Jimma University

Jimma Institute of Technology

School of Biomedical Engineering

Biomedical Imaging Program

Automatic Diagnosis of Parkinson's Disease Using EMG Signals from Different Hand Movements

---

A Thesis submitted to School of Graduate Studies of Jimma Institute of Technology, Jimma University, in Partial Fulfillment for the Degree of Master of Science in Biomedical Imaging

By: Hamdia Murad

Advisors:

Main Advisor: Gizeaddis L. Simegn (Ph.D.)

Co-Advisor: Abel Worku Tessema (MSc.)

December, 2019 G.C.

Jimma, Ethiopia

Declaration

I declare this research with the title of **"Automatic Diagnosis of Parkinson's Disease using EMG Signal from Different Hand Movements**" as my original work and I assure it with my signature.

Name:	Sign	Date
Hamdia Murad	_____	_____

On behalf of the School of Biomedical Engineering at Jimma Institute of Technology, we the advisors of this research with the title of **"Automatic Diagnosis of Parkinson's Disease using EMG Signal from Different Hand Movements**" and We, the evaluators, confirm that this research is approved as MSc. thesis for the student.

Advisors

Name:	Sign	Date
Gizeaddis L. Simegn (Ph.D.)	_____	_____
Abel Worku (MSc.)	_____	_____

Examiners

Name:	Sign	Date
_____	_____	_____
_____	_____	_____

Chair Person

Name:	Sign	Date
_____	_____	_____

Abstract

Parkinson's disease (PD) is the second most common neurodegenerative disease that affects wide range of productive individuals worldwide. It is neurological disorder characterized by muscle rigidity, tremors, uncontrolled movement, change in speech and sleep disorders. These problems arise because of loss of substance called dopamine that act as a messenger between two brain areas, the substantia nigra and the corpus striatum to produce smooth and controlled movements.

The common approach to diagnose PD is through clinical assessment of the patient, which is highly subjective and time consuming. Electromyography (EMG) recordings can be used for diagnosis of PD. However, highly experienced experts are required to interpret the signals, which is complex and time-consuming procedure. These manual procedures are prone to error and may lead to misdiagnosis. Many researchers designed automated systems to solve this problem but they have their own pitfalls such as, achieving limited accuracy, using small number of data sets and sticking with binary classification.

In this research, a reliable, accurate and automatic system for early detection and classification of PD using EMG signals is developed. A total of 1000 EMG signal data were collected from flexor carpi radialis and biceps muscles of 15 PD patients and 10 healthy control subjects at JUMC using SCU-7 EMG system. And the signal was analyzed using MATLAB 2018. Data augmentation for collected signals was performed by adding white noise with SNR value of 90 and 100. The raw EMG signal was denoised by applying an infinite impulse response notch and Butterworth filters. Then features in time and frequency were extracted and feature reduction algorithm was applied to discard irrelevant features. Finally, the selected features were sent to the model to train the four-class classification. Support vector machine was used to classify the features of the signal into four (normal, early, moderate and advanced) classes for each hand movements.

The performance of the system was evaluated at testing phase and a promising result has been found. 90%, 91.7%, 95% and 96.6% overall classification accuracies were obtained for elbow flexion by 90^0 without load, elbow flexion by 90^0 with load, touching shoulder and wrist pronation, respectively. The proposed system will be used as a decision support system for physicians, especially those in low resource setting by detecting PD at early stage and classifying it's level. This will have a great impact in reducing the disease progression and the mortality rate due to PD.

Keywords: - Parkinson's disease, Electromyogram, Detection, Classification, Detection, SVM

Acknowledgment

Firstly, I am humbly grateful to ALLAH for guiding and helping me through all the way and I take this opportunity to express my sincere gratitude to my advisors; Dr. Gizeaddis Lamesgin and Mr. Abel Worku for their encouragements and help to do my thesis on this area and for being my advisors.

I would also extend my special thanks to Jimma Medical center staffs and my clinical collaborators, Dr. Hiwot Amare, Dr. Meakel Belay, Sister Fantaye Wgira, Sister Kebebush Shanko and Mr. Filagot Buzayew (BME) for the time they devoted for me, valuable guidance and willingness to involve in this research.

I would also like to acknowledge the head of the school of Biomedical Engineering, Mr. Wasihun Alemayehu and school academic quality assurance and CBE coordinator Mr. Solomon Gebru, for helping me during the research and facilitating all the necessary administrative tasks related to the thesis work.

My gratitude also goes to Mr. Kokeb Desse, Mrs. Genet Tadesse, Mr. Mohammed Aliy, Mr. Ahmed Ali, Mr. Henok Birhanu (ECE staff) and all BME school staffs whose offices were always open whenever a trouble spot or had a question about my research and for guiding me in the right direction.

Finally, I would like to thank my husband Dr. Mohammed Ibrahim, all my families and my friends for their unfailing support and encouragements.

Acronyms

ANN	Artificial Neural Network
CNS	Central Nervous System
EEG	Electroencephalogram
EMG	Electromyogram
ET	Essential tremor
FFT	Fast Fourier transform
HD	Huntington disease
KNN	K-nearest negbours
KLT	Karhunen-loeve-transform
LDA	Linear discriminant analysis
MAV	Mean absolute value
MDF	Median frequency
MFL	Maximum fractal length
MLP	Multi-Layer Perception
NDD	Neurodegenerative Disease
PD	Parkinson's disease
RBF	Radial basis function
RMS	Root mean square
sEMG	Surface electromyogram
SNR	Signal to noise ration
SN	Substantia nigra

SVM	Support Vector Machine
WL	Wave length
WPA	Wavelet package analysis
ZC	Zero crossing

Table of Contents Page

List of Table

List of Figures

Preface

This thesis is segregated into five chapters. From **chapter 1,** the reader will get the overall introduction about neurodegenerative and Parkinson's disease, EMG machine and selected muscles. This chapter also give highlight about the basic problem, the purpose and significance of this thesis. The next chapter, **Chapter 2**, discusses the works performed in the area of EMG signal recording, processing and classification for Parkinson's disease diagnosis. **Chapter 3** explains the general procedures, materials and methods used in the research. The tests performed and results found in this thesis are discussed and analyzed in detail in **Chapter 4.** The last chapter, **chapter 5,** summarizes all the chapters, discusses the main achievements of the research and leaves a clue to be addressed in the future.

Chapter One

1. Background of the Research

1.1. Neurodegenerative Disease

Neurodegenerative disease is a general terminology that describes a group of neurological disorder characterized by affecting the central nervous system. These heterogeneous categories of diseases have a major impact on human health. Even though the pathophysiology of the disease group is distinct based on the brain part the disease affects [1], according to recent findings these diseases share lots of common degenerative processes that cause neuronal death, leading to functional impairments and dysfunctions of brain parts [2]. Irreversible distraction and progressive deterioration of brain functions, because these diseases directly affect the neurons within the brain, can cause the death of functional neurons in the central nervous system [3].

These diseases are grouped as, Parkinson's, Alzheimer's, Huntington's, Amyotrophic Lateral Sclerosis and other diseases. Patients with these kinds of disease, experience a cognitive decline over a long period and these symptoms include gait abnormalities, problems with speech, and memory loss due to progressive cognitive deterioration that affects memory, thinking, behavior, language, calculation learning and emotion capacity [2]. Physical, mental and social impact of these diseases extend far beyond the obvious characteristics of the diseases' symptoms [3].

Main risk factors for these diseases are genetic causes and increasing age. Some other possible risk factors are gender, endocrine conditions, oxidative stress, inflammation, stroke, hypertension, diabetes, smoking, head trauma, depression, infection, tumors, vitamin deficiencies, immune and metabolic conditions, and environmental factors [4].

1.2. Parkinson's Disease

Parkinson's disease was found by James Parkinson in 1817. He medically described this disease as neurological disorder for the first time. But, the disease was known in ancient Indian and ancient Chinese before 1000 BC [5].

Parkinson's disease is the fastest growing of neurological disorders which is the leading source of disabilities [6]. The major cases of PD are sporadic, that account about 85-90%. However, some cases are related with positive family history. It is the second most common neurodegenerative disorder and the most common movement disorder. An estimated number of 10 million people have PD worldwide [7] and 4% of the people are under the age of 50 [8]. According to study findings, PD is more common in men than in women [5]. Figure below shows the distribution of Parkinson's disease globally.

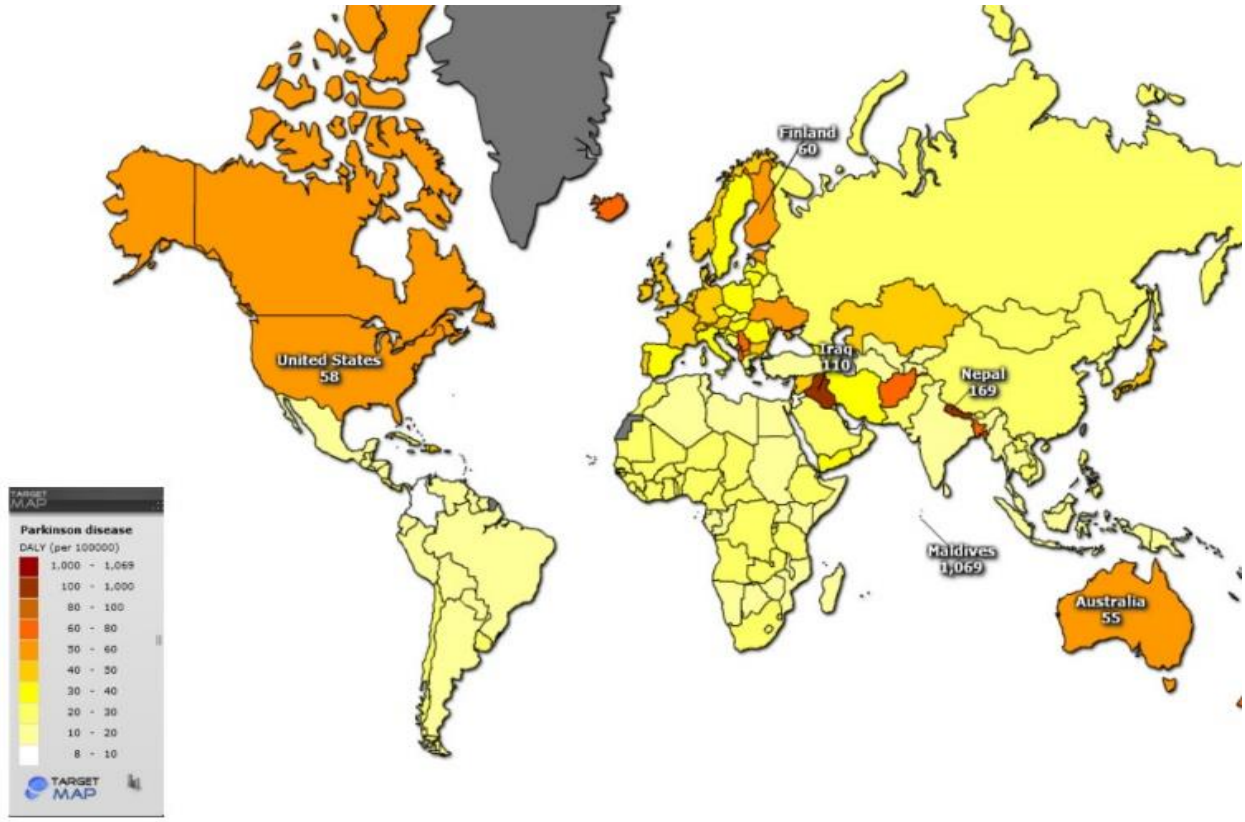


Figure 1:1: Global prevalence of Parkinson's disease, 2016 [6]

Most people who develop PD are 60 years of age or older. Since overall life expectancy is rising, the number of individuals with Parkinson's disease will increase in the future. Adult-onset Parkinson's disease is the most common, but early-onset Parkinson's disease (onset between 21-40 years), and juvenile-onset Parkinson's disease (onset before age 21) can occur in some cases [9].

1.2.1. Parkinson's Disease in Ethiopia

Although there is insufficient information on the prevalence and epidemiology of PD in the continent of Africa, confounding effects show that there are low case ascertainments and high selective mortality due to Parkinson's disease in Africa [10].

The finding of studies indicated that Ethiopia is one of the countries with high prevalence of PD in Africa. A case study that was done for one year in Tikur Anbesa (Black Lion) Hospital showed that 47.7% of movement disorder cases were identified as Parkinson's disease [11]

Even though, the prevalence of PD patients in Ethiopia is high, it remains under recognized and undertreated disease in the country. Additionally, Ethiopian patients with PD experience highest rates PD pain which is related with poverty and ethnicity [12]. Ethiopia is one of the few target African countries for the Michael J. Fox foundation for Parkinson's disease research that encourages researches that focuses on understanding the role of genes in the onsets, its progression, and the cellular effects of the disease [13].

PD has caused multilayered psycho-social and economic impact on patients, family members as well as their close relatives. It was observed that the majority of the patients living with PD either experienced early retirement, self-stigmatization, introvert behaviors, subjective experience of time, lonesomeness, depression, social anxiety and etc., divorce, frequent quarrel with family members, refrain oneself from social gathering, unable to deliver the social and communal commitments, wrong interpretation of PD, unable to make love and give birth are some of the daily challenges of the people with PD [14].

1.2.2. Anatomical Basis and Symptoms of Parkinson's Disease

The human brain which functions as the center for the control of all the parts of human body is a highly specialized organ that allows a human being to adapt and endure varying environmental conditions. Substantia nigra and basal ganglia are brain structures, located in mid part of human brain. It produces an important substance called dopamine for correct function of central nervous system. The produced substance (neurotransmitter dopamine) is responsible for planning and controlling of body movements. In addition, the basal ganglia is responsible for activating and

inhibiting the body circuits based on the message [15]. Figure 1.2 shows the position of basal ganglia and substantia nigra in midbrain.

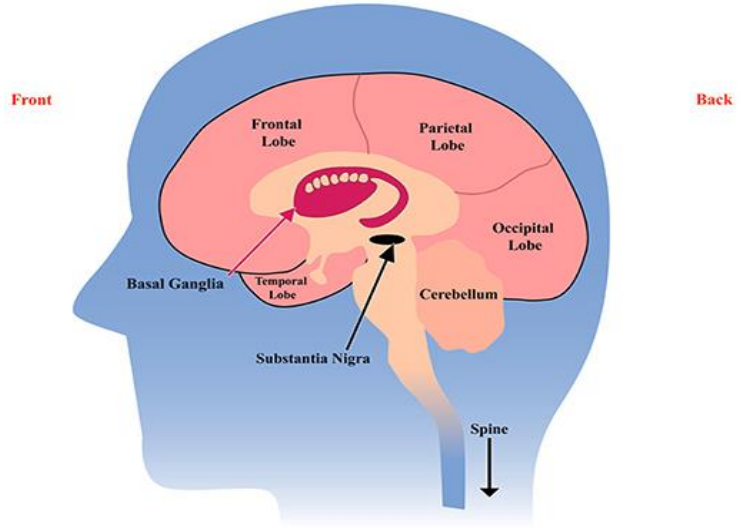


Figure 1:2: Anatomical position of basal ganglia and substantia nigra [16].

Impulse starts in the motor cortex parts of central nerves system for body movements. These signals are transmitted from one motor unit to another by means of dopamine by moving fast from brain to spinal cord and finally to respective muscle. When PD affects the brain, communication between these parts become ineffective [15]. Individuals with PD lack dopamine. This substance acts as a messenger between two brain areas, the substantia nigra and the corpus striatum to produce smooth and controlled movements [17].

The most common symptoms of PD are motor symptoms such as, bradykinesia, rest tremor, rigidity, postural disturbances, falls, freezing of gait, speech and swallowing difficulties [17, 18]. And the spectrum of nonmotor symptoms encompasses constipation, bladder dysfunction, hallucinations, difficulty in concentration, dribbling, dysphagia, confusion, fatigue, impulse control disorders, memory problems, mood disorders, orthostatic hypotension, paranoia, sensation of breathlessness, sleep disturbances, sweating, and sexual dysfunction [19].

Symptoms of PD are caused by lack of dopamine due to the loss of dopamine-producing cells in the substantia nigra. When the amount of dopamine is too low, communication between the substantia nigra and corpus striatum becomes ineffective, and movement becomes impaired; the greater the loss of dopamine, the worse the movement-related symptoms [20]. Parkinson's disease

patients may lose 60% to 80% or more of the dopamine-producing cells in the brain. Characteristics of PD are progressive loss of muscle control, which leads to trembling of the limbs and head while at rest, stiffness, slowness, and impaired balance. As symptoms worsen, it may become difficult to walk, talk, and complete simple tasks [20]. The degree of impairment and progression of this disease is varying from person to person based on different factors. Some people with Parkinson's disease may become disabled much more quickly whereas, others live long productive lives [21].

1.2.3. Challenges of Parkinson's Disease Diagnosis

PD has no definitive way to diagnosis. General diagnosis is done through assessment of motor symptoms and using medical history of patients. The neurologist looks for the presence or absence of the main symptoms of PD such as, tremor, bradykinesia and rigidity by looking Facial expressions, signs of tremor while the patient is at rest, conditions of the patient while standing up from sitting and how the patient regain balance. The other method is physician prescribe a medication and based on response of the medication they decide whether the patient has Parkinson's disease or not. These methods need highly experienced neurologists because some symptoms are common to other diseases such like, essential tremor and ALS [22].

Additionally, neuroimaging test such as, magnetic resonance imaging commonly diffusion, perfusion, and functional MRI, positron emission tomography scan (PET), computed tomography scan (CT) used as supportive ways for diagnosis of PD. These neuroimaging modalities uses to assess changes in striatal dopamine and to detect subclinical nigral pathology. The use of the electroencephalography (EEG) and electromyography (EMG) machines are also considered as approaches for screening of neuromuscular diseases [23]. But, the interpretation of the images and signals needs highly experienced neurologists, these procedures are expensive, uncertain for this disease and they are found only in higher hospitals.

Diagnosis of PD remains challenging because most methods are dependent on knowledge and experience of physicians. The methods are prone to diagnosis errors thus, around 25% of Parkinson's patients are misdiagnosed [24].

1.2.4. Rating Scale and Stages of Parkinson's Disease

According to Parkinson's disease foundation [21] and, Hoehn and Yahr (H&Y) rating scale [25], Parkinson's disease can be classified into five stages. Physicians use H&Y is simple rating scale, found by Dr. Margaret M. Hoehn in 1967, to describe the severity of the disease based on motor symptoms of the patient. According to both rating scales, stage 1 and 2 are considered as early-level, 3 mid-level, and 4 and 5 advanced-level of PD.

Stage one is initial stage of PD, at this stage, the patient starts to show some minimal symptoms of functional impairments. Symptoms such as, tremor, changes in posture, walking and facial expressions occur only on one side of the body. At stage two, symptoms start to become bilateral and get worse. Patients start showing symptoms like, tremor, rigidity of truncal muscles and other movement problems. Third stage is mid-stage, the individual starts facing difficulties to perform daily activities. Loss of balance, falls and slowness of movements are more common in this stage. At fourth stage, the individual starts to present fully developed and severe symptoms. Standing without assistant may be possible but movement require a walker and the person needs help in daily activities. The final or fifth stage is most advanced and debilitating stage. Stiffness of legs makes standing and walking impossible for the patient with this stage. The person becomes fully dependent for any activities. At this stage the person starts experiencing hallucinations and delusions [26].

Generally, body muscles are highly affected by PD in all levels of the disease because of improper functioning of substantia nigra and basal ganglia. Movement rigidity and stiffness, joint freezing, postural instability and tremor are the most common but the degree is different based on the stage of the disease [27]. Additionally, different research findings suggested that Parkinson's disease also causes muscle weakness. Although the specific cause of muscle weakness is not known [28].

Loss of movement control and improper dexterity pattern in upper limbs such as hand and finger movements are common motor disorders in patients with Parkinson's disease [29]. Most of hand actions are done by joint movements. Thus, the effect of PD on muscles can be analyzed by measuring the activities of upper limb (arm) wrist and elbow joint muscles such as, biceps brachii, flexor carpi ulnaris, flexor carpi radialis and extensor carpi radialis brevis [30, 31, 32]. Typically, selected hand muscles, biceps brachii and flexor carpi radialis are illustrated in section 1.3.

1.3. Biceps brachii and flexor carpi radialis Muscles

Biceps brachii (BB) muscle simply known as biceps muscle, is one of the main muscles in human anatomy and the muscle emerges from the shoulder joint and its insertion is at the elbow joint. This muscle is involved in most hand movements. There are many movements that take place at the elbow and shoulder joints with which the BB play a major role. Most of elbow joint movements or actions are done by biceps brachii muscle along with brachialis and brachioradialis. Additionally, biceps with coracobrachialis and anterior deltoid muscles are responsible for shoulder joint actions [33].

Flexor carpi radialis is a muscle of the human forearm which is required for dorsiflexion and hijack. It is a combination of fast twitch fibers of hand [34]. This muscle is a relatively thin muscle located on the anterior part of the forearm. This muscle emerges on the medial epicondyle of the humerus and closes to the wrist area. Flexor carpi radials is responsible for hand and finger movements [35]. Figure 1.3 shows the anatomical positions of flexor carpi radialis and biceps muscles.

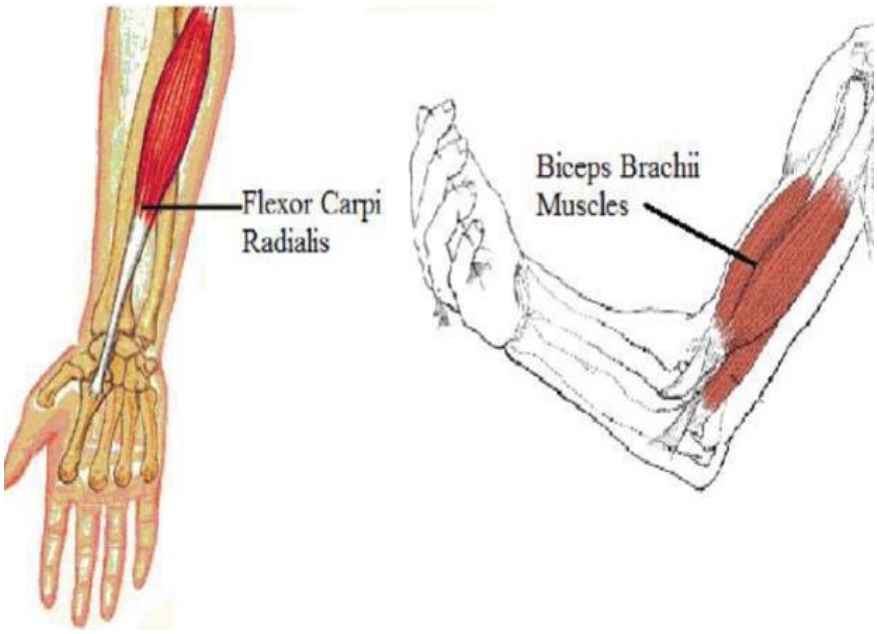


Figure 1:3: Anatomical positions of flexor carpi radialis and biceps brachii (biceps) [36]

Measuring myoelectric activities of these muscles provide important information to assess the several muscle disorders and electromyography is a common procedure for monitoring effect of

neuromuscular disorders on different muscle cells. The detail about EMG machine is described in section 1.4.

1.4. Electromyography

The term EMG was discovered in 1922 that describes myoelectric activities of muscle contractions. Contraction of muscle creates complex muscle action potential (MAP), and an electrode properly placed with respect to the muscle senses and records the MAPs as electrical waveforms called electromyograph. This can be done in two ways: either intramuscular where needle electrodes are inserted into the muscle of interest which is preferable during studying small muscles or extra muscular (surface electromyography procedures), where electrodes are attached to the skin surface above the muscle of interest which is preferable during studying large muscles. Typically, surface electromyography (sEMG) provides non-invasive way of measuring muscle activations and currently available sEMG machine technologies are portable, this makes the procedure convenient [37].

EMG activity (measured in microvolts) is linearly related to the amount of muscle contraction as well as the number of contracted muscles. In other words, the stronger the muscle contraction and the higher the number of activated muscles, the higher the recorded voltage amplitude will. Surface EMG machine incorporates; electrodes to sense myoelectric activities of muscles, analog to digital converter to convert analog signal into digital, amplifier to amplify the signal and computer system to display and analyze the recorded signals [38]. The recording procedure is illustrated in Figure 1.4.

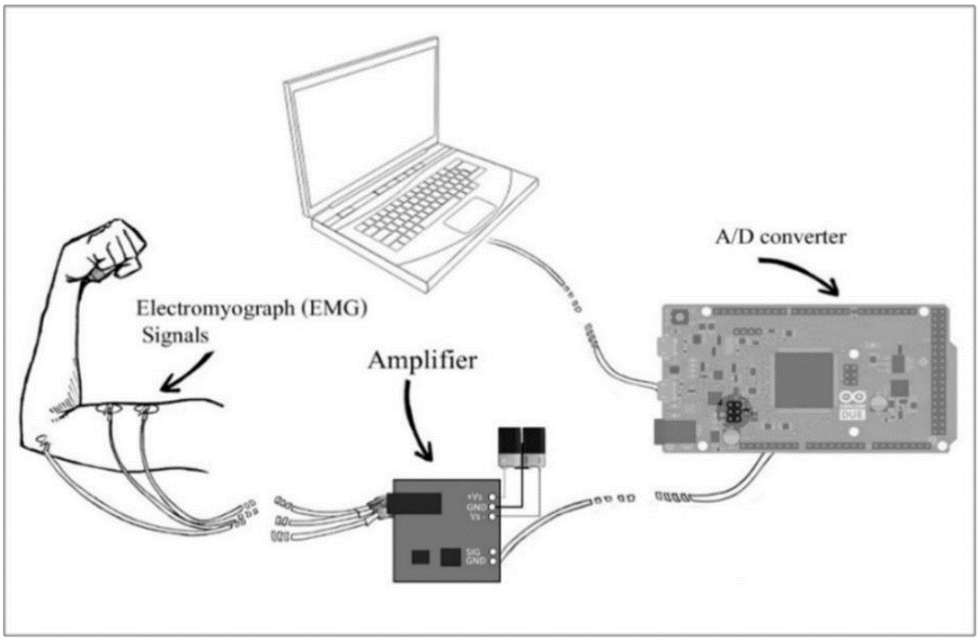


Figure 1:4: EMG signal recording procedure [38].

In the diagnosis of PD, surface EMG is informative procedure used to obtain relevant quantitative difference and can be used as biomarker for early detection of PD [39]. However, EMG signal activity pattern are strikingly irregular and complex even under normal muscle tones [30]. This makes interpretation of the signal difference in normal and pathological conditions like, PD and other neuromuscular disorders difficult for medical experts.

1.5. Problem Statement

Currently, several people are suffering with Parkinson's disease worldwide. Because of lack of awareness, the disease remains unrecognized and untreated in developing countries like Ethiopia.

People living with Parkinson 's disease are exposed to multilayered challenges such as, psychological, social and economic problems in their day to day life. These challenges are prolonged throughout their life because of uncurable nature of the disease. Apart from this, there is no standardized way of diagnosing this disease. The common approach to diagnose this disease is through clinical assessment of the patients' sign and symptoms. However, the disease starts showing symptoms at late stage and this may restrict the efficiency of the treatment.

In addition to this, the current diagnosis method is highly dependent on the knowledge and experience of the experts which may lead to subjective diagnosis errors and cause other medical complications on the patient. The presence of few neurologist experts and limited resource, especially in the developing world, makes the challenge severe. To solve this problem, several computer aided systems have been developed. However, most literatures are found to have their own gaps, such as, accuracy, use of small dataset and computational expensiveness.

Therefore, a system, which solves such a crucial problem, will be highly needed, especially for low resource settings. In this research, an automatic system has been proposed to identify PD from EMG signals at the early stage.

1.6. Objectives

1.6.1. General objective

The main objective of this study is to design an automatic diagnosis of PD using EMG signals from different hand movements.

1.6.2. Specific Objectives

- ✓ To detect PD automatically from EMG signal.
- ✓ To determine the severity of the disease from EMG signal.
- ✓ To evaluate the performance of the system.
- ✓ To compare the performance of the trained multi class SVM model with other machine learning approaches.
- ✓ To determine the effective hand movement for PD detection.

1.7. Significance of the study

This study mainly focused on designing an automated system that can diagnose PD at early stage with optimized accuracy and reliability. It will also provide a comfortable diagnosis system for patients as well as for physicians. Detecting the disease at early stage have positive medication outcome.

This system will minimize possible errors and decrease the work load of neurologists, by making the diagnoses procedure more efficient. It will give insight for researchers about which hand movement and muscle signal is effective for diagnosing PD.

This will give a direction to EMG machine manufacturers during designing stage to integrate such algorithms to the system which simplifies and supports the neurologists in the process of diagnosing the disease.

Chapter Two

2. Automated Diagnosis Methods of PD Using EMG signal

For automated Diagnosis of PD using EMG signals, signal recording, preprocessing, feature extraction and classification of the EMG signals are followed as main tasks. The research community has been interested in the development of computer aided systems in this area to introduce a better way of signal analysis and classification. Some of recent findings on Parkinson's disease detection using EMG signals are discussed in this section.

2.1. Muscles for EMG Signal Recording

PD affects several body muscles that causes functional impairments. Before signal recording, selecting proper muscle should be considered to get relevant information for intended purpose. Some previous studies used EMG to quantify myoelectric activities and better understand the effect of Parkinson's disease on performance of upper and lower limb muscles while the subjects are performing different tasks [30-32, 40-43] .

The difference between healthy individuals and patients with PD can be analyzed using myoelectric signals recordings from tibialis anterior, gastrocnemius medialis and gastrocnemius lateralis while the subjects performing different muscle tasks. Pertinent information are found from these muscles because these muscles take the major role for lower limb activities [40, 41, 42].

Other types of muscle that are important for PD and other neuromuscular disorder detection are Upper limb muscles. These muscles give relevant information to detect effect of Parkinson's disease on muscles with activities that require their efforts such as, wrist flexion, elbow flexion and carrying objects [30, 31]. Biceps brachii muscle is the main muscle that uses to characterize the morphology and pattern of the signal the elbow flexion with different loading conditions such as, without and with different weights [30, 31]. EMG signals acquired from extensor and flexor muscles like, extensor carpi radialis brevis and flexor carpi ulnaris muscles uses to quantify the intensity of tremor in patients with PD and to differentiate the disease from essential tremor. The difference between healthy individuals and patients with Parkinson's disease is significant while

the subjects are performing different actions such as joint movements, resting, holding, carrying weight [32, 43].

2.2. Preprocessing

Surface EMG signal is prone for high frequency and low frequency component noises that contaminate the quality of signals. It is exposed for low frequency artifacts such as, motion artifact, power line interference (60Hz or 50Hz) and thermal noises of the amplification system itself [44]. And the inherent noises of the equipment may cause high frequency noises up to several thousand Hertz on the signal [45]. To eliminate these effects several studies come up with different preprocessing and noise removal approaches.

High pass filters, low pass filters [43, 46] and band pass filters [31, 32] with different cut-off frequencies are common and easy approaches to eliminate the effects of high and low frequency noises from EMG signals. The lower corner frequencies used for high pass filters are 1Hz -20Hz and upper corner frequencies used for low pass filters are 70Hz-500Hz for above mentioned studies. Applying moving average filter is also important to remove low frequency and high frequency components [47].

Discrete Wavelet Transform [47] and Haar mother wavelet analysis [43] can be applied to further increase the quality and the visibility of the tremor bursts in EMG signals. Additionally, rectification, normalization and segmentation are also further preprocessing that uses to enhance the quality and quantity of the signal [32].

The cut-off frequencies of high pass and low pass filters should be selected carefully. Improper selection of cut-off frequencies may lead to wrong filtration and may cause missing the relevant part of signals. The recommended usable energy band of EMG signal is 20Hz-500Hz [48]. So, this should be taken in to consideration during EMG signal denoising.

2.3. Features Extraction

Feature extraction is a process of retrieving important information from a raw signal for analysis and grouping of classes [49]. Determination of the feature type depends on the desired information needed from the data. Features can be classified as time domain and frequency domain features. Several studies have been done to determine important features for PD detection.

Time domain features such as, variance, skewness, kurtosis, root mean square, crossing rate and recurrence rate showed separability property for Parkinson's disease [40] and to classify the intensity of tremor in PD patients [32]. These features along with entropy and energy can be used to differentiate PD with other related disease like essential tremor [43]. Other time domain features that have potential to differentiate patients with Parkinson's disease [41] and to classify the severity [42] of PD are Integrated EMG, Log detector, mean absolute value, waveform length, and mean power.

Complexity and regularity determination features are also important to determine the effect of Parkinson's disease on muscles. These features are under time domain features and the difference between the signal of PD patients and healthy subjects can be analyzed by extracting non-linear features such as, sample entropy (SampEn), correlation dimension (CD), percent of determination (DET%), recurrence rate (REC%) and RMS of EMG signals [30]. Some different feature vectors such as, histogram values, crossing rate expansion values, concatenated CR expansion are presented in [31]. Frequency features such as, dominant frequency, mean frequency, median frequency and total power are also relevant to analyze the difference between signals from healthy individuals and patients with PD [40]. Features play an important role on separability of categories and classification performance of classifiers. So, the feature type and the information that retrieved from that feature should be known before feature extraction step.

2.4. Classification

The EMG signal features in different domains are fed into classifier to determine the differences between different categories. Numerous approaches are deployed to classify EMG data for PD detection such as, artificial neural networks (ANN), fuzzy logic (FL), multilayer perceptron (MLP), support vector machines (SVM), linear discriminant analysis (LDA), and K-nearest neighbor (KNN).

Linear Discriminant Analysis (LDA) and K-means clustering approaches showed limited classification performance but K-means clustering better to differentiate patients with Parkinson's disease and healthy individuals with high re-substitution and classification errors [47].

Research findings showed the use of binary support vector machine (SVM) for discriminating between healthy subjects and patients with PD [40], additionally SVM can differentiate PD from

other related diseases like essential tremor [43] using statistical features of EMG signals. However, the classification performance is better to differentiate Parkinson's disease from healthy subjects [40] than differentiating it from other related diseases [43] that have similar features.

Multilayer artificial neural network (ANN) model is applicable to perform binary classification for EMG signal time domain and frequency domain features of healthy individuals and patients with Parkinson's disease [41] showed comparable performance with binary SVM. Multilayer ANN is also capable to perform separation on non-linear features estimated from EEG and EMG signals such as, Approximate Entropy, Correlation Dimension, Fractal dimension and Largest Lyapunov Exponent [50]. To solve the pitfall of binary classification ANN model can be trained to classify EMG signal features to the disease severity level classification [42] but with limited accuracy. Artificial neural network and fuzzy logic also combined to give better separation accuracy between neuromuscular disorders into normal, myopathy, Huntington disease(HD) and PD using spectral and temporal features extracted from EMG signals of patients from each muscular disorder [51]. Despite this combined module give improved classification accuracy, sensitivity of the model for PD (73.3%) is not satisfactory.

Fast Orthogonal Search Algorithm classifier is also applicable to classify the stage of tremor intensity using EMG signal features [32]. Parallel structure model using K-fold approach gives very accurate classification performance than single structure model using K-fold approach. Despite improved classification accuracies maintained in such studies, considering only tremor intensity of the signal is not sufficient to grade the severity level of PD. Considering only tremor intensity may lead to false diagnosis result to those diseases have similar characteristic with PD such as, essential tremor, orthostatic tremor, cerebellar disease, peripheral neuropathy and alcohol withdrawal, this may reduce the reliability of the model. In addition to this using parallel structure model using K-fold approach uses five separate models to classify the method is computationally complex and expensive. Model selection should be done based on intended purpose for successful classification or to map the extracted features into specific classes. The general summery of some related works is illustrated in Table 2.1.

Table 2-1: Related work summery

Paper	Data	Preprocessing	Features	Model	Classes	Acc.
[40]	5 HC 5 PD	-	Skew, Var, Kurt, RMS, MNF, MedF, Dominant F, Total power	SVM	Binary	91%
[41]	8 HC 13 PD	-	Mav, log detector, WL, Skew, Var, Kurt, RMS, MNF, MedF, IEMG Total power, mean power	ANN	Binary	88.4%
[42]	8 HC 13 PD	-	Mav, log detector, WL, Skew, Var, Kurt, RMS, MNF, MedF, IEMG Total power, mean power	ANN	normal, possible, probable & definite	71%
[43]	11 ET 13 PD	Low pass 70 Hz High pass 20 Hz Haar-wavelet analysis	Mean, standard deviation, skew, kurt, entropy, energy & RMS	SVM	ET& PD	83%
[32]	12 PD	band pass filter 10Hz-500 Hz	RMS, mean variable, kurt, crossing rate and recurrence rate	Orthogonal Search Algorithm	Tremor intensity levels	88.57% & 99.25%
[51]	260 signals (online)	-	MNF, mean power, IEMG, total power MAV, SSI, VAR, wavelet mean, wavelet energy, wavelet power, wavelet SD	ANN and fuzzy logic combined	normal, myopathy, HD and PD	Overall=96.47% For PD 73.33%
[47]	15 PD 15 HC	Moving average filtering & WPA	Tabulated frequency domain powers	LDA & k-means clustering	PD & Normal	Isloss = 21.11% & loss = 18.9%.

2.5. Gap Analysis of Related Works

Though a variety of studies were done on EMG signal processing to analyze the effect of Parkinson's disease on a signal from muscle, most of the recent works focused on feature

extraction without denoising the signal [40, 41, 42 51]. Because EMG is sensitive to many artifacts, noise removal is an intensive step in EMG signal processing. The frequency band of important EMG signal is 20Hz-500Hz [48]. Following improper pre-processing or denoising steps [43] has been observed in some literatures. Using limited number of features is another gap that is identified in some literatures [40, 51].

Several automated systems are developed to minimize human intervention and errors related with human knowledge gaps, still there are many problems that remained unsolved. Despite understanding the disease stage and progress is critical for the patients and physicians, most of the recent works focused on designing binary classification systems [40, 41, 43, 47, 51] and small number of datasets used. And some of the studies done in this area [42, 43, 47, 51] achieved limited accuracies (71%-91%) even for binary classifications, which is critical in medical areas.

Chapter Three

3. Materials and Methods

In this section, different procedures that are adopted for the achievement of this research will be briefly demonstrated including the experimental setup, materials, and methodology of the research.

3.1. General methodology

The following sections demonstrates the workflow starting from the data collection to the classification of the signal to normal, early, moderate and advanced levels of PD. First the signal was recorded from patients with PD from each level and age-matched healthy control subject then, it was processed and analyzed in steps, preprocessing, feature extraction, feature selection and feature classification using MATLAB 2018. Figure: 3:1 shows the overall block diagram of the procedure used for this study.

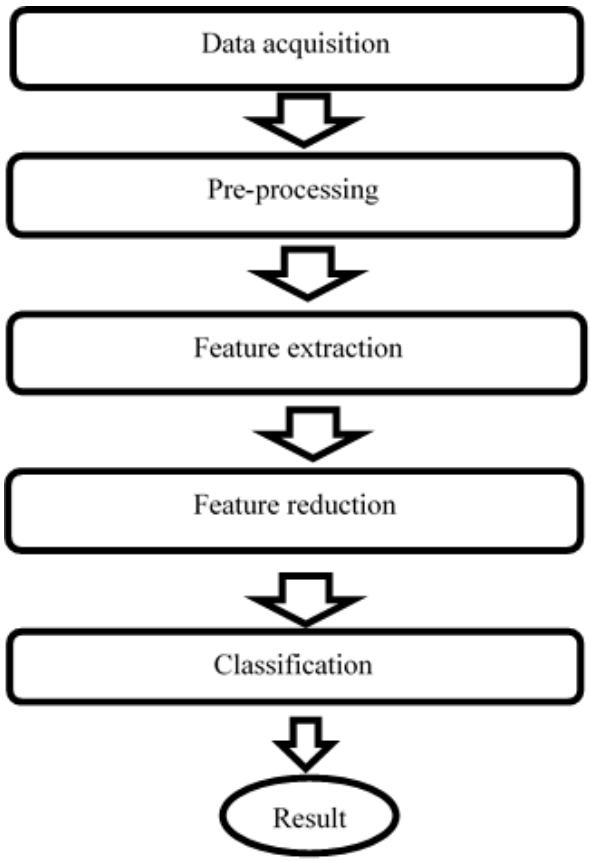


Figure 3:1: Block diagram of the procedure the study

3.2. Data Collection

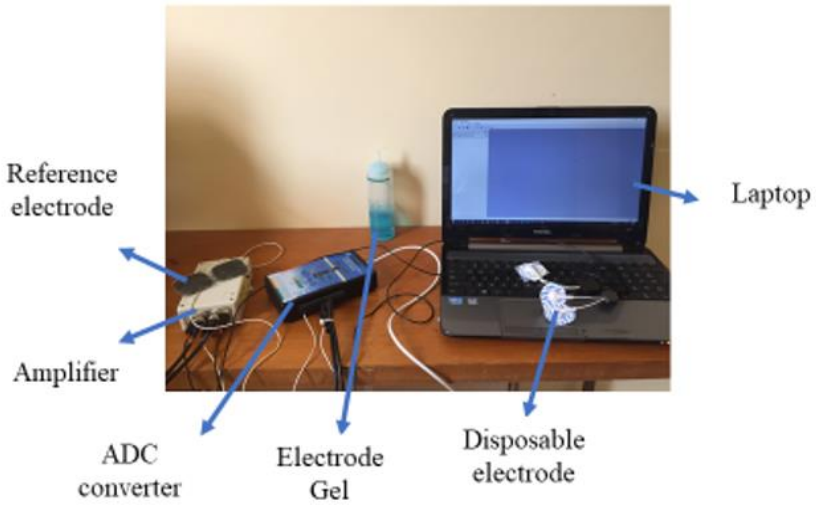
3.2.1. Subjects

10 healthy subjects (6 male and 4 female, age between 38 and 51), and 15 patients (13 male and 2 female, with age 40-54) with PD 5 for early-level, 5 for mid-level and 5 for advanced-level of the disease were selected to record their muscle activity while performing selected hand movements: elbow flexed with 90 degree, elbow flexed by 90 degree with 1 kg load, wrist pronation and touching shoulder using SCU-7 EMG system. It is surface EMG machine that incorporates reusable and disposable electrodes, amplifier, analog to digital converter and computer system. (some specification of the system are included in appendix part). The age group is selected taking into consideration the age of the patients with PD, and healthy subjects with no known history of neuromuscular diseases were selected. All patients tested were on medication.

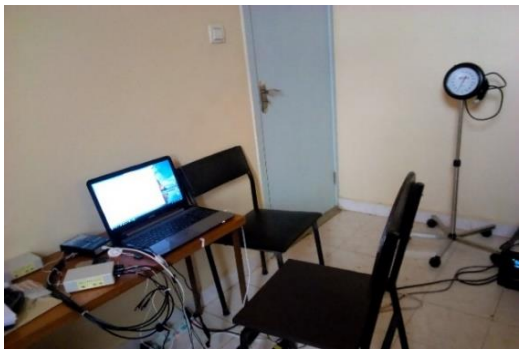
3.2.2. Recording Set-up and Preparation

Before starting the recording procedure, the clinical setup was prepared at Jimma University Medical Center, Neurologic Clinic. EMG machine (SCU-7 EMG system) with disposable surface electrodes were used. Before electrodes attachment to the skin, the skin of the subject was cleaned using sandpaper to remove dead cells and some small hairs on skin. This procedure helps to decrease impedance mismatching between electrodes and skin. An electrode gel was applied on the skin to remove motion artifacts and to increase signal to noise ratio (SNR).

The electrodes were attached on the skin over flexor Carpi radialis and Biceps muscles, as shown in figure 6: (e). These muscle types are selected based on their role on upper limb movements and because these muscles are situated most superficially. The reference electrode was attached on the skin over elbow. In general, the electrode placement followed the SENAIM guidelines. The procedure followed in the preparation is illustrated in Figure 3:2 (a-e).



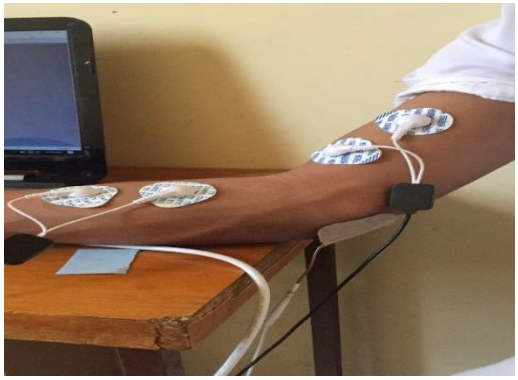
a



b

c

d



e

Figure 3:2: Preparation procedures: (a & b) clinical setup, (c) removing dead cells, (d) applying jel and (e) electrode placement

3.2.3. EMG Signal Recording

EMG signal was recorded using SCU-7 EMG system with hand movements of the right hand for ten trials. During signal recording two channels with two electrodes each were used but the signals were recorded from one channel. This means, during wrist pronation signals were recorded from the channel (two electrodes) attached skin over flexor carpi radialis muscle and during other three hand movements (elbow flexion with 90-degree, elbow flexed by 90 degree with 1 kg load and touching the shoulder specifically, the acromion part of the acromioclavicular joint of shoulder) signals were recorded from channel (two electrodes) attached over the skin of biceps muscle. Figure 3:3 (a-d) shows mentioned hand movements for this study.

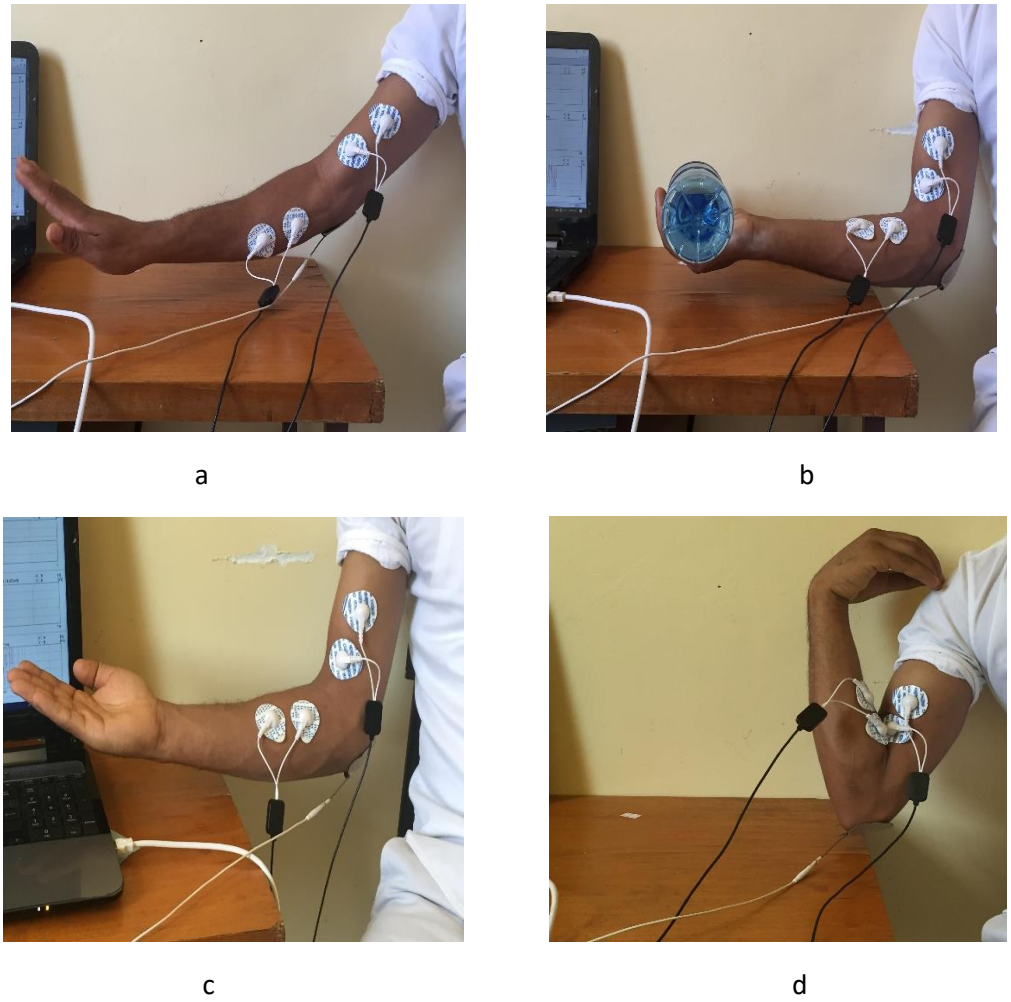


Figure 3:3: Studied hand movements: (a) wrist pronation, (b) elbow flexion by 90 degree with 1 kg load, (c) elbow flexion by 90 degree and (d) touching shoulder.

3.2.4. Data Augmentation

Data augmentation is the process of generating new samples by transforming the available data and increase the number of data samples in order to improve the accuracy and robustness of classifiers. Geometric transformation (rotation, translation and shear mapping) and noise addition are two types of data augmentation approaches in image processing. Since time series signals are one dimensional, geometric transformation are not suitable for them [52]. Hence, noise addition has been selected as data augmentation procedure for this study. In addition to this, EMG signal is highly random and non-stationary signal. If random noise is added, it will increase the randomness of the signal [53]. By taking this in to consideration, white gaussian noise with signal to noise ratio 100 and 90 was added to augment the signal.

3.3. Pre-Processing

The collected EMG signals were filtered using iirnotch filter to reject 50Hz power line interference. 6^{th} order Butterworth low pass and high pass filters were applied to reject high frequency noises above 500 Hz and low-frequency noises below 20 Hz because recommended usable frequency band of surface EMG is between 20Hz and 500Hz [48] after denoising of the signal fast fourier transform was applied to the time-series signal in order to extract frequency domain features.

3.4. Feature Extraction

Feature extraction is the alteration of the raw signal data into an important information highlighting the relevance of data [49]. For this study Fourteen different features have been extracted from the preprocessed EMG signals. The detail of extracted features is discussed below.

3.4.1. Average Amplitude Change

Average amplitude change (AAC) feature is easily implementable and it is expressed as the mean or average value of the difference between two consecutive waveform length of EMG signals [54]. It can be defined as:

$$\text{AAC} = \frac{1}{N}\sum_{i=1}^{N-1} |X_{i+1} - X_i| \quad \text{......} \quad (3.1)$$

where, X=input signal & N = number of signals

3.4.2. Entropy

Entropy is a measure of complexity, irregularity and randomness of signals. It is important feature for detection and diagnosis of different diseases using biological signals [55]. The entropy used for this study is based on Shannon and spectral entropies [56] which can be mathematically calculated as:

$$\text{Entropy} = \sum_{i=1}^{N}(P_i \log_2 P_i) \quad \text{(3.2)}$$

where, P = The probability distribution of the signal & N = number of signals

3.4.3. Integrated EMG (IEMG)

Integrated EMG feature is related to the information of EMG signal sequence firing point and describes the area under the curve of the rectified EMG signal. IEMG can be simplified and expressed as the summation of the absolute values of the EMG amplitude [57]. Mathematically, this can be expressed as:

$$IEMG = \sum_{i=1}^{N} |X_i| \quad \text{(3.3)}$$

where, X=input signal & N = number of signals

3.4.4. Kurtosis

Kurtosis is a time domain-based feature. It is known as a statistical method that used to describe the distribution and characteristic that identifies the tendency of peak data [58]. Kurtosis level data is determined by comparing the peak of the curve inclination data, distribution and standard curve. It is the descriptor of the shape of a probability distribution and can be defined as equation (3.4)

$$Kurtosis(x) = \frac{\mathrm{E}\,[(\mathrm{X}-\mu)^4]}{(\mathrm{E}\,[(\mathrm{X}-\mu)^2])^2} \quad \text{(3.4)}$$

where, X=input signal, μ= mean & E = expected value

3.4.5. Maximum Fractal Length

Maximum fractal length (MFL) is a recently used and modified form EMG feature. MFL is very important feature to measure the activation of low-level muscle contraction [59] and mathematically it can be calculated as:

$$\text{MFL} = \sqrt{(\sum_{i=1}^{N-1}(X_{i+1} - X_i))^2} \quad \text{(3.5)}$$

where, X=input signal & N = number of signals

3.4.6. Mean Absolute Value

Mean absolute value (MAV) is one of the most commonly used EMG features. It is a modified form of IEMG feature. MAV feature is an average of absolute value of the EMG signal amplitude in a segment [60], which can be defined as:

$$MAV = \frac{1}{N}\sum_{i=1}^{N} |X_i| \quad \text{(3.6)}$$

where, X=input signal & N = number of signals

3.4.7. Simple Square Integral

It describes the total energy of the EMG signal. It can be expressed as a summation of square values of the EMG signal amplitude and this parameter is generally defined as an energy index [61], which can be expressed as:

$$SSI = \sum_{i=1}^{N} {X_i}^2 \quad \text{(3.7)}$$

where, X=input signal & N = number of signals

3.4.8. Waveform Length

Waveform length describes cumulative length of the EMG waveform over the time segment [59]. The other name of waveform length feature in some literatures is wavelength (WAVE). It can be defined by equation (3.8).

$$\text{WL} = \sum_{i=1}^{N-1} |X_{i+1} - X_i| \quad \text{(3.8)}$$

where, X=input signal & N = number of signals

3.4.9. Root Mean Square

Root mean square (RMS) feature represents the square root of the mean power of EMG signal. This feature relates to non-fatigue muscle contraction and constant force [61]. And it can be expressed as in equation (3.9)

$$RMS = \sqrt{\frac{1}{N}\sum_{i=1}^{N} {X_i}^2} \quad \text{(3.9)}$$

where, X=input signal and N = number of signals.

3.4.10. Skewness

Skewness is one of the EMG signal features and categorized under time-domain functions defined as the inclination distribution data. It is a measure of the symmetry of the probability distribution of a real-valued random variable about its mean. The data is said to have a normal distribution when the location of the average value, the median value, and data mode on a line in the curve if these values are not located in one line in the curve occurs the skewness [62]. Equation (3.10) was used to define the skewness of the signal.

$$\text{Skewness}(x) = \frac{\text{E}\,[(\text{X}-\mu)^3]}{(\text{E}\,[(\text{X}-\mu)^2])^{3/2}} \quad \text{(3.10)}$$

where, X = signals, E = expected value & μ = mean

3.4.11. Variance

It is the power index of EMG signal. It is also an average of square values of the deviation of that variable; however, mean value of EMG signal is close to zero. Hence, the variance of the EMG signal can be defined as

$$\text{VAR} = \frac{1}{N-1}\sum_{i=0}^{N} {X_i}^2 \quad \text{(3.11)}$$

where, X=input signal & N = number of signals

3.4.12. Zero Crossing

Zero crossing (ZC) is a measure of times that amplitude values of the EMG signal cross zero amplitude level. It measures the frequency shift and shows the number of signal sign varying [61]. Equation (3.12) define zero crossing feature.

$$ZC = \sum_{i=1}^{N-1}[sgn(-X_i * X_{i+1}) \cap |X_i - X_{i+1}| \geq threshold] \quad \text{..... (3.12)}$$

$$Sgn(X) = \begin{cases} 1, & \text{if } X \geq threshold \\ 0, & \text{otherwise} \end{cases}$$

where, X=input signal & N = number of signals

3.4.13. Mean Frequency

As Frequency defines the number of occurrences of a repeating event per unit time; mean frequency is the average value of the input signal's frequency. This feature estimates the mean frequency of the signal in a specified time length [20] and for this study, Fast Fourier transform was applied before extracting this and one more frequency domain feature. Mean frequency can be expressed using the equation (3.13).

$$\text{Mean freq} = \frac{\sum_{i=1}^{N} X_i}{N} \quad \text{..... (3.13)}$$

where, X=input signal & N = number of signals

3.4.14. Median Frequency

Median frequency (MDF) is a value of frequency that separate the higher half frequency of the input signal from the lower half [60]. And it can be mathematically defined as:

$$MDF = \frac{1}{2}\sum_{i=1}^{N} X_i \quad \text{..... (3.14)}$$

where, X=input signal & N = number of signals

3.5. Feature Reduction

After extracting the fourteen features, misleading features have to be reduced and only significant features have to be retained for classification. This reduces the computational cost and makes the algorithm time efficient. The final algorithm should have a minimum number of features and should have maximum accuracy. So, features are ranked and only the best features are used for classification using ReliefF algorithm [63]. The selected features have the potential to discriminate between the four classes of signals namely; normal, early, moderate and advanced levels of the

disease. This is done in order to reduce the dimensionality, redundancy and computational load. The features have been reduced using ReliefF algorithm.

Nine best features are selected out of all the extracted features which can do classification with higher accuracy. There are various methods for feature reduction process. For this study ReliefF algorithm was applied. This algorithm finds the weights of features in the case of multiclass categorical variable. The algorithm penalizes the features that give different values to neighbors of the same class, and rewards features that give different values to neighbors of different classes. ReliefF algorithm first set the value of weight of features to zero then the weight increase or decrease based on their contribution [64]. The weight can be calculated as

If X_r and X_q are in the same class,

$$W_j^i = W_j^{i-1} - \frac{\Delta_j(X_r,\ X_q)}{m} \quad \text{(3.15)}$$

If X_r and X_q are in different classes,

$$W_j^i = W_j^{i-1} + \frac{s_{lq}}{1-s_{lr}} . \frac{\Delta_j(X_r,\ X_q)}{m} . d_{rq} \quad \text{(3.16)}$$

where,

W_j^i is the weight of the feature at the i^{th} iteration.

$s_{lq}\ and\ s_{lr}$ are the prior probabilities of the classes to which $X_q\ and\ X_r$ belong respectively.

$\Delta_j(X_r,\ X_q)$ is the difference in the value of the feature between observations $X_r\ and\ X_q$

m is the number of iterations specified.

d_{rq} is a distance function.

ReliefF algorithm is selected for this study because it is simple to implement, effective and widely used approach to feature rank [63, 65]. More importantly ReliefF algorithm is not limited for to class problems [66].

3.6. Classification of the Signal using Multi SVM Model

After feature extraction and reduction, classification was followed. Mainly multiclass Support vector machine (SVM) was used to classify the feature set in to target classes. Training feature sets were given to the classifier for training then test feature sets were given for the trained classifier for classification during the testing phase.

SVM is categorized under supervised machine learning algorithms and main objective of the first designed SVM was to perform binary classification problems. But in real world, problems require multiclass classification. Consequently, modifying this system into multiclass SVM was needed. Multi-class SVM problems are commonly decomposed into a series of binary classifications [67].

Multi-class SVM is an extension of binary SVM. the idea of using a hyperplane to separate the data into binary classification sounds well when there are only two classes or targets. For Multiclass classification, one-vs-all approach [68] is commonly used method that constructs by fitting one classifier for one class. Each category is split out and all of the other categories are merged together, and the class which classifies the test data with greatest margin is selected. It categorizes an *m* class targets into *m* binary problems. The learning step of the classifiers is done by the whole training data, considering the patterns from the particular class as positives and all other groups as negatives. Figure 3:4 shows how one versus all approach construct four binary classifiers for four class problem.

In this research, Multi-SVM was selected because the Parkinson's disease needs multiclass classification, it is easy to train, more straightforward to implement, it is a non-linear classifier and it is efficient classifier for small number of datasets.

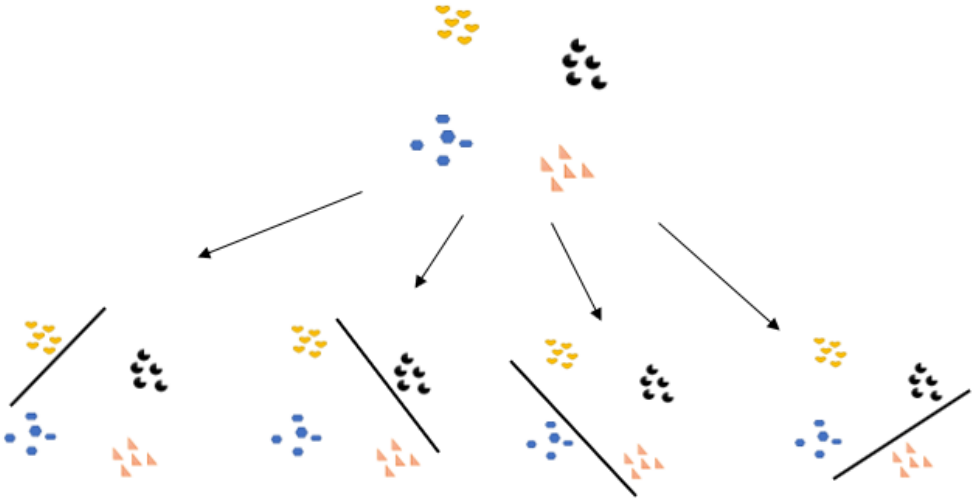


Figure 3:4: One versus all approach for four class classification

The classification is done using the set of selected feature set. The feature set is divided into a training set consisting of 480 signals which is 80% of the total signal data for one hand movement and a test set consisting of 120 signals which is 20%. The training dataset was used to train the classifier and the test samples are given as input for the classifier to test the performance of the classifier. Signals from different hand movements are treated as separate datasets and separate SVM classifiers were trained. In addition to the main model (multi-SVM), multi class KNN and LDA models were trained with the same dataset to compare the result with the main model. The general methodology of the algorithm is shown diagrammatically in the figure 3:5.

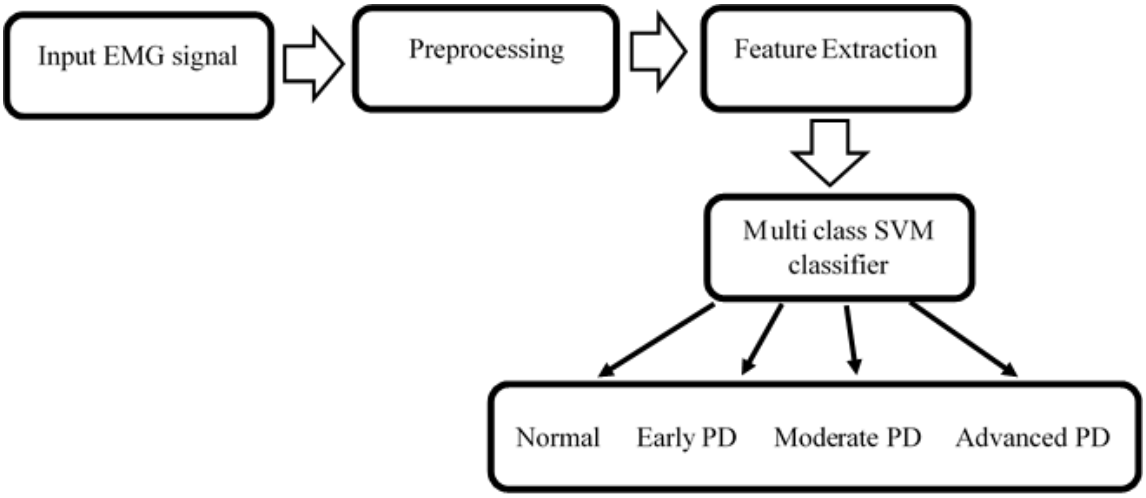


Figure 3:5: Block diagram of the developed system

3.7. Graphical User Interface (GUI)

Graphical user interface (GUI) is a model that allows the user to interact or communicate with a system easily without needs to know any programming language and memorize the commands. The user should only interact to graphical user interface elements such as buttons, icons and cursors to make smooth communication with the system.

For this study, after training the model to classify four classes, GUI was developed to make the system easy to use, implementable and user friendly. The developed GUI incorporates 'load button' which is used to load the signal, 'hand movement selection menu' that enables the user to select the hand gesture to which the signal recorded from, and 'result button' that makes the system to process and display the result. The 'exit' button is used to stop the program. Figure 3:6 shows the general layout of the developed GUI for this system.

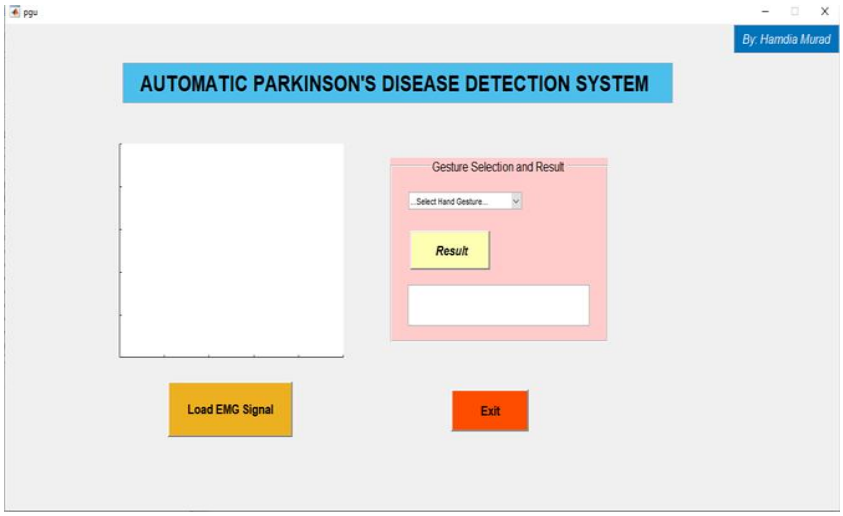


Figure 3:6: GUI model of the system

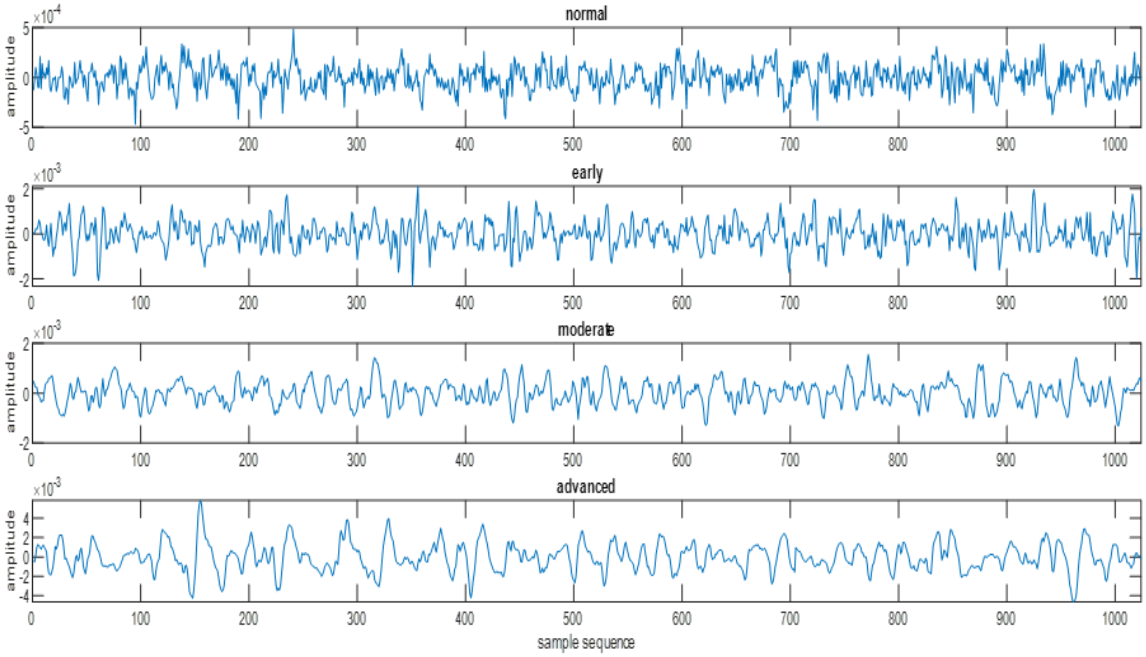
Chapter Four

4. Results and Discussion

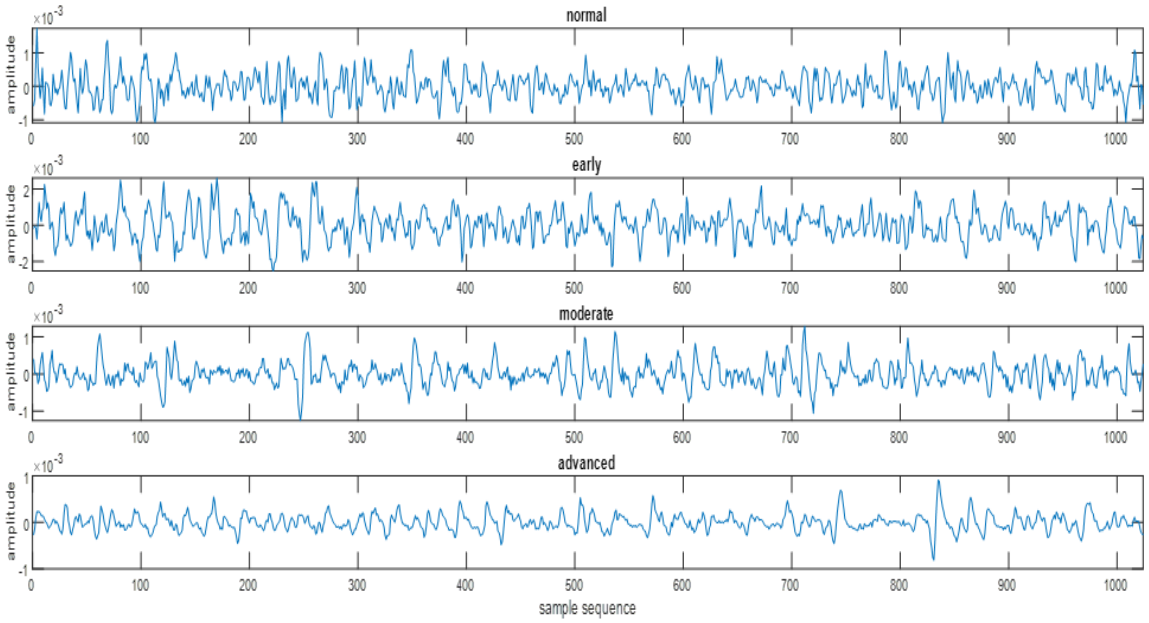
4.1. Results

4.1.1. EMG Signal Data

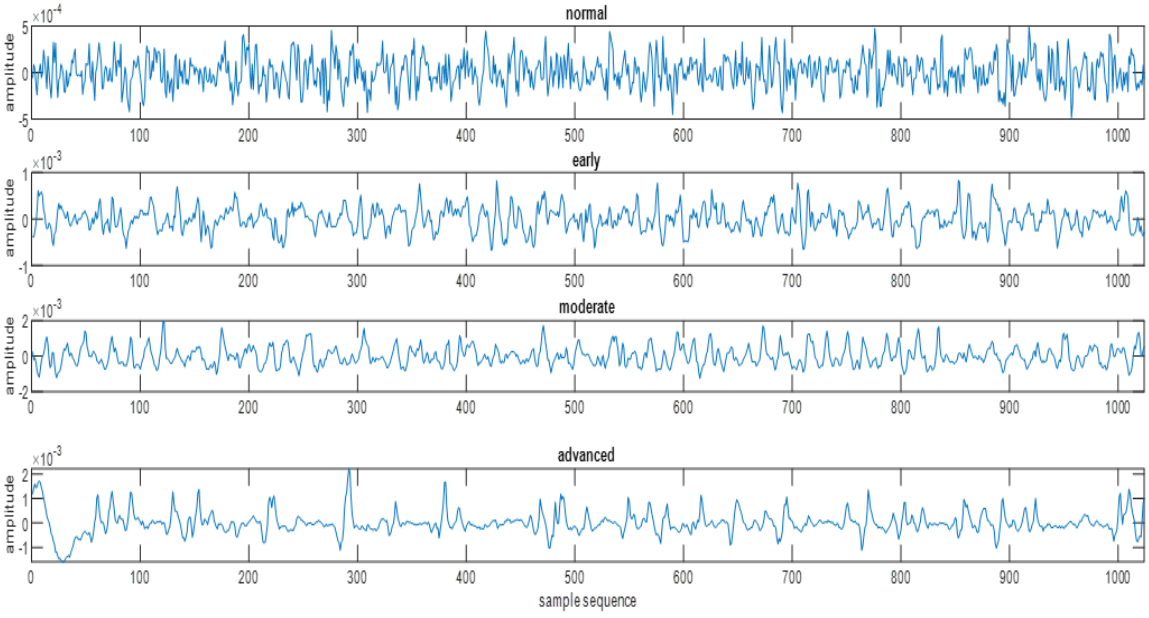
A total of 1000 myoelectric signal data were recorded from all subjects while the subjects performing selected hand movements. For each hand movement 250 EMG signal data that include all four classes of the disease were recorded. A sample of EMG signal collected from each class for selected hand movements is presented in Figures 4:1 (a-d).



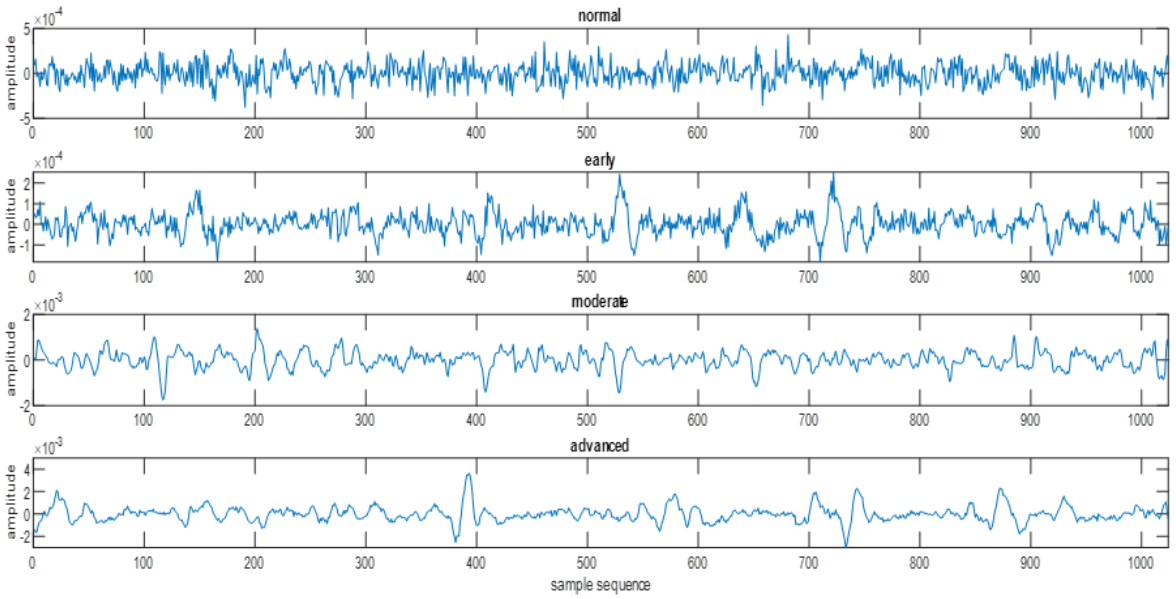
a


normal
early
moderate
advanced
amplitude
sample sequence

b


normal
early
moderate
advanced
amplitude
sample sequence

c



d

Figure 4:1: Sample signal. (a) for elbow flexed with 90-degree, (b) for elbow flexed by 90 degree with 1 kg load, (c) wrist pronation and (d) touching the shoulder (All are for Normal, Early, Moderate and advanced level from top to bottom respectively)

4.1.2. Data Augmentation

By taking the randomness of EMG signal in to consideration, white gaussian noise with signal to noise ratio of 100 and 90 was added to augment the signal. After noise addition, the signal maintains it's characteristics. Figure: 4:2 shows the sample of augmented signals with mentioned SNR. After data augmentation a total of 2400 raw signals (600 for each hand movement) has been obtained. The detail of the collected dataset is illustrated in Table 4-1.

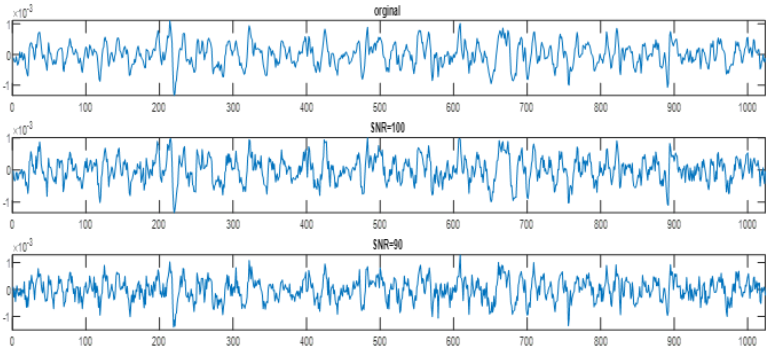


Figure 4:2: Augmented signals.

Table 4-1: Description of the dataset

	Wrist pronation	Touching shoulder	Elbow flexion without load	Elbow flexion with load	Total
Normal	150	150	150	150	600
Early	150	150	150	150	600
Moderate	150	150	150	150	600
Advanced	150	150	150	150	600
Total	600	600	600	600	2400

4.1.3. Preprocessing

4.1.3.1. De-noising

Before extraction of important features, the recorded EMG signals were denoised thoroughly. iirnotch filter was used to remove 50Hz power line interference followed by Low-frequency trends were removed from EMG using high pass filter (sixth order Butterworth), the high-pass cut-off frequency was 20 Hz. High frequency artifacts were removed using low-pass filter (sixth order Butterworth) with cutoff frequency 500Hz. The effect of denoising may not be clearly visible but it shows significant difference on classification performance of the model. Figure 4:3 shows sample of original and denoised signals.

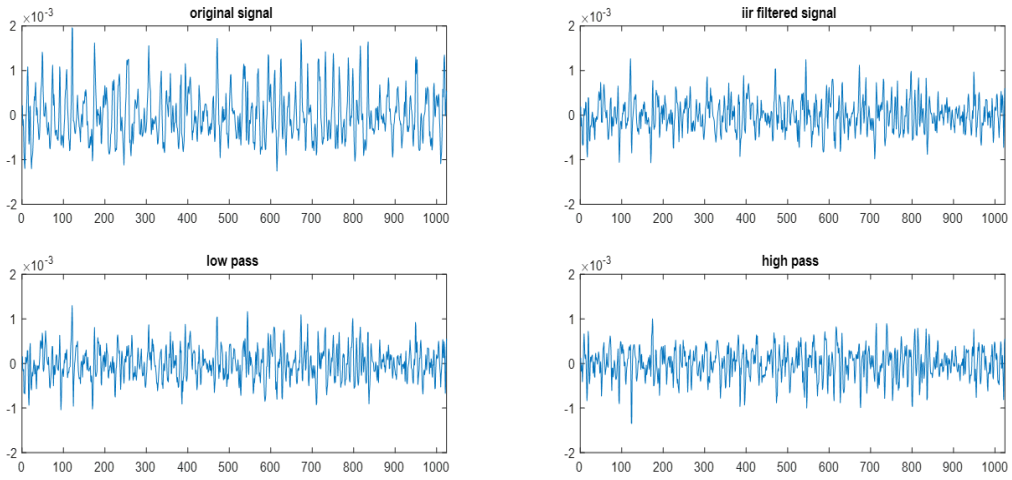


Figure 4:3: Denoising procedure.

4.1.3.2. Fast Fourier Transform (FFT)

After denoising Fast Fourier transform was applied to extract frequency domain features. Time domain features were directly extracted from time series signal. Figure 4:4 shows sample of FFT applied signals that recorded during the subjects flexing elbow with 90 degree for all classes.

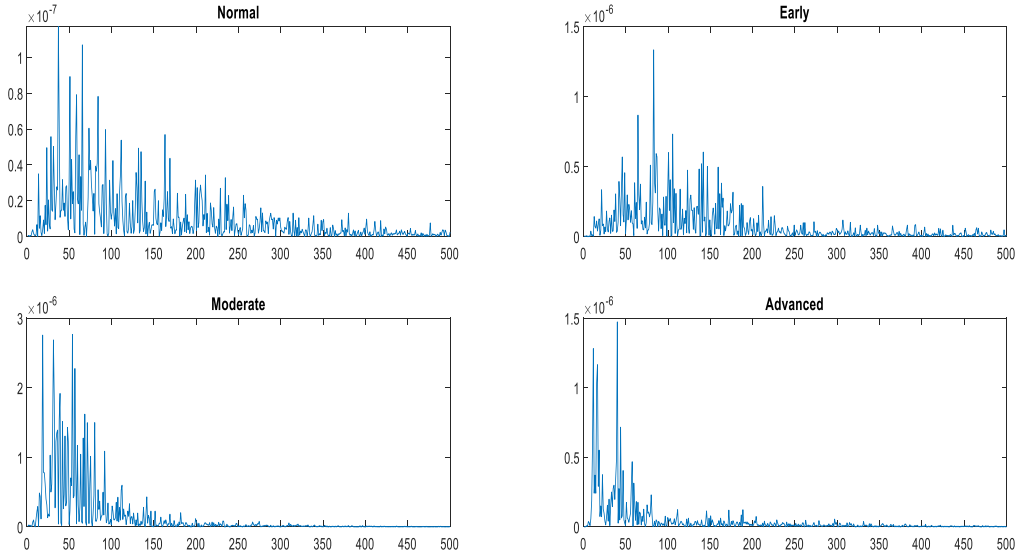
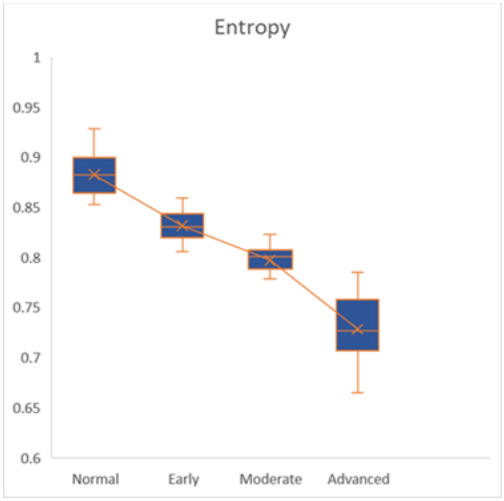


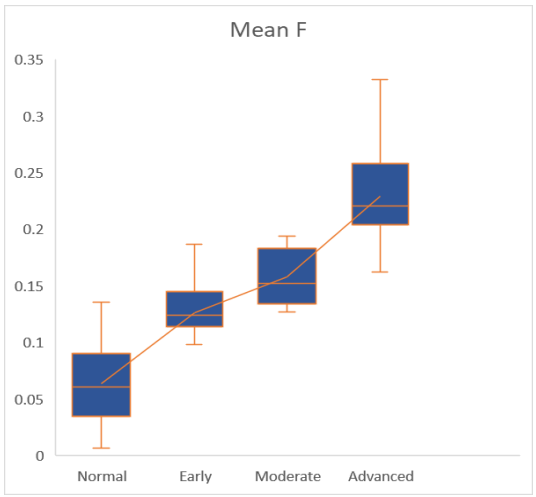
Figure 4:4: FFT of 90-degree signal: For healthy, early, moderate and advanced levels

4.1.4. Feature Extraction

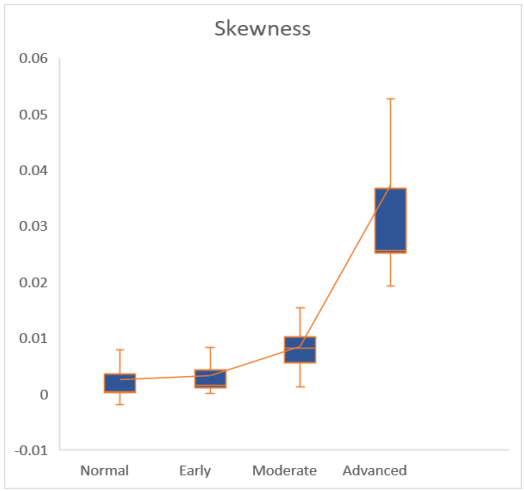
Different features have been extracted in time domain and frequency domain. A total of 14 features have been extracted from datasets of each hand movements. Several MATLAB built-in functions and formulas were used to calculate the 14 features. Figure 14: shows box plot of some features extracted from wrist pronation signal dataset for all classes. Features extracted from the signals of remaining three hand movements' shows similar characteristics. Some features showed significant differences between classes while the other didn't. The boxplots (Figure 4:5 (1-8)) represent the distribution of the value of 8 features from 14 extracted features versus level of the disease during feature extraction.



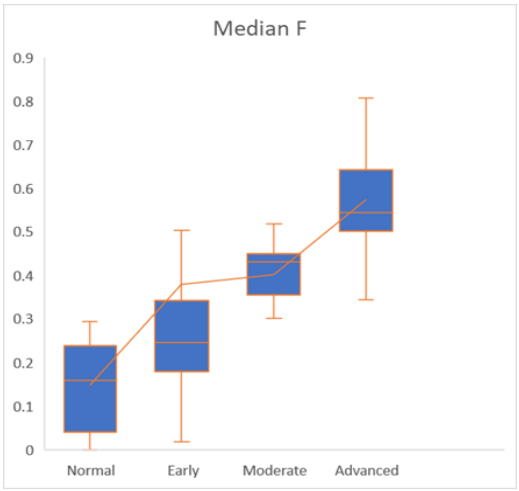
1



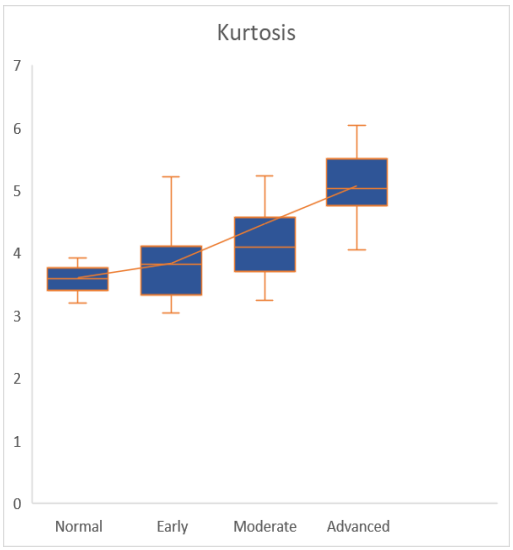
2



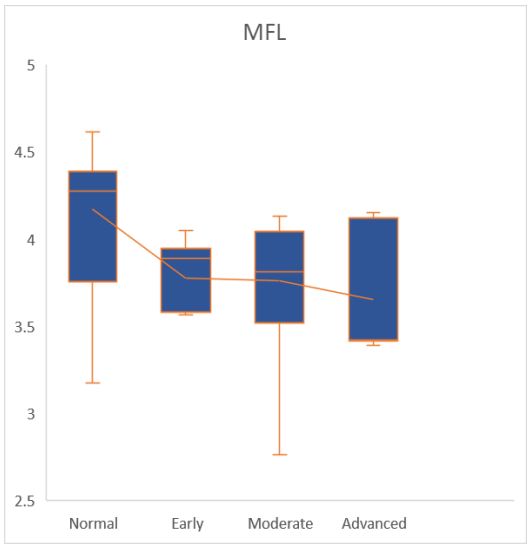
3



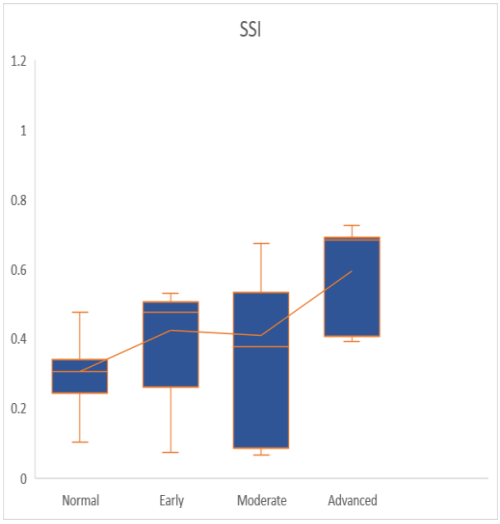
4



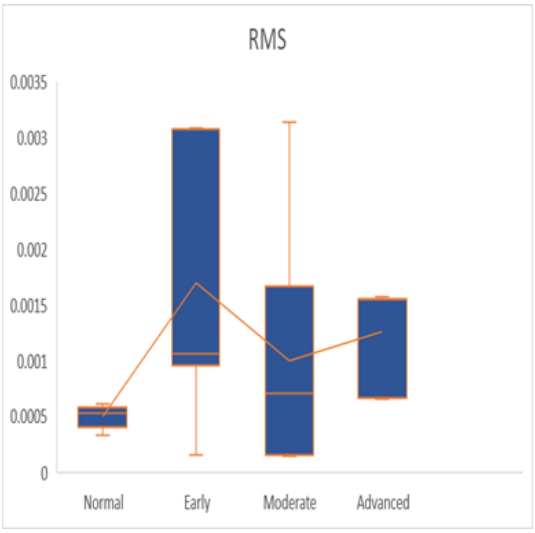
5



6



7



8

Figure 4:5: Comparison for distribution of some features' value versus stage of the disease. (1) Entropy, (2) Mean frequency, (3) Skewness, (4) median frequency, (5) Kurtosis, (6) MFL, (7) SSI, (8) RMS

4.1.5. Feature Reduction

The features have been reduced into nine using ReliefF algorithm. The weight and the rank of selected features are shown in Table 4-2. After applying ReliefF algorithm, only nine features have been selected. As demonstrated in the table best ranked feature are found to be entropy, mean frequency, skewness, and median frequency. Kurtosis, maximum fractal length, root mean square, integrated EMG and mean absolute value got the middle ranks. The last ranks with nearly zero or negative weight were hold by SSI, AAC, variance, waveform length and zero crossing. These last ranked features are omitted.

Table 4-2: List and rank of selected features

Rank	Features	Feature extraction domain	Weight
1	Entropy	Time	0.3408
2	Mean frequency	Frequency	0.2613
3	Skewness	Time	0.2531
4	Median frequency	Frequency	0.2350
5	Kurtosis	Time	0.2294
6	Maximum fractal length	Time	0.1223
7	Root mean square	Time	0.1131
8	Integral EMG	Time	0.1019
9	MAV	Time	0.1016

4.1.6. Classification

The training dataset (with 480 signals sample) was used to train the classifier and the test samples (with 120 signal samples) were given as input for the classifier to test the performance of the classifier. Signals from different hand movements are treated as separate datasets and separate SVM classifiers were trained. In addition to the main model (multi-SVM), multi class KNN and LDA models were trained to compare the result with the main model.

After the model training, the multiclass model was cross-validated to evaluate the generalization performance for all hand movements. In-sample classification errors and out-of-sample classification errors were computed before testing the prediction of the class for unknown or unseen data. According to the cross-validation result the in-sample loss and the out-of-sample loss did not show high difference. This indicates that the model is neither under fitted nor over fitted and the model has good generalization performance. Table below shows the training losses of the model for each hand movements.

Table 4-3: Generalization loss of Multi class SVM

Multi-SVM	Wrist pronation	Touching shoulder	Elbow flexion 90^0 with load	Elbow 90^0 without load
isLoss	0.0182	0.0219	0.0408	0.0347
oosLoss	0.0201	0.0224	0.0424	0.0356

In order to analyze the efficiency of the algorithm, it was evaluated using different performance evaluation methods such as, accuracy, sensitivity, specificity. Confusion matrices list below are visualizations of the performance of the algorithm. Confusion matrix shows the number of true positives, true negatives, false positives and false negatives. By using confusion matrix, the performance measure parameters like, accuracy, sensitivity and specificity can be calculated using the formulas given below.

$$\text{Accuracy} = \frac{TN+TP}{TN+TP+FN+FP} x\ 100\% \quad \ldots\ldots (4.1)$$

$$\text{Sensitivity} = \frac{TP}{TP+FN} x 100\% \quad \ldots\ldots (4.2)$$

Specificity= $\frac{TN}{TN+FP} x100\%$ ……. (4.3)

where, TP = True positive, FP= False positive, TN = True negative and FN= False negative

Tables 4-4, 4-5, 4-6, and 4-7 show the confusion matrix for wrist pronation., elbow flexion with 90 degree without load, touching shoulder and elbow flexion with 90 degree with 1kg load respectively. The confusion matrix has also been made according to the classification of test samples for classification using multiclass SVM for all four hand movements. The test set contains 120 signals (30 normal. 30 early, 30 moderate and 30 advanced level). Different accuracies, sensitivities and specificities are obtained for different hand movements.

Table 4-4: Confusion matrix for wrist pronation

Prediction / Actual	Normal	Early	Moderate	Advanced	Sensitivity
Normal (30)	30	0	0	0	100%
Early (30)	0	27	2	1	90%
Moderate (30)	0	1	29	0	96.7%
Advanced (30)	0	0	0	30	100%
Specificity	100%	98.9%	96.7%	98.9%	Over all Ac=96.6%

According to the confusion matrix of the test result, 3 early stage subjects are misclassed, 2 of them as moderate and 1 as advanced level Parkinson's disease. From moderate level, class 1 subjects are misclassified as early stage PD. Normal and advanced level classes were successfully

classified with no misclassification. The classifier gives overall accuracy of 96.6%, average specificity 98.6% and average sensitivity 96.7% for unseen or test data during wrist pronation.

Table 4-5: Confusion matrix for elbow flexion by 90 degree without load

Prediction / Actual	Normal	Early	Moderate	Advanced	Sensitivity
Normal (30)	28	2	0	0	93.3%
Early (30)	3	26	1	0	86.7%
Moderate (30)	0	0	28	2	93.3%
Advanced (30)	0	0	4	26	86.7%
Specificity	96.7%	97.8%	94.4%	97.8%	Ac=90

According to the confusion matrix of the test result, 2 healthy subjects are mis classed as early, from early stage class 4 subjects are misclassified, 3 of them as normal and 1 as moderate level of PD. From moderate level class 2 subjects are mis classified as advanced level and from advanced class 4 subjects are misclassified as moderate. Generally, the classifier gives overall accuracy 90%, average specificity 96.7% and average sensitivity 90% for test data during elbow flexion and without load.

Table 4-6: Confusion matrix for touching the shoulder.

Prediction / Actual	Normal	Early	Moderate	Advanced	Sensitivity
Normal (30)	30	0	0	0	100%
Early (30)	2	26	0	2	86.7%
Moderate (30)	0	2	28	0	93.3%
Advanced (30)	0	0	0	30	100%
Specificity	97.8%	97.8%	100%	97.8%	Acc=95

According to the confusion matrix, 4 early stage subjects are mis classed 2 of them as normal and 2 of them as advanced level of Parkinson's disease. From moderate level class 2 subjects are mis classified as early. Normal and advanced level classes were successfully classified with no misclassification. The classifier gives overall accuracy 95%, average specificity 98.3% and average sensitivity 95% for unseen test data during the subjects touching their shoulder.

Table 4-7: Confusion matrix elbow flexion by 90 degree with load

Prediction / Actual	Normal	Early	Moderate	Advanced	Sensitivity
Normal (40)	25	2	2	1	83.3%
Early (30)	4	26	0	0	86.7%
Moderate (30)	0	1	29	0	96.7%
Advanced (30)	0	0	0	30	100%
Specificity	95.5%	96.7%	97.8%	98.9%	Acc=91.7%

According to the confusion matrix (Table 4-7), 5 healthy subjects are mis classified 2 of them as early, 2 of them as moderate and 1 of them as advanced level of Parkinson's disease. From early level class 4 subjects are mis classified as normal and from moderate class only 1 subject was classified as early. No misclassification for advanced level. The classifier gives overall accuracy 91.7%, average specificity 97.2% and average sensitivity 91.67% for unseen data during elbow flexion and with 1 kg load.

Based on the test results, the summery of the performance of used model (multiclass SVM) is shown in the Figure 4:6.

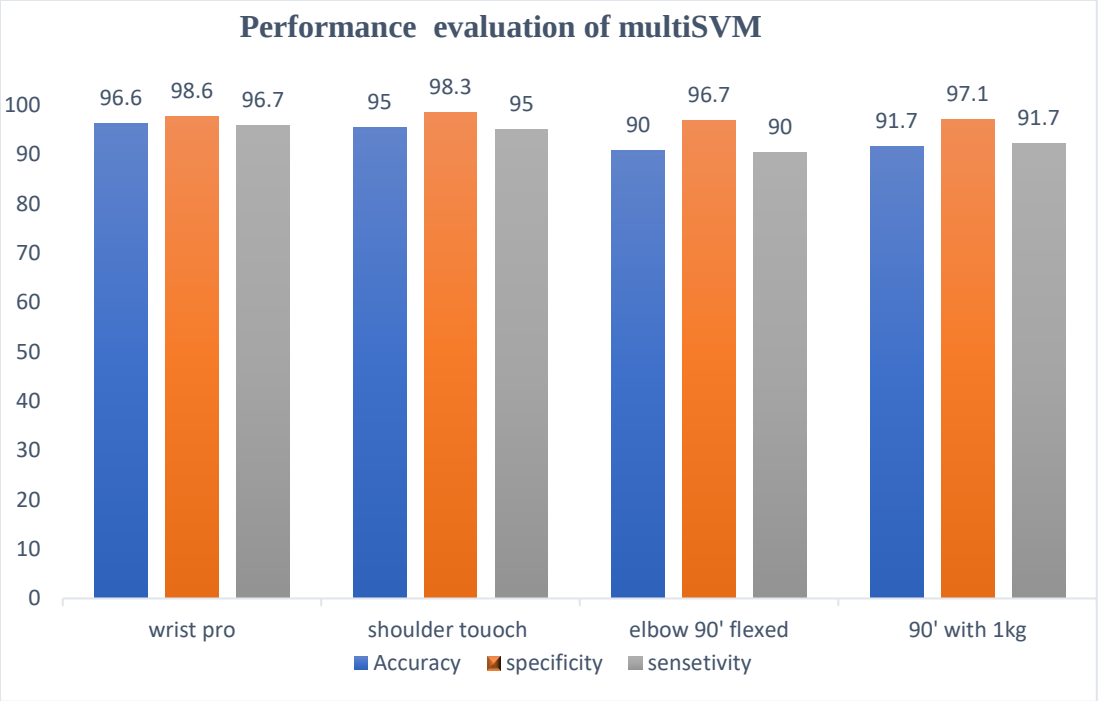


Figure 4:6: Summery for the performance of multi class SVM

In addition to multi class SVM with rbf, SVM polynomial function, SVM with gaussian function, KNN and LDA were trained and the performance of the models were evaluated. The accuracies of these models were compared with the main model (multiclass SVM). Table 8 shows the comparison. According to the accuracy, specificity and sensitivity measured, multiclass SVM with rbf is found to have promising performance to classify EMG signals for the detection and severity determination of PD compared to KNN, LDA, SVM with polynomial and gaussian kernel function. The comparison is illustrated in Table 4-8.

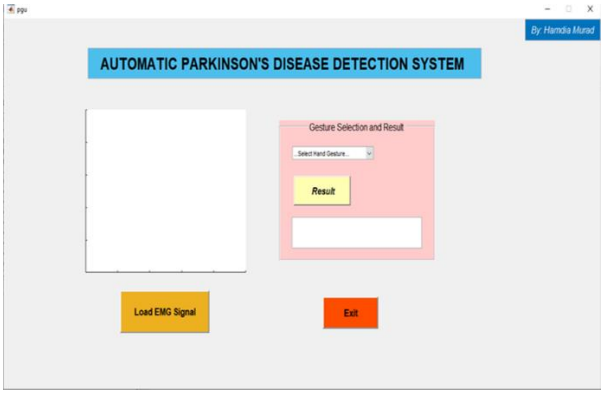
Table 4-8: Comparison of different models

	Wrist pro			**Touching shoulder**			**Elbow flexed by 90^0 without load**			**Elbow flexed by 90^0 with load**		
	Ac	**Sp**	**Se**	**Ac**	**Sp**	**Se**	**Ac**	**Sp**	**Se**	**Ac**	**Sp**	**Se**
LDA	70	89	70	67	88	67.5	63.3	87.4	66.5	72.5	87	72.8
KNN	91.6	95	91.8	90	96	90.4	70	89	71	86	93	86.7
SVM (Polynomial)	92	96	92.5	81.5	92	81.7	83.3	94	83.1	85	94	85.1
SVM (Gaussian)	92.5	95.6	90.5	88.2	94	88.3	86.6	94.2	86.8	88.3	95	88.4
SVM (rbf)	**96.6**	**98.6**	**96.7**	**95**	**98.3**	**95**	**90**	**96.7**	**90**	**91.7**	**97.1**	**91.7**

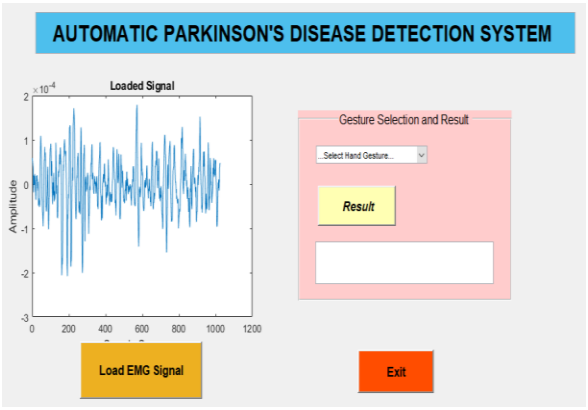
Where, Ac= accuracy, Sp = specificity and Se = sensitivity described in percent.

4.1.7. Graphical user interface (GUI)

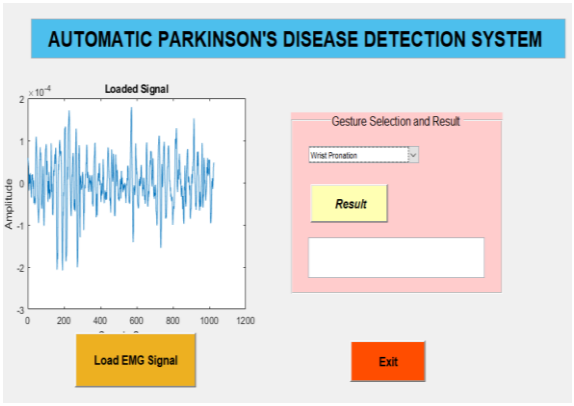
The developed was tested and it works properly. The ‘load signal’ button enables the user to browse and load and display the saved signal, the ‘gesture selection menu’ enables the user to choose the hand movement from which the signal recorded, the ‘result’ button makes the result to be displayed on the display. Finally, the user can close the system using exit button. Figure 4:7 (a-d) shows the pictures for each step. The model also tested with respect to response time and convenience for the user. The response time is 4 seconds and the GUI was checked, it is convenient and simple. The outputs of each steps are illustrated in figure below.



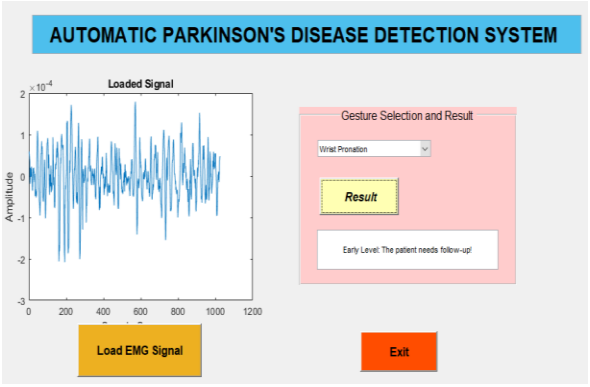
a



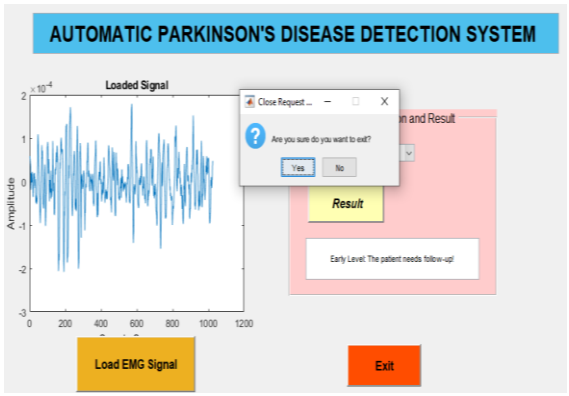
b



c



d



e

Figure 4:7: Test result of GUI. (a) The GUI model, (b) Loading the signal, (c) Selecting hand movement, (d) Displaying the result and (e) Exiting the model

4.2. Discussion

PD is the first most common movement disorder and the second most common neurodegenerative disease that affects around 10 million people worldwide [7] . It causes multilayered challenges for patients living with this disease because of lack of smooth and controlled movement. The main symptoms of PD are classified in to motor and non-motor symptoms [6]. Motor symptoms include: muscle rigidity, tremors, uncontrolled movement and changes in speech and non- motor symptom encompasses sleep disorders, hallucination, constipation, bladder dysfunction, difficulty in concentration, dysphagia, episodes of confusion, fatigue, impulse control disorders, memory problems, mood disorders and sexual dysfunction [19].

PD causes multilayered challenges on patients and their family members additionally; challenges are prolonged throughout their live because of uncurable nature of the disease. Apart from this there is no standardized way of diagnosing this disease. The common approach to diagnose this disease is through clinical assessment of the patients' sign and symptoms. The methods are highly subjective and need the intervention of highly specialized experts. EEG and EMG signals are used for detection of this disease but the signals have complex and irregular natures this also need experts to interpret. So, designing an automated system is highly supportive for health care service improvement in this area.

In this study four different hand movements have been studied to give insight about the effect of Parkinson's disease on muscle activities during contraction. Additionally, this study demonstrated that different time and frequency domain characteristics of EMG signals to distinguish patients with Parkinson's disease from normal groups and study severity level of the disease using four hand gestures.

All patients participated in this study were on medication. Some of the patients especially those with severe level of the disease got difficulties to perform and repeat the instructed hand movements during signal recording. This is because of muscle rigidity, tremor, muscle fatigue and bradykinesia. A total of 1000 signals (250 signals data for each hand movement) were found. This means 100 signals from 10 healthy subjects, 50 signals from 5 early stage patients, 50 signals from 5 mid stage patients and 50 signals from 5 severe stage patients for each hand movements were recorded.

Data augmentation was performed by adding white Gaussian noise with SNR value of 100 and 90 in order to have more data which is significant to protect the model from underfitting and to improve classification accuracy. Half of the healthy subject signals were augmented by adding white gaussian noise with SNR 100 then 150 signals were found and signals for other classes were augmented by adding white gaussian noise with SNR 100 and 90 then the signals were tripled to 150 for each hand movement. The number of signals for patients were less compared to the healthy subjects. So, half of healthy subjects' signals were doubled and the patients' signals were tripled to make a balance between signals of each class. As shown in figure 4:2 the signal maintains its characteristics except presentation of some distortions.

In this study, Butterworth high pass and low pass filters with lower and upper corner frequencies of 20Hz and 500Hz were applied to remove high frequency and low frequency artifacts out of usable range of EMG signal energy band. Additionally, notch filter with cut off frequency 50Hz was used to reject AC power line interference. 50 Hz was selected because it is 50 Hz in Ethiopia case. Even if the effect of filtering was not clearly visible on the signal as presented in figure: 4:3 (a-d), the difference on the performance of the model was significant. The model was trained and tested with extracted features by bypassing this step, compared with the performance of the model along with this step and the difference was significant.

EMG signal provides us valuable information, if it can be quantified by means of features. For this study different features were proposed. It is found that from entropy, mean frequency, kurtosis, skewness and median frequency) gave a clear boundary to distinguish one class from other classes as indicated in Figure 4:5 (1-5) Patients with PD presented low entropy value compared with the healthy groups. Similar result was found in study [69]. This study showed that when severity level increases the entropy value, became decreasing which is reduced entropy value than healthy subjects. This can be interpreted as the EMG signal complexity decreases with the disease severity increase. Other features such as, mean frequency, median frequency, skewness and kurtosis value of signal from Parkinson's disease patients showed increased value related to healthy control groups. P.Kugler et al, [40] found the higher value of mean frequency and kurtosis features of EMG signals recorded from lower limb muscles of PD patients than healthy control subjects. This study also supports the finding and also this study revealed that the value of these features increases while the disease level increases.

The distinguishing capability of remaining four from selected features, was not much satisfactory. Specially to classify the disease severity. Normal subjects presented low value of mean absolute value, integrated EMG, and root mean square. The highest values of these features account for early stage PD level and the value decrease when the severity of the disease stage increases. Other five features (zero crossing, waveform length, variance, simple square integral and average amplitude change) that didn't show significant separation property between different classes, have been rejected using ReliefF algorithm.

The first aim of this study is to detect PD and to classify the level of the disease from the upper limb muscles' myoelectric signals. The present results show the promising separability performance of features using multiclass SVM model and this model shows improved classification performance compared to recent studies [40 , 43, 70] those uses SVM for PD detection and this study showed improved performance with the use of minimum number of feature sets compared to related works [41, 42, 51] that minimize time, cost and complexity of the system. The classification accuracy is improved and this model is able to classify the severity of the disease into three levels.

The second aim of this research is to determine most effective hand movement and muscle type to detect Parkinson's disease and classify the level, the extracted features were tested with multi class SVM. As shown in Tables (4-4, 4-5, 4-6 and 4-7) the multi class SVM model with rbf gives different classification accuracies, sensitivities and specificities for different hand movements. The model gives less sensitivity for heathy subjects during elbow flexion with load which is 83.3% and for patients with advanced level of PD during elbow flexion without load which is 86.7%. Relative to other classes less sensitivity was presented for subjects with early stage PD in all cases, the average sensitivity for all hand movements is 87.5%. Additionally, the model gave highest sensitivity for advanced level of the disease in all cases except elbow flexion without load, the average sensitivity of this class for all hand movements is 96.5%.

Generally, multi class SVM presented over all classification accuracies 90% for flexion by 90^0 without load, 91.7% for flexion by 90^0 with load, 95% for touching shoulder and 96.6% for wrist pronation. Even though all hand movements give promising results, best performance of the model was achieved from the EMG signal acquired from flexor carpi radialis and biceps muscles during wrist pronation and touching the shoulder respectively. This is could be because of the force

needed for wrist pronation and touching the shoulder is higher related to energy needed for elbow flexion movements. But, to deduct this it needs further investigations.

The third aim of this study is to compare the performance of multiclass SVM model with other models. Performance of multi class SVM classifier with rbf kernel function was compared with multi class SVM classifier with polynomial kernel function, multi class SVM classifier with gaussian kernel function, KNN and LDA with the same signal datasets. As shown in Table 8, Multi class SVM with radial basis function (rbf) and with (k fold = 10) showed the promising classification performance compared to all evaluated classifiers.

The proposed system provided robust and promising results that can be clinically applicable with some modifications the GUI model makes the system easy, user friendly and implementable. Thus, it can be used as a decision support system for physicians, especially for those in low resource setting by detecting PD at early stage and classifying it's level and will have a great impact in reducing the disease progression and the mortality rate due to PD.

Chapter Five

5. Conclusion and Recommendations

5.1. Conclusion

An automated system was proposed to detect and classify severity of PD using surface EMG signals recorded from flexor carpi radialis and biceps muscles for four different hand movements. The recorded signals pass through filtering steps to eliminate the effect of noises. Features were extracted and feature reduction algorithm was applied to select relevant features. Then multiclass SVM was trained as a classifier and the performance was evaluated using performance evaluation metrics. According to the results, the proposed method has the best overall performance compared with state-of-art of Parkinson's disease detection methods. In addition to the performance improvement, this study showed that signals recorded during touching the shoulder and wrist pronation have best separability property compared to elbow flexion movements with load and without load for Parkinson's disease detection. Since the selected movements are easy and the recording procedures are simple and painless, the proposed system will be comfortable for patients as well as for physicians. Thus, it can provide more reliable way of diagnosis for Parkinson's disease, severity determination and medication progress.

5.2. Recommendations

Even though PD has five definite stages based on H&Y and UPDRS rating scale procedures, this study was conducted to classify PD in to three levels because of lack of highly specialized experts. Thus, this study can be extended to definite stages classification by making highly specialized medical experts to involve on rating of the patients and labeling of the signals by taking non-motor symptoms in to account in addition to motor symptoms followed in this study. The rating scale procedure followed is H & Y rating scale which sticks on motor symptoms only so, including UPDRS rating scale procedures during the disease level determination for future studies may help for definite classification of the disease in to five stages but this needs highly specialized neurologists.

Additionally, signal recording was done using two electrodes for each channel (to measure activity of one muscle), but assessing the muscle activities by increasing the number of electrodes and channels will increase the accuracy and reliability of the system. By considering mentioned limitations, other studies can be done for further classification to assess the level and progress the disease.

Reference

[1] A. D. Gitler, P. Dhillon, and J. Shorter, “Neurodegenerative disease: Models, mechanisms, and a new hope,” *DMM Dis. Model. Mech.*, vol. 10, no. 5, pp. 499–502, 2017.

[2] S. S. M. Saba, sheikh, safia, Ejazul haque, “Neurodegenerative Diseases : Multifactorial Conformational Diseases and Their Therapeutic Interventions,” *Hindawi Publ. Corp.*, vol. 2013, pp. 1–8, 2013.

[3] P. Batista and A. Pereira, “Quality of Life in Patients with Neurodegenerative Diseases,” *J. Neurol. Neurosci.*, vol. 7, no. 1, pp. 1–7, 2016.

[4] R. C. Brown, A. H. Lockwood, and B. R. Sonawane, “Neurodegenerative Diseases : An Overview of Environmental Risk Factors,” *enviromental Heal. Perspect.*, vol. 1250, no. 9, pp. 1250–1256, 2005.

[5] C. G. Goetz, “The history of Parkinson’s disease: Early clinical descriptions and neurological therapies,” *Cold Spring Harb. Perspect. Med.*, vol. 1, no. 1, pp. 1–16, 2011.

[6] F. Bill and M. G. Foundation, "Articles Global , regional , and national burden of Parkinson ’ s disease" , 1990 – 2016 : a systematic analysis for the Global Burden of Disease Study 2016,” pp. 939–953, 2016.

[7] N. Ball, W. P. Teo, S. Chandra, and J. Chapman, “Parkinson’s Disease and the Environment,” *Front. Neurol.*, vol. 10, March, 2019.

[8] G. Alves, E. B. Forsaa, K. F. Pedersen, M. Dreetz Gjerstad, and J. P. Larsen, “Epidemiology of Parkinson’s disease,” *J. Neurol.*, vol. 255, no. SUPPL. 5, pp. 18–32, 2008.

[9] Y. Iwasa, I. Saito, and C. Fujii, “Investigation of the Treatment and Living Assistance Needed by Patients with Young- Onset Parkinson ’ s Disease,” vol. 64, no. 5, pp. 180–188, 2018.

[10] U. Williams, O. Bandmann, and R. Walker, “Parkinson ’ s Disease in Sub-Saharan Africa : A Review of Epidemiology , Genetics and Access to Care,” vol. 11, no. 2, pp. 53–64, 2018.

[11] J. H. Bower, Z. Melaku, M. Teshome and G. Zenebe, “Frequency of Movement Disorders

in an Ethiopian University Practice," Movement disorders, vol. 20, no. 9, pp. 9–13, 2005.

[12] J. O. Hirsi, Y. M. Yifru, G. Z. Metaferia, and J. H. Bower, "Prevalence of pain in patients with Parkinson's disease in Addis Ababa, Ethiopia," *Park. Relat. Disord.*, vol. 61, pp. 214–218, 2019.

[13] M. Champan,, " Fox Foundation Awards $5M to Support Genetic Studies of Parkinson's in Africa, Asia and India", August 29, 2019. Accessed on: November 18, 2019. Accessed on: November 18, 2019. [Online]. Available: https://parkinsonsnewstoday.com/2019/08/29/fox-foundationawards-5-million-to-genetic parkinsons-studies-in-africa-east-asia-india/.

[14] H. Asfaw, "The Psychosocial and Economic Impact of Parkinson's Disease on People Infected and Affected by the Disease," Addis Abba University, 2014.

[15] A. Galvan and T. Wichmann, ""Pathophysiology of Parkinsonism," Clin Neurophysiol, vol. 119, no. 7, pp. 1459–1474, 2008.

[16] T. Berman and A. Bayati, "What are Neurodegenerative Diseases and How Do They Affect the Brain?," *Front. Young Minds*, vol. 6, 2018.

[17] L. M. Bollinger, C. E. Cowan, and T. P. LaFontaine, "Parkinson's disease: Epidemiology, pathophysiology, and exercise intervention," *Strength Cond. J.*, vol. 34, no. 2, pp. 50–54, 2012.

[18] F. Magrinelli, A. Picelli, P. Tocco, A Federico, A. Roncari, N. Smania, G. Zanette and S. Tamburin "Pathophysiology of Motor Dysfunction in Parkinson's Disease as the Rationale for Drug Treatment and Rehabilitation," Hindawi Publishing Corp., vol. 2016, 2016.

[19] G. Bronner and D. B. Vodusek, "Management of sexual dysfunction in Parkinson's disease," *Ther. Adv. Neurol. Disord.*, vol. 4, no. 6, pp. 375–383, 2011.

[20] G. Ayano, "Parkinson's Disease: A Concise Overview of Etiology, Epidemiology, Diagnosis, Comorbidity and Management," *J. Neurol. Disord.*, vol. 4, no. 6, 2016.

[21] S. N. Heyn and C. P. Davis, " Parkinson's disease ". Accessed on: November 20, 2019. [Online]. Available: https://www. medicinet.com/parkinsons-disease.net/diagnosis/

article.htm.

[22] E. Downward, "How is Parkinson's Disease Diagnosed?", March, 2017. Accessed on: November 20, 2019. [Online]. Available: https://parkinsonsdisease.net/diagnosis/.

[23] American Parkinson's disease association, "Diagnosing Parkinson's disease". Accessed on: November 21, 2019. [Online]. Available: https://www.apdaparkinson.org/what-is-parkinsons/diagnosing/.

[24] R.Evans, "Difficulties in Diagnosing Parkinson's Disease-Varsity Branding ," April 24, 2018. [Online]. Accesed on: November 21, 2019. Available: https://www.varsitybranding.com/2018/04/24/difficultiesdiagnosing-parkinsons-disease/.

[25] C. G. Goet Goetz, G. Christopher W. Rascol, O. Sampaio, C. Stebbins, T. G. Counsell, C. Giladi, N. Holloway, R. G. Moore, C. G. Wenning, G. K. Yahrb and M. D. SeidL,., "Movement Disorder Society Task Force report on the Hoehn and Yahr staging scale: Status and recommendations," *Mov. Disord.*, vol. 19, no. 9, pp. 1020–1028, 2004.

[26] R. Barmore, "Parkinson's Foundation- Better Lives". [Online]. Accessed on: December 1, 2019. Available: https://www.parkinson.org/Understanding-Parkinsons/What-is Parkinsons/Stages-of-Parkinsons/.

[27] C. Spears, "Parkinson's Foundation- Better Lives". Accessed on: December 4, 2019. [Online]. Available: http://parkinson.org/symptoms-of- Parkinson/.

[28] R. Cano-De-La-Cuerda, M. Pérez-De-Heredia, J. C. Miangolarra-Page, E. Muñoz-Hellín, and C. Fernández-De-Las-Peñas, "Is there muscular weakness in parkinson's disease?," *Am. J. Phys. Med. Rehabil.*, vol. 89, no. 1, pp. 70–76, 2010.

[29] T. V. Correa, V. L. Santos de Britto, and C. Lourenco Correa, "Progressive Muscular Strength Protocol for the Functionality of Upper Limbs in Individuals with Parkinson's disease: Protocol Study," *Int. J. Phys. Med. Rehabil.*, vol. 05, no. 06, pp. 23–25, 2017.

[30] A. Y. Meigal, S. M. Rissanen, and P. Mika, "Non-linear EMG parameters for differential and early diagnostics of Parkinson ' s disease," *www.frontier.org*, vol. 4, pp. 1–8, 2013.

[31] S. Rissanen, M. Kankaanp, M. P. Travainen, J. Nuuteinen, I. M. Tarka, O. Airaksinen and P. A. Karjalainen, "Analysis of surface EMG signal morphology in Parkinson ' s disease."

[32] R. Moghadam and M. Homam, "Quantification of Parkinson Tremor Intensity Based On EMG Signal Analysis Using Fast Orthogonal Search Algorithm," *ijeee.iust.ir*, vol. 14, no. 2, pp. 106–115, 2018.

[33] D. Landin, M. Thompson, and M. R. Jackson, "Actions of the Biceps Brachii at the Shoulder: A Review," *J. Clin. Med. Res.*, vol. 9, no. 8, pp. 667–670, 2017.

[34] C. A. Pressure, F. C. R. H-reflex, S. Without, and N. Deficits, "Without Neurological Deficits '," no. 1990, pp. 565–579, 2006.

[35] R. Barański and A. Kozupa, "Hand grip-EMG muscle response," *Acta Phys. Pol. A*, vol. 125, no. 4 A, pp. 7–10, 2014.

[36] B. P. and K. G. Yamik Mangukiya, "Electromyography(EMG) sensor controlled Assistive Orthotic Robotic Arm for Forearm Movement," *IEEE*, 2017.

[37] M. Wehner, "Man to Machine, Applications in Electromyography," *EMG Methods for Evaluating Muscle and Nerve Function*, InTech, 2012. .

[38] O. Dorgham, I. Al-Mherat, J. Al-Shaer, S. Bani-Ahmad, and S. Laycock, "Smart system for prediction of accurate surface electromyography signals using an artificial neural network," *Future Internet*, 2019. .

[39] M. Linderman, "Electromyography Signals as Biomarkers for Parkinson's Disease, April 1, 2015. Accessed on: August 10, 2019. [Online]. Available: https://www.michaeljfox.org/grant/electromyographysignals- biomarkers-parkinsons-disease.

[40] P. Kugler and J. K. C. Jaremenko, J. Schlachetzki and J. Winkler, "Automatic Recognition of Parkinson ' s Disease Using Surface Electromyography During Standardized Gait Tests," *IEEE*, no. 4208, pp. 5781–5784, 2013.

[41] P. Pekerja, P. Kelapa, and S. Bagerpang, "Electromyography Gait Test for Parkinson Disease Recognition Using Artificial Neural Network Classification in Indonesia," *Momentum*, vol. 2013, no. 2, pp. 23–27, 2013.

[42] F. Putri, W. Caesarendra, E. D. Pasmanasari, M. Ariyanto, and J. D. Setiawan, “Parkinson Disease Detection Based on Voice and EMG Pattern Classification Method for Indonesian Case Study,” *JEMMME*, vol. 3, no. 2, pp. 87–98, 2018.

[43] N. H. Ghassemi *et al.*, “Combined Accelerometer and EMG Analysis to Differentiate Essential Tremor from Parkinson’s Disease,” no. 1, pp. 1–4.

[44] C. J. De Luca, L. Donald Gilmore, M. Kuznetsov, and S. H. Roy, “Filtering the surface EMG signal: Movement artifact and baseline noise contamination,” *J. Biomech.*, vol. 43, no. 8, pp. 1573–1579, 2010.

[45] Amrutha N and Arul V H, “A Review on Noises in EMG Signal and its Removal,” *Int. J. Sci. Res. Publ.*, vol. 7, no. 5, p. 23, 2017.

[46] M. W. Flood, B. R. Jensen, A. S. Malling, and M. M. Lowery, “Increased EMG intermuscular coherence and reduced signal complexity in Parkinson’s disease,” *Clin. Neurophysiol.*, vol. 130, no. 2, pp. 259–269, 2019.

[47] A. K. Bhoi, “Classification and Clustering of Parkinson ’ s and Healthy Control Gait Dynamics Using LDA and K-means,” *INT.J.BIOAUTOMATION*, vol. 21, no. 1, pp. 19–30, 2017.

[48] John R.Giudicessi, BA.Michael J.Ackerman., “Frequency analysis of lower extremity electromyography signals for the quantitative diagnosis of dystonia,” *J Electromyogr Kinesiol*, vol. 23, no. 1, pp. 1–7, 2014.

[49] C. Spiewak, “A Comprehensive Study on EMG Feature Extraction and Classifiers,” *Open Access J. Biomed. Eng. Biosci.*, vol. 1, no. 1, pp. 1–10, 2018.

[50] R. A. M. B. Geman, Oana, Jonni Guiller Ferreira MADEIRA4, “Joint EEG - EMG Signal Processing for Identification of the Mental Tasks in Patients with Neurological Diseases,” *IEEE*, pp. 1598–1602, 2016.

[51] S. Shijiya and P. Thomas, “A Modular Approach to Finer Classification of EMG Signals,” *IJAREEIE*, vol. 5, no. 10, pp. 7893–7900, 2016.

[52] A. Sakai, Y. Minoda, and K. Morikawa, “Data augmentation methods for machine-learning-

based classification of bio-signals," *BMEiCON 2017 - 10th Biomed. Eng. Int. Conf.*, vol. 2017-Janua, pp. 1–4, 2017.

[53] R. T. B, T. Ege, and K. Y. B, "Data Augmentation for EEG-Based Emotion Recognition with Deep Convolutional Neural Networks," *Springer Int. Publ.*, vol. 1, pp. 352–356, 2018.

[54] A. Fougner, "Proportional Myoelectric Control of a Multifunction Upper-limb Prosthesis Anders Lyngvi Fougner Master," pp. 1–55, 2007.

[55] T. R. Lee, Y. H. Kim, and P. S. Sung, "Spectral and entropy changes for back muscle fatigability following spinal stabilization exercises," *J. Rehabil. Res. Dev.*, vol. 47, no. 2, pp. 133–142, 2010.

[56] M. Borowska, "Entropy-based algorithms in the analysis of biomedical signals," *Stud. Logic, Gramm. Rhetor.*, vol. 43, no. 56, pp. 21–32, 2015.

[57] C. Spiewak, "A Comprehensive Study on EMG Feature Extraction and Classifiers," *Open Access J. Biomed. Eng. Biosci.*, vol. 1, no. 1, pp. 17–26, 2018.

[58] K. Nazarpour, A. H. Al-Timemy, G. Bugmann, and A. Jackson, "A note on the probability distribution function of the surface electromyogram signal," *Brain Res. Bull.*, vol. 90, no. 1, pp. 88–91, 2013.

[59] A. Phinyomark, P. Phukpattaranont, and C. Limsakul, "Fractal analysis features for weak and single-channel upper-limb EMG signals," *Expert Syst. Appl.*, vol. 39, no. 12, pp. 11156–11163, 2012.

[60] J. Too, A. R. Abdullah, T. N. S. T. Zawawi, N. M. Saad, and H. Musa, "Classification of EMG Signal Based on Time Domain and Frequency Domain Features," *Int. J. Hum. Technol. Interact.*, vol. 1, no. 1, pp. 2590–3551, 2017.

[61] A. Phinyomark, P. Phukpattaranont, and C. Limsakul, "Feature reduction and selection for EMG signal classification," *Expert Syst. Appl.*, vol. 39, no. 8, pp. 7420–7431, 2012.

[62] M. Arozi, F. T. Putri, M. Ariyanto, M. Khusnul Ari, Munadi, and J. D. Setiawan, "Electromyography (EMG) signal recognition using combined discrete wavelet transform based adaptive neuro-fuzzy inference systems (ANFIS)," *AIP Conf. Proc.*, vol. 1788, pp.

1–9, 2017.

[63] S. F. Rosario and Tangadurai, “RELIEF: Feature Selection Approach,” *Int. J. Innov. Res. Dev.*, vol. 4, no. 11, pp. 218–224, 2015.

[64] F. Falceto and K. Gawedzki, “Theoretical and Empirical Analysis of ReliefF and RReliefF,” *Commun. Math. Phys.*, vol. 159, no. 3, pp. 549–579, 1994.

[65] Z. Wang *et al.*, “Application of ReliefF algorithm to selecting feature sets for classification of high resolution remote sensing image,” *Int. Geosci. Remote Sens. Symp.*, vol. 2016-Novem, no. 41271013, pp. 755–758, 2016.

[66] R. P. L. DURGABAI and R. B. Y, “Feature Selection using ReliefF Algorithm,” *Ijarcce*, vol. 3, no. 10, pp. 8215–8218, 2014.

[67] H. Li, F. Qi, and S. Wang, “A comparison of model selection methods for multi-class support vector machines,” *Lect. Notes Comput. Sci.*, vol. 3483, no. IV, pp. 1140–1148, 2005.

[68] D. Roth, “Multiclass Classification Binary to multiclass,” *CS 446 Mach. Learn. Fall*, pp. 1–13, 2016.

[69] M. W. Flood, B. R. Jensen, A. S. Malling, and M. M. Lowery, “Increased Emg Intermuscular Coherence and Reduced Signal Complexity in Parkinson’s Disease,” *Clin. Neurophysiol.*, pp. 1–31, 2018.

[70] H. Jeon, W. Lee, H. Park, H. J. Lee, S. K. Kim, H. B. Kim, B. Joen and K. S. Park, “Automatic classification of tremor severity in Parkinson’s disease using awearable device,” *Sensors (Switzerland)*, vol. 17, no. 9, pp. 1–14, 2017.

Appendix

SCU-7 EMG System Specifications

SCU-7 EMG system Amplifier module specifications*:*

Some specifications of SCU-7 EMG system amplifier used in the study are illustrated in table below.

Table 0-1: Specification of SCU-7 EMG system

No.	Specification	Description
1	Amplification	100K, 1K, 10K
2	Maximum output voltage range	± 10 Volt
3	Channel frequency response	ET1-5: 13± 1.3 Hz to 463±27.3 Hz ET6: 11 Hz - 3.78 KHz
4	System noise (RTI)	< 1.0 uV (rms)
5	Power requirements	9 V battery
6	Operating temperature	15 to 40 C (59 to 104 F)

Specification of Electrodes

The electrode used for this study is multipurpose electrode that can be used for patient monitoring, ECG and EMG. Some of specifications of disposable multipurpose electrodes used in this study are illustrated in table below

Table 0-2: Specification of multipurpose electrode

No	Specification	detail
1	Gel type	Sticky
2	Dimensions	Electrode Size: 4.06cm x 3.45cm Skin Contact Size: 4.06cm x 3.45cm
3	Lifetime	Recommended max application time: 1-time use Sealed pouch: 2 years
4	Electrode Materials	Backing Material: Foam Connector: Stainless Steel Snap Sensor Material: Silver/silver-chloride coated plastic
5	Operating Temp	15^0 C -30^0 C

Sample MATLAB source code

```
% % % % Matlab functions of some features % % % %
% Entropy function
function E = Entropy(x)
NFFT = length(x);
Y = fft(x,NFFT);
S = abs(Y(1:NFFT/2+1))/NFFT;
S = S./sum(S);
S=S(find(S~=0));
p = S/sum(S);
E = -sum(p .* log(p))/log(NFFT);
end
% zero crossing
function ZC=ZeroC(x,thres)
N=length(x);
ZC=0;
thres=0.001;
for i=1:N-1
  if ((x(i) > 0 && x(i+1) < 0) || (x(i) < 0 && x(i+1) > 0))...
&&(abs(x(i)-x(i+1)) >= thres)
    ZC=ZC+1;
  end
end
end
% average amplitude change.
function AAC=FAAC(X)
N=length(X); Y=0;
for i=1:N-1
  Y=Y+abs(X(i+1)-X(i));
end
AAC=Y/N;
end
% variance
function VAR=FVAR(x)
N=length(x); VAR =(1/(N-1))*sum(x.^2);
end
% SSI
function SSI=FSSI(X)
SSI=sum(X.^2);
End
```

Extracting feature vector, training and testing for shoulder touch

Train 1

```
% extracting feature vector for training SVM with sigals during
% shoulder touch
clear all;
close all;
clc;
% assign aray for data
datasht=[];
% Load the signal folder
```

```
for i=1:480
k=dlmread(['C:\Users\Hamdi\Documents\MATLAB\hamdisht\' num2str(i) '.txt']);
% fs=1024;
% l=length(k);
%Apply filters
% notch filter
[b,a]=iirnotch(0.1,0.1,1);
f=filter(b,a,k);
%Low pass filter
[b1,a1]=butter(6,0.97,'low');
f1=filter(b1,a1,f);
% high pass filter
[b2,a2]=butter(6,0.04,'high');
f2=filter(b2,a2,f1);
% Taking FFT
fft1= fft(f2,1024);
Mag= fft1.*conj(fft1)/1024;
f= 1000/1024*(1:512);
% figure
% plot(f,Mag(1:512)
% % % % % feature extraction % % % %
absolute =(abs(f2));
% % mav
mav=mean(absolute);
% % RMS
rms=sqrt(mean(absolute.^2));
% %IEMG;
iemg=sum(absolute);
% kurosis
kur = kurtosis(f2);
% skeness
sk = skewness(f2);
% Entropy
En= Entropy(f2);
% maximum fractal length
mfl=FMFL(f2);
% normalized mean and median frequencies
meanf = meanfreq(Mag);
medf= medfreq(Mag);
featurevector1=[en,meanf,sk,medf,kur,mfl,rms,iemg,mav];
datasht=[datasht;featurevector1];
save shttrainfeatures.mat datasht
end
% Train the classifier
% load the feature vector
load shttrainfeatures.mat;
% Extract the data saved as datasht
featuressht = datasht;
% Load the excel file contains the label of classes
namesht=csvread('hamdishttrainlabel.csv');
% SPECIFY CLASSIFIER PARAMETERS
svmParams = templateSVM('KernelFunction','rbf', 'KernelScale', 'auto',
'Standardize', 1);
rng(1)
% TRAIN THE CLASSIFIER MODEL
finalmodelsht = fitcecoc(featuressht, namesht,'Learners', svmParams,
'Coding', 'onevsall' );
```

```
save finalmodelsht

Test 1
% extracting feature vector for testing SVM with sigals during
% shoulder touch
% assign aray for data
datashttest=[];
% Load the signal folder
for i=481:600
k=dlmread(['C:\Users\Hamdi\Documents\MATLAB\hamdisht\' num2str(i) '.txt']);
% fs=1024;
% l=length(k);
%Apply filters
% notch filter
[b,a]=iirnotch(0.1,0.1,1);
f=filter(b,a,k);
%Low pass filter
[b1,a1]=butter(6,0.9,'low');
f1=filter(b1,a1,f);
% high pass filter
[b2,a2]=butter(6,0.04,'high');
f2=filter(b2,a2,f1);
% Taking FFT
fft1= fft(f2,1024);
Mag= fft1.*conj(fft1)/1024;
f= 1000/1024*(1:512);
% figure
% plot(f,Mag(1:512)
% % % % % feature extraction %  % % %
absolute =(abs(f2));
% % mav
mav=mean(absolute);
% % RMS
rms=sqrt(mean(absolute.^2));
% %IEMG;
iemg=sum(absolute);
% kurosis
kur = kurtosis(f2);
% skeness
sk = skewness(f2);
% Entropy
En= Entropy(f2);
% maximum fractal length
mfl=FMFL(f2);
% normalized mean and median frequencies
meanf = meanfreq(Mag);
medf= medfreq(Mag);
featurevector1=[en,meanf,sk,medf,kur,mfl,rms,iemg,mav];
datashttest=[datashttest;featurevector1];
save shttestfeatures.mat datashttest
end
% test with extracted test feature vector
% LOAD THE FEATURE SET
load testfeaturessht.mat;
% EXTRACT THE DATA
nnsh = datatestsht;
%Load the Model;
```

```
load finalmodelsht
Result = predict(finalmodelsht, nnsh);
Testclassname=csvread('hamdishttestlabel.csv');
confusionsht=confusionmat(Testclassname,Result)
save confusionsht
```