AUTOMATED CLASSIFICATION SYSTEM FOR NORMAL AND ALS EMG SIGNALS BASED ON ITERATIVE FILTERING

M.Tech. Thesis

By

RICHA SINGH

DISCIPLINE OF ELECTRICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY INDORE

JULY 2019

AUTOMATED CLASSIFICATION SYSTEM FOR NORMAL AND ALS EMG SIGNALS BASED ON ITERATIVE FILTERING

A THESIS

Submitted in partial fulfillment of the
requirements for the award of the degree
of
Master of Technology
in
Electrical Engineering
with specialization in
Communication and Signal Processing

by

RICHA SINGH

DISCIPLINE OF ELECTRICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY INDORE
JULY 2019

INDIAN INSTITUTE OF TECHNOLOGY INDORE

CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in the thesis entitled **"AUTOMATED CLASSIFICATION SYSTEM FOR NORMAL AND ALS EMG SIGNALS BASED ON ITERATIVE FILTERING"** in the partial fulfillment of the requirements for the award of the degree of **MASTER OF TECHNOLOGY** and submitted in the **DISCIPLINE OF ELECTRICAL ENGINEERING, Indian Institute of Technology Indore**, is an authentic record of my own work carried out during the time period from July 2018 to July 2019 under the supervision of Prof. Ram Bilas Pachori, Discipline of Electrical Engineering, IIT Indore.

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

Signature of the student with date

(RICHA SINGH)

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

Signature of the Supervisor of
M.Tech. thesis (with date)

(Prof. Ram Bilas Pachori)

RICHA SINGH has successfully given her M.Tech. Oral Examination held on **01-07-2019.**

Signature of Supervisor of M.Tech. thesis
Date:

Convener, DPGC
Date:

Signature of PSPC Member #1
Date:

Signature of PSPC Member #2
Date:

Acknowledgements

I take this opportunity to express a deep sense of gratitude towards my thesis supervisor **Prof. Ram Bilas Pachori** for providing excellent guidance, encouragement and inspiration throughout the project work. Without his invaluable guidance, this work would never have been a successful one. Furthermore, I would like to express my sincere gratitude to my PSPC members **Dr. Vivek Kanhangad** and **Dr. Neminath Hubballi** for their valuable suggestions.

I would like to thank all the faculty members and the staff at IIT Indore for their cooperation throughout my study and thesis work. I would also like to thank the Discipline of Electrical Engineering for providing all the facilities, resources, and research environment required for the completion of this work.

I would like to thank all the members of Signal Analysis Research Lab for their suggestions and helpful discussions during this work. In particular, I would like to thank Ms. Pratistha Chandra for her assistance at various levels of this work.

I am grateful to all my friends for their constant support and making my stay at the institute enjoyable.

I am especially grateful to my family members for their invaluable support and strong belief in me.

RICHA SINGH
(mt1702102007)
M.Tech. (Communication and Signal Processing)
Discipline of Electrical Engineering, IIT Indore

Dedicated to my Family

Abstract

Electromyogram (EMG) signals are proved very useful in identification of neuromuscular diseases. In proposed work, we came up with a new method for the analysis and classification of normal and abnormal EMG signals to identify neuromuscular diseases. First, we have obtained all motor unit action potentials (MUAPs) from EMG signals. Extracted MUAPs are then decomposed using iterative filtering decomposition method. Intrinsic mode functions (IMF) obtained from iterative filtering method, are considered for analysis and classification purpose. Features like Euclidean distance quadratic mutual information (ED-QMI), Cauchy-Schwartz quadratic mutual information (CS-QMI), cross information potential (CIP) and correntropy (COR) are computed for each level of IMFs separately. For the analysis of EMG signals, statistical analysis has been performed by the Kruskal-Wallis statistical test. From the results obtained after analysis process, we have observed that the iterative filtering decomposition method is better and provides statistical significant difference in normal and ALS EMG signals. For classification, the calculated features are given as an input to the three different classifiers: repeated incremental pruning to produce error reduction (JRip) rules classifier, reduces error pruning (REP) tree classifier and random forest classifier for the classification of normal and ALS EMG signals. The results obtained from classification process show that this classification method is very efficient and provided very accurate classification of normal and ALS EMG signals.

Contents

List of Figures

List of Tables

List of Abbreviations

Acc	Accuracy
ALS	Amyotrophic lateral sclerosis
ANN	Artificial neural network
ANOVA	Analysis of variance
CIP	Cross information potential
CNN	Convolutional neural network
COR	Correntropy
CS-QMI	Cauchy-Schwartz quadratic mutual information
CWT	Continuous wavelet transform
DWT	Discrete wavelet transform
ED-QMI	Euclidean distance quadratic mutual information
EEG	Electroencephalogram
EMD	Empirical mode decomposition
EMG	Electromyogram
EVDHM	Eigenvalue decomposition of Hankel matrix
FDA	Food and drug administration
FEBANN	Feed-forward error propagation artificial neural network
FN	False negative
FP	False positive
ID3	Iterative dichotomiser 3
IF	Iterative filtering
IMF	Intrinsic mode function
ITL	Information-theoretic learning
KC	Kolmogorov complexity
KNN	K-nearest neighborhood
MFCC	Mel-frequency cepstral coefficient
MI	Mutual information
MST	Minimum spanning tree
MU	Motor unit
MUAP	Motor unit action potential
MUP	Motor unit potential

MUPT	Motor unit potential train
NINDS	National institute of neurological disorder and stroke
NMD	Neuromuscular diseases
PCL	Potential class
PDF	Probability density function
QMI	Quadratic mutual information
QMI_{CS}	Cauchy-Schwartz quadratic mutual information
QMI_{ED}	Euclidean distance quadratic mutual information
REP	Reduces error pruning
RIPPER	Repeated incremental pruning to produce error reduction
rms	Root mean square
sEMG	Surface electromyogram
Sen	Sensitivity
SOD1	Superoxide dismutase 1
Spe	Specificity
SVM	Support vector machines
TN	True negative
TP	True positive
TQWT	Tunable-Q factor wavelet transform
var	Variance
WEKA	Waikato environment for knowledge analysis
WHO	World health organization
WNN	Wavelet neural network

Chapter 1

Introduction

Human beings are apparently the most complex life forms on this planet. Millions of tiny components, each with its very own identity, working simultaneously in an organized way. The human body is a solitary structure but it is composed of millions of smaller structures of cells, tissues, organs and various systems such as nervous, skeletal, muscular systems [1].

The muscle system comprises of different muscle types, each of which plays a vital part in the body's function. When nerves and muscles work together, the system they make is recognized as the neuromuscular system. The neuromuscular system contains all the muscles in the body and nerves work for them. When human body makes any movement, it needs communication in between the brain and the muscles [2]. The nervous system provides that link for communication between brain and muscles. Nerves have cells known as neurons which deliver messages from the brain to the muscles through the spinal cord. These neurons are known as motor neurons. The motor neurons are capable of discharging a chemical, which is grabbed by the muscle fiber. This informs the muscle fiber to contract, which makes the muscles move. Due to any reason, if neurons became unhealthy and dead then communication between brain and muscles breaks down [3, 4]. This condition is known as neuromuscular disorder or neuromuscular disease.

Neuromuscular diseases (NMD) are mostly described by progressive muscular impairment resulting in loss of ability of movement, being wheelchair-bound, swallowing difficulties [5, 6] weakness of respiratory muscles and lastly, death due

to respiratory failure. According to statistics, about 1 in 3500 of the population around the world is expected to have a NMD present in childhood or in later life [7]. Progressive NMD are characterized by muscle impairment which gets worse over months and resulting a significantly reduced life expectancy and eventually death in a couple of years. Amyotrophic lateral sclerosis (ALS) is an example of such type of disease [8]. According to statistics, ALS is the cause of five deaths in every 100,000 people of ages 20 or older [9]. Till now, there is no cure available for ALS. So it is desirable to detect the ALS at early stage to increase the life expectancy of patient suffering from ALS.

The objective of this thesis is to distinguish normal and ALS electromyogram signals using iterative filtering method. This chapter provides a brief insight to NMDs, ALS, EMG signals, and various methods of EMG signal classification.

1.1 Human body structure

For proper understanding of NMDs, we have to first know about human body structure and parts of body which are directly involve in NMDs. So in this section, we will discuss about neuromuscular system and skeletal muscles and how they affect the human body that causes NMDs.

1.1.1 The neuromuscular system

The nervous system and muscles working combinedly to permit movement, is known as the neuromuscular system [10]. The nervous system functions as both the controlling and communicating system of the body. This system includes numerous excitable linked cells called neurons that interact quickly and specifically with distinct components of the body through electrical signals. The nervous system has three primary components which are: the brain, the spinal cord and the peripheral nerves. Neurons are the fundamental structural unit of the nervous system which can have variable shape and size. Neurons are extremely specific cells that transmit messages from one portion of the body to another in the form of nerve impulses [11].

A muscle consists of groups of specific cells capable of contraction and relaxation. These specific cells have the primary purpose of generating forces, motions, and the capacity to interact such as speech, writing, or other forms of expression. Muscle tissues have elasticity and extensibility properties. Muscle tissues are capable of receiving and responding to stimulation and can truncate or contract [11]. There are three kinds of muscle tissue that can be recognized based on composition, contractile characteristics, and control systems: (i) skeletal muscle, (ii) soft muscle, (iii) cardiac muscle. The EMG is often used for skeletal muscle study.

1.1.2 Skeletal muscles

Skeletal muscle is such a voluntary muscle which implies that its functions can be actively controlled. It is linked to the bone and builds a separate organ of muscle tissue, blood vessels, tendons and nerves that surrounds our bones and makes it possible to move [12]. A skeletal muscle corresponds to the combination of various bundles of cells combined to each other called muscle fibers. Muscle fibers are cylindrical and generally have more than one nucleus [13]. Skeletal muscles function in a multitude of ways. Skeletal muscle's primary role is to transform chemical energy into mechanical energy to generate strength and power, keep posture, and produce motion through contraction and relaxation that affects the body's operations [14].

Skeletal contractions of the muscle pull tendons that are connected to the bones. If muscle contraction creates shortening of the muscle, then bone moves and thus the body part attached to that bone also moves [15]. Skeletal muscle contraction is initiated by impulses in the neurons to the muscle and is generally controlled voluntarily. Skeletal muscle fibers are well supplied with neurons for contraction. This particular sort of neuron is called a "motor neuron" and it approaches muscle tissue but is not linked to it in fact. Usually one motor neuron is able to stimulate many muscle fibers [11]. The electrical activities of muscle contraction and relaxation can be evaluated by EMG.

1.2 Neuromuscular diseases

NMD is a very wide name that includes many diseases and disorders that impair the functioning of the muscles or their immediate control of the nervous system, either directly or indirectly [16, 17]. Problems with central nervous control can trigger muscle stiffness or some degree of paralysis, depending on the nature and place of the trouble. Parkinson's disorder, various sclerosis, Huntington's disorder, and Creutzfeldt-Jakob disorder are some examples of central disorders. Spinal muscle atrophies are lower motor neuron diseases, while ALS is a combined disease of upper and lower motor neuron.

There may be various explanations for neuromuscular illnesses such as: autoimmune disorders, genetic disorders, certain types of collagen disease, exposure to environmental chemicals, poisoning including heavy metal poisoning. A mutation in your genes can also cause NMDs. In most of the cases, cause of NMD is unknown. There is no cure for many NMDs. However, medications can help to improve symptoms, boost mobility and can extend the life of patient suffering from NMD [18, 19].

1.2.1 Amyotrophic lateral sclerosis

ALS is a motor neuron disease that selectively impacts motor neurons. Motor neurons are regarded as the cells that regulate the body's voluntary muscles. Voluntary muscles make motions such as speaking, chewing, and walking. ALS is described by rigid muscles, twitching muscles, and gradually deteriorating weakness due to reduction of muscles in size. This makes it difficult for the patient to talk, swallow, breathe and the patient eventually dies [20].

ALS affects the upper motor neurons as well as the lower motor neurons. Both types of neurons either degenerate or die, and no longer send messages to muscle. Muscles gradually weaken, start twitching, and waste away (atrophy) as they are unable to operate. The brain ultimately loses its capacity to initiate and regulate voluntary motions [21]. ALS is a progressive disease, meaning that with time the symptoms get worse. Most individuals with ALS die from respiratory failure,

generally within 3 to 5 years of initial symptoms. About 10 % of individuals with ALS, however, can survive for 10 years or more [22, 23].

ALS is a worldwide prevalent NMD. ALS can affect people of all races and ethnic backgrounds. Although the disease may strike at any age, symptoms develop most frequently between the ages of 55 and 75 years. Men are slightly more probable to develop ALS than females. A majority of ALS instances with no obviously related risk variables appear to happen at random. The cause is not known in 90 to 95 % of instances, but it is thought that genetic and environmental factors are involved. [24]. Approximately 5-10 % of all instances of ALS are familial, which means that an person acquires the illness from their parents. Generally speaking, the family form of ALS needs only one parent to transmit the disease-responsible gene. Mutations in more than a dozen genes have been discovered to cause familial ALS. The National Institute of Neurological Disorders and Stroke (NINDS) found that some instances of familial ALS were correlated with mutations in the SOD1 (superoxide dismutase 1) gene. Researchers proposed that exposure to toxins during warfare or severe physical practice may be grounds for the danger of developing ALS [20, 24].

No medical test can provide a definitive ALS diagnosis. However, the existence of symptoms of upper and lower motor neurons firmly indicates the disease's existence. There is currently no cure for ALS and there is no efficient therapy for stopping or reversing disease progression. However, some medicines and treatments are accessible that can help manage symptoms, prevent unnecessary complications and make it easy to live with the disease [20]. Riluzole and edaravone have been approved by the U.S. Food and Drug Administration (FDA) to treat ALS. Riluzole is believed to decrease motor neuron harm by reducing glutamate concentrations [25]. In addition to the medication, some therapies such as physical therapy, speech therapy, nutritional support and breathing support are helpful in increasing the lifespan of ALS patients.

1.3 Electromyogram signal

The EMG signal is a biomedical signal that measures the electrical potential produced in the muscles in the form of voltage and current during its contraction and relaxation [26]. The nervous system always controls muscle operations such as contraction and relaxation. Therefore, the EMG signal is a complex signal which is controlled by nervous system and depends on the anatomical and physiological characteristics of the muscle [11].

An EMG detects the electrical potential produced by muscle cells when the cells are electrically or neurologically activated. EMG signals can be evaluated to identify medical abnormalities, activation level or to evaluate biomechanics of human or animal movement [27, 28]. Fig. 1.1 displays a normal EMG signal plot.

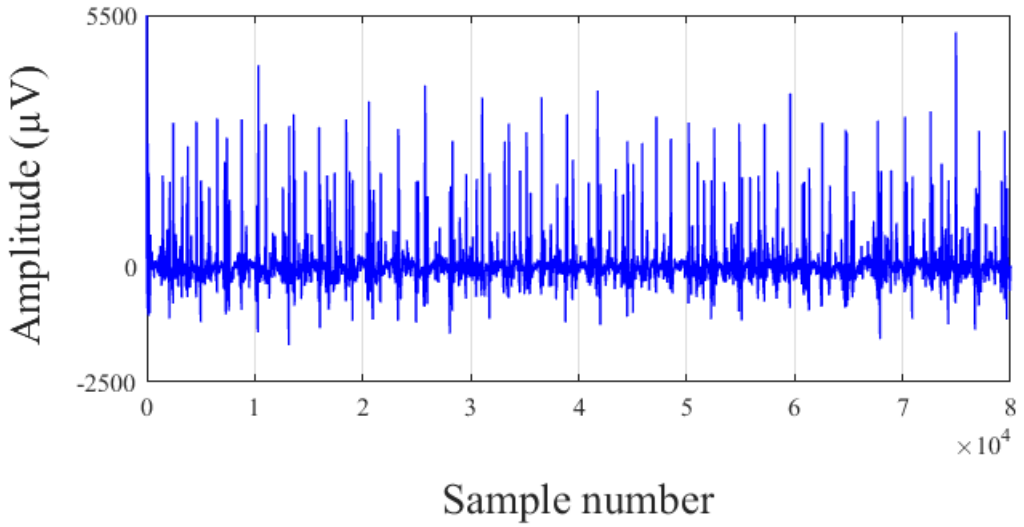

Figure 1.1: Plot of normal EMG signal.

1.3.1 The motor unit action potentials

Motor unit action potential (MUAP) plays a major part in the analysis of EMG signals. A motor unit is defined as one motor neuron and all the innervating muscle fibers to it. When a motor unit fires, the impulse is transferred from motor neuron to the muscle and is known as an action potential. The neuromuscular junction is recognized as the region where the nerve contacts the muscle. After the action potential is transferred throughout the neuromuscular junction, an action

potential is elicited in all the innervated muscle fibers of that specific motor unit. The summation of all this electrical activity is called as a motor unit action potential [29].

EMG signals are basically comprised of superimposed MUAPs from several motor units (MUs). For a proper analysis, the measured EMG signals must be decomposed into their constituent MUAPs [30]. MUAPs collected from various motor units tend to have distinct shapes, whereas MUAPs collected from the same motor unit by the same electrode are typically analogous. MUAP size and shape are predominantly based on the location of the electrode with respect to the fibers and may therefore appear to be distinct if the electrode shifts position. In addition, the amplitude of the action potential relies on multiple variables such as: muscle fiber diameter, distance between active muscle fiber, and electrode filtering characteristics in human muscle tissue [31]. The peak-to-peak amplitude of an MUAP in a normal muscle may range from a few microvolt with a typical value of 500 µV to 5mV potential [32]. MUAPs shapes and firing rates in EMG signals provide a significant source of information for neuromuscular disorder diagnosis.

1.3.2 Nature and electrical characteristics of EMG signals

The EMG signals are characterized as non-stationary and non-linear signals and influenced by the basic and utilitarian qualities of muscles [33]. The EMG signal appears random in nature. The electrical source is the potential of muscle membrane which is about $-90\,\text{mV}$ [34]. The EMG potential measured, ranges from $50\,\mu\text{V}$ to $30\,\text{mV}$ which also depends on the muscle being observed. Typically, the repetition rate of firing of motor unit of a muscle is about $7\,\text{Hz} - 20\,\text{Hz}$ and depends on the factors like: size of the muscle, previous axonal damage and other factors. The damage to motor units can be estimated between the range of $450\,\text{mV}$ to $780\,\text{mV}$ [35]. Fig. 1.2 displays characteristics of EMG signal.

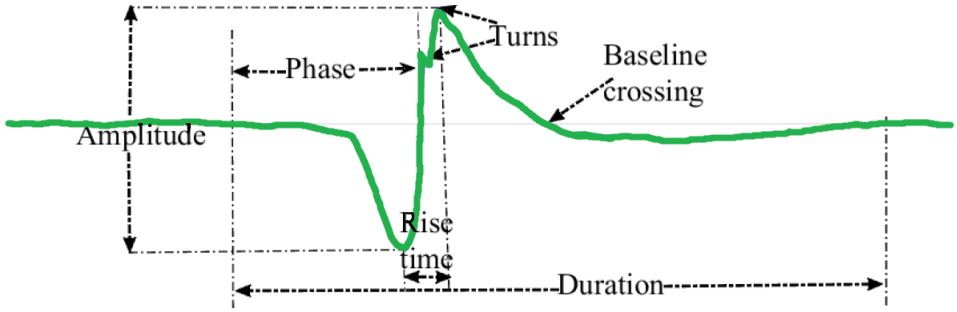

Figure 1.2: Characteristics of EMG signal [36].

According to De Luca [37], the energy distribution of EMG signal is generally within the range of 0 Hz to 400 Hz in frequency domain. The most of dominant components lie within the range of 50 Hz − 150 Hz. Outside the 0 Hz − 500 Hz frequency range, EMG signals are unusable as the energy is less than electrical noise level [38].

1.3.3 Recording of EMG signal

To obtain accurate EMG signals, it is very important to record the signals correctly. Skin preparation is a very significant phase before the needle electrode is inserted. Typically, skin preparation includes just cleaning the skin with an alcohol pad. The real positioning of the needle electrode can be challenging and depends on a number of variables such as: particular muscle choice and muscle size. The EMG signal is weaker when an individual has more body fat. The ideal location for placing the EMG sensors is at the belly of the muscle i.e. longitudinal midline [39].

There are two approaches to record EMG signal: (a) non-invasive electrode method (b) invasive electrode method. When EMG signals are obtained from electrodes directly positioned on the skin, this is known as non-invasive electrode method and EMG signal obtained from this method is known as surface EMG signal (sEMG). Either a pair of electrodes or an array of multiple electrodes can be used to record surface EMG. Minimum two electrodes must needed as EMG recording gives the potential difference between two distinct electrodes.

Sometimes EMG signals are obtained through wire or needle electrodes

positioned directly into the muscle, this method of recording EMG signal is known as invasive electrode method and EMG signal obtained from this method is known as intramuscular EMG signal. To conduct intramuscular EMG recording, a range of distinct kinds of electrodes can be used. The easiest technique is a monopolar needle electrode. This may be either a fine wire inserted into a muscle with a reference surface electrode or two fine wires inserted into muscle referenced to each other [11, 40].

1.3.4 Processing of EMG signal

The signal recorded for single contraction using single electrode can mathematically be represented as follows [41]:

$$E_t = \sum_{k=1}^{N} M_k(t) + n(t) \tag{1.1}$$

where M_k describes the voltage contributing function measured over the time of contraction, of a motor unit potential k from the total N MUPs. M termed as motor unit potential train (MUPT) which is the summation of all MUPs and the $n(t)$ shows the presence of noise. Because the raw EMG signal carries noise, the raw EMG signal must be processed. So raw EMG provides valuable information but it is in useless form, the information is only useful if it can be quantified. In order to obtain the correct and meaningful EMG signal, different signal processing techniques are applied to raw EMG signal.

The raw EMG signal recorded with help of electrodes is first picked up and amplified. As a first phase amplifier, a differential amplifier is typically used. Additional amplification phases may follow if necessary. Filtering is performed after amplification to filtered out low-frequency or high-frequency noises or other available artifacts. Before the signal is displayed or stored, the analog to digital converter converts the signal into a digital form. The user is often interested in the signal amplitude. The signal is therefore often rectified and averaged in some format to show the amplitude of EMG signal [42].

1.4 Existing methodologies

To diagnose the neuromuscular diseases various methodologies have been proposed by researchers which are based on analysis of EMG signal such as artificial neural network based analysis, machine learning and pattern classification algorithms [43], numerous transforms, various decomposition techniques, feature extraction through different techniques, different type of classifiers have been used to analyze the EMG signal in the past.

Various concepts for EMG signal classification such as Bayesian techniques [44, 45, 46, 47, 48], neural network modeling [49, 50], support vector machines (SVM) [51], and multilayer perceptron (MLP) [52] have been used. To diagnose neuromuscular diseases, a technique based on continuous wavelet transform (CWT) has been suggested in which an automatic system for classification of EMG signals [53]. Segmentation is performed by CWT and for the classification multi-channel artificial neural network (ANN) has been used. This method is effective to reduce the effect of noise but failed to classify overlapping MUPs.

A technique based on wavelet neural networks (WNN) and feedforward error backpropagation artificial neural networks (FEBANN) classifiers has been developed and accuracies achieved by this technique are 98.7 % and 88 % respectively [54]. A method based on empirical mode of decomposition (EMD) has been employed for decomposing the EMG signal into its natural components and the amount of irregularity within each component is determined by Kolmogorov complexity (KC). After extraction of features, SVM classifier has been used to classify neuropathy, myopathy, and normal subjects [55, 56].

For the detection of ALS disease, various methods have been developed. A method has been proposed for calculation of the features to distinguish abnormal EMG signal in which neuromuscular disease classification has been done by calculating mel frequency cepstrum coefficient (MFCC) of MUAPs [57]. Tunable-Q wavelet transformation (TQWT) based features for classification of ALS and healthy EMG signals has been presented in [58]. A method for classification of amyotrophic lateral sclerosis disease based on convolutional neural network (CNN)

and reinforcement sample learning algorithm has been proposed in [59]. Another wavelet transform based technique for characterizing EMG signals in terms of singularity was provided in [60]. A method based on EMD for feature extraction was used on the EMG signal directly [59]. For EMG signal analysis, the wavelet packet transform was also applied [61]. By applying attributes directly to non-overlapping frames obtained from EMG signals, classification of NMDs can be conducted [62]. A frame can have multiple MUAPs, attributes are evaluated for further assessment on each MUAP [63].

Another technique for the detection of abnormal EMG signals based on TQWT was provided. In this method, MUAPs are decomposed into their components and then features are calculated straight from MUAPs and decomposed components [64]. In another technique, high energy discrete wavelet transform (DWT) coefficients with maximum value were extracted from the EMG data frame-by-frame and K-nearest neighbors (KNN) has been used as classifier [65]. Pattern classification technique with two-fold features extraction technique also has been used for EMG signal classification [66]. The outlier based method also has been presented [67]. EMG signal classification has also been done by fuzzy logic [68]. EMG signal analysis has also been done by eigenvalue decomposition of Hankel matrix (EVDHM) in [16].

1.5 Motivation

There are numerous NMDs that affect the spinal cord, muscles and nerve cells in the brain and hence affect the movement of human body. These disorders can result in spasticity or some degree of paralysis. World health organization (WHO) studies show that around 1 billion individuals worldwide suffer from NMDs. NMD like ALS is a progressive NMD which gets worse over time and results in a significantly reduced life expectancy and eventually death within a few years.

According to statistics, ALS is the cause of five deaths in every 100,000 people of ages 20 or older. There is no cure available for ALS till now. Early detection and diagnosis of these illnesses through clinical examination is therefore essential to understand the nature of the disease and possible therapy of these illnesses. EMG

signal is a very useful tool for ALS diagnosis. Visual identification of EMG signals is a difficult and time-consuming activity that motivates the designing of automatic identification methods for ALS and normal EMG signals analysis and classification.

There are different techniques previously have proposed for the analysis and classification of various NMDs. Although all the above described methods are good for diagnosis of NMDs but there is always a scope to develop a more accurate and reliable method which can be more effective to diagnose these diseases with less complexity. So keeping these things in mind, a method has been proposed in this thesis which requires less number of features to be extracted for the analysis and classification of EMG signals to distinguish between normal and ALS EMG signals. The proposed method can be very useful for assisting doctors in the diagnosis of NMDs.

1.6 Objective

The goal of this thesis is to utilize the iterative filtering method for the decomposition of EMG signals to distinguish between normal and abnormal EMG signals and to get high accuracy of classification by computing less number of parameters.

For EMG signal analysis and classification, many features have been used previously. And for classification, various classifiers have been used for a number of distinguishing parameters. In this modern era, there is a need to develop techniques which would be less complex and give results in a more accurate way. So this is the objective of the presented work in this thesis, that even by computing one or two parameters, we can easily distinguish among different classes of EMG signals and can classify EMG signals of different classes with high accuracy.

1.7 Thesis organization

The rest of the thesis is organized systematically and in easily understandable manner .

Chapter 2 presents the background of iterative filtering method and also provides the algorithm for the same.

Chapter 3 provides the detailed information about database which is used for our experiment and study on EMG signals. It also presents the proposed methodology for EMG signal analysis and classification in detail.

Chapter 4 presents the obtained results and graphs. Results are organized in the form of tables for easy observation.

Chapter 5 discusses about the conclusions drawn from the whole study. It also gives the insight to the future scope of this work.

1.8 Summary

This chapter gives a insight of skeletal muscles, neuromuscular system of human body and NMDs. In this chapter, we have discussed about EMG signals and MUAPs. The electric characteristics and nature of EMG signal and recording and processing of EMG signals are discussed briefly. The motivation for doing this thesis work and what are the objectives of this research work are also discussed in this chapter. At the end, thesis organisation is presented.

Chapter 2

Iterative filtering

This chapter presents a brief description about the Iterative filtering (IF) method and its significance in decomposition of non-linear and non-stationary signals. We also discussed why traditional empirical mode decomposition (EMD) method is replaced by IF method for decomposition.

2.1 Introduction to iterative filtering

Analysis of non-linear and non-stationary signals like EMG signal is a challenging task. The analysis methods need to be local, adaptive and stable in order to extract features from these signals. Several decomposition techniques for analyzing non-linear and non-stationary signals have been proposed over the past years. The decomposition of the signal can be achieved in two ways: through iteration or optimization [69]. Huang introduced the first such iterative algorithm, the EMD [70], in 1998. The purpose of this method is iteratively decomposing a signal into a finite sequence of simple components known as IMFs with separated frequency in the time domain. However, this method is unstable under perturbations. So one other method known as iterative filtering was proposed as an alternative algorithm for EMD [71].

IF is an iterative approach which decomposes a nonlinear and non-stationary signal into a finite number of simple oscillatory components [71]. The obtained components from the IF method are known as IMFs. An IMF is a function which

satisfies following two conditions [70]:

1. The number of extrema and the number of zero crossings must be equal or at most vary by one.

2. Considering an upper envelope that connects all the local maxima and a lower envelope that connects all the local minima of the function, their mean must be zero at any point.

The iterative structure used in EMD method is known as sifting process which is extremely adaptive to data as a minute change in data results in different EMD. In this iterative process, the mean operation of the upper envelope and the lower envelope gives the moving average which is given by cubic splines that connects local maxima and local signal minima respectively. Since cubic splines are utilized in the iteration, this method is not stable under perturbations. Since the EMD method inspires IF, algorithm of IF method utilizes a similar structure as EMD method, But IF uses moving average in the sifting process which is determined by the convolution of that signal with low pass filters. IF method remains stable under perturbations [69]. An additional advantage of IF over traditional EMD is that it can be extended easily to higher dimensions.

2.2 Algorithm for iterative filtering

For iterative process, there are two methods for calculation of the moving average. In first one, the mean function of the upper and lower envelope is taken, which are given by cubic splines that connects local maxima and local minima of the signal. This method is similar as EMD process and suffers from unstability under perturbations. Second method calculates the moving average determined by the convolution of signal with a low pass filter (like double average filter). In IF method, the second method is used to calculate the moving average.

The process of extracting IMFs from a signal using IF is explained as follows:

For a given signal $s(t)$ where $t \in \Re$, let an operator R to make $R(s)$ which determines the moving average filter of signal $s(t)$. If $h(\tau)$ denotes a double average filter then the moving average of signal $s(t)$ can be determined by [69] as:

$$R(s)(t) = \int_{-m}^{m} s(t+\tau)h(\tau)d\tau \tag{2.1}$$

where double averaging filter $h(\tau)$, can be given by the following equation:

$$h(\tau) = \frac{(m+1-\tau|)}{(m+1)^2}, \quad t \in [-m,\ m] \tag{2.2}$$

If $s_1 = s$ and consider a operator $O_1, n(s_n) = s_n R_1, n(s_n) = s_n + 1$ which catches the fluctuations of s_n, The 1^{st} IMF can be shown by $I_1 = \lim_{n\to\infty} O_1, n(s_n)$, where R_1, n relies upon the mask length m_n, which defines the filter length at step n. To obtain I_2 (the second IMF), apply the operators O to the remainder signal $s - I_1$. Similarly, we obtained the $q - th$ IMF as $I_q = \lim_{n\to\infty} O_{q,n}(r_n) = r_{n+1}$, where $r_1 = s - I_1 - ... - I_{q1}$. The IF method stops when $r = s - I_1 - ... - I_q$, $q \in \mathbb{N}$ turns into a trend signal i.e. the remaining signal r has at most one local maxima or minima. Thus the signal s(t) can be represented in terms of decomposition as follows:

$$s(t) = \sum_{j=1}^{q} I_j(t) + r(t) \tag{2.3}$$

where q represents the number of IMFs into which the signal is decomposed and $r(t)$ is the remainder trend signal.

IF algorithm contains two nested loops: an inner loop and an outer loop. Inner loop is to compute each single IMF and outer loop is to derive all the IMFs. Algorithm for IF method is shown below [69]:

Algorithm 1 IF

IMF $= \{\}$
while the number of extrema of $s \geq 2$ **do**
    $s_1 = s$
    while the stopping criterion is not satisfied **do**
        compute the filter length m_n for s_n
        $s_{n+1}(t) = s_n(t) - \int_{-m_n}^{m_n} s_n(t+\tau)h_n(\tau)d\tau$
        $n = n + 1$
    end while
    IMF $=$ **IMF** $\cup \{s_n\}$
    $s = s - s_n$
end while
IMF $=$ **IMF** $\cup$ **{s}**

The mask length m_n can be computed using the following equation [71] :

$$m_n = 2\lfloor \beta \frac{N}{q} \rfloor \tag{2.4}$$

where β is a parameter generally fixed around 1.6. N is the total number of sample points in signal $s_n(t)$ and q is number of its extreme points. $\lfloor . \rfloor$ rounds a positive number to the nearest integer closer to zero.

The mask length m_n can be updated at each step of the inner loop. But in the implemented algorithm, the mask length is computed for the first step and the same value is used for remaining steps. Using same mask length at all steps of inner loop ensures that the IMFs obtained from this method have a proper group of instantaneous frequencies. To get that, O and R should not be dependent on step number n. So, the 1^{st} IMF can be shown as $I_1 = \lim_{n\to\infty} O^n(s)$. Where $O(s) = sR(s)$ and $R(s)$ is given by $R(s)(t) = \int_{-m}^{m} s(t+\tau)u(\tau)d\tau$, with mask length m calculated initially for the internal loop. and $u(\tau)$ represents any appropriate filter function.

The executed algorithm has some termination criterion for internal loop, and we did not take $n = \infty$. Termination criteria can be defined as follows:

$$\alpha = \frac{||I_{1,n} - I_{1,n-1}||_2}{||I_{1,n-1}||_2} \tag{2.5}$$

To stop the algorithm, certain threshold value of α can be used as a stop criteria [70, 72] or the maximum number of iterations can be set in the inner loop. Convergence of the inner loop of algorithm is guaranteed for periodic signals [70] and was studied for l^∞ functions in [73].

2.3 Summary

In this chapter, we discussed about the IF method and its importance in decomposition of non-linear and non-stationary signals. We also discussed the advantages of IF method over tradition EMD method for decomposition and why we are using IF method for decomposition. The algorithm for IF method is also

explained in detail.

Chapter 3

Database and proposed methodology

This chapter contains the detailed information about the database taken for our study. Also, this chapter provides information about preprocessing required for further analysis of EMG signals briefly and also explains the proposed methodology for analysis and classification of normal and ALS EMG signals in details.

3.1 Database

The database of EMG signals is taken from EMGLAB which is available online publicly [74]. Miki Nikolic had recorded EMG signals for quantitative analysis of MUAPs under normal conditions. For signal recording, a concentric needle electrode with area of $0.07\,\mathrm{mm}^2$ is used and one surface ground electrode was placed on the limb. For monitoring the signal, a slight and constant contraction was created in the muscle. The electrode was put at one location for 11.2 seconds and then moved to another location.

The EMG signals obtained from the needle were amplified by the amplification factor of 4000. Since recorded EMG signals were analog in nature, so they had to be converted in digital form. To digitize the signals, a 16 bit analog to digital converter is used [74]. This data has the sampling rate of 23437.5 Hz. After this, the signals are filtered by a bandpass filter with lower and higher cut off frequencies of 2 Hz and 10 kHz respectively.

In order to study the properties of EMG signals and analyze the status of NMD,

the signals were recorded from five muscles: abductor pollicisbrevis, biceps brachii, vastus medialis, tibialis anterior, deltoideus, and tensor fascial latae . The EMG signals were recorded under following usual conditions:

- The signals were recorded at low voluntary (just above the threshold) and constant level of contraction.

- Visual and audio feedbacks were used to monitor the quality of the signal.

- The EMG signals were recorded at three insertion depths (deep, medium, and low) from five locations in the muscle [75].

The material consisted of three groups: a normal control group, a group of myopathic patients, and a group of ALS patients. For our study, we are using normal group data and ALS group data. The normal control group consists of 10 normal subjects aged between 21 to 37 years, out of which 4 were females and 6 were males. 6 out of 10 were in very good physical shape, and the rest, except one, were generally in good shape. There were no traces or history of NMDs in any subject belongs to the normal control group. There were 8 patients in the ALS group ; 4 females and 4 males aged between 35 to 67 years. In addition to ALS-compatible clinical and electrophysiological symptoms, 5 of them died within a few years from the beginning of the disease which support the diagnosis of ALS [75]. The brachial biceps and medial vastus muscles were used in this study because they were the most frequently investigated muscles in the two patient groups.

3.2 Proposed methodology

In this section, a detailed description of proposed methodology for analyzing and classifying NMDs has been discussed. Fig. 3.1 and Fig. 3.2 show the block diagrams of proposed methodology for analysis and classification of EMG signals respectively.

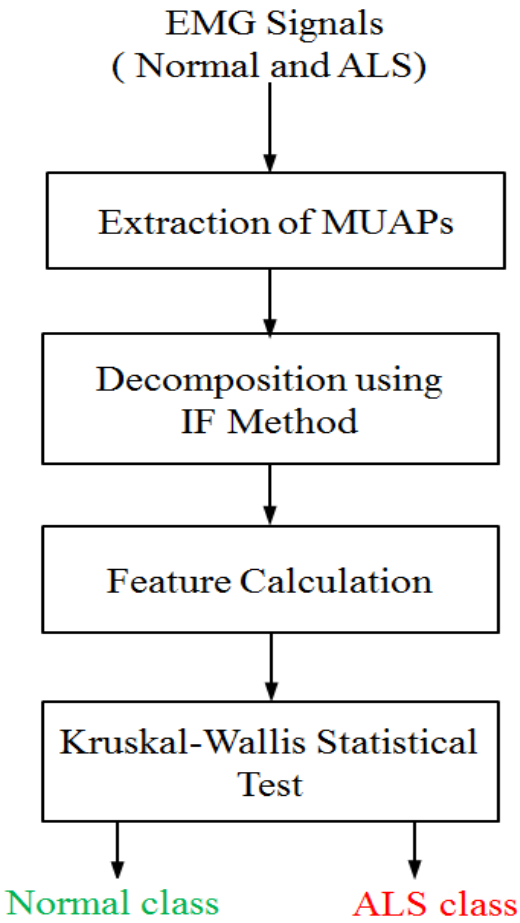

Figure 3.1: Block diagram for EMG sigal analysis

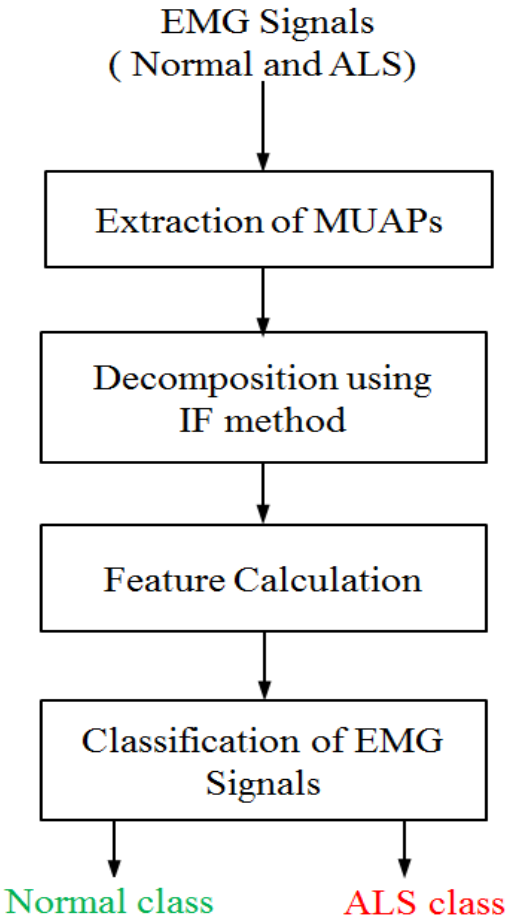

Figure 3.2: Block diagram for EMG signal classification

First step is to extract MUAPs from EMG signals. After obtaining the MUAPs, IF method was applied to decompose the MUAPs obtained from EMG signals. The IMFs are obtained from the decomposition of MUAPs. Euclidean distance quadratic mutual information (ED-QMI), Cauchy Schwarz quadratic mutual information (CS-QMI), cross information potential (CIP) and Correntropy (COR) features have been computed for all IMFs. Thereafter, statistical analysis and classification have been performed.

3.2.1 MUAP extraction

Extracting the MUAPs from the EMG signal involves detecting and identifying potential from all motor units. If more than one MUAP is superimposed, each of them should be identified separately so that we can obtain individual MUAPs

and their firing pattern [76]. MUAPs extraction has been done in three stages: (i) segmentation (ii) clustering and (iii) resolution [77]. All three stages are explained below:

- **Segmentation:** First, EMG signal is partitioned into several time intervals, then time intervals having MUAPs were searched. These time intervals are known as segments. There may be one MUAP or many superimposed MUAPs in a segment [76]. Segments containing superimposed MUAPs are known as compound segments. Time intervals which do not contain any MUAPs are called baseline [77]. For segmentation process, a window w_d of length 5.6 ms is applied along the entire EMG signal $z(n)$. Thereafter, the variance of signal inside that window is calculated. The variance can be calculated by following equation [78] :

$$\text{var(k)} = \frac{1}{w-1}\sum_{j=-q}^{q} z^2(k+j) - \left(\frac{1}{w-1}\sum_{j=-q}^{q} z(k+j)\right)^2 \tag{3.1}$$

  Where $var(k)$ denotes the variance of signal $z(n)$ at the k^{th} sample and calculated for the sample range from $-q$ to q. And w represents the segment length [76].

  A new segment is detected only when the variance calculated inside the window is greater than a detection threshold. The detection threshold can be estimated from the amplitude density function of normalized variance signal. MUAPs of similar size are represented by the local maximum of density function [76, 78]. The detection threshold can therefore be described as the first local minimum to be searched from the origin, where the smallest MUAPs should be separated from the noise. The method goes on until the variance becomes less than a delimiting threshold. The following equation estimates the segment-delimiting threshold (T_L) [77]:

$$T_L = 0.15(T_D - M_b) + M_b \tag{3.2}$$

  where T_D represents the detection threshold, and M_b represents the mean of the baseline variance and can be estimated by pre-segmenting the EMG signal

with $T_L = T_D/3$ [77].

- **Clustering:** In clustering process, the similar looking segments are clustered into groups. Segments which have similar shapes are assumed to be MUAPs from the same motor unit [78]. There may be several segments in a group. A group which contains five or more segments is known as a potential class (PCL). A template is chosen from each PCL to represent that PCL, These selected templates represent the active MUAPs in the EMG signal [78].

  The algorithm used for clustering process is a modified nearest-neighbor clustering algorithm. Which is based on a minimum spanning tree (MST) and uses heuristically determined tree cutting thresholds to form clusters of similar-looking segments successively. A distance measure is used to determine whether two segments are similar in shape. The distance between two segments each having only one MUAP generated by the same motor unit, should be small. Whereas the distance between two segments containing MUAPs from different motor units or between one simple and one compound segments or between two compound segments, should be large. The distance between two segments $d(s_1, s_2)$ can be determined by following equation [76, 78]:

$$d(s_1, s_2) = \frac{\text{var}(e)}{\text{rms}(s_1) + \text{rms}(s_2)}. \tag{3.3}$$

  Where s_1 and s_2 denotes the two segments to be compared and e is the residue signal after subtracting the two segments s_1 and s_2. var and rms represent the variance and root mean square value respectively. Distance $d(s_1, s_2)$ uses the variance of the residual after alignment and subtraction of the two segments and normalized with the sum of the rms values for the segments. Before measuring distances, it is necessary to time align the segments. If the lengths of the two segments are unequal, they are made equal by padding zeros to the shortest segment [79].

- **Resolution:** This is the last stage of MUAP extraction process. False templates and compound segments are resolved in this stage [79]. Some segments do not belong to any PCL, these segments are considered to be

compound segments. Some of the PCLs may consist of compound segments that look very similar, these PCLs are known as false PCLs and their template is known as false template [77, 79]. The false template and compound segments are resolved with the help of the templates from the PCLs. The false PCLs are resolved first, after that the compound segments are resolved [77].

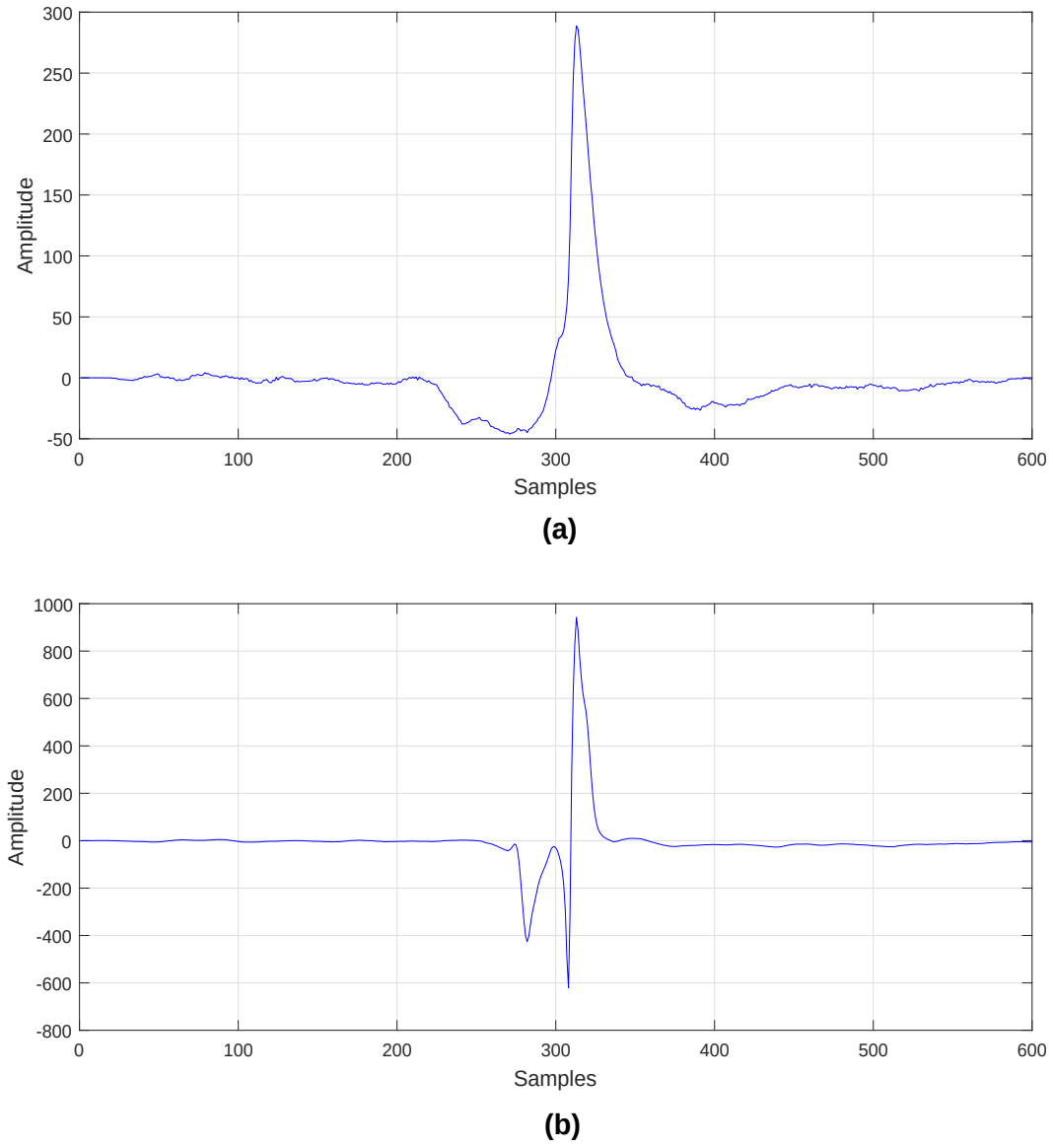

Figure 3.3: Plots of extracted MUAPs from (a) Normal EMG signal, (b) ALS EMG signal.

All MUAPs have been extracted from EMG signals for both normal and ALS classes with the help of EMGLAB software [80]. EMGLAB is a Matlab program designed to view and decompose EMG signals into MUAP trains. It provides a convenient graphical interface for displaying and editing results. The software is publicly available at http://www.emglab.net. Fig. 3.3 shows the plots of extracted MUAPs for both normal and ALS EMG signals.

3.2.2 Decomposition using iterative filtering method

Signal decomposition is the process of resolving a composite signal into its constituent parts in such a way that the original signal can be reconstructed from those parts. The aim of signal decomposition is separation of simple components from composite signals, so that it is easy to extract essential features for further analysis of EMG signal. Since EMG signals are complicated, it is not a good idea to extract features from signal directly. So decomposition of EMG signals is a very important step to get useful and differentiating properties for analysis and classification purpose.

Several Techniques have been presented for the decomposition of an EMG signal into its constituent components. In this thesis work, IF is the technique that have been used to decompose EMG signals. IF is an iterative approach which allows to decompose a given non-linear and non-stationary signal into a finite number of simple components. The obtained components from the IF method are known as IMFs. The IF process has already been discussed in detail in previous chapter.

By using IF method, MUAPs extracted from EMG signals have been decomposed into IMFs. We have used the MATLAB code of the IF for decomposition of EMG signals which is publicly available at http://www.mathworks.com/matlabcentral/fileexchange/53405-iterative-filters.
We have obtained first six IMFs for both normal and ALS class EMG signals. Extracted IMFs for normal and ALS EMG signals are shown in Fig. 3.4 and Fig. 3.5 respectively.

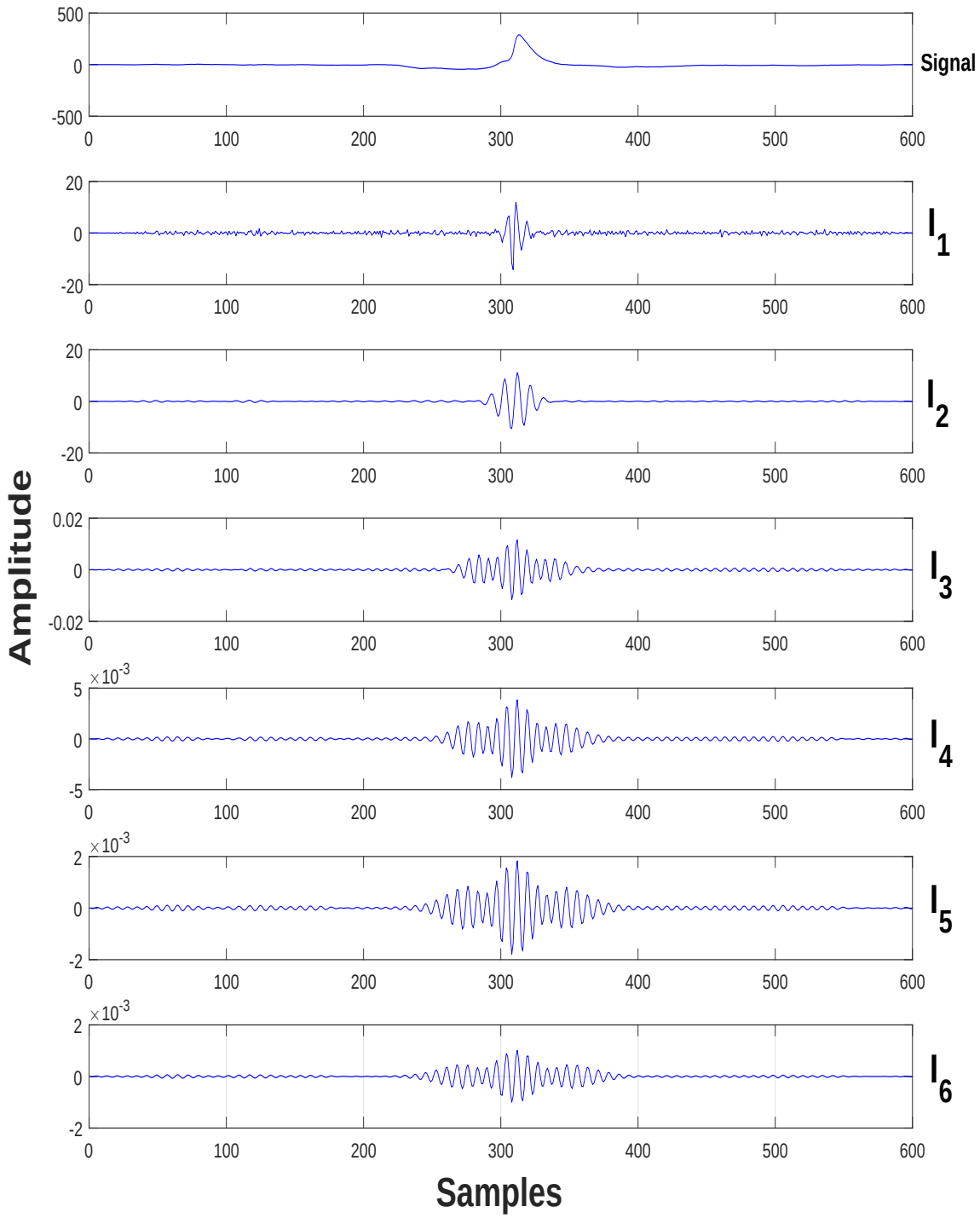

Figure 3.4: Plots of a normal EMG signal and it's first six obtained IMFs.

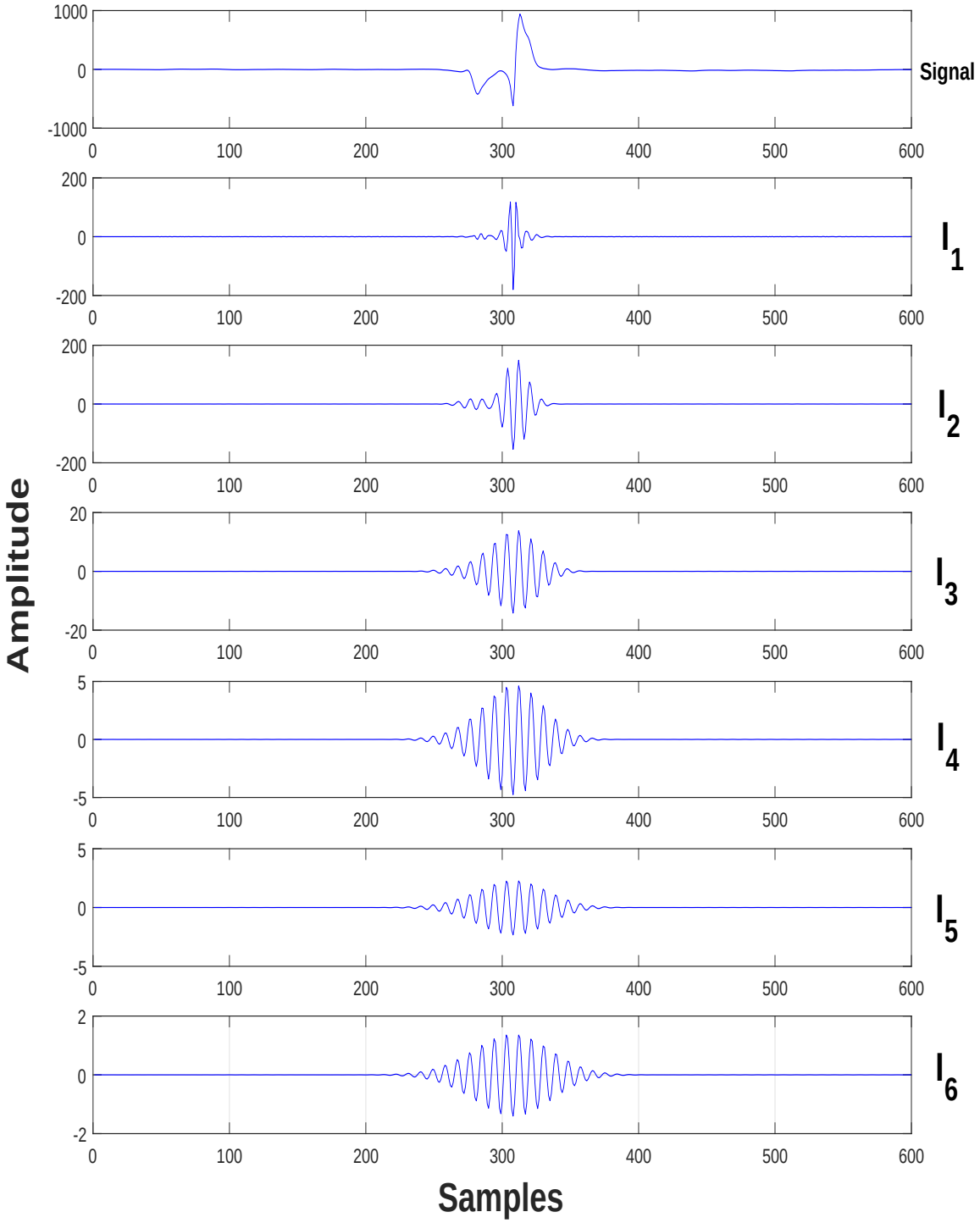

Figure 3.5: Plots of a ALS EMG signal and it's first six obtained IMFs.

3.2.3 Feature calculation

Since EMG signals are complex in nature, it is not a good approach to apply features on the EMG signal directly to distinguish between normal and abnormal EMG signals. Hence the features are applied on the decomposed components of EMG signal to uniquely distinguish normal and ALS data. In this work, we have studied four features ED-QMI, CS-QMI, CIP and COR for each level of IMFs separately, which are extracted from MUAP signal. In this section we have discussed about these features briefly.

- **Euclidean distance quadratic mutual information**

  The mutual information (MI) of two random variables quantifies the amount of information obtained about one random variable through observing the other random variable. When MI measured with only simple quadratic form of probability density functions (PDFs), then it is known as quadratic mutual information. So basically, QMI is a measure for statistical dependency between random variables [81, 82]. The Euclidean distance is straightforward distance measure for two PDFs, and can be defined as follow [83]:

$$D_{ED}(f,z) = \int (f(y) - z(y))^2 dy \tag{3.4}$$

  Where D_{ED} denotes the Euclidean distance between two PDFs $f(y)$ and $z(y)$.

  The squared distance between the joint pdf and the factorized marginal PDF is known as the ED-QMI. ED-QMI for two random variables Y_1 and Y_2 can be defined as [83]:

$$QMI_{ED}(Y_1,Y_2) = D_{ED}\Big[f_{Y_1Y_2}(y_1,y_2), f_{Y_1}(y_1)f_{Y_2}(y_2)\Big] \tag{3.5}$$

  where $f_{Y_1Y_2}(y_1,y_2)$ is denotes the joint PDF of Y_1 and Y_2. And $f_{Y_1}(y_1)$ and $f_{Y_2}(y_2)$ are the marginal PDFs of Y_1 and Y_2 respectively.

- **Cauchy-Schwartz quadratic mutual information**

CS-QMI is a variant of QMI which is based on the Cauchy-Schwartz distance between two PDFs. We can calculate the Cauchy-Schwartz distance between two PDFs by following equation [83]:

$$D_{CS}(f,z) = log\left[\frac{\int f^2(y)dy \int z^2(y)dy}{\int f(y)z(y)dy}\right] \tag{3.6}$$

Where D_{CS} denotes the Cauchy-Schwartz distance between two PDFs $f(y)$ and $z(y)$. Based on Cauchy-Schwartz distance, CS-QMI for two random variables Y_1 and Y_2 can be defined as follows [83]:

$$QMI_{CS}(Y_1,Y_2) = D_{CS}\Big[f_{Y_1Y_2}(y_1,y_2), f_{Y_1}(y_1)f_{Y_2}(y_2)\Big] \tag{3.7}$$

where $f_{Y_1Y_2}(y_1,y_2)$ denotes the joint PDF of Y_1 and Y_2. And $f_{Y_1}(y_1)$ and $f_{Y_2}(y_2)$ are the marginal PDFs of Y_1 and Y_2 respectively.

- **Cross information potential**

CIP characterizes the similarity between two PDFs [84]. The information among distinct random variables is characterized by CIP based on multidimensional PDFs. CIP can be used to quantify the divergence or the cross-covariance between the two random variables for two-dimensional PDFs [85]. The following equation can be used to calculate CIP [86]:

$$CIP(A,B) = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}k(a_i - b_j) \tag{3.8}$$

where A and B denotes two random variables data sets with independent and identically distributed (iid) sample sets $\{a_1.....a_N\}$ and $\{b_1.....b_N\}$ respectively, and N is the total number of samples. a_i is the i$^{\text{th}}$ sample of the data set A and b_j is the j$^{\text{th}}$ sample of the data set B. $k(a_i - b_j)$ represents the kernel function.

In this present work, we have used the Gaussian kernel with kernel size (σ) is equal to 1 [87]. We have used the ITL toolbox for the calculation of CIP. In this ITL toolbox, incomplete Cholesky decomposition is used to calculate

CIP. Previously, CIP feature has been used for characterization of coronary artery disease in [87].

- **Correntropy**

Two different words correlation and entropy are combined together to create a new term correntropy [88]. COR is the measure of probability of closeness between two random variables in a neighborhood of the joint space, in a specific window controlled by the kernel size [89]. COR has many properties that directly quantify the PDF of data. Suppose a and b are two random variables, then its parzen PDF can be given by following equation [88]:

$$P_\sigma = \frac{1}{N}\sum_{i=1}^{N} k_\sigma(a-b) \tag{3.9}$$

where $k_\sigma(a-a_i)$ represents the Gaussian kernel function. Gaussian kernel is widely used kernel function which can be given by following equation:

$$k_\sigma(a-b) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(a-b)^2}{2\sigma^2}\right]. \tag{3.10}$$

The COR determines the similarity between the signal and the delayed samples of the signal [89]. COR function for two random variable a and b can be defined as follow [88]:

$$COR(a,b) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(a-b)^2}{2\sigma^2}\right]. \tag{3.11}$$

Previously, COR feature been used for classification of focal electroencephalogram (EEG) signal [90, 91].

In this thesis work, we have calculated all four features of decomposed IMFs for both classes separately by taking first sample as a reference from both normal and ALS class signals. We have used the ITL toolbox for the calculation of all four features, with kernel size is equal to 1. ITL toolbox is available online at http://www.sohanseth.com/Home/codes.

3.2.4 Analysis using Kruskal-Wallis statistical test

Kruskal-Wallis statistical test is also known as Kruskal-Wallis H test or one-way analysis of variance (ANOVA) on ranks [92, 93]. It is a non-parametric method to test whether samples originate from the same distribution [94, 95, 96]. Non-parametric means that the test does not rely on data belonging to any particular parametric family of probability distributions. Kruskal-Wallis statistical test uses rank of data instead of data value [97, 98].

Kruskal-Wallis statistical test is considered as non- parametric equivalent of the ANOVA. It is an extension of the Mann-Whitney U test in which only two classes can be compared [99], but Kruskal-Wallis statistical test allows us to compare two or more than two independent samples of equal or different sample sizes. The Kruskal-Wallis test shows whether there is a significant difference between the groups, but it will not show which specific groups are statistically significantly different from each other [100]. The Kruskal-Wallis statistical test looks for the median of groups to determine whether they are different. Unlike the one-way ANOVA, Kruskal-Wallis test can also be used for ordinal data.

In Kruskal-Wallis statistical test, H statistic is used [101]. The hypotheses for test are H_0 and H_1, where H_0 represents that population medians are equal and H_1 is represents that population medians are not equal. Steps to perform Kruskal-Wallis statistical test are as follows [102]:

Step 1: First, data is sorted in ascending order for all the samples and then grouped together in one combined set.

Step 2: Assign ranks to all data points in combined set.

Step 3: Add up different ranks for each sample.

Step 4: H statistic is estimated by following equation:

$$H = \left[\frac{12}{S(S+1)} \sum_{q=1}^{n} \frac{{R_q}^2}{S_q} \right] - 3(S+1) \tag{3.12}$$

where, S=sum of sample sizes for all samples,

n = number of samples,

R_q = sum of ranks in the q_th sample,

S_q= size of the q$^{\text{th}}$ sample.

Step 5: The critical Chi-square (χ^2) value is calculated and then compared with H value obtained from step 4.

Step 6: The p-value is estimated by probability of H being less than or equal to Chi-square i.e. $P_r(\chi^2_{n-1} \geq H)$.

In this thesis work, The Kruskal-Wallis statistical test is carried out on the features calculated from both normal and ALS classes. The data from both classes is combined into a single series, and a rank is given to all data points in that combined set. We carried out Kruskal-Wallis test for obtaining the statistical significance($p < 0.05$) of features in diffrent oscillatory levels of the MUAPs obtained from EMG signals. The Kruskal-Wallis test has also been studied for EEG signal analysis in [103], and for coronary artery disease identification in [104].

3.2.5 Classification of normal and ALS EMG signals

Classification may refer to the process of predicting the class of given data points, it involves building a model of classes from a set of records which contain class labels and known as training data set. The resulting classification model will be capable of predicting labels for a new set of unlabeled records that have the same features as the training data. Each distinct label value is referred to as a class. An algorithm that implements the process of classification is known as a classifier [105].

In this proposed work, we have used three classifiers viz. JRip rules classifier, REP tree Classifier and random forest, to classify the normal and ALS class EMG signals. We have used 10-fold cross validation method to classify the data. Cross-validation is a technique for evaluating predictive models by dividing the original sample into a training set to train the model and a test set for evaluation. In 10-fold cross-validation, the original sample is randomly divided into 10 equal size sub-samples. From those 10 sub-samples, a single sub-sample is retained as the validation data for testing the model, and the remaining 9 sub-samples are used as

training data. The cross-validation process is then repeated 10 times (the folds). The 10 results from the folds can then be averaged to produce a single estimation. The advantage of this method is that all observations are used for both training and validation, and each observation is used for validation exactly once [106]. The classifiers used in this proposed work, are discussed below briefly:

- **JRip rules classifier:**

  JRip is one of the basic and popular classifier for classification among different classes. JRip applies a propositional rule learner which is known as "Repeated incremental pruning to produce error reduction"(RIPPER), as proposed in [107].

  In JRip algorithm, first all classes are examined in increasing size. After examination of classes, an initial set of rules for the class is generated using incremental reduced error JRip. After that all the examples of a particular decision in the training data treated as a class, and a set of rules that cover all the members of that class has obtained. Thereafter it proceeds to the next class and does the same. Repetition is going on until all classes have been covered [108, 109].

- **REP tree classifier:**

  REP tree classifier is a fast decision tree learning algorithm. It works on the principle of computing the information gain with entropy and minimizing the error arising from variance [110]. REP tree applies regression tree logic and generates multiple trees in altered iterations. Thereafter, it selects the best tree among all generated trees and considers it as representative.

  REP tree classifier builds a decision/regression tree using variance and information gain. The pruning of that decision tree is done by using reduced-error pruning with back fitting method. It sorts numeric attribute values once at the start of the model preparation. This algorithm also deals with the missing values by dividing the corresponding instances into pieces, in a similar manner to C4.5 algorithm [111].

- **Random forest classifier:** The random forest classifier is a set of decision tree classifiers, where each classifier is generated using a random vector which is sampled independently from the input vector [112]. To classify an input vector, each tree casts a unit vote for the most popular class, and then it takes the average of the votes received from different decision trees in order to decide the final class of the test objects [113] . It basically merges the multiple decision trees to create a wide diversity which leads to the better classification accuracy [114]. Random forest classifier works efficiently even on the large data sets.

  The random forest classifier uses randomly selected features or a combination of features at each node to grow a tree. For the best split, only selected features are searched at each node [112]. Each time using a combination of features, a tree is grown to the maximum depth on new training data. To generate a random forest classifier, two user defined parameters are required: (i) the number of features used at each node to generate a tree and (ii) the number of trees to be grown.

  These fully grown trees are not pruned and this is one of the major advantages of the random forest classifier over other decision tree methods. Studies suggest that as the number of trees increases, the generalization error always converges even without pruning the tree and over-fitting is not an issue due to the strong law of large numbers [115, 116].

In this proposed work, the popular machine learning toolbox Waikato Environment for Knowledge Analysis (WEKA) has been used for classification. It is a collection of machine learning algorithms, for data mining tasks. WEKA is open source software issued under the General Public License that can be downloaded from https://www.cs.waikato.ac.nz/ml/weka/.

3.3 Summary

In this chapter, we provided the detailed information about the database taken for our study. MUAPs are extracted from Normal and ALS EMG signals and the whole

MUAP extraction process is explained in detail. The proposed methodology for analysis and classification of normal and ALS EMG signals is explained in detail. Four features ED-QMI, CS-QMI, CIP and COR are calculated and discussed briefly. For analysis, we used Kruskal-Wallis statistical test for (p <0.05), which is also explained in this chapter. For classification, three different classifiers JRip rules classifier, REP tree classifier and random forest classifier are used and discussed briefly.

Chapter 4

Results and discussion

This chapter shows the results of EMG signal analysis using Kruskal-Wallis statistical test in the form of tables and box plots. It also presents the results of the classification of normal and ALS EMG signals by using three classifiers, in the tabular form.

4.1 Analysis results

For analysis of EMG signals, we have been carried out Kruskal-Wallis statistical test on four features ED-QMI, CS-QMI, CIP and COR separately, which are extracted from normal and ALS EMG signals. All four features mentioned above are computed for six IMFs i.e. $QMI_{ED}1$ to $QMI_{ED}6$, $QMI_{CS}1$ to $QMI_{CS}6$, CIP_1 to CIP_6 and COR_1 to COR_6 and considered for Kruskal-Wallis test for obtaining the statistical significance ($p < 0.05$). Table 4.1 presents the p-value obtained from ED-QMI feature for six IMFs. Table 4.2 presents the p-value obtained from CS-QMI feature for six IMFs. Table 4.3 presents the p-value obtained from CIP feature for six IMFs. Table 4.4 presents the p-value obtained from COR feature for six IMFs. For all the features, p-value is less than 0.05 and either zero or very close to zero. These results show that we can easily distinguish between normal and ALS EMG signals with the proposed methodology.

Table 4.1: p-value for QMI_{ED} feature of first six IMFs.

Features	$QMI_{ED}1$	$QMI_{ED}2$	$QMI_{ED}3$	$QMI_{ED}4$	$QMI_{ED}5$	$QMI_{ED}6$
p-value	2.2524×10^{-109}	6.7073×10^{-64}	0	0	5.2049×10^{-275}	1.9133×10^{-231}

Table 4.2: p-value for QMI_{CS} feature of first six IMFs.

Features	$QMI_{CS}1$	$QMI_{CS}2$	$QMI_{CS}3$	$QMI_{CS}4$	$QMI_{CS}5$	$QMI_{CS}6$
p-value	8.2216×10^{-107}	5.3181×10^{-75}	0	0	2.1800×10^{-276}	2.1512×10^{-217}

Table 4.3: p-value for CIP feature of first six IMFs.

Features	CIP_1	CIP_2	CIP_3	CIP_4	CIP_5	CIP_6
p-value	0.0407	1.5295×10^{-283}	0	1.3511×10^{-290}	1.1558×10^{-284}	0

Table 4.4: p-value for COR feature of first six IMFs.

Features	COR_1	COR_2	COR_3	COR_4	COR_5	COR_6
p-value	7.3062×10^{-25}	0	1.3189×10^{-318}	6.0976×10^{-277}	2.0751×10^{-280}	4.9407×10^{-324}

Box plot is a simple way of representing statistical data on a plot in which a rectangle is drawn to represent the second and third quartiles, usually with a vertical line inside to indicate the median value. The lower and upper quartiles are shown as horizontal lines either side of the rectangle.

The box plots obtained for ED-QMI feature for first six IMFs are shown in Figs. 4.1, 4.2, 4.3, 4.4, 4.5, 4.6. The box plots obtained for CS-QMI feature for first six IMFs are presented in Figs. 4.7, 4.8, 4.9, 4.10, 4.11, 4.12. The box plots obtained for CIP feature for first six IMFs are shown in Figs. 4.13, 4.14, 4.15, 4.16, 4.17, 4.18 and Figs. 4.19, 4.20, 4.21, 4.22, 4.23, 4.24 represent the box plots obtained for COR feature for first six IMFs.

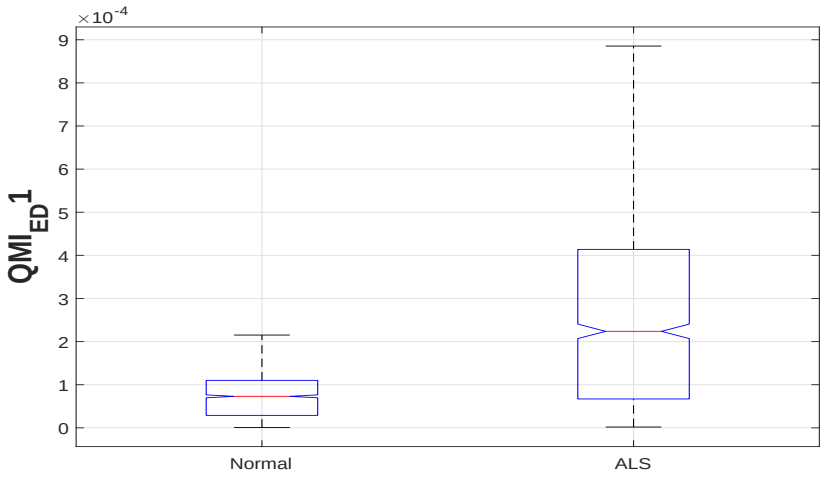

Figure 4.1: Box plot of QMI_{ED} feature of 1^{st} IMFs for normal and ALS EMG signals.

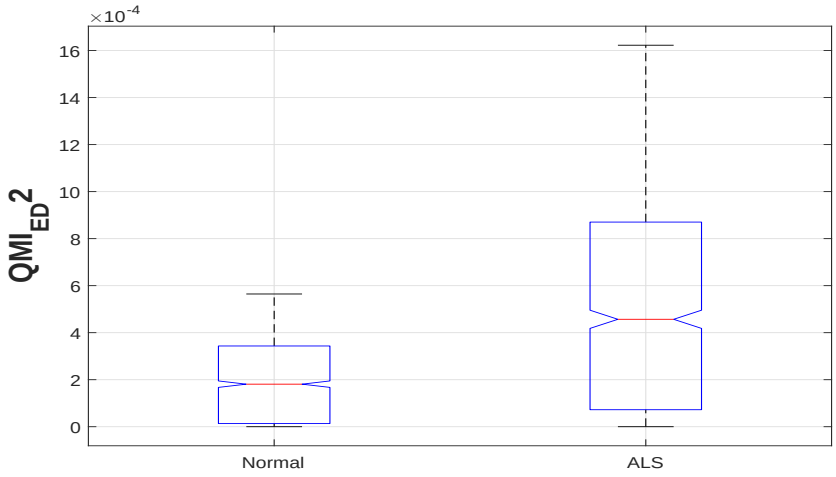

Figure 4.2: Box plot of $\mathrm{QMI_{ED}}$ feature of 2^{nd} IMFs for normal and ALS EMG signals.

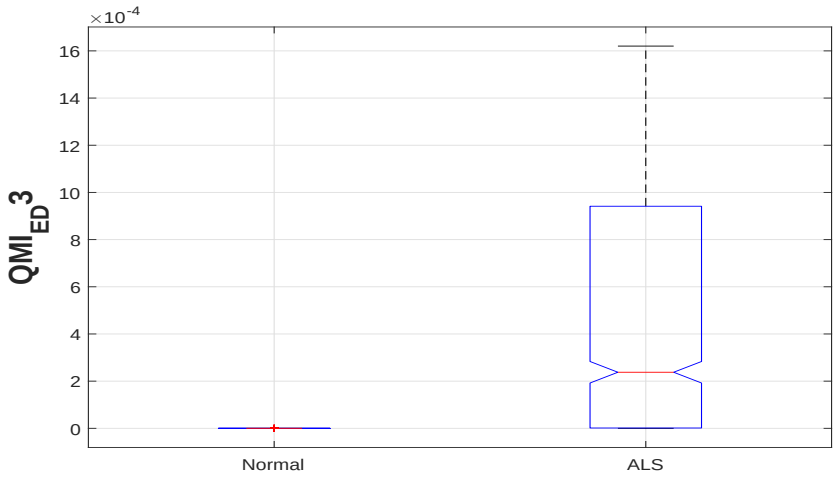

Figure 4.3: Box plot of $\mathrm{QMI_{ED}}$ feature of 3^{rd} IMFs for normal and ALS EMG signals.

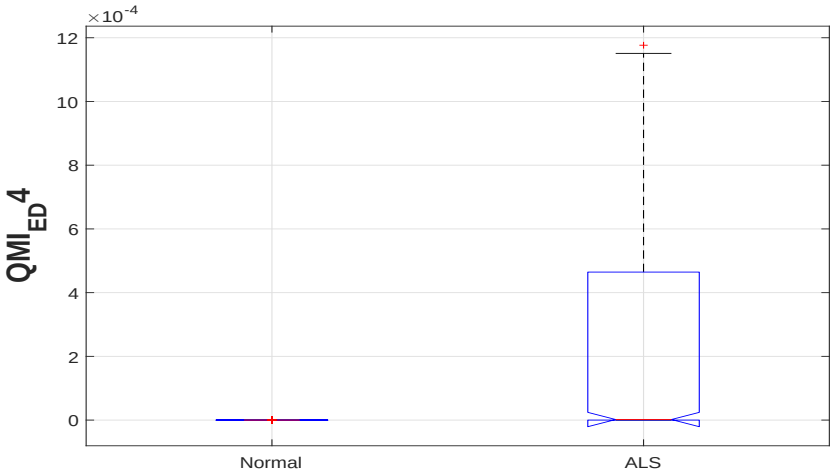

Figure 4.4: Box plot of QMI_{ED} feature of 4^{th} IMFs for normal and ALS EMG signals.

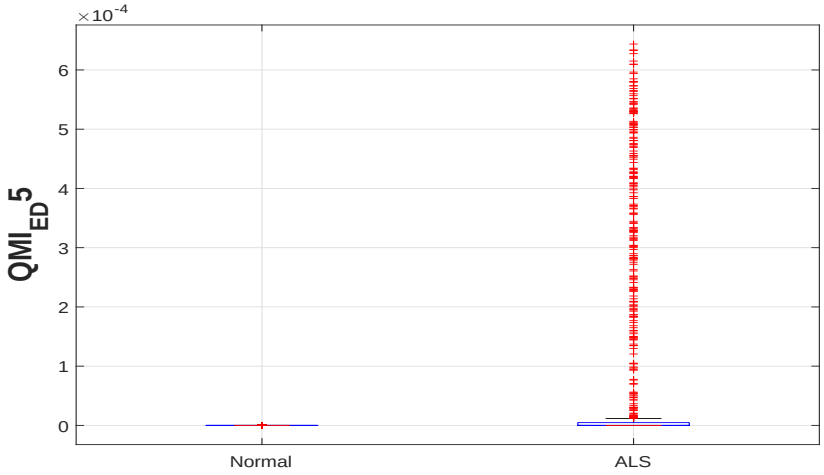

Figure 4.5: Box plot of QMI_{ED} feature of 5^{th} IMFs for normal and ALS EMG signals.

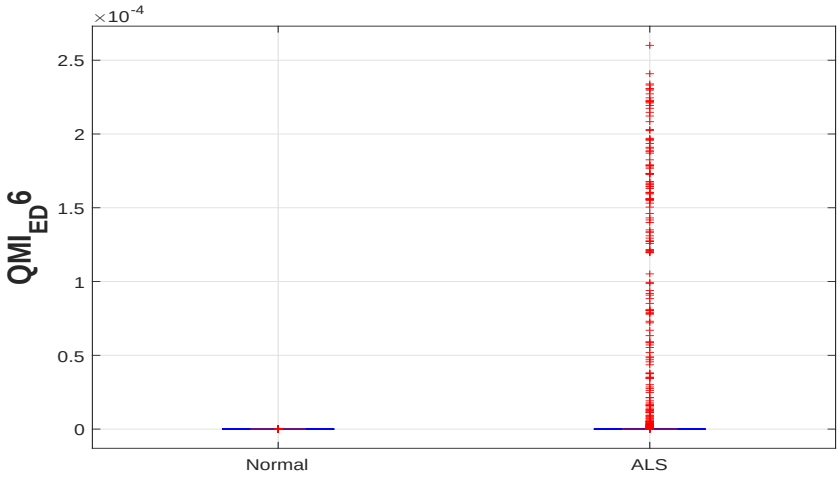

Figure 4.6: Box plot of $\mathrm{QMI}_{\mathrm{ED}}$ feature of 6^{th} IMFs for normal and ALS EMG signals.

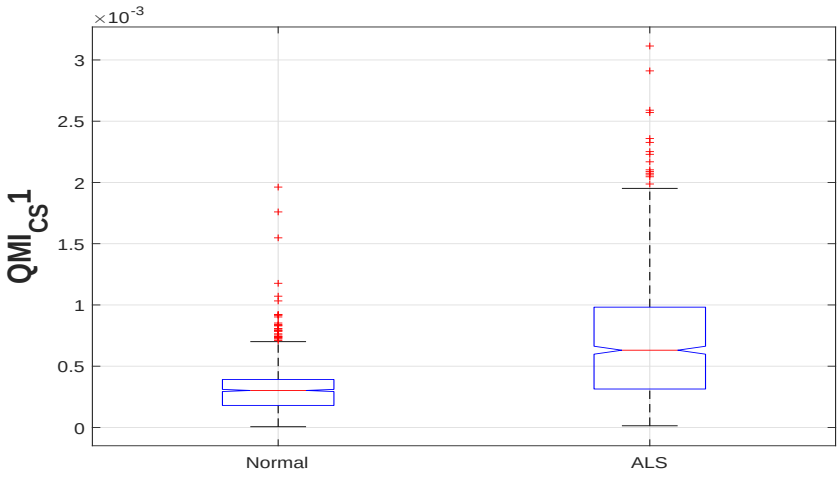

Figure 4.7: Box plot of $\mathrm{QMI}_{\mathrm{CS}}$ feature of 1^{st} IMFs for normal and ALS EMG signals.

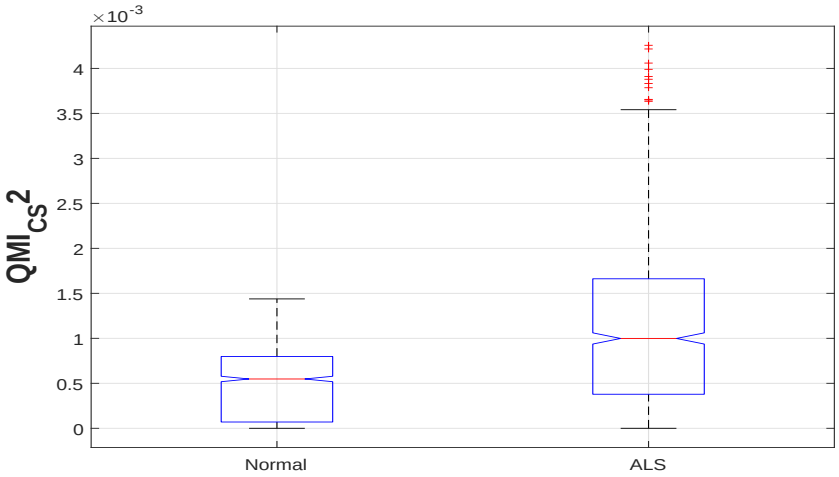

Figure 4.8: Box plot of QMI_{CS} feature of 2nd IMFs for normal and ALS EMG signals.

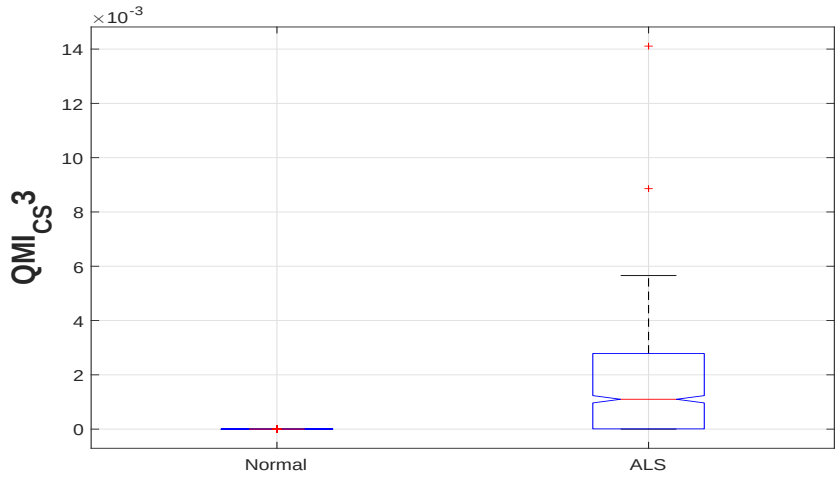

Figure 4.9: Box plot of QMI_{CS} feature of 3rd IMFs for normal and ALS EMG signals.

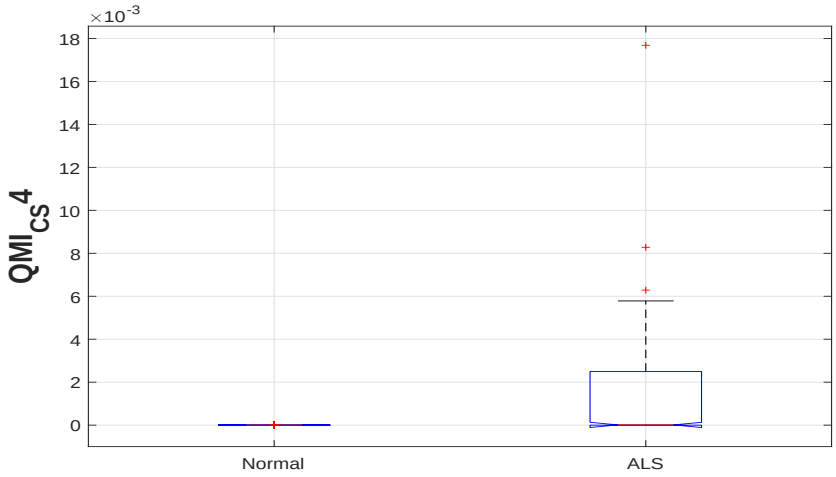

Figure 4.10: Box plot of $\mathrm{QMI_{CS}}$ feature of 4^{th} IMFs for normal and ALS EMG signals.

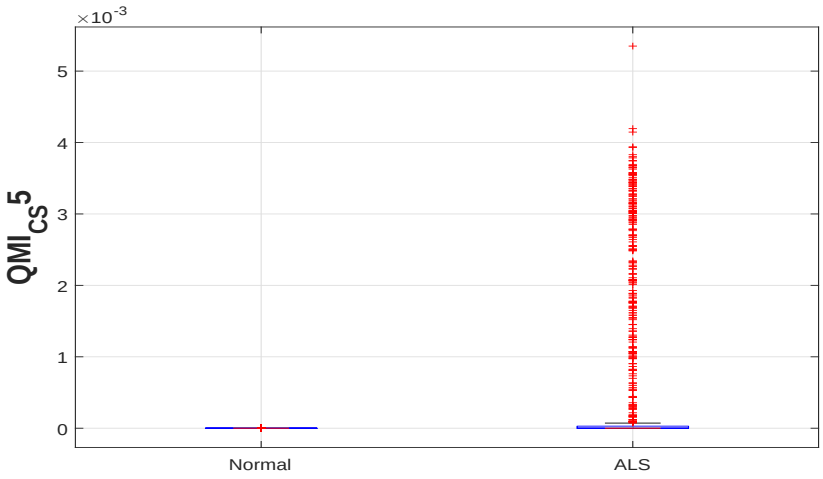

Figure 4.11: Box plot of $\mathrm{QMI_{CS}}$ feature of 5^{th} IMFs for normal and ALS EMG signals.

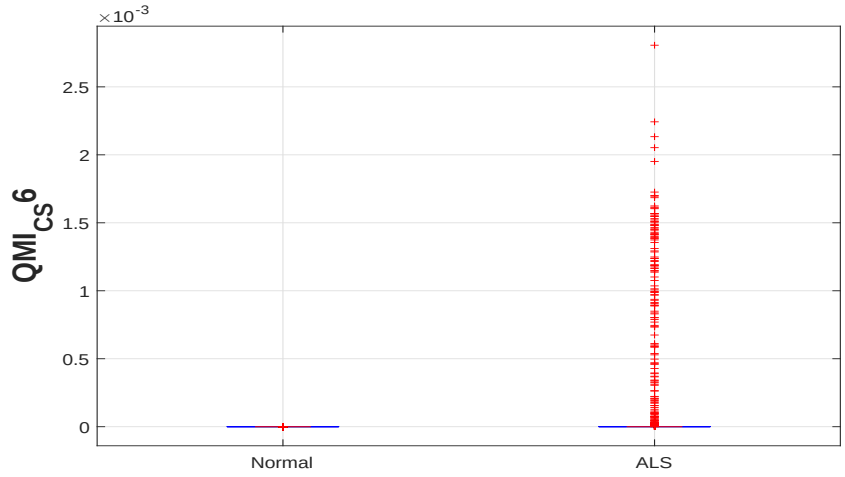

Figure 4.12: Box plot of $\mathrm{QMI_{CS}}$ feature of 6^{th} IMFs for normal and ALS EMG signals.

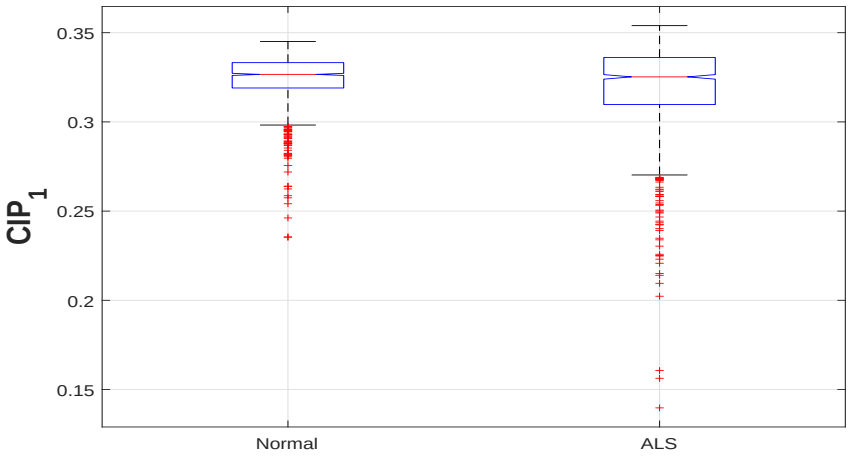

Figure 4.13: Box plot of CIP feature of 1^{st} IMFs for normal and ALS EMG signals.

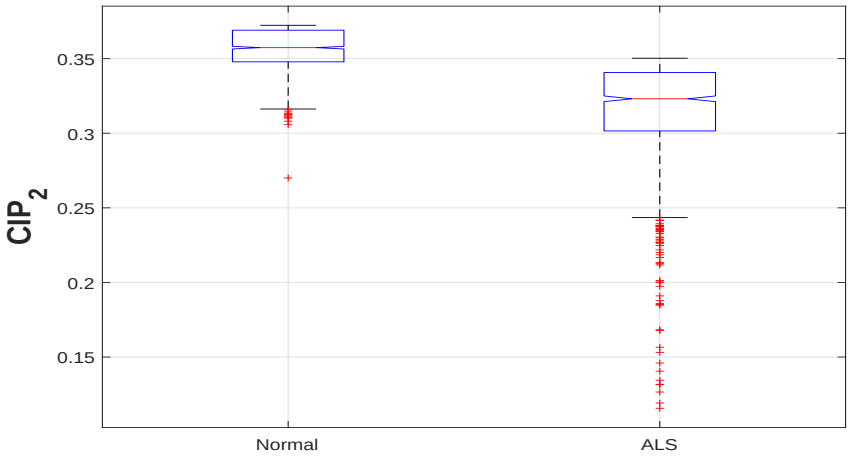

Figure 4.14: Box plot of CIP feature of 2nd IMFs for normal and ALS EMG signals.

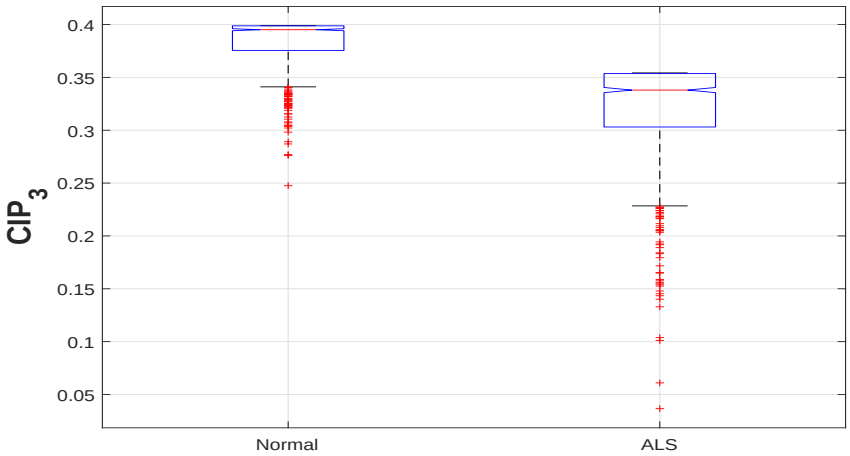

Figure 4.15: Box plot of CIP feature of 3rd IMFs for normal and ALS EMG signals.

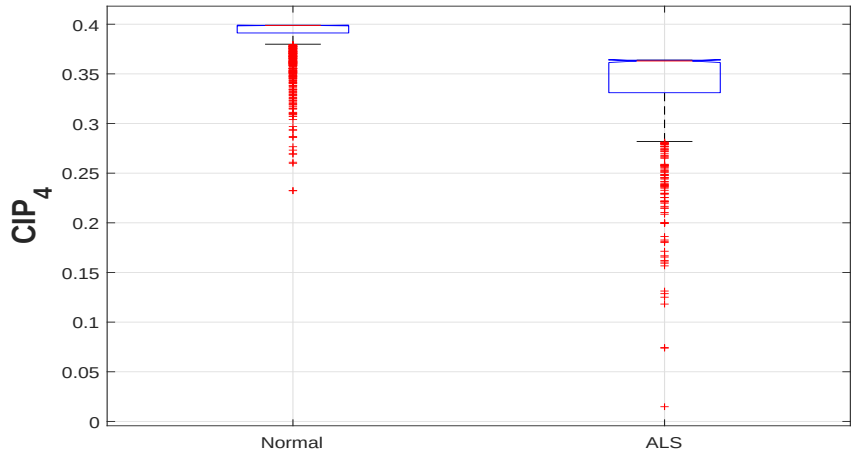

Figure 4.16: Box plot of CIP feature of 4$^{\text{th}}$ IMFs for normal and ALS EMG signals.

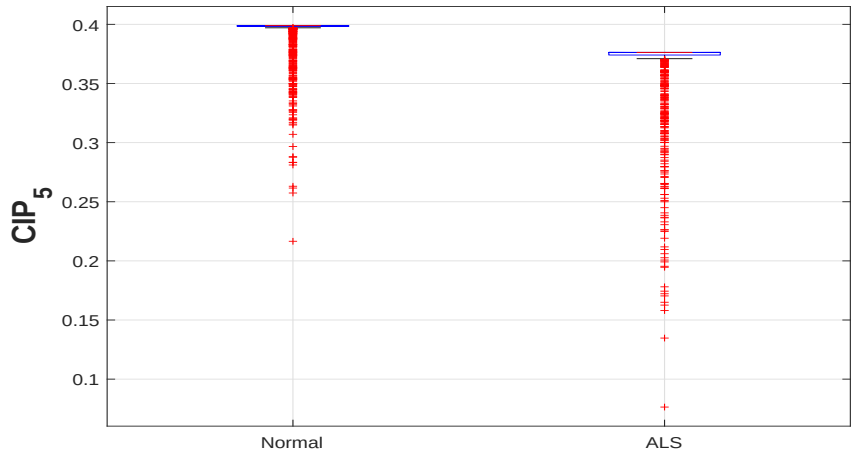

Figure 4.17: Box plot of CIP feature of 5$^{\text{th}}$ IMFs for normal and ALS EMG signals.

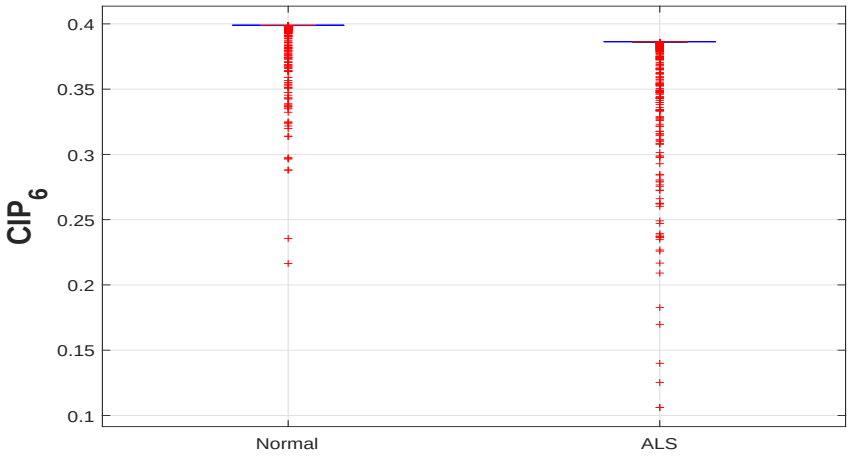

Figure 4.18: Box plot of CIP feature of 6$^{\text{th}}$ IMFs for normal and ALS EMG signals.

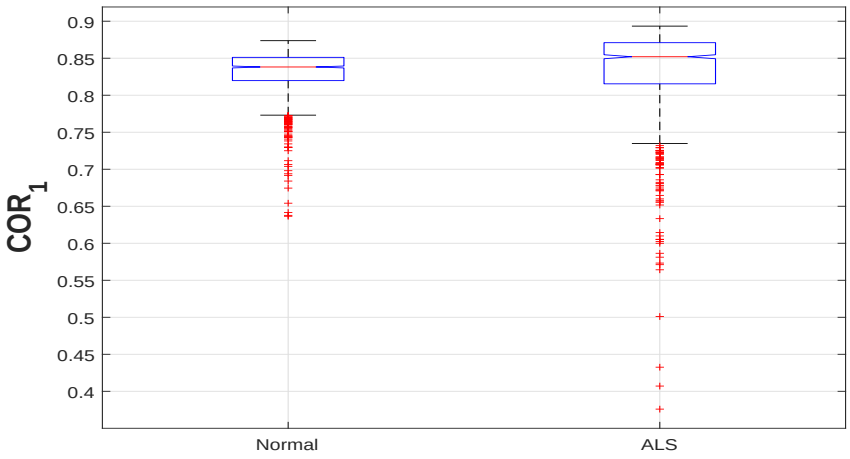

Figure 4.19: Box plot of COR feature of 1$^{\text{st}}$ IMFs for normal and ALS EMG signals.

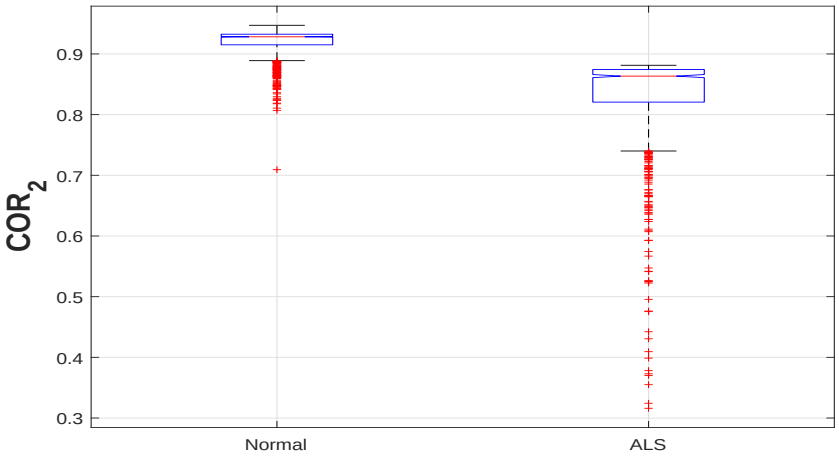

Figure 4.20: Box plot of COR feature of 2nd IMFs for normal and ALS EMG signals.

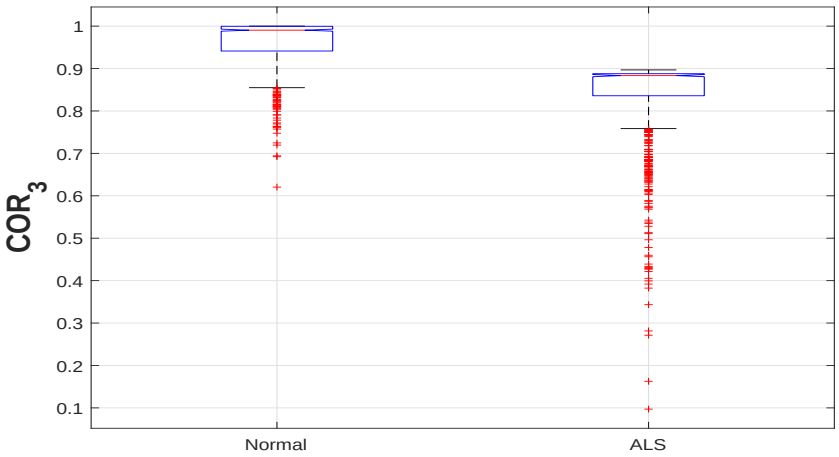

Figure 4.21: Box plot of COR feature of 3rd IMFs for normal and ALS EMG signals.

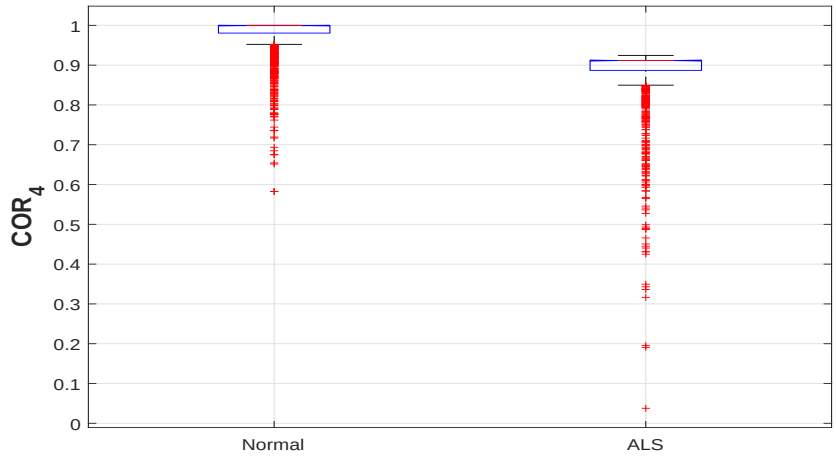

Figure 4.22: Box plot of COR feature of 4$^{\text{th}}$ IMFs for normal and ALS EMG signals.

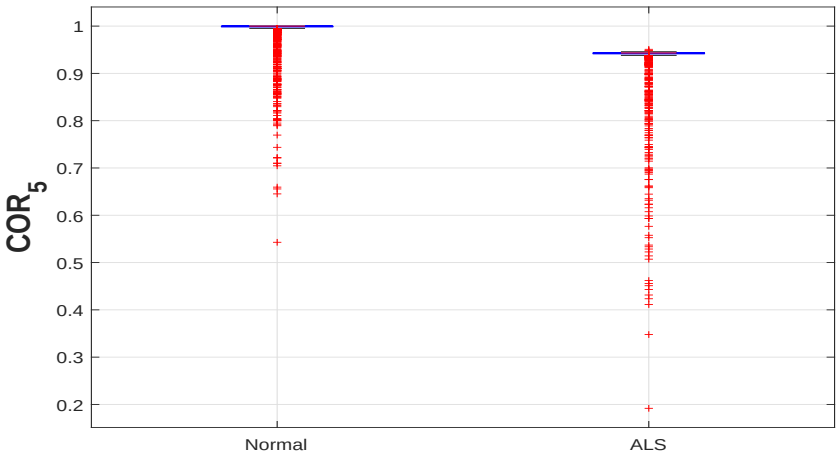

Figure 4.23: Box plot of COR feature of 5$^{\text{th}}$ IMFs for normal and ALS EMG signals.

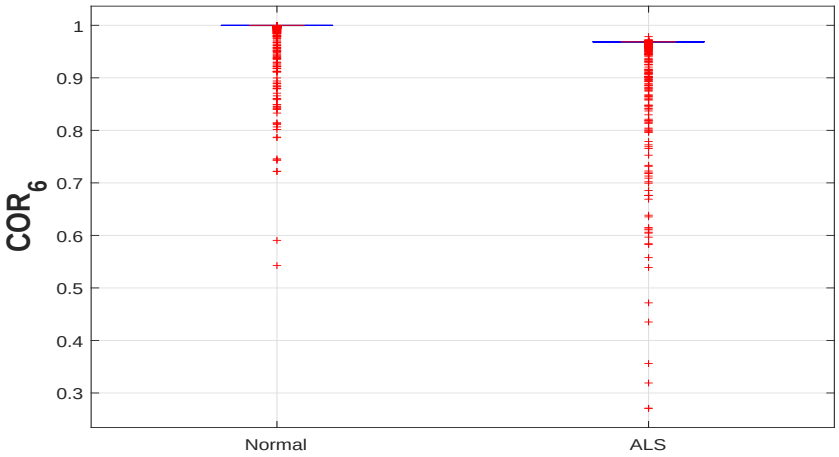

Figure 4.24: Box plot of COR feature of 6^{th} IMFs for normal and ALS EMG signals.

4.2 Classification results

For EMG signal classification, we have given all four calculated features i.e. ED-QMI, CS-QMI, CIP and COR separately in three different classifiers: JRip rules classifier, REP tree classifier and random forest classifier, which have been already discussed in previous chapter. Classification have been done using 10-fold cross validation method and three distinguishing parameters sensitivity (Sen), Specificity (Spe) and accuracy (Acc) have been calculated from the features. To calculate these three parameters, the data after classification is divided into four sets: true positive (TP), true negative (TN), false positive (FP), false negative (FN). For two classes normal and ALS, these four terms can be defined as follows:

1. TP: the number of cases actually belong to ALS class and correctly identified as ALS class.

2. TN: the number of cases actually belong to normal class and correctly identified as normal class.

3. FP: the number of cases actually belong to normal class but incorrectly identified as ALS class.

4. FN: the number of cases actually belong to ALS class but incorrectly identified as normal class.

On the basis of above four terms, the three parameters viz. Acc, Sen and Spe can be defined as follows:

- Acc: The Acc can be defined as the ability to differentiate the ALS class and Normal class correctly. It is the ratio of the number of correct assessment to the number of all assessments. Mathematically, it can be represented as:

$$\text{Acc} = \frac{TP + TN}{TP + TN + FP + FN} \tag{4.1}$$

- Sen: The Sen can be defined as the ability to determine the ALS class correctly. It is the the ratio of the true positive assessment to the total number of positive

assessments. Mathematically, it can be represented as:

$$\text{Sen} = \frac{TP}{TP + FN} \tag{4.2}$$

- Spe: The Spe can be defined as the ability to determine the Normal class correctly. It is the the ratio of the number of true negative assessment to the total number of negative assessments. Mathematically, it can be represented as:

$$\text{Spe} = \frac{TN}{TN + FP} \tag{4.3}$$

We have calculated Acc, Sen and Spe for all four features ED-QMI, CS-QMI, CIP and COR obtained from all six IMFs, with three different classifiers: JRip rules classifier, REP tree classifier and random forest classifier. The three parameters viz. Acc, Sen and Spe obtained from classification process are shown below in the tabular form.

Table 4.5 shows the classification results when we have used only one feature to classify Normal and ALS EMG signals. The results are shown for five different features: (i) ED-QMI feature calculated for 3^{rd} IMFs (ii) CS-QMI feature calculated for 3^{rd} IMFs (iii) CIP feature calculated for 6^{th} IMFs (iv) COR feature calculated for 2^{nd} IMFs (v) COR feature calculated for 6^{th} IMFs. Among all features, the above mentioned five features are giving the best results.

Table 4.6 shows the classification results when we have used two features to classify Normal and ALS EMG signals. We have tried different combinations of two features and the below mentioned five feature combinations are giving the best results. The results are shown for five different combination of two features which are: (i) CIP feature calculated for 6^{th} IMFs and COR feature calculated for 2^{nd} IMFs (ii) COR feature calculated for 2^{nd} IMFs and COR feature calculated for 6^{th} IMFs (iii) CIP feature calculated for 3^{rd} IMFs and CIP feature calculated for 6^{th} IMFs (iv) ED-QMI feature calculated for 3^{rd} IMFs and COR feature calculated for 2^{nd} IMFs (v) CS-QMI feature calculated for 3^{rd} IMFs and COR feature calculated for 2^{nd} IMFs.

Table 4.5: Classification results by using only one feature.

Features	Classifiers	Spe (%)	Sen (%)	Acc (%)
$QMI_{ED}3$	JRip	99.7	95.3	97.8
	REP tree	99.6	95.3	97.7
	Random forest	97.8	95.9	97.0
$QMI_{CS}3$	JRip	99.6	95.3	97.7
	REP tree	99.4	95.7	97.8
	Random forest	96.8	96.0	96.4
CIP_6	JRip	94.7	100	96.9
	REP tree	94.7	99.6	96.8
	Random forest	96.6	95.1	95.9
COR_2	JRip	94.2	99.8	96.5
	REP tree	94.0	99.9	96.5
	Random forest	94.5	93.3	94.0
COR_6	JRip	94.7	99.6	96.7
	REP tree	94.7	99.9	96.8
	Random forest	96.7	94.2	95.6

Table 4.6: Classification results by using two features.

Features	Classifiers	Spe (%)	Sen (%)	Acc (%)
CIP_6+COR_2	JRip	99.9	99.8	99.8
	REP tree	99.7	99.9	99.7
	Random forest	99.9	99.8	99.8
COR_2+COR_6	JRip	99.8	99.6	99.7
	REP tree	99.6	99.8	99.7
	Random forest	99.8	99.7	99.7
CIP_3+CIP_6	JRip	99.3	99.9	99.5
	REP tree	99.3	99.9	99.5
	Random forest	99.4	99.6	99.4
$QMI_{ED}3$+COR_2	JRip	98.7	97.1	98.0
	REP tree	98.4	97.1	97.8
	Random forest	98.7	98.3	98.5
$QMI_{CS}3$+COR_2	JRip	98.5	97.2	97.9
	REP tree	98.8	96.6	97.8
	Random forest	98.7	98.1	98.4

Table 4.7: Performance Comparison with other methods.

Features	Classifiers	Acc (%)
MFCC based method, (2014) [57]	KNN	92.5
TQWT based method, (2018) [58]	LS-SVM	95
CNN based method, (2017) [59]	CNN	96.8
Proposed method	Random forest	99.8

Table 4.7 shows the performance comparison of our proposed method with other previously existing methods for classification of EMG signals, which indicates that our proposed method is better for the classification of normal and ALS EMG signals.

4.3 Summary

In this chapter, we presented the results obtained from analysis and classification process for four features: ED-QMI, CS-QMI, CIP and COR of normal and ALS EMG signals in the form of tables and graphs. For analysis, Kruskal-Wallis statistical test is performed for $p < 0.05$. The p values obtained for all four features are presented in the form of tables separately. And the box plots obtained from all four features are also presented in this chapter. For classification of EMG signals, we have used three different classifiers: JRip rules classifier, REP tree classifier and random forest classifier. In this chapter we discussed about three classification parameters briefly. These three parameters have been calculated for all features. The best classification results are shown in the form of tables for when we are using only one feature for classification and for when we are using two features for classification.

The obtained analysis and classification results show that our proposed method is very much capable of distinguishing between normal and ALS EMG signals and very accurate for classify EMG signals.

Chapter 5

Conclusion and future work

In this proposed work, we have been presented a methodology for analysis of EMG signals and for classification of normal and ALS EMG signals. The extracted MUAPs from EMG signals are decomposed into first six IMFs for both classes and then ED-QMI, CS-QMI, CIP and COR features have been calculated.

Kruskal-Wallis statistical test is performed to analyze EMG signals for p is less than 0.05. We have obtained $p = 0$ for $\text{QMI}_{\text{ED}}3$, $\text{QMI}_{\text{ED}}4$, $\text{QMI}_{\text{CS}}3$, $\text{QMI}_{\text{CS}}4$, CIP_3, CIP_6 and COR_2. For remaining features also, p-values obtained are very less than 0.05 and very close to zero. The box plot obtained for these features are also signifies the difference in the ALS and normal signals. From these observations, we can easily say that this proposed methodology is good enough to analyze the EMG signals and with this proposed method, we can easily distinguish between normal and ALS EMG signals.

For EMG signal classification, we have given ED-QMI, CS-QMI, CIP and COR features in three different classifiers: JRip rules classifier, REP tree classifier and random forest classifier. Classification have been done using 10-fold cross validation method and three distinguishing parameters Sen, Spe and Acc have been calculated from the features. The classification results obtained show that even with only one feature we got the maximum Acc of 97.8%, and the highest Spe and Sen obtained are 99.7% and 100% respectively. The Acc is increased upto 99.8% when we use only two features for classification and highest Spe and Sen obtained are 99.9% and 99.9% respectively. These classification results show that our proposed methodology

is good enough to classify normal and ALS EMG signals very accurately.

On the basis of results obtained for analysis and classification of EMG signals, it can be easily concluded that IF decomposition method is very efficient with even less number of features extracted. This methodology can help in differentiating abnormal and normal EMG for diagnosis of subjects.

In future, the proposed method can be studied for analysis and classification of other biomedical signals corresponding to normal and abnormal classes.

REFERENCES

[1] National cancer institute, Surveillance, Epidemiology and End Results (SEER) Program.training program modules. https://training.seer.cancer.gov/anatomy/body/ .

[2] E. Mtui, M.J.T. FitzGerald, and G. Gruener, "Clinical Neuroanatomy and Neuroscience", Saunders , 2011.

[3] E.R. Kandel and J.H. Schwartz, "Principles of Neural Science", Appleton & Lange: McGraw Hill, 2012.

[4] M. Anthea, J. Hopkins, C.W. McLaughlin, S. Johnson, M.Q. Warner, D LaHart, and J.D. Wright, "Human Biology and Health", Prentice Hall, 1993.

[5] C.Perrin, J.N. Unterborn, C.D. Ambrosio, et al., "Pulmonary complications of chronic neuromuscular diseases and their management".CrossRefPubMedWeb of ScienceGoogle Scholar,vol. 29, pp. 5-27, 2004.

[6] M. Hill, T. Hughes, C. Milford, "Treatment for swallowing difficulties (dysphagia) in chronic muscle disease",Cochrane Database Syst Rev 2004;2:CD004303 PubMedGoogle Scholar. 2004.

[7] A.E.H. Emery,"Population frequencies of inherited neuromuscular diseases - A world survey",Neuromuscular Disorders, vol. 1, no. 1, pp. 19-29, 1991.

[8] European respiratory journal, ERS publications. https://erj.ersjournals.com/.

[9] G. DeMaagd and A. Philip, "Parkinsons disease and its management: part 1: disease entity, risk factors, pathophysiology, clinical presentation, and diagnosis",Pharmacy and Therapeutics, vol. 40, no. 8, pp. 504-532, 2015.

[10] Neuromuscular diseases in the MDA Program, Muscular Dystrophy Association USA, https://www.betterhealth.vic.gov.au/health/ConditionsAndTreatments /neuromuscular-disorders.

[11] M.B.I. Raez, M.S. Hussain, F.Mohd.Yasin, "Techniques of EMG signal analysis: detection, processing, classification and applications", vol. 8, pp. 11-35, 2006.

[12] Science Human body tribute. http://sciencehumanbodytribute.weebly.com/muscular-system.html.

[13] Birbrair, Alexander; Zhang, Tan; Wang, Zhong-Min; Messi, Maria Laura; Enikolopov, Grigori N.; Mintz, Akiva; Delbono, Osvaldo. "Role of Pericytes in Skeletal Muscle Regeneration and Fat Accumulation". Stem Cells and Development. 22 (16): 2298-2314. doi:10.1089/scd.2012.0647. ISSN 1547-3287. PMC 3730538. PMID 23517218, 2013.

[14] W. Frontera and J. Ochala, "Skeletal muscle: A brief review of structure and function",Calcified Tissue International, vol. 96, 2014.

[15] S.J. Valberg, Skeletal Muscle Function, Clinical Biochemistry of Domestic Animals, 2008.

[16] R.R. Sharma , P. Chandra , R.B. Pachori, " Electromyogram Signal Analysis Using Eigenvalue Decomposition of the Hankel Matrix". Advances in Intelligent Systems and Computing, vol 748. Springer, Singapore, 2019.

[17] M. Hill, "The Neuromuscular Junction Disorders". Journal of Neurology, Neurosurgery & Psychiatry. 74 (90002): ii32-ii37. doi:10.1136/jnnp.74.suppl 2.ii32. PMC 1765619. PMID 12754327, 2003.

[18] Kraker, Jessica, A. Zivkovic, Sasa, "Autoimmune Neuromuscular Disorders". Current Neuropharmacology. 9 (3): 400-8. doi:10.2174/157015911796558000. PMC 3151594. PMID 22379454, 2011.

[19] "Neuromuscular Disorders: MedlinePlus". www.nlm.nih.gov. Retrieved 2016-04-24.

[20] "Amyotrophic Lateral Sclerosis (ALS) Fact Sheet", NINDS, Publication date June 2013.NIH Publication No. 16-916, 2013.

[21] "Motor Neuron Diseases Fact Sheet," NINDS. NIH Publication No. 12-5371. Publication date March, 2012.

[22] E.V. Hobson, C.J. McDermott "Supportive and symptomatic management of amyotrophic lateral sclerosis". Nature Reviews. Neurology. 12 (9): 526-38. doi:10.1038/nrneurol.2016.111. PMID 27514291, 2016.

[23] M.C. Kiernan, S. Vucic, B.C. Cheah, M.R. Turner, A. Eisen, O. Hardiman, J.R. Burrell, M.C. Zoing, "Amyotrophic lateral sclerosis". Lancet. 377 (9769): 942-55. doi:10.1016/s0140-6736(10)61156-7. PMID 21296405, 2011.

[24] O. Hardiman, A. Al-Chalabi, A. Chio, E.M. Corr, G. Logroscino, W. Robberecht, P.J.Shaw, Z. Simmons, L.H. van den Berg, "Amyotrophic lateral sclerosis". Nature Reviews. Disease Primers. 3 (17071): 17071. doi:10.1038/nrdp.2017.71. PMID 28980624, 2017.

[25] R.G.Miller, J.D. Mitchell, D.H. Moore, "Riluzole for amyotrophic lateral sclerosis (ALS)/motor neuron disease (MND)". The Cochrane Database of Systematic Reviews, 3(3) : $CD001447.doi$: $10.1002/14651858.CD001447.pub3.PMID22419278$, 2012.

[26] W.F. Brown, C.F. Bolton, and M.J. Amino, Neuromuscular Function and Disease, WB Saunders, 2002.

[27] D. Gordon, E. Robertson, G.E. Caldwell, J. Hamill, G. Kamen, S.N. Whittlesey,"Research Methods in Biomechanics". Champaign, IL: Human Kinetics Publ., 2004.

[28] Electromyography at the US National Library of Medicine Medical Subject Headings (MeSH).

[29] F. Buchtal, H. Schmalbruch, "Motor Unit of Mammalian Muscle". Physiological Reviews. 60 (1): PP.90 - 142, 1980.

[30] S. Shahid,"Higher Order Statistics Techniques Applied to EMG Signal Analysis and Characterization", Ph.D. thesis, University of Limerick, Ireland, 2004.

[31] M.J. Asmussen, V. Tscharner, and B.M. Nigg, "Motor unit action potential clustering- theoretical consideration for muscle activation during a motor task," Frontiers in Human Neuroscience, vol. 12, no. 15, 2018.

[32] J.V. Basmajian,"Muscles Alive - Their Functions Revealed by Electromyography", The Williams & Wilkins Company, Baltimore, 1962.

[33] A. N. Paithane, D. S. Bormane. "Analysis of nonlinear and nonstationary signal to extract the features using Hilbert Huang transform." Computational Intelligence and Computing Research (ICCIC), 2014 IEEE International Conference on. IEEE, 2014.

[34] B.M. Nigg & W. Herzog, "Biomechanics of the Musculo-Skeletal system". Wiley. Page:349, 1999.

[35] Patterson, R. John, "Fitwise". Castillo. Brian T. Retrieved 24 June 2009.

[36] L. Shaw, and B. Sangeeta, "Online EMG signal analysis for diagnosis of neuromuscular diseases by using PCA and PNN", International Journal Of Engineering Science and Technology, vol. 4, pp. 4453 - 4459, 2012.

[37] D. Luca, C. J.,"Surface electromyography: Detection and recording". DelSys Incorporated, 2002.

[38] J. Wang, L. Tang, "Surface EMG Signal Amplification and Filtering", International Journal of Computer Applications (0975 - 8887) ,Vol.82 - No1, November 2013.

[39] American Association of Neuromuscular & Electrodiagnostic Medicine. www.aanem.org.

[40] Yunfen Wu and Maria Angeles Martinez Martinez and Pedro Orizaola Balaguer,"Overview of the Application of EMG Recording in the Diagnosis and Approach of Neurological Disorders",10.5772/56030, 2013

[41] E. Schulte, D. Farina, R. Merletti, G. Rau, and C. Disselhorst-Klug,"Inuence of muscle fibre shortening on estimates of conduction velocity and spectral frequencies from surface electromyographic signals", Medical and Biological Engineering and Computing, vol. 42, no. 4, pp. 477 - 486, 2004.

[42] Ismail Saad, Nur Husna Bais, Bun Seng C., Mohd. Zuhir H., N. Bolong "Electromyogram (EMG) Signal Processing Analysis for Clinical Rehabilitation Application",2012.

[43] J. Yousefi and A. Hamilton-Wright, "Characterizing EMG data using machine-learning tools", Computers in Biology and Medicine , vol. 51,pp. 1-13, 2014.

[44] G. Pfeiffier,"The diagnostic power of motor unit potential analysis: an objective Bayesian approach", Muscle Nerve, vol. 22, no. 5, pp. 584-591, 1999.

[45] G. Pfeiffier, "Sequential diagnostic decisions in quantitative motor-unitpotential analysis: a Bayesian approach", Electroencephalography and Clinical Neurophysiology, vol. 99, no. 6, pp. 306, 1996.

[46] G. Pfeiffer and K. Kunze, "Discriminant classification of motor unit potentials (MUPs) successfully separates neurogenic and myopathic conditions. A comparison of multi and univariate diagnostical algorithms for MUP analysis", Electroencephalography and Clinical Neurophysiology/Electromyography and Motor Control, vol. 97, no. 5, pp. 191-207, 1995.

[47] Z. Barutcuoglu and C. DeCoro, "Hierarchical shape classification using Bayesian aggregation", IEEE International Conference on Shape Modeling and Applications, pp. 44, 2006.

[48] Y. Xiang, B. Pant, A. Eisen, M.P. Beddoes, and D. Poole, "Multiply sectioned Bayesian networks for neuromuscular diagnosis", Artificial Intelligence in Medicine, vol. 5, no. 4, pp. 293-314, 1993.

[49] E.W. Abel, P.C. Zacharia, A. Forster, and T.L. Farrow, "Neural network analysis of the EMG interference pattern", Medical Engineering & Physics, vol. 18, no. 1, pp. 12-17, 1996.

[50] B. Pandey and R.B. Mishra, "An integrated intelligent computing model for the interpretation of EMG based neuromuscular diseases", Expert Systems with Applications, vol. 36, no. 5, pp. 9201-9213, 2009.

[51] C.D. Katsis, D.I. Fotiadis, A. Likas, and I. Sarmas, "A novel EMG decomposition method that discovers the number of MUAP clusters", Proceedings of the 2nd European Medical and Biological Engineering Conference , pp. 400-401, 2002.

[52] C.I. Christodoulou and C.S. Pattichis,"Unsupervised pattern recognition for the classification of EMG signals",IEEE Transactions on Biomedical Engineering, vol. 46, no. 2, pp. 169-178, 1999.

[53] M. Gazzoni, D. Farina, and R. Merletti,"A new method for the extraction and classification of single motor unit action potentials from surface EMG signals", Journal of Neuroscience Methods, vol. 136, no. 2, pp. 165-177, 2004.

[54] A. Subasi, M. Yilmaz, and H.R. Ozcalik, "Classification of EMG signals using wavelet neural network", Journal of Neuroscience Methods, vol. 156, no. 1, pp. 360-367, 2006.

[55] 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", The Royal Society, vol. 454, no. 1971, pp. 903-995, 1998.

[56] S. Kouchaki, R. Boostani, S. Shabani, and H. Parsaei, "A new feature selection method for classification of EMG signals",The 16th CSI International Symposium on Artificial Intelligence and Signal Processing, pp. 585-590, 2012.

[57] A.B.M.S.U. Doulah, S.A. Fattah, "Neuromuscular disease classification based on mel frequency cepstrum of motor unit action potential". In: Electrical Engineering and Information & Communication Technology (ICEEICT), International Conference, pp. 1-4, 2014.

[58] P.U. Kiran, N. Abhiram, S. Taran, V. Bajaj, "TQWT Based Features for Classification of ALS and Healthy EMG Signals". American Journal of Computer Science and Information Technology. vol.6 No.2:19. doi: 10.21767/2349-3917.100019, (2018).

[59] A. Sengur, Y. Akbulut, Y. Guo, V. Bajaj, "Classification of amyotrophic lateral sclerosis disease based on convolutional neural network and reinforcement sample learning algorithm". Health Information Science and Systems. doi:10.1007/s13755-017-0029-6, vol. 5(1), 9. (2017).

[60] E.W. Abel, H. Meng, A. Forster, and D. Holder,"Singularity characteristics of needle EMG IP signals", IEEE Transactions on Biomedical Engineering, vol. 53, no. 2, pp. 219-225, 2006.

[61] K. Englehart, B. Hudgin, and P.A. Parker, "A wavelet-based continuous classification scheme for multifunction myoelectric control", IEEE Transactions on Biomedical Engineering, vol. 48, no. 3, pp. 302-311, 2001.

[62] N.F. Guler and S. Kocer, "Classification of EMG signals using PCA and FFT", Journal of Medical Systems, vol. 29, no. 3, pp. 241-255, 2005.

[63] M. Nikolic and C. Krarup, "EMGTools, an adaptive and versatile tool for detailed EMG analysis", IEEE Transactions on Biomedical Engineering, vol. 58, no. 10, pp. 2707-2718, 2011.

[64] D. Joshi, A. Tripathi, R. Sharma, and R.B. Pachori, "Computer aided detection of abnormal EMG signals based on tunable-Q wavelet transform",4th International Conference on Signal Processing and Integrated Networks (SPIN), pp. 544-549, 2017.

[65] A.B.M.S.U. Doulah, S.A. Fattah, W.P. Zhu, and M.O. Ahmad, "Wavelet domain feature extraction scheme based on dominant motor unit action potential of EMG signal for neuromuscular disease classification", IEEE Transactions on Biomedical Circuits and Systems, vol. 8, no. 2, pp. 155-164, 2014.

[66] A. Hazarika, L. Dutta, M. Barthakur, and M. Bhuyan, "Two-fold feature extraction technique for biomedical signals classification", International Conference on Inventive Computation Technologies, vol. 2, pp. 1-4, 2016.

[67] E. Stalberg, C. Bischoff, and B. Falck, "Outliers, a way to detect abnormality in quantitative EMG", Muscle & Nerve, vol. 17, pp. 392-399, 1994.

[68] O. Ulkir, G. Gokmen, and E. Kaplanoglu, "EMG signal classification using fuzzy logic", Balakan Journal of Electrcical and Computer Engineering, vol. 5, no. 2, pp. 97-101, 2017.

[69] A. Cicone, J. Liu, and H. Zhou, "Adaptive local iterative filtering for signal decomposition and instantaneous frequency analysis", arXiv preprint arXiv:1411.6051, 2014.

[70] 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", Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, 454(1971):903, 1998.

[71] L. Lin, Y. Wang, and H. Zhou. "Iterative filtering as an alternative algorithm for empirical mode decomposition", Advances in Adaptive Data Analysis, 1(4):543-560, 2009.

[72] N. E. Huang, M. L. C. Wu, S. R. Long, S. S. P. Shen, W. Qu, P. Gloersen, and K. L. Fan, "A confidence limit for the empirical mode decomposition and Hilbert spectral analysis", Proceedings of the Royal Society of London, Series A: Mathematical, Physical and Engineering Sciences, vol. 459, pp. 2317-2345, 2003.

[73] Y. Wang, G. W. Wei, and S. Yang, "Iterative filtering decomposition based on local spectral evolution kernel", Journal of Scientific Computing, vol. 50, no. 3, pp. 629-664, 2012.

[74] K.C. McGill, Z.C. Lateva, and H.R. Marateb,"EMGLAB: An interactive EMG decomposition program", Journal of Neuroscience Methods, vol. 149, no. 2, pp. 121-133, 2005.

[75] K.C. McGill , Z.C. Lateva, and M.E. Johanson,"Validation of a computer-aided EMG decomposition method", Proceeding IEEE Eng Med Biol Soc Conference, pp.4744-4747, 2004.

[76] V.M. Koch, K.C. McGill, and H.A. Loeliger, "Resolution of superpositions in EMG signals using belief propagation: results for the known constituent problem", IEEE Engineering in Medicine and Biology Society, 2006.

[77] M. Nikolic and C. Krarup, "EMGTools, an Adaptive and Versatile Tool for Detailed EMG Analysis", IEEE Transactions on Biomedical Engineering, vol. 58, pp.2707-2718, 2011.

[78] K.C. McGill, "Optimal resolution of superimposed action potentials", IEEE Transactions on Biomedical Engineering, vol. 49, pp. 640-650, 2002.

[79] K.C. McGill, K.L. Cummins, and L.J. Dorfman, "Automatic decomposition of the clinical electromyogram", IEEE Transactions on Biomedical Engineering, vol. 32, pp. 470-477, 1985.

[80] K.C. McGill, Z.C. Lateva, H.R. Marateb, "EMGLAB: an interactive EMG decomposition program." J Neurosci Methods, vol. 149(2), pp. 121-33, 2005.

[81] Voot Tangkaratt, Hiroaki Sasaki, and Masashi Sugiyama," Direct Estimation of the Derivative of Quadratic Mutual Information with Application in Supervised Dimension Reduction", arXiv:1508.01019v1, 5 Aug 2015.

[82] J. C. Principe, D. Xu, Q. Zhao, and J. W. Fisher, "Learning from examples with information theoretic criteria.", VLSI Signal Processing, 26(1-2):61-77, 2000.

[83] Dongxin xu ,"Energy, entropy and information potential for neural computation", pp. 31-33, 1999.

[84] J.W. Xu, A.R.C. Paiva, I. Park, and J.C. Principe, "A reproducing kernel Hilbert space framework for information-theoretic learning",IEEE Transactions on Signal Processing, vol. 56, no. 12, pp. 5891-5902, 2008.

[85] J. C. Principe," Information Theoretic Learning: Renyi's Entropy and Kernel Perspectives" . pp.358, 2010.

[86] H. Tang, and H. Li,"Information theoretic learning: Renyi's entropy and kernel perspectives", IEEE Computational Intelligence Magazine, vol. 6, no. 3, pp. 60-62, 2011.

[87] M. Kumar, R. Pachori, and U.R. Acharya,"Characterization of coronary artery disease using exible analytic wavelet transform applied on ECG signals", Biomedical Signal Processing and Control, vol. 31, pp. 301-308, 2017.

[88] W. Liu, P.P. Pokharel, and J.C. Principe, "Correntropy: properties and applications in non-gaussian signal processing", IEEE Transactions on Signal Processing, vol. 55, no. 11, pp. 5286-5298, 2007.

[89] A. Gunduz and J.C. Principe,"Correntropy as a novel measure for nonlinearity tests," Signal Processing, vol. 89, no. 1, pp. 14-23, 2009.

[90] M. Dalal, M. Tanveer, and R. Pachori, "Automated identification system for focal EEG signals using fractal dimension of FAWT based sub-bands signals," International Conference on Machine Intelligence and Signal Processing, 2017.

[91] R. Sharma, M. Kumar, R.B. Pachori, and U.R. Acharya, "Decision support system for focal EEG signals using tunable-Q wavelet transform," Journal of Computational Science, vol. 20, pp. 52-60, 2017.

[92] P.E. McKight and J. Najab, Kruskal-Wallis Test, Corsini Encyclopedia of Psychology, 2010.

[93] T.P. Hettmansperger, Statistical Methods Based on Ranks, Wiley, New York, 1984.

[94] C. Siegel, Castellan, "Nonparametric Statistics for the Behavioral Sciences (Second ed.)", New York: McGraw-Hill. ISBN 0070573573, 1998.

[95] Corder, W. Gregory, Foreman, I. Dale, "Nonparametric Statistics for Non-Statisticians", Hoboken: John Wiley & Sons. pp. 99-105. ISBN 9780470454619, 2009.

[96] Kruskal, Wallis, "Use of ranks in one-criterion variance analysis". Journal of the American Statistical Association. 47 (260): 583-621. doi:10.1080/01621459.1952.10483441, 1952.

[97] E. Ostertagova, O. Ostertag, and J. Kovac , "Methodology and Application of the Kruskal-Wallis Test," Applied Mechanics and Materials, vol. 611, pp. 115-120, 2014.

[98] J.S. Maritz, "Distribution-Free Statistical Methods", Chapman and Hall Mathematics Series, CRC Press, 1995.

[99] Mann, B. Henry, Whitney, R. Donald, "On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other", Annals of Mathematical Statistics. 18 (1): 50-60. doi:10.1214/aoms/1177730491. MR 0022058. Zbl 0041.26103, 1947.

[100] J.D. Spurrier,"On the null distribution of the KruskalWallis statistic, "Journal of Nonparametric Statistics, vol. 15, no. 6, pp. 685-691, 2003.

[101] S.B. Green and N.J. Salkind, Using SPSS for Windows and Macintosh : Analyzing and Understanding Data, Pearson Prentice Hall, 2014.

[102] K. Wallis, "Use of ranks in one-criterion variance analysis," Journal of the American Statistical Association, vol. 47, no. 260, pp. 583-621, 1952.

[103] A. Bhattacharyya, R.B. Pachori, and U.R Acharya, "Tunable-Q qavelet transform based multivariate sub-band fuzzy entropy with application to focal EEG signal analysis,"Entropy, vol. 19, no. 3, 2017.

[104] M. Kumar, R.B. Pachori, and U.R. Acharya,"An efficient automated technique for CAD diagnosis using exible analytic wavelet transform and entropy features extracted from HRV signals," Expert Systems with Applications Journal, Elsevier, vol. 63, no. C, pp. 165-172, 2016.

[105] B. G. Becker, "Visualizing decision table classifiers", Proceedings IEEE Symposium on Information Visualization (Cat. No.98TB100258), pp. 102-105, 1998.

[106] Joaquin Vanschoren, J. N. van Rijn, Bernd Bischl, and Luis Torgo. "OpenML: networked science in machine learning". SIGKDD Explorations 15(2), pp 49-60, 2013.

[107] W. W.Cohen, "Fast effective rule induction". In:12th Intl.conf. on Machine learning,pp.115-123, 1995.

[108] Parsania, Dr. Vaishali & Jani, Nalin & Bhalodiya, Navneet. "Applying NaÃŕve bayes, BayesNet, PART, JRip and OneR Algorithms on Hypothyroid Database for Comparative Analysis", IJDI-ERET. 3, 2014.

[109] R. Anil, R.P. Aharwal ,D. Meghna, S.P. Saxena, R. Manmohan. "J48 and JRIP Rules for E-Governance Data", International Journal of Computer Science and Security (IJCSS), Vol. 5 : Issue (2) , 2011.

[110] I.H. Witten, and E. Frank, "Data mining: practical machine learning tools and techniques-2nd ed.", the United States of America, Morgan Kaufmann series in data management systems, 2005.

[111] Dr. B. Srinivasan, P.Mekala, "Mining Social Networking Data for Classification Using REPTree", International Journal of Advance Research in Computer Science and Management Studies, Vol.2, pp-155-160, 2014.

[112] M. Pal , "Random forest classifier for remote sensing classification", International Journal of Remote Sensing, vol. 26, pp. 217-222, 2005.

[113] Tin Kam Ho, "Random decision forests", Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, pp. 278-282, 1995.

[114] Tao Shi and Steve Horvath,"Unsupervised learning with random forest predictors", Journal of Computational and Graphical Statistics, vol. 15, pp. 118-138, 2006.

[115] L. Breiman, "Random forests-random features", Technical Report 567, Statistics Department, University of California, Berkeley, 1999.

[116] W. Feller, "An Introduction to Probability Theory and its Application", vol.1, 3rd edn, New York Wiley, 1968.