, F. A. (2010). Automatic detection and localization of myocardial infarction using back propagation neural networks. In *4th International Conference on Bioinformatics and Biomedical Engineering* (pp. 1–4). Chengdu,China.

Asgari, S., Mehrnia, A., & Moussavi, M. (2015). Automatic detection of atrial fibrillation using stationary wavelet transform and support vector machine. *Computers in Biology and Medicine*, *60*, 132–142.

Asyali, M. H. (2003). Discrimination power of long-term heart rate variability measures. In *Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 200–203). Cancun, Mexico.

Azar, A. T., & El-Said, S. A. (2014). Performance analysis of support vector machines classifiers in breast cancer mammography recognition. *Neural Computing and Applications*, *24*, 1163–1177.

Babaeizadeh, S., Gregg, R. E., Helfenbein, E. D., Lindauer, J. M., & Zhou, S. H. (2009). Improvements in atrial fibrillation detection for real-time monitoring. *Journal of Electrocardiology*, *42*, 522–526.

Babaoglu, I., Findik, O., & Ülker, E. (2010). A comparison of feature selection models utilizing binary particle swarm optimization and genetic algorithm in determining coronary artery disease using support vector machine. *Expert Systems with Applications*, *37*, 3177–3183.

Babaoğlu, I., Findik, O., & Bayrak, M. (2010). Effects of principle component analysis on assessment of coronary artery diseases using support vector machine. *Expert Systems with Applications*, *37*, 2182–2185.

Baim, D. S., Colucci, W. S., Monrad, E. S., Smith, H. S., Wright, R. F., Lanoue, A., Gauthier, D. F., Ransil, B. J., Grossman, W., & Braunwald, E. (1986). Survival of patients with severe congestive heart failure treated with oral milrinone. *Journal of the American College of Cardiology*, *7*, 661–670.

Bajaj, V., & Pachori, R. B. (2012). Classification of seizure and nonseizure EEG signals using empirical mode decomposition. *IEEE Transactions on Information Technology in Biomedicine*, *16*, 1135–1142.

Bandt, C., & Pompe, B. (2002). Permutation entropy: a natural complexity measure for time series. *Physical Review Letters*, *88*, 174102.

Banerjee, S., & Mitra, M. (2010). ECG feature extraction and classification of anteroseptal myocardial infarction and normal subjects using discrete wavelet transform. In *International Conference on Systems in Medicine and Biology* (pp. 55–60). Kharagpur, India.

Banerjee, S., & Mitra, M. (2014). Application of cross wavelet transform for ECG pattern analysis and classification. *IEEE Transactions on Instrumentation and Measurement*, *63*, 326–333.

Bayram, I. (2013). An analytic wavelet transform with a flexible time-frequency covering. *IEEE Transactions on Signal Processing*, *61*, 1131–1142.

Bayram, I., & Selesnick, I. W. (2009). Frequency-domain design of overcomplete rational delation wavelet transforms. *IEEE Transactions on Signal Processing*, *57*, 2957–2972.

Bhati, D., Sharma, M., Pachori, R. B., & Gadre, V. M. (2017). Time–frequency localized three-band biorthogonal wavelet filter bank using semidefinite relaxation and nonlinear least squares with epileptic seizure EEG signal classification. *Digital Signal Processing*, *62*, 259–273.

Bhattacharyya, A., Pachori, R. B., & Acharya, U. R. (2017a). Tunable-Q wavelet transform based multivariate sub-band fuzzy entropy with application to focal EEG signal analysis. *Entropy*, *19*, 99.

Bhattacharyya, A., Pachori, R. B., Upadhyay, A., & Acharya, U. R. (2017b). Tunable-Q wavelet transform based multiscale entropy measure for automated classification of epileptic EEG signals. *Applied Sciences*, *7*, 385.

Bhattacharyya, A., Sharma, M., Pachori, R. B., Sircar, P., & Acharya, U. R. (2016). A novel approach for automated detection of focal EEG signals using empirical wavelet transform. *Neural Computing and Applications*, *10.1007/s00521-016-2646-4*, 1–11.

Biagetti, G., Crippa, P., Curzi, A., Orcioni, S., & Turchetti, C. (2014). A Multiclass ECG beat classifier based on the truncated KLT representation. In *European Modelling Symposium* (pp. 93–98). Pisa, Italy.

Bigger, J. T., Fleiss, J. L., Steinman, R. C., Rolnitzky, L. M., Schneider, W. J., & Stein, P. K. (1995). RR variability in healthy, middle-aged persons compared with patients with chronic coronary heart disease or recent acute myocardial infarction. *Circulation*, *91*, 1936–1943.

BIOPAC Systems Canada, Inc. (2010). Acqknowledge 4.1, BIOPAC Systems, Inc. http://www.biopac.ca/Acqknowledge_40.htm.

Bousseljot, R., Kreiseler, D., & Schnabel, A. (1995). Nutzung der EKG-Signaldatenbank CARDIODAT der PTB über das Internet. *Biomedizinische Technik/Biomedical Engineering*, *40*, 317–318.

Box, J. F. (1987). Guinness, gosset, fisher, and small samples. *Statistical Science*, *2*, 45–52.

Bozzola, P., Bortolan, G., Combi, C., Pinciroli, F., & Brohet, C. (1996). A hybrid neuro-fuzzy system for ECG classification of myocardial infarction. In *Computers in Cardiology* (pp. 241–244). Indianapolis, IN, USA.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5–32.

Chang, P. C., Hsieh, J. C., Lin, J. J., Chou, Y. H., & Liu, C. H. (2009). A hybrid system with hidden Markov models and Gaussian mixture models for myocardial infarction classification with 12-lead ECGs. In *11th IEEE International Conference on High Performance Computing and Communications* (pp. 110–116). Seoul, South Korea.

Chazal, P. d., O'Dwyer, M., & Reilly, R. B. (2004). Automatic classification of heartbeats using ECG morphology and heartbeat interval features. *IEEE Transactions on Biomedical Engineering*, *51*, 1196–1206.

Chen, S.-W. (2000). A two-stage discrimination of cardiac arrhythmias using a total least squares-based prony modeling algorithm. *IEEE Transactions on Biomedical Engineering*, *47*, 1317–1327.

Chen, S.-W. (2002). A wavelet-based heart rate variability analysis for the study of nonsustained ventricular tachycardia. *IEEE Transactions on Biomedical Engineering*, *49*, 736–742.

Chen, W., Wang, Z., Xie, H., & Yu, W. (2007). Characterization of surface EMG signal based on fuzzy entropy. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *15*, 266–272.

Chen, Y.-H., & Yu, S.-N. (2012). Selection of effective features for ECG beat recognition based on nonlinear correlations. *Artificial Intelligence in Medicine*, *54*, 43–52.

Chua, K. C., Chandran, V., Acharya, U. R., & Lim, C. M. (2008). Cardiac state diagnosis using higher order spectra of heart rate variability. *Journal of Medical Engineering & Technology*, *32*, 145–155.

Chui, C. K. (1992). *An Introduction to Wavelets*. Wavelet Analysis and Its Applications 1 (1st ed.). USA: Academic Press.

Cohen, L. (1989). Time-frequency distributions-A review. *Proceedings of the IEEE*, *77*, 941–981.

Cottrell, C. (2012). Atrial fibrillation part 1: pathophysiology. *Practice Nursing*, *23*, 16–21.

Crippa, P., Curzi, A., Falaschetti, L., & Turchetti, C. (2015). Multi-class ECG beat classification based on a Gaussian mixture model of Karhunen-Loève transform. *International Journal of Simulation: Systems, Science and Technology*, *16*, 2.1–2.10.

Cromwell, L., Weibell, F. J., & Pfeiffer, E. (2012). *Biomedical Instrumentation and Measurement*. PHI.

Dash, S., Chon, K. H., Lu, S., & Raeder, E. A. (2009). Automatic real time detection of atrial fibrillation. *Annals of Biomedical Engineering*, *37*, 1701–1709.

Derryberry, D. R., Schou, S. B., & Conover, W. J. (2010). Teaching rank-based tests by emphasizing structural similarities to corresponding parametric tests. *Journal of Statistics Education*, *18*, 1–19.

Du, X., Rao, N., Qian, M., Liu, D., Li, J., Feng, W., Yin, L., & Chen, X. (2014). A novel method for real-time atrial fibrillation detection in electrocardiograms using multiple parameters. *Annals of Noninvasive Electrocardiology*, *19*, 217–225.

Dua, S., Du, X., Sree, S. V., & V. I., T. A. (2012). Novel classification of coronary artery disease using heart rate variability analysis. *Journal of Mechanics in Medicine and Biology*, *12*, 1240017.

Duda, R. O., Hart, P. E., & Stork, D. G. (2000). *Pattern Classification*. (2nd ed.). John Willey & Sons.

Engin, M. (2004). ECG beat classification using neuro-fuzzy network. *Pattern Recognition Letters*, *25*, 1715–1722.

Fujita, H., Acharya, U. R., Sudarshan, V. K., Ghista, D. N., Sree, S. V., Eugene, L. W. J., & Koh, J. E. (2016). Sudden cardiac death (SCD) prediction based on nonlinear heart rate variability features and SCD index. *Applied Soft Computing*, *43*, 510–519.

Fuster, V., & et. al. (2011). ACCF/AHA/HRS focused updates incorporated into the ACC/AHA/ESC 2006 guidelines for the management of patients with atrial fibrillation. *Circulation*, *123*, e269–e367.

Gamero, L. G., Vila, J., & Palacios, F. (2002). Wavelet transform analysis of heart rate variability during myocardial ischaemia. *Medical and Biological Engineering and Computing*, *40*, 72–78.

Gao, R. X., & Yan, R. (2011). *Wavelets-Theory and Applications for Manufacturing*. (1st ed.). Springer US.

German-Sallo, Z. (2014). Wavelet transform based HRV analysis. *Procedia Technology*, *12*, 105–111.

Gillis, A. M., Krahn, A. D., Skanes, A. C., & Nattel, S. (2013). Management of atrial fibrillation in the year 2033: New concepts, tools, and applications leading to personalized medicine. *Canadian Journal of Cardiology*, *29*, 1141–1146.

Giri, D., Acharya, U. R., Martis, R. J., Sree, S. V., Lim, T.-C., VI, T. A., & Suri, J. S. (2013). Automated diagnosis of coronary artery disease affected patients using LDA, PCA, ICA and discrete wavelet transform. *Knowledge-Based Systems*, *37*, 274–282.

Goldberger, A. L., Amaral, L. A., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., Mietus, J. E., Moody, G. B., Peng, C.-K., & Stanley, H. E. (2000). Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals. *Circulation*, *101*, e215–e220.

Graaff, K. V. d., Rhees, R., & Palmer, S. (2010). *Schaum's Outlines; Human Anatomy and Physiology*. Schaum's Outline Series (3rd ed.). McGraw-Hil.

Gupta, V., Priya, T., Yadav, A. K., Pachori, R. B., & Acharya, U. R. (2017). Automated detection of focal EEG signals using features extracted from flexible analytic wavelet transform. *Pattern Recognition Letters*, *94*, 180–188.

Guzzetti, S., Mezzetti, S., Magatelli, R., Porta, A., Angelis, G. D., Rovelli, G., & Malliani, A. (2000). Linear and non-linear 24 h heart rate variability in chronic heart failure. *Autonomic Neuroscience*, *86*, 114–119.

Hadase, M., Azuma, A., Zen, K., Asada, S., Kawasaki, T., Kamitani, T., Kawasaki, S., Sugihara, H., & Matsubara, H. (2004). Very low frequency power of heart rate variability is a powerful predictor of clinical prognosis in patients with congestive heart failure. *Circulation Journal*, *68*, 343–347.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software: An update. *ACM: SIGKDD Explorations Newsletter*, *11*, 10–18.

Han, J., Dong, F., & Xu, Y. Y. (2009). Entropy feature extraction on flow pattern of gas/liquid two-phase flow based on cross-section measurement. *Journal of Physics: Conference Series*, *147*, 012041.

Haraldsson, H., Edenbrandt, L., & Ohlsson, M. (2004). Detecting acute myocardial infarction in the 12-lead ECG using Hermite expansions and neural networks. *Artificial Intelligence in Medicine*, *32*, 127–136.

Hayano, J., Sakakibara, Y., Yamada, M., Ohte, N., Fujinami, T., Yokoyama, K., Watanabe, Y., & Takata, K. (1990). Decreased magnitude of heart rate spectral components in coronary artery disease. Its relation to angiographic severity. *Circulation*, *81*, 1217–1224.

Huang, C., Ye, S., Chen, H., Li, D., He, F., & Tu, Y. (2011). A novel method for detection of the transition between atrial fibrillation and sinus rhythm. *IEEE Transactions on Biomedical Engineering*, *58*, 1113–1119.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N.-C., Tung, C. C., & Liu, H. H. (1998). The empirical mode decomposition and the

Hilbert spectrum for nonlinear and nonstationary time series analysis. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, *454*, 903–995.

Inan, O. T., Giovangrandi, L., & Kovacs, G. T. A. (2006). Robust neural-network-based classification of premature ventricular contractions using wavelet transform and timing interval features. *IEEE Transactions on Biomedical Engineering*, *53*, 2507–2515.

Ince, T., Kiranyaz, S., & Gabbouj, M. (2009). A generic and robust system for automated patient-specific classification of ECG signals. *IEEE Transactions on Biomedical Engineering*, *56*, 1415–1426.

Iyengar, N., Peng, C., Morin, R., Goldberger, A. L., & Lipsitz, L. A. (1996). Age-related alterations in the fractal scaling of cardiac interbeat interval dynamics. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, *271*, R1078–R1084.

Jang, J.-S. R., Sun, C.-T., & Mizutani, E. (1997). *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*. Londan, UK: Pearson.

Ji, L., Li, P., Liu, C., Wang, X., Yang, J., & Liu, C. (2016). Measuring electromechanical coupling in patients with coronary artery disease and healthy subjects. *Entropy*, *18*, 153.

Jiang, K., Huang, C., Ye, S., & Chen, H. (2012). High accuracy in automatic detection of atrial fibrillation for Holter monitoring. *Journal of Zhejiang University-SCIENCE B*, *13*, 751–756.

Jong, T.-L., Chang, B., & Kuo, C.-D. (2011). Optimal timing in screening patients with congestive heart failure and healthy subjects during circadian observation. *Annals of Biomedical Engineering*, *39*, 835–849.

Kadambe, S., Murray, R., & Boudreaux-Bartels, G. F. (1999). Wavelet transform-based QRS complex detector. *IEEE Transactions on Biomedical Engineering*, *46*, 838–848.

Karamanos, K., Nikolopoulos, S., Hizanidis, K., Manis, G., Alexandridi, A., & Nikolakeas, S. (2006). Block entropy analysis of heart rate variability signals. *International Journal of Bifurcation and Chaos*, *16*, 2093–2101.

Karimi, M., Amirfattahi, R., Sadri, S., & Marvasti, S. A. (2005). Noninvasive detection and classification of coronary artery occlusions using wavelet analysis of heart sounds with neural networks. In *The 3rd IEE International Seminar on Medical Applications of Signal Processing* (pp. 117–120). London, UK.

Kaur, I., Rajni, R., & Marwaha, A. (2016). ECG signal analysis and arrhythmia detection using wavelet transform. *Journal of The Institution of Engineers (India): Series B*, *97*, 499–507.

Khaled, A. S., Owis, M. I., & Mohamed, A. S. (2006). Employing time-domain methods and Poincaré plot of heart rate variability signals to detect congestive heart failure. *BIME Journal*, *6*, 35–41.

Khandoker, A. H., Lai, D. T. H., Begg, R. K., & Palaniswami, M. (2007). Wavelet-based feature extraction for support vector machines for screening balance impairments in the elderly. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *15*, 587–597.

Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence* (pp. 1137–1143). Montreal, Quebec, Canada.

Korurek, M., & Dogan, B. (2010). ECG beat classification using particle swarm optimization and radial basis function neural network. *Expert Systems with Applications*, *37*, 7563–7569.

Kraskov, A., Stögbauer, H., & Grassberger, P. (2004). Estimating mutual information. *Phys. Rev. E*, *69*, 066138.

Ladavich, S., & Ghoraani, B. (2015). Rate-independent detection of atrial fibrillation by statistical modeling of atrial activity. *Biomedical Signal Processing and Control*, *18*, 274–281.

Lahiri, T., Kumar, U., Mishra, H., Sarkar, S., & Roy, A. D. (2009). Analysis of ECG signal by chaos principle to help automatic diagnosis of myocardial infarction. *Journal of Scientific and Industrial Research*, *68*, 866–870.

Lake, D. E., & Moorman, J. R. (2011). Accurate estimation of entropy in very short physiological time series: The problem of atrial fibrillation detection in implanted ventricular devices. *American Journal of Physiology-Heart and Circulatory Physiology*, *300*, H319–H325.

Lee, H. G., Noh, K. Y., & Ryu, K. H. (2007). Mining biosignal data: coronary artery disease diagnosis using linear and nonlinear features of HRV. In T. Washio, Z.-H. Zhou, J. Z. Huang, X. Hu, J. Li, C. Xie, J. He, D. Zou, K.-C. Li, & M. M. Freire (Eds.), *Emerging Technologies in Knowledge Discovery and Data Mining* (pp. 218–228). Springer Berlin Heidelberg volume 4819 of *Lecture Notes in Computer Science*.

Lee, J., Reyes, B. A., McManus, D. D., Mathias, O., & Chon, K. H. (2013). Atrial fibrillation detection using an iPhone 4S. *IEEE Transactions on Biomedical Engineering*, *60*, 203–206.

Li, H., Liang, H., Miao, C., Cao, L., Feng, X., Tang, C., & Li, E. (2016). Novel ECG signal classification based on KICA nonlinear feature extraction. *Circuits, Systems, and Signal Processing*, *35*, 1187–1197.

Liu, B., Liu, J., Wang, G., Huang, K., Li, F., Zheng, Y., Luo, Y., & Zhou, F. (2015). A novel electrocardiogram parameterization algorithm and its application in myocardial infarction detection. *Computers in Biology and Medicine*, *61*, 178–184.

Liu, G., Wang, L., Wang, Q., Zhou, G., Wang, Y., & Jiang, Q. (2014). A new approach to detect congestive heart failure using short-term heart rate variability measures. *Plos One*, *9*, e93399.

Lu, H. L., Ong, K., & Chia, P. (2000). An automated ECG classification system based on a neuro-fuzzy system. In *Computers in Cardiology* (pp. 387–390). Cambridge, USA.

Maestri, R., Pinna, G. D., Accardo, A., Allegrini, P., Balocchi, R., D'addio, G., Ferrario, M., Menicucci, D., Porta, A., Sassi, R., Signorini, M. G., La Rovere, M. T., & Cerutti, S. (2007). Nonlinear indices of heart rate variability in chronic heart failure patients: Redundancy and comparative clinical value. *Journal of Cardiovascular Electrophysiology*, *18*, 425–433.

Maheshwari, S., Pachori, R. B., & Acharya, U. R. (2017a). Automated diagnosis of glaucoma using empirical wavelet transform and correntropy features extracted from fundus images. *IEEE Journal of Biomedical and Health Informatics*, *21*, 803–813.

Maheshwari, S., Pachori, R. B., Kanhangad, V., Bhandary, S. V., & Acharya, U. R. (2017b). Iterative variational mode decomposition based automated detection of glaucoma using fundus images. *Computers in Biology and Medicine*, *88*, 142–149.

Malliani, A., Lombardi, F., Pagani, M., & Cerutti, S. (1994). Power spectral analysis of cardiovascular variability in patients at risk for sudden cardiac death. *Journal of Cardiovascular Electrophysiology*, *5*, 274–286.

Mant, J., Fitzmaurice, D. A., Hobbs, F. D. R., Jowett, S., Murray, E. T., Holder, R., Davies, M., & Lip, G. Y. H. (2007). Accuracy of diagnosing atrial fibrillation on electrocardiogram by primary care practitioners and interpretative diagnostic software: Analysis of data from screening for atrial fibrillation in the elderly (SAFE) trial. *British Medical Journal*, *335*, 380.

Markides, V., & Schilling, R. J. (2003). Atrial fibrillation: Classification, pathophysiology, mechanisms and drug treatment. *Heart*, *89*, 939–943.

Martis, R. J., Acharya, U. R., Mandana, K. M., Ray, A. K., & Chakraborty, C. (2013a). Cardiac decision making using higher order spectra. *Biomedical Signal Processing and Control*, *8*, 193–203.

Martis, R. J., Acharya, U. R., & Min, L. C. (2013b). ECG beat classification using PCA, LDA, ICA and discrete wavelet transform. *Biomedical Signal Processing and Control*, *8*, 437–448.

McDarby, G., Celler, B. G., & Lovell, N. H. (1998). Characterising the discrete wavelet transform of an ECG signal with simple parameters for use in automated diagnosis. In *Proceedings of the 2nd International Conference on Bioelectromagnetism* (pp. 31–32). Melbourne, Australia.

McKight, P. E., & Najab, J. (2010). *Kruskal-Wallis Test: Corsini Encyclopedia of Psychology*. John Wiley and Sons.

Melillo, P., Luca, N. D., Bracale, M., & Pecchia, L. (2013). Classification tree for risk assessment in patients suffering from congestive heart failure via long-term heart rate variability. *IEEE Journal of Biomedical and Health Informatics*, *17*, 727–733.

Mishu, M. M. H., Hossain, A. B. M. A., & Emon, M. E. A. (2014). Denoising of ECG signals using dual tree complex wavelet transform. In *17th International Conference on Computer and Information Technology (ICCIT)* (pp. 379–382).

Moody, G. B., & Mark, R. G. (1983). A new method for detecting atrial fibrillation using R-R intervals. *Computers in Cardiology*, *10*, 227–230.

Musialik-Łydka, A., Sredniawa, B., & Pasyk, S. (2003). Heart rate variability in heart failure. *Kardiologia Polska*, *58*, 10–16.

Mussalo, H., Vanninen, E., Ikäheimo, R., Laitinen, T., Laakso, M., Länsimies, E., & Hartikainen, J. (2001). Heart rate variability and its determinants in patients with severe or mild essential hypertension. *Clinical Physiology*, *21*, 594–604.

Narin, A., Isler, Y., & Ozer, M. (2014). Investigating the performance improvement of HRV indices in CHF using feature selection methods based on backward elimination and statistical significance. *Computers in Biology and Medicine*, *45*, 72–79.

National Heart, Lung and Blood Institute (2015a). What Is Coronary Heart Disease? https://www.nhlbi.nih.gov/health/health-topics/topics/cad/(accessed on 2015-11-25).

National Heart, Lung and Blood Institute (2015b). What is heart failure? https://www.nhlbi.nih.gov/health/health-topics/topics/hf/(accessed on 2017-01-15).

Nolan, J., Batin, P. D., Andrews, R., Lindsay, S. J., Brooksby, P., Mullen, M., Baig, W., Flapan, A. D., Cowley, A., Prescott, R. J., Neilson, J. M. M., & Fox, K. A. A. (1998). Prospective study of heart rate variability and mortality in chronic heart failure. *Circulation*, *98*, 1510–1516.

Nowak, J., Ylean, M., Hagerman, I., Nyquist, O., & Sylven, C. (1993). Electrocardiogram signal variance analysis in the diagnosis of coronary artery disease– A comparison with exercise stress test in an angiographically documented high prevalence population. *Clinical Cardiology*, *16*, 671–682.

Ortiz, M. R., Bojorges, E. R., Aguilar, S. D., Echeverria, J. C., Gonzalez-Camarena, R., Carrasco, S., Gaitan, M. J., & Martinez, A. (2005). Analysis of high frequency fetal heart rate variability using empirical mode decomposition. In *Computers in Cardiology* (pp. 675–678). Lyon, France.

Osowski, S., Hoai, L. T., & Markiewicz, T. (2004). Support vector machine-based expert system for reliable heartbeat recognition. *IEEE Transactions on Biomedical Engineering*, *51*, 582–589.

Osowski, S., & Linh, T. H. (2001). ECG beat recognition using fuzzy hybrid neural network. *IEEE Transactions on Biomedical Engineering*, *48*, 1265–1271.

Pachori, R. B. (2008). Discrimination between ictal and seizure-free EEG signals using empirical mode decomposition. *Research Letters in Signal Processing*, *2008*, 1–5.

Pachori, R. B., Avinash, P., Shashank, K., Sharma, R., & Acharya, U. R. (2015). Application of empirical mode decomposition for analysis of normal and diabetic RR-interval signals. *Expert Systems with Applications*, *42*, 4567–4581.

Pachori, R. B., Hewson, D., Snoussi, H., & Duchene, J. (2009). Postural time-series analysis using empirical mode decomposition and second-order difference plots. In

2009 IEEE International Conference on Acoustics, Speech and Signal Processing (pp. 537–540).

Pachori, R. B., Kumar, M., Avinash, P., Shashank, K., & Acharya, U. R. (2016). An improved online paradigm for screening of diabetic patients using RR-interval signals. *Journal of Mechanics in Medicine and Biology*, *16*, 1640003.

Pachori, R. B., & Patidar, S. (2014). Epileptic seizure classification in EEG signals using second-order difference plot of intrinsic mode functions. *Computer Methods and Programs in Biomedicine*, *113*, 494–502.

Pan, J., & Tompkins, W. J. (1985). A real-time QRS detection algorithm. *IEEE Transactions on Biomedical Engineering*, *32*, 230–236.

Patidar, S. (2014). Tunable-Q wavelet transform based methodologies for analysis and classification of cardiac signals. *Ph.D. Thesis*, Discipline of Electrical Engineering, Indian Institute of Technology Indore, Indore.

Patidar, S., & Pachori, R. B. (2014). Classification of cardiac sound signals using constrained tunable-Q wavelet transform. *Expert Systems with Applications*, *41*, 7161–7170.

Patidar, S., Pachori, R. B., & Acharya, U. R. (2015a). Automated diagnosis of coronary artery disease using tunable-Q wavelet transform applied on heart rate signals. *Knowledge-Based Systems*, *82*, 1–10.

Patidar, S., Pachori, R. B., & Garg, N. (2015b). Automatic diagnosis of septal defects based on tunable-Q wavelet transform of cardiac sound signals. *Expert Systems with Applications*, *42*, 3315–3326.

Patidar, S., Pachori, R. B., Upadhyay, A., & Acharya, U. R. (2017). An integrated alcoholic index using tunable-Q wavelet transform based features extracted from EEG signals for diagnosis of alcoholism. *Applied Soft Computing*, *50*, 71–78.

Pazos-López, P., Peteiro-Vázquez, J., Carcía-Campos, A., García-Bueno, L., de Torres, J. P. A., & Castro-Beiras, A. (2011). The causes, consequences, and treatment of left or right heart failure. *Vascular Health and Risk Management*, *7*, 237–254.

Pecchia, L., Melillo, P., Sansone, M., & Bracale, M. (2011). Discrimination power of short-term heart rate variability measures for CHF assessment. *IEEE Transactions on Information Technology in Biomedicine*, *15*, 40–46.

Poddar, M. G., Kumar, V., & Sharma, Y. P. (2015). Automated diagnosis of coronary artery diseased patients by heart rate variability analysis using linear and non-linear methods. *Journal of Medical Engineering & Technology*, *39*, 331–341.

Ponikowski, P., Anker, S. D., AlHabib, K. F., Cowie, M. R., Force, T. L., Hu, S., Jaarsma, T., Krum, H., Rastogi, V., Rohde, L. E., Samal, U. C., Shimokawa, H., Budi Siswanto, B., Sliwa, K., & Filippatos, G. (2014). Heart failure: preventing disease and death worldwide. *European Society of Cardiology (ESC) Heart Failure*, *1*, 4–25.

Prabhakaran, D., Jeemon, P., & Roy, A. (2016). Cardiovascular diseases in India: Current epidemiology and future directions. *Circulation*, *133*, 1605–1620.

Qiao, S., & Zhou, P. (2007). Wavelet and wavelet packet transform analysis in the ECG signals of atrial fibrillation. In *IEEE/ICME International Conference on Complex Medical Engineering* (pp. 1766–1769).

Quinlan, J. R. (1986). Induction of decision trees. *Machine learning*, *1*, 81–106.

Quinlan, J. R. (1993). *C4.5: programs for machine learning*. San Francisco, CA, USA: Morgan Kaufmann.

Richman, J. S., & Moorman, J. R. (2000). Physiological time–series analysis using approximate entropy and sample entropy. *American Journal of Physiology Heart and Circulatory Physiology*, *278*, H2039–H2049.

Rodenas, J., Garcia, M., Alcaraz, R., & Rieta, J. J. (2015). Wavelet entropy automatically detects episodes of atrial fibrillation from single-lead electrocardiograms. *Entropy*, *17*, 6179–6199.

Roman, J., Vilacosta, I., Castillo, J., Rollan, M., Hernandez, M., Peral, V., Garcimartin, I., de la Torre, M. M., & Fernandez-Aviles, F. (1998). Selection of the optimal stress test for the diagnosis of coronary artery disease. *Heart*, *80*, 370–376.

Safdarian, N., Dabanloo, N. J., & Attarodi, G. (2014). A new pattern recognition method for detection and localization of myocardial infarction using T-wave integral and total integral as extracted features from one cycle of ECG signal. *Journal of Biomedical Science and Engineering*, *7*, 818–824.

Sanchis-Gomar, F., Perez-Quilis, C., Leischik, R., & Lucia, A. (2016). Epidemiology of coronary heart disease and acute coronary syndrome. *Annals of Translational Medicine*, *4*, 256.

Schreck, D. M., Ng, L., Schreck, B. S., Bosco, S. F., Allegra, J. R., Zacharias, D., & Wortzel, J. V. (1988). Detection of coronary artery disease from the normal resting ECG using nonlinear mathematical transformation. *Annals of Emergency Medicine*, *17*, 132–134.

Shahbazi, F., & Asl, B. M. (2015). Generalized discriminant analysis for congestive heart failure risk assessment based on long-term heart rate variability. *Computer Methods and Programs in Biomedicine*, *122*, 191–198.

Shahnaz, C., Anowar, T. B., Rafi, R. H. M., Ahmmed, I., & Fattah, S. A. (2015). Cardiac beat classification based on wavelet analysis of empirical mode decomposed ECG signals. In *IEEE Region 10 TENCON Conference* (pp. 1–6). Macao, China.

Sharma, L. N., Tripathy, R. K., & Dandapat, S. (2015a). Multiscale energy and eigenspace approach to detection and localization of myocardial infarction. *IEEE Transactions on Biomedical Engineering*, *62*, 1827–1837.

Sharma, M., Dhere, A., Pachori, R. B., & Acharya, U. R. (2017a). An automatic detection of focal EEG signals using new class of time–frequency localized orthogonal wavelet filter banks. *Knowledge-Based Systems*, *118*, 217–227.

Sharma, M., & Pachori, R. B. (2017). A novel approach to detect epileptic seizures using a combination of tunable-Q wavelet transform and fractal dimension. *Journal of Mechanics in Medicine and Biology*, *17*, 1740003.

Sharma, M., Pachori, R. B., & Acharya, U. R. (2017b). A new approach to charac-

terize epileptic seizures using analytic time-frequency flexible wavelet transform and fractal dimension. *Pattern Recognition Letters*, *94*, 172–179.

Sharma, R., Kumar, M., Pachori, R. B., & Acharya, U. R. (2017c). Decision support system for focal EEG signals using tunable-Q wavelet transform. *Journal of Computational Science*, *20*, 52–60.

Sharma, R., & Pachori, R. B. (2015). Classification of epileptic seizures in EEG signals based on phase space representation of intrinsic mode functions. *Expert Systems with Applications*, *42*, 1106–1117.

Sharma, R., Pachori, R. B., & Acharya, U. R. (2015b). An integrated index for the identification of focal electroencephalogram signals using discrete wavelet transform and entropy measures. *Entropy*, *17*, 5218–5240.

Sharma, R., Pachori, R. B., & Acharya, U. R. (2015c). Application of entropy measures on intrinsic mode functions for the automated identification of focal electroencephalogram signals. *Entropy*, *17*, 669–691.

Sharma, R., Pachori, R. B., & Upadhyay, A. (2017d). Automatic sleep stages classification based on iterative filtering of electroencephalogram signals. *Neural Computing and Applications*, *28*, 2959–2978.

Sharma, R. R., Kumar, M., & Pachori, R. (2017e). Automated CAD identification system using time-frequency representation based on eigenvalue decomposition of ECG signals. In *International Conference on Machine intelligence and Signal processing 22-24 December*. Indore, India.

Sharma, R. R., & Pachori, R. B. (2018). Time-frequency representation using IEVDHM-HT with application to classification of epileptic EEG signals. *IET Science, Measurement & Technology*, *12*, 72–82.

Singh, P., & Pachori, R. B. (2017). Classification of focal and nonfocal EEG signals using features derived from fourier-based rhythms. *Journal of Mechanics in Medicine and Biology*, *17*, 1740002.

Slocum, J., Sahakian, A., & Swiryn, S. (1992). Diagnosis of atrial fibrillation from surface electrocardiograms based on computer-detected atrial activity. *Journal of Electrocardiology*, *25*, 1–8.

Small, M., Yu, D. J., Grubb, I., Simonotto, J., Fox, K. A. A., & Harrison, R. G. (2000). Automatic identification and recording of cardiac arrhythmia. In *Computers in Cardiology* (pp. 355–358). Cambridge, MA, USA.

Sokunbi, M. O., Fung, W., Sawlani, V., Choppin, S., Linden, D. E. J., & Thome, J. (2013). Resting state fMRI entropy probes complexity of brain activity in adults with ADHD. *Psychiatry Research: Neuroimaging*, *214*, 341–348.

Sood, S., Kumar, M., Pachori, R. B., & Acharya, U. R. (2016). Application of empirical mode decomposition-based features for analysis of normal and CAD heart rate signals. *Journal of Mechanics in Medicine and Biology*, *16*, 1640002.

Stein, P. K., Kleiger, R. E., Domitrovich, P. P., Schechtman, K. B., & Rottman, J. N. (2000). Clinical and demographic determinants of heart rate variability in patients post myocardial infarction: Insights from the cardiac arrhythmia suppression trial (CAST). *Clinical Cardiology*, *23*, 187–194.

Sun, L., Lu, Y., Yang, K., & Li, S. (2012). ECG analysis using multiple instance learning for myocardial infarction detection. *IEEE Transactions on Biomedical Engineering*, *59*, 3348–3356.

Suykens, J. A. K., & Vandewalle, J. (1999). Least squares support vector machine classifiers. *Neural Processing Letters*, *9*, 293–300.

Tang, B., Dong, S., & Song, T. (2012). Method for eliminating mode mixing of empirical mode decomposition based on the revised blind source separation. *Signal Processing*, *92*, 248–258.

Tateno, K., & Glass, L. (2001). Automatic detection of atrial fibrillation using the coefficient of variation and density histograms of RR and $\triangle$ RR intervals. *Medical and Biological Engineering and Computing*, *39*, 664–671.

Thakre, T. P., & Smith, M. L. (2006). Loss of lag-response curvilinearity of indices of heart rate variability in congestive heart failure. *BioMed Central (BMC) Cardiovascular Disorders*, *6*, 27.

Theodoridis, S., & Koutroumbas, K. (2003). Feature selection. In *Pattern Recognition* (pp. 163–205). San Diego, USA: Academic Press. (2nd ed.).

Thygesen, K., & et al (2012). Third universal definition of myocardial infarction. *European Heart Journal*, *33*, 2551–2567.

Tiwari, A. K., Pachori, R. B., Kanhangad, V., & Panigrahi, B. K. (2017). Automated diagnosis of epilepsy using key-point-based local binary pattern of EEG signals. *IEEE Journal of Biomedical and Health Informatics*, *21*, 888–896.

Tripathy, R. K., & Dandapat, S. (2016). Detection of cardiac abnormalities from multilead ECG using multiscale phase alternation features. *Journal of Medical Systems*, *40*, 143.

Vaseghi, M., & Shivkumar, K. (2008). The role of the autonomic nervous system in sudden cardiac death. *Progress in Cardiovascular Diseases*, *50*, 404–419.

Veselkov, K. A., Pahomov, V. I., Lindon, J. C., Volynkin, V. S., Crockford, D., Osipenko, G. S., Davies, D. B., Barton, R. H., Bang, J.-W., Holmes, E., & Nicholson, J. K. (2010). A metabolic entropy approach for measurements of systemic metabolic disruptions in patho-physiological states. *Journal of Proteome Research*, *9*, 3537–3544.

Warlar, R., & Eswaran, C. (1991). Integer coefficient bandpass filter for the simultaneous removal of baseline wander, 50 and 100 Hz interference from the ECG. *Medical and Biological Engineering and Computing*, *29*, 333–336.

Wong, N. D. (2014). Epidemiological studies of CHD and the evolution of preventive cardiology. *Nature Reviews Cardiology*, *11*, 276–289.

World Health Organization: Cardiovascular diseases (2015). http://www.who.int/mediacentre/factsheets/fs317/en/ (accessed 2016-03-01).

Xu, D., & Erdogmuns, D. (2010). Renyi's entropy, divergence and their nonparametric estimators, in: Information Theoretic Learning: Renyi's Entropy and Kernel Perspectives. (pp. 47–102). Springer New York.

Xu, J.-W., Paiva, A. R. C., Park, I., & Principe, J. C. (2008). A reproducing kernel hilbert space framework for information-theoretic learning. *IEEE Transactions on Signal Processing*, *56*, 5891–5902.

Ye, C., Kumar, B. V. K. V., & Coimbra, M. T. (2012). Heartbeat classification using morphological and dynamic features of ECG signals. *IEEE Transactions on Biomedical Engineering*, *59*, 2930–2941.

Yeap, T. H., Johnson, F., & Rachniowski, M. (1990). ECG beat classification by a neural network. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 1457–1458). Philadelphia, PA, USA.

Yu, S.-N., & Lee, M.-Y. (2012). Bispectral analysis and genetic algorithm for congestive heart failure recognition based on heart rate variability. *Computers in Biology and Medicine*, *42*, 816–825.

Zanin, M., Zunino, L., Rosso, O. A., & Papo, D. (2012). Permutation entropy and its main biomedical and econophysics applications: A review. *Entropy*, *14*, 1553–1577.

Zavar, M., Rahati, S., Akbarzadeh-T, M.-R., & Ghasemifard, H. (2011). Evolutionary model selection in a wavelet-based support vector machine for automated seizure detection. *Expert Systems with Applications*, *38*, 10751–10758.

Zhang, C., Li, B., Chen, B., Cao, H., Zi, Y., & He, Z. (2015). Weak fault signature extraction of rotating machinery using flexible analytic wavelet transform. *Mechanical Systems and Signal Processing*, *64–65*, 162–187.

Zhao, Z., & Ma, C. (2008). An intelligent system for noninvasive diagnosis of coronary artery disease with EMD-TEO and BP Neural Network. In *International Workshop on Education Technology and Training and International Workshop on Geoscience and Remote Sensing.* (pp. 631–635). Shanghai, China.

Zhou, X., Ding, H., Ung, B., Pickwell-MacPherson, E., & Zhang, Y. (2014). Automatic online detection of atrial fibrillation based on symbolic dynamics and Shannon entropy. *BioMedical Engineering OnLine*, *13*, 18.

List of publications

Publications from thesis

Journal papers:

1. Kumar, M., Pachori, R.B., & Acharya, U.R. (2016). An efficient automated technique for CAD diagnosis using flexible analytic wavelet transform and entropy features extracted from HRV signals. *Expert Systems with Applications*, *63*, 165–172.

2. Kumar, M., Pachori, R.B., & Acharya, U.R. (2017). Characterization of coronary artery disease using flexible analytic wavelet transform applied on ECG signals. *Biomedical Signal Processing and Control*, *31*, 301–308.

3. Kumar, M., Pachori, R.B., & Acharya, U.R. (2017). Automated diagnosis of myocardial infarction ECG signals using sample entropy in flexible analytic wavelet transform framework. *Entropy*, *19*, 488.

4. Kumar, M., Pachori, R.B., & Acharya, U.R. (2017). Use of accumulated entropies for automated detection of congestive heart failure in flexible analytic wavelet transform framework based on short-term HRV signals. *Entropy*, *19*, 92.

5. Kumar, M., Pachori, R.B., & Acharya, U.R. (2018). Automated diagnosis of atrial fibrillation ECG signals using entropy features extracted from flexible analytic wavelet transform. *Biocybernetics and Biomedical Engineering*, *38*, 564–573.

Publications apart from thesis

Journal papers:

1. Sood, S., Kumar, M., Pachori, R.B., & Acharya, U.R. (2016). Application of empirical mode decomposition-based features for analysis of normal and

CAD heart rate signals. *Journal of Mechanics in Medicine and Biology*, *16*, 1640002.

2. Pachori, R.B., Kumar, M., Kora, S., Avinash, P., & Acharya, U.R. (2016). An improved online paradigm for screening of diabetic patients using RR interval signals, *Journal of Mechanics in Medicine and Biology*, *16*, 1640003.

3. Sharma, R., Kumar, M., Pachori, R.B., & Acharya, U.R. (2017). Decision support system for focal EEG signals using tunable-Q wavelet transform. *Journal of Computational Science*, *20*, 52–60.

4. Sharma, R.R., Kumar, M., & Pachori, R.B. (2019). Joint time–frequency domain–based CAD disease sensing system using ECG signals, *IEEE Sensors Journal*, DOI: 10.1109/JSEN.2019.2894706.

Conference papers:

1. Sharma, R.R., Kumar, M., & Pachori, R.B. (2017). Automated CAD identification system using time-frequency representation based on eigenvalue decomposition of ECG signals. *International Conference on Machine Intelligence and Signal Processing*, 22-24 December, Indore, India.