

Analysis and Classification of Vibroarthographic Signal Using Nonstationary Signal Processing Techniques

THESIS

Submitted in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

by

Nalband Saif Dilavar Shehnaz

Under the Supervision of

Dr. A Amalin Prince

Co - Supervisor

Dr. Anita Agrawal



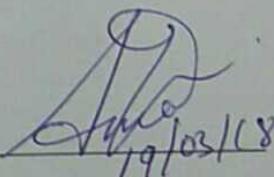
BITS Pilani
Pilani | Dubai | Goa | Hyderabad

**BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE
PILANI (RAJASTHAN) INDIA**

2018

Certificate

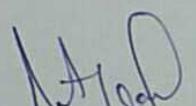
This is to certify that the thesis entitled, "Analysis and Classification of Vibroarthographic Signals Using Nonstationary Signal Processing Techniques" and submitted by SAIF NALBAND ID No. 2013PHXF0008G in partial fulfillment of the requirements of PhD. Thesis embodies the work done by him under my supervision.



19/03/18

Supervisor

DR. A. AMALIN PRINCE
Asst. Professor,
BITS-Pilani K.K. Birla Goa Campus
Date: 19/03/18



~~19/03/18~~

Co-Supervisor

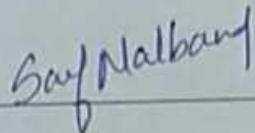
DR. ANITA AGRAWAL
Asst. Professor,
BITS-Pilani K.K. Birla Gos Campus
Date: 19/03/18

Declaration of Authorship

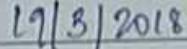
I, SAIF NALBAND, declare that this Ph.D thesis titled, 'Analysis and Classification of Vibroarthrographic Signals Using Nonstationary Signal Processing Techniques' and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:



Date:



“It does not matter how slowly you go as long as you do not stop. ”

Confucius

Abstract

The knee joint is one of the largest and complex joints in the human body. It can withstand a significant amount of strain and injury risks in routine and sports activities. A survey conducted by Indian orthopedic association (IOA) concluded that knee joint disorders are growing among Indians at an alarming rate. Though there are invasive and noninvasive techniques available for early diagnosis of knee joint disorders, these techniques are not feasible for a routine checkup. Interpretation of the vibration or the audio signals emitted from the patella femoral joint during the active knee joint movements is known as vibroarthrography (VAG). VAG signals are nonlinear, nonstationary and multicomponent in nature. VAG signals could lead to safe, cost effective, objective and noninvasive clinical tool for early detection, localization and quantification of the articular cartilage disorder.

The work carried out in this thesis focuses on discriminating between the healthy and unhealthy VAG signals by using different methodologies. The analysis has been carried out using different preprocessing techniques, feature extraction and selection algorithms and these are followed by pattern classification.

In this thesis, nonstationary linear signal processing technique namely wavelet packet decomposition (WPD) has been proposed for the analysis of VAG signals. The VAG signals are decomposed into subband signals using WPD. The VAG signals are reconstructed by identifying the significant subband signals. Entropy based features and recurrence quantification analysis (RQA) parameters are extracted from subband signals and reconstructed VAG signals. Entropy based features namely approximate entropy (ApEn), sample entropy (SampEn), Shannon entropy (ShEn), Rényi entropy (ReEn), Tsallis entropy (TsEn) and permutation entropy (PeEn) are studied as feature extraction techniques. RQA based parameters consisting of recurrence rate (RR), Determinism (DET), Entropy (ENTR) and averaged diagonal line length (DLL) are extracted as features. A total of 60 features were extracted which consisted of 50 features from subband signals and ten from reconstructed VAG signals. In this study, least square support vector machine (LS-SVM) is used as a classifier for pattern classification. The feature vector of 10 features extracted from reconstructed VAG signals provided better classification accuracy of 84.27%.

To identify the most significant features and to remove the redundant features, feature selection algorithm has been used. Feature selection algorithm namely genetic algorithm (GA) and apriori algorithm are studied for computing the optimal feature set. The classification performance was compared using LS-SVM. The feature vector of 60 features extracted from subband signals and reconstructed VAG signals gave a classification accuracy of 69.66%. Using feature selection algorithms, eight features were selected from GA which gave a classification accuracy of 82.02%. The highest classification accuracy of 85.39% was obtained using five features selected by the

a priori algorithm. Results also illustrated that entropy based feature performed better with respect to RQA parameters.

The nonlinear nonstationary signal processing technique namely complete ensemble empirical mode decomposition with adaptive white noise (CEEMDAN) has been studied. It is based on empirical mode decomposition (EMD) and decomposes the VAG signals into intrinsic mode functions (IMFs). Six entropy based features have been extracted from the reconstructed VAG signals. A simple feature extraction technique based on central tendency measure (CTM) has also been computed. The pattern classification using LS-SVM illustrated that the feature vector of entropy based features gave a classification accuracy of 84.27%. PeEn performed better with respect to other entropies by giving classification accuracy of 86.52% but highest classification accuracy was obtained by CTM of 87.64%. The computed time complexity of PeEn and CTM having an order of $O(N)$ provided a simple, robust and low computational feature extraction technique along with better classification accuracy.

Time-frequency analysis of VAG signal has been studied for the diagnosing knee-joint disorders. Smoothed pseudo Wigner–Ville distribution (SPWVD) and Hilbert-Huang transform (HHT) has been used for time frequency analysis of VAG signals. HHT consists of EMD for computing intrinsic mode functions (IMFs) and Hilbert transform (HT). But in this study, CEEMDAN has been used for computing IMFs. The time-frequency representation of the proposed method is considered as a time-frequency image. Statistical features such as mean, standard deviation, skewness and kurtosis are extracted from the time-frequency images. Pattern classification has been carried out by LS-SVM. CEEMDAN-HHT performed better with respect to SPWVD by giving the highest accuracy of 88.76%. The proposed method based on time-frequency technique provides an effective tool for the analysis of VAG signals.

In this thesis, a comparative analysis of VAG signals has been carried out using different methodologies. These methodologies consist of nonstationary signal processing, feature extraction techniques and pattern classification using machine learning technique. This analysis has been carried out with an objective to develop a computer aided diagnostic system for noninvasive detection of knee joint disorders.

Acknowledgements

It is my privilege to take this opportunity to humbly state my gratitude to those people whom I met and interacted with in the course of my Ph.D. thesis. First of all, I would like to sincerely thank my supervisor Dr. A. Amalin Prince (HoD, EEE Department, BITS, Pilani - K. K. Birla Goa Campus) for his support, and most importantly, patience, with all aspects of my Ph.D. thesis. For this, I would always be indebted. Dr. A Amalin Prince has always contributed insightful guidance, thoughtful discussions, and timely encouragement in the due course of my Ph.D. thesis. During my research work, he provided me with high flexibility in various scientific research and creative thinking. I am honoured to say thanks to my beloved supervisor Dr. A Amalin Prince for his firm supports around the clock.

I would like to thank my co-supervisor Dr. Anita Agrawal for constant encouragement. She provide a firm support and helped me in my initial stages of Ph.D journey. She had played a significant role in improving my scientific writing skills. I have been privilege to be as teaching assist under her course and this learning has been very valuable for me as a student.

My deepest gratitude to the members of the doctoral advisory committee, Prof. Raghurama Gunaje (Director, BITS, Pilani - K. K. Birla Goa Campus), and Prof. K R Anupama without whose knowledge and assistance this study would not have been successful. I also express my sincere thanks to Dr. Gautam Bacher, Convener, Departmental Research Committee for providing me with the facilities to conduct my research work at BITS, Pilani - K. K. Birla Goa Campus. I would also like to thank Dr. M.K.Deshmukh (Ex-HoD) and Nitin Sharma (Ex-Convener, Departmental Research Committee). I am also very thankful to my departmental faculty for their help and support. I would like to personally thank Dr. Gautam Bacher for his constant encouragement and support throughout my journey of thesis.

I am extremely grateful to Prof. Souvik Bhattacharyya (Vice Chancellor, BITS- Pilani), Prof. S. K. Verma (Dean, Academic Research, Ph. D. Programme, BITS, Pilani) and Prof. P. K. Das (Associate Dean, Academic Research, BITS, Pilani - K. K. Birla Goa Campus). I would like to thank past directors of BITS, Pilani - K. K. Birla Goa Campus, Prof. K. E Raman and Prof. Sasikumar Punnekatt. Collective and individual acknowledgements are due to all my colleagues and friends, within the BITS-Pilani community, who have directly or indirectly helped me in my work.

I would like to thank Prof. R.M.Rangayyan (Professor Emeritus of Electrical and Computer Engineering at the University of Calgary, Calgary, Alberta, Canada.) for helping, encouraging and providing us the data set. I would like to thank the reviewers of journals/ conference who have helped me to improve my research.

I would like to thank Mr. Vijaykumar Patil and “Reconfigurable Computing lab” for providing me the facilities. I would like to thank my lab colleagues Mr. Jay Kumar and Mr. Gibin George for their support. I would also like to thank Mr. Adtiya Sundar, Mr. Valliappan C.A and other students who have worked in the lab.

I also thank large number of persons and friends Mr. Praveen GB, Mr. Geetesh Mishra, Mr. Gautam Kumar, Mr. Chetan Bandodkar, Mrs. Pinky, my football buddies and research scholars for moral support. Special thanks to members of Goan Royal Riders (GRR) especially Raju Saini and Toastmaster International Margao for making my stay in Goa a memorial one.

I would forever be indebted to my parents, parents-in-law, sister and my cats for their constant love, confidence, prayers, strong support and good wishes. I would like to thank my wife for having patience, for her love, understanding and support throughout my Ph.D journey.

I thank to the Almighty Lord for all the wonderful things and surrounding life he is giving to me every moment and hope his blessing will be continue on my journey.

Nalband Saif Dilavar Shehnaz

Contents

Certificate	i
Declaration of Authorship	ii
Abstract	iv
Acknowledgements	vi
Contents	viii
List of Figures	xi
List of Tables	xiii
Abbreviations	xiv
1 Introduction	1
1.1 Background	1
1.1.1 Knee joint disorders in India	3
1.2 Motivation	4
1.3 Aim Of Thesis	5
1.4 Objective Of The Thesis	6
1.5 Organization Of Thesis	8
2 Literature Review	9
2.1 Introduction	9
2.2 VAG Acquisition Techniques & Datasets	10
2.3 Data Set Used In Thesis	11
2.3.1 Characteristics Of VAG Signal	13
2.4 Preprocessing Techniques Of VAG Signals	14
2.4.1 Linear Signal Preprocessing Techniques	14
2.4.1.1 Preprocessing VAG Signal Using Wavelet Packet Decomposition	16
2.4.2 Nonlinear Signal Processing Techniques	16
2.4.2.1 Preprocessing Technique Using CEEMDAN	18

2.4.3	Time-Frequency Analysis	19
2.4.3.1	Smoothed Pseudo Wigner–Ville Distribution (SPWVD) & Hilbert–Huang Transform (HHT)	21
2.5	Feature Extraction and Feature Selection Techniques	22
2.5.1	Feature Extraction	22
2.5.1.1	Proposed Feature Extraction Techniques	24
2.5.2	Feature Selection	25
2.5.2.1	Genetic & Apriori Algorithm For Feature Selection	26
2.6	Pattern Classification	27
2.6.0.1	Least Square Support Vector Machine For Pattern Classification	30
2.7	Gaps In Existing Literature	31
2.8	Conclusion	32
3	Analysis of VAG signals using Wavelet Packet Decomposition	33
3.1	Introduction	33
3.2	Wavelet Packet Decomposition (WPD)	34
3.3	Feature Extraction	36
3.3.1	Entropy Based Feature Extraction	36
3.3.1.1	Approximate Entropy	36
3.3.1.2	Sample Entropy	37
3.3.1.3	Shannon Entropy	38
3.3.1.4	Rényi Entropy	39
3.3.1.5	Tsallis Entropy	39
3.3.1.6	Permutation Entropy	39
3.3.2	Recurrence Quantification Analysis (RQA)	40
3.4	Results	42
3.4.1	Feature Extraction	44
3.4.2	Classification	51
3.5	Discussion	55
3.6	Conclusion	56
4	Analysis Of VAG Signals Using Feature Selection Algorithms	57
4.1	Introduction	58
4.2	Methodology	59
4.2.1	Data Set Preparation	59
4.2.2	Genetic Algorithm	59
4.2.3	Apriori Algorithm	61
4.3	Results & Discussion	62
4.3.1	Feature Selection	63
4.3.1.1	Genetic Algorithm	63
4.3.1.2	Apriori Algorithm	63
4.3.2	Classification	64
4.4	Conclusion	66
5	Analysis of VAG Signals Using Complete Ensemble Empirical Mode Decomposition With Adaptive Noise	68
5.1	Introduction	69
5.2	Methodology	70

5.2.1	CEEMDAN	70
5.2.2	Hilbert-Huang Transform	72
5.2.2.1	Hilbert Transform	72
5.2.3	Feature Extraction	73
5.2.3.1	Entropy Based Feature Extraction	73
5.2.3.2	Central Tendency Measurement (CTM)	73
5.3	Results	74
5.3.1	Feature Extraction	75
5.3.1.1	Entropy	75
5.3.1.2	CTM	77
5.3.2	Classification	79
5.4	Discussion	81
5.5	Conclusion	84
6	Time-Frequency Analysis of VAG signals	85
6.1	Introduction	86
6.2	Methods	87
6.2.1	Wavelet Packet Decomposition	87
6.2.2	SPWVD	87
6.2.3	Hilbert Huang Transform	88
6.2.4	Feature Extraction	88
6.3	Results	89
6.3.1	Preprocessing	89
6.3.2	Feature Extraction	90
6.3.3	Classification	93
6.4	Discussion	95
6.5	Conclusion	96
7	Conclusion and Future Scope	98
7.1	Outcome Of The Thesis	98
7.2	Contribution Of The Thesis	100
7.3	Future Scope	102
Bibliography		103
List Of Publications		122
Biography		123

List of Figures

1.1	Anatomy of Knee joint	1
1.2	Statistical Measures	3
1.3	Flow chart for diagnosis of knee joint disorders using VAG signals	5
1.4	Thesis work flow	7
2.1	VAG data acquisition system. Image taken from Wu Y. Knee joint vibroarthrographic signal processing and analysis. Springer; 2015 Jan 29 [6].	12
2.2	Sample raw VAG signal	13
2.3	Filtered VAG signal using double cascaded moving average filter	15
3.1	Flow chart for diagnosis of knee joint disorders using VAG signals	33
3.2	Wavelet decomposition	35
3.3	An example of recurrence plot for raw VAG signal	40
3.4	Wavelet decomposition of VAG signal for (a) healthy and (b)unhealthy subject .	42
3.5	Reconstructed VAG signals	43
3.6	An example of recurrence plot for reconstructed VAG signal	44
3.7	Boxplot of ApEn	47
3.8	Boxplot of SampEn	47
3.9	Boxplot of ShEn	48
3.10	Boxplot of ReEn	48
3.11	Boxplot of TsEn	49
3.12	Boxplot of PeEn	49
3.13	Boxplot of RQA-RR	50
3.14	Boxplot of RQA-DET	50
3.15	Boxplot of RQA-ENTR	51
3.16	Boxplot of RQA-DLL	51
3.17	ROC plot for extracted features obtained from subband signals	54
3.18	ROC plot for extracted features obtained from reconstructed signals	55
4.1	Flow chart for diagnosis of knee joint disorders using feature selection	57
4.2	ROC plot of features selected from FS	66
5.1	Flow chart for diagnosis of knee joint disorders using CEEMDAN	69
5.2	VAG signal of (a) healthy subject and (b) unhealthy subject	73
5.3	IMFs obtained from VAG signal of (a) healthy subject and (b) unhealthy subject	74
5.4	Reconstructed VAG signal of (a) healthy subject and (b) unhealthy subject . .	75
5.5	Boxplots of VAG signal of both healthy and unhealthy	76
5.6	Analytical representation of VAG signal (a) healthy subject (b) unhealthy subject and reconstructed VAG signal of (c) healthy subject and (d) unhealthy subject .	77

5.7	Analytical representation of IMF5, IMF6 & IMF7 for healthy and unhealthy subject	78
5.8	Boxplot for CTM	79
5.9	Performance of ROC of extracted features	81
6.1	Flow chart for time frequency analysis of knee joint disorders using VAG signals	85
6.2	VAG signal of (a) healthy and (b) unhealthy subject	89
6.3	Wavelet decomposition of VAG signal for (a) healthy and (b) unhealthy subject .	89
6.4	IMFs of VAG signals obtained using CEEMDAN for (a) healthy and (b) unhealthy subject	90
6.5	The reconstructed VAG signal obtained from SPWVD (a) healthy subject and (b) unhealthy subject. Time-frequency representation of c) healthy subject and d) unhealthy subject VAG signal	91
6.6	The reconstructed VAG signal obtained from CEEMDAN (a) healthy subject and (b) unhealthy subject. Time-frequency representation of c) healthy subject and d) unhealthy subject VAG signal	92
6.7	Box plots for features extracted from SPWVD and CEEMDAN-HHT	93
6.8	ROC plot for statistical features obtained from time frequency distribution	95

List of Tables

2.1	Linear signal preprocessing techniques	15
2.2	Nonlinear Signal Processing Techniques	18
2.3	Time Frequency Techniques	20
2.4	Feature Extraction	24
2.5	Feature Selection	26
2.6	Classifier	30
3.1	Decomposition of the VAG Signal into Sub-bands	36
3.2	Statistical measures of entropy and RQA based features extracted from reconstructed VAG signals	44
3.3	Statistical measures of reconstructed VAG signals for RQA parameters	45
3.4	Statistical measures of six entropies extracted from subband signals	46
3.5	Classification performance of extracted features from subband signals	52
3.6	Classification performance of extracted features from reconstructed VAG signals	53
4.1	Comparison of classification performance of FS algorithm	65
5.1	Statistical measures of reconstructed VAG signals for entropy based features . . .	76
5.2	Statistical measure for CTM	79
5.3	Classification performance of extracted features from reconstructed VAG signal .	80
5.4	Complexity of the extracted features	82
5.5	Comparison of the proposed methodology with the existing non-linear studies .	83
6.1	Statistical measures of extracted features using SPWVD and CEEMDAN-HHT .	93
6.2	Classification of features extracted using SPWVD by LS-SVM	94
6.3	Classification of features extracted using CEEMDAN-HHT by LS-SVM	94
6.4	Comparison of the proposed methodology with the existing TFD	96
7.1	Summary of thesis methodologies	99
7.2	Summary of results from the proposed work in thesis	100

Abbreviations

ACC	Accuracy
ApA	Apriori Algorithm
ApEn	Approximate Entropy
CAD	Computer Aided Diagnostic
CEEMDAN	Complete Ensemble Empirical Mode Decomposition with added noise
CTM	Central Tendency Measure
ECG	Electrocardiography
EEG	Electroencephalogram
EEMD	Ensemble Empirical Mode Decomposition
Entr	Entropy
EMD	Empirical Mode Decomposition
EMG	Electromyography
FDR	False Discovery Detection
FS	Feature Selection
GA	Genetic Algorithm
HHT	Hilbert Huang Transform
IMF	Intrinsic Mode Functions
LS-SVM	Least Square Support Vector Machine
MCC	Matthews Correlation Coefficient
NPV	Negative Predictive Value
OA	osteoarthritis
PeEn	Permutation Entropy
PPV	Positive Predictive Value
ReCon	Reconstructed
ReEn	Rényi Entropy

AUC-ROC	Area Under The Receiver Operating Characteristic
RQA	Recurrence Quantification Analysis
SampEn	Sample Entropy
SB	subband
SEN	Sensitivity
ShEn	Shannon Entropy
SPE	Specificity
SPWVD	Smooth Wigner Ville Distribution
TFD	Time Frequency Distribution
TsEn	Tsallis Entropy
VAG	vibroarthrography
WPD	Wavelet Packet Decomposition
WT	Wavelet Transform

To my parents.

Chapter 1

Introduction

The introduction of many minds into many fields of learning along a broad spectrum keeps alive questions about the accessibility, if not the unity, of knowledge

Edward Levi

1.1 Background

The knee joint of the human body, also known as the tibiofemoral joint, is one of the largest and most complex joint structure. The knee joint comprises of bones (fibula, femur, tibia and patella), tendons, cartilages and a joint capsule (figure 1.1). The knee is a synovial hinge joint.



FIGURE 1.1: Anatomy of Knee joint

The knee cap which is also known as patella covers and protects the articular surface of knee joint. The joint capsule provides lubrication and strength to the surrounding of the knee joint. Ligaments surrounding the joint capsule helps to provide a stable structure and align the bones correctly. The patella is supported by the patellar ligament which extends from the patella to the tibial tuberosity of the tibia. There are two external ligaments known as the medial collateral ligament (MCL) and the lateral collateral ligament (LCL). MCL connects the medial side of the femur to tibia while LCL connects the lateral side of the femur to fibula. The primary function of these ligaments is to resist the valgus and varus force on the knee joint. The internal ligaments known as the anterior and posterior cruciate ligaments help to maintain a stable alignment of the knee.

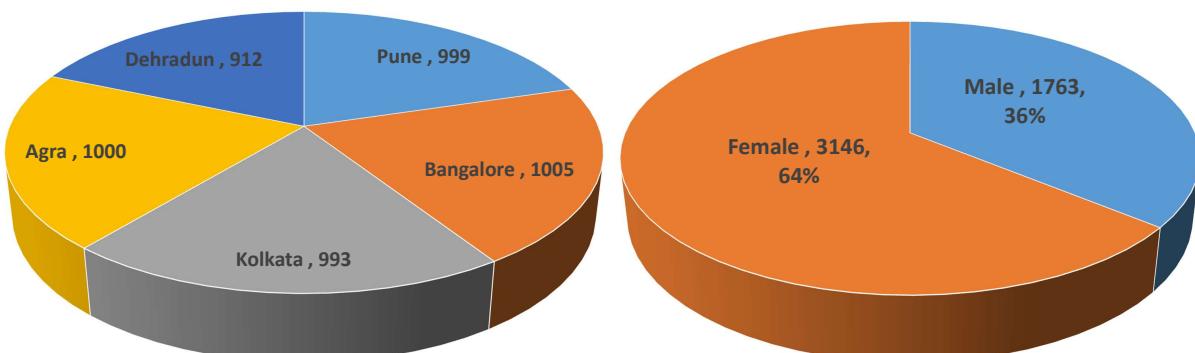
Apart from joint capsule and ligaments, several other structures surrounding the knee joints are present which provide a cushion and protects the knee joints. Synovial fluid is found in the cavity of synovial joint. The synovial fluid is responsible for reducing the friction between the patella and femur. The articular fat pads tissues present near the knee joint act as a cushion to knee joint from the external stress. The infrapatellar fat pad is the largest of the pads surrounding the knee joint. It provides flexible movements of extension and flexion during the leg movements. It also acts as a shock absorber to the anterior surface of the knee. These tissues, ligaments and bones provide a stable structure and hold the bones in correct posture and alignment. Hence, the knee joint is able to withstand human body weight while allowing the movements of the leg relative to the thigh. The knee joint can resist a significant amount of strain and stress. Consequently, the knee joints are prone to injury risks in routine life activities and especially in sports activities.

Arthritis is one of the most common diseases that occur in the knee joints. It can be broadly classified into osteoarthritis (OA) and inflammatory arthritis. OA is caused mainly due to degeneration of protective cartilage or due to wear and tear of the muscles. Therefore, movements of the leg become quite restricted since the bones of the knee joint brush against each other, thereby causing severe pain to the subject, which varies among the subjects. But in the cases related to inflammatory arthritis, the immune system of the body fails to protect tissues, muscles ligaments etc. from infection. This leads to uncontrollable inflammation in joints and damaging tissue, muscle. Rheumatoid arthritis, ankylosing spondylitis, juvenile idiopathic arthritis and psoriatic arthritis are examples of inflammatory arthritis. There are other factors involved in knee joint disorders which are aging, obesity, physical activities, previous knee injury, and regular climbing of stairs [1].

1.1.1 Knee joint disorders in India

OA is the most common knee joint disorder detected among Indians. By projecting its numbers at the present rate, India may become the OA hub with more than 65 million cases by 2025 according to studies contacted by medical professionals [2]. “In the last few decades, Indians in the age-group of 30 to 50 years are falling prey to OA and it continues to have a serious impact on the lives of elderly people” said Mudit Khanna, consultant Orthopedics at Wockhardt Hospital.

A recent study concluded by Prognosis management and research consultants Pvt. Ltd in collaboration with the Indian orthopedic association conference (IOACON) Agra, made a critical study on OA. The survey was conducted among the Indian population and the factors associated with it [2]. The survey was carried out in 5 major cities across India, namely, Pune, Bangalore, Kolkata, Agra and Dehradun. Figure 1.2 (a) shows a pie chart diagram of the number of subjects participated and the percentage of subjects having OA. A total of 4909 subjects had participated in the study.



(c) Age distribution of subject having OA

FIGURE 1.2: Statistical Measures

The survey was also carried out on the basis of gender. The study inferred that the occurrence of OA was quite prevalent in cases of female subjects in comparison to male subjects. From the figure 1.2 (b), it can be observed that the occurrence of OA among females is quite higher than males. The figure elaborates that among 4909 subjects 64% of subjects were female while 36% were male. A study was also carried out with subjects consisting a wide range of ages having knee joint disorders. This can be observed in figure 1.2 (c), the bar chart provides the description of the subjects with the different range of ages having OA. Thereby, these studies provide motivation for the prognosis of knee joint disorders.

1.2 Motivation

Procedures based on invasive and noninvasive techniques are available in detecting the deterioration of knee joint disorders. A minimally invasive surgical procedure such as arthroscopy is generally preferred for diagnosis of knee joint disorder. Arthroscopy is a standard procedure which provides a comparatively low-risk estimation of knee joint disorders. However, it has a limitation that it cannot be applied to the subject's knees, which is in the highly degenerated state due to arthritis ligament instability. Such procedure cannot be applied for a routine checkup since repetitive surgical invasions could be susceptible to bacterial infection.

Noninvasive techniques such as computed tomography (CT) and magnetic resonance imaging (MRI) are some of the popular procedures in asserting knee joint disorders [3]. Though MRI and CT provide a good image contrast, it only displays anatomically. Therefore, these procedures are highly observer dependent and this limits the diagnosis of knee-joint pathology. The cost, availability and portability are some of the other major limitations of these procedures. Therefore, these procedures cannot be used in the routine investigation or monitoring especially cases related to rehabilitation programs.

Furthermore, the above mentioned procedures are only able to provide a static information of knee joint disorders. Therefore, these techniques fail to provide dynamic aspects related to knee joint movements [4]. These limitations and drawbacks of semi-invasive and noninvasive procedures have motivated researchers for an alternative solution which would be noninvasive, cost effective and configurable with computer-aided diagnosis (CAD). This could provide an early diagnostic system for knee joint disorders. The auditory and deceleration/acceleration signals from knee joints offer researchers noninvasive technique for diagnosis of knee joint disorders [5]. These signals are nonlinear and nonstationary in nature and are known as vibroarthrographic signals (VAG). VAG signals could provide an assessable procedure for the diagnosis of knee joint disorders since VAG signals of a pathological condition of knee joint are highly complex waveform and large variation of frequencies [6]. These VAG signals contain pathological aspects of knee joint disorders and could define the characteristic of VAG signal. Thus, VAG signals along with

signal processing techniques could provide safe, cost effective, objective and noninvasive clinical application for early diagnosis of knee joint disorders, thereby, reducing the surgery cost and it would enhance the rehabilitation programs for monitoring the progress.

1.3 Aim Of Thesis

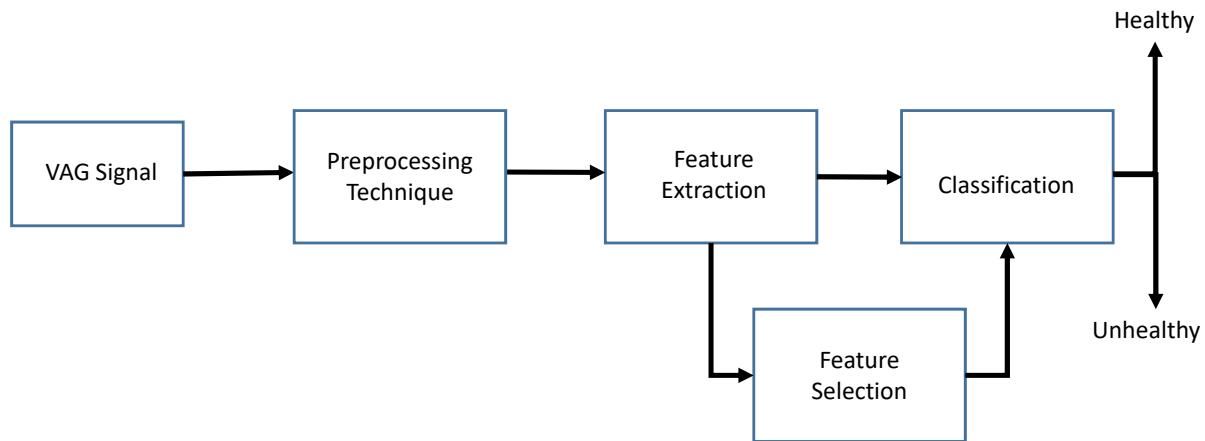


FIGURE 1.3: Flow chart for diagnosis of knee joint disorders using VAG signals

A typical CAD system involved in the diagnosis of knee joint disorder using VAG signal has been presented in figure 1.3. These are the basic steps of pattern classification involved in the analysis of signal or image processing [7].

- **Data acquisition:** Initially, the VAG signals are acquired from the subjects and the signals are amplified.
- **Preprocessing Technique:** The next step involves the preprocessing, in which the signals are prepared for the analysis. In this step, different signal processing techniques are applied for removing any external noises, distortion etc.
- **Feature extraction:** Feature extraction is followed after preprocessing. The extracted features gives the characteristic and the behaviour of the signals.
- **Classification:** Last step is pattern classification using machine learning algorithms. The extracted features are given as input to the classifier algorithm in order to compute the performance of the proposed methodology.
- **Feature selection:** The intermediate step between classification and features extraction is feature selection. This step is involved whenever a larger number of features have been extracted. In such cases feature selection is required to select most significant and relevant features.

The main aim of the thesis are

- To develop methodologies for computer aided diagnostic (CAD) system for early diagnosis of knee joint disorders.
- Methodologies are developed for the analysis of VAG signals using nonstationary signal processing techniques, feature extraction techniques and pattern classification.

1.4 Objective Of The Thesis

The aim of the thesis is achieved with specific objectives. These objectives are as follows:

- 1: *To analyse the classification performance of VAG signals using nonstationary linear signal processing technique and to extract entropy and recurrence quantification analysis(RQA) based features.*

This objective is achieved by decomposing the VAG signal using wavelet packet decomposition (WPD) into subband signals. The subband signals consist of different frequencies. The significant subband signals of particular frequencies are identified and VAG signals are reconstructed. Features based on entropies and RQA parameters are extracted from subband and reconstructed VAG signals. Entropies namely approximate entropy (ApEn), sample entropy (SampEn), Shannon entropy (ShEn), Rényi entropy (ReEn), Tsallis entropy (TsEn) and Permutation entropy (PeEn) are extracted. RQA based parameters namely recurrence rate (RR), Determinism (DET), Entropy (ENTR) and averaged diagonal line length (DLL) are extracted in this study. The pattern classification is carried out using LS-SVM.

- 2: *To identify the most significant and relevant features using feature selection algorithms for building effective classification model.*

This objective is achieved by using feature selection algorithms for choosing the most significant and relevant features and discarding the irrelevant features. In this work, genetic algorithm (GA) and apriori algorithm (ApA) have been utilized as feature selection algorithms and their performance have been compared.

- 3: *To analyse the classification performance of VAG signals using nonstationary nonlinear signal processing techniques and to extract entropy based features and simple feature extraction technique based on central tendency measures (CTM).*

This study proposes to use a nonstationary nonlinear signal prepossessing technique known as complete ensemble empirical mode decomposition with adaptive white noise (CEEMDAN) to decompose the VAG signals into intrinsic mode functions (IMFs). Approximate entropy

(ApEn), sample entropy (SampEn), Shannon entropy (ShEn), Rényi entropy (ReEn), Tsallis entropy (TsEn) and permutation entropy (PeEn) along with simple feature extraction technique based on central tendency measures (CTM) are explored.

- 4: To analyse VAG signals by time-frequency distribution (TFD) using smoothed pseudo Wigner–Ville distribution (SPWVD) and CEEMDAN-Hilbert-Huang transform (CEEMDAN-HHT) and to extract statistical features.

The objective is achieved by analysing VAG signals using time-frequency analysis namely SPWVD and CEEMDAN-HHT for diagnosing knee-joint disorders. The time-frequency representation of the proposed methods is considered as a time-frequency image. Statistical features namely mean, standard deviation, skewness and kurtosis are extracted from the time-frequency images. A pattern classification is performed to compare the effectiveness of the proposed methods.

This thesis can be summarized in the flow diagram as shown in the figure 1.4.

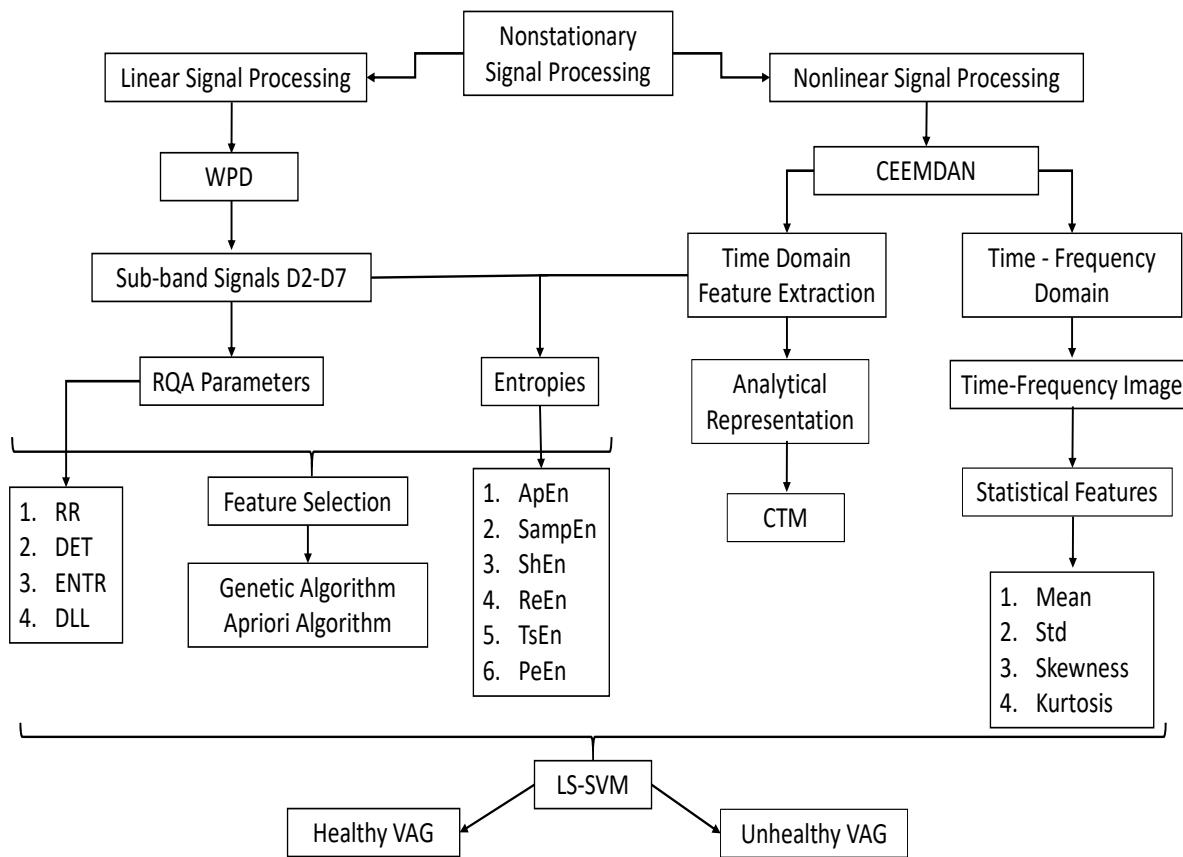


FIGURE 1.4: Thesis work flow

1.5 Organization Of Thesis

The thesis comprises of seven chapters and contents of each chapters has been briefly described below:

- Chapter 1 is an introductory chapter that briefly describes the problem domain. It briefly introduces the motivation for studying the knee joint disorders. It also gives an outline of various existing diagnostic methods. This chapter introduces how the computer aided diagnostic system would be developed for a noninvasive and cost effective method for diagnosing knee joint disorders using VAG signals.
- Chapter 2 gives literature review on VAG signals that sustain the significance of the work. It gives literature review on the various methodologies applied by the researchers. The literature review consists of different signal processing techniques, feature extraction techniques and feature selection techniques. Dataset used in this research has also been elaborated in this chapter. It also discusses the gaps in the existing research.
- Chapter 3 discusses the nonstationary linear technique using WPD. The VAG signals are decomposed into subband signals of particular frequencies. Entropy and RQA based features are discussed in this chapter. A pattern classification is carried out using LS-SVM.
- Chapter 4 introduces the aspects of features selection techniques. These techniques are used for selecting the most significant and relevant features obtained using the techniques in chapter 3. Genetic algorithm and apriori algorithm as feature selection algorithm are explored. The performance of features selected by the feature selection algorithms is compared.
- Chapter 5 discusses the nonstationary and nonlinear signal processing technique such as CEEMDAN. It briefly explains the advantages of CEEMDAN over EMD based techniques. Entropy based feature extraction techniques are carried out followed by pattern classification. A simpler feature extraction based on central tendency measurement is also explored.
- Chapter 6 gives the general discussion on time-frequency distribution. It introduces the methodology of SPWVD and CEEMDAN-HHT. The statistical features such as mean, standard deviation, skewness and kurtosis obtained from time-frequency representation are briefly explained. A comparative study between SPWVD and CEEMDAN-HHT is also carried out.
- Chapter 7 finally, concludes the thesis and provides suggestions for future research.

Chapter 2

Literature Review

Literature review is a process of finding the sources of relevant material for particular topic or subjects

Galvan (2005)

In this chapter, an extensive literature survey has been carried out on VAG signals. Methodologies consisting of data acquisition of VAG signals, signal processing techniques, feature extraction, feature selection and pattern classification has been covered. The data acquisition carried out for acquiring VAG signal by different researcher groups are discussed and data set selected in this thesis has been elaborated. This chapter also presents an outline of various signal processing techniques applied to VAG signals by researchers. These techniques are based on linear and nonlinear techniques. Different feature extraction techniques studied by researchers have been reviewed. The significance of feature selection algorithms in the analysis of VAG signals has been covered. This is followed by a review on pattern classification of VAG signals using machine learning techniques. Gaps in the existing literature have been identified and presented in this chapter. A review of the proposed methodology has been covered.

2.1 Introduction

The knee joint is one of the most complex, strongest and most important joints in the human body. It can withstand the weight of human body while allowing lower leg movements relative to the thigh. The knee joint can resist a significant strain and injury risks in routine activities and sports activities. Therefore, there is degradation of muscles, tissues and cartilage surrounding the knee joint. Other factors which cause degradation are ageing, obesity and physical activity. Thus, there is a need to interpret these deteriorations of the knee joint during the early stages.

As discussed in the previous chapter, there are existing invasive and noninvasive techniques available for early diagnosis of knee joint disorders. Noninvasive techniques include X-rays, CT scans, MRI etc. are available, but these are notably expensive for a routine check-up and it only provides static information about knee joint disorders. Therefore, a noninvasive, cost-effective and portable method for early diagnosis of knee joint disorders using VAG signals has been explored.

VAG signals are audible sound or vibration emitted from knee joint when a piezoelectric sensor is placed on the patella. These sounds or vibration are emitted due to the friction created between the bones and the surrounding tissues and ligaments. The wear and tear of these tissues and ligaments are caused due to ageing, daily activities or sports activities. Therefore, the smooth movements are restricted due the friction created between tissues, bones and ligaments. These signals contain information that could be used to characterize certain pathological aspects of the knee joint. VAG signals along with computer aided diagnostic system could lead to a safe, objective and noninvasive clinical tool for early detection, localization, and quantification of the articular cartilage disorder [8]. Diagnosis carried out using VAG signals could reduce the need for diagnostic surgery and would enhance the rehabilitation program in monitoring knee joints.

In this section, a brief review on VAG signal processing techniques is presented. An extensive research has been carried out in the area of VAG signals as a potential tool for diagnosis of the knee joint disorders.

2.2 VAG Acquisition Techniques & Datasets

Researchers have used various sensors such as contact microphone, electro-stethoscope, and accelerometers to acquire VAG signals with a varying degrees of success. The first attempt to record the knee joint sound was by Erb in 1933. He placed a microphone on mid-patella and graphically recorded the knee joint sounds. Researchers have used highly sensitive microphone placed on mid patella to record the sounds originating from knee joints. A team led by Kim *et al.* has used an electro-stethoscope (Hanbyul Meditec, Korea) for acquiring knee joint sounds [9], [10]. The stethoscope was placed on the mid-patella considering the structure of knee and it also provided minimal muscle interferences. The acoustic sounds acquired from the sensor were characterized by acoustic parameters such as fundamental frequency, mean amplitude of the pitches, and jitter and shimmer of the signals. In this experiment, the sample size was restricted to 17 subjects (6 healthy and 11 unhealthy subjects). Unhealthy subjects suffered from knee joint disorder such as ruptured wounds of the meniscus and OA. Results concluded that sound parameters such as jitter, shimmer, and fundamental frequency were comparatively higher for unhealthy subjects with respect to healthy subjects.

Literature survey concluded that microphones cannot be considered as an effective sensor. Microphone-based studies continued to be limited by the shortcomings of all acoustic systems, chiefly the poor frequency response to low-frequency signals and distortion caused by background noise. Mollan and his co-workers designed various data acquisition system for acquiring the knee joint vibration signals [11]. Their studies concluded that accelerometers are better than acoustic transducers because of their frequency response and small size. The accelerometers which are mounted to the skin are sensitive to vibration at their point of application, and therefore the problem of background noise, which affects microphone-based systems, is minimised. Their small size and the fact that they can be securely fixed avoids skin friction noise. Cable noise is removed by the use of pre-amplification. The study concluded that VAG signals of unhealthy subject have major frequency components (300 to 600 Hz) as compared to signals of the healthy subject (25 to 320 Hz). Since then, many researchers have reported the feasibility and usefulness of this technique as a potential use for the noninvasive diagnosis of knee joint disorders.

A research team of Hongzhi Chen and Lik-Kwan Shark from applied digital signal and image processing research center (ADSIP), University of Central Lancashire, Preston, UK carried out preliminary studies on knee joints [12], [13]. Their study focused on the potential use of acoustic emission(AE) with a biomarker for the analysis of knee joint disorders. 34 healthy subjects and 19 unhealthy subjects having OA participated in this experiment. The study was carried to classify the knee joint disorders with respect to age and degeneration of knee joint [13].

Baczkowicz *et al.* used accelerometer to study age related quality of patellofemoral joint (PFJ) motion in the VAG [14], [15]. The experiment setup was carried out for 220 subjects and these subjects were distributed into 5 groups. Each group consisted of subjects belonging to a particular range of ages. The subjects were part of a prevention program of locomotor system dysfunctions at Institute of Physiotherapy, Opole University of Technology. Subjects with any knee joint disorders were excluded from the study. Baczkowicz have *et al.* also studied the impact of joint motion using VAG signal on subjects having knee joint disorders [14], [16], [17]. Subjects involved in this study included were suffering from OA, lateral patellar compression syndrome and chondromalacia. The results concluded that signals recorded from unhealthy subjects possessed significantly higher values of parameters (variance of mean squared,entropy,power spectral density bands) with respect to healthy subjects.

2.3 Data Set Used In Thesis

R.M. Rangayyan and his research team is one of the most prominent research team working in the area of VAG signals. Rangayyan *et al.* carried out the data acquisition of VAG signals using a miniature accelerometer (Model 3115A, Dytran Instruments, Inc., Chatsworth, CA, USA). The setup for data/signal acquisition of knee joints was developed at the University of Calgary,

Canada by research group led by Rangayyan [6], [18]. A block diagram of data acquisition setup is shown in figure 2.1. The experiment protocol was given ethical clearance by Conjoint Health

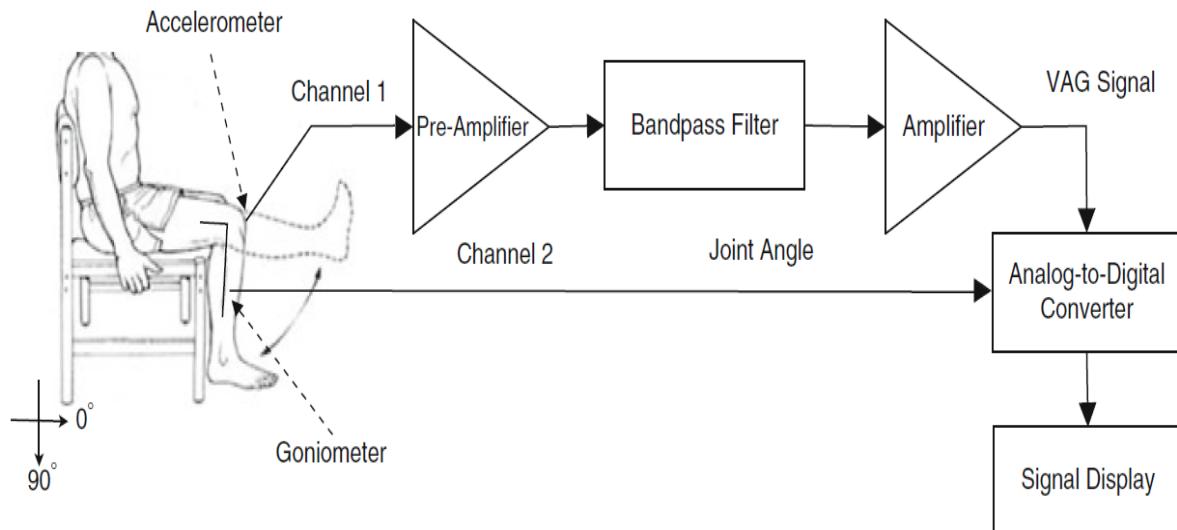


FIGURE 2.1: VAG data acquisition system. Image taken from Wu Y. Knee joint vibroarthrographic signal processing and analysis. Springer; 2015 Jan 29 [6].

Research Ethics Board of the University of Calgary. The experiment consisted of 89 subjects. The subjects were consulted and examined by a physician. 51 subjects (22 male, 29 female, age 28 ± 9.5 years) were considered as healthy since they were not affected by any abnormalities in knee joint pains or disorders. 38 subjects (20 male, 18 female, age 35 ± 13.8 years) suffered from the various knee joint disorders such as tibial chondromalacia, anterior cruciate ligament tear, meniscal tear etc. as reported by the physician. The protocol consisted of a simple experiment in which the subjects were told to sit on a rigid table and were instructed to float their legs freely. The position of the subject was fixed in such a manner that leg movements could be performed without any discomfort. At the mid-patella, a miniature accelerometer sensor was placed to acquire a signal from the knee joint. The physician instructed the subjects to flex and extend their leg for a duration of 4 secs. The electric voltage generated by the sensor were acquired in terms of acceleration and deceleration due to the knee joint movements of the leg. National Instrument's LabView software and data acquisition board were used for amplifying and digitizing the acquired VAG signal. The VAG signal was prefiltered (10 Hz to 1 kHz) and amplified. 2kHz was the sampling frequency of VAG signal and it was digitized with 12-bit resolution [18].

An example of an acquired raw VAG signal is shown in figure 2.2. The figure (a) and (b) are the VAG signals of a healthy and unhealthy subject respectively. As observed from figure 2.2 (a), the VAG signal of a healthy person produces little or no vibration since the supporting tissues and muscles of the knee joint are smooth and aren't degraded. Whereas in the case of an unhealthy person being affected by OA or other diseases, audible grinding sounds and vibration can be

observed in figure 2.2 (b). This may be due to loss of cartilage tissues and supporting ligaments of knee joints which restricts the smooth movements. Rangayyan *et al.* had concluded that sharp burst audible sound appeared in the range of 0-200Hz from subjects who had meniscal lesions. Subjects suffering from mild chondromalacia had a VAG signal in the frequency 0-300Hz while subjects suffering from severe chondromalacia had low frequency components (0-100 Hz) due to wear and tear of cartilage tissues [4].

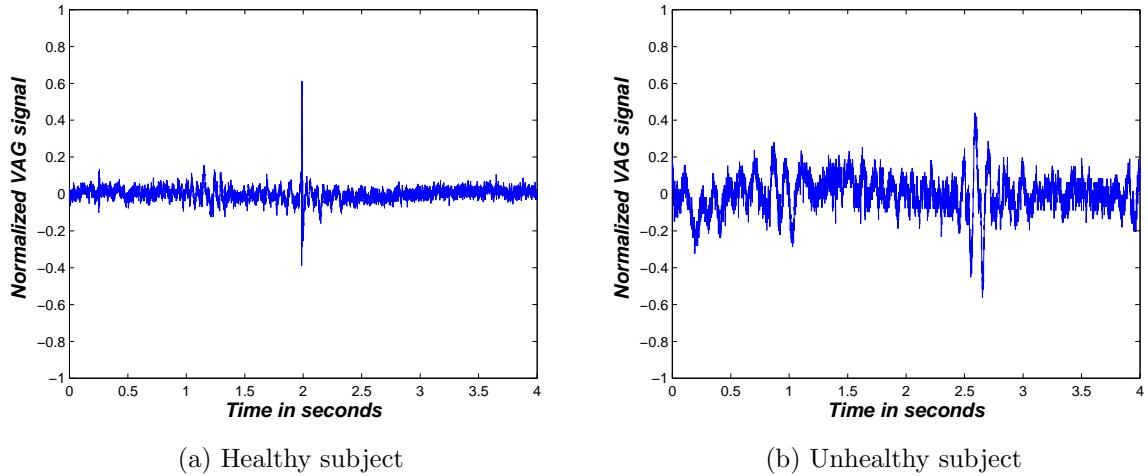


FIGURE 2.2: Sample raw VAG signal

In our study, the analysis of VAG signal has been carried out using the data set acquired by Rangayyan *et al.* Prior permission was obtained from Rangayyan for utilizing the VAG data set. This VAG data set was selected in our study since various studies have been carried using the same data set. Therefore, the performance of our proposed work with the existing methodology can be compared.

2.3.1 Characteristics Of VAG Signal

The characteristics of VAG signal are as follows:

- i. VAG signal is a multicomponent signal since it contains several vibrations or acoustic components. These vibrations occur due to friction produced between bones and surrounding tissues/cartilage of knee joint.
- ii. VAG signal is nonstationary in nature. This can be observed from figure 2.2; the acquired VAG signals of the healthy and unhealthy subject consists of different amplitudes and frequencies.
- iii. VAG signal is aperiodic and nonlinear.

Considering these characteristics of VAG signal, an effective advanced signal processing technique followed by pattern classification is required for developing a CAD system. Following section explores the existing signal processing techniques and the feature extraction methodology used in the analysis of VAG signals.

2.4 Preprocessing Techniques Of VAG Signals

The task of the signal preprocessing procedure is to remove the artefacts in the raw signal. Its main objective is to extract essential information about the signal and thereby making a proper inference about the behaviour of the signal. Biomedical signals are susceptible to contamination from artefacts such as baseline wander, external noise, muscle interferences etc. Therefore, preprocessing is an indispensable step in the analysis of biomedical signals. Since the VAG signals are acquired using an accelerometer, they too are contaminated by various artefacts. Therefore, it is important that these artefacts are eliminated prior any analysis since these artefacts could mislead the results of analysis [19].

Since VAG signals are nonlinear and nonstationary in nature, various researchers have used linear and nonlinear signal processing technique. Linear signal processing techniques assume that input signals are periodic and linear in nature. Therefore, the analysis carried out using linear signal processing techniques may result in loss of important information. Whereas, nonlinear methods do not assume the input data/signal as linear or stationary. The nonlinear methods are suitable for nonlinear and nonstationary signals. But, the nonlinear methods are computationally complex in nature. Researchers have extensively worked in the methodology related to the frequency domain, time domain, and time-frequency analysis.

Various researchers have used different methodologies for the analysis of VAG signals. These methodologies includes preprocessing techniques, feature extraction techniques and followed by pattern classification for differentiating between healthy and unhealthy VAG signals. The following section presents various linear and nonlinear signal processing techniques used on VAG signals. A literature review of the proposed signal processing techniques has been presented in the following section.

2.4.1 Linear Signal Preprocessing Techniques

An initial study on VAG signals was carried out using traditional linear signal processing techniques[20]. Zhang *et al.* had used an adaptive cancellation of muscle contraction interference and artefacts [21]. This technique was based on a two-stage least-mean-squares (LMS) adaptive filter. The external noise generated from the accelerometer and its environment were removed

TABLE 2.1: Linear signal preprocessing techniques

Year	Author	Techniques	Features	Classifier	Accuracy%
1994	Zhang[21]	two-stage LMS	-	-	-
1997	Krishnan[22]	adaptive segmentation using RLSL	AR coefficients, VMS	Logistic classification	68.9
2013	Wu[23]	adaptive time-delay neural filter	-	-	-
2012	Cai[24]	DCMA	-	-	-

in the first stage. In the second stage, the adaptive filter was able to remove the muscle interference from the signal. The work was further investigated by Krishnan *et al.* using adaptive segmentation based on the recursive least squares lattice (RLSL) algorithm [22]. In the segmentation process, the VAG signals were broken down into series of quasi-stationary or stationary segments. The segmented signals were assumed as stationary and thereby features were extracted from these segmented signals. Yunfeng Wu used an adaptive time-delay neural filter which estimated the noise present in the VAG signal by convolving neuron with synaptic weights over a reference input [23]. Suxian Cai *et al.* proposed a double cascaded moving average filter for removing artefacts from VAG signals [24]. The first layer of the filter has two moving average operators of the same order. The second layer smooths the coefficients obtained from the first layer. Figure 2.3 shows the filtered VAG signal using a double cascaded filter. In this thesis, a double cascaded moving average (DCMA) filter has been used for 89 VAG signals as a mandatory filtering step. Therefore, prior carrying out the analysis, the VAG signals are filtered using a double cascaded moving average filter. The summary of literature survey carried on VAG signals using linear signal processing techniques has been shown in the table 2.1.

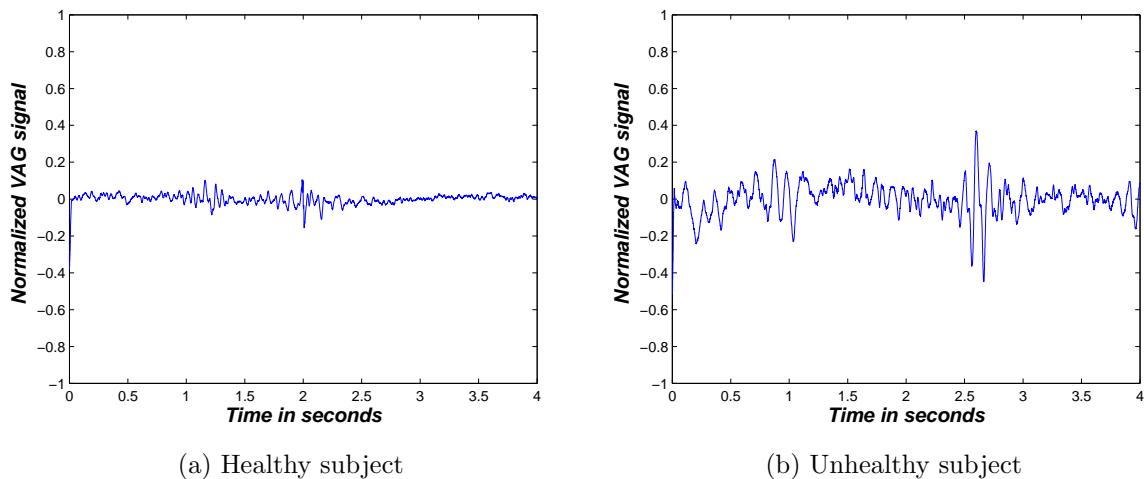


FIGURE 2.3: Filtered VAG signal using double cascaded moving average filter

Work carried out by above researchers assumed that the VAG signals are stationary and applied linear signal processing techniques. The analysis carried out using these techniques provided poor classification performance and may mislead the results. Therefore, nonstationary linear processing techniques for the analysis of VAG signals has not been extensively studied for the analysis of VAG signals.

2.4.1.1 Preprocessing VAG Signal Using Wavelet Packet Decomposition

Wavelet transform (WT) is a popular nonstationary signal processing techniques and has been widely used in research community. WT divides the input data or signals into different frequency components [25]. WT has been studied where input data is nonstationary, aperiodic and nonlinear. WT has been applied in the analysis of biomedical signals such as EEG, EMG etc. [26]–[31]. Recent studies have concluded that wavelet based decomposition of EEG signals into subband signals of different frequencies provided more information about the neuronal activities [32]. The study also stated that the characteristics were distinctly evident in the subband signals when compared to original EEG signals. Considering the similarity between EEG and VAG signals as nonstationary and nonlinear nature, a technique of wavelet packet decomposition (WPD) of VAG signal has been investigated.

Coifman *et al.* introduced WPD also known as wavelet packets or subband tree [33]. It is a wavelet transform in which the input signal is given to bank of filters. It disintegrates a given signal into detail approximation coefficients at the first stage of decomposition [34]. These approximation coefficients are further disintegrated into approximation coefficients in successive stages. Hence the signal breaks down like a tree structure. The WPD has been unexplored in the literature review for the analysis of VAG signal. Therefore, WPD of VAG signal in this dissertation has been studied. There are various wavelet functions available for selecting the mother wavelet function. We have selected the mother wavelet function based upon the correlation between VAG signals and coefficients of WPD.

2.4.2 Nonlinear Signal Processing Techniques

Since VAG signals are nonstationary and nonlinear by nature, nonlinear signal processing techniques are explored for the analysis. With the evolution of digital signal processing, various researchers have used different signal processing techniques for the analysis of VAG signals. Comon *et al.* introduced independent component analysis (ICA) for the analysis of complex signals [35]. ICA is based on a statistical approach that transforms multicomponent data into simpler components that are statistically not dependent on each other. An example of ICA is listening to one's conversation in a noisy environment, also known as “cocktail party problem”. In this method, the input signal/data are assumed as a linear or nonlinear combination of

variables. These variables are assumed as mutually independent and are nongaussian in nature and these variables are called as independent components of the signal/data. ICA is an extension of principal component analysis method [36]. ICA has been used in various application domain such as imaging processing, speech and audio signal processing, document databases and financial data for stock markets [37]. ICA has been extensively applied in the analysis of biomedical signal. The biomedical signals are acquired using contact sensors or multiple sensors and therefore, these acquired biomedical signals are contaminated by various artefacts. Hence ICA has been used for removing these artefacts present in biomedical signals [38], [39]. ICA has been used as preprocessing technique for removing the artefacts present in such as EEG [40]. Chen *et al.* applied ICA for the analysis of VAG signals [41]. VAG signals are contaminated with artefacts such as muscle interference, baseline wanders etc. analysis of contaminated VAG signals would be difficult. Therefore, ICA acts as preprocessing step in removing these artefacts and thereby improving the quality of VAG signals [41].

Variational mode decomposition (VMD) is another nonlinear signal processing method used as a preprocessing technique. VMD was proposed by Dragomiretskiy *et al.* as a novel technique for decomposing the signal into an ensemble of band-limited IMFs [42]. VMD is based on a nonrecursive model and modes are extracted concurrently. In this method, the input signal is reconstructed by building the model of an ensemble of modes with respective central frequencies. Sundar *et al.* used VMD for the denoising of VAG signals and followed by Wiener entropy for computing the correlation of signals [43].

Huang *et al.* had proposed a signal processing technique for nonstationary nonlinear signals known as the EMD [44]. EMD decomposes the signal into a number of intrinsic mode functions (IMFs). Wu *et al.* have used EMD by removing the baseline wander and other external noises [45]. IMFs signals are oscillations residing over the signals. It defines the only a single mode of oscillation and multi waves components are restricted. Wu *et al.* had used EMD for removing artefacts/other noises by using confidence limit criterion and probability density function technique. The significant IMFs were identified and the VAG signal was reconstructed. Turn count method was used to identify the fluctuating part of the signal. Results illustrated that quality of VAG signal improved considerably from artefacts [45]. However, EMD suffers from one prominent drawback of “mode mixing”. In order to overcome this drawback, Huang proposed a modified version of EMD called as “ensemble empirical mode decomposition” (EEMD) [46]. The analysis of VAG signal using EEMD was carried out by Wu *et al.* to remove the baseline wander and other external noises [47]. EEMD decomposes the VAG signals into a number of IMFs and detrended fluctuation analysis (DFA) algorithm is applied on each of the IMFs to calculate fractal scaling index [47]. DFA algorithm helps in inferring the anti-correlated and long-range correlation components. The identified IMFs helps in reconstructing the VAG signals [46]. The summary of literature survey carried on VAG signals using nonlinear signal processing techniques has been shown in the table 2.2.

TABLE 2.2: Nonlinear Signal Processing Techniques

Year	Author	Techniques	Features	Classifier	Accuracy%
2012	*Chen[41]	ICA	root mean square	-	-
2015	Sundar[43]	VMD	-	-	-
2013	Wu[45]	EMD	turn counts	-	-
2013	Wu[47]	EEMD	-	-	-
2016	Wu[48]	EEMD	Entropy, EAM	LS-SVM	83.56

*Different dataset

Therefore considering the drawbacks of EMD based methods, nonstationary nonlinear signal processing techniques using complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) could be used for the analysis of VAG signals.

2.4.2.1 Preprocessing Technique Using CEEMDAN

VAG signals are nonlinear and nonstationary in nature and therefore advance signal processing techniques are required for the analysis. EMD provides a vital signal processing techniques for nonlinear and nonstationary data analysis. EMD has found widespread usage in the application related to biomedical studies, geophysical studies [49], [50]. EMD was applied by Wu *et al.* for carrying out the analysis of VAG signals. But the frequent occurrence of mode mixing is one of the major flaws of EMD. In mode mixing, a particular IMF is composed of signals with widely disparate scales, or a signal identical scale may reside in different IMF components. Huang *et al.* applied intermittency test for removing the problem of mode mixing but wasn't successful [51]. Huang proposed a new noise-assisted data analysis (NADA) method known as EEMD for removing the problem of mode mixing [46]. In this method, a true IMF is defined as the mean of the ensemble of trails which consists of a signal and added Gaussian white noise. Wu *et al.* applied EEMD for the analysis of VAG signals [47] along with DFA algorithm. However, EEMD still contains the drawbacks such as

1. Presence of residual noise in the reconstructed signal and the modes
2. Different number of modes were obtained by realizing different input signal plus Gaussian white noise.

Hence, in order to remove these drawbacks complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) was proposed [52]. Considering the nature of VAG signal, this thesis proposes to use CEEMDAN in computing IMFs for the analysis of VAG signal [52]. CEEMDAN provides an improvement in minimizing the reconstructed error, provides a fixed number of modes for different signals and minimizes residual noise present in the modes [52].

CEEMDAN has been successfully reported in the analysis of nonstationary and nonlinear data, especially in cases of biomedical signals [53]–[55], vibrational signals from rolling element bearings [56], wind speed forecasting [57]. Hence, considering these studies, CEEMDAN as one of the preprocessing technique has been proposed for the analysis of VAG signals in this thesis.

2.4.3 Time-Frequency Analysis

Time frequency analysis of the input data/signal is represented in the frequency and time domain. It is usually considered for a signal which possesses time-varying frequency content. Hence interpretation of these signal in time-frequency domain allows intermittent, local, transient or multicomponent of signals to be interpreted clearly [58]. Therefore these signal are represented by time-frequency distribution (TFD). TFD infers the energy distribution of the signal in time and frequency domain. Hence extracting features from TFD could provide distinct properties or characteristics of the data/signal in the time-frequency domain. Various time-frequency techniques are available such as short-time Fourier transform (STFT) or Gabor transform (GT), the continuous Morlet wavelet transforms (CMWT), and the Hilbert transform (HT) [58]. Time-frequency analysis is one of the most active research fields in digital signal processing. Analysis of biomedical signal using time-frequency techniques is helpful in understanding the physiological aspects of the body [59], [60]. Various researchers have used time-frequency analysis for VAG signals.

Sridhar Krishnan *et al.* used wavelet matching pursuit (MP) algorithm [34], [61] for time-frequency analysis of VAG signals [62]. Features such as instantaneous energy, instantaneous energy spread, instantaneous mean frequency and instantaneous mean frequency spread were extracted from TFD. This methodology was carried out on 37 subjects (19 healthy and 18 unhealthy) and did not require any prior knowledge of any degeneration of articular cartilage pathology. Result from MP gave an accuracy of 77.8% [62]. Krishnan *et al.* have used traditional short-time Fourier transform (STFT) for time-frequency analysis of VAG signals [63], [64]. STFT provided the frequency distribution of VAG signal with time domain. Fast Fourier transform (FFT) with Hamming window of 128 samples was used to compute the spectrogram of VAG signals. Due to the low resolution of STFT, the variation of VAG signal could not be observed adequately. Krishnan *et al.* further continued the work in the analysis of VAG signals by using adaptive time-frequency analysis [63]. Minimum cross-entropy optimization was carried out on wavelet MP algorithm and Wigner-Ville distribution for obtaining TFD of VAG signals [34], [61]. Features interpreting the dynamics of the VAG signals such as energy, energy spread, frequency, and frequency spread were extracted. This was followed by pattern classification which gave an accuracy of 68.9% [63]. Umapathy and Krishna further carried out the analysis of VAG signal [65]. In this study, wavelet packet decomposition and local discriminant bases algorithm were

TABLE 2.3: Time Frequency Techniques

Year	Author	Techniques	Features	Classifier	Accuracy%
1997	Krishnan[62]	wavelet matching pursuit	energy, frequency	discriminant analysis	77.8
2000	Krishnan[63]	adaptive matching pursuit	energy, frequency	LRA	68.90
2006	Umapthy[65]	modified local discriminant	normalized node energy, correlation coef.	LDA	79.8
2009	Wu[66]	wavelet MP	MP atoms TC	LS-SVM	73.03
2009	*Kim[9]	dynamic time warping	energy and frequency	NN	91.4
2013	*Chen[67]	HHT	-	NN	85.3
2015	*Backowiz[15]	STFT	-	-	-

* Different dataset

used for the analysis of VAG signals. The performance of this methodology gave an accuracy of 79.8% with linear-discriminant-analysis (LDA) as classifier [65].

Wu *et al.* further carried out investigation by comparing features from TFD obtained from wavelet MP decomposition with respect to time domain features [66]. Features such as the number of wavelet MP atoms and significant turns were extracted and pattern classification of 73.03% was obtained using LS-SVM [66]. Kim *et al.* (different dataset) had segmented the VAG signals using dynamic time warping (DTW) for TFD [9]. Features such as energy, energy spread, frequency and frequency spread were extracted from 1408 segments (healthy 1031, unhealthy 377) [9] and classification accuracy of 91.4% was obtained using a back-propagation neural network (BPNN). Chen *et al.* used nonlinear signal processing techniques such as Hilbert-Huang transform (HHT) for the analysis of VAG signals using different dataset [67]. The resultant gave a classification accuracy of 85.3%. The time-frequency analysis related to age-specific degradation of knee joint disorders was carried out by Baczkowicz *et al.* [15]. The discrete Fourier transform was applied to segmented VAG signals to obtain short time spectrum. The summary of literature survey carried on VAG signals using time frequency techniques has been shown in the table 2.3.

From table 2.3 it can be observed that limited studies has been carried out in the analysis of VAG signals using time frequency distribution. Therefore, the time frequency analysis can be further investigated using advanced time-frequency techniques namely smoothed pseudo Wigner–Ville distribution (SPWVD) and Hilbert-Huang transform (HHT) for the analysis of VAG signals.

2.4.3.1 Smoothed Pseudo Wigner–Ville Distribution (SPWVD) & Hilbert-Huang Transform (HHT)

Time-frequency distribution provides wider aspects in the analysis of the biomedical signal which are nonlinear and nonstationary in nature. Considering the characteristics of VAG signals, nonstationary signal processing techniques has been carried out. This thesis compares the performance of smoothed pseudo Wigner–Ville distribution (SPWVD) [68] and Hilbert-Huang transform (HHT)[44] for the time-frequency analysis of VAG signals. Two approaches based on linear and nonlinear signal processing techniques has been proposed. In this work, we propose to use nonstationary linear signal processing technique namely WPD and nonstationary nonlinear signal processing technique namely CEEMDAN as preprocessing techniques.

In first method, SPWVD has been proposed for obtaining the time-frequency distribution (TFD). SPWVD overcomes the problem of interference of “cross terms” in Wigner-Ville distribution. This work proposes to apply WPD to decompose the VAG signal into subband signals of different frequencies. The dominant subband are identified and the VAG signal is reconstructed. Hence, time-frequency representation of reconstructed VAG is obtained by using SPWVD and is considered as the time-frequency image. The statistical features such as mean, standard deviation, skewness, and kurtosis are extracted. SPWVD has been successfully studied for the analysis of biomedical signals [69], wind turbine fault diagnostic [70] acoustic and vibration signals in detection of gear failures [71].

Moreover, considering the nonlinear, nonstationary and multicomponent nature of VAG signal, a modified version of HHT known as CEEMDAN-HHT has been proposed. Traditional HHT consist of empirical mode decomposition (EMD) [44] and Hilbert transform (HT). EMD decomposes the signal into a number of IMFs. But this thesis proposes to use CEEMDAN for obtaining IMFs [52]. The VAG signal is reconstructed by dominant IMFs and the time-frequency representation is obtained by applying HT on reconstructed VAG signals. This time-frequency representation is considered as time-frequency image and statistical features consisting of mean, standard deviation, skewness, and kurtosis are extracted from time-frequency image. HHT has been successfully applied for the application related to geophysical studies [49], biomedical studies, vibration fault tolerance etc. [50], [72], [73].

Therefore, a comparative study of different methodologies (nonstationary linear signal processing technique and nonstationary nonlinear signal processing technique) have been carried out and their performance has been compared.

2.5 Feature Extraction and Feature Selection Techniques

Feature extraction and feature selection techniques provide data to be represented in reduced subset [7]. In the following subsection,a brief review of feature extraction and feature selection techniques carried in the analysis of VAG signals have been presented. This section also includes the proposed methodology for feature extraction and selection studied in this thesis.

2.5.1 Feature Extraction

Feature extraction is an important aspect in any pattern classification. It refers to processes in which input data is transformed into a set of features. These features possess different distinct properties of the input data which helps in differentiating among the input pattern [74]. Brain Ripley defined feature extraction as “Feature extraction is generally used to mean the construction of linear combinations of continuous features which have good discriminatory power between classes” [75]. To develop a computer-aided diagnosis (CAD) system for any biomedical application, feature extraction techniques plays an important role. The features obtained from VAG signals could be utilized in discriminating and detecting abnormalities in VAG signals. Researchers have successfully carried out the analysis of VAG signals by various feature extraction methodologies. These extracted features are significant since they are correlated with the degraded articular cartilage. Initial studies carried out by various researchers did not assume VAG signals as a nonstationary signal. Hence, some of the traditional feature extraction techniques were applied to VAG signal [20].

Reddy *et al.* have used power spectral analysis on VAG signals acquired from subjects having OA, rheumatoid arthritis, and chondromalacia patella [76]. Their study concluded that VAG signals of these three cases behaved differently in between 100-500Hz. Researcher used signal processing techniques for segmenting the VAG signals. In the segmentation process, the VAG signals were broken down into series of quasi-stationary or stationary segments and features were extracted from segmented signals. Tavitha *et al.* applied fixed segmentation and linear prediction modelling on segmented VAG signals and three features (model parameters, dominant poles, spectral power ratio) were extracted [77]. The adaptive segmentation was further modified by Moussavi by using recursive least squares (RLS) algorithm [78]. Krishna *et al.* further carried out series of work with adaptive segmentation and computing autoregressive coefficients using different methods [8], [22], [79], [80]. Features were extracted from the segmented signal of healthy and unhealthy VAG signals. The data set used by the above mentioned researchers had used the VAG signals acquired by Rangayyan and pattern classification of the accuracy of 75.6% was obtained in distinguishing healthy and unhealthy VAG signals. But the processes of segmentation required clinical information gathered from arthroscopy. Therefore, this drawback does not make segmentation a feasible approach.

Yunfeng Wu and Rangayyan have worked extensively on the analysis of VAG signal using statistical parameters [18], [81], [82]. The VAG signals were analysed by simple signal processing techniques followed by simple features extraction techniques. Therefore, the need for segmentation and dependency of clinical information was removed. Wu *et al.* computed the histogram of VAG signals. In order to characterize the differences observed between the histograms of healthy and unhealthy VAG signals, features based on statistical parameters such as moments, form factors and turn counts were extracted. The measures are based upon the moments of the probability density function (PDF) of the given VAG signals. Wu further carried the work by computing the probability density function of VAG signal by Parzen-window to represent the statistical characteristics of VAG signals [83]. Kullback-Leibler distance and statistical parameters were computed. The classification accuracy of 77.53%, sensitivity of 71.05%, and specificity of 82.35% were obtained using neural network using a radial basis function [83]. Wu *et al.* further tested these statistical features derived from kernel-based bivariate probability distribution modeling estimation [84]. This density distribution showed a significant difference between healthy and unhealthy VAG signals and gave classification accuracy of 86% [84]. This work was further carried forward by computing two more additional features in the time domain. Features such as a number of atoms obtained from wavelet matching pursuit decomposition and the number of prominent signals turn count with fixed threshold [85]. Dynamic weighted fusion (DWF) based classifier fusion method gave classification accuracy of 88.76% [85]. Rangayyan *et al.* have used fractal analysis to characterize spectral energy of healthy and unhealthy VAG signals. The fractal dimension (FD) was obtained from the power spectrum along with 1/f model [86]. Classification accuracy using area under the receiver operating characteristic curve (ROC-AUC) of 0.74 was obtained. Yang *et al.* carried out the analysis of VAG signal by extracting features using detrended fluctuation analysis algorithm and averaged envelope amplitude of VAG signals [87]. The work done by Wu using simple feature extraction techniques [18], [83]–[87] did not carry preprocessing using double cascaded moving average filter. Therefore, the presence of noise (low frequency components) and baseline wander had an effect on classification performance. The summary of literature survey carried on feature techniques has been shown in the table 2.4.

In initial studies, features were mostly extracted from the segmented signals using linear signal processing techniques. The segmentation was either fixed or adaptive. The segmented signals were considered stationary and features such as AR coefficients, poles etc. were extracted. Later, the segmentation was removed by representing signals in PDF and statistical features were extracted. Therefore, the experiment lacked the necessary signal preprocessing procedures to remove random noise and baseline wander artefacts. Such a phenomenon would ruin the effectiveness of threshold-based signal analysis especially for feature extraction techniques. This may mislead the classification performance. Entropy, which provides the measure of disorder within a system is used in biomedical applications as feature extraction techniques. Similarly, recurrence quantification analysis (RQA) has been studied in the analysis of nonlinear data

TABLE 2.4: Feature Extraction

Year	Author	Techniques	Features	Classifier	Accuracy%
1995	Tavitha[77]	Fixed segmentation & linear prediction modelling	model parameters, dominant poles, spectral power ratio	-	-
1996	Moussavi[78]	recursive least squares	AR coefficients, VMS, age	leave-one-out	71.1
1997	Rangayyan[79]	adaptive segmentation using RLSL	Dominant poles cepstral coef.	LRA	75.6
1997	Krishnan[22]	adaptive segmentation using RLSL	AR coef.,VMS	leave-one-out	68.9
2007	Wu[18]	PDF + histogram	FF,mean,TC	NN	0.82 AUC
2010	Wu [83]	Parzen Window	FF,mean,TC	NN	77.53
2013	Wu[84]	Bivariate Probability Distribution Modeling	FF,mean,TC	maximal posterior probability	86.6
2013	Wu[85]	MP segmented	MP atoms & TC	DWF	88
2013	Wu[86]	Power Spectrum	Fractal dimension	RBF NN	0.74 AUC
2014	Wu[87]	bivariate Gaussian	fractal scaling	Bayesian decision	0.95 AUC

especially in cases related to EEG signals and geographical studies. From literature survey it has been inferred that, entropy based features and RQA parameters have not been used as feature extraction in the analysis of VAG signals. Therefore entropy based techniques and RQA parameters could be investigated.

2.5.1.1 Proposed Feature Extraction Techniques

Feature extraction is an important step in pattern classification. The characteristics of VAG signals are obtained by feature extraction techniques. This thesis presents various feature extraction techniques with respect to linear and nonlinear signal processing technique. These techniques are yet to be explored for the analysis of VAG signals and they are as follows:

Entropy based feature extraction: Entropy based feature extraction techniques has been proposed in this study. Entropy is the measure of the energy of the system or is a quantified measure of the complexity of the signals. This work proposes to use entropy based features involving approximate entropy (ApEn), sample entropy (SampEn), Shannon entropy (ShEn), Rényi entropy (ReEn), Tsallis entropy (TsEn) and permutation entropy (PeEn). Reported

literature revealed that entropy based features have been extensively studied in the analysis of biomedical signals [88]–[91].

Recurrence Quantification Analysis (RQA): The recurrence quantification analysis (RQA) is a method to analyse the nonlinear data. It quantifies the duration and number of recurrences of a dynamical system presented by its state space trajectory. For time series, the concept of recurrence was introduced by Eckmann and developed by Webber and Zbilut [92] by means of the recurrence plot, a visual tool designed to display recurring patterns and to investigate nonstationary patterns. Recurrence plots (RP) are binary symmetric of $T \times T$ arrays. Several recurrence quantification measures have been introduced and successfully applied for the analyses of nonstationary data. This research presents four primary measures defined by Webber *et al.* [93]. They are recurrence rate (RR), determinism (DET), entropy (ENTR) and averaged diagonal line length (DLL).

Central tendency measure (CTM): Features are also extracted from the analytical signal representation of VAG signal. It is obtained by implementing Hilbert transform on IMFs. Central tendency measure (CTM) is computed from the analytical signal representation of VAG signal. CTM is a method used to sum up the observable data in the plots [94]. CTM is used to quantify the variability of the signal.

Statistical parameters: Features are also extracted from the time-frequency analysis of the VAG signals. As mention in the previous section, the time-frequency representation of reconstructed VAG is obtained by using SPWVD and HHT. This representation is considered as the time-frequency image. Thereby the statistical features such as mean, standard deviation, skewness, and kurtosis are extracted.

Hence, in this thesis, different features extraction techniques has been covered. The objective of studying various feature extraction techniques is to study the different parameters in characterizing VAG signal. This would help us in developing an efficient methodology for the computer aided diagnostic system.

2.5.2 Feature Selection

Feature selection refers to selecting a small subset of features that is relevant [95], [96]. The raw input data/signal is represented by large features and thereby consisting of high dimension vector. As a result, this could cause serious effects in building an effective classification model[97], [98]. Therefore reducing these number of irrelevant/redundant features can improve the classification model more effectively [74]. The advantages of feature selection techniques are understanding the input data, reducing the storage requirements, reducing learning and training time for building

classification, reducing the high dimension vector and improving the prediction performance [74]. The feature selection techniques are broadly classified as follows [99]–[101]:

1. filters
2. wrappers
3. embedded
4. hybrid methods

Various techniques available for feature extraction techniques have been covered in the review articles [102]–[104]. Feature selection techniques have been widely used in the application field of bioinformatics [105], biomedical signal analysis [106]–[110], classification of power quality disturbances [111], [112] etc.

Wu *et al.* had used feature selection technique for six statistical features [18], [81], [82], [113]. Rangayyan *et al.* have used sequential forward selection (SFS) and sequential backward selection (SBS) methods as feature selection methods for the analysis of VAG signals [86]. Features were extracted from fractional analysis parameters and spectral variability along with features from the previous study such as statistical parameters was consider [82]. The summary of literature survey carried on feature selection techniques has been shown in the table 2.5.

TABLE 2.5: Feature Selection

Year	Author	Techniques	Features	Classifier	Accuracy%
2008	Wu [82]	PDF + histogram	FF,mean, turn counts,VMS	NN	0.91 AUC
2013	Wu[86]	power spectral analysis	FF,mean, turn counts, VMS,fractal dimension	RBF NN	0.92 AUC

From table it can be observed that only limited studies has been carried out using feature selection algorithms. In above studies the feature selection algorithm was applied on less number of features and feature selection has not been applied on larger set of features. Therefore feature selection algorithm could be used on larger set of features. Therefore feature selection algorithm is presented in detail for selecting significant features.

2.5.2.1 Genetic & Apriori Algorithm For Feature Selection

Feature selection algorithm has also been proposed in order to identify the most stable, significant and discriminate features. Different techniques of feature selection have been utilized for improving the accuracy of the classifier by discarding redundant and irrelevant features. These

techniques have been employed in the domain of medical imaging, biomedical signals, and power-quality disturbance evaluations. In this thesis, k nearest neighbour based apriori algorithm (ApA) [114] and genetic algorithm (GA) [115] has been used for selecting the most stable and significant features from the extracted features. The feature sets obtained by feature selection algorithms are given as input to classifiers.

Genetic algorithm: GA is one of the most popular feature selection techniques used for improving the performance of data mining. GA is an evolutionary technique and sufficiently good results have been obtained with respect to other non-evolutionary and traditional feature selection methods. GA has been successfully implemented for large-scale feature selection problems. GA has been used as feature selection algorithm in various application related to biomedical signals [116], [117], credit risk assessment [118], text categorization [119].

Apriori algorithm: Apriori uses a “bottom” up approach. Apriori algorithm has been widely used in data mining [120]. Apriori algorithm is based upon association rule learning over a database. It identifies the most frequent individual’s items and extending them to larger sets. The associating rules help in selecting frequent items of the set. The algorithm terminates when no further frequent items are selected. Breadth-first search and a Hash tree are used for counting the candidate key set effectively. Apriori algorithm has been widely used as feature reduction algorithm in wide application ranging from data mining, imaging, wind forecasting etc. [121]–[124].

Hence, apriori algorithm and GA as feature selection techniques has been proposed in this study. This work is presented in chapter 4 of the thesis.

2.6 Pattern Classification

Pattern classification refers to classifying the input data based upon the characteristics or features extracted from the input data/signal into particular classes [7]. The main task of pattern classification is to generate a decision boundary to separate the input data/signal into classes or group depending upon the extracted features. Classification methods can be broadly categorized into

- a. supervised learning
- b. unsupervised learning
- c. reinforcement learning.

In supervised learning, a training set of the input data is available to build the classification model whereas unsupervised learning involves clustering of data based upon the similarity.

Reinforcement learning is a classification model built using the learning continuously by trial and error method [125], [126]. The analysis of biomedical signal is concluded by classifying the given input data/signal into particular classes. This is an important step since a particular diagnostic inference can be made [127]. For building an effective classification model, it is recommended that classes or sets are disjoint distinctly. Therefore, feature extraction and feature selection techniques are important aspects in building the classification model. Smaller the number of features, the classification model is simpler to build [7]. Various classification techniques have been developed with the advancement of machine learning algorithms. These techniques have been used in various application of pattern classification in image processing, signal processing, text processing etc. The effectiveness of the classifier techniques can be evaluated using different performance parameters[128]. The performance of the classification model is evaluated by following parameters. They are as sensitivity (SEN), specificity (SPF), accuracy (ACC), positive predictive value (PPV), negative predictive value (NPV), Matthews correlation coefficient (MCC) and false discovery rate (FDR)[125], [126]. And they are evaluated as follows:

- 1) Sensitivity (SEN): It measures the proportion of the positive sample to the test data and is defined as:

$$SEN = \frac{TP}{(TP + FN)} \times 100 \quad (2.1)$$

- 2) Specificity (SPF): It measures the proportion of the negative samples to the test data and is defined as:

$$SPF = \frac{TN}{(TN + FP)} \times 100 \quad (2.2)$$

- 3) Accuracy (ACC): It is defined as the ratio of the samples correctly classified to the total number of samples and is defined as:

$$ACC = \frac{TP + TN}{(Total)} \times 100 \quad (2.3)$$

- 4) Positive predictive value (PPV): It is the ratio of true positive values to the total number of positive samples identified by the classifier and is defined as:

$$PPV = \frac{TP}{(TP + FP)} \times 100 \quad (2.4)$$

- 5) Negative predictive value (NPV): It is the ratio of true negative values to the total number of negative samples identified by the classifier and is defined as:

$$PPV = \frac{TN}{(TN + FN)} \times 100 \quad (2.5)$$

- 6) Matthews correlation coefficient (MCC): For larger values of MCC, classifier performance will be better. The MCC can be defined as :

$$MCC = \frac{(TP \times TN - FN \times FP)}{\sqrt{(TP + FP)(TN + FN)(TN + FP)(TP + FP)}} \quad (2.6)$$

- 7) False Discovery Detection (FDR): Its is the ratio of false positive to total number of positive samples identified by classifier and is defined as:

$$FDR = \frac{FP}{(TP + FP)} \quad (2.7)$$

where TP=True Positive, TN= True Negative FP=False Positive, FN= False Negative.

Another method to evaluate the performance of the classifier is receiver operating characteristic (ROC) curve. ROC curve is a graphical representation of true positive rate (TPR, also named sensitivity) against false positive rate (FPR, FPR = 1-specificity) [129]. The area under the ROC curve (ROC-AUC) is an effective way of comparing the performance of different features or classifiers. The area under ROC curve (ROC-AUC) provides a measure of how considerably a parameter would be able to classify between two diagnostic groups. Larger of the ROC area emphasises a better classification accuracy. Completely random discrimination would give points along a diagonal line from the left bottom to the top right corners and generate an area of 0.5 under the curve. Perfect discrimination between classes would yield points in the upper left corner of the ROC space and generate unity area under the ROC curve [130], [131].

Researchers have successfully applied machine learning techniques for pattern classification of VAG signal. Moussavi *et al.* had used nearest neighbour, unsupervised cluster analysis, discriminant analysis and logistic classification from statistical package for social sciences (SPSS) software. Results illustrated that classification performance of nearest neighbour and cluster analysis were quite poor in comparison with others. Logical regression analysis performed better with respect to other classifier [77], [78]. Krishna also used logical regression analysis for pattern classification [8], [22], [63]. Krishna *et al.* have used the leave-one-out (LOO) method classification for features extracted from adaptive segmentation [80]. Umapathy had used linear discriminant analysis (LDA) based classifier using SPSS and validation was performed by LOO method and a classification of 80% was obtained [65], [132]. Wu *et al.* have used neural network with radial basis function and Fisher's linear discriminant analysis (FLDA) with statistical features (the turns count and variance of the mean squared values) as input [18], [81], [82]. A screening efficiency of up to 0.8570 was achieved in terms of AUC-ROC. Tingting *et al.* applied linear and nonlinear strict 2-surface proximal (S2SP) classifiers [133] on statistical features extracted from VAG signals [18], [82], [83]. The classification performance using linear S2SP classifier gave AUC-ROC of 0.82 while nonlinear S2SP classifier gave AUC-ROC of 0.95 [113],

TABLE 2.6: Classifier

Year	Author	Techniques	Features	Classifier	Accuracy%
2007	Mu[133]	PDF + histogram	FF,mean,TC	Strict 2 surface classifier	0.95 AUC
2011	Wu[136]	Parzen Window	FF,mean, TC	LS-SVM	80.90
2014	Wu[137]	PDF + histogram	FF,mean,TC	k-nearest neighbour (k-NN)	80

[133]. Kim *et al.* used a back-propagation neural network technique for classifying acoustic based features [9], [134], [135].

Wu *et al.* used least square support vector machines and multiple classifier systems (MCS) based on a recurrent neural network (RNN) [136]. Statistical parameters extracted as features [18], [81], [82] from VAG signals were given as input to these classifiers. Results inferred that MCS gave a classification accuracy of 80.9% and AUC-ROC of 0.9484 [136]. Wu *et al.* used Fisher's linear discriminant analysis (FLDA), support vector machine (SVM) with polynomial kernels, and the maximal posterior probability decision criterion on statistical features [85]. Results illustrated that maximal posterior probability decision criterion gave the highest classification accuracy of 86.67% and AUC-ROC of 0.9096 among FLDA (accuracy: 81.33%, AUC-ROC: 0.8564) and SVM (accuracy: 81.33%, AUC-ROC: 0.8533) [85]. Classification based on dynamic weighted fusion (DWF) method was applied by Cai *et al.* [85]. Two features such as turn count and a number of atoms were derived from the wavelet matching pursuit decomposition along with the statistical parameters [18]. A classification accuracy of 88.76% and ROC-AUC of 0.9515 was obtained [85]. Kaizhi used *k*-nearest neighbour (*k*-NN) algorithm on statistical features extracted from VAG signals and (*k*-NN) algorithm gave classification accuracy of 80% [137]. Yang *et al.* used Bayesian decision rule on features extracted by detrended fluctuation analysis algorithm and averaged envelope amplitude of VAG signals [87]. A classification accuracy of 88% and AUC-ROC of 0.957 was obtained. Wu *et al.* have used support vector machine (SVM) for the analysis of VAG signals and gave a classification accuracy of 0.8356 and AUC-ROC of 0.9212 was computed [48]. The summary of literature survey carried on classifier used in the analysis of VAG signals has been shown in the table 2.1-2.6. It was concluded that LS-SVM provides better classification performance among different classifier since it avoids the problem of overfitting.

2.6.0.1 Least Square Support Vector Machine For Pattern Classification

Support Vector Machine (SVM) is a machine learning algorithm based on statistical learning of Vapnik-Chervonenkis dimension [138]. SVM was proposed by Vapnik and has been used for nonlinear estimation, function estimation, density estimation and classification [138]. One of

the major draw back of SVM is its high computation load. LS-SVM singles out this liability by computing linear equation instead of solving the convex for quadratic programming problem in case of SVM. Due to the ease of its implementation, it has widely used in the area of adaptive signal processing [139].

Kernel based learning is a recent development in this area[139]. This thesis proposes to use LS-SVM with radial basis function (RBF) as a kernel function for the binary classification of VAG signals. LS-SVM is upgraded version of SVM and has been widely used in various application domain [140], [141]. These applications include biomedical signals, vibration signals from the ball bearing, the financial stock market prediction etc. [142]–[144]. In this thesis, LS-SVM with RBF kernel function as classifier has been proposed.

2.7 Gaps In Existing Literature

Based on literature survey carried in the analysis of VAG signals, following research gaps have been identified:

- ✓ From literature survey it was concluded that the analysis of VAG signal using linear signal processing techniques was carried out. In these techniques the VAG signals were segmented and assumed stationary. The experiment lacked the necessary signal preprocessing procedures to remove random noise and baseline wander artefacts. Such a phenomenon would devastate the effectiveness of threshold-based signal analysis and results could be misleading. Nonstationary linear signal processing namely WPD has been used in various application especially related to biomedical studies. The decomposition of signal into subband signals are highly correlated with applications. For example the subband signals obtained from EEG signal were related to neural activities. Therefore, considering similar characteristics of EEG and VAG signals, the analysis of VAG signal using nonstationary linear signal processing technique using WPD can be investigated.
- ✓ In the reported literature survey, it was reviewed that nonlinear methods based on EMD were studied as preprocessing technique for the analysis of VAG signals. But the presence of mode mixing and noises still persisted in these methods. To eliminate these drawbacks, a nonstationary nonlinear signal processing techniques based on EMD known as CEEMDAN can be used. Therefore, the analysis of VAG signals could be carried out using CEEMDAN as preprocessing step.
- ✓ In the reported literature survey, features were extracted from the segments of VAG signals and features such as autoregressive modelling coefficients, poles, statistical moments etc. were extracted. These features were extracted from linear signal processing methods.

Entropy which provides the dynamic of the system and nonlinear features namely RQA have been studied as feature extraction techniques in various other applications especially in the analysis of biomedical studies, geographical studies etc. Therefore, entropy based features, RQA parameters and simple feature extraction techniques can be carried for computing the characteristics of VAG signals.

- ✓ Limited studies have been carried out using feature selection in the analysis of VAG signal. In the reported studies, few features were extracted and traditional feature selection algorithm was applied. Therefore, features selection techniques such as genetic and apriori algorithm can be used for a set of large feature elements for building effective classification model.
- ✓ Few studies have been reported in the analysis of VAG signal using time frequency distribution. Most of the studies has been carried out using linear signal processing techniques. These techniques are based on classical Fourier analysis which assumes that signals are infinite in time or periodic. Thereby the spectrum of the signal using these methods (eg: STFT, CWT etc.) are poorly represented in term of sharpness and contrast. Therefore, the spectrum representation of VAG signals can be improved using advance time frequency distribution namely SPWVD and CEEMDAN-HHT. Hence, statistical features could be extracted from TFD of VAG signals and these techniques have been unexplored in the analysis of VAG signals. Therefore, time-frequency analysis using SPWVD and CEEMDAN-HHT can be carried out for the analysis of VAG signals.

2.8 Conclusion

This chapter briefly reviews various signal processing techniques, feature extraction, feature selection and classification methods on VAG signals. Gaps in the existing literature were identified and literature survey on the proposed methodology has been explained briefly.

Chapter 3

Analysis of VAG signals using Wavelet Packet Decomposition

The wavelet transform is a tool that cuts up data or functions or operators into different frequency components, and then studies each component with a resolution matched to its scale.

Ingrid Daubechies (1992)

3.1 Introduction

This chapter presents the analysis of VAG signals using wavelet packet decomposition (WPD). The VAG signals are decomposed into subband signals using WPD. The subband signals are based on frequencies where each subband signal represents a particular frequency.

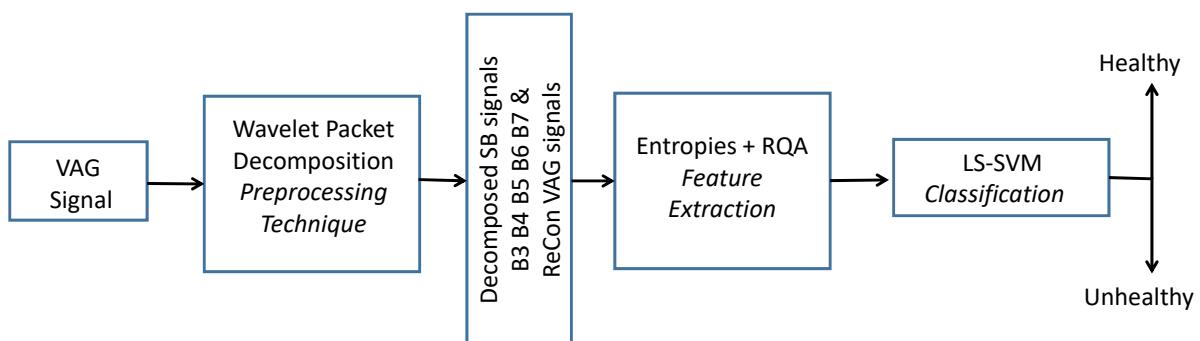


FIGURE 3.1: Flow chart for diagnosis of knee joint disorders using VAG signals

Entropy based features and recurrence quantification analysis (RQA) parameters are extracted from subband signals. A pattern classification is carried out using LS-SVM. The flowchart of the work carried out in this chapter is shown in figure 3.1.

This chapter has been organised as follows. The section 3.2 and 3.3 gives the theoretical background of the proposed methodology. Section 3.2 describes the preprocessing technique using WPD. Feature extraction techniques based on entropy and RQA parameters has been covered in section 3.3. Results obtained from the methodology has been presented in section 3.4 and discussion in section 3.5. Finally, the chapter concludes with section 3.6.

3.2 Wavelet Packet Decomposition (WPD)

Wavelet packet decomposition was introduced by Coifman, Meyer, and Wickerhauser [33]. The dilation property of WT makes it quite popular for various application such as detection of irregularities, time-frequency analysis, feature extraction, waveform representations, segmentation and compression of digital data. The dilation property and translation helps to adjust the frequency band and it identifies the position of low frequency and high frequency components. The multi-resolution analysis and time localization ability are the two important characteristics of WPD which motivates to analyse nonstationary signals.

The discrete wavelet transform (DWT) of a given signal is given as

$$C(j, k) = \sum_{n \in Z} x(n)\psi_{j,k}(n) \quad (3.1)$$

where $x(n)$ is the signal to be analysed, $\psi_{j,k}(n)$ is the discrete wavelet function, j is the scale factor and k is the displacement.

The wavelet transform divides the signal into high frequency and low frequency components. The low frequency component is called an approximation space of the original signal while the difference of the approximations (V_j) between two successive decompositions is called a vector of detail space (W_j). The multi-resolution analysis of the signal is obtained by successively decomposing the low frequency components.

WPD is an extended work of WT in which the signals are passed through the bank of filters. Here in WPD, the approximation space (V_j) and the detail space (W_j) are further decomposed successively. As a result, the input signal is iteratively filtered at each successive stage. At each stage of filtering, the frequency is divided into subband. The successive splitting of vector spaces is represented by a tree. Mallat computed the equations for obtaining the coefficients [34]. For any node in binary tree labelled by (j,p) , where $j-L \geq 0$ is the depth of the node and p is the

number of nodes that are on its left at the same depth $j-L$. The two wavelet packet orthogonal bases at the children nodes are defined by

$$\psi_{j+1}^{2p} = \sum_{-\infty}^{\infty} h[n] \psi_j^{2p}(t - 2^j n) \quad (3.2)$$

$$\psi_{j+1}^{2p+1} = \sum_{-\infty}^{\infty} g[n] \psi_j^{2p}(t - 2^j n) \quad (3.3)$$

since $\psi_j^{2p}(t - 2^j n)_{n \in Z}$, therefore

$$h[n] = \langle \psi_{j+1}^{2p}(u), \psi_j^{2p}(t - 2^j n) \rangle \quad (3.4)$$

$$g[n] = \langle \psi_{j+1}^{2p+1}(u), \psi_j^{2p}(t - 2^j n) \rangle \quad (3.5)$$

where $h[n]$ and $g[n]$ are low pass and high pass filters respectively.

Hence figure 3.2 (a) and (b) shows the level 3 WPD of the signal. The complete binary tree is produced in the one-dimensional case. As observed, detail and approximation spaces are obtained by splitting the original signal by low pass and high pass filter bank successively. The resultant vectors are obtained and represented in a binary tree as shown in figure 3.2. In this

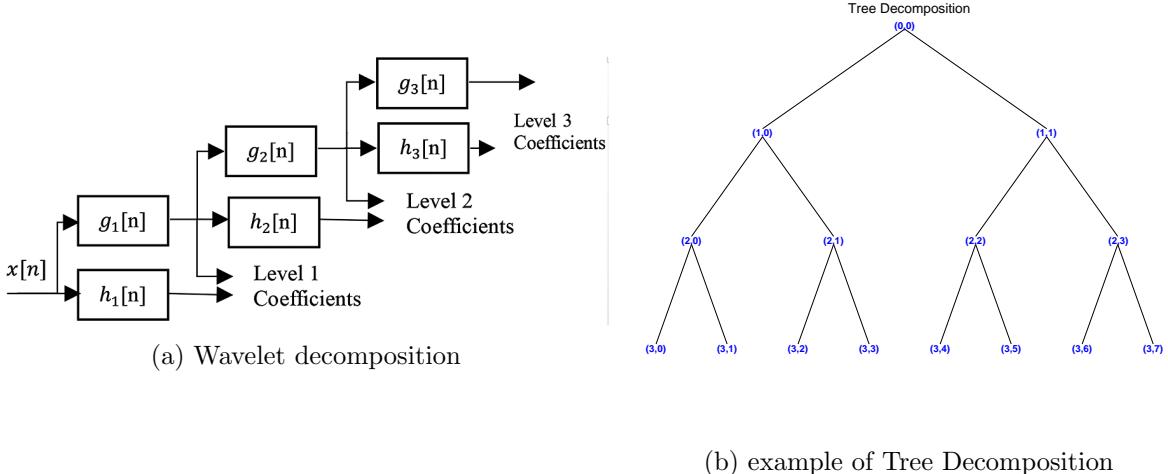


FIGURE 3.2: Wavelet decomposition

work, the maximum available frequency is 1 kHz since the VAG signal has been sampled at 2k Hz. The number of levels of WPD is selected based upon the VAG signal's dominant frequencies. The levels are selected with respect to the required frequencies that are eventually related to the wavelet coefficients. The VAG signal is decomposed into ten wavelet based subband signals ($B1, B2, B3, B4, B5, B6, B7, B8, B9, B10, A10$).

TABLE 3.1: Decomposition of the VAG Signal into Sub-bands

Decomposed Signal	Frequency	Decomposed Signal	Frequency
B1	500-1000 Hz	B6	15.62-31.25 Hz
B2	250-500 Hz	B7	7.81-15.62 Hz
B3	125-250 Hz	B8	3.90-7.81 Hz
B4	62.5-125 Hz	B9	1.95-3.90 Hz
B5	31.25-62.5 Hz	B10	1-1.95 Hz
A10	0-1 Hz		

The frequency range 50 Hz-500 Hz captures the VAG [18]. The subband signal $B8, B9, B10, A10$ represents slow baseline variations mainly caused by sweating and different artefacts, while the subband signals $B1$ capture high-frequency artefacts such as muscular contractions and environment noise[145]. These high frequency subband and low frequencies signals $B1, B2, B8, B9, B10, A10$ aren't investigated for evaluating features. For the analysis of physiological signals especially for biomedical signal processing, an efficient wavelet such as Daubechies 4(db4) was considered as the mother wavelet function [32].

3.3 Feature Extraction

3.3.1 Entropy Based Feature Extraction

3.3.1.1 Approximate Entropy

ApEn was initially developed to analyse medical data, such as heart rate[146], [147]. For a given N data points from a time series $x(n)=x(1), x(2), \dots, x(N)$ where N is the length, Two parameters, the embedding dimension m and a tolerance window r are specified. Usually, m is set to be 1, 2 or 3 and r is set to be some percentage value of the standard deviation of the amplitude of time series. ApEn can be calculated in following these steps [146], [147]

Step 1. Form m -dimensional vectors $X_m(1), \dots, X_m(N-m+1)$ defined by $X_m(i)=[X_m(i), X_m(i+1) \dots, X_m(i+m-1)], i=1, 2, \dots, N-m+1$. Each vector is viewed as a template.

Step 2. The distance between each of the corresponding templates is denoted as $d[X_m(i), X_m(j)]$, is computed as the maximum absolute difference between their scalar components:

$$d[X_m(i), X_m(j)] = \max_{k=0,\dots,m-1}(|x(i+k) - x(j+k)|) \quad (3.6)$$

Step 3. For a given template $X_m(i)$, count the number of template matching, denoted as, i.e. the number of satisfying the distance $d[X_m(i), X_m(j)]$ Then

$$A_i^m(r) = \frac{1}{N-m+1} A_i \quad (3.7)$$

Step 4. Define $\phi_i^m(r)$

$$A_i^m(r) = \frac{1}{N-m+1} \sum_i^{N-m+1} \ln A_i^m(r) \quad (3.8)$$

Step 5. Increase the dimension to $m + 1$ and follow the steps 1 - 4 to compute $A_i^{m+1}(r)$ and $A^{m+1}(r)$

$$A_i^{m+1}(r) = \frac{1}{N-m} \sum_i^{N-m} \ln A_i^{m+1}(r) \quad (3.9)$$

Step 6. The approximate entropy is defined as

$$ApEn(m, r) = \lim_{N \rightarrow \infty} [A^m(r) - A^{m+1}(r)] \quad (3.10)$$

Step 7. For a finite length of data points N, the approximate entropy is estimated as

$$ApEn(m, r, N) = [A^m(r) - A^{m+1}(r)] \quad (3.11)$$

where

$$A^m(r) = \frac{1}{N-m+1} \sum_i^{N-m+1} \ln A_i^m(r) \quad (3.12)$$

3.3.1.2 Sample Entropy

The drawback of ApEn is biasing caused by the self matching, this was reduced by the concept of SampEn. J.S. Richman *et al.* proposed the concept of SampEn [148]. SampEn generally demonstrates the corresponding consistency where ApEn doesn't and SampEn is independent of record length. SampEn has been applied successfully in biomedical application especially, in the analysis of the epileptic activities. Detection of seizure were identified based on computing SampEn by Song *et al.* [149]. SampEn is computed by the following equation. Mathematically,

the computation of SampEn has the same steps 1-3 as in case of ApEn.

$$A^m(r) = \frac{1}{N-m+1} A_i^m(r) \quad (3.13)$$

$$SampEn(m, r) = \lim_{N \rightarrow \infty} [ln A^m(r) - A^{m+1}(r)] \quad (3.14)$$

$$SampEn(m, r, N) = ln[A^m(r) - A^{m+1}(r)] \quad (3.15)$$

In this work, the values of $m=2$, $r=0.2*\text{std}(\text{data})$ and $N = 8000$ (total data length of VAG signal) are set in the calculation of approximate and SampEn. The parameters selected for computing ApEn and SampEn were concluded by Jennifer *et al.* [150] on the clinical data. These parameters provided an optimal setting for accurately estimating ApEn and SampEn. These parameters are considered for biomedical signals especially for the nonlinear and nonstationary signals such as EEG signal[151]–[154]. Considering the similar nature of VAG signals as nonlinear and nonstationary, the values of m , r and N were chosen.

3.3.1.3 Shannon Entropy

The complexity of measurement data or measure of information content is carried out by Shannon Entropy [155]. Claude Shannon proposed Shannon entropy as a measure of the information, content or complexity of data. Shannon has proposed ShEn for use in communication domain. High value of entropy is often associated with more randomness and less system order. In the present work, ShEn is computed using power spectrum density (PSD) of the VAG signal in order to evaluate the dynamic complexity of the VAG signals. Steps involved in computing ShEn are as follows

1. The spectrum of the signal is computed by applying Fourier transform (Fast Fourier Transform).
2. Power spectrum is obtained by computing the Fourier's transform of the signal. E_r denotes the power level of the frequency component.
3. Probability density function is obtained by normalizing the computed PSD.

$$e_f = \frac{E_f}{\sum E_r} \quad (3.16)$$

4. The ShEn is given as

$$ShEn = \sum_f e_f \log\left(\frac{1}{e_f}\right) \quad (3.17)$$

3.3.1.4 Rényi Entropy

Rényi [156] introduced Rényi entropy which is a mathematical generalization of Shannon entropy. Rényi defines the information measure in more general terms that would preserve additivity individual events. The ReEn in VAG signal represents the statistics as indices of diversity. The spikes or the high peaks of occurrence in abnormal VAG signals would be inferred as these events. The Rényi entropy can be defined as :

$$ReEn(\chi) = \frac{1}{1-\chi} \left(\sum_f e_f^\chi \right), \chi > 0, \chi \neq 1 \quad (3.18)$$

Where χ is the enabling measurement of uncertainty among the distribution and due to which ReEn is much more flexible. In our study, the value of χ is taken as 2.

3.3.1.5 Tsallis Entropy

Tsallis entropy (TsEn) was presented by Constantino Tsallis which is a generalization of the standard Boltzmann–Gibbs entropy [157]. Tellenbach et al. prove that TsEn is also known as non-additive entropy, is useful for characterizing measures with non-Gaussian trends. TsEn is used for the analysis of VAG signals since VAG signals are characterized by multicomponent signals and non-Gaussian trends. TsEn is useful in the analysis since TsEn provides strong correlations between the different microstates in a system. TsEn can signify the complexity of the VAG signal. Tsallis Entropy characterizes the behaviour emulated by VAG signals in spikes and bursts. The uncertainties could be modelled by the empirical value f . TsEn defined as follows

$$TEn = \frac{1 - \sum_{j=1}^K d_j^f}{f - 1} \quad (3.19)$$

The parameter f is called as non-extensibility index, d_j is the probability of j^{th} state and $b \neq 1$. Since VAG signals are nonstationary and nonlinear, therefore it would be reasonable to consider VAG signals as a sub extensive system (i.e., $b > 1$). In the present work, $f = 2$ has been set for conducting simulations. TsEn has been widely used in the analysis of biomedical signal processing especially in cases of ECG [158] and EEG [91], [159] recently. Recent work concluded that TsEn could provide a descriptive information, particularly in cases of identifying the spikes or burst EEG analysis.

3.3.1.6 Permutation Entropy

Permutation entropy is another complexity measurement for stationary or non-stationary, deterministic or stochastic, regular or noisy or any arbitrary time series data[160]. It is based

on comparing the neighbouring values and PE is evaluated based on PE index associated for a pattern of order. Simplicity, robustness with respect to non-linear transformations and quick computation are the major advantages of PE. The scalar time series $x(t)_t^T$ is embedded into an s -dimensional space $Z_t = [z(t), z(t + L), \dots, z(t + (s-1)L)]$, where s is called the embedding dimension and L the embedding delay time.

If w_j represents the number of occurrences of the j^{th} symbol in the time series, then the probability of occurrence of the j^{th} permutation can be estimated as

$$p_j = \frac{w_j}{N - s + 1} \quad (3.20)$$

where N is length of the signal. And thus PeEn can be defined as

$$PeEn = - \sum_{k=1}^K p_j \log(p_j) \quad (3.21)$$

The authors, Bandit and Pompe suggested values of $s = 3, 4, 5\dots 7$ and yet they inferred that the values $s = 3$ or 4 might still be small. In this work, the values of s and t_d are 3 and 1 respectively since these values have been reported in various analysis of biomedical signals [161], [162].

3.3.2 Recurrence Quantification Analysis (RQA)

Non-linear data analysis is been carried out by an advanced technique known as recurrence plot (RP)[92], [93]. Recurrence plots is a graphical technique to locate hidden recurring patterns, non-stationary and structural changes. Recurrence plots (RPs) are binary symmetric of $T \times T$ arrays, in which a spot is placed at (g, h) when the distance between Y_g and Y_h is less than a

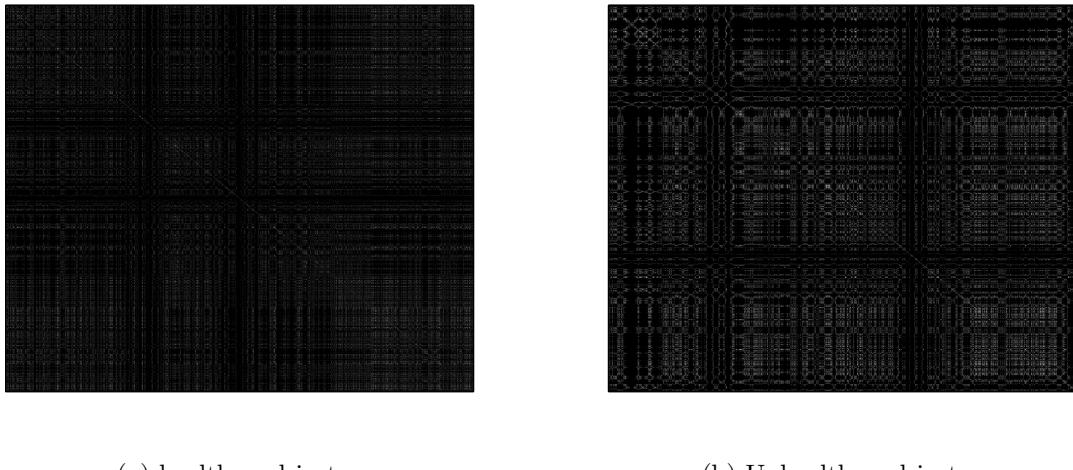


FIGURE 3.3: An example of recurrence plot for raw VAG signal

prescribed value $\in [92], [93]$.

$$R_{g,h} = \begin{cases} 0 & \text{if } \|Y_g - Y_h\| < \epsilon \\ c & \text{if } \|Y_g - Y_h\| \geq \epsilon \end{cases} \quad (3.22)$$

RQA was framed to provide a quantification measure about recurrence plot since enumerating graphically would be onerous. The figure 3.3 shows the RQA plot for the raw VAG signal of a healthy and unhealthy subject. Four primary variables including %REC, %DET, ENTR and DLL are evaluated in the present study for the quantification of recurrences [92], [93].

- Recurrence rate (%REC): The recurrence rate is correlated as the probability that a particular state will reappear.

$$\%REC = \frac{1}{T^2} \sum_{g,h=1}^T R(g,h). \quad (3.23)$$

- Determinism (%DET): Passivity of a dynamic system is defined by %DET. It is also the percentage of recurrence points which form diagonal lines. These lines represent epochs of similar time evolution of states of the system. Hence, DET indicates the predictability of the system

$$\%DET = \frac{\sum_{l=l_{min}}^T lQ(b)}{\sum_{g,h}^T R_{g,h}} \quad (3.24)$$

$Q(b)$ is the histogram of the length b of the diagonal lines

- ENTR: The complexity or the Shannon entropy is computed as ENTR in recurrence plot .

$$ENTR = - \sum_{l=l_{min}}^T Q(b) \ln Q(b) \quad (3.25)$$

where

$$Q(b) = \frac{W(b)}{\sum_{b=b_{min}}^T Q(b)} \quad (3.26)$$

- Average diagonal line length: The average length of the diagonal lines:

$$DLL = \frac{\sum_{b=b_{min}}^T bQ(b)}{\sum_{b=b_{min}}^T Q(b)} \quad (3.27)$$

In this work, to estimate the RQA, calculation were based on the work by Grassberger *et al.* where $c=1$, a delay of $\tau = 1$, a radius of $\epsilon = 15$ and distance norm=Euclidean are chosen [163].

3.4 Results

The proposed work consisting of preprocessing, feature extraction and classification has been carried out in MATLAB 2015b. After pre-filtering using a double cascaded moving average filter, a WPD is carried out by decomposing the VAG signals into subband signals. Since the sampling rate of VAG is 2k, the frequency available would be 1 kHz. Hence, VAG signals are decomposed into 10 levels of subband signals ($B_1, B_2, B_3, B_4, B_5, B_6, B_7, B_8, B_9, B_{10}, a_{10}$) and each subband signals consists of a particular frequency. Figure 3 (a) and (b) represents the subband signals obtained from WPD of healthy and unhealthy respectively. The subband signals B_1, B_2, B_8, B_9 and B_{10} as observed in figure 3.4 were ignored. B_8, B_9 and B_{10} represented artefacts such as baseline wander, while the subband signals B_1 and B_2 indicated the presence of high-frequency components.

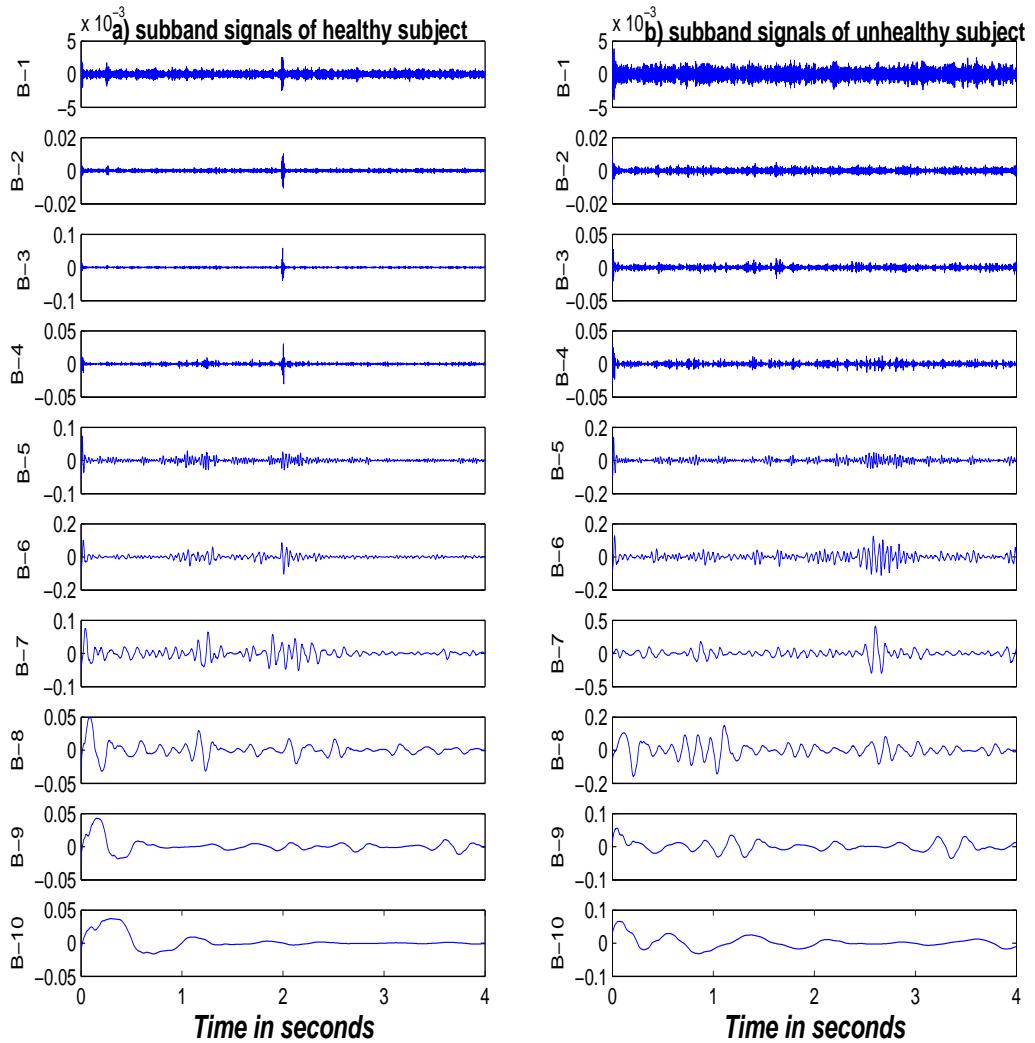


FIGURE 3.4: Wavelet decomposition of VAG signal for (a) healthy and (b)unhealthy subject

The VAG signal is reconstructed from the subband signal (*B3, B4, B5, B6 and B7*) and is shown in figure 3.5. Hence the feature based on six entropies and four RQA parameters are

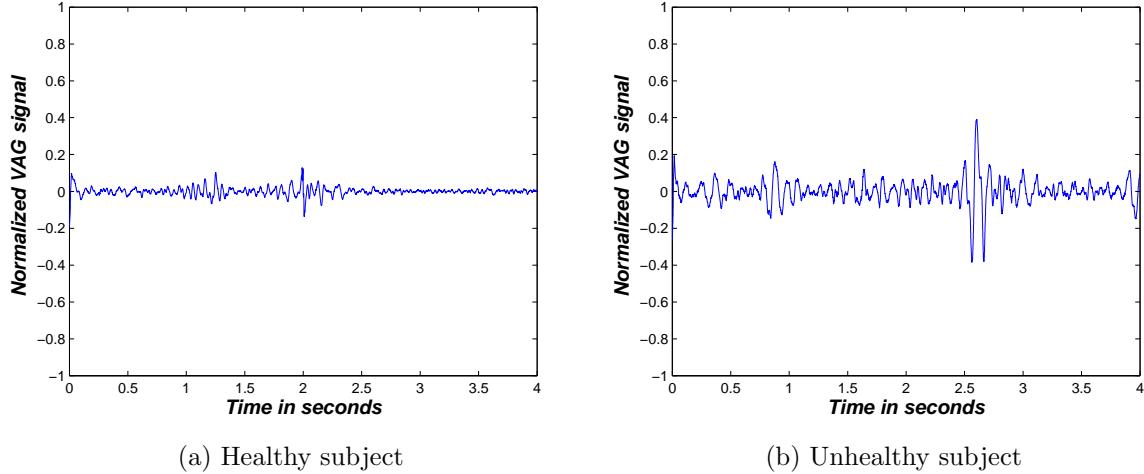


FIGURE 3.5: Reconstructed VAG signals

extracted from subband signals (*B3, B4, B5, B6 and B7*) and reconstructed VAG signal. These procedures are followed for 89 samples of VAG signals. The feature set comprised of fifty features were extracted from subband signals *B3, B4, B5, B6 and B7*:

- 1) ApEn from subband signal B3 to B7.
- 2) SampEn from subband signals B3 to B7.
- 3) ShEn from subband signals B3 to B7.
- 4) ReEn from subband signals B3 to B7.
- 5) TsEn from subband signals B3 to B7.
- 6) PeEn from subband signals B3 to B7.
- 7) RQA-RR from subband signals B3 to B7.
- 8) RQA-DET from subband signals B3 to B7.
- 9) RQA-ENTR from subband signals B3 to B7.
- 10) RQA-DLL from subband signals B3 to B7.

Similarly, ten features (six entropies + four RQA parameters) were extracted from reconstructed VAG signals. Figure 3.6 shows the RQA plot for the reconstructed VAG signal of a healthy and unhealthy subject.

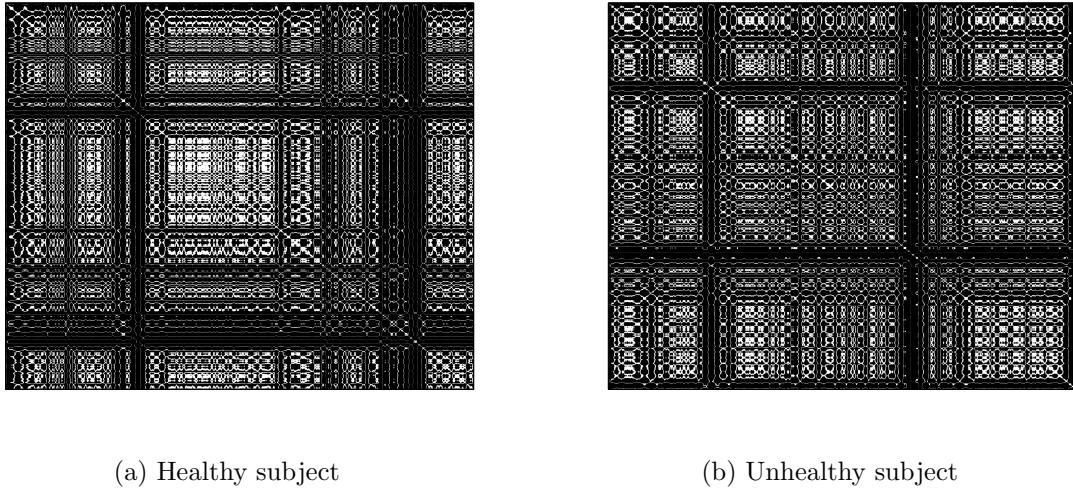


FIGURE 3.6: An example of recurrence plot for reconstructed VAG signal

3.4.1 Feature Extraction

Table 3.2 shows the mean and standard deviation of entropy and RQA based features extracted from reconstructed VAG. From table 3.2 it can be observed that, the mean and standard deviation of ApEn is 2.0113 ± 0.1303 for unhealthy subject which is 1.3 times that of healthy subjects (1.7792 ± 0.1650). Similarly, though the mean of SampEn of healthy (2.3083) class is larger than unhealthy class (2.0744), the huge variation can be observed since the standard deviation of unhealthy class is two times the healthy class. Similar inference are observed in table 3.2 for other entropy based features and RQA parameters.

TABLE 3.2: Statistical measures of entropy and RQA based features extracted from reconstructed VAG signals

Features	Healthy Class	Unhealthy Class	
	Mean \pm Std	Mean \pm Std	P value
ApEn	1.7792 ± 0.1650	2.0113 ± 0.1303	2.421E-08
SampEn	2.3083 ± 0.1908	2.0744 ± 0.3727	1.7E-05
ShEn	5.0713 ± 0.4934	5.3087 ± 0.5355	0.0828
ReEn	4.4365 ± 0.5220	4.6323 ± 0.5849	0.1671
TsEn	0.9864 ± 0.0082	0.9886 ± 0.0066	0.1671
PeEn	1.7047 ± 0.1963	1.8260 ± 0.1467	0.0071
RR	0.2524 ± 0.0737	0.2885 ± 0.1391	0.3445
DET	0.9974 ± 0.0016	0.9967 ± 0.0026	0.2846
ENTR	3.3780 ± 0.5787	3.3725 ± 0.4149	0.6854
DLL	17.1224 ± 14.6281	14.7414 ± 6.3774	0.7632

Kruskal-Wallis test was performed on the extracted features in order to infer the discrimination between the two distribution class. Table 3.2 gives the p -value computed for extracted features from reconstructed VAG signals. The return p -value signify that lower the p-value, the better is the feature as it is different from the feature data. A p -value less than 0.01 is taken to indicate that the difference between the means is statistically highly significant. The p -value signify the discrimination between the two distributions. The p -value for ApEn, SampEn and PeEn are 2.421E-08, 1.7E-05 and 0.0071 which showed high significant among the features. The table 3.2 inferred that features extracted from reconstructed VAG signals were statistically significant. Similarly, table 3.3 and 3.4 shows the mean, standard deviation and p -value of RQA and entropy based features extracted from reconstructed and subband signals respectively. From table 3.2, 3.3 and 3.4, it can be observed that statistically significance are seen in the extracted features. From table it can be also observed that large variation are observed from features extracted from unhealthy class. This variation is due to the presence of the high irregular vibrational components of VAG signals. This is caused mainly due to the irregular movement of the tibia, patella and surrounding tissues and muscles.

TABLE 3.3: Statistical measures of reconstructed VAG signals for RQA parameters

Features	Subband	Healthy Class	Unhealthy Class	p -value
		Mean \pm Std	Mean \pm Std	
RR	B3	0.1765 \pm 0.1772	0.2239 \pm 0.2288	0.4359
	B4	0.2450 \pm 0.2000	0.1990 \pm 0.1306	0.5128
	B5	0.3107 \pm 0.1918	0.2626 \pm 0.1120	0.3392
	B6	0.3177 \pm 0.1659	0.2735 \pm 0.0984	0.2446
	B7	0.3038 \pm 0.1167	0.2608 \pm 0.0686	0.0417
	B3	0.9146 \pm 0.0400	0.9063 \pm 0.0346	0.4799
	B4	0.9414 \pm 0.0294	0.9370 \pm 0.0250	0.6254
DET	B5	0.9859 \pm 0.0089	0.9843 \pm 0.0089	0.4059
	B6	0.9973 \pm 0.0018	0.9966 \pm 0.0018	0.1051
	B7	0.9997 \pm 0.0003	0.9997 \pm 0.0002	0.8926
	B3	1.8505 \pm 0.5405	1.7367 \pm 0.4086	0.4929
	B4	2.0760 \pm 0.5821	1.9426 \pm 0.3438	0.6476
ENTR	B5	2.6444 \pm 0.4331	2.8004 \pm 0.6818	0.4671
	B6	3.4563 \pm 0.6132	3.2964 \pm 0.4037	0.2941
	B7	3.9469 \pm 0.2782	4.0881 \pm 0.3908	0.0723
	B3	5.4733 \pm 3.3934	6.8783 \pm 6.1558	0.3661
	B4	5.7170 \pm 1.9525	7.4314 \pm 7.0469	0.4929
DLL	B5	9.1549 \pm 4.4132	14.6638 \pm 19.9002	0.4671
	B6	15.3268 \pm 8.1079	21.5304 \pm 24.1569	0.2941
	B7	25.1575 \pm 9.4067	32.2625 \pm 22.5412	0.0315

TABLE 3.4: Statistical measures of six entropies extracted from subband signals

Features	Subband	Healthy Class	Unhealthy Class	<i>p</i> -value
		Mean \pm Std	Mean \pm Std	
ApEn	B3	0.6374 \pm 0.0779	0.5768 \pm 0.1772	0.0407
	B4	0.4766 \pm 0.0753	0.4380 \pm 0.1348	0.1192
	B5	0.3367 \pm 0.0974	0.3127 \pm 0.1336	0.2619
	B6	0.1668 \pm 0.0631	0.2048 \pm 0.1287	0.7951
	B7	0.0940 \pm 0.0481	0.1221 \pm 0.0806	0.1802
	B3	0.6392 \pm 0.0794	0.5810 \pm 0.1881	0.0417
	B4	0.4775 \pm 0.0760	0.4414 \pm 0.1366	0.1242
SampEn	B5	0.3372 \pm 0.0978	0.3146 \pm 0.1338	0.2754
	B6	0.1668 \pm 0.0633	0.2053 \pm 0.1296	0.7951
	B7	0.0941 \pm 0.0482	0.1219 \pm 0.0817	0.2125
	B3	6.5880 \pm 0.1234	6.5770 \pm 0.0997	0.8192
	B4	6.3005 \pm 0.1639	6.3561 \pm 0.2216	0.1051
	B5	5.4372 \pm 0.2094	5.5314 \pm 0.2409	0.0629
	B6	4.9574 \pm 0.2361	5.1058 \pm 0.2637	0.0100
ShEn	B7	4.3413 \pm 0.2308	4.4315 \pm 0.2812	0.2013
	B3	6.0097 \pm 0.1074	6.0102 \pm 0.1157	0.7951
	B4	5.7294 \pm 0.4702	5.6884 \pm 0.2323	0.0904
	B5	4.8876 \pm 0.2629	4.9946 \pm 0.3001	0.1096
	B6	4.6020 \pm 0.3174	4.4298 \pm 0.2770	0.0368
	B7	3.8594 \pm 0.2907	3.9519 \pm 0.3363	0.2941
	B3	0.9975 \pm 0.0003	0.9975 \pm 0.0003	0.7951
TsEn	B4	0.9962 \pm 0.0034	0.9965 \pm 0.0008	0.0904
	B5	0.9922 \pm 0.0019	0.9929 \pm 0.0019	0.1096
	B6	0.9895 \pm 0.0034	0.9876 \pm 0.0036	0.0368
	B7	0.9780 \pm 0.0075	0.9797 \pm 0.0070	0.2941
	B3	1.9952 \pm 0.0152	1.9945 \pm 0.0197	0.5468
	B4	1.7056 \pm 0.0142	1.7049 \pm 0.0126	0.2941
	B5	1.3277 \pm 0.0204	1.3291 \pm 0.0203	0.7317
PeEn	B6	1.1460 \pm 0.0289	1.1520 \pm 0.0255	0.3186
	B7	0.9869 \pm 0.0186	0.9926 \pm 0.0155	0.0774

Graphical representation of extracted features from subband signals and reconstructed VAG signals is shown by box-plot. Figure 3.7 represents the box-plot for ApEn. Figure 3.7(a) is the boxplot for ApEn extracted from reconstructed VAG signal while figure (b)-(f) represents the box-plots for subband signals (*B3, B4, B5, B6 and B7*). As observed from figure 3.7, the ApEn

distributions of unhealthy VAG signals marked larger variation with respect to the healthy VAG signals. Similarly, box-plots for other entropy based features and RQA parameters have been shown in figure 3.8-3.16 and similar inference can be observed from these features. Thus, the statistical measures obtained inferred that a significant distinction is observed between healthy and unhealthy reconstructed VAG signals. Considering these results, the extracted features are given as inputs in building an effective classification model.

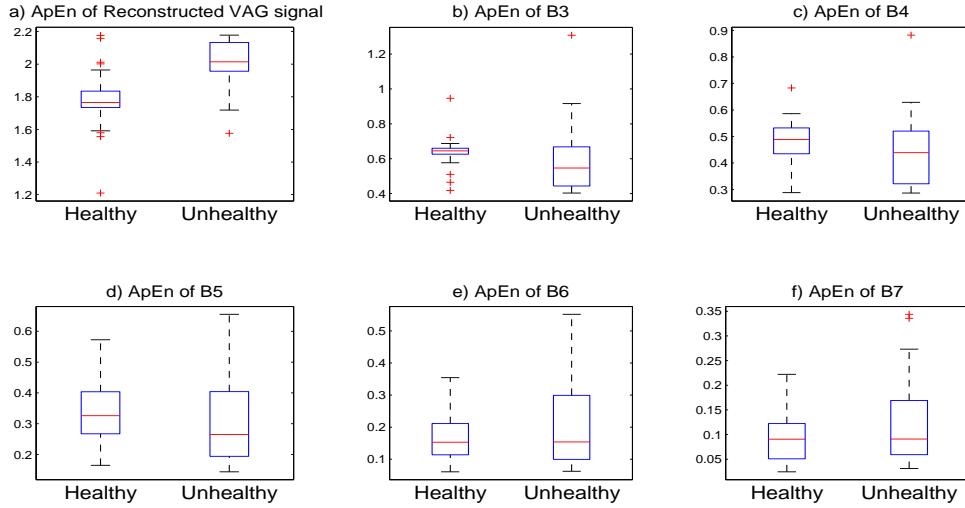


FIGURE 3.7: Boxplot of ApEn

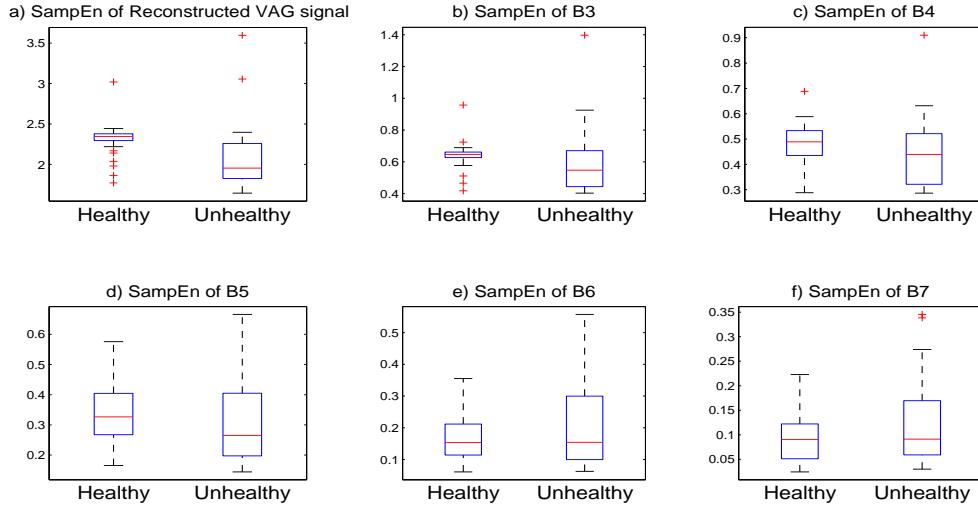


FIGURE 3.8: Boxplot of SampEn

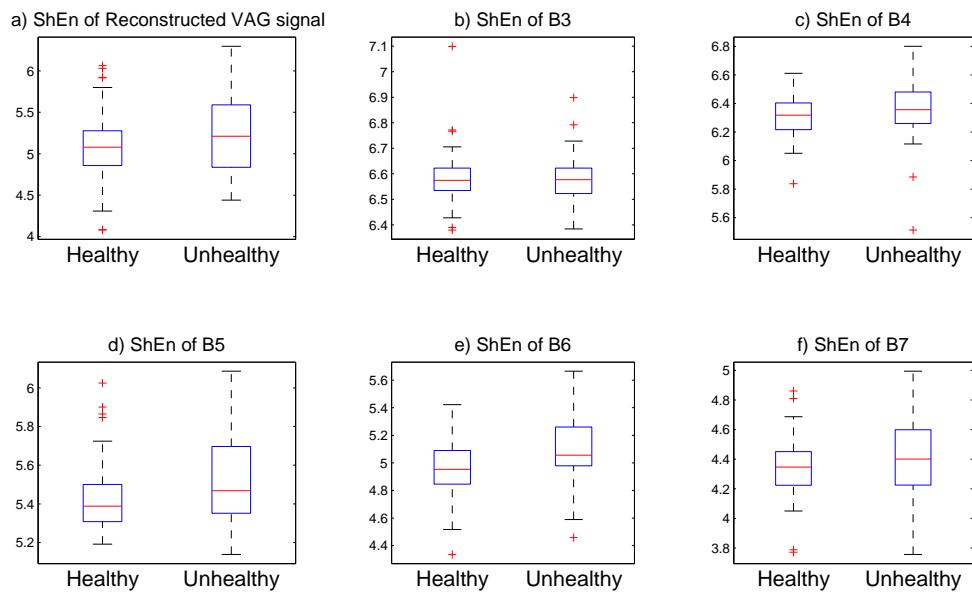


FIGURE 3.9: Boxplot of ShEn

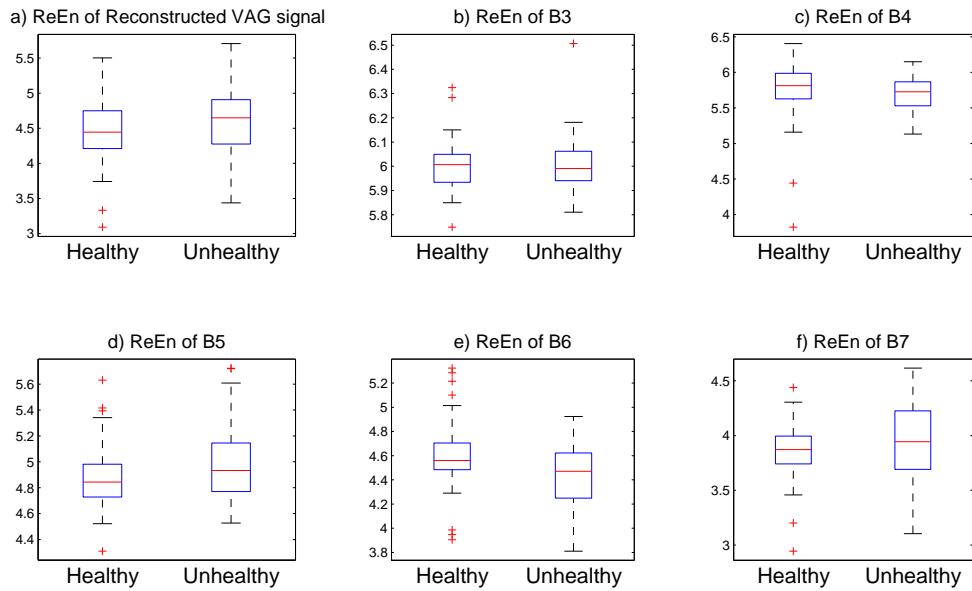


FIGURE 3.10: Boxplot of ReEn

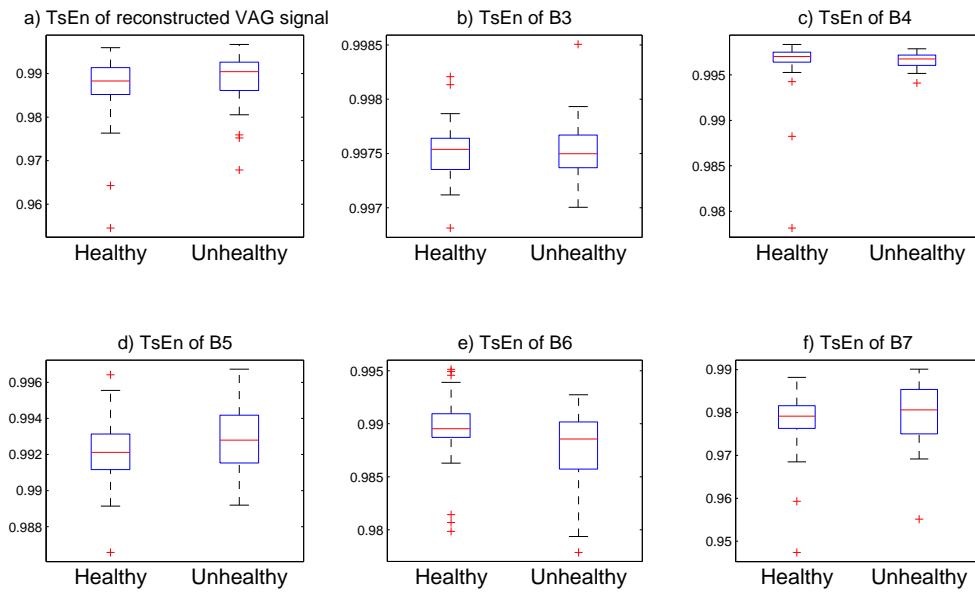


FIGURE 3.11: Boxplot of TsEn

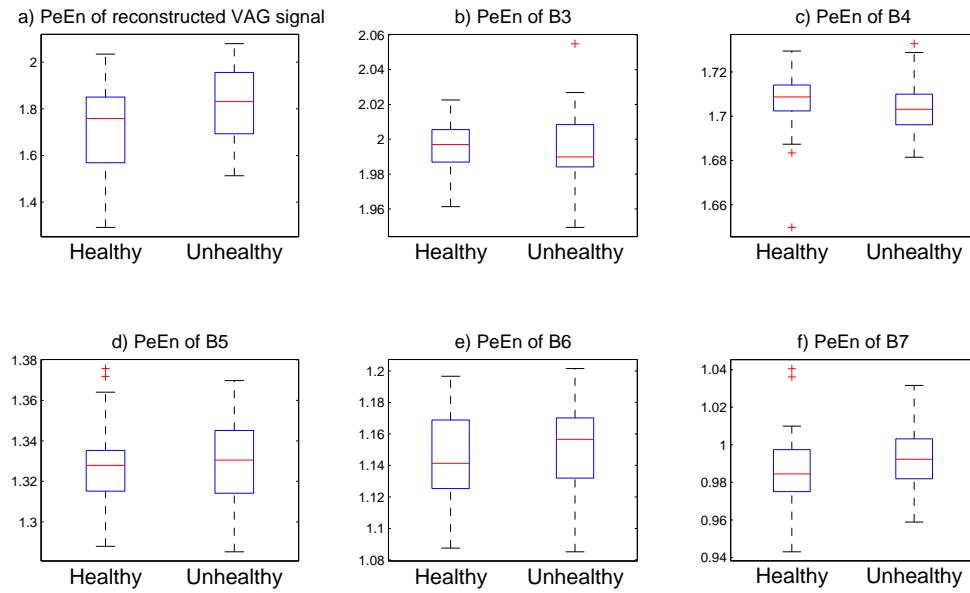


FIGURE 3.12: Boxplot of PeEn

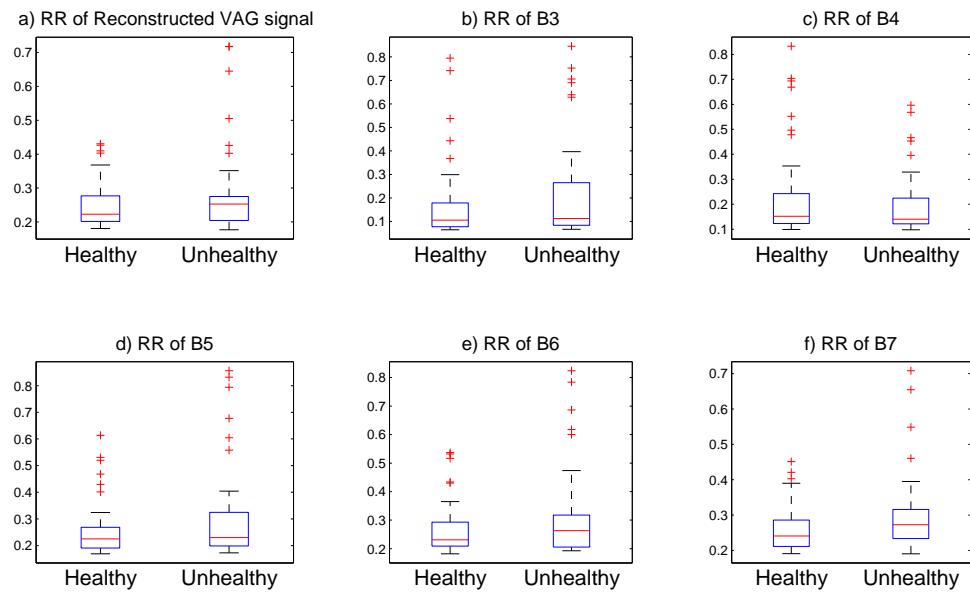


FIGURE 3.13: Boxplot of RQA-RR

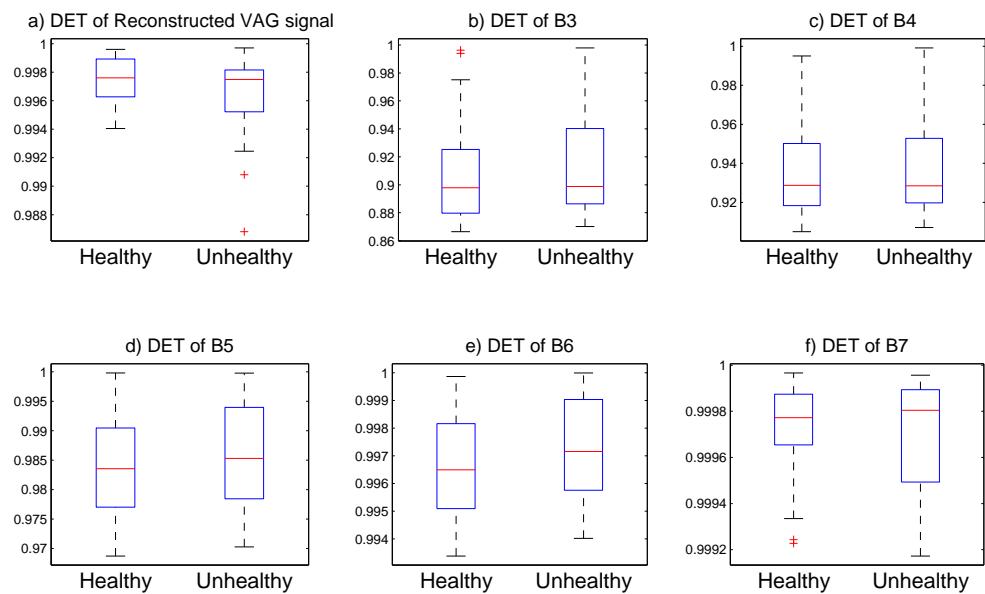


FIGURE 3.14: Boxplot of RQA-DET

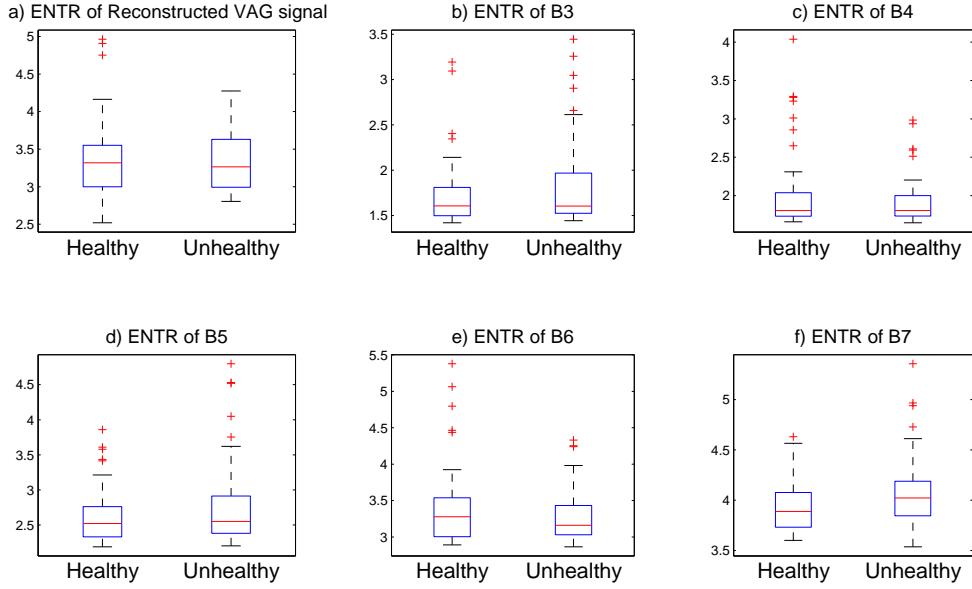


FIGURE 3.15: Boxplot of RQA-ENTR

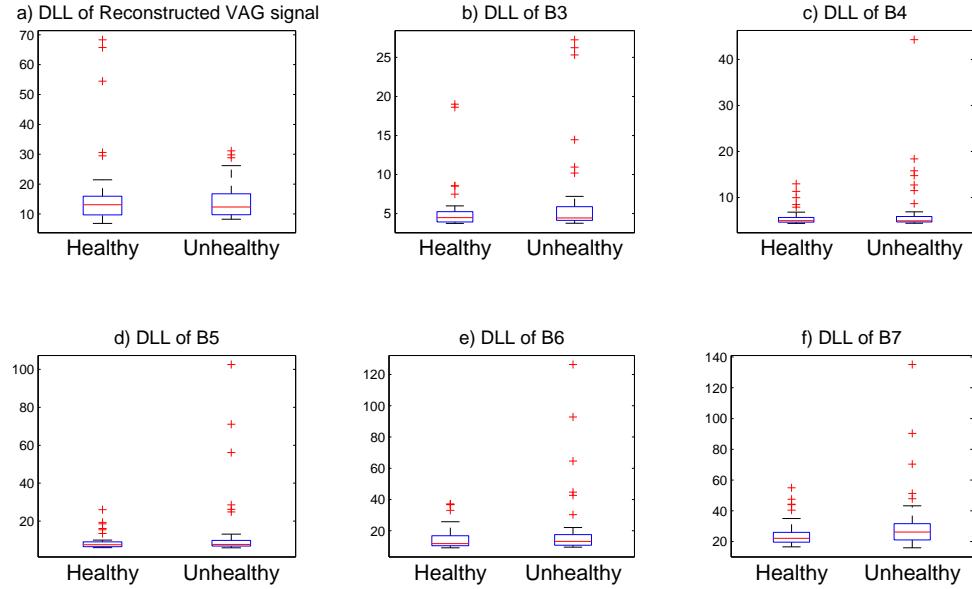


FIGURE 3.16: Boxplot of RQA-DLL

3.4.2 Classification

In this work, LS-SVM has been used for classification of VAG signals. LS-SVM toolbox by Brabanter *et al.* has been used for implementing LS-SVM [164]. Six entropy based features and four RQA parameters extracted from the reconstructed VAG signal and subband signals are

given as inputs to LS-SVM. The dataset consisted of 89 VAG samples. Training set has been set to 60% of healthy and unhealthy VAG signals and remaining signals were considered as a test set. Targets belonging to unhealthy VAG signals are chosen as “one” and healthy signals is set to “zero”. A tenfold cross-validation procedure has been used to evaluate the optimal classification performance of the classifier [165], [166]. Radial basis function (RBF) has been used as a kernel function for the LS-SVM classifier. Command '(tunelssvm)' from “LS-SVMlab Toolbox” was used to determine the tuning parameters (e.g. the regularization and kernel parameters) of the LS-SVM. The tuning parameters of LS-SVM is a two step procedure. Initially, coupled simulated annealing (CSA) algorithm is used for determining the parameters of LS-SVM. In the next step, the parameters are given to “simplex” or “gridsearch” algorithm for determining the optimized regularization constant and kernel parameters. In our study, the classification results were found to be more effective using gridsearch algorithm. The classification has been carried out for 200 iterations and the best classification results were stored. The performance of the classifier was hugely based on area under the receiver operating curve (AUC-ROC), SEN, SPF and ACC. Other parameters such as PPV, NPV, MCC and FDR were also considered for the evaluation purpose.

Table 3.5 gives the classification performance of entropy based features extracted from subband signals (*B3, B4, B5, B6 and B7*). The first row gives the classification performance of ApEn. The ApEn extracted from subband signals (*B3, B4, B5, B6 and B7*) are given as a feature vector to LS-SVM. Hence, the feature vector of ApEn consists of five elements as

$$ApEn = [ApEn \text{ of } B3, ApEn \text{ of } B4, ApEn \text{ of } B5, ApEn \text{ of } B6, ApEn \text{ of } B7] \quad (3.28)$$

TABLE 3.5: Classification performance of extracted features from subband signals

Features	ACC	SEN	SPE	NPV	PPV	MCC	FDR	AUC-ROC
ApEn	0.8315	0.8205	0.8400	0.8101	0.8302	0.8571	0.2000	0.7663 ± 0.0524
SampEn	0.8539	0.8824	0.8158	0.8738	0.8484	0.8378	0.1346	0.7642 ± 0.0526
ShEn	0.6292	0.9216	0.2368	0.7402	0.4672	0.6923	0.3816	0.6780 ± 0.0574
ReEn	0.6404	0.9412	0.2368	0.7500	0.4721	0.7500	0.3766	0.7033 ± 0.0559
TsEn	0.6854	0.9216	0.3684	0.7705	0.5827	0.7778	0.3380	0.6899 ± 0.0577
PeEn	0.8764	0.9412	0.7895	0.8972	0.8620	0.9091	0.1429	0.7920 ± 0.0583
RR	0.7865	0.9423	0.5676	0.8376	0.7313	0.8750	0.2462	0.7848 ± 0.0565
DET	0.7753	0.8085	0.7381	0.7917	0.7725	0.7750	0.2245	0.9187 ± 0.0268
ENTR	0.7978	0.9804	0.5526	0.8475	0.7361	0.9545	0.2537	0.8668 ± 0.0391
DLL	0.7079	0.8431	0.5263	0.7679	0.6662	0.7143	0.2951	0.6367 ± 0.0627
All Features	0.8090	0.9216	0.6579	0.8468	0.7786	0.8621	0.2167	0.8606 ± 0.0397

Similarly, the classification has been carried out for other features and their performance is shown in table no 3.5. The classification has been also carried out by selecting all features as a feature vector. In table 3.5, the “All features” consisted of six entropies based features and four RQA parameters extracted from subband signals. Therefore, the feather vector of “All features” consisted of 50 elements as

$$\text{All Features} = [\text{six entropies from } B3 - B7, \text{ four RQA paramters from } B3 - B7] \quad (3.29)$$

From table 3.5 it can be observed that “All features” gives an accuracy of 80.90% with a sensitivity of 92.12% and specificity of 65.79%. SampEn gives a classification accuracy of 85.39% with lowest FDR of 0.1346. It can be observed that PeEn performed well with respect to other entropies and gives the highest classification accuracy of 87.64% with FDR of 0.1429. Entropies such ApEn and SampEn provided a better pattern classification with an accuracy of 83.15% and 85.39% respectively. From the table 3.5 it can be observed that the performance of entropy based features were better with respect to RQA parameters.

A similar pattern classification has been carried out with features extracted from reconstructed VAG signals. Table 3.6 shows the classification performance of extracted features from reconstructed VAG signals. The feature vector “All features” consists of

$$\text{All Features} = [\text{ApEn}, \text{SampEn}, \text{ShEn}, \text{ReEn}, \text{TsEn}, \text{PeEn}, \text{RR}, \text{DET}, \text{ENTR}, \text{DLL}] \quad (3.30)$$

Therefore, “All features” feature vector consists of ten feature elements (six entropies and four RQA parameters) obtained from reconstructed VAG signals.

TABLE 3.6: Classification performance of extracted features from reconstructed VAG signals

Features	ACC	SEN	SPE	NPV	PPV	MCC	FDR	ROC
ApEn	0.7528	0.7843	0.7105	0.7843	0.7465	0.7105	0.2157	0.7791 ± 0.0514
SampEn	0.8202	0.8235	0.8158	0.8400	0.8197	0.7750	0.1429	0.8540 ± 0.0417
ShEn	0.6629	0.7647	0.5263	0.7222	0.6344	0.6250	0.3158	0.6501 ± 0.0583
ReEn	0.8315	0.9804	0.6316	0.8696	0.7869	0.9600	0.2188	0.6692 ± 0.0588
TsEn	0.7753	0.8163	0.7250	0.8000	0.7693	0.7632	0.2157	0.6011 ± 0.0601
PeEn	0.8132	0.9508	0.5333	0.8722	0.7121	0.8421	0.1944	0.8989 ± 0.0236
RR	0.7978	0.7419	0.8276	0.7188	0.7836	0.8571	0.3030	0.7069 ± 0.0551
DET	0.7528	0.7353	0.7636	0.6944	0.7493	0.8235	0.3421	0.6383 ± 0.0596
ENTR	0.6517	0.7647	0.5000	0.7156	0.6183	0.6129	0.3276	0.7312 ± 0.0528
DLL	0.7079	0.8824	0.4737	0.7759	0.6465	0.7500	0.3077	0.7461 ± 0.0548
All Features	0.8427	0.9020	0.7632	0.8679	0.8297	0.8529	0.1636	0.8560 ± 0.0427

From the table 3.6, it can be observed that “*All features*” gives the highest classification accuracy of 84.27% with FDR of 0.1636. It can be also inferred that entropy based features performed well with respect to RQA parameters. SampEn gives a classification accuracy of 82.02% with low FDR of 0.1429. The classification results inferred that the proposed methodology with feature extraction technique could distinguish between healthy and unhealthy VAG signals.

The performance of the classification model has been measured graphically by AUC-ROC. Figure 3.17 gives the AUC-ROC of features extracted from subband signals. Figure 3.17 (a) gives AUC-ROC plot for entropy based features (b) represents the AUC-ROC plot for RQA parameters. AUC-ROC for “*All features*” was 0.8606 ± 0.0397 (ACC: 80.90% from table no 3.5). The AUC-ROC of other features has been presented in table 3.5. From table 3.5 and figure 3.17 (a) it can be observed that PeEn which gave the highest classification accuracy of 87.64% had AUC-ROC of 0.7920 ± 0.0583 . Though the AUC-ROC of RQA parameter DET was highest of 0.9187 ± 0.0268 but its classification accuracy was 77.53% (Table 3.5). A similar observation can be observed for other RQA parameters.

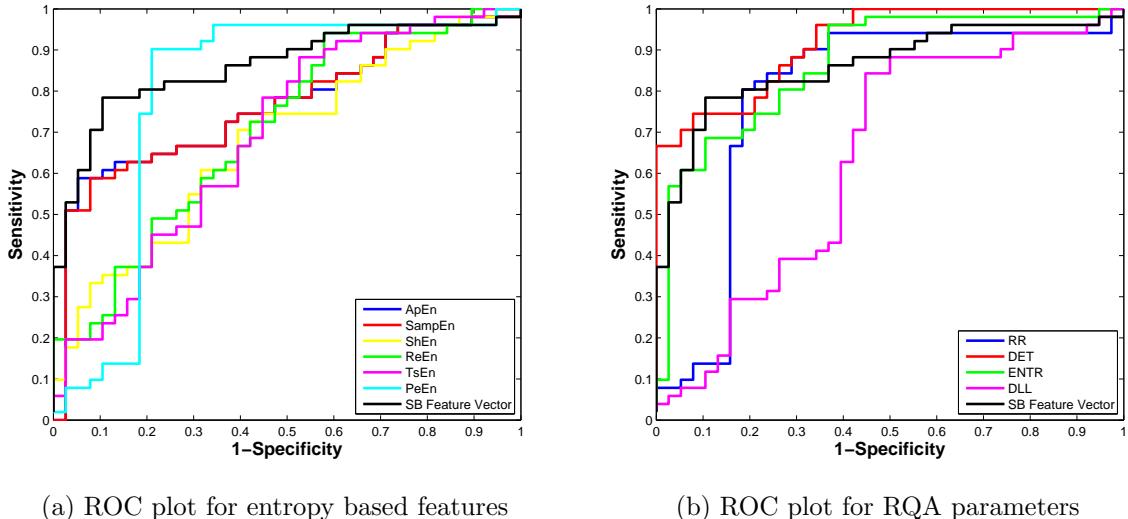


FIGURE 3.17: ROC plot for extracted features obtained from subband signals

Similarly, figure 3.18 gives the AUC-ROC of features extracted from reconstructed VAG signals. while 3.18(a) represents AUC-ROC plot for entropy based features (b) represents the AUC-ROC plot for RQA parameters. From table 3.6 and figure 3.18 (a) it can be observed that the feature vector “*All features*” gave AUC-ROC of 0.8560 ± 0.0427 with a classification accuracy of 84.27%. While entropies performed better with respect to RQA parameters as observed from table 3.6 and figure 3.18. PeEn gives AUC-ROC of 0.8989 ± 0.0236 . From table 3.5 and 3.6 and figure 3.17 and 3.18 it can be observed that LS-SVM as a classifier effectively categorized a larger number of VAG signals between healthy and unhealthy classes.

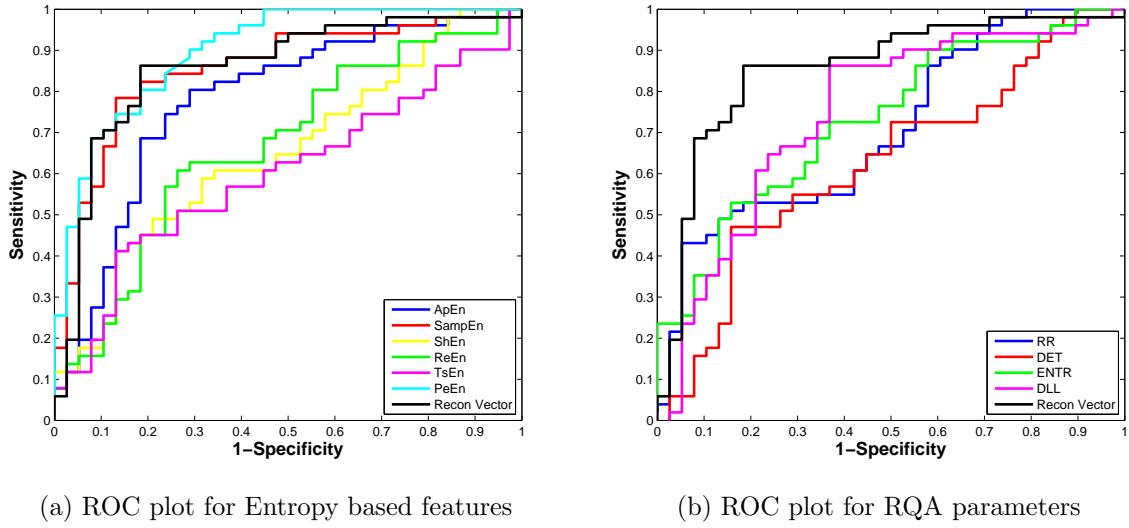


FIGURE 3.18: ROC plot for extracted features obtained from reconstructed signals

3.5 Discussion

The work carried out in this chapter focused on developing a nonstationary methodology for CAD system. The methodology proposed in this chapter has been contributed by the analysis of VAG signal using WPD followed by feature extraction techniques. Six entropies and four RQA parameters has been studied. The VAG signals has been decomposed by WPD into subband signals as observed in figure 3.4. Features has been extracted from subband signals (*B3, B4, B5, B6 and B7*) and reconstructed VAG signals. A total of 50 features has been extracted from subband signal while 10 features has been extracted from reconstructed VAG signals. A total of sixty features were extracted in this study. Statistical analysis has been carried out for the extracted features and are shown in the table 3.2, 3.3 and 3.4. From the table 3.2, 3.3 and 3.4, it has been observed that large variation of statistical measures (mean and standard deviation) exist between healthy and unhealthy classes. For example, from table 3.2, the mean and standard deviation of ApEn (reconstructed VAG signals) for healthy VAG signal is 1.7792 ± 0.1650 while for unhealthy VAG signal it is 2.0113 ± 0.1303 . Furthermore, the graphical representation of the extracted features has been carried out using boxplot. The distribution of features extracted from the unhealthy data was larger in comparison to healthy data as observed in figures 3.5 - 3.16. For example, the box-plot for SampEn extracted from reconstructed VAG signal as observed in figure 3.6 (a), showed that the distribution of unhealthy VAG signals (std:0.3727 from table 3.2) has been larger with respect to healthy VAG signals (std: 0.1908 from table 3.2). Similar observations has been observed from the boxplot of other extracted features. The statistical measures and boxplot of extracted features gave a favourable inference in discriminating healthy and unhealthy VAG signals. This was followed by pattern classification by using LS-SVM.

A detailed classification analysis has been carried out using features extracted from subband signals and reconstructed as observed from table 3.5 and 3.6 respectively. Classification results shows that “All features” (no of features: 50) of subband signals gives a classification accuracy of 80.90% (table 3.5) with AUC-ROC of 0.8606 ± 0.0397 while 84.27% was obtained for reconstructed VAG signals (no of features: 10) with AUC-ROC of 0.8560 ± 0.0427 . The pattern classification for individual features extracted from subband signals and reconstructed VAG signals has been carried out and is shown in table 3.5 and 3.6. It can be observed that entropy based features performed better with respect to RQA parameters. Though PeEn extracted from subband signals (five feature elements) gives the highest classification accuracy of 87.64% but it provided relatively poor AUC-ROC of 0.7920 ± 0.0583 . From this study, it can be concluded that, the feature vector “All features” (no of features: 10) extracted from reconstructed VAG signals performed better in terms of classification accuracy (ACC: 84.27%) and AUC-ROC (0.8560 ± 0.0427).

The proposed methodology provided a better classification accuracy in distinguishing healthy and unhealthy VAG signals. The proposed work provided a better methodology in analysing nonlinear, nonstationary and multicomponent VAG signals using entropy based methods and RQA parameters. Hence, a computer aided diagnostic system could be developed with the proposed methodology for noninvasive diagnosis of knee-joint disorders.

3.6 Conclusion

In this chapter, a nonstationary linear signal processing technique based on wavelet transform has been studied. Initially, the VAG signals were decomposed into subband signals of different frequencies by WPD. The VAG signals were reconstructed by identifying the significant subband signals. Six entropy based features namely ApEn, SampEn, ShEn, ReEn, TsEn, PeEn and four RQA parameters namely recurrence rate (RR), determinism (DET), entropy (ENTR) and diagonal length (DLL) were extracted from subband and reconstructed VAG signals distinctly. Fifty features were extracted from the subband signals (B_3, B_4, B_5, B_6, B_7) and ten from reconstructed VAG signal. The feature vector “All features” (no of features: 10) extracted from reconstructed VAG signals provided better classification accuracy of 84.27% and AUC-ROC of 0.8560 ± 0.0427 in this study. The results illustrated that entropy based features performed better with respect to the features extracted from RQA since they were able to classify the signal between healthy and unhealthy more precisely.

Chapter 4

Analysis Of VAG Signals Using Feature Selection Algorithms

Coming up with features is difficult, time-consuming, requires expert knowledge. “Applied machine learning” is basically feature engineering.

Andrew Ng (2010)

In this chapter, the merits of feature selection algorithm in the analysis of VAG signals has been studied. In any pattern classification problem, features consist of essential properties in representing the data. From the previous chapter, various methodologies have been carried out for extracting features. Most of these extracted features tend to be irrelevant and may lead to over-fitting. Therefore in order to reduce the size of feature set, feature selection algorithm has been presented in this work.

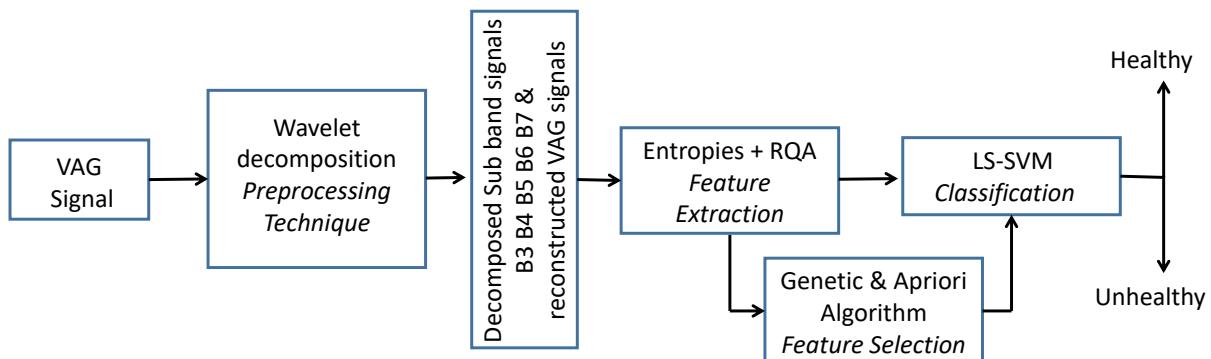


FIGURE 4.1: Flow chart for diagnosis of knee joint disorders using feature selection

This chapter has been organized as follows. Importance of feature selection algorithm is discussed in the first section. Section 4.2 discusses the proposed feature selection algorithm used in this study. Finally, the performance of the feature selection algorithm has been presented in section 4.3 along with the discussion. This chapter concludes with section 4.4.

4.1 Introduction

In pattern classification, a feature represents individual property or characteristics of any data. Extracting essential properties which are distinctive and informative is one of the important steps in describing and choosing the features. In a supervised learning, these features as a “feature vector” helps in building an effective algorithm for modelling classification, regression and pattern recognition. The quality and quantity of the features are important aspects of building a classification model. In the initial study of machine learning and pattern recognition, the number of features was quite restrictive whereas, in the recent development, the number of features has expanded drastically.

Foremost, many of these methodologies suffer from “curse of dimensionality”. That is, as the number of features in a feature vector increases, the computational complexity of the model increases abruptly. One of the prominent causes of this phenomena is that most of these features are irrelevant and insignificant. These insignificant and unessential features hamper the classification model, thereby do not contribute to discriminate among the classes [167]. These features do not provide an effective prediction and thereby increasing the computational and storage complexity. Hence feature selection algorithm provides an aid in reducing the computational complexity, understanding the data, improving the prediction performance and lowering the effects of “curse of dimensionality”. Various techniques have evolved in removing redundant, insignificant and unessential features. The feature selection algorithm provides a solution in building an effective classification model. The feature selection algorithm selects the most stable, significant and discriminant features, thereby selecting an optimal feature subset. The insignificant and unessential features are discarded and efficient classification model is built. As the dimension of the feature vector is reduced, it decreases the computational complexity and increases the performance of the classifiers. This thesis proposes to use genetic algorithm (GA) and apriori algorithm (ApA) as feature selection technique to select the optimal feature subset.

4.2 Methodology

4.2.1 Data Set Preparation

The features extracted from the previous chapter were used for the analysis. The VAG signals were decomposed by WPD into subband signals. Six entropy based features and four RQA based features were extracted from the subband signals. The features were also extracted from the reconstructed VAG signals. Hence a total of 60 features were extracted and they can be summarised as follows

- 1) ApEn from subband signal B3 to B7 and from reconstructed VAG signals.
- 2) SampEn from subband signals B3 to B7 and from reconstructed VAG signals.
- 3) ShEn from subband signals B3 to B7 and from reconstructed VAG signals.
- 4) ReEn from subband signals B3 to B7 and from reconstructed VAG signals.
- 5) TsEn from subband signals B3 to B7 and from reconstructed VAG signals.
- 6) PeEn from subband signals B3 to B7 and from reconstructed VAG signals.
- 7) RQA-RR from subband signals B3 to B7 and from reconstructed VAG signals.
- 8) RQA-DET from subband signals B3 to B7 and from reconstructed VAG signals.
- 9) RQA-ENTR from subband signals B3 to B7 and from reconstructed VAG signals.
- 10) RQA-DLL from subband signals B3 to B7 and from reconstructed VAG signals.

4.2.2 Genetic Algorithm

One of the feature selection techniques used for improving the performance of data mining is the genetic algorithm (GA). Sufficiently good results can be obtained with GA when compared with many nonevolutionary and traditional feature selection methods. The comparison is based on higher classification accuracy and improved instance selection as shown by Cano *et al.* [168]. GA has been successfully implemented for large scale feature selection problems. A solution to a given problem is represented in the form of a string, called “chromosome”. It consists of a set of elements called as “genes”, which holds a set of values of the optimization variables. Once an initial population is created, the individuals of this population are estimated through a fitness function. The fitness function gives the description of the goodness of a chromosome. In a genetic algorithm, the function to be optimized by minimization or maximization the “fitness”. Based on this fitness function, chromosomes are selected and some genetic operators such as

mutation, and crossover are applied over the selected chromosomes. These chromosomes evolve better generation till it reaches the global optimum solution. Very briefly, in a GA basic steps involved are as follows.

1. Population: The values of the single variables are drawn to randomly select a certain number of experimental conditions (chromosomes). Creation of the initial population: a set with N_{pop} chromosomes is generated Cr for the initial population, in which each chromosome is a vector $Cr = [i_1 \dots i_{N_n}]$ containing N_n features, indexes i are randomly generated without repeated elements[116]. In our study 3, 4, 5, 6, 7 and 8 have been chosen as size of the population. An individual's features or genes are contained in each of the chromosome.

2. Objective Function: The applied GA fitness function is based on the principle of Max-Relevance and Min-Redundancy - the objective is that the outputs of the selected features present discriminant power, avoiding redundancy. The application of the principle of Max-Relevance and Min-Redundancy corresponds to searching the set of feature indexes i that satisfies the maximization problem: $\text{Max } \phi$. The maximization of ϕ is a convenient approach to simultaneously maximize V and minimize P .

$\phi = V - P$ is the objective function based on Max-Relevance/Min-Redundancy [169].

$$V = \frac{1}{N} \sum_{i=1}^{N_n} I(yh_i; y) \quad (4.1)$$

is the relevance of a set with N_n features, and

$$P = \frac{1}{N} \sum_{i=1}^{N_n} \sum_{j=1}^{N_n} I(yh_i; yh_j) \quad (4.2)$$

is the redundancy between features.

Notice that “ V ” is the mean value of the mutual information $I(yh_i ; y)$ between extracted features and outputs whereas “ P ” is the mean value of mutual information $I(yh_i; yh_j)$ between extracted features.

3. Crossover: A new population is obtained from the actual population by performing a certain number of “matings” between chromosomes. These are randomly chosen in such a way that the best individuals have a higher probability of being selected. Each gene of a new individual is taken from one of the parents in a random process. Two new chromosomes are created when an exchange of genes takes place during the crossover, and each has a mixture of the experimental conditions which characterize the “parents”. These off springs can then replace some elements of the population to which the two parents belonged. After fitness evaluation, the individuals are organized based on their fitness values by the algorithm, ranking the indexes of the individuals

into a vector, from the best to the worst. In order to create a new generation, the crossover operator is applied. The fitness value gives each individual a probability of reproduction. More adapted individuals (i.e. the individuals in higher positions of the ranking) have more probability of participating in the crossover operation. The blend alpha crossover (BLX- α) crossover genetic operator [170] is defined by:

$$v[k + 1] = \text{round}(\alpha v_1[k] + (1 - \alpha)v_2[k]) \quad (4.3)$$

Where k is the generation, v_1 and v_2 are parents of model wand $\alpha \in [0, 1]$, with normal density probability density function (PDF). The blend alpha crossover (BLX- α) for real vectors creates new offspring by sampling a new value in the range $[min_i - d * \alpha, max_i + d * \alpha]$ at each position i . Here min_i and max_i are the smaller and larger value of the two parents at position i and d is $max_i - min_i$.

4. Termination criteria: When the set of features outputs (y_{h_i}) is mutually exclusive and totally correlated to the target output y , the maximum value is attained. In other words, the idea is to take advantage of the diversity between neurons that were trained by means of different methods. The termination criteria is kept as 80. In this work, the selective pressure parameter, used to speed up the convergence of the GA was set equal to 6.

4.2.3 Apriori Algorithm

In data mining, the apriori algorithm is one of the most influential algorithm. This was proposed by Agrawal *et al.* and is based on the association property [114]. Association rules in data mining are to search the hidden relationships among the attributes. Association rule is a pattern which states that when “X” occurs then the probability of occurrence of “Y” increases [171]. It removes the strange relationship, repeating pattern, association or ordinary links among the dataset.

Let $K = K_1, K_2, \dots, K_m$ be the items in a set and let “R” be the set of transaction “T” in the database. Suppose “A” and “B” be the item sets such that $A \subset K, B \subset K$ and $A \cap B = \emptyset$. The association rules can be applied to an expression such as $A \rightarrow B$. The association rule over the transaction depends on two factor: support and confidence. The rate of “A” contained in “R” is called as support and is denoted as $Sup(A, R)$. It infers the statistical measures of the “R”. Its importance increases with increasing value of statistical measures. The confidence is the rate of both “A” and “B” are contained in “R” and is represented as $Conf(A \Rightarrow B)$ [171]. That is,

$$Conf(A \Rightarrow B) = \frac{Sup(A \cup B, R)}{Sup(A, R)} \quad (4.4)$$

In the data mining, the association rules are applied to the transaction to have specific minimum support ($Minsup$) and confidence ($Minconf$). Agrawal *et al.* proposed Apriori algorithm for association rules in larger databases. Apriori algorithm has two stages. The initial step involves the scanning the database to identify the item set whose support value is greater than $Minsup$. This is followed by implementing the association rules which must follow

$$Sup(A \cup B, R) \geq Minsup \quad \text{and} \quad Conf(A \Rightarrow B) \geq Minconf \quad (4.5)$$

The apriori algorithm is recursive in nature since it scans the database iteratively to find the recurring item. In the first scan elements having minimum support are detected. In the following next scans, the frequent elements found in the proceeding scans are used to the obtained new probable item as candidate sets. In the meanwhile, the support values are computed and the sets which have $Minsup$ become the frequent item set for that particular transaction. The frequent item set is again processed for next transaction as candidate keys. The procedure is repeated unless new frequent items are not identified. The pseudo code of apriori algorithm is presented below.

Algorithm 1: Apriori Algorithm

```

Data:  $C_k$ : candidate item set of size k
Data:  $L_k$ : frequent item set of size k
 $L_1$ =frequent items ;
for  $k = 1; L_k \neq \phi; k++$  do
    do begin ;
     $C_{k+1}$  = candidates generated from  $L_k$  ;
    foreach each transaction  $t$  in database do
        increment the count of all candidates in  $C_{k+1}$  ;
        that are contained in  $t$  ;
         $L_{k+1}$  = candidates in  $C_{k+1}$  ;
        with minsupport ;
    end
end
```

Apriori algorithms are the most popular algorithms used in the first stage of association rule mining. Apriori algorithms were used effectively to find the associations among various application of data mining such as text clustering, image processing and web mining [172]–[175].

4.3 Results & Discussion

Feature selection is used to improve the early diagnosis of knee joint disorders. Sixty features extracted from the subband and reconstructed signals characterises a VAG signal. Among these

sixty features, most of them are insignificant and irrelevant features. These features are also quite large and create a curse of high dimensionality. Thereby, most of these extracted features do not contribute to building an efficient classification model. Thus, to reduce this dimensionality and improve the classification model, genetic algorithm and apriori algorithm as feature selection algorithm have been used. The feature selection algorithm is able to choose an optimal feature subset and thereby reduces the size of the feature set. This simulation has been carried out in MATLAB©2015b.

4.3.1 Feature Selection

4.3.1.1 Genetic Algorithm

The extracted features of 89 VAG signals were given as input to the genetic algorithm. The genetic algorithm was able to choose most significant, relevant and stable features and discard the irrelevant features. Results illustrated that the genetic algorithm has been able to determine eight features which showed the maximum recurrence. The optimal feature set constituted of the following features:

- 1) ApEn of subband B7.
- 2) SampEn of subband B3.
- 3) SampEn of subband B7.
- 4) TsEn of subband B3.
- 5) TsEn of subband B7.
- 6) PeEn of subband B4.
- 7) ApEn of reconstructed signal.
- 8) SampEn of reconstructed signal.

4.3.1.2 Apriori Algorithm

Similarly, the apriori algorithm has been used as feature selection algorithm for sixty features. In this procedure, a *K-means* clustering followed by the apriori algorithm has been applied on the extracted features. *K-means* clustering algorithm divide the each of the features into two classes and these two classes were labelled by single bit encoding. The extracted features belonging to particular class were labelled as zeros and vice versa for the other class. Further analyses were carried out on the encoded feature vectors for recurrences. Features that acted as the primary

key was chosen as most significant, stable and relevant features. The feature set consisted of five features, which showed the maximum recurrence. The compressed feature set constituted the following extracted features:

- 1) ApEn of subband B6.
- 2) ShEn of subband B6.
- 3) PeEn of subband B5.
- 4) ENTR (RQA) of subband B5.
- 5) SampEn of reconstructed signal.

As observed from the features selected from the genetic and apriori algorithm, the entropy based features were more significant with respect to RQA based features. In both feature selection algorithms, ApEn and SampEn provided a significant presence in the optimal feature set.

4.3.2 Classification

LS-SVM toolbox by Brabanter *et al.* has been used for implementing LS-SVM [164]. Training set has been set to 60% of healthy and unhealthy VAG signals and remaining signals were considered as a test set. Targets belonging to unhealthy VAG signals are chosen as 1 and healthy signals are set to 0. A tenfold cross-validation procedure has been used in order to assess the classification performance of the classifier [165]. Radial basis function (RBF) has been used as a kernel function for the LS-SVM classifier. “Gridsearch” algorithm has been used for determining the optimized regularization constant and kernel parameters. The classification has been carried out for 200 iterations and the best classification results were stored. The performance of the classifier was largely based on the area under the receiver operating curve (AUC-ROC), SEN, SPF and ACC. Other parameters such as PPV, NPV, MCC and FDR are also considered for the evaluation purpose. The classification analysis was carried out with the feature vector as follows:

1. Fifty features were extracted from the subband VAG signals which consists of 50 features.
2. Ten features were extracted from reconstructed VAG signals which consists of 10 features.
3. A combined feature vector consisting of features extracted from subband and reconstructed VAG signals consists of 60 features.
4. Features selected by GA consists of eight features
5. Features selected by apriori algorithm consists of five features

TABLE 4.1: Comparison of classification performance of FS algorithm

Attributes	All features vector	Features subband	Features Recon	FS from GA	FS from ApA
No of features	60	50	10	8	5
ACC	0.6966	0.8090	0.8427	0.8202	0.8539
SEN	0.58882	0.9216	0.9020	0.7254	0.9272
SPE	0.8421	0.6579	0.7632	0.9473	0.7352
NPV	0.6038	0.8468	0.8679	0.72	0.862
PPV	0.8333	0.7786	0.8297	0.9487	0.85
MCC	0.4337	0.8621	0.8529	0.6707	0.6868
FDR	0.1667	0.2167	0.1636	0.0512	0.15
AUC-ROC	0.7822 ± 0.0471	0.8606 ± 0.0397	0.8560 ± 0.0427	0.8754 ± 0.0347	0.9252 ± 0.0261

Table 4.1 shows the classification performance among all features vector (60*89), features extracted from subband signals of VAG signals (50*89), features extracted from the reconstructed VAG signal (10*89) and optimal feature set selected from GA (8*89) and apriori algorithm (5*89). From the table 4.1, it can be observed that “*All feature vector*” which consisted of sixty feature subset gives a classification accuracy of 69.66%. The classification accuracy of feature vector “*Features SB*” of subband signals gives an accuracy of 80.90% while the classification accuracy of feature vector “*Features Recon*” extracted from reconstructed VAG signals gives 84.27%. The highest classification accuracy has been obtained for the features selected by the apriori algorithm with 85.39% which consisted of five features only. GA selected eight features as an optimal feature set which gives a classification accuracy of 82.02%. The optimal feature set selected by the feature selection algorithms provided better classification performance as observed from the table 4.1. It can be seen from the table that, as the number of features decreases the classification accuracy increases.

The classification performance of the five feature vectors has been graphically represented by AUC-ROC. Figure 4.5 gives the comparative AUC-ROC for the optimal feature set and other feature vectors. The feature vector “All feature vector” gives AUC-ROC of 0.7822 ± 0.0471 which is lowest among the other feature vectors. The feature vector “Features SB” and “Features Recon” gives AUC-ROC of 0.8606 ± 0.0397 and 0.8560 ± 0.0427 respectively. Feature selected by GA gives AUC-ROC of 0.8754 ± 0.0347 while highest AUC-ROC has been obtained from features selected by the apriori algorithm (0.9252 ± 0.0261). Therefore, as observed from the table the AUC-ROC of both feature selection algorithm performed well with respect to other feature vectors. The classification analysis carried out provided the inference that feature selection plays an important role in building effective classification by discarding the irrelevant and redundant features.

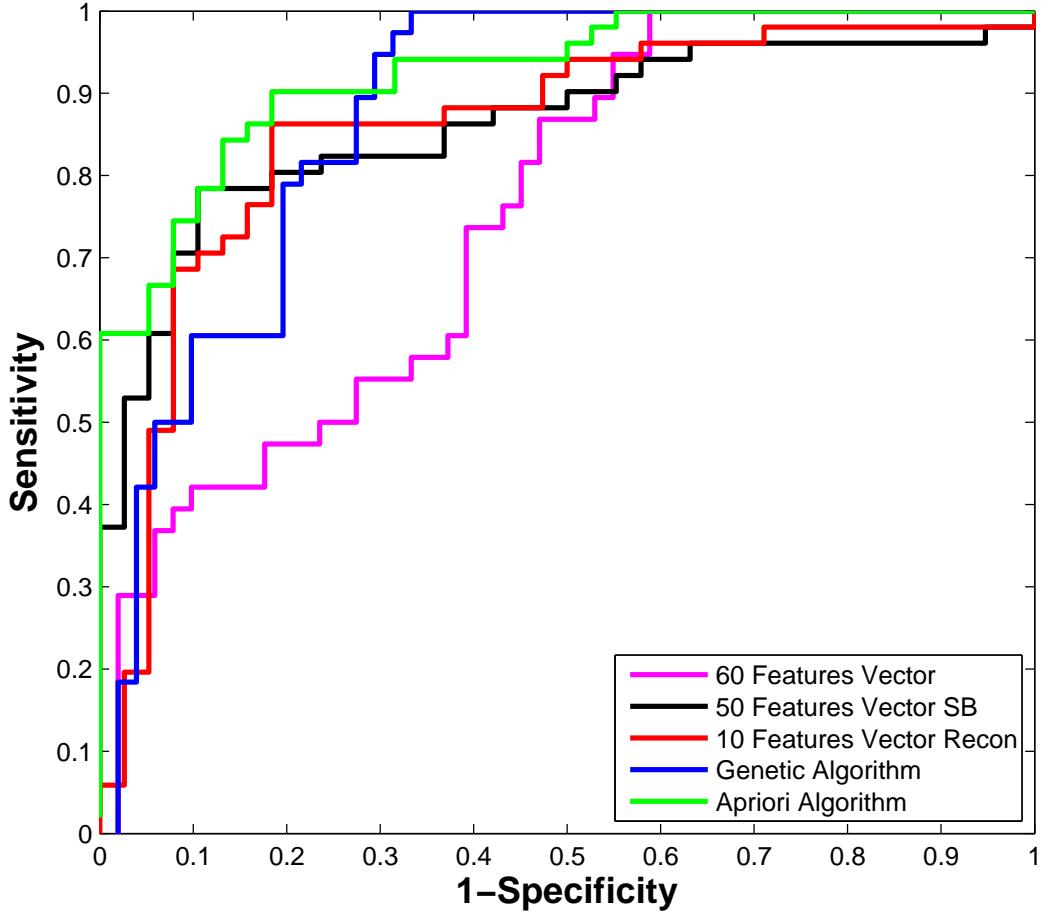


FIGURE 4.2: ROC plot of features selected from FS

4.4 Conclusion

In this chapter feature selection algorithm has been studied to discard the irrelevant and redundant features. To select the most significant, relevant and stable features, two feature selection algorithm namely genetic and apriori algorithm has been studied. The performance of the feature selection algorithm was evaluated by LS-SVM. The results illustrated that feature selected by feature selection algorithms performed better with respect to other feature vectors. The genetic algorithm selected eight features namely ApEn of subband B7, SampEn of subband D3, SampEn of subband D7, TsEn of subband D3, TsEn of subband D7, PeEn of subband D4, ApEn and SampEn of the reconstructed signal. While apriori algorithm selected five features which consisted of ApEn of subband B6, ShEn of subband B6, PeEn of subband B5, ENTR (RQA) of subband B5 and SampEn of the reconstructed signal. Furthermore, the features selected from the feature selection algorithms also inferred that, entropy based features were prominent features with respect to RQA parameters. The highest classification of 85.39% and

AUC-ROC of 0.9252 ± 0.0261 was obtained by five features selected by apriori algorithm. The proposed methodology in this works could find a potential application in the knee joint pathology for early detection and monitoring.

Chapter 5

Analysis of VAG Signals Using Complete Ensemble Empirical Mode Decomposition With Adaptive Noise

A necessary condition to represent nonlinear and nonstationary data is to have an adaptive basis. A priori defined function cannot be relied on as a basis, no matter how sophisticated the basis function might be.

Norden E. Huang (2005)

From the previous chapters it has been concluded that VAG signal is multicomponent, nonlinear and nonstationary in nature. The analysis carried out by traditional signal processing cannot provide proper inferences of VAG signals since most of the traditional signal processing techniques are based on Fourier. These techniques assume that the system is linear and input data is stationary or periodic in nature and therefore improper inference could be assumed using these techniques. In this chapter, nonlinear signal processing techniques such as empirical mode decomposition (EMD) based techniques have been studied for the analysis of VAG signals. In the previous chapter, it has been inferred that, entropy based features were prominent features with respect to RQA based parameters. Therefore, entropy based feature extraction techniques has been continued to be focused in this study. Additionally, a simple feature extraction technique based on central tendency measure (CTM) has been computed.

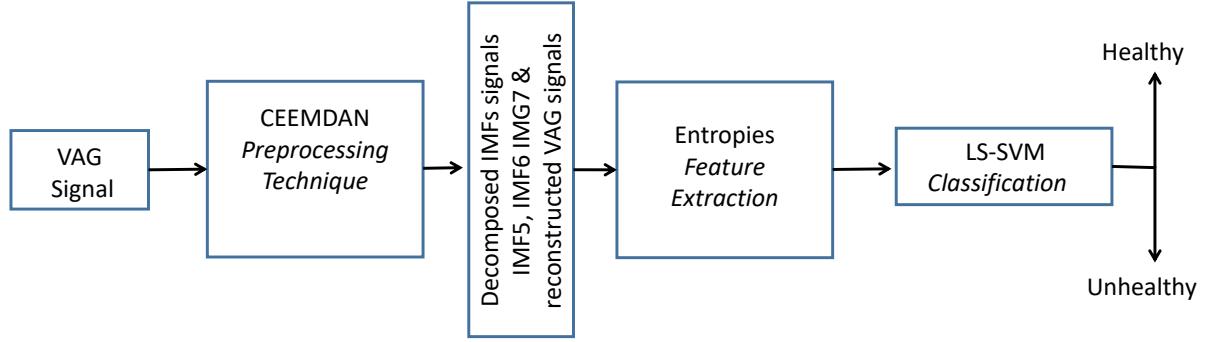


FIGURE 5.1: Flow chart for diagnosis of knee joint disorders using CEEMDAN

The chapter has been organized as follows: section 5.1 introduces the techniques of nonlinear signal processing techniques. The next section briefly describes the methodology studied in this chapter and followed by results and discussion section. The chapter concludes with section 5.4.

5.1 Introduction

The characteristics obtained from VAG signals could be utilized in discriminating and detecting abnormalities in VAG signals. Researchers have successfully carried out the analysis of VAG signals by various linear signal processing techniques. In the previous chapter, the analysis of VAG signal was based on wavelet analysis which is essentially a Fourier analysis with an adjustable window function. It is a linear method but applicable on nonstationary signals. The VAG signals were decomposed into subband signals by WPD. But these subband signals were still contaminated with artefacts. These artefacts included external noise, muscle contraction interference, baseline wanders etc. These artefacts could mislead in inferring the results. Considering these drawbacks and the characteristics of VAG signals, nonlinear signal processing techniques have been explored in this work. The literature survey also revealed that nonlinear techniques accomplished by feature extraction techniques could improve in discriminating between healthy and unhealthy VAG signals.

Huang *et al.* introduced a nonlinear and nonstationary signal processing technique known as the empirical mode decomposition (EMD) [44]. EMD has been applied for the data analysis of a large number of an application involving nonstationary and nonlinear data such as biomedical, finance, physical and geological engineering. Wu *et al.* applied EMD for the analysis of VAG signal for detecting chondromalacia patellae [45]. However, EMD has one major disadvantage of “mode mixing”. In mode mixing, a particular IMF is composed of signals with widely disparate scales, or a signal identical scale may reside in different IMF components. To overcome this problem “ensemble empirical mode decomposition” (EEMD) was proposed [46]. Wu *et al.* applied EEMD for the analysis of VAG signal [47]. EEMD has been studied as an improved preprocessing

technique for VAG signals. EEMD overcomes the problem of “mode mixing” by adding white noise along with the original signal. In computing EEMD, EMD is performed on the original signal but the true IMFs are obtained through an ensemble mean.

In this study, a complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) for the analysis of VAG signal [52] has been covered. CEEMDAN provides an improvement in minimizing the reconstructed error, providing a fixed number of modes for different signals and minimizes residual noise present in the mode [52]. CEEMDAN has been successfully reported in the analysis of nonstationary [54], [176] and nonlinear data, especially in the application related to biomedical signals [52]. Therefore, in this chapter CEEMDAN as a preprocessing technique has been presented. Entropy, which is a quantified measure of the complexity is investigated. Entropy based features as studied in the previous chapter namely ApEn, SampEn, ShEn, ReEn, TsEn and PeEn has been used in this chapter as feature extraction techniques. In the proposed work, the VAG signals are decomposed into intrinsic mode functions (IMF) by CEEMDAN. The VAG signals are reconstructed by identifying the significant IMFs and entropy based features are extracted. Finally, the extracted features are given to classifier for evaluating their performance. LS-SVM as a classifier is proposed in the present study. Moreover, the time complexity of feature extraction techniques has been computed and compared. Hence, the computational complexity of the entire proposed system has been computed. The proposed methodology could provide an aid in developing a noninvasive and low cost diagnostic system for knee joint disorders.

5.2 Methodology

5.2.1 CEEMDAN

Empirical mode decomposition (EMD) is a data-dependent approach suitable for decomposition of a nonlinear and nonstationary signal into symmetric, amplitude and frequency modulated (AM-FM) components known as intrinsic mode functions (IMFs) [44]. IMFs of the signal is considered when following two conditions are fulfilled: (a) the number of maxima/minima (extrema) and the number of zero crossing should be equal or differ by one (b) the mean of upper and lower envelopes (local mean) must be zero. In the case of EEMD, the IMFs are computed as the mean of the corresponding IMFs obtained from an ensemble of the original signal with additional finite variance white noise [46]. EEMD overcomes the problem of “mode mixing” as EEMD is a multiple trial process and each trial has the a procedure similar to EMD, except that the input signal is a mixture of the original signal and a finite Gaussian white noise [46]. The scale separation capabilities of EEMD enables to eliminate the problem of mode mixing. But there are some disadvantages of EEMD and they are as follows (i) the decomposition of the signal is usually incomplete and residual noise are attached in the modes (ii) different modes are

produced by the addition of noise with the signal. The number of sifting process increases with the increase in the number of trials [52]. In order to overcome these disadvantages, CEEMDAN has been reported.

The steps for CEEMDAN are as follows: Let $R(\cdot)$ be the operator producing local means of the signal. $U_k(\cdot)$ be the operator producing $k_{(th)}$ mode obtained by the EMD. And let $v^{(i)}$ be the white Gaussian noise added having zero mean and variance as 1 [52].

- 1) The first residue is computed from local means of I realization $x^{(i)} = x + \beta U_1(v^{(i)})$ by EMD

$$r_1 = \langle R(x^{(i)}) \rangle \quad (5.1)$$

- 2) The first mode is calculated for $k=1$

$$d_1 = x - r_1 \quad (5.2)$$

- 3) The second residue is computed as the average of local means $r_1 + \beta U_2(v^{(i)})$ and second mode is computed as

$$d_2 = r_1 - r_2 = r_1 - \langle R(r_1 + \beta_1 R_2(v^{(i)})) \rangle \quad (5.3)$$

- 4) The K_{th} residue is calculated as

$$\langle R(r_{k-1} + \beta_{k-1} U_k(v^{(i)})) \rangle \quad (5.4)$$

- 5) k_{th} mode is evaluated as

$$d_k = r_{k-1} - r_k \quad (5.5)$$

- 6) repeat from step 4 for next k.

The CEEMDAN algorithm's parameters are optimized by setting the amplitude of added Gaussian noise and complete ensemble number by statistical rule suggested by Huang [46].

$$\beta_n = \frac{\beta}{\sqrt{EN}} \quad (5.6)$$

Here EN is the ensemble number, β is the added Gaussian noise and β_n std (standard deviation) of error which is the difference between the input data and its IMFs. The stopping criteria can be set either by means of energy parameter or by fixing the shifting iteration. Hence, in the present study, the following parameters are set as

- a) The complete ensemble number is set as $EN = 100$. This was the mean of IMFs obtained from 100 trials of CEEMDAN.
- b) 0.2 times the standard deviation (SD) of VAG signal is fixed as added Gaussian noise.
- c) The maximum number of shifting iteration NS is set to 5000.

5.2.2 Hilbert-Huang Transform

N.E Huang *et al.* has proposed HHT method for effectively analysing the non-stationary and aperiodic signals [44]. It consists of two part: Obtaining IMFs using EMD and then applying the Hilbert transform. The most important part of HHT is decomposing the signal into a finite number of distinct oscillations called IMFs. In this study, IMFs have been computed using CEEMDAN.

5.2.2.1 Hilbert Transform

Inorder to obtain the analytical representation of the signal, instantaneous phase and amplitudes is computed by applying Hilbert transform to each IMF obtained from EEMD. For an arbitrary time series data $x(t)$, its Hilbert transform $v(t)$ as

$$v(t) = \frac{1}{\pi} C.V \int_{-\infty}^{\infty} \frac{v(\tau)}{(t - \tau)} d\tau \quad (5.7)$$

Where $C.V$ shows the principal of Cauchy value. The analytic signal of $w(t)$ is defined as follows:

$$z(t) = x(t) + jv(t) = Z(t)e^{j\phi(t)} \quad (5.8)$$

The amplitude $Z(t)$ and instantaneous phase $\phi(t)$ are defined as

$$Z(t) = \sqrt{x(t)^2 + v(t)^2} \quad (5.9)$$

$$\phi(t) = \arctan \frac{v(t)}{x(t)} \quad (5.10)$$

The instantaneous frequency $\omega(t)$ is obtained by differentiation:

$$\omega(t) = \frac{d(\phi(t))}{dt} \quad (5.11)$$

HHT has been successfully applied for the application related to geophysical studies [49], biomedical studies, vibration fault tolerance etc. [72], [177].

5.2.3 Feature Extraction

5.2.3.1 Entropy Based Feature Extraction

Entropy based features were considered for this study. Six entropy such as ApEn, SampEn, ShEn, ReEn, TsEn and PeEn were extracted from the reconstructed VAG signals. The details explanation of entropy based features has been presented in chapter 3.

5.2.3.2 Central Tendency Measurement (CTM)

CTM is a method to sum up the observable data in the plots [94]. CTM is used to quantify the variability of the signal. CTM is computed by initially finding the circular region of particular radius ' r '. Then, dividing the number of data points existing within that circular region by the total number of data points. In the analytical representation of IMFs, the real part of the signal is plotted against the imaginary part of the signal in *z-plane*. CTM is exploited to compute the circular region, which is the area of the analytic representation of IMF in the *z-plane*. For computing CTM, let r be the radius of the central area and N be the total number of points [94]. Then, the CTM for analytic signal $w[t]$ is given as

$$CTM = \frac{\sum_{n=1}^N \Delta(d_i)}{N} \quad (5.12)$$

$$\Delta(d_i) = \begin{cases} 1 & \text{if } \sqrt{\text{real}(w[t])^2 + \text{imag}(w[t])^2} < r \\ 0 & \text{otherwise} \end{cases} \quad (5.13)$$

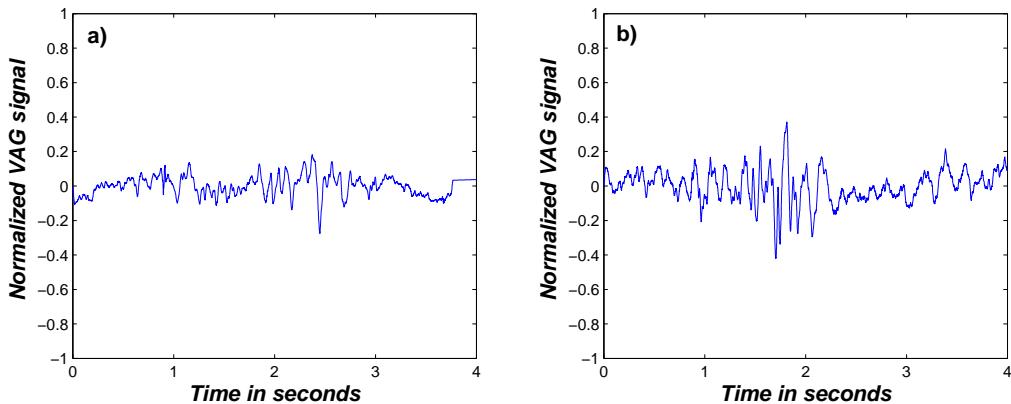


FIGURE 5.2: VAG signal of (a) healthy subject and (b) unhealthy subject

5.3 Results

The current study comprising of preprocessing (CEEMDAN), feature extraction and classification technique has been implemented in MATLAB 2015a. Figure 5.2 (a) and(b) shows the VAG signals of a healthy and unhealthy subject respectively as a sample. The CEEMDAN decomposes VAG signals into a set of IMFs as observed in figure 5.3. From figure 5.3 it can be observed that different IMFs reveal different degrees of dynamics involved in the VAG signal. IMF1 - IMF3 represents the high frequency oscillation while IMF4 - IMF12 represents the low frequency oscillation. The high frequency oscillation IMFs (IMF1 and IMF2) have flat envelopes and thereby do not represent pathological conditions of knee-joint disorders [6] while low frequency oscillation IMFs (IMF9 and IMF10) are the external noise or artefacts.

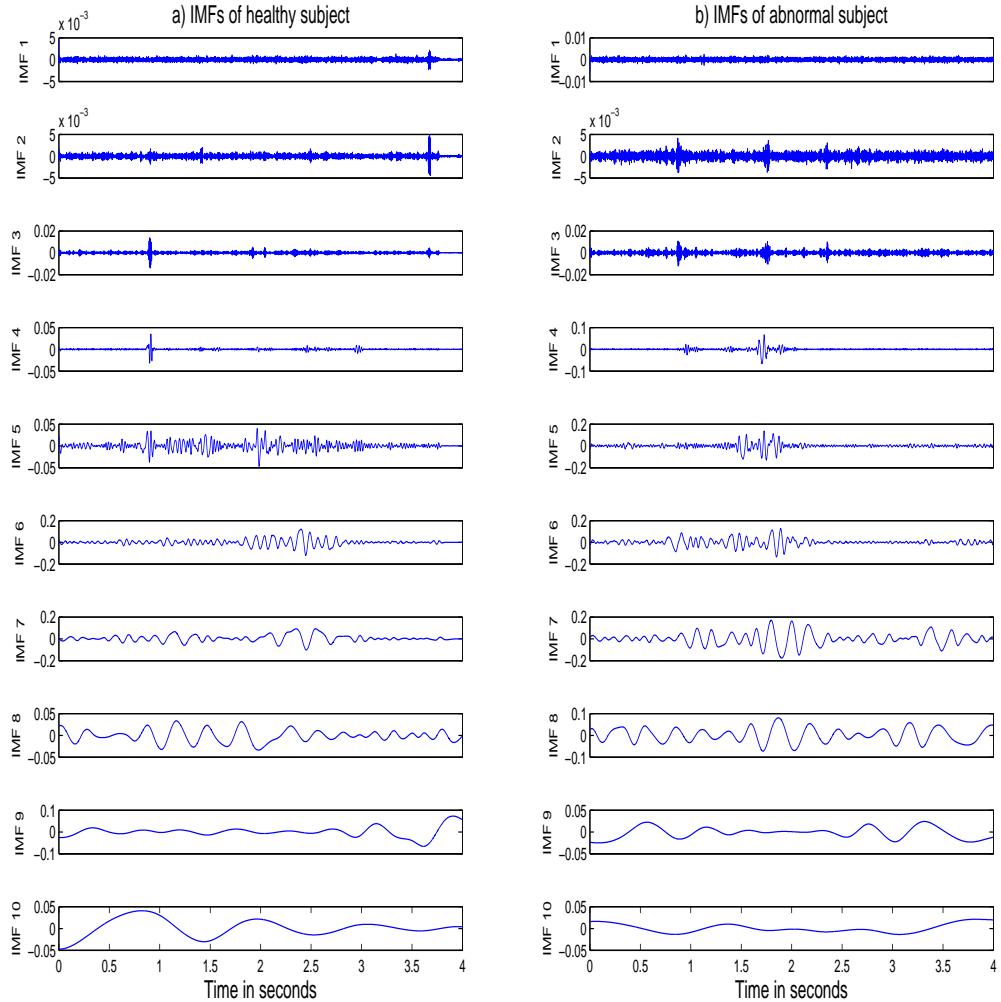


FIGURE 5.3: IMFs obtained from VAG signal of (a) healthy subject and (b) unhealthy subject

The detrended fluctuation analysis (DFA) algorithm is computed to analyse the correlation properties of each decomposed IMFs [6]. IMFs which contain the prominent information are identified based on the computed fractal scaling index. As suggested from the study by Wu [6], the IMFs with anti-correlation value between ($0 < \alpha < 0.5$) and baseline wander related IMFs (IMF9-IMF10) have been ignored. Therefore, IMF5, IMF6 and IMF7 are considered as significant data of VAG signal. The obtained α values for IMF5, IMF6 and IMF7 are 1.302, 1.622 and 1.88 respectively. The VAG signals are reconstructed from the summation of these IMFs (IMF5, IMF6 and IMF7). Figure 5.4 (a) and (b) shows the reconstructed VAG signals for healthy and unhealthy subject respectively.

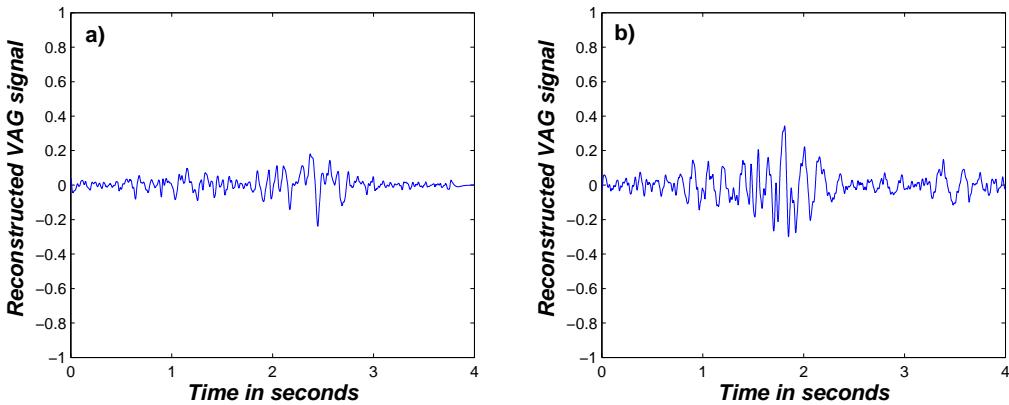


FIGURE 5.4: Reconstructed VAG signal of (a) healthy subject and (b) unhealthy subject

5.3.1 Feature Extraction

5.3.1.1 Entropy

Entropy based features are extracted from reconstructed VAG signals. These features are extracted for 38 unhealthy and 51 healthy VAG signals. Table 5.1 shows the mean and standard deviation of entropy based features of reconstructed VAG. From table 5.1 it can be observed that entropy based features are relatively high for unhealthy VAG signals with respect to healthy VAG signals. The mean and standard of ApEn for healthy and unhealthy signals are 0.000178 ± 0.00019 and 0.000301 ± 0.00023 respectively. As observed for feature ApEn, the mean of unhealthy class is 1.7 times the healthy class. The large variation observed in features extracted from unhealthy VAG signals is due to the presence of the high irregular vibrational components of VAG signals. This is caused mainly due to the irregular movement of the tibia, patella and surrounding tissues and muscles. In case of unhealthy VAG signals, a comparatively larger friction is created between the bones and surrounding tissues, thereby restricting the smooth movements of the knee joint and containing multiple component vibrations. The features extracted from the

VAG signals signifies these characteristics. Similar inference can be observed for the rest of the features.

TABLE 5.1: Statistical measures of reconstructed VAG signals for entropy based features

Features	mean \pm standard deviation		
	healthy	unhealthy	p-value
ApEn	0.000178 ± 0.000191	0.000301 ± 0.00023	0.012657
SampEn	0.000182 ± 0.000201	0.000299 ± 0.000221	0.014217
ShEn	0.239857 ± 0.165505	0.321193 ± 0.161272	0.009697
ReEn	0.053276 ± 0.041809	0.073854 ± 0.041279	0.002975
TsEn	0.051084 ± 0.038967	0.070421 ± 0.038454	0.015059
PeEn	0.888321 ± 0.034445	0.967527 ± 0.024158	0.000429

Kruskal-Wallis test has been performed on the extracted features in order to infer the discrimination between the two distribution class. Table 5.1 gives the *p*-value computed for extracted features from reconstructed VAG signals. The *p*-value obtained for the extracted features provided discrimination among two classes. For example, the *p*-value obtained for ApEn and SampEn were 0.012657 and 0.014217 respectively. The *p*-value and the statistical analysis inferred that there is significant discrimination in the distribution of features extracted between healthy and unhealthy classes.

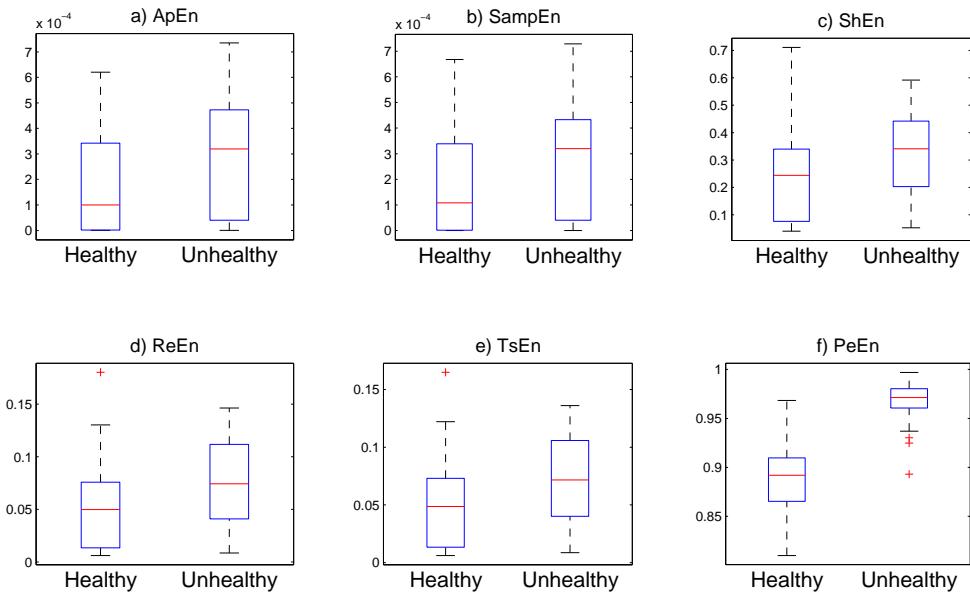


FIGURE 5.5: Boxplots of VAG signal of both healthy and unhealthy

Graphical display of the distribution of entropy based features extracted from the reconstructed VAG signals has been presented as boxplot in figure 5.5. Figure 5.5 shows the boxplot for a) ApEn

b) SampEn and c)ShEn d) ReEn e) TsEn and f)PeEn respectively. As observed from boxplot, unhealthy distributions of extracted features show a comparatively large variation with respect to the class of healthy VAG signals. Thus, the results obtained from the statistical analysis provided a significant distinction between healthy and unhealthy VAG signals. Considering these results, the extracted features are given as inputs in building an effective classification model.

5.3.1.2 CTM

Hilbert transform is applied to the VAG signal, IMFs (5,6 and 7) and reconstructed VAG signal in order to represent it in analytical form. It is obtained by plotting the imaginary part of the signal against the real part of the signal in a *z-plane*. Figure 5.6 shows the analytical representation of raw VAG and reconstructed VAG signal while figure 5.7 shows the analytical representation of individual IMFs. From figure 5.6 (a) and (b) it can be observed that analytical representation of VAG signals does not possess any geometrical shape. As a result, it is quite difficult to analyse the raw VAG signal. These irregular shapes signify the presence of artefacts and multicomponent signals. Whereas in figure 5.6 (c) and (d), prominent circular curves are observed for reconstructed VAG signal. Similar circular curves are observed for analytical

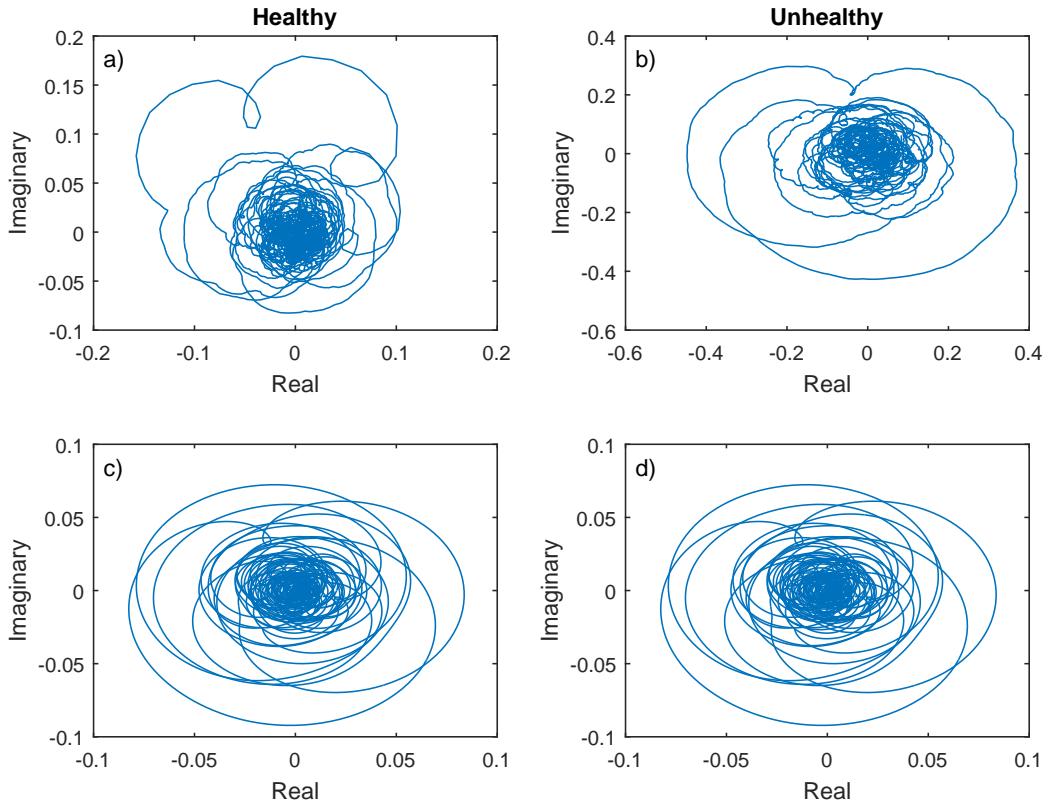


FIGURE 5.6: Analytical representation of VAG signal (a) healthy subject (b) unhealthy subject and reconstructed VAG signal of (c) healthy subject and (d) unhealthy subject

representation of IMFs (5,6 and 7) in figure 5.7. These circular curves possess proper individual center for a particular circular curve. It can be also observed that there are many circular curves having individual centers. This might indicate the presence of distinct frequency components in the VAG signal. Considering these observation, the uniqueness described in the analytical representation of reconstructed VAG signal has been extracted as features.

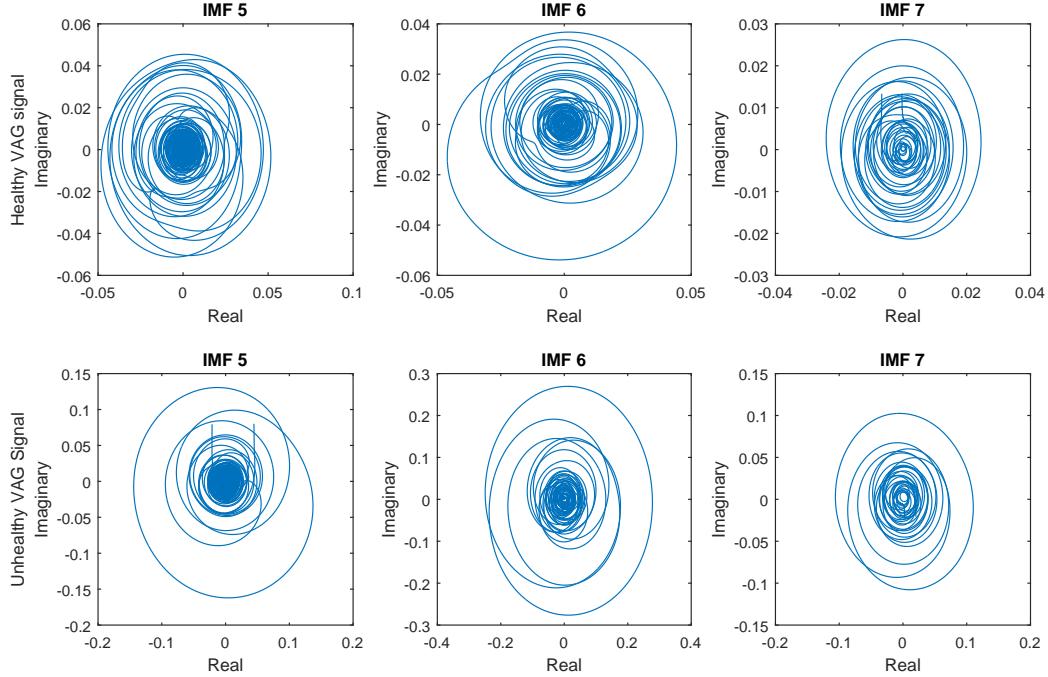


FIGURE 5.7: Analytical representation of IMF5, IMF6 & IMF7 for healthy and unhealthy subject

In order to extract the feature from the analytical representation of reconstructed VAG signals, the radius of the circular curve of the reconstructed VAG signal in the *z-plane* has been identified. Then, the area of the reconstructed VAG signal in the *z-plane* is computed with 95% CTM utilization of data points. That is, the number of data points enclosed within the area of the circle is computed. This provides a significant difference in surface area between unhealthy and healthy VAG signals. The area parameter as a feature has been computed for 89 reconstructed VAG signals. The computed area for healthy subjects were found to be smaller in comparison with the area computed for unhealthy subjects. The larger area could signify the presence of high amplitude for unhealthy subjects.

The statistical analysis of the area as feature computed from the reconstructed VAG signal has been carried out. Table 5.2 provides the mean and standard deviation obtained from feature extracted from reconstructed VAG signals. From the table, it can be observed that the area computed from the unhealthy VAG signal has a larger variation with respect to healthy VAG signals. Similarly, figure 5.8 shows the boxplot of extracted feature. From the boxplot, it can be

Healthy	Unhealthy
mean \pm std dev	mean \pm std dev
0.2369 ± 0.1078	0.2952 ± 0.0894
p-value	
0.0138	

TABLE 5.2: Statistical measure
for CTM

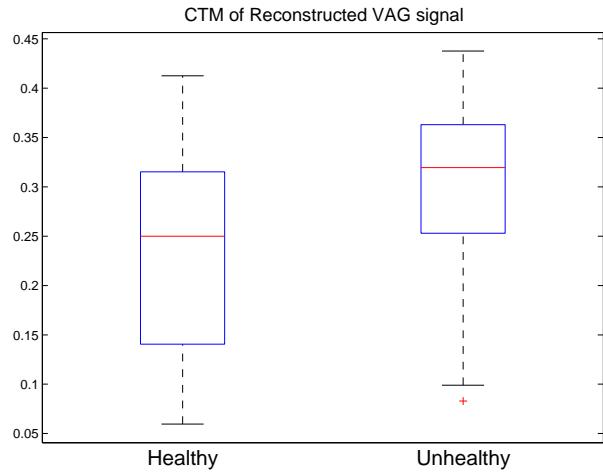


FIGURE 5.8: Boxplot for CTM

inferred that the area computed from unhealthy VAG signals consisted of larger distribution with respect to healthy VAG signals. The extracted features have been given as input to the classifier.

5.3.2 Classification

In this work, LS-SVM has been used for classification of VAG signals. LS-SVM toolbox by Brabanter *et al.* has been used for implementing LS-SVM [164]. Six entropy based features and area computed from CTM have been extracted from the reconstructed VAG signal. The dataset consisted of 89 VAG samples. Training set has been set to 60% of healthy and unhealthy VAG signals and remaining signals were considered as a test set. Targets belonging to unhealthy VAG signals are chosen as 1 and healthy signals are set to 0. A tenfold cross-validation procedure has been used in order to assess the classification performance of the classifier [165]. Radial basis function (RBF) has been used as a kernel function for the LS-SVM classifier. The parameters of LS-SVM were optimized using gridsearch algorithm. The classification has been carried out for 200 iterations and the best classification results were stored. The performance of the classifier was hugely based on area under the receiver operating curve (AUC-ROC), SEN, SPF and ACC. Other parameters such as PPV, NPV, MCC and FDR was also considered for the evaluation purpose.

Table 5.3 gives the classification performance of VAG signal reconstructed from dominant IMFs. The dominant IMFs (5,6 and 7) has been identified by DFA algorithm and reconstructed VAG signal has been formed by the summation of IMF5, IMF6 and IMF7. Six entropy based features (ApEn, SampEn, ShEn, ReEn, TsEn and PeEn) and CTM extracted from the reconstructed VAG signal has been given as an input to LS-SVM. For example, from table 5.3, the first row

TABLE 5.3: Classification performance of extracted features from reconstructed VAG signal

Features	ACC	SEN	SPE	PPV	NPV	MCC	FDR	AUC
ApEn	0.7640	0.8947	0.5313	0.7727	0.7391	0.4670	0.2273	0.7647 ± 0.0521
SampEn	0.8202	0.9444	0.6286	0.7969	0.8800	0.6228	0.2031	0.7672 ± 0.0506
ShEn	0.7528	0.8947	0.5000	0.7612	0.7273	0.4391	0.2388	0.6434 ± 0.0590
ReEn	0.7191	0.8793	0.4194	0.7391	0.6500	0.3409	0.2609	0.6583 ± 0.05851
TsEn	0.7079	0.8421	0.4688	0.7385	0.6250	0.3361	0.2615	0.6929 ± 0.0624
PeEn	0.8652	0.9107	0.7879	0.8793	0.8387	0.7082	0.1207	0.7864 ± 0.0474
Vector	0.8427	0.9107	0.7273	0.8500	0.8276	0.6575	0.1500	0.7812 ± 0.0507
CTM	0.8764	0.8947	0.8627	0.8293	0.9167	0.7517	0.1707	0.7023 ± 0.0506

signifies the classification performance of ApEn and similar classification have been carried out for the rest of the features. A vector of entropy based features has been formed by combining the six features.

$$\text{vector} = [\text{ApEn}, \text{SampEn}, \text{ShEn}, \text{ReEn}, \text{TsEn}, \text{PeEn}]. \quad (5.14)$$

In table 5.3, the “*vector*” label in “*Feature*” column represents classification performance of a feature vector consisting of six entropies (ApEn, SampEn, ShEn, ReEn, TsEn and PeEn) as elements. From table 5.3 it can be seen that using SampEn as an input parameter, LS-SVM gives a sensitivity of 94.44% which is highest among the other features and gives an accuracy of 82.02%. The performance of SampEn is better with respect to ApEn (ACC: 76.40%). The highest accuracy of 86.52% has been obtained by using PeEn as a parameter. MCC for PeEn is also highest among entropy based features with 0.7082 and provided a lowest FDR rate of 0.1207. The vector of entropies (ApEn, SampEn, ShEn, ReEn, TsEn and PeEn) gives an accuracy of 84.26%, MCC of 0.6544 and FDR of 0.1355. The area computed with CTM as feature extraction performed well with respect to other entropy based features. It gives the highest accuracy of 87.64% with highest MCC of 0.7517.

A receiver operating characteristic (ROC) curve is another effective measure of the classifier. It is a graphical representation of false positive rate (FPR, $\text{FPR} = 1 - \text{specificity}$) vs true positive rate (TPR, also named sensitivity). The area under the ROC (AUC-ROC)curve provides how adequately a classifier technique would be able to diagnose among different classes [130], [131]. The ROC plot for reconstructed VAG signals is given in figure 5.9. The area obtained from the vector of entropies gives AUC-ROC of 0.7812 ± 0.0507 while highest AUC-ROC has been obtained for PeEn (AUC: 0.7864 ± 0.0474). The AUC-ROC computed for CTM is 0.7023 ± 0.0506 .

From table 5.3 and figure 5.9 it can be observed that PeEn gives the highest accuracy of 86.52%

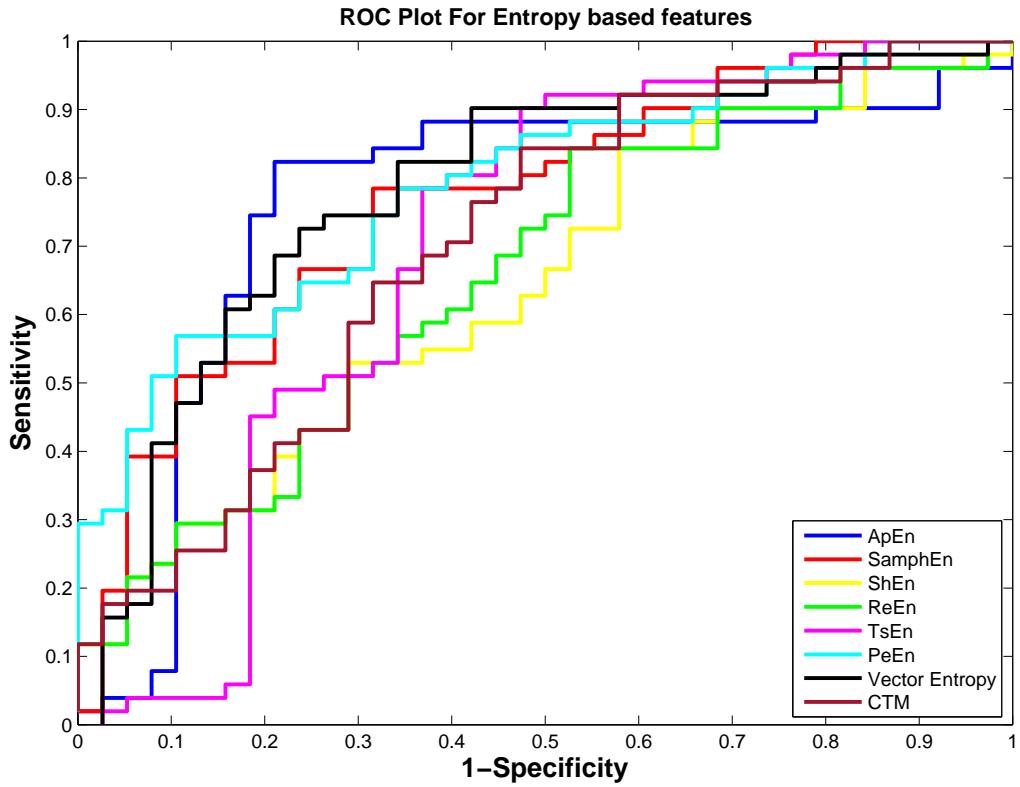


FIGURE 5.9: Performance of ROC of extracted features

and ROC curve (AUC: 0.7864 ± 0.0474). While area computed from CTM gives 0.7023 ± 0.0506 . The obtained results implied that LS-SVM as a classifier precisely categorized a larger number of VAG signals between healthy and unhealthy classes.

5.4 Discussion

The traditional signal processing techniques fail in the analysis of the nonstationary and nonlinear signals since methodologies a) assume that signals are linear and periodic in nature b) it is based on Fourier. Considering these aspects, nonlinear signal processing techniques has been carried out. This work has focused on feature extraction methodologies and classification. Simple extraction technique based on the area computed by CTM has been also covered in the study. In this work, VAG signals were decomposed into IMFs using CEEMDAN. Figure 5.3 (a) and (b) shows the IMFs obtained from the VAG signal for healthy and unhealthy subject respectively. The VAG signal was reconstructed by identifying the dominant IMFs (5,6 and 7) and the reconstructed VAG signal can be observed in figure 5.4.

The analysis has been carried out for 89 VAG signals. Different entropy based features (ApEn, SampEn, ShEn, ReEn, TsEn and PeEn) and CTM were computed for obtaining the characteristic

of VAG signal. The features were extracted from reconstructed VAG signals. Table 5.1 and 5.2 gave the statistical analysis of extracted features. As observed from the table, features extracted from unhealthy VAG signals were higher with respect to normal VAG signals. These statistical results motivated in building an effective classification model. The extracted features were given as input to LS-SVM.

This work compared the performance of extracted features. The classification performance has been presented in table 5.3. PeEn gives the highest classification accuracy of 86.52% with AUC-ROC curve of 0.7864 ± 0.0474 with respect to other entropies. But the highest classification accuracy was obtained by the feature extracted by CTM which gives a classification accuracy of 87.64% as observed in table 5.3. The vector consisting of all entropies also provided a better classification accuracy of 84.26% and AUC-ROC of 0.7812 ± 0.0507 .

Furthermore, the computational complexity of the proposed system consists of preprocessing technique, feature extraction technique and classification has been computed. In this work, the preprocessing (CEEMDAN) and classification (LS-SVM) remain constant while feature extraction techniques differ with respect to different entropies and CTM. The computational complexity of the whole system would be summation of CEEMDAN, extracted features and LS-SVM which is given as

$$\text{Complexity of whole system} = O(\text{CEEMDAN}) + O(\text{Features}) + O(\text{LS-SVM}) \quad (5.15)$$

The computational complexity of CEEMDAN would be approximately equivalent to EMD/EEMD. Wang *et al.* had computed the computational complexity of EMD/EEMD [178] and it is $O(N \log N)$. Olivier stated that computational complexity of LS-SVM depends on the sample size and no of features given as input to the classifier. The time complexity of LS-SVM computed is $O(N_2)$ [179], [180].

TABLE 5.4: Complexity of the extracted features

Features	ApEn	SampEn	ShEn	
Complexity	$O(N^2)$	$O(N^{3/2})$	$O(N \log(N))$	
Features	ReEn	TsEn	PeEn	CTM
Complexity	$O(N \log(N))$	$O(N \log(N))$	$O(N)$	$O(N)$

Table 5.4 shows the computational complexity of extracted features. ShEn, ReEn and TsEn entropies are computed by fast Fourier transform (FFT) which has the complexity of $O(N \log(N))$. Hence the complexity of ShEn, ReEn and TsEn are considered as $O \log(N)$. The complexity of ApEn is $O(N^2)$ [181] and Yu-Hsiang *et al.* computed the complexity of SampEn as $O(N^{3/2})$ [182]. PeEn and CTM provided simple computation having complexity of $O(N)$.

PeEn and CTM provides a high classification accuracy with a low computational complexity of $O(N)$. Thus, a simple, robust and low computational diagnostic system with PeEn or CTM as feature extraction technique could be designed. The complexity of the proposed system would be the summation of the computational complexity of CEEMDAN, PeEn/CTM, and LS-SVM. Hence, the computational complexity of the proposed system would be

$$O(N \log N) + O(N) + O(N^2) \approx O(N^2) \quad (5.16)$$

The proposed methodology has been compared with the previous studies related to nonlinear signal processing techniques and is shown in table 5.5. From the table, it can be observed that our proposed method with CEEMDAN and PeEn/CTM as feature extraction technique performs better with respect to other methods. PeEn and CTM have given the highest accuracy of 86.52% with MCC of 0.7082 and 87.64% respectively with respect to other nonlinear studies. Jein has used HHT method for VAG signals [183] but this analysis was carried on different dataset. While Wu *et al.* had used three entropies (symbolic entropy (SyEn), approximate entropy (ApEn), fuzzy entropy (FuzzyEn)) and three distinct envelope amplitude parameters (the mean, standard deviation, and root-mean-squared (RMS) values) of distinct signals [48] which gave an accuracy of 83.56% with highest sensitivity of 0.9444. Proposed method also performed well with respect to work of Nalband *et al.* [184] where only three entropies (Tsallis entropy, permutation entropy and spectral entropy) were studied without removing artefacts with double cascaded moving average filter.

TABLE 5.5: Comparison of the proposed methodology with the existing non-linear studies

	EMD[183]	EEMD[48]	EEMD[184]	PeEn	CTM
ACC	85.30%	83.56%	86.52%	86.52%	87.64%
SEN	NA	0.9440	0.9412	0.9107	0.8947
SPE	NA	0.8000	0.7632	0.7878	0.8627
MCC	NA	0.6599	0.6227	0.7082	0.7516
Classifier	NA	SVM	Random Forest	LS-SVM	LS-SVM

Nonlinear signal processing technique followed by entropy based feature extraction and CTM has been studied. The proposed methodology provided a better classification accuracy in distinguishing healthy and unhealthy VAG signals. The performance of features was compared and their computational complexity was also considered. Furthermore, the computational complexity of the proposed system was also computed.

5.5 Conclusion

In this work, the analysis of VAG signals has been carried out using nonlinear nonstationary signal processing technique “CEEMDAN”. It decomposes the VAG signals into IMFs. The VAG signals were reconstructed by identifying the significant IMFs. Entropy based features and area computed by CTM are extracted from reconstructed VAG signals. The extracted features provided distinct characteristics between healthy and unhealthy VAG classes and these features were given as inputs to the classifier. LS-SVM with RBF as kernel function has been used as classifier and the performance of extracted features has been evaluated. PeEn gives an accuracy of 86.52% among the six entropies but the highest classification accuracy has been obtained by CTM (ACC: 87.64%). The computed time complexity of PeEn/CTM having an order of $O(N)$ provided a simple, robust and low computational feature extraction technique along with good classification accuracy. The proposed methodology which includes CEEMDAN as preprocessing techniques, PeEn/CTM as feature extraction and LS-SVM as a classifier, provides an effective tool for developing computer aided diagnostics system for knee joint disorders.

Chapter 6

Time-Frequency Analysis of VAG signals

God is a frequency. Stay tuned.

Alan Cohen

In this chapter, the analysis of the VAG signal has been carried out using the time-frequency techniques. The analysis carried on nonstationary and nonlinear signals by traditional signal processing assumes the linearity nature of the signal. Thereby, the inference of these signal is either interpreted in the time domain or in the frequency domain. Therefore, time frequency analysis has been explored which provides time frequency content of the signal simultaneously. This chapter compares the performance of smoothed pseudo Wigner–Ville distribution (SPWVD) and Hilbert-Huang transform (HHT) for the time frequency analysis of VAG signals. Figure 6.1 shows the proposed methodology of the chapter.

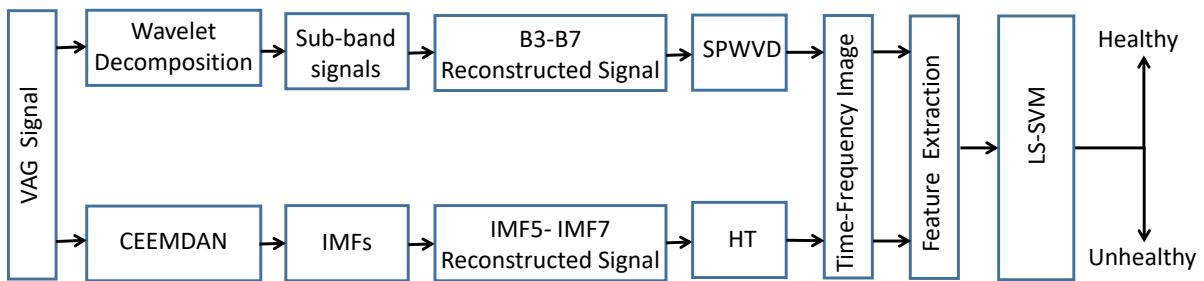


FIGURE 6.1: Flow chart for time frequency analysis of knee joint disorders using VAG signals

The chapter has been organized as follows: section one introduces the techniques of time frequency techniques in the analysis of the signal. Section 2 covers a brief theoretical background of the

proposed work. The results and discussion of the proposed work have been explained in section 3 and finally, section 4 concludes our proposed study.

6.1 Introduction

The essential steps in signal processing for any nonlinear, nonstationary and multicomponent signals are analysis, detection and classification. The analysis of signal has been carried out in the time domain and frequency domain. Most of the signals are multicomponent having time variant properties and have different frequencies. Techniques based on these domains are quite restrictive in making the physical inferences. Hence, time variant and frequency domain are required to be observed simultaneously and this technique is known as “time frequency analysis”. The algorithms which compute the time frequency analysis is called as “time-frequency techniques”. Most popular time frequency techniques used in biomedical are short-time Fourier transform (STFT) or Gabor transform (GT), the continuous Morlet wavelet transform (CMWT) and the Hilbert transform (HT). Time frequency analysis has been widely used for many application of engineering.

Time-frequency analysis has been studied for the analysis of various biomedical signals such as ECG, EEG, EMG etc. [185]–[187]. Time-frequency has been also carried out for the analysis of VAG signals. Krishna *et al.* applied time-frequency distribution by optimization of minimum cross-entropy [63]. The time-frequency distribution was obtained by matching pursuit decomposition algorithm. Features such as energy, energy spread, frequency, and frequency spread were computed and compared between healthy and unhealthy VAG signals. A similar analysis of time-frequency distribution was carried out but Keo Sik Kim *et al.* by segmenting VAG signals with dynamic time warping (DTW) and singular value decomposition (SVD) for denoising [9]. Chen *et al.* have used independent component analysis for VAG signals [67]. The time frequency analysis related to age specific was carried out by Baczkowicz *et. al* [15]. Discrete Fourier transform was applied to segmented VAG signals to obtain short time spectrum. Therefore, this chapter continue to explore the advance time frequency analysis of VAG signals.

In the current investigation, we propose to use nonstationary linear signal processing techniques namely as wavelet packet decomposition (WPD) [33], [34] and nonstationary nonlinear signal processing technique such complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) [52]. These techniques are used as preprocessing techniques for the analysis of VAG signals.

This chapter explains and compares the performance of smoothed pseudo Wigner–Ville distribution (SPWVD) [68] and Hilbert-Huang transform (HHT)[188] for the time frequency analysis of VAG signals. SPWVD overcomes the problem of interference of cross terms in Wigner-Ville

distribution. This chapter proposes to apply WPD on VAG signal into subband signals of different frequency. The dominant subband signals are identified and VAG signal is reconstructed [189]. Hence, time-frequency representation of reconstructed VAG is obtained by using SPWVD and is considered as the time-frequency image. Thereby the statistical features such as mean, standard deviation, skewness and kurtosis are extracted. SPWVD has been successfully studied for the analysis of biomedical signals [69], wind turbine fault diagnostic [70] acoustic and vibration signals in detection of gear failures [71].

Moreover, considering the nonlinear, nonstationary and multicomponent nature of VAG signal, a modified version of Hilbert-Huang Transform (HHT) [188] known as CEEMDAN-HHT has been proposed. Traditional HHT consists of empirical mode decomposition (EMD) [44] and Hilbert transform (HT). EMD decomposes the signal into a number of intrinsic mode functions (IMF). But one of the major drawback of EMD is the mode mixing, in which a single IMF consists of signals having different scales [46]. This thesis proposes to use CEEMDAN for obtaining IMFs [52]. This has been presented in chapter 5, section 5.2.1. The VAG signal is reconstructed by dominant IMFs and the time frequency representation is obtained by HT. This time-frequency representation is considered as time-frequency image and statistical features consisting of mean, standard deviation, skewness and kurtosis are extracted from time-frequency image. The extracted features obtained by using SPWVD and CEEMDAN-HHT are given as input to LS-SVM as a classifier to evaluate their performance. Therefore, a comparative study of different methodologies (nonstationary linear signal processing technique and nonstationary nonlinear signal processing technique) have been carried out and their performance has been compared.

6.2 Methods

6.2.1 Wavelet Packet Decomposition

The WPD algorithm disintegrates a given signal into detail coefficients and approximation to derive at the first stage of decomposition [34]. Thus, these approximation coefficients are further disintegrated into approximation and coefficients in successive stages. Hence the signal breaks down like a tree. Detailed explanation has been given in the chapter 3 section 3.2.

6.2.2 SPWVD

Wigner-Ville distribution (WVD) is quadratic or nonlinear time-frequency analysis in which it distributes the energy of the signal along the variables time and frequency [190]. Initially, it was introduced in the domain of quantum mechanics as Wigner distribution but later it was

used in the framework of signal analysis as Wigner-Ville distribution. Moreover, WVD has better theoretical properties, but it has a major drawback as it produces interferences terms or cross terms between different components of multi-component signals. It means that there are additional components present in the signal. In order to remove this, a separable smoothing window is introduced which is independent of time and frequency and it is called as smooth pseudo Wigner-Ville distribution [68]. Here, the time-frequency resolution can be controlled independently. Hence, in this work smooth pseudo Wigner-Ville distribution (SPWVD), the improved version of WVD has been investigated for the nonstationary and nonlinear VAG signals.

6.2.3 Hilbert Huang Transform

Detailed explanation has been presented in Chapter 5 section 5.2.2.

6.2.4 Feature Extraction

The time-frequency representation using SPWVD and HHT is considered as a time-frequency image and features are extracted as the characteristics of the signal [191]–[193]. The pixel intensity is used as the measure of calculated features. Each pixel corresponds to a particular time and frequency which represents the power of the signal at any instance of frequency. Statistical features such as mean, standard deviation, skewness and kurtosis are extracted from time-frequency image. Hence, these statistical value gives an overview of the essential information as attributes of the signal[191]. The statistical features of VAG signals are computed as follows:

$$\text{Mean} : \mu_p = \frac{1}{N} \sum_{n=1}^N I_p(n) \quad (6.1)$$

$$\text{Standarddeviation} : \sigma_p = \sqrt{\frac{1}{N} \sum_{n=1}^N (I_p(n) - \mu_p)^2} \quad (6.2)$$

$$\text{Skewness} : S_p = \sqrt{\frac{\frac{1}{N} \sum_{n=1}^N (I_p(n) - \mu_p)^3}{(\sigma_p)^3}} \quad (6.3)$$

$$\text{Kurtosis} : K_p = \sqrt{\frac{\frac{1}{N} \sum_{n=1}^N (I_p(n) - \mu_p)^4}{(\sigma_p)^4}} \quad (6.4)$$

Here I_p is pixel intensity and $n \in \{1, 2, 3, 4, \dots, N\}$ where N is the total number of pixels in the image.

6.3 Results

6.3.1 Preprocessing

Figure 6.2 represents a sample VAG signals of healthy and unhealthy subjects. A double cascaded moving average filter was applied on VAG signals to remove the artefacts and baseline wander [24]. The proposed work consisting of preprocessing, feature extraction and classification have been carried out in MATLAB©2015b. After filtering, the VAG signals are decomposed into wavelet coefficients and IMFs using WPD and CEEMDAN respectively. A WPD is carried out by decomposing the VAG signals into subband signals. Since the sampling rate of VAG is 2k, the frequency available would be 1 Khz. Hence, VAG signals are decomposed into 10 levels of subband signals ($B_1, B_2, B_3, B_4, B_5, B_6, B_7, B_8, B_9, B_{10}, a_{10}$) and each subband signals consists of particular frequency[189]. Figure 6.3 (a) and (b) represents the subband signals obtained from wavelet packet decomposition of healthy and unhealthy VAG signal respectively.

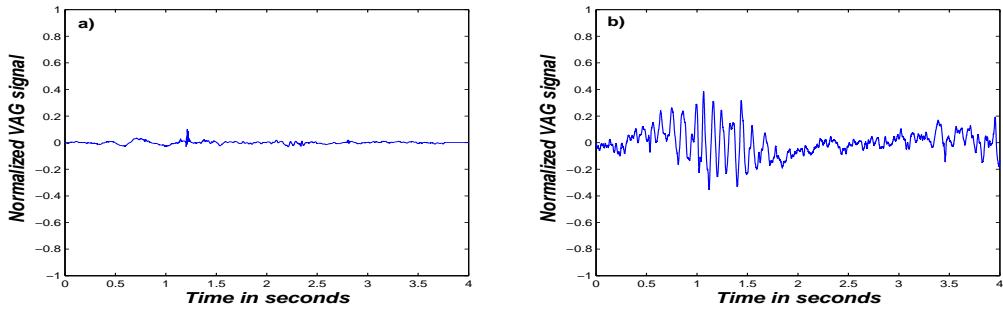


FIGURE 6.2: VAG signal of (a) healthy and (b) unhealthy subject

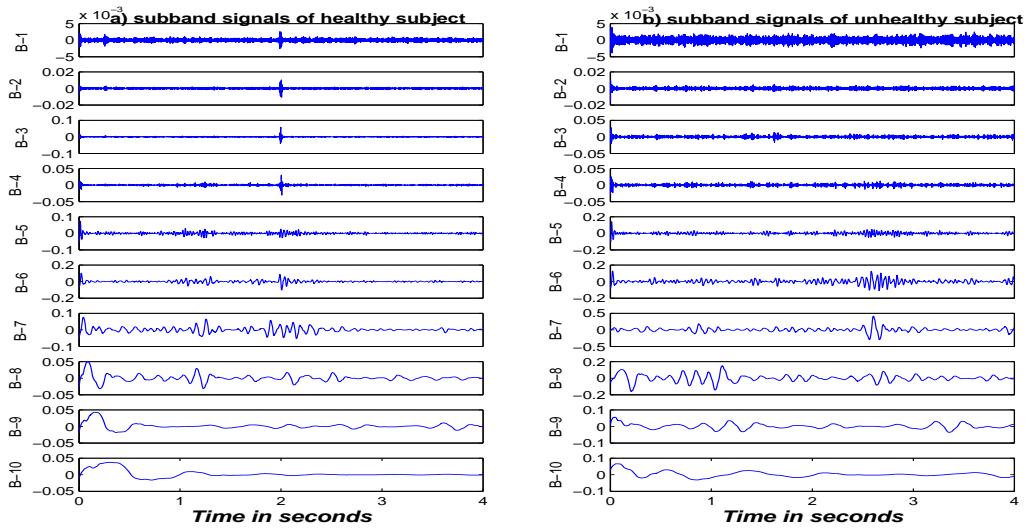


FIGURE 6.3: Wavelet decomposition of VAG signal for (a) healthy and (b) unhealthy subject

Similarly, Figure 6.4 represents the decomposed signal obtained from CEEMDAN. Figure 6.4 (a) and (b) shows the IMFs obtained using CEEMDAN. It can be observed from the figure that different dynamics involved in VAG signals are observed in different IMFs for healthy and unhealthy subject respectively. High frequency oscillation is observed in IMF1-IMF3 and low frequency oscillation can be seen in IMF4-IMF10. As observed, flat envelopes can be seen in IMF1 and IMF2 for both the cases and can be discarded for the study as it doesn't signify any pathological condition.

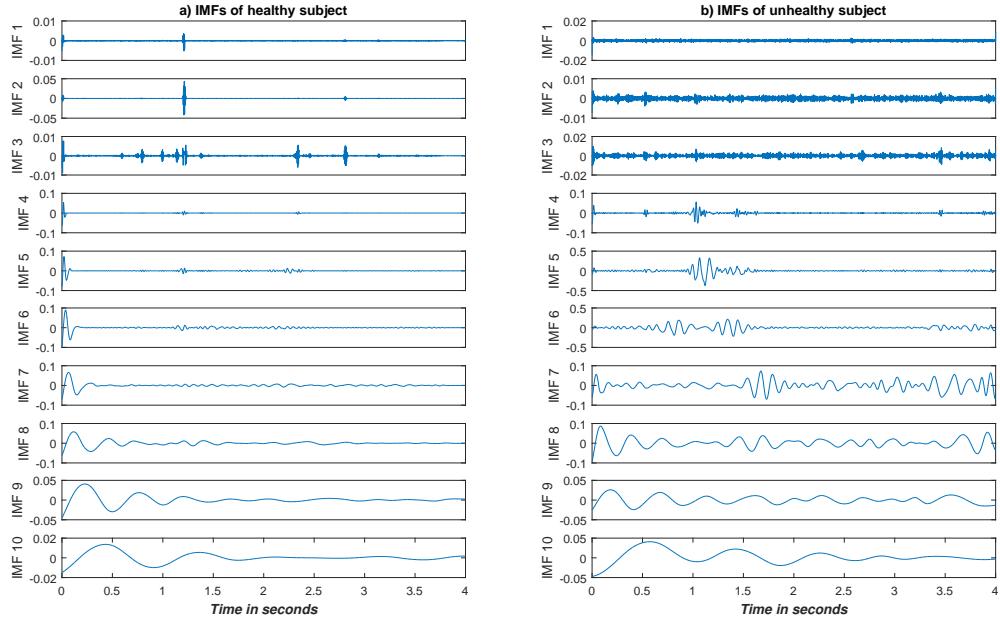


FIGURE 6.4: IMFs of VAG signals obtained using CEEMDAN for (a) healthy and (b) unhealthy subject

6.3.2 Feature Extraction

VAG signals are reconstructed from the subband signals obtained from WPD. From the figure 6.3, the subband signals $B1$, $B2$, $B8$, $B9$ and $B10$ were ignored and were not considered for the study. $B8$, $B9$ and $B10$ represented artefacts such as baseline wander, while the subband signals $B1$ and $B2$ indicated the presence of high-frequency components. And thus, VAG signal was reconstructed using subband signals $B3$, $B4$, $B5$, $B6$ and $B7$ as shown in figure 6.5(a) and (b). Figure 6.5(c) and (d) represents the time-frequency representation using SPWVD. Hence, attributes such as mean, standard deviation, skewness and kurtosis are computed from time-frequency image for 89 VAG signals.

Identifying the correct IMFs for VAG signal is carried out by computing detrended fluctuation analysis (DFA). The DFA is computed to obtain the correlation properties of IMFs. The fractal

scaling index (α) obtained from DFA helps in identifying the correct IMFs. The α value of the IMFs in the range between ($0 < \alpha < 0.5$) and baseline wander related IMFs (IMF9-IMF10) have been ignored as per the study concluded by Wu[45], [184]. Hence, from the computed α value, IMF5, IMF6 and IMF7 are considered as correct IMFs of VAG signal and their α values are 1.302, 1.622 and 1.88 respectively. Hence, the VAG signal is reconstructed from the correct IMFs and is shown in figure 6.6 (a) and (b) for healthy and unhealthy VAG signal respectively. Figure 6.6 (c) and (d) represents the time-frequency representation of the reconstructed VAG signals. This representation is considered as time-frequency image and individual attributes such as mean, standard deviation, skewness and kurtosis are computed. From figure 6.6 (a) and (b) it can be observed that the variation of frequency is highly variant in case of unhealthy VAG signals as compared to healthy VAG signal. Statistical features were extracted from 89 VAG signals. As observed from figure 6.5 and 6.6, the time-frequency representation of VAG signal using CEEMDAN provides a sharper distinguishable spectrum than time-frequency representation of VAG signal using SPWVD. Thereby, the characteristics of the time-frequency representation obtained from above two methods are computed by extracting statistical features. This analysis

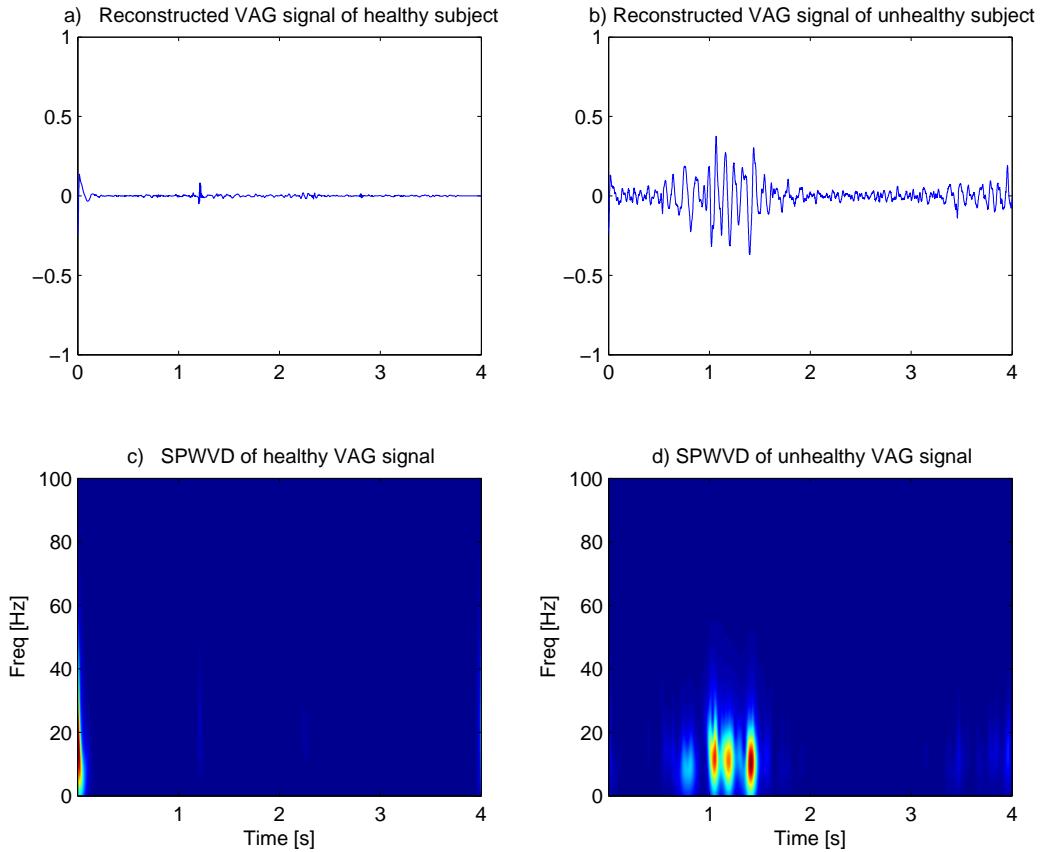


FIGURE 6.5: The reconstructed VAG signal obtained from SPWVD (a) healthy subject and (b) unhealthy subject. Time-frequency representation of c) healthy subject and d) unhealthy subject VAG signal

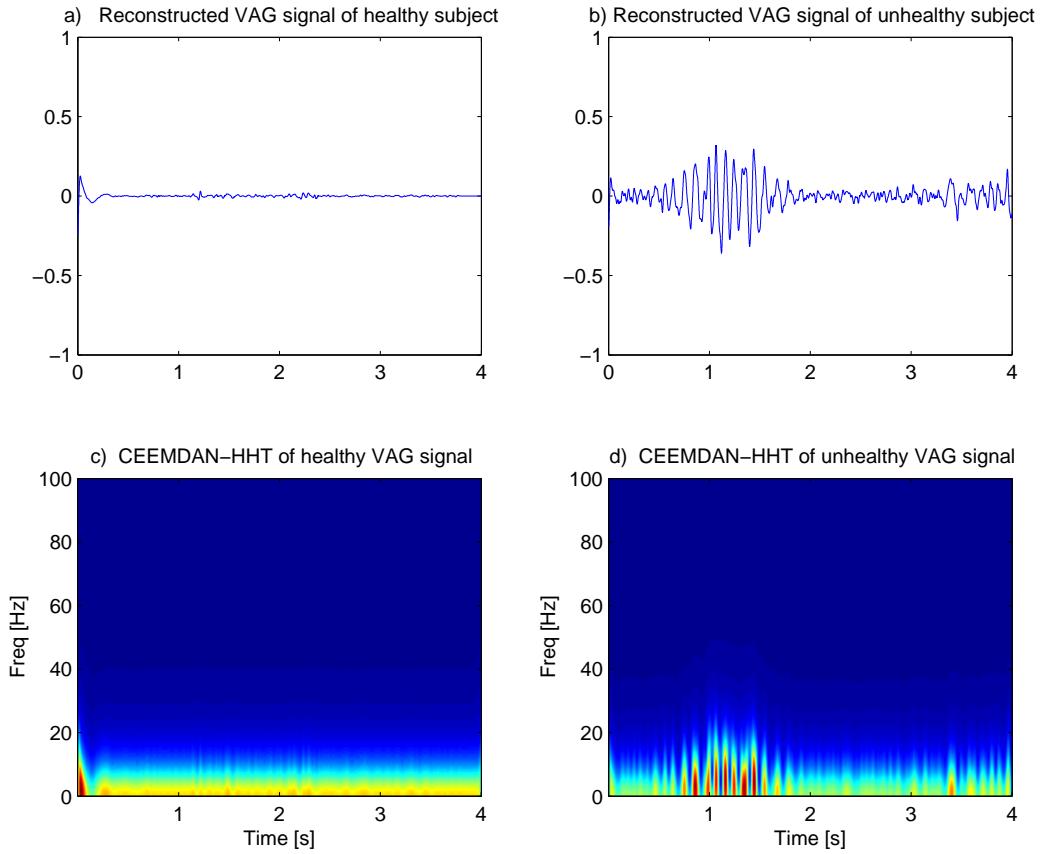


FIGURE 6.6: The reconstructed VAG signal obtained from CEEMDAN (a) healthy subject and (b) unhealthy subject. Time-frequency representation of c) healthy subject and d) unhealthy subject VAG signal

has been carried out for 89 VAG signal and hence a comprehensive study has been carried out.

A statistical test was also performed on the extracted features. Kruskal-Wallis test was carried out to infer the discriminability of extracted features among the healthy and unhealthy class of distribution. Table 6.1 refer the mean, standard deviation and p -value computed for extracted features from SPWVD and CEEMDAN-HHT. From the table 6.1, it can be seen that there is a clear distinction between the healthy and unhealthy group of VAG signals. Figure 6.7 provides the box plots for extracted features for SPWVD and CEEMDAN-HHT. In both cases of SPWVD and CEEMDAN-HHT, it can be observed that there is a significant variation among the healthy and unhealthy group. The statistical features of unhealthy VAG signals provided larger variation with respect to healthy VAG signals. Hence, the statistical analysis provided a significant difference between the groups and these extracted features are given as input to the classifier in building an effective classification model.

TABLE 6.1: Statistical measures of extracted features using SPWVD and CEEMDAN-HHT

Features	class	Normal	Abnormal	<i>p-value</i>
		Mean \pm Std	Mean \pm Std	
SPWVD				
Mean	Normal	0.3461 \pm 0.0662	0.3330 \pm 0.0766	0.3237
Standard	Normal	1.6886 \pm 0.3238	1.6224 \pm 0.3766	0.3136
skewness	Normal	6.7910 \pm 0.0900	6.8137 \pm 0.1127	0.2663
Kurtosis	Normal	52.7752 \pm 2.1699	53.4333 \pm 2.6143	0.2050
CEEMDAN-HHT				
Mean	Normal	0.3587 \pm 0.0769	0.3373 \pm 0.0670	0.1167
STD	Normal	1.7578 \pm 0.3705	1.6550 \pm 0.3224	0.1293
skewness	Normal	7.0215 \pm 0.3311	7.1490 \pm 0.3979	0.0629
Kurtosis	Normal	58.7469 \pm 8.8681	62.2163 \pm 10.6879	0.0559

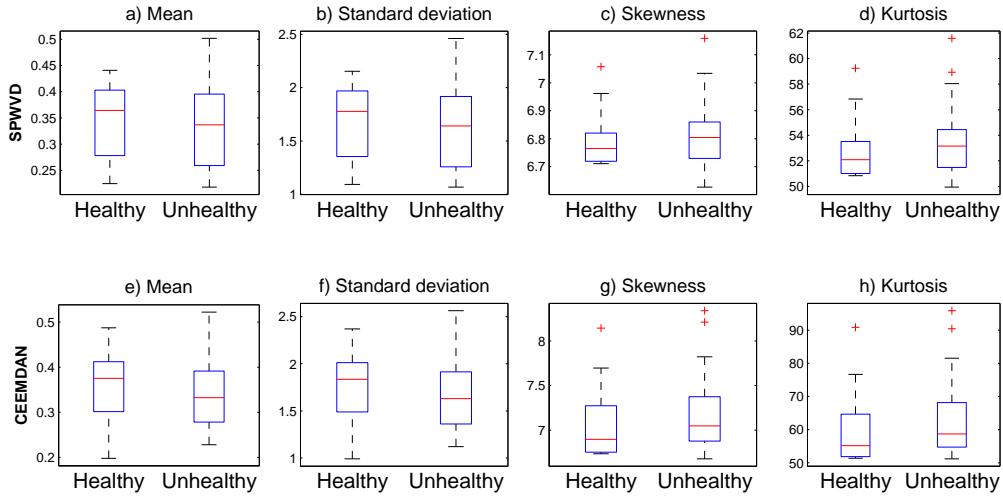


FIGURE 6.7: Box plots for features extracted from SPWVD and CEEMDAN-HHT

6.3.3 Classification

The extracted features from SPWVD and CEEMDAN-HHT are given as inputs to LS-SVM. Various kernel such as linear, polynomial, quadratic and radial basis function (RBF) are available as kernel functions. In our study, RBF as kernel function has been used since it provides optimal results. The kernel parameters were evaluated based on trial and error approach. Regularization constant was selected appropriately for avoiding over-fitting and 10 cross validation was performed. The performance of the proposed system was evaluated based on parameters such as sensitivity (SEN), specificity(SPE), accuracy(ACC), Matthews correlation

TABLE 6.2: Classification of features extracted using SPWVD by LS-SVM

	Vector	Mean	Std	Skewness	Kurtosis
ACC	0.7753	0.7865	0.6667	0.7416	0.6180
SEN	0.9091	0.8800	0.8000	0.7586	0.6429
SPE	0.7313	0.7500	0.6389	0.7333	0.6133
PPV	0.5263	0.5789	0.3158	0.5789	0.2368
NPV	0.9608	0.9412	0.9388	0.8627	0.9020
MCC	0.5585	0.5724	0.3343	0.4661	0.1886
FDR	0.4737	0.4211	0.6842	0.4211	0.7632
AUC-ROC	0.7699 ± 0.0573	0.8945 ± 0.0622	0.7281 ± 0.0593	0.7384 ± 0.0620	0.6744 ± 0.0594

TABLE 6.3: Classification of features extracted using CEEMDAN-HHT by LS-SVM

	Vector	Mean	Std	Skewness	Kurtosis
ACC	0.8202	0.8876	0.8764	0.7978	0.7865
SEN	0.8750	0.8333	0.9655	0.7500	0.7209
SPE	0.7895	0.9362	0.8333	0.8367	0.8478
PPV	0.7000	0.9211	0.7368	0.7895	0.8158
NPV	0.9184	0.8627	0.9804	0.8039	0.7647
MCC	0.6410	0.7766	0.7569	0.5901	0.5746
FDR	0.3000	0.0789	0.2632	0.2105	0.1842
AUC-ROC	0.8994 ± 0.0579	0.9383 ± 0.0615	0.9494 ± 0.0617	0.8127 ± 0.0617	0.8519 ± 0.0568

coefficient (MCC), false discovery detection rate (FDR), negative predictive value (NPV) and positive predictive value (PPV).

Table 6.2 and 6.3 provides the classification performance of LS-SVM using features extracted from time-frequency image of SPWVD and CEEMDAN-HHT respectively. The classification performance has been evaluated with features given as individual input and as a vector of all features (mean, standard deviation, skewness and kurtosis) to LS-SVM. The performance of features of CEEMDAN-HHT given as a vector provides an accuracy of 82.02% with a sensitivity of 87.50% and specificity of 78.95%. MCC is 0.6410 with low FDR of 0.3000 for a vector of features as observed in table 6.3. But the highest classification accuracy has been obtained for the feature “mean” which gives an accuracy of 88.76% with a highest MCC value of 0.7766 and low FDR of 0.0789 using CEEMDAN-HHT (table 6.3). The highest accuracy obtained using SPWVD gives 77.53% for a vector of all features. Which is comparably less than features extracted from CEEMDAN-HHT (ACC = 82.02% for a vector of features). As observed from the table 6.2 and 6.3, CEEMDAN-HHT provided better classification performance with respect to SPWVD. The other parameters of classification also inferred that the features extracted from CEEMDAN-HHT provided better pattern classification than SPWVD.

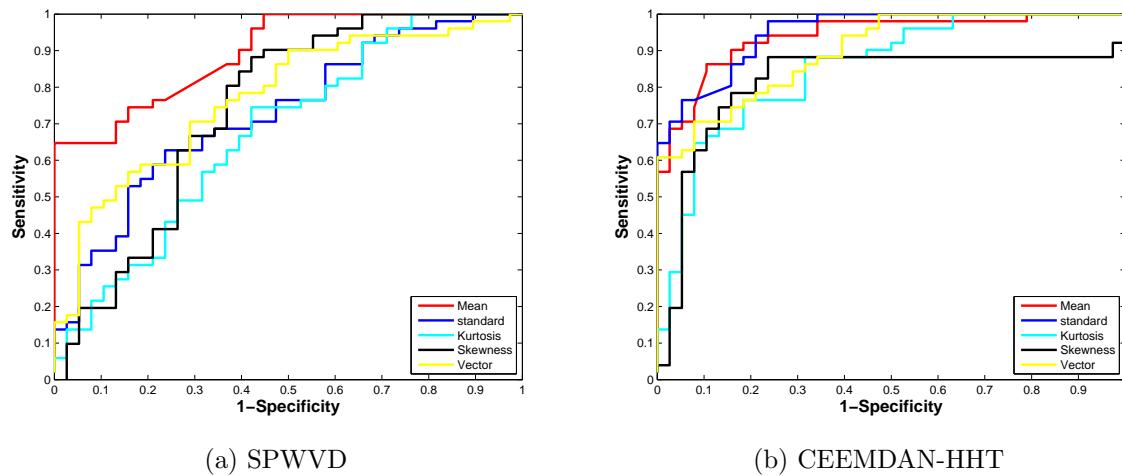


FIGURE 6.8: ROC plot for statistical features obtained from time frequency distribution

Graphical representation of classifier was evaluated using receiver operating characteristic (ROC). The area under the curve (AUC) provides the effectiveness of the classifier. The AUC-ROC plot for individual features and vector of features extracted from SPWVD and CEEMDAN-HHT are shown in figure 6.8 (a) and (b) respectively. The highest (AUC-ROC = 0.8519) was obtained for a vector of features extracted from CEEMDAN-HHT. These parameters implied that using LS-SVM classifier technique, a larger number of VAG signals were precisely categorized into healthy and unhealthy classes for features extracted from CEEMDAN-HHT.

6.4 Discussion

This study demonstrated the comparative analysis of time-frequency distribution of VAG signals. The time-frequency representation has been computed by SPWVD and CEEMDAN-HHT which provides the energy distribution of VAG signals as observed in figure 6.5 and 6.6. Hence, it can be inferred that the unhealthy VAG signal provide a larger variation as compared to VAG signal of the healthy subjects. The cause of this large variation might be due to wear and tear of the muscle tissues associated with the knee joint. As a result, this did not allow the free leg movements in performing extension and flexion. While comparing figure 6.5 and 6.6, it can be observed that the time frequency image of HHT is sharper than time frequency image of SPWVD. It also provides a better variety especially in the case of the unhealthy VAG signals. These variations were evident through features extracted from time frequency images of SPWVD and CEEMDAN-HHT. These variations were quite evident in Table 6.1 which provided the statistical analysis of extracted features. The graphical display of extracted feature using boxplot has been provided by figure 6.7, indicates a large variation between the healthy and unhealthy group. And finally, the features has been given as input to build a classification model using

LS-SVM. The performance of SPVWD and CEEMDAN-HHT has been compared in table 6.2 and 6.3. Results indicated that CEEMDAN-HHT provided a much better pattern classification.

TABLE 6.4: Comparison of the proposed methodology with the existing TFD

Methodology	MP [62]	MP [63]	MLD [65]	EMD-HHT [67]	CEEMDAN
ACC	77.8	68.9	79.8	85.3	88.76
SEN	NA	NA	NA	NA	0.8334
SPE	NA	NA	NA	NA	0.9362
MCC	NA	NA	NA	NA	0.7766
Classifier	DA	LR	LDA	NA	LS-SVM
No of samples	37	90	89	35	89

Table 6.4 represents comparison of the proposed work with the existing work reported in the time frequency analysis of VAG signals. Initial study carried out by Krishnan *et al.* consisted of 37 samples of VAG signals [62]. The time frequency distribution was obtained by matching pursuit method (MP) method and it gave a classification accuracy of 77.8%. Krishnan *et al.* continued his work by constructing adaptive time frequency distribution by minimizing the cross optimization of time frequency distribution computed by the MP method [63]. The work consisted of 90 samples of VAG signals which gave a classification accuracy of 68.9% using logistic regression. Umapathy *et al.* further carried out the time frequency analysis by computing TFD using modified local discriminant (MLD) [65]. It gave a classification accuracy of 79.8% using linear discriminant analysis. Chen *et al.* carried out the analysis using traditional HHT (using EMD for computing IMFs) which gave a classification accuracy of 85.3% [67]. Chen used different dataset which consisted of 35 (12 healthy subjects and 23 unhealthy subjects) samples of VAG signal. Different data acquisition system had been applied for acquiring VAG signals (different accelerometers sensor, DAQ board, sampling frequency). The dataset consisted of few samples of VAG signals and therefore, comparing the results of different dataset would not be ideal. We have compared the performance of nonlinear techniques (EMD-HHT) used by Chen *et al.* on the analysis of VAG signals. From table 6.4 it can be observed that, the proposed method of CEEMDAN-HHT performed better with respect to other methodologies. The proposed methodology (CEEMDAN-HHT) gives the highest classification accuracy of 88.76% using LS-SVM.

6.5 Conclusion

This study proposed a computer aided diagnostic system using time-frequency analysis of VAG signals. Time-frequency based techniques namely SPWVD and CEEMDAN-HHT were studied and their performances were evaluated. The time-frequency representation has been

considered as time-frequency image and statistical features based on image pixies were extracted. Results concluded that time frequency technique using CEEMDAN-HHT provided a prominent distinction between the healthy and unhealthy group. Furthermore, the pattern classification results also concluded that CEEMDAN-HHT performed better with respect to SPWVD giving highest accuracy of 88.76% and MCC of 0.7766 with low FDR of 0.0789. The proposed work provided a better methodology in analysing nonlinear, nonstationary and multicomponent VAG signals using time-frequency techniques. Hence, a computer aided diagnostic system could be developed with the proposed methodology for non-invasive diagnosis of knee joint disorders.

Chapter 7

Conclusion and Future Scope

I am turned into a sort of machine for observing facts and grinding out conclusions.

Charles Darwin

7.1 Outcome Of The Thesis

The knee-joint is one of the most complex, the strongest and most important joints in the human body. Knee joints can tolerate a moderate amount of stress without significant damage due to daily work routine activities. But they are unable to resist the rotational forces that are commonly related to sports activities. These activities are quite susceptible to knee joint disorders. Therefore, early diagnosis of knee-joint diseases provides researchers an opportunity for investigation. Procedures based on noninvasive and invasive techniques are available for the diagnosis of knee-joint diseases. Arthroscopy, a semi-invasive based procedure does provide an aide. But arthroscopy does carry a considerable amount of risk, such as bleeding into the knee-joint, damage to the tissues/muscles or infection in the knee-joint. Noninvasive based medical imaging approaches such as fMRI, MRI, Computer Tomography, X-Rays do provide the early detection of knee-joint disease, however they are expensive and are usually not recommended for a regular checkup [3]. In addition, their ineffectiveness in displaying the dynamic aspects of an injured knee-joint disease is one of the major drawbacks of these procedures [5]. Considering these constraints, an alternative solution which is noninvasive, and low cost is required for early diagnosis of knee-joint disorders.

During active movements such as flexion and extension, the auditory sounds or the vibration emitted from the mid patella is called as vibroarthrographic signal (VAG) [45]. These signals

contain information that could be used to characterize certain pathological aspects of the knee joint diseases. The VAG signals are nonstationary, nonlinear and multicomponent in nature. Therefore, the analysis of VAG signal using signal processing is an important aspect in the diagnosis of knee joint disorders. Various researchers has used different methodologies in the analysis of VAG signals. The study has been limited in classifying the healthy and unhealthy VAG signals.

TABLE 7.1: Summary of thesis methodologies

	Methodology 1	Methodology 2	Methodology 3	Methodology 4
Chapter	3	4	5	6
Preprocessing	WPD	WPD	CEEMDAN	WPD-SPWVD & CEEMDAN-HHT
Feature extraction	Entropies + RQA	Entropies + RQA	Entropies	Statistical Features
Feature Selection	-	GA & ApA	-	-
Classifier	LS-SVM			

The main objectives of this research is to develop a framework methodology for a computer-aided diagnostic system for early diagnosis of knee joint disorders. VAG signals are nonlinear, nonstationary and multicomponent in nature. Considering these characteristics, methodologies based on nonstationary signal processing techniques, followed by feature extraction techniques and pattern classification have been carried out. This study has been limited in classifying the healthy and unhealthy VAG signals. The summary of proposed methodology has been shown in table 7.1 and they are briefly explained as follows:

1. In the first approach, a linear signal processing such as wavelet packet decomposition (WPD) has been considered for the analysis of VAG signals. Six entropy based features and four recurrence quantification analysis (RQA) parameters were considered in this study. A pattern classification was carried out using LS-SVM. In order to remove the irrelevant and insignificant features, feature selection algorithm has been carried out. Genetic algorithm and the apriori algorithm has been considered for this study.
2. In the second approach, nonlinear signal processing technique based on EMD known as CEEMDAN has been studied. CEEMDAN as preprocessing technique has been utilised as a mandatory procedure for further analysis of VAG. Similarly, six entropy based features have been considered for this study. A simple feature extraction technique based on CTM has been considered in this work.

TABLE 7.2: Summary of results from the proposed work in thesis

Features	No of features	ACC	SEN	MCC	AUC-ROC	Chapter
Entr + RQA (SB+ReCon)	60	69.66%	0.5888	0.4337	0.7822 ± 0.0471	
Entr + RQA (SB)	50	80.90%	0.9216	0.8621	0.8606 ± 0.0397	3
Entr + RQA (ReCon)	10	84.27%	0.9020	0.8529	0.8560 ± 0.0427	
FS by GA	8	82.02%	0.7254	0.6707	0.8754 ± 0.0347	4
FS by ApA	5	85.39%	0.9272	0.6868	0.9252 ± 0.0261	
Entropies (vector)	6	84.27%	0.9107	0.6575	0.7812 ± 0.0507	
PeEn	1	86.52%	0.9107	0.7082	0.7864 ± 0.0474	5
CTM	1	87.64%	0.8947	0.7517	0.7023 ± 0.0506	
SPWVD	4	77.53%	0.9091	0.5585	0.7699 ± 0.0573	
CEEMDAN-HHT	4	82.02%	0.8750	0.6410	0.8994 ± 0.0579	6
Mean(CEEMDAN-HHT)	1	88.76%	0.8333	0.7766	0.9383 ± 0.0615	

3. Time frequency analysis of VAG signal has been considered in this thesis work. SPWVD and CEEMDAN-HHT has been considered for the TFD. Statistical features were extracted from TFD as characteristics of the VAG signals.

7.2 Contribution Of The Thesis

This study provides different framework methodologies for the analysis of VAG signal. Table 7.2 provided the overall contribution of the thesis. The contribution of the thesis can be summarized as follows.

1. The analysis of VAG signals has been carried out using nonstationary linear signal processing technique such as WPD. The VAG signals have been decomposed into subband signal. Entropy based features and recurrence quantification analysis(RQA) based features have been extracted from the subband signals and reconstructed VAG signals. To classify and evaluate the performance of the proposed methodology, LS-SVM has been used for distinguishing healthy and unhealthy VAG signals. This analysis has been carried out using feature extracted from subband signals and from reconstructed signal. From table 7.2 it can be observed that a vector of 60 features extracted from subband signals and reconstructed VAG signals gives an accuracy of 69.66%. A vector of 50 features extracted from subband signal gives an accuracy of 80.90% while features (no of features: 10) extracted from reconstructed VAG signal gives a better classification accuracy of 84.27% with AUC-ROC of 0.8560 ± 0.0427 .

2. Feature selection algorithm has been used to select the most relevant and significant features from 60 features extracted from subband signals and reconstructed VAG signals. GA and ApA algorithm has been used as feature selection algorithm. Results from the feature selection algorithm inferred that entropy based features provided more discrimination between healthy and unhealthy VAG signals, since most of the features selected by feature selection algorithms were entropy based features rather than RQA parameters. The feature vector (no of features: 8) selected from GA gives a classification accuracy of 82.02% while the highest classification accuracy of 85.39% and AUC-ROC of 0.9252 ± 0.0261 was obtained from feature vector (no of features: 5) selected by apriori algorithm as observed from table 7.2.
3. The analysis of VAG signals has been carried out using nonstationary nonlinear signal processing techniques. In this work, technique based on EMD, namely, CEEMDAN has been used as a prepossessing step. The VAG signals has been reconstructed by identifying the significant IMFs and feature extraction has been carried out from the reconstructed VAG signals. In this study entropy based features has been studied since it was concluded that entropy based features performed better with respect to RQA parameters in the previous chapter. The feature vector of entropy based feature gives an accuracy of 84.27%. A simple feature extraction technique based on CTM gives the highest classification accuracy of 87.64%.
4. The analysis of VAG signals has been studied using time-frequency distribution. Time-frequency based techniques such as SPWVD and CEEMDAN-HHT has been studied and their performances has been evaluated. The time-frequency representation has been considered as time-frequency image and statistical features based on image pixies has been extracted. Results concluded that time frequency technique using CEEMDAN-HHT provided a prominent distinction between the healthy and unhealthy group. Furthermore, the pattern classification results also concluded that CEEMDAN-HHT with feature as “mean” performed better with respect to other features by giving the highest accuracy of 88.76% and AUC-ROC of 0.9383 ± 0.0615 .

Furthermore, from table 7.2 it can be also concluded that in this thesis the time frequency analysis using CEEMDAN-HHT performed better with respect to other methodologies. Moreover, the proposed work using CEEMDAN-HHT provided a better methodology in analysing nonlinear, nonstationary and multicomponent VAG signals using CEEMDAN-HHT. Hence, a computer aided diagnostic system could be developed with proposed methodology for noninvasive diagnosis of knee joint disorders.

7.3 Future Scope

The noninvasive detection of knee joint disorders by VAG signals could provide a low cost diagnostic system. In the current investigation, different methodologies related to preprocessing, feature extraction and selection have been proposed. This was followed by a pattern classification using machine learning technique. This work may be extended in the following direction

1. The data set consists of 89 samples which is comparatively small. The proposed work could be carried out with a larger data set. Different researchers working in this area can develop a common data acquisition system. The acquired data from different researchers could be shared in cloud based environment so that larger database of knee joint disorder can be developed.
2. Deep learning based convolution neural network (CNN) can be used for the analysis of larger database. Recently, CNN has evolved as the most important research area, this requires a large training set.
3. Data acquisition could be carried out by placing multiple sensors across the knee joint. The signals obtained from the sensors could help in diagnosing the degree of knee joint disorders. Various disease related to knee joint could be identified. With this type of data acquisition the grades of knee joint disorders can be identified and this can help in the diagnosis, especially in the rehabilitation program.
4. Nonlinear signal processing techniques proposed in this thesis can be used for developing a prototype knee joint diagnostics system. This instrument could be available in a physician's clinic and can provide diagnostic information about the knee joint disorders. This device can provide a cost effective tool in healthcare industry.

Bibliography

- [1] M. Fransen, L. Bridgett, L. March, D. Hoy, E. Penserga, and P. Brooks, “The epidemiology of osteoarthritis in Asia,” *International Journal of Rheumatic Diseases*, vol. 14, no. 2, pp. 113–121, 2011, ISSN: 17561841. DOI: [10.1111/j.1756-185X.2011.01608.x](https://doi.org/10.1111/j.1756-185X.2011.01608.x).
- [2] C. P. Pal, P. Singh, S. Chaturvedi, K. K. Pruthi, and A. Vij, “Epidemiology of knee osteoarthritis in india and related factors,” *Indian journal of orthopaedics*, vol. 50, no. 5, p. 518, 2016.
- [3] J. R. Crim, B. J. Manaster, and Z. S. Rosenberg, *Diagnostic and Surgical Imaging Anatomy: Knee, ankle, foot*, 2nd Edition. Elsevier, 2017, p. 608, ISBN: 9780323477802.
- [4] C. B. Frank, R. Rangayyan, and G. D. Bell, “Analysis of Knee Joint Sound Signals for Non-Invasive Diagnosis of Cartilage Pathology,” *IEEE Engineering in Medicine and Biology Magazine*, vol. 9, no. 1, pp. 65–68, 1990, ISSN: 07395175. DOI: [10.1109/51.62910](https://doi.org/10.1109/51.62910).
- [5] G. F. McCoy, J. D. McCrea, D. E. Beverland, W. G. Kernohan, and R. Mollan, “Vibration arthrography as a diagnostic aid in diseases of the knee. A preliminary report.,” *The Journal of bone and joint surgery. British volume*, vol. 69, no. 2, pp. 288–293, 1987, ISSN: 02680033. DOI: [10.1016/0268-0033\(87\)90023-4](https://doi.org/10.1016/0268-0033(87)90023-4).
- [6] Y. F. Wu, *Knee Joint Vibroarthrographic Signal Processing and Analysis*. Springer-Verlag, 2015, pp. XIV, 81. DOI: [10.1007/978-3-662-44284-5](https://doi.org/10.1007/978-3-662-44284-5).
- [7] A. R. Webb, *Statistical pattern recognition*, Second Edition. John Wiley & Sons, 2003.
- [8] R. Rangayyan, S. Krishnan, G. G. D. Bell, C. B. Frank, and K. O. Ladly, “Parametric representation and screening of knee joint vibroarthrographic signals,” *IEEE Transactions on Biomedical Engineering*, vol. 44, no. 11, pp. 1068–1074, 1997, ISSN: 00189294. DOI: [10.1109/10.641334](https://doi.org/10.1109/10.641334).
- [9] K. S. Kim, J. H. Seo, J. U. Kang, and C. G. Song, “An enhanced algorithm for knee joint sound classification using feature extraction based on time-frequency analysis,” *Computer methods and programs in biomedicine*, vol. 94, no. 2, pp. 198–206, 2009.
- [10] K. S. Kim, J. H. Seo, and C. G. Song, “An acoustical evaluation of knee sound for non-invasive screening and early detection of articular pathology,” *Journal of Medical Systems*, vol. 36, no. 2, pp. 715–722, 2012, ISSN: 01485598. DOI: [10.1007/s10916-010-9539-3](https://doi.org/10.1007/s10916-010-9539-3).

- [11] G. Kernohan and R. Mollan, "Microcomputer analysis of joint vibration," *Journal of Microcomputer Applications*, vol. 5, no. 4, pp. 287–296, 1982, ISSN: 07457138. DOI: 10.1016/0745-7138(82)90018-5. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/0745713882900185>.
- [12] Mascaro, J. Prior, L.-K. Shark, J. Selfe, P Cole, and J. Goodacre, "Exploratory study of a non-invasive method based on acoustic emission for assessing the dynamic integrity of knee joints," *Medical engineering & physics*, vol. 31, no. 8, pp. 1013–1022, 2009.
- [13] L.-K. Shark, H. Chen, and J. Goodacre, "Knee acoustic emission: A potential biomarker for quantitative assessment of joint ageing and degeneration," *Medical Engineering and Physics*, vol. 33, no. 5, pp. 534–545, 2011, ISSN: 13504533. DOI: 10.1016/j.medengphy.2010.12.009.
- [14] D. Baczkowicz and K. Krecisz, "Zastosowanie vibroartrografia w diagnostyce narza{ogonek}du ruchu - Doniesienie wste{ogonek}pne," *Ortopedia Traumatologia Rehabilitacja*, vol. 15, no. 5, pp. 407–416, 2013, ISSN: 15093492. DOI: 10.5604/15093492.1084242.
- [15] D. Baczkowicz, E. Majorczyk, and K. Krecisz, "Age-related impairment of quality of joint motion in vibroarthrographic signal analysis," *BioMed Research International*, vol. 2015, 2015, ISSN: 23146141. DOI: 10.1155/2015/591707.
- [16] D. Baczkowicz and E. Majorczyk, "Joint motion quality in vibroacoustic signal analysis for patients with patellofemoral joint disorders," *BMC musculoskeletal disorders*, vol. 15, no. 1, p. 426, 2014, ISSN: 1471-2474. DOI: 10.1186/1471-2474-15-426.
- [17] D. Baczkowicz and E. Majorczyk, "Joint Motion Quality in Chondromalacia Progression Assessed by Vibroacoustic Signal Analysis," *PM and R*, vol. 8, no. 11, pp. 1065–1071, 2016, ISSN: 19341482. DOI: 10.1016/j.pmrj.2016.03.012.
- [18] R. Rangayyan and Y. F. Wu, "Screening of knee-joint vibroarthrographic signals using statistical parameters and radial basis functions," *Medical and Biological Engineering and Computing*, vol. 46, no. 3, pp. 223–232, 2008, ISSN: 01400118. DOI: 10.1007/s11517-007-0278-7.
- [19] R. Rangayyan, "Filtering for removal of artifacts," in *Biomedical Signal Analysis*, Second Edition. John Wiley And Sons, Inc., 2015, pp. 91–232, ISBN: 9780470544204. DOI: 10.1109/9780470544204.ch3. [Online]. Available: <http://dx.doi.org/10.1109/9780470544204.ch3>.
- [20] R. M. Rangayyan, "Analysis of nonstationary signals," in *Biomedical Signal Analysis*, Second Edition. John Wiley And Sons, Inc., 2015, pp. 469–570, ISBN: 9780470544204. DOI: 10.1109/9780470544204.ch8. [Online]. Available: <http://dx.doi.org/10.1109/9780470544204.ch8>.

- [21] Y. T. Zhang, R. Rangayyan, G. Bell, and C. Frank, “Adaptive cancellation of muscle contraction interference in vibroarthrographic signals,” *IEEE Transactions on Biomedical Engineering*, vol. 41, no. 2, pp. 181–191, 1994, ISSN: 15582531. DOI: 10.1109/10.284929.
- [22] S. Krishnan, R. Rangayyan, G. G. D. Bell, C. B. Frank, and K. O. Ladly, “Adaptive filtering, modelling and classification of knee joint vibroarthrographic signals for non-invasive diagnosis of articular cartilage pathology,” *Medical and Biological Engineering and Computing*, vol. 35, no. 6, pp. 677–684, 1997, ISSN: 0140-0118. DOI: 10.1007/BF02510977.
- [23] Y. Wu, S. Cai, M. Lu, S. Yang, F. Zheng, N. Xiang, J. He, and Z. Zhong, “Noise Cancellation in Knee Joint Vibration Signals Using a Time-Delay Neural Filter and Signal Power Error Minimization Method,” vol. 8, 2013, pp. 912–919. DOI: 10.4156/jcit.vol8.issue5.106.
- [24] S. Cai, Y. Wu, N. Xiang, Z. Zhong, J. He, L. Shi, and F. Xu, “Detrending knee joint vibration signals with a cascade moving average filter,” in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 2012, pp. 4357–4360, ISBN: 9781424441198. DOI: 10.1109/EMBC.2012.6346931.
- [25] D. Ingrid, *Ten lectures on wavelets: CBMS-NSF Regional Conference Series in Applied Mathematics*. SIAM, Philadelphia, USA, 1992, p. 377.
- [26] D. Zennaro, P. Wellig, V. M. Koch, G. S. Moschytz, and T. Laubli, “A software package for the decomposition of long-term multichannel emg signals using wavelet coefficients,” *IEEE Transactions on Biomedical Engineering*, vol. 50, no. 1, pp. 58–69, 2003, ISSN: 0018-9294. DOI: 10.1109/TBME.2002.807321.
- [27] L. L. Chen, J. Zhang, J. Z. Zou, C. J. Zhao, and G. S. Wang, “A framework on wavelet-based nonlinear features and extreme learning machine for epileptic seizure detection,” *Biomedical Signal Processing and Control*, vol. 10, no. 1, pp. 1–10, 2014, ISSN: 17468094. DOI: 10.1016/j.bspc.2013.11.010.
- [28] S. Abbaspour, A. Fallah, M. Lindén, and H. Gholamhosseini, “A novel approach for removing ecg interferences from surface emg signals using a combined anfis and wavelet,” *Journal of Electromyography and Kinesiology*, vol. 26, pp. 52–59, 2016. DOI: <https://doi.org/10.1016/j.jelekin.2015.11.003>.
- [29] Y. Zhang, B. Liu, X. Ji, and D. Huang, “Classification of eeg signals based on autoregressive model and wavelet packet decomposition,” *Neural Processing Letters*, pp. 1–14, 2016.
- [30] M. Tarata, M. Serbanescu, D. Georgescu, D. Alexandru, and W. Wolf, “A new wavelet based approach for the noninvasive dynamic detection of m1 & m2 emg stretch reflex response components – preliminary results,” in *XIV Mediterranean Conference on Medical and Biological Engineering and Computing 2016: MEDICON 2016, March 31st-April 2nd 2016, Paphos, Cyprus*. Cham: Springer International Publishing, 2016, pp. 46–49, ISBN: 978-3-319-32703-7. DOI: 10.1007/978-3-319-32703-7_10.

- [31] J. Kevric and A. Subasi, “Comparison of signal decomposition methods in classification of eeg signals for motor-imagery bci system,” *Biomedical Signal Processing and Control*, vol. 31, pp. 398–406, 2017, ISSN: 17468108. DOI: [10.1016/j.bspc.2016.09.007](https://doi.org/10.1016/j.bspc.2016.09.007).
- [32] S. Xie and S. Krishnan, “Wavelet-based sparse functional linear model with applications to EEGs seizure detection and epilepsy diagnosis,” *Medical and Biological Engineering and Computing*, vol. 51, no. 1-2, pp. 49–60, 2013, ISSN: 01400118. DOI: [10.1007/s11517-012-0967-8](https://doi.org/10.1007/s11517-012-0967-8).
- [33] R. R. Coifman, Y. Meyer, and V. Wickerhauser, “Wavelet analysis and signal processing,” in *In Wavelets and their Applications*, 1992, pp. 153–178.
- [34] S. Mallat, *A Wavelet Tour of Signal Processing The Sparse Way*. Academic press, 2008, ISBN: 9780123743701.
- [35] P. Comon, “Independent component analysis, A new concept?” *Signal Processing*, vol. 36, no. 3, pp. 287–314, 1994, ISSN: 01651684. DOI: [10.1016/0165-1684\(94\)90029-9](https://doi.org/10.1016/0165-1684(94)90029-9).
- [36] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent component analysis*. John Wiley & Sons, 2004, vol. 46.
- [37] A. Hyvärinen and E. Oja, “Independent component analysis: algorithms and applications,” *Neural networks*, vol. 13, no. 4-5, pp. 411–430, 2000, ISSN: 0893-6080. DOI: [10.1016/S0893-6080\(00\)00026-5](https://doi.org/10.1016/S0893-6080(00)00026-5). arXiv: [1504.05070](https://arxiv.org/abs/1504.05070).
- [38] C. J. James and C. W. Hesse, “Independent component analysis for biomedical signals,” *Physiological measurement*, vol. 26, no. 1, R15, 2004.
- [39] T.-P. Jung, S. Makeig, T.-W. Lee, M. J. McKeown, G. Brown, A. J. Bell, and T. J. Sejnowski, “Independent component analysis of biomedical signals,” in *Proc. Int. Workshop on Independent Component Analysis and Signal Separation*, 2000, pp. 633–644.
- [40] A. Delorme and S. Makeig, “Eeglab: An open source toolbox for analysis of single-trial eeg dynamics including independent component analysis,” *Journal of neuroscience methods*, vol. 134, no. 1, pp. 9–21, 2004.
- [41] J.-C. Chen, P.-C. Tung, S.-F. F. Huang, S.-L. L. Lin, S.-W. Wu, and S.-L. L. Lin, “Extraction and screening of knee joint vibroarthrographic signals using the independent component analysis method,” *International Journal of Innovative Computing, Information and Control*, vol. 8, no. 11, pp. 7501–7518, 2012.
- [42] K. Dragomiretskiy and D. Zosso, “Variational mode decomposition,” *IEEE Transactions on Signal Processing*, vol. 62, no. 3, pp. 531–544, 2014, ISSN: 1053587X. DOI: [10.1109/TSP.2013.2288675](https://doi.org/10.1109/TSP.2013.2288675).

- [43] A. Sundar, C. Das, and V. Pahwa, “Denoising knee joint vibration signals using variational mode decomposition,” in *Information Systems Design and Intelligent Applications: Proceedings of Third International Conference INDIA 2016, Volume 1*, S. C. Satapathy, J. K. Mandal, S. K. Udgata, and V. Bhateja, Eds. New Delhi: Springer India, 2016, pp. 719–729, ISBN: 978-81-322-2755-7. DOI: [10.1007/978-81-322-2755-7_74](https://doi.org/10.1007/978-81-322-2755-7_74).
- [44] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, “The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis,” in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, The Royal Society, vol. 454, 1998, pp. 903–995.
- [45] Y. Wu, S. Cai, F. Xu, L. Shi, and S. Krishnan, “Chondromalacia patellae detection by analysis of intrinsic mode functions in knee joint vibration signals,” in *World Congress on Medical Physics and Biomedical Engineering May 26-31, 2012, Beijing, China*, M. Long, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 493–496, ISBN: 978-3-642-29305-4. DOI: [10.1007/978-3-642-29305-4_130](https://doi.org/10.1007/978-3-642-29305-4_130).
- [46] Z. Wu and N. E. Huang, “Ensemble Empirical Mode Decomposition: a Noise-Assisted Data Analysis Method,” *Advances in Adaptive Data Analysis*, vol. 01, no. 01, pp. 1–41, 2009, ISSN: 1793-5369. DOI: [10.1142/S1793536909000047](https://doi.org/10.1142/S1793536909000047).
- [47] Y. F. Wu, S. Yang, F. Zheng, S. Cai, M. Lu, and M. Wu, “Removal of artifacts in knee joint vibroarthrographic signals using ensemble empirical mode decomposition and detrended fluctuation analysis.,” *Physiological measurement*, vol. 35, no. 3, pp. 429–39, 2014. DOI: [10.1088/0967-3334/35/3/429](https://doi.org/10.1088/0967-3334/35/3/429).
- [48] Y. Wu, P. Chen, X. Luo, H. Huang, L. Liao, Y. Yao, M. Wu, and R. Rangayyan, “Quantification of knee vibroarthrographic signal irregularity associated with patellofemoral joint cartilage pathology based on entropy and envelope amplitude measures,” *Computer Methods and Programs in Biomedicine*, vol. 130, pp. 1–12, 2016, ISSN: 18727565. DOI: [10.1016/j.cmpb.2016.03.021](https://doi.org/10.1016/j.cmpb.2016.03.021).
- [49] N. E. Huang and Z. Wu, “A review on hilbert-huang transform : Method and its applications to geophysical studies,” *October*, vol. 46, no. 2007, pp. 1–23, 2008, ISSN: 87551209.
- [50] N. E. Huang, *Hilbert-Huang transform And Its Applications*. World Scientific, 2014, vol. 16.
- [51] N. E. Huang, Z. Shen, and S. R. Long, “A new view of nonlinear water waves: The hilbert spectrum,” *Annual Review of Fluid Mechanics*, vol. 31, no. 1, pp. 417–457, 1999, ISSN: 0066-4189. DOI: [10.1146/annurev.fluid.31.1.417](https://doi.org/10.1146/annurev.fluid.31.1.417).
- [52] M. A. Colominas, G. Schlotthauer, and M. E. Torres, “Improved complete ensemble EMD: A suitable tool for biomedical signal processing,” *Biomedical Signal Processing and Control*, vol. 14, no. 1, pp. 19–29, 2014, ISSN: 17468108. DOI: [10.1016/j.bspc.2014.06.009](https://doi.org/10.1016/j.bspc.2014.06.009).

- [53] M. A. Colominas, G. Schlotthauer, and M. E. Torres, “Complete ensemble emd and hilbert transform for heart beat detection,” in *VI Latin American Congress on Biomedical Engineering CLAIB 2014, Paraná, Argentina 29, 30 & 31 October 2014*, A. Braidot and A. Hadad, Eds. Springer International Publishing, 2015, pp. 496–499, ISBN: 978-3-319-13117-7. DOI: [10.1007/978-3-319-13117-7_127](https://doi.org/10.1007/978-3-319-13117-7_127).
- [54] A. R. Hassan and M. I. H. Bhuiyan, “Computer-aided sleep staging using Complete Ensemble Empirical Mode Decomposition with Adaptive Noise and bootstrap aggregating,” *Biomedical Signal Processing and Control*, vol. 24, pp. 1–10, 2016, ISSN: 17468108. DOI: [10.1016/j.bspc.2015.09.002](https://doi.org/10.1016/j.bspc.2015.09.002).
- [55] A. R. Hassan and A. Subasi, “Automatic identification of epileptic seizures from EEG signals using linear programming boosting,” *Computer Methods and Programs in Biomedicine*, vol. 136, pp. 65–77, 2016, ISSN: 18727565. DOI: [10.1016/j.cmpb.2016.08.013](https://doi.org/10.1016/j.cmpb.2016.08.013).
- [56] Y. Lei, Z. Liu, J. Ouazri, and J. Lin, “A fault diagnosis method of rolling element bearings based on CEEMDAN,” *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, vol. 0, no. 28, pp. 1–12, 2015, ISSN: 0954-4062. DOI: [10.1177/0954406215624126](https://doi.org/10.1177/0954406215624126).
- [57] Y. Ren, P. Suganthan, and N. Srikanth, “A comparative study of empirical mode decomposition-based short-term wind speed forecasting methods,” *IEEE Transactions on Sustainable Energy*, vol. 6, no. 1, pp. 236–244, 2015.
- [58] B. Boashash, *Time frequency signal analysis and processing: a comprehensive reference*, Second edition. Academic Press, 2015, p. 770, ISBN: 0080443354.
- [59] P. Addison, J. Walker, and R. Guido, “Time–frequency analysis of biosignals,” *IEEE engineering in medicine and biology magazine*, vol. 28, no. 5, 2009.
- [60] M. Wacker and H. Witte, “Time-frequency techniques in biomedical signal analysis,” *Methods of information in medicine*, vol. 52, no. 4, pp. 279–296, 2013.
- [61] S. G. Mallat and Z. Zhang, “Matching pursuits with time-frequency dictionaries,” *IEEE Transactions on signal processing*, vol. 41, no. 12, pp. 3397–3415, 1993.
- [62] S. Krishnan, R. Rangayyan, G. G. D. Bell, and C. B. Frank, “Time-frequency signal feature extraction and screening of knee joint vibroarthrographic signals using the matching pursuit method,” in *Engineering in Medicine and Biology Society, 1997. Proceedings of the 19th Annual International Conference of the IEEE*, IEEE, vol. 3, 1997, pp. 1309–1312.
- [63] S. Krishnan, R. Rangayyan, G. D. Bell, and C. B. Frank, “Adaptive time-frequency analysis of knee joint vibroarthrographic signals for noninvasive screening of articular cartilage pathology,” *IEEE Transactions on Biomedical Engineering*, vol. 47, no. 6, pp. 773–783, 2000.

- [64] Y. Wu, S. Krishnan, and R. Rangayyan, “Computer-aided diagnosis of knee-joint disorders via vibroarthrographic signal analysis: A review,” *Critical ReviewsTM in Biomedical Engineering*, vol. 38, no. 2, 2010.
- [65] K. Umapathy and S. Krishnan, “Modified local discriminant bases algorithm and its application in analysis of human knee joint vibration signals,” *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 3, pp. 517–523, 2006, ISSN: 00189294. DOI: 10.1109/TBME.2005.869787.
- [66] Y. Wu and S. Krishnan, “Classification of knee-joint vibroarthrographic signals using time-domain and time-frequency domain features and least-squares support vector machine,” in *Digital Signal Processing, 2009 16th International Conference on*, IEEE, 2009, pp. 1–6.
- [67] J.-C. Chen, P.-C. Tung, S.-F. Huang, S.-W. Wu, S.-L. Lin, and K.-L. Tu, “Extraction and screening of knee joint vibroarthrographic signals using the empirical mode decomposition method,” *International Journal of Innovative Computing, Information and Control*, vol. 9, no. 6, pp. 2689–2700, 2013.
- [68] F. Hlawatsch, T. G. Manickam, R. L. Urbanke, and W. Jones, “Smoothed pseudo-Wigner distribution, Choi-Williams distribution, and cone-kernel representation: Ambiguity-domain analysis and experimental comparison,” *Signal Processing*, vol. 43, no. 2, pp. 149–168, 1995, ISSN: 01651684. DOI: 10.1016/0165-1684(94)00150-X.
- [69] W. Liu and B. Tang, “A hybrid time-frequency method based on improved Morlet wavelet and auto terms window,” *Expert Systems with Applications*, vol. 38, no. 6, pp. 7575–7581, 2011, ISSN: 09574174. DOI: 10.1016/j.eswa.2010.12.107.
- [70] B. Tang, W. Liu, and T. Song, “Wind turbine fault diagnosis based on Morlet wavelet transformation and Wigner-Ville distribution,” *Renewable Energy*, vol. 35, no. 12, pp. 2862–2866, 2010, ISSN: 09601481. DOI: 10.1016/j.renene.2010.05.012.
- [71] N. Baydar and A. Ball, “A comparative study of acoustic and vibration signals in detection of gear failures using Wigner-Ville distribution,” *Mechanical Systems and Signal Processing*, vol. 15, no. 6, pp. 1091–1107, 2001, ISSN: 08883270. DOI: 10.1006/mssp.2000.1338.
- [72] F. M. Villalobos-Castaldi, J. Ruiz-Pinales, N. C. K. Valverde, and M. Flores, “Time-frequency analysis of spontaneous pupillary oscillation signals using the hilbert-huang transform,” *Biomedical Signal Processing and Control*, vol. 30, pp. 106–116, 2016, ISSN: 17468108. DOI: 10.1016/j.bspc.2016.06.002.
- [73] Y. Lei, J. Lin, Z. He, and M. J. Zuo, “A review on empirical mode decomposition in fault diagnosis of rotating machinery,” *Mechanical Systems and Signal Processing*, vol. 35, no. 1, pp. 108–126, 2013.
- [74] I. Guyon and A. Elisseeff, “An introduction to feature extraction,” in *Feature extraction*, Springer, 2006, pp. 1–25.

- [75] B. D. Ripley, *Pattern Recognition and Neural Networks*, 2nd Edition. New York, NY, USA: Cambridge University Press, 2007.
- [76] N. P. Reddy, B. M. Rothschild, M. Mandal, V. Gupta, and S. Suryanarayanan, “Non-invasive acceleration measurements to characterize knee arthritis and chondromalacia,” *Annals of Biomedical Engineering*, vol. 23, no. 1, pp. 78–84, 1995, ISSN: 00906964. DOI: 10.1007/BF02368303.
- [77] S. Tavathia, R. Rangayyan, C. B. Frank, G. D. Bell, and K. O. Ladly, “Analysis of knee vibration signals using linear prediction,” *IEEE transactions on biomedical engineering*, vol. 39, no. 9, pp. 959–70, 1992, ISSN: 0018-9294. DOI: 10.1109/10.256430.
- [78] Z. M. K. Moussavi, R. Rangayyan, G. D. Bell, C. B. Frank, K. O. Ladly, and Y. T. Zhang, “Screening of vibroarthrographic signals via adaptive segmentation and linear prediation modeling,” *IEEE Transactions on Biomedical Engineering*, vol. 43, no. 1, pp. 15–23, 1996, ISSN: 00189294. DOI: 10.1109/10.477697.
- [79] S. Krishnan, R. Rangayyan, G. D. Bell, C. B. Frank, and K. O. Ladly, “Screening of knee joint vibroarthrographic signals by statistical pattern analysis of dominant poles,” *Proceedings of 18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 3, no. 2, pp. 968–969, 1996, ISSN: 05891019. DOI: 10.1109/IEMBS.1996.652664.
- [80] S. Krishnan, R. M. Rangayyan, G. D. Bell, C. B. Frank, and K. O. Ladly, “Recursive least-squares lattice-based adaptive segmentation and autoregressive modeling of knee joint vibroarthrographic signals,” in *Proceedings of 1996 Canadian Conference on Electrical and Computer Engineering*, vol. 1, 1996, pp. 339–342, ISBN: 0-7803-3143-5. DOI: 10.1109/CCECE.1996.548106. [Online]. Available: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=548106>.
- [81] R. Rangayyan and Y. Wu, “Analysis of knee-joint vibroarthrographic signals using statistical measures,” in *Proceedings - IEEE Symposium on Computer-Based Medical Systems*, 2007, pp. 377–382, ISBN: 0769529054. DOI: 10.1109/CBMS.2007.23.
- [82] R. M. Rangayyan and Y. Wu, “Analysis of vibroarthrographic signals with features related to signal variability and radial-basis functions,” *Annals of Biomedical Engineering*, vol. 37, no. 1, pp. 156–163, 2009, ISSN: 00906964. DOI: 10.1007/s10439-008-9601-1.
- [83] R. M. Rangayyan and Y. Wu, “Screening of knee-joint vibroarthrographic signals using probability density functions estimated with Parzen windows,” *Biomedical Signal Processing and Control*, vol. 5, no. 1, pp. 53–58, 2010, ISSN: 17468094. DOI: 10.1016/j.bspc.2009.03.008.

- [84] Y. Wu, S. Cai, S. Yang, F. Zheng, and N. Xiang, "Classification of Knee Joint Vibration Signals Using Bivariate Feature Distribution Estimation and Maximal Posterior Probability Decision Criterion," *Entropy*, vol. 15, no. 4, pp. 1375–1387, 2013, ISSN: 1099-4300. DOI: 10.3390/e15041375. [Online]. Available: <http://www.mdpi.com/1099-4300/15/4/1375/>.
- [85] S. Cai, S. Yang, F. Zheng, M. Lu, Y. Wu, and S. Krishnan, "Knee joint vibration signal analysis with matching pursuit decomposition and dynamic weighted classifier fusion," *Computational and Mathematical Methods in Medicine*, vol. 2013, 2013, ISSN: 1748670X. DOI: 10.1155/2013/904267.
- [86] R. Rangayyan, F. Oloumi, Y. Wu, and S. Cai, "Fractal analysis of knee-joint vibroarthrographic signals via power spectral analysis," *Biomedical Signal Processing and Control*, vol. 8, no. 1, pp. 23–29, 2013, ISSN: 17468094. DOI: 10.1016/j.bspc.2012.05.004. [Online]. Available: <http://dx.doi.org/10.1016/j.bspc.2012.05.004>.
- [87] S. Yang, S. Cai, F. Zheng, Y. Wu, K. Liu, M. Wu, Q. Zou, and J. Chen, "Representation of fluctuation features in pathological knee joint vibroarthrographic signals using kernel density modeling method," *Medical Engineering and Physics*, vol. 36, no. 10, pp. 1305–1311, 2014, ISSN: 18734030. DOI: 10.1016/j.medengphy.2014.07.008.
- [88] Cirugeda-Roldan, C.-F. D, Miro-Martinez, Oltra-Crespo, Vigil-Medina, and Varela-Entrecanales, "A new algorithm for quadratic sample entropy optimization for very short biomedical signals: Application to blood pressure records.,," *Computer methods and programs in biomedicine*, vol. 114, no. 3, pp. 231–9, 2014, ISSN: 1872-7565. DOI: 10.1016/j.cmpb.2014.02.008.
- [89] R. Acharya, O. Faust, V. Sree, G Swapna, R. J. Martis, N. A. Kadri, and J. S. Suri, "Linear and nonlinear analysis of normal and CAD-affected heart rate signals.,," *Computer methods and programs in biomedicine*, vol. 113, no. 1, pp. 55–68, 2014, ISSN: 1872-7565. DOI: 10.1016/j.cmpb.2013.08.017.
- [90] K. Natarajan, M. L. Choo, A. Rajendra, and S. Puthusserypady, "Entropies for detection of epilepsy in eeg," *Computer methods and programs in biomedicine*, vol. 80, no. 3, pp. 187–94, 2005, ISSN: 0169-2607. DOI: 10.1016/j.cmpb.2005.06.012.
- [91] R. Sharma, R. B. Pachori, and R. Acharya, "An integrated index for the identification of focal electroencephalogram signals using discrete wavelet transform and entropy measures," *Entropy*, vol. 17, no. 8, pp. 5218–5240, 2015, ISSN: 10994300. DOI: 10.3390/e17085218.
- [92] C. L. Webber and J. P. Zbilut, "Dynamical assessment of physiological systems and states using recurrence plot strategies," *Journal of Applied Physiology*, vol. 76, no. 2, pp. 965–973, 1994, ISSN: 8750-7587, 1522-1601.

- [93] J. Gao and H. Cai, “On the structures and quantification of recurrence plots,” *Physics Letters, Section A: General, Atomic and Solid State Physics*, vol. 270, pp. 75–87, 2000, ISSN: 03759601. DOI: [10.1016/S0375-9601\(00\)00304-2](https://doi.org/10.1016/S0375-9601(00)00304-2).
- [94] M. E. Cohen, D. L. Hudson, and P. Ć. Deedwania, “Applying continuous chaotic modeling to cardiac signal analysis,” *IEEE Engineering in Medicine and Biology Magazine*, vol. 15, no. 5, pp. 97–102, 1996. DOI: [10.1109/51.537065](https://doi.org/10.1109/51.537065).
- [95] K. Kira and L. A. Rendell, “The feature selection problem: Traditional methods and a new algorithm,” in *AAAI*, vol. 2, 1992, pp. 129–134. DOI: [10.1016/S0031-3203\(01\)00046-2](https://doi.org/10.1016/S0031-3203(01)00046-2).
- [96] P. Langley, “Selection of Relevant Features in Machine Learning,” in *In Proceedings of the AAAI Fall Symposium on Relevance*, Elsevier, 1994, pp. 140–144, ISBN: 0004-3702. DOI: [10.1.1.43.4648](https://doi.org/10.1.1.43.4648).
- [97] G. H. John, R. Kohavi, K. Pfleger, *et al.*, “Irrelevant features and the subset selection problem,” in *Machine learning: proceedings of the eleventh international conference*, 1994, pp. 121–129. DOI: [10.1145/2663598.2629474](https://doi.org/10.1145/2663598.2629474).
- [98] M. Dash and H. Liu, “Feature selection for classification,” *Intelligent data analysis*, vol. 1, no. 1-4, pp. 131–156, 1997. DOI: [10.3233/IDA-1997-1302](https://doi.org/10.3233/IDA-1997-1302).
- [99] N. Kwak and C. H. Choi, “Input feature selection by mutual information based on Parzen window,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 12, pp. 1667–1671, 2002, ISSN: 01628828. DOI: [10.1109/TPAMI.2002.1114861](https://doi.org/10.1109/TPAMI.2002.1114861).
- [100] G. Chandrashekhar and F. Sahin, “A survey on feature selection methods,” *Computers and Electrical Engineering*, vol. 40, no. 1, pp. 16–28, 2014, ISSN: 00457906. DOI: [10.1016/j.compeleceng.2013.11.024](https://doi.org/10.1016/j.compeleceng.2013.11.024).
- [101] M. Bennasar, Y. Hicks, and R. Setchi, “Feature selection using Joint Mutual Information Maximisation,” *Expert Systems with Applications*, vol. 42, no. 22, pp. 8520–8532, 2015, ISSN: 09574174. DOI: [10.1016/j.eswa.2015.07.007](https://doi.org/10.1016/j.eswa.2015.07.007).
- [102] A. L. Blum and P. Langley, “Selection of relevant features and examples in machine learning,” *Artificial intelligence*, vol. 97, no. 1, pp. 245–271, 1997.
- [103] J. Tang, S. Alelyani, and H. Liu, “Feature selection for classification: A review,” *Data Classification: Algorithms and Applications*, p. 37, 2014. DOI: [10.1.1.409.5195](https://doi.org/10.1.1.409.5195).
- [104] V. Kumar and S. Minz, “Feature selection : A literature review,” *Smart Computing Review*, vol. 4, no. 3, pp. 211–229, 2014. DOI: [10.6029/smarter.2014.03.007](https://doi.org/10.6029/smarter.2014.03.007).
- [105] Y. Saeys, I. Inza, and P. Larrañaga, “A review of feature selection techniques in bioinformatics,” *Bioinformatics*, vol. 23, no. 19, pp. 2507–2517, 2007. DOI: [10.1093/bioinformatics/btm344](https://doi.org/10.1093/bioinformatics/btm344).

- [106] L. Zoubek, S. Charbonnier, S. Lesecq, A. Buguet, and F. Chapotot, “Feature selection for sleep/wake stages classification using data driven methods,” *Biomedical Signal Processing and Control*, vol. 2, no. 3, pp. 171–179, 2007, ISSN: 17468094. DOI: [10.1016/j.bspc.2007.05.005](https://doi.org/10.1016/j.bspc.2007.05.005).
- [107] M. D’Alessandro, R. Esteller, G. Vachtsevanos, A. Hinson, J. Echauz, and B. Litt, “Epileptic seizure prediction using hybrid feature selection over multiple intracranial EEG electrode contacts: a report of four patients,” *IEEE Transactions on Biomedical Engineering*, vol. 50, no. 5, pp. 603–615, 2003, ISSN: 0018-9294. DOI: [10.1109/TBME.2003.810706](https://doi.org/10.1109/TBME.2003.810706).
- [108] A. Phinyomark, P. Phukpattaranont, and C. Limsakul, “Feature reduction and selection for EMG signal classification,” *Expert Systems with Applications*, vol. 39, no. 8, pp. 7420–7431, 2012, ISSN: 09574174. DOI: [10.1016/j.eswa.2012.01.102](https://doi.org/10.1016/j.eswa.2012.01.102).
- [109] J. Liu, “Feature dimensionality reduction for myoelectric pattern recognition: A comparison study of feature selection and feature projection methods,” *Medical Engineering and Physics*, vol. 36, no. 12, pp. 1716–1720, 2014, ISSN: 18734030. DOI: [10.1016/j.medengphy.2014.09.011](https://doi.org/10.1016/j.medengphy.2014.09.011).
- [110] T. Mar, S. Zaunseder, J. P. Martínez, M. Llamedo, and R. Poll, “Optimization of ECG classification by means of feature selection,” *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 8, pp. 2168–2177, 2011, ISSN: 00189294. DOI: [10.1109/TBME.2011.2113395](https://doi.org/10.1109/TBME.2011.2113395).
- [111] C. Y. Lee and Y. X. Shen, “Optimal feature selection for power-quality disturbances classification,” *IEEE Transactions on Power Delivery*, vol. 26, no. 4, pp. 2342–2351, 2011, ISSN: 08858977. DOI: [10.1109/TPWRD.2011.2149547](https://doi.org/10.1109/TPWRD.2011.2149547).
- [112] B. K. Panigrahi and R. Pandi, “Optimal feature selection for classification of power quality disturbances using wavelet packet-based fuzzy k-nearest neighbour algorithm,” *IET Generation, Transmission & Distribution*, vol. 3, no. 3, pp. 296–306, 2009, ISSN: 1751-8687. DOI: [10.1049/iet-gtd:20080190](https://doi.org/10.1049/iet-gtd:20080190).
- [113] T. Mu, A. K. Nandi, and R. Rangayyan, “Screening of knee-joint vibroarthrographic signals using the strict 2-surface proximal classifier and genetic algorithm,” *Computers in Biology and Medicine*, vol. 38, pp. 1103–11, 2008, ISSN: 0010-4825 (Print) 0010-4825 (Linking). DOI: [10.1016/j.combiomed.2008.08.009](https://doi.org/10.1016/j.combiomed.2008.08.009).
- [114] R. Agrawal and R. Srikant, “Fast Algorithms for Mining Association Rules in Large Databases,” in *Journal of Computer Science and Technology*, vol. 15, 1994, pp. 487–499, ISBN: 1558601538. DOI: [10.1007/BF02948845](https://doi.org/10.1007/BF02948845).
- [115] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*. 1989, vol. Addison-We, p. 432, ISBN: 0201157675. DOI: [10.1007/s10589-009-9261-6](https://doi.org/10.1007/s10589-009-9261-6).

- [116] O. Ludwig and U. Nunes, “Novel maximum-margin training algorithms for supervised neural networks,” *IEEE Transactions on Neural Networks*, vol. 21, no. 6, pp. 972–984, 2010, ISSN: 10459227. DOI: 10.1109/TNN.2010.2046423.
- [117] I. Rejer, “Genetic algorithm with aggressive mutation for feature selection in bci feature space,” *Pattern Analysis and Applications*, vol. 18, no. 3, pp. 485–492, 2015.
- [118] S. Oreski and G. Oreski, “Genetic algorithm-based heuristic for feature selection in credit risk assessment,” *Expert systems with applications*, vol. 41, no. 4, pp. 2052–2064, 2014.
- [119] A. S. Ghareb, A. A. Bakar, and A. R. Hamdan, “Hybrid feature selection based on enhanced genetic algorithm for text categorization,” *Expert Systems with Applications*, vol. 49, pp. 31–47, 2016.
- [120] M. Kantardzic, *Data mining: concepts, models, methods, and algorithms*. John Wiley & Sons, 2011.
- [121] L. Hanguang and N. Yu, “Intrusion Detection Technology Research Based on Apriori Algorithm,” in *Physics Procedia*, vol. 24, 2012, pp. 1615–1620. DOI: 10.1016/j.phpro.2012.02.238.
- [122] H. Erişti, Ö. Yıldırım, B. Erişti, and Y. Demir, “Optimal feature selection for classification of the power quality events using wavelet transform and least squares support vector machines,” *International Journal of Electrical Power & Energy Systems*, vol. 49, pp. 95–103, 2013, ISSN: 01420615. DOI: 10.1016/j.ijepes.2012.12.018.
- [123] N. Sharma and H. Om, “Extracting Significant Patterns for Oral Cancer Detection Using Apriori Algorithm,” *Intelligent Information Management*, vol. 06, no. 02, pp. 30–37, 2014, ISSN: 2160-5912. DOI: 10.4236/iim.2014.62005.
- [124] Z. Guo, D. Chi, J. Wu, and W. Zhang, “A new wind speed forecasting strategy based on the chaotic time series modelling technique and the Apriori algorithm,” *Energy Conversion and Management*, vol. 84, pp. 140–151, 2014, ISSN: 01968904. DOI: 10.1016/j.enconman.2014.04.028.
- [125] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. New York: John Wiley, Section, 2001, p. 680. DOI: 10.1007/BF01237942.
- [126] C. M. Bishop, *Pattern Recognition and Machine Learning*, 4. Springer, New York, 2006, p. 738. DOI: 10.1111/1.2819119.
- [127] R. Rangayyan, “Pattern classification and diagnostic decision,” in *Biomedical Signal Analysis*, Second Edition. John Wiley And Sons, Inc., 2015, pp. 571–632, ISBN: 9780470544204. DOI: 10.1109/9780470544204.ch9. [Online]. Available: <http://dx.doi.org/10.1109/9780470544204.ch9>.

- [128] A. T. Azar and S. A. El-Said, "Performance analysis of support vector machines classifiers in breast cancer mammography recognition," English, *Neural Computing and Applications*, vol. 24, no. 5, pp. 1163–1177, 2014, ISSN: 0941-0643. DOI: 10.1007/s00521-012-1324-4.
- [129] S. H. Park, J. M. Goo, and C.-H. Jo, "Receiver Operating Characteristic (ROC) Curve: Practical Review for Radiologists," *Korean Journal of Radiology*, vol. 5, no. 1, p. 11, 2004, ISSN: 1229-6929. DOI: 10.3348/kjr.2004.5.1.11.
- [130] M. H. Zweig and G. Campbell, "Receiver-operating characteristic (roc) plots: A fundamental evaluation tool in clinical medicine.," *Clinical Chemistry*, vol. 39, no. 4, pp. 561–577, 1993, ISSN: 0009-9147. eprint: <http://clinchem.aaccjnls.org/content/39/4/561.full.pdf>. [Online]. Available: <http://clinchem.aaccjnls.org/content/39/4/561>.
- [131] M. Greiner, D. Pfeiffer, and R. Smith, "Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests," vol. 45, no. 1-2, pp. 23–41, 2000, ISSN: 01675877. DOI: 10.1016/S0167-5877(00)00115-X.
- [132] K. Umapathy and S. Krishnan, "A signal classification approach using time-width vs frequency band sub-energy distributions," in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, vol. V, 2005, pp. 477–480, ISBN: 0780388747. DOI: 10.1109/ICASSP.2005.1416344.
- [133] T. Mu, A. K. Nandi, and R. Rangayyan, "Strict 2-surface proximal classification of knee-joint vibroarthrographic signals," in *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*, 2007, pp. 4911–4914, ISBN: 1424407885. DOI: 10.1109/IEMBS.2007.4353441.
- [134] C. G. Song, K. S. Kim, and J. H. Seo, "Non-invasive Monitoring of Knee Pathology based on Automatic Knee Sound Classification," in *Proceedings of the World Congress on Engineering and Computer Science 2009*, vol. 1, 2009, pp. 444–448, ISBN: 978-988-17012-6-8.
- [135] K. S. Kim, C. G. Song, J. H. Seo, and K. S. Kim, "An efficient algorithm to improve feature extraction and classification of knee joint sound," in *2009 ICME International Conference on Complex Medical Engineering, CME 2009*, IEEE, 2009, pp. 1–6, ISBN: 9781424433162. DOI: 10.1109/ICCME.2009.4906598.
- [136] Y. Wu and S. Krishnan, "Combining least-squares support vector machines for classification of biomedical signals: A case study with knee-joint vibroarthrographic signals," *Journal of Experimental & Theoretical Artificial Intelligence*, vol. 23, no. 1, pp. 63–77, 2011. DOI: 10.1080/0952813X.2010.506288.
- [137] K. Liu, X. Luo, S. Yang, S. Cai, F. Zheng, and Y. Wu, "Classification of knee joint vibroarthrographic signals using k-nearest neighbor algorithm," in *Canadian Conference on Electrical and Computer Engineering*, IEEE, 2014, ISBN: 9781479930999. DOI: 10.1109/CCECE.2014.6900933.

- [138] V. N. Vapnik and A. Y. Chervonenkis, *On the Uniform Convergence of Relative Frequencies of Events to Their Probabilities*, 1971. DOI: [10.1137/1116025](https://doi.org/10.1137/1116025).
- [139] V. N. Vapnik, *The Nature of Statistical Learning Theory*. IEEE Transactions on Neural Networks, 1995, vol. 8, p. 1564, ISBN: 0387945598. DOI: [10.1109/TNN.1997.641482](https://doi.org/10.1109/TNN.1997.641482).
- [140] J. A. Suykens and J. Vandewalle, “Least Squares Support Vector Machine Classifiers,” *Neural Processing Letters*, vol. 9, pp. 293–300, 1999, ISSN: 13704621. DOI: [10.1023/A:1018628609742](https://doi.org/10.1023/A:1018628609742).
- [141] J. A. Suykens, J. De Brabanter, L. Lukas, and J. Vandewalle, “Weighted least squares support vector machines: Robustness and sparse approximation,” *Neurocomputing*, vol. 48, pp. 85–105, 2002, ISSN: 09252312. DOI: [10.1016/S0925-2312\(01\)00644-0](https://doi.org/10.1016/S0925-2312(01)00644-0).
- [142] S. Dong and T. Luo, “Bearing degradation process prediction based on the PCA and optimized LS-SVM model,” *Measurement*, vol. 46, no. 9, pp. 3143–3152, 2013, ISSN: 02632241. DOI: [10.1016/j.measurement.2013.06.038](https://doi.org/10.1016/j.measurement.2013.06.038).
- [143] M. Kumar, R. B. Pachori, and U. R. Acharya, “Characterization of coronary artery disease using flexible analytic wavelet transform applied on ECG signals,” *Biomedical Signal Processing and Control*, vol. 31, pp. 301–308, 2017, ISSN: 17468094. DOI: [10.1016/j.bspc.2016.08.018](https://doi.org/10.1016/j.bspc.2016.08.018).
- [144] J. Chorowski, J. Wang, and J. M. Zurada, “Review and performance comparison of svm-and elm-based classifiers,” *Neurocomputing*, vol. 128, pp. 507–516, 2014.
- [145] Shoeb, Ali and Edwards, Herman and Connolly, Jack and Bourgeois, Blaise and Ted Treves, S and Guttag, John, “Patient-specific seizure onset detection,” *Epilepsy & Behavior*, vol. 5, no. 4, 483–498, 2004.
- [146] S.Pincus, “Approximate entropy as ameasure of system complexity,” *Proceedings of the National Academy of Sciences*, vol. 88, no. 6, pp. 2297–2301, 1991. DOI: [10.1073/pnas.88.6.2297](https://doi.org/10.1073/pnas.88.6.2297).
- [147] S. M. Pincus and R. R. Viscarello, “Approximate entropy: a regularity measure for fetal heart rate analysis.,” *Obstetrics and gynecology*, vol. 79, pp. 249–255, 1992, ISSN: 0029-7844.
- [148] J. S. Richman and J. R. Moorman, “Physiological time-series analysis using approximate entropy and sample entropy.,” *American journal of physiology. Heart and circulatory physiology*, vol. 278, H2039–H2049, 2000, ISSN: 0363-6135.
- [149] Y. Song, J. Crowcroft, and J. Zhang, “Automatic epileptic seizure detection in EEGs based on optimized sample entropy and extreme learning machine,” *Journal of Neuroscience Methods*, vol. 210, pp. 132–146, 2012, ISSN: 01650270. DOI: [10.1016/j.jneumeth.2012.07.003](https://doi.org/10.1016/j.jneumeth.2012.07.003).

- [150] J. Yentes, N. Hunt, K. Schmid, J. Kaipust, D. McGrath, and N. Stergiou, “The appropriate use of approximate entropy and sample entropy with short data sets,” English, *Annals of Biomedical Engineering*, vol. 41, no. 2, pp. 349–365, 2013, ISSN: 0090-6964. DOI: 10.1007/s10439-012-0668-3.
- [151] D. E. Lake, J. S. Richman, M. P. Griffin, and J. R. Moorman, “Sample entropy analysis of neonatal heart rate variability,” *American journal of physiology. Regulatory, integrative and comparative physiology*, vol. 283, R789–R797, 2002, ISSN: 0363-6119. DOI: 10.1152/ajpregu.00069.2002.
- [152] N. Burioka, M. Miyata, G. Cornélissen, F. Halberg, T. Takeshima, D. T. Kaplan, H. Suyama, M. Endo, Y. Maegaki, T. Nomura, Y. Tomita, K. Nakashima, and E. Shimizu, “Approximate entropy in the electroencephalogram during wake and sleep.,” *Clinical EEG and neuroscience : official journal of the EEG and Clinical Neuroscience Society (ENCS)*, vol. 36, pp. 21–24, 2005, ISSN: 1550-0594. DOI: 10.1177/155005940503600106.
- [153] L. Guo, D. Rivero, and A. Pazos, “Epileptic seizure detection using multiwavelet transform based approximate entropy and artificial neural networks,” *Journal of Neuroscience Methods*, vol. 193, no. 1, pp. 156–163, 2010, ISSN: 01650270. DOI: 10.1016/j.jneumeth.2010.08.030.
- [154] D Abásolo, J Escudero, R Hornero, C Gómez, and P Espino, “Approximate entropy and auto mutual information analysis of the electroencephalogram in Alzheimer’s disease patients,” *Medical & Biological Engineering & Computing*, vol. 46, no. 10, pp. 1019–1028, 2008, ISSN: 1741-0444. DOI: 10.1007/s11517-008-0392-1. [Online]. Available: <http://dx.doi.org/10.1007/s11517-008-0392-1>.
- [155] C. E. Shannon, “A mathematical theory of communication,” *The Bell System Technical Journal*, vol. 27, no. July 1928, pp. 379–423, 1948, ISSN: 07246811. DOI: 10.1145/584091.584093. arXiv: 9411012 [chao-dyn].
- [156] a Rényi, “On measures of entropy and information,” *Entropy*, vol. 547, no. c, pp. 547–561, 1961, ISSN: 15205207. DOI: 10.1021/jp106846b. arXiv: 1101.3070.
- [157] S Robert, “The tsallis entropy of natural information [j],” *Physica A: Statistical Mechanics and Its Applications*, vol. 386, no. 1, pp. 101–118, 2007.
- [158] M. E. Torres and L. G. Gamero, “Relative complexity changes in time series using information measures,” *Physica A: Statistical Mechanics and its Applications*, vol. 286, no. 3-4, pp. 457–473, 2000, ISSN: 0378-4371. DOI: 10.1016/S0378-4371(00)00309-5.
- [159] M. O. Mendez, I. Chouvarda, A. Alba, A. M. Bianchi, A. Grassi, E. Arce-Santana, G. Milioli, M. G. Terzano, and L. Parrino, “Analysis of A-phase transitions during the cyclic alternating pattern under normal sleep.,” *Medical & biological engineering & computing*, 2015, ISSN: 1741-0444. DOI: 10.1007/s11517-015-1349-9.

- [160] C. Bandt and B. Pompe, “Permutation entropy: A natural complexity measure for time series,” *Physical review letters*, vol. 88, no. 17, p. 174102, 2002.
- [161] M. Zanin, L. Zunino, O. A. Rosso, and D. Papo, “Permutation entropy and its main biomedical and econophysics applications: A review,” *Entropy*, vol. 14, no. 8, pp. 1553–1577, 2012.
- [162] K. Kalpakis, S. Yang, P. F. Hu, C. F. Mackenzie, L. G. Stansbury, D. M. Stein, and T. M. Scalea, “Permutation entropy analysis of vital signs data for outcome prediction of patients with severe traumatic brain injury,” *Computers in biology and medicine*, vol. 56, pp. 167–174, 2015.
- [163] S. T. Grassberger Peter and C. Schaffrath, “Nonlinear time sequence analysis,” *International Journal of Bifurcation and Chaos*, vol. 01, no. 03, pp. 521–547, 1991. DOI: [10.1142/S0218127491000403](https://doi.org/10.1142/S0218127491000403).
- [164] K. De Brabanter, P. Karsmakers, F. Ojeda, C. Alzate, J. De Brabanter, K. Pelckmans, B. De Moor, J. Vandewalle, and J. A. Suykens, *LS-SVMLab Toolbox User’s Guide: version 1.7*. Katholieke Universiteit Leuven, 2010.
- [165] R. Kohavi, “A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection,” in *International Joint Conference on Artificial Intelligence*, vol. 14, 1995, pp. 1137–1143, ISBN: 1-55860-363-8. DOI: [10.1067/mod.2000.109031](https://doi.org/10.1067/mod.2000.109031).
- [166] V. K. Mishra, V. Bajaj, A. Kumar, and G. K. Singh, “Analysis of als and normal emg signals based on empirical mode decomposition,” *IET Science, Measurement & Technology*, vol. 10, no. 8, pp. 963–971, 2016.
- [167] I. Guyon and A. Elisseeff, “An Introduction to Variable and Feature Selection,” *Journal of Machine Learning Research (JMLR)*, vol. 3, no. 3, pp. 1157–1182, 2003, ISSN: 00032670. DOI: [10.1016/j.aca.2011.07.027](https://doi.org/10.1016/j.aca.2011.07.027). arXiv: [1111.6189v1](https://arxiv.org/abs/1111.6189v1).
- [168] J. Cano, F. Herrera, and M. Lozano, “Using evolutionary algorithms as instance selection for data reduction in kdd: An experimental study,” *Evolutionary Computation, IEEE Transactions on*, vol. 7, no. 6, pp. 561–575, 2003, ISSN: 1089-778X. DOI: [10.1109/TEVC.2003.819265](https://doi.org/10.1109/TEVC.2003.819265).
- [169] H. Peng, F. Long, and C. Ding, “Feature selection based on mutual information: Criteria of Max-Dependency, Max-Relevance, and Min-Redundancy,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1226–1238, 2005, ISSN: 01628828. DOI: [10.1109/TPAMI.2005.159](https://doi.org/10.1109/TPAMI.2005.159). arXiv: [f](https://arxiv.org/abs/f).
- [170] O. Ludwig, U. Nunes, R. Araújo, L. Schnitman, and H. A. Lepikson, “Applications of information theory, genetic algorithms, and neural models to predict oil flow,” *Communications in Nonlinear Science and Numerical Simulation*, vol. 14, no. 7, pp. 2870–2885, 2009, ISSN: 10075704. DOI: [10.1016/j.cnsns.2008.12.011](https://doi.org/10.1016/j.cnsns.2008.12.011).

- [171] R. Agrawal, H. Mannila, R. Srikant, H. Toivonen, A. I. Verkamo, *et al.*, “Fast discovery of association rules,” *Advances in knowledge discovery and data mining*, vol. 12, pp. 307–328, 1996.
- [172] R. Chitta, A. K. Jain, and R. Jin, “Sparse kernel clustering of massive high-dimensional data sets with large number of clusters,” in *Proceedings of the 8th Workshop on Ph.D. Workshop in Information and Knowledge Management*, ser. PIKM ’15, Melbourne, Australia: ACM, 2015, pp. 11–18, ISBN: 978-1-4503-3782-3. DOI: [10.1145/2809890.2809896](https://doi.org/10.1145/2809890.2809896).
- [173] F. Beil, M. Ester, and X. Xu, “Frequent term-based text clustering,” *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining - KDD ’02*, p. 436, 2002, ISSN: 158113567X. DOI: [10.1145/775047.775110](https://doi.org/10.1145/775047.775110).
- [174] M. Wu, L. Wang, M. Li, and H. Long, “An approach based on the SIR epidemic model and a genetic algorithm for optimizing product feature combinations in feature fatigue analysis,” *Journal of Intelligent Manufacturing*, vol. 26, no. 1, pp. 199–209, 2013. DOI: [10.1007/s10845-013-0773-7](https://doi.org/10.1007/s10845-013-0773-7).
- [175] N. Sharma and H. Om, “Extracting Significant Patterns for Oral Cancer Detection Using Apriori Algorithm,” *Intelligent Information Management*, vol. 06, no. 02, pp. 30–37, 2014, ISSN: 2160-5912. DOI: [10.4236/iim.2014.62005](https://doi.org/10.4236/iim.2014.62005).
- [176] I. Sadek, J. Biswas, V. F. S. Fook, and M. Mokhtari, “Automatic heart rate detection from FBG sensors using sensor fusion and enhanced empirical mode decomposition,” *2015 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, pp. 349–353, 2015. DOI: [10.1109/ISSPIT.2015.7394358](https://doi.org/10.1109/ISSPIT.2015.7394358).
- [177] H. G. Chen, Y. J. Yan, and J. S. Jiang, “Vibration-based damage detection in composite wingbox structures by HHT,” *Mechanical Systems and Signal Processing*, vol. 21, no. 1, pp. 307–321, 2007, ISSN: 08883270. DOI: [10.1016/j.ymssp.2006.03.013](https://doi.org/10.1016/j.ymssp.2006.03.013).
- [178] Y.-H. Wang, C.-H. Yeh, H.-W. V. Young, K. Hu, and M.-T. Lo, “On the computational complexity of the empirical mode decomposition algorithm,” *Physica A: Statistical Mechanics and its Applications*, vol. 400, pp. 159–167, 2014, ISSN: 03784371. DOI: [10.1016/j.physa.2014.01.020](https://doi.org/10.1016/j.physa.2014.01.020).
- [179] O. Chapelle, “Training a Support Vector Machine in the Primal,” *Neural Computation*, vol. 19, no. 5, pp. 1155–1178, 2007, ISSN: 0899-7667. DOI: [10.1162/neco.2007.19.5.1155](https://doi.org/10.1162/neco.2007.19.5.1155).
- [180] L. Bottou and J. L. Chih, “Support vector machine solvers,” *Large scale kernel machines*, pp. 1–27, 2007. DOI: [10.1.1.127.511](https://doi.org/10.1.1.127.511).
- [181] G. Manis, “Fast computation of approximate entropy,” *Computer methods and programs in biomedicine*, vol. 91, no. 1, pp. 48–54, 2008.

- [182] Y.-H. Pan, Y.-H. Wang, S.-F. Liang, and K.-T. Lee, “Fast computation of sample entropy and approximate entropy in biomedicine,” *Computer methods and programs in biomedicine*, vol. 104, no. 3, pp. 382–396, 2011.
- [183] J. C. Chen, P. C. Tung, S. F. Huang, S. W. Wu, S. L. Lin, and K. L. Tu, “Extraction and screening of knee joint vibroarthrographic signals using the empirical mode decomposition method,” *International Journal of Innovative Computing, Information and Control*, vol. 9, no. 6, pp. 2689–2700, 2013, ISSN: 13494198.
- [184] S. Nalband, R. Sreekrishna, and A. A. Prince, “Analysis of Knee Joint Vibration Signals Using Ensemble Empirical Mode Decomposition,” *Procedia Computer Science*, vol. 89, pp. 820–827, 2016, ISSN: 18770509. DOI: 10.1016/j.procs.2016.06.067.
- [185] J. Lee, D. D. McManus, P. Bourrell, L. Sörnmo, and K. H. Chon, “Atrial flutter and atrial tachycardia detection using Bayesian approach with high resolution time-frequency spectrum from ECG recordings,” *Biomedical Signal Processing and Control*, vol. 8, no. 6, pp. 992–999, 2013, ISSN: 17468094. DOI: 10.1016/j.bspc.2013.04.002.
- [186] A. Subasi and M. K. Kiymik, “Muscle fatigue detection in EMG using time-frequency methods, ICA and neural networks,” *Journal of Medical Systems*, vol. 34, no. 4, pp. 777–785, 2010, ISSN: 01485598. DOI: 10.1007/s10916-009-9292-7.
- [187] S. Dash, K. H. Shelley, D. G. Silverman, and K. H. Chon, “Estimation of respiratory rate from ECG, photoplethysmogram, and piezoelectric pulse transducer signals: A comparative study of time-frequency methods,” *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 5, pp. 1099–1107, 2010, ISSN: 00189294. DOI: 10.1109/TBME.2009.2038226.
- [188] N. E. Huang and S. S. P. Shen, *Hilbert-Huang Transform and Its Applications*. World Scientific Publishing Co. Pte. Ltd., 2014, vol. 5, p. 399, ISBN: 9789814508230. DOI: 10.1142/9789812703347.
- [189] S. Nalband, A. Sundar, A. A. Prince, and A. Agarwal, “Feature selection and classification methodology for the detection of knee-joint disorders,” *Computer Methods and Programs in Biomedicine*, vol. 127, pp. 94–104, 2016, ISSN: 01692607.
- [190] S. Ghofrani, a Ayatollahi, and M. B. Shamsollahi, “Wigner – Ville Distribution,” *Proceedings of the IEEE*, vol. 39, no. 2, pp. 2–2, 1993, ISSN: 00678856. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/20160789>.
- [191] R. Gonzalez, *Digital Image Processing*. Pearson Education, 2009, ISBN: 9788131726952.
- [192] V. Bajaj and R. B. Pachori, “Automatic classification of sleep stages based on the time-frequency image of EEG signals,” *Computer Methods and Programs in Biomedicine*, vol. 112, no. 3, pp. 320–328, 2013, ISSN: 01692607. DOI: 10.1016/j.cmpb.2013.07.006.

- [193] K. Fu, J. Qu, Y. Chai, and Y. Dong, “Classification of seizure based on the time-frequency image of EEG signals using HHT and SVM,” *Biomedical Signal Processing and Control*, vol. 13, pp. 15–22, 2014, ISSN: 17468108. DOI: [10.1016/j.bspc.2014.03.007](https://doi.org/10.1016/j.bspc.2014.03.007).

List Of Publications

Journal Publications

- 1) S. Nalband, CA.Valliappan, A.A. Prince, and A. Agrawal. “Time-frequency based feature extraction for the analysis of vibroarthographic signals”, (Accepted) *Computers and Electrical Engineering, Elsevier Publication*, 2018, doi:<https://doi.org/10.1016/j.compeleceng.2018.02.046> **SCI: Impact Factor: 1.57**
- 2) S. Nalband, A.A. Prince, and A. Agrawal. “An entropy based feature extraction and classification of Vibroarthographic signal using improved complete ensemble empirical mode decomposition with adaptive noise,” *IET Science, Measurement & Technology*, December 2017, DOI: 10.1049/iet-smt.2017.0284 **SCI: Impact Factor: 1.26**
- 3) S. Nalband, A. Sundar, A.A. Prince, and A. Agrawal. “Feature selection and classification methodology for the detection of knee-joint disorders,” Computer methods and programs in biomedicine, 127, 94-104, 2016. *Elsevier Publication*, **SCI: Impact Factor: 2.503**

Conference Publications

- 1) S.Nalband, CA.Valliappan, R.Gupta, A.A.Prince, and A.Agrawal, “Feature Extraction and Classification of Knee Joint Disorders Using Hilbert Huang Transform,” 14th Annual IEEE Conference of ECTI Society, (ECTI-CON 2017), June 27-30, 2017 Phuket, Thailand, IEEE. (**Scopus**)
- 2) S. Nalband, R.R.Sreekrishna R.R. and A.A.Prince “Analysis of Knee Joint Vibration Signals Using Ensemble Empirical Mode Decomposition,” Elsevier Procedia Computer Science, vol.89, pp. 820-827, 2016. (**Scopus**)
- 3) R.R.Sreekrishna, S. Nalband and A.A.Prince. “Real Time Cascaded Moving Average Filter for Detrending of Electroencephalogram Signals,” 5th IEEE International Conference on Communication and Signal Processing (ICCSP’16), p.p. 0745 - 0750, April 2016, Chennai, India, IEEE. (**Scopus**)

Biography

Biography Of Candidate

Saif Nalband is pursing PhD from Department of Electrical Electronics Engineering, BITS-Pilani, KK Birla Goa Campus, India. He completed his Bachelor of Engineering from University of Mumbai and masters from RV College, Bangalore. His areas of interest includes biomedical signal processing and machine learning.

Biography Of Supervisor

A. Amalin Prince completed his PhD from BITS-Pilani, Rajasthan in Jan 2011. He obtained his Master of Engineering in Applied Electronics from Sathyabama Institute of Science and Technology, Chennai, India, in 2005. He had completed his Bachelor of Engineering in the Electronics and Communication Engineering from The Indian Engineering College, Tamil Nadu, India, in 2002. He is currently working as Assistant Professor in the Department of Electrical Electronics Engineering , BITS-Pilani, KK Birla Goa Campus, India. His area of research includes FPGA based signal processing systems, VLSI signal processing, Reconfigurable computing system and Microsystems.

Biography Of Co-Supervisor

Anita Agrawal completed her PhD from Govt. College of Engineering, Amravati, India in 2009. She is currently working as Assistant Professor in the department of Electrical Electronics Engineering, BITS-Pilani, KK Birla Goa Campus, India. She is a recipient of the distinguished faculty award at the BITSAA first global meet in New Delhi in 2011. She has a vast teaching experience of 24 years. Her research interest lies in microelectronics, biomedical signal processing and medical image processing.