Computational Intelligence based EMG for Automated Classification of Foot Drop Rehabilitation

By

SAHAR ADIL ABBOUD

Dissertation Submitted in partial fulfillment of the requirements for the Degree of

Doctor of Philosophy

Supervisor: Assoc. Prof. Adel Al-Jumaily

School of Biomedical Engineering

Faculty of Engineering and Information Technology

University of Technology Sydney (UTS)

August 2019

Certificate of Original Authorship

I, Sahar Adil Abboud, declare that this thesis, is submitted in fulfillment of the requirements for the award of Doctorate of Philosophy in the School of Biomedical Engineering at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise referenced or acknowledged. IN addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution.

This research is supported by the Australian Government Research Training Program.

Signature of Candidate

Production Note:
Signature removed
prior to publication.

6/5/2019

Abstract

Foot drop is a complication that arises from the weakness that occurs in specific muscles in the ankle and foot; such as the Anterior Tibialis muscle (AT) during the foot flexion and extension. A lesion in the Lower Motor Neuron (LMN) will cause Foot Drop. Foot drop has been found to arise in 52% to 67% of patients with spinal Upper Motor Neuron (UMN) pathology. Foot Drop (FD) is a common disorder and is not specific to age; it affects around 1% of women and 2.8 % of men. The affected muscles impact on the motion of the ankle and foot both downward and upward. To overcome this problem and to improve rehabilitation devices, this thesis introduces several methods to improve the performance of the Myoelectric Pattern Recognition (M-PR) in both the offline and real data experiments. The thesis proposes a new M-PR system that will work satisfactorily on both the healthy and non-healthy leg by classifying movements in the offline experiments. The thesis describes the state-of-the-art procedures for M-PR and studies all the conceivable structures for the best M-PR features and classifiers.

This thesis presents new classifiers to develop the performance of the M-PR. The Self-Advised-Support Vector Machine (SA-SVM) modified from a single class to multi-class. Developments in methodology lead to more significant applications for overall use. The thesis adapted and altered label classification methods resulting in a new classification named Label Self- Advised-Support Vector Machine LSA-SVM. Further, a development to LSA-SVM is to upgrade from a single class into Multi-LSA-SVM and then to evolve the methodology to match Extreme Learning Machine (ELM) with LSA-SVM to acquire a new rapid method, named ELM-LSA-SVM. For the real data experimental option, a collected data from using the sEMG device from Foot Drop Patients in Metro Rehabilitation Hospital in Sydney, Australia using Ethical Approval (UTS HREC NO.ETH15-0152).

The collected used to apply the latest and fastest method to FD patients to use the myoelectric pattern recognition (M-PR) for leg movement detection. The experimental results for the EMG dataset and benchmark datasets exhibit its benefits. Furthermore, the experimental results on UCI datasets indicate that ELM-LSA-SVM achieves the best performance when working together with LSA-SVM and SVM. The whole sets of experimental results are encouraging as recorded and reported in the thesis.

Acknowledgments

During the Ph.D. study, have had the privilege of working with many people who have made time at the University of Technology, Sydney an enjoyable and intellectually stimulating experience.

I want to thank all the people who have helped along the way and contributed to this thesis. I am especially grateful to the supervisor, Associate Professor Dr. Adel Al-Jumaily. He has been very motivating throughout these years. While working with Dr. Adel, I have learned how to pursue research with intellectual rigor and how to evaluate work and improve research skills critically.

I want to thank friends and colleagues who have encouraged and supported to accomplish a degree.

As a final point, I would like to acknowledge the unending love of my family who has been a constant source of support. Their faith in and encouragement has helped throughout life.

Dedication

This thesis dedicated to the five most important people in life, without their support, this journey would have been just a dream. Special thanks to mother, Sammera A.Gh., father, Adil Abboud (Allah place him in heaven), bother Jameel, Daughter Noor and my very best friends for being my inspiration and encouraging me throughout my career.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AFO Ankle Foot Orthotic
ANN Artificial Neural Network
AR Auto-Regressive (model)
SA-SVM Self Advised Support Vector Machine
BCI Brain-Computer Interfaces
BPF Band-Pass Filtering
CNS Central Nervous System
COM Center of Mass
CPN Common Peroneal Nerve
CSA Cross-Sectional Area
DOF Degree of Freedom
DSP Digital Signal Processor
DM Data Mining
ECG Electrocardiograph
eCGS Comfortable Gait Speed
EEG Electroencephalography
EI Efficiency Index
EMG Electromyography
ES Electrical Stimulation
ESCS Epidural Spinal Cord Stimulation
FES Functional Electrical Stimulation
FLC Free Fuzzy Logic Controller
GRF Ground Reaction Force
GRNN General Regression Neural Network
HGO Hip Guidance Orthotics
HRF Handle Reaction Force
IEMG Integrated EMG
iFES-LCE isokinetic Leg Cycle Ergometer
IOFL Input Output Feedback Linearised
KAFO Knee-Ankle-Foot Orthosis

KNN K Nearest Neighbour

MA Moving Average

macro MUP macro Motor Unit Potential

MeCFES Myoelectrically Controlled Functional Electrical Stimulator

MFPV Muscle Fibre Propagation Velocity

MGH FAC Massachusetts General Hospital Functional Ambulation Classification

MLD Multi Labelling Dimensionality

NFES Non Function Electrical Stimulation

NIH National Institutes of Health

NMES Neuromuscular Electrical Stimulation

PCI Physiological Cost Index

PD Parkinson's Disease

RF Rectus Femoral

RGO Reciprocating Gait Orthotics

RMS Root Mean Square

ROM Range Of Motion

RTS Randomly Timed Stimulation

RTT Robotic Treadmill Training

SCI Spinal Cord Injury

SCKAFO Stance-Control Knee Ankle Foot Orthosis

SCS Spinal Cord Stimulation

SD Standard Deviation

sEMG surface EMG

SOMF Second Order Moment Function

SThD Single Threshold detection

TA Tibias Anterior

TLSI Thoracolumbar Spinal Injury

TO Toe-Off

VL Vastus Lateralis

VM Vastus Medialis

VR Virtual Reality

VSNR Virtual Signal-to-Noise Ratio

WPT Wavelet Packet Transform

WSTT Weight-Supported Treadmill Training.

Chapter 1

Introduction

Developing classifiers that used for the electromyography (EMG) signal for the foot drop rehabilitation application and devices have shown the success in the laboratory environment. However, the acceptance proportion of the rehabilitation device on the clinical application is deficient. An essential part of achieving real-time environment application is required understanding the foot drop problem and use an accurate real-time classifier. This chapter will present the motivations of this thesis and give an outline of the problem that will study briefly in the literature review chapter. The questions relevant to the thesis will identify as well as the reasons for their significance, and then provide a chapter-by-chapter summary as per the thesis.

1.1 Background

Stroke represents as one of the significant health problems. One of the effects of the stroke is foot drop. Foot drop is a weakness that occurs in specific muscles in ankle and foot such as Anterior Tibialis, Gastrocnemius, Plantaris and Soleus muscles. Foot flexion and extension usually generated by Lower Motor Neuron (LMN). Foot drop manifests in 52% to 67% of patients with spinal upper motor neuron (UMN) [1]. Foot Drop (FD) is a common disorder without specifying to age and affects around 1% of women and 2.8 % of men [2]. Bio-signal control systems like electromyography (EMG) used for Leg Rehabilitation devices that including the Foot Drop, one of such system is Function Electrical Stimulation (FES) [3]. This thesis has proposed several methods and algorithms to improved automated rehabilitation classification and testing them in a real-time.

Function Electrical Stimulation (FES) as a therapeutic model stated as neuron modulation, which is the uppermost classification of interferences used in restorative neurology. Developed from physiological and biomedical engineering in the1960s, the therapeutic model is a modulation of the Central Nerve System (CNS) movement rather than the permanent alteration of the CNS structure. Vladimir Liberson's functional electrotherapy changed its name to functional electrical stimulation (FES), and they used it for patients with foot drop after stroke events[4]. One of the traditional rehabilitation devices such as the Hanger Walk Aid (WA) [5] can control by a hand switch, a foot sensor or a tilt sensor. All three of these are used to set up the unit controller. WA automatically records the number of hours for each day and the number of stimulus trains it carries. The modern and most advanced FD rehabilitation devices are **Bioness, NESS L300Plus** which is used to investigate the effects of daily peroneal and thigh muscle. It used on the temporal aspects of gait act in individuals with hemiparesis who have the walking ability and demonstrates ankle and knee dysfunction [6, 7].

Dealing with user's purpose by the myoelectric signal is an exciting task, giving its user many forms of aid. The main benefit is to generate a smooth interface between the human and the rehabilitation device. This interface will permit the device to be a natural extension of the human body. This mechanism can be complete by capturing the message held by muscles from the human brain and then transmitted to the device allowing processing with suitable technique and methodology.

Electromyography (EMG) that record myoelectric signal from muscle activity has been widely used to detect the user's intention beforehand [8]. The EMG electrodes are placed in the human's limb either in an invasive or non-invasive way. A wide range of people does not like to plant electrodes inside the body; instead, preferring them to use of the surface EMG (sEMG). However, the sEMG has numerous disadvantages such as the crosstalk from other muscles and its robustness. Furthermore, it is hard to acquire a myoelectric signal from deeper muscles so that it is not easy to deal with myoelectric signal processing using EMG.

Myoelectric pattern recognition (M-PR) or EMG-based pattern recognition (EMG-based PR) can treat with various limb activities. M-PR implemented in the Leg motion recognition for years, and it indicates supports above EMG-based non-pattern recognition. The surface EMG based Leg motion recognition has a big challenge that required final clarifier which response to EMG signal nature and environmental change. The performance of myoelectric pattern recognition M-PR is decreasing depending on changes in the EMG signal. For example, electrode shift can demote the performance of the M-PR up to 25 – 30 % [9]. Besides, the varied environment of testing may change the characteristic of EMG signals. For instance, the trial test done today will have various setups with the experiment the day after or maybe a couple of days after it. This fact raises concerns about the M-PR that can adapt to changes in EMG signals.

Support vector machines (SVMs) are a machine learning model based on arithmetical learning theory [10]. For a long time, support vector machine (SVM) has widely used in classification applications because of their astounding classification proficiency and its variations. SVM have two keys of learning feature: a first feature where SVM over a nonlinear feature mapping function φ(x), the training sample data first mapped into a higher dimensional feature space. Second feature: the typical optimization technique is applied, to discover the explanation by maximizing the separate margin of two various classes in the feature space, leading to minimizing train errors.

Many Bio-medical applications have applied SVMs techniques to classify data or to diagnose diseases. For example by developing an automated diagnosis system for skin cancer [11] and detection and prediction of sleep apnoea [12]. Another area of application is the National Health and Nutrition Examination Survey (NHANES). NHANES used probability data sample of US residents to construct SVM and Logistic Regression classification models for two classification systems: 1) people who have diabetes (diagnosed or undiagnosed) versus people do not have diabetes, 2) people who have undiagnosed diabetes or primer diabetes versus people have not had diabetes.

The standard SVM disregards the train data that is not separate linearly by the kernels through the training stage, which occurs during the outline of the tolerance parameters in the impartial function and restrictions. For this reason, it will be classified incorrectly if the data is similar or matching the misclassified data that appears in the test set. Ti is happened when the data that is close to the misclassified data is unspecified. Thus resulting in a misclassification that is not sensible and not controlled [13].

Many researchers adapted versions of SVM which are present in chapter 3, targeting to rise classification efficiency and performance for some applications. **Use of the label classification on expert knowledge is one of the methods that was used without increased cost since there will be no additional of an extra parameter.**

1.2 Thesis Motivations

This thesis proposes a Computational Intelligence for automated classification of Foot Drop Rehabilitation with machine learning methods. A machine learning methods such as Support Vector Machines or Extreme Learning Machines have made several experimental achievements [14] which is frequently overtaken by other learning methods in a wide range of occupations. Moreover, various machine learning techniques with different inputs also studied in the range of FD problem.

Consequently, for the detection of Foot Drop, input signals will be tested to choose the best and most significant one by using the support vector machine such as a classifier.

It is essential to develop computational tools to overcome the problem and to improve the reliability of the diagnosis process for automated diagnosis that operates on a quantitative measure. Such tools can facilitate objective mathematical judgment complementary to that of medical experts, and help them to identify the affected areas more efficiently with more accurately diagnosing and less wasting in time of treatment while trying not to lose the bounder for working in real time standard.

Foot Drop affects many people and the main inspiration behind this thesis **applying minimum recording signals (EMG) to discover or detect Foot Drop correct as it occurs automatically.** Another inspiration is to find **the best set of input signals and features to detect Foot Drop. Also discover the best machine learning system to study Foot Drop?**

1.3 Contributions of the Thesis

This thesis contributes to knowledge theoretically and practically. The main objective of this research is to develop computational intelligence based EMG for automated classification of foot drop rehabilitation.

These contributions can combat problems that occur in real-time applications. To achieve the main purpose for this research several methods and algorithms have been examined with real-time. This thesis will present and discuss all theoretical and practical contributions and explain the acquirement briefly for collecting sEMG data signal from Foot Drop Patients in Rehabilitation Hospital in Sydney, Australia using Ethical Approval (UTS HREC NO.ETH15-0152).

The following are the main contributions of the work presented in this thesis:

- The first contribution proposed to develop the performance of the Self-Advised-Support Vector Machine (SA-SVM) by modified from single class to multi-class. Developments in methodology lead to more significant applications for overall use. (Chapter4).
- The second contribution proposed the novel a recognition system that is an integration of label classification methods with SA-SVM to get LSA-SVM for two-class classification. (Chapter4).
- The third contribution is to develop LSA-SVM from two classes into multi-classes and compared it with three sets of a dataset. (Chapter 5).
- The fourth contribution in this thesis also contains theoretical and practical structures, which describe the performance of engaging ELM method with LSA-SVM to get new method ELM-LSA-SVM. (Chapter5)

1.4 Organization of the Thesis

The main purpose of this research is to introduce novel classifiers to develop the performance of the M-PR for Foot Drop Rehabilitation devices, which race the problems that occur in the real-time application. To achieve the main object of this research proposes various methodologies and then examines them with the real-time application. This thesis presented and discusses all contributions, and explains the acquirement briefly for collecting sEMG data signal from Foot Drop Patients in Rehabilitation Hospital in Sydney, Australia using Ethical Approval (UTS HREC NO.ETH15-0152).

Chapter 2 this chapter discussed the Foot Drop problem and its related subjects such as Lower Limb Anatomy, diseases, and complications of the lower limb. Detection and of FD problem using sEMG characteristics electromyography reviewed. Microcontroller system and M-PR stages for leg movement recognition are studied. Furthermore, some conventional approaches and methods use the machine learning system, and their previous applications in Foot Drop are studied and reviewed. Also, functional electrical stimulation (FES) and most common rehabilitation devices used for FD patients are studied. At the end of the chapter, dynamic simulation software for different movement of the musculoskeletal system discussed, this simulation named Ope-Sim software, which is simulated and evaluate FD anatomy.

Chapter 3 addresses some Learning Machine methodologies that used in this research like SVM, SA-SVM, ELM, and Label classification methods. For these methods, a review of the Support Vector Machine, Self-Advise SVM, Extreme Learning Machine and Label Classification methods are studies to develop auto-control and a diagnostic system for foot drop.

Chapter 4 describes the data collected from the hospital based on the design procedure that got the ethical approval to collect data from FD patients in Metro Rehabilitation Hospital in Sydney, Australia using Ethical Approval (UTS HREC NO.ETH15-0152). Compliments the novel idea behind Myoelectric Pattern Recognition

In addition to the above **Chapter 4,** proposed a new based SVM classifier method named Label Self-Advised Support Vector Machine (LSA-SVM) and upgraded a Self-Advised Support Vector Machine (SA-SVM) to multi-classes to classify leg motion recognition. Overall, LSA-SVM methods could be classified as leg movements if it is healthy or not with an accuracy of 99.06 % comparable to known classifiers such as SA-SVM, SVM. Therefore, LSA-SVM could improve the performance of the advised-based SVM.

The implementation of *m*SA-SVM and LSA-SVM for Myoelectric leg motion recognition is a **contribution-1 of the thesis**. Other contributions that projected in the related section is *L*SA-SVM. The advantage of LSA-SVM over SA-SVM exists on working with Label data instead of data value itself, therefore, this is **contribution-2.**

Further to this Myoelectric leg motion classification, LSA-SVM has applied in a vast collection of classification problems using UCI machine learning data set. The experimental results show that LSA-SVM can apply to a wide range of data set from small to large-sized data. Overall, LSA-SVM is a promising classifier for several classification applications, particularly for Myoelectric pattern recognition.

Chapter 5 addresses the third and fourth contributions that had developed in this thesis. Firstly, LSA-SVM has modified from a single class to Multiclass; this development allowed further application in multi-label classification. Secondly, match the base ELM with LSA-SVM to obtain a novel method that was speedier, named ELM-LSA-SVM.

The real-time data experiments apply the faster method in the real-time application of Myoelectric pattern recognition (M-PR) for leg movement detection. The experimental results on the sEMG signal data and benchmark datasets exhibit its benefits as we can see in table 6.1 and 6.2 all kinds of results for all novel methods. Furthermore, the experimental results on UCI datasets indicate that ELM-LSA-SVM achieves the best performance when combined with LSA-SVM. The entire experimental results recorded and display in chapter 5.

To estimate the recommended *m*SA-SVM, LSA-SVM and ELM- LSA-SVM the experiment embrace 13 datasets from the UCI machine learning repository [15, 16]. These databases nominated from the most public benchmarks for classification and diagnosis. The diversity of these databases supports the authentication in this study. The numbers of instances and the attributes of each database had shown in the Tables in each chapters. It should be note that for a dataset with multi-class we used 6 datasets while for two classes were selected 7 datasets [17].

1.5 Publication

The following papers have published as a part of the result of the research presented in this thesis:

Conference

1- Adil, S., A.A. Jumaily, and K. Anam, "AW-ELM-based Crouch Gait recognition after ischemic stroke" in 2016 5th International Conference on Electronic Devices, Systems and Applications (ICEDSA), 2016. DOI

10.1109/ICEDSA.2016.7818552, ISSN 2159-2055. Which is partial publication in chapter 4 as contribution 1.

2- Adil, S., et al. “Extreme learning machine based sEMG for drop-foot after stroke detection” in 2016 Sixth International Conference on Information Science and Technology (ICIST), 2016. DOI 10.1109/ICIST.2016.7483378, Which has a partial publication in chapter 5 as contribution 4.

3- Hosseini, S., A. Al-Jumaily, and Sahar A. Abboud. “Active force control system for hand tremor suppression by different actuators.” in 2016 5th International Conference on Electronic Devices, Systems and Applications (ICEDSA), 2016. DOI 10.1109/ICEDSA.2016.7818487.

4- Marwa Ibrahim1, S.A.a.A.A.-J., Superimposed Signal Representation for Deep Learning in Biosignal Processing. The First MoHESR and HCED Iraqi Scholars Conference in Australasia, 2017.

Journal

Sahar Adil, Adel Al-Jumaily, “ Label Self -Advised Support Vector Machine (LSA-SVM),” Expert Systems with Applications. (under-reviewed in IEEE Journal of Biomedical and Health Informatics**), 2019.** Which is the main publication in chapter 4 as contribution 2 and chapter 5 as contribution 3.

Chapter 2

Overview of Foot Drop and Rehabilitation Techniques

2.1 Introduction

To understand how rehabilitation devices are helpful for patients with Foot Drop and how to improve their quality of treatment. First required to understand how the muscles of the lower limbs activate, function and review its anatomy. Then provide a brief sequential history of devices including the most recent devices and technological advances. Then, we must analyze the most critical methodologies that applied to these devices.

This chapter briefly covers concepts of the anatomy of the lower limb including the muscles involved in simple movements. The following section will present basic knowledge about the biological theory of Electromyography (EMG) and Function Electrical stimulation (FES). Other methods such as different Myoelectric Controller (MEC) techniques will also be examined. The chapter then presents a literature review on the Foot Drop (FD) rehabilitation devices using EMG and FES as their control. Finally, the chapter will highlight the limitations and disadvantages of the existing Myoelectric Controller and overcome those limitations and disadvantages with the subsequent research. In addition, functional electrical stimulation (FES) and most common rehabilitation devices used for FD patients are study. At the end of the chapter, dynamic simulation software for different movement of the musculoskeletal system will be discussed. Open Sim software simulation that simulates and evaluate the FD anatomy will review in this chapter.

2.2 The lower limb anatomy and bio-signal

2.2.1 The lower limb anatomy

The anatomy of the lower limb (Pelvic girdle, Thigh, Leg, Ankle, and Foot) needs to be reviewed to understand the structure of the lower limb and the correlated muscles so that the controller for the lower limb rehabilitation device can designed appropriately. The first section explains the bones, joints, and movement of the lower limb. After that explains, the muscles are motivating the lower limb.

2.2.1.1 Bones and Joints

The lower limb is attached to the body through the pelvic girdle Figure 2.1a. The Pelvic Girdle consists of two groups of bones: the right and left Coxal bones (hip bones), these join each other anteriorly and the sacrum joins posteriorly to form a ring of bone. The male pelvis can be differentiated from the female pelvis because it is usually larger, where the female pelvis tends to be more extensive, and the subpubic angle is higher in the female. This increased size of these openings supports and accommodates the fetus during childbirth.

The lower limb contains the four set of bones: thigh, leg, ankle, and the foot. The most significant notes for each set will be explained below [18]:

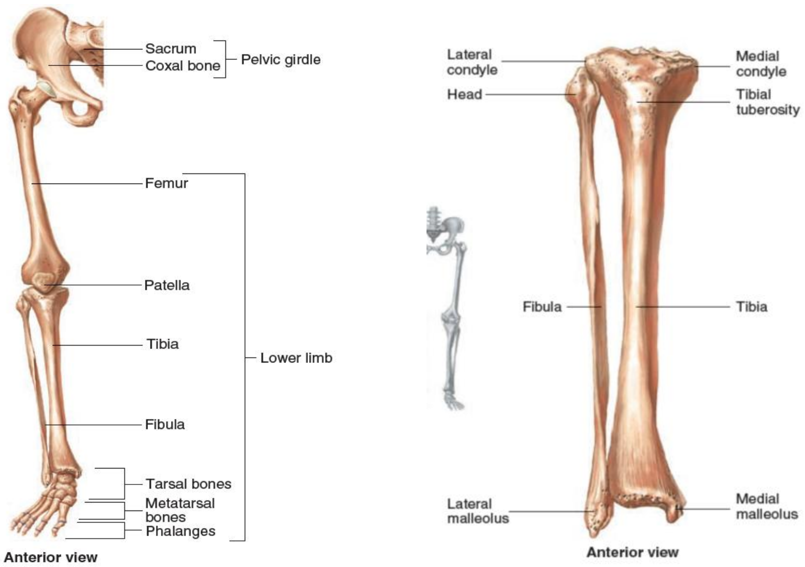
1- **Thigh:** the **thigh** is the area between the hip and the knee which contains one bone named Femur. The head of the Femur attaches (articulates) with the acetabulum of the hip bone. At the distal end of the Femur, the condyles attach with the Tibia. The Medial Epicondyle and Lateral Epicondyle which form the condyles and is the point of bone attachment. The Femur can be recognized by the humerus which consists of its long neck placed from the head to the Greater Trochanters. The Trochanters are points of muscle attachment. The Patella (kneecap) is positioned inside the major tendon of the anterior thigh muscles and supports the tendon to bend over the knee.

2- **Leg: T**he **leg** is the area between the knee and the ankle Fig.2.1b, and has two bones, the Tibia and Fibula. The Tibia is the large one and has the most weight which is the bearing bone of the leg.

The curved Condyles of the Femur reside on the surface condyles on the proximal end of the Tibia. Tibial Tuberosity is distal to the condyles of the Tibia, on its anterior surface, where the muscles of the anterior thigh attach. The fibula joined with the Femur where it is head connected to the proximal end of the Tibia. Distal ends of the Tibia and Fibula cover a part of the socket that matches with a bone of the ankle or Talus. A distinction can be seen from each side of the ankle known as Medial Malleolus of the Tibia and the Lateral malleolus of the Fibula.

3- **Ankle:** includes seven tarsal bones (sole). The Tarsal bones are the Talus (ankle bone), Calcaneus (heel), Cuboid, and Navicular, and the medial, intermediate, and lateral Cuneiforms. The Talus matches with the Tibia and Fibula to cover the ankle joint, and the Calcaneus covers the heel.

4- **Foot: T**he **Metatarsal bones** and Phalanges of the foot are organized and numbered similarly to the Metacarpal bones and Phalanges of the Hand. To some extent, the metatarsal bones are longer than the metacarpal bones, while the phalanges of the foot are noticeably shorter than those of the hand. There are three main A**rches** in the foot, shaped by the placement of the Tarsal bones and Metatarsal bones, and held in place by tendons. Two longitudinal arches cover from the heel to the ball of the foot, and a bridge-like arch covers across the foot. The arches utility is similar to that of the springs of a car, permitting the foot to stretch and spring back.



(a) (b)

Figure 2. 1 Bones and Joins for lower limb: (a) Bones of the Pelvic Girdle and Right Lower Limb, (b) Leg bones [18]

2.2.1.2 Types of Joint Movement

Types of movement at certain joints correlated to the structure of that joint. Some joints are restricted to only one type of movement, while the other joints permitted the movement to several directions. All kinds of movements are categorized depending on the anatomical position. Since most movements associated with movements in the opposite direction, they commonly demonstrated in pairs. **Flexion** and **Extension** are every opposite day movements that can be described in different ways and each description has an exception.

Flexion moved a part of the body in the anterior position from the frontal plane whereas extension moves a part in the posterior position from the frontal plane. An exception to this movement happens in the knee, where the flexion moves the leg in a posterior position, and extension moves it into the anterior position. **Plantar Flexion** is a movement of the foot at the plantar surface (sole) when the person stands-up on their toes. While **Dorsiflexion** is the movement of the foot toward the shin when the person is walking on their heels. **Abduction** is the movement away from the middle, or mid-sagittal plane and **Adduction** is the movement towards the middle plane.

An example of this is to move the legs away from the middle line of the body like the outward movement of jumping jacks is called Abduction, while bringing the legs back together is Adduction.

2.2.1.3 Muscles and nerves

The coordination of the interior and exterior muscles is needed to perform the movements of the leg. These muscles control the Flexion and Extensions of the leg. The leg is divided into four sections to describe the muscle compartments, and its actions in leg movement: Anterior, Lateral, Posterior (Superficial) and Deep Posterior sections Figure 2.2 [19].

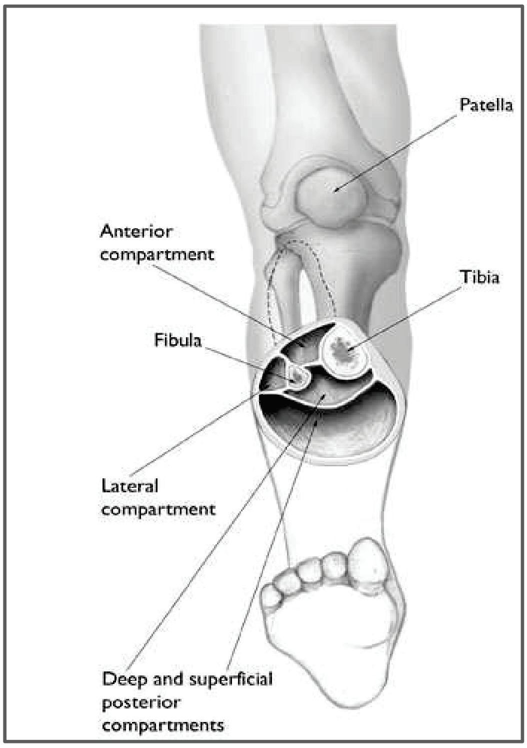


Figure 2. 2 Structure of the four sections of Leg [19]

- **Anterior**: Contains the muscles that Dorsiflex the ankle and extend the toes. It includes the Anterior Tibial Nerve and Anterior Tibial (AT) Artery which supply the AT with the blood. Innervation of this area has the Deep peroneal (fibular) nerve. The muscles of the Anterior are Tibialis Anterior (TA), Extensor Hallucis Longus (EHL), Extensor Digitorum Longus (EDL) and Peroneus Tertius (PT) muscles as shown in Fig. 2.3 [20].
- **Lateral**: Includes the muscles that turn foot outwards, and contains the superficial Peroneal nerve, there is no major artery in this area to supply the blood. So the sharp division that arises from the Peroneal artery is the division from Posterior Tibia artery. The nerve supply is from the Superficial Peroneal nerve (Fibular). The muscles for this area is Peroneus Longus and Peroneus Brevis.
- **Posterior (Superficial)**: Contains the Plantar flexion muscles, the Short Saphenous Vein, the Peroneal branch of the common Peroneal nerve, the Medial cutaneous nerve of the calf and Sural nerve. Moreover, it is innervat by the Tibial nerve. The muscles which included in this area are Gastrocnemius, Plantaris and Soleus muscles.

- **Deep Posterior**: Contains the Posterior Tibial Artery as its blood supply, the Tibial nerve to innervate the muscles and the muscles which are: Popliteus (Pop), Flexor Digitorum Longus (PL), Flexor Hallucis Longus (FHL) and Tibialis Posterior (TP).

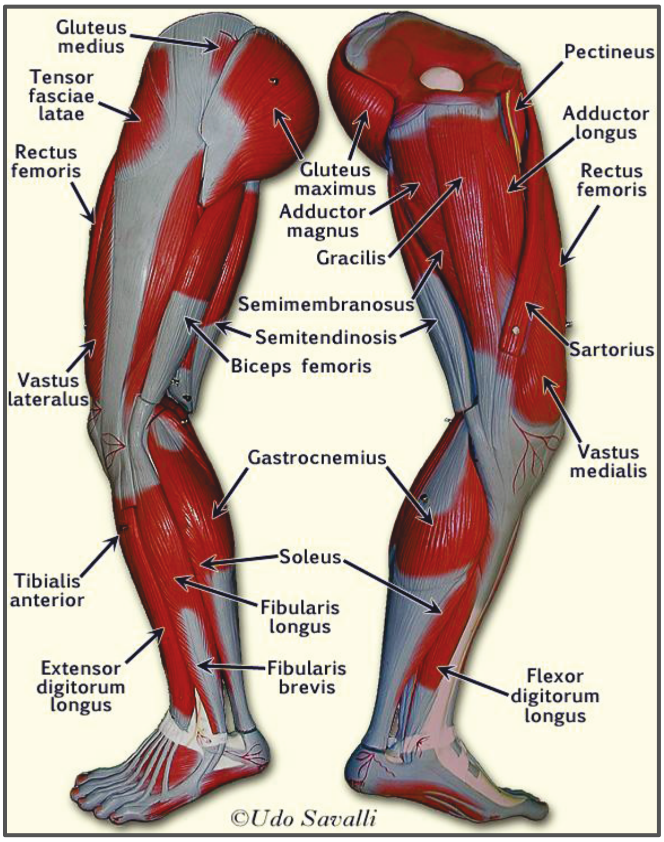


Figure 2. 3 The muscles shape and places [20]

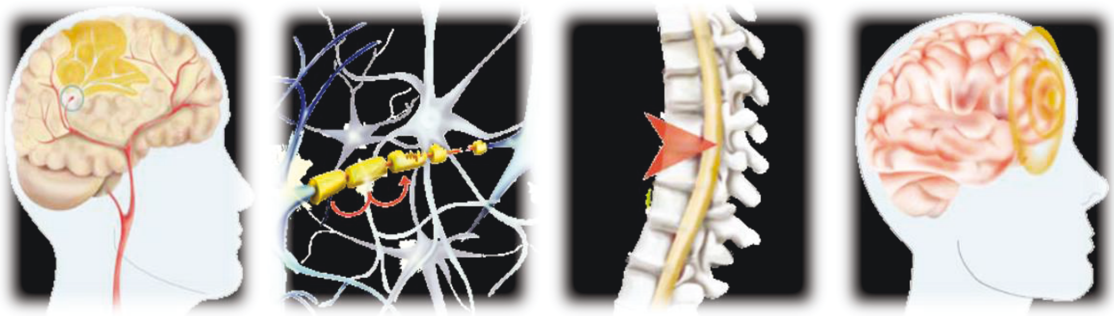
2.2.2 Lower limb diseases and problems

The diseases that affect the lower limb such as (Stroke, Multiple sclerosis, Incomplete spinal cord injuries, and Traumatic brain injuries) causes unexpected problems or loss of function. These problems such as Foot Drop will be explained in this section. This section will also explain how we can diagnose the Foot Drop problem.

2.2.2.1 Foot Drop (FD)

Drop Foot or Foot Drop (FD) is the disorder of the stimulus in muscles in the lower limb or Leg and ankle. FD leads to the weakness or paralysis of the lower limb and the patient being unable to lift the foot correctly. It disrupts one or both feet causing the toes to drag during gait. To avoid the dragging of the toes, patients will habitually lift their knee higher in the affected leg while they walk. This gait pattern puts them at risk for falling or infection due to loading on the knee joint which is sensitive. This problem results from damage to the central nervous system [21, 22], such as:-

1- Stroke;
2- Multiple sclerosis (MS);
3- Incomplete spinal cord injuries;
4- Traumatic brain injuries as shown in Figure 2.4.



(1) Stroke **(2) MS** **(3) SCI** **(4) TBI**

Figure 2. 4 Diseases causes Foot Drop

2.2.2.2 Diagnosing Foot Drop

Foot Drop gait can categorize as slow action of muscles and unsynchronized lower limb movements. Also, FD defines as residual muscle weakness, abnormal move interactions and spasticity effect on changing the gait forms. These problems lead to reduced body balancing, more risk of dropping, also increase energy spending through walking. FD gait is asymmetric with reducing involve lower limb range of motion (ROM) and walk speed. An Equinovarus Foot (clubfoot feet) distortion exists and compromising heel strike while person's walk. They also, describing that a normal function of the foot and ankle as a group of three serial rockers: the heel, ankle, and forefoot [23].

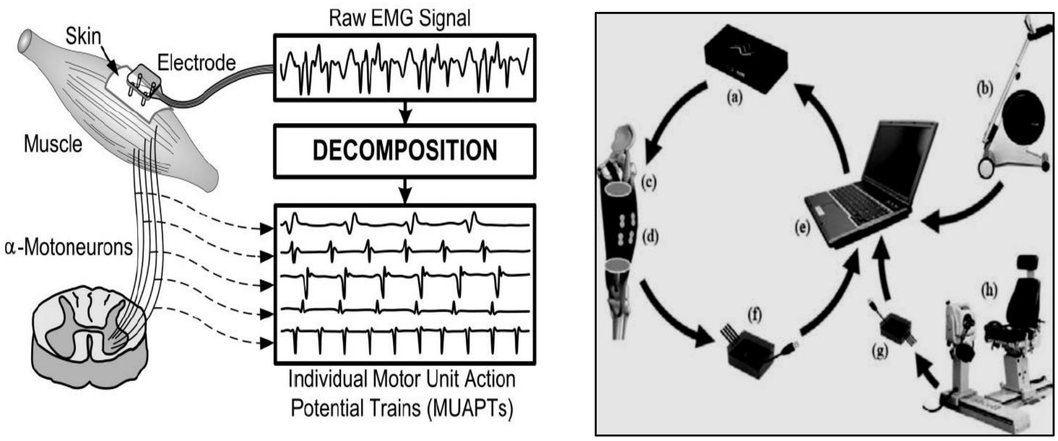
Electromyography devices (EMG) is used to diagnose, detect, and analyze the electrical signal created during muscle contraction [24]. This signal is known as an Electromyogram (EMG) or Myoelectric signal (MES) [25]. The Anterior and Exterior muscles motivate the leg and the Foot. These muscles generate electrical signals called Myoelectric signal (MES) and measured by Electromyography (EMG) devices.

2.3 Electromyography (bio-signal)

This section presents the way of detecting FD problems using EMG then clarifies the characteristics of the EMG signals, then the methods that record these signals.

2.3.1. Detection using EMG signal

Estigoni EH et al.[26] designed the UniSyd e2MG system as a multipurpose signal completion stage for both evoked and controlled EMG's collected during clinical studies. The e2MG can record the analog and serial data for other devices as well. Thus it will give it the ability to integrate with other medical devices. It is produced to deliver an easy path for selected data, clearing it from errors, outside interfere or wave artifacts. It can also produce a correct synchronism between the (9) available channels. Figure 2.5 shows the UniSyd e2MG components, where Figure 2.5a appears a graphic outline of the decomposition surface EMG signal into its essential motor unit action capacities [27]. Figure 2.5b describes the UniSyd e2MG system where (e) is Purpose-designed software, (a) is controls the neuromuscular stimulator and (f) reading data from the EMG electronic module, (g) from the second module that connects to (h) which is a Biodex dynamometer and from (b) a leg cycle ergometer. While (c) is skin surface electrodes are used to evoke muscle contractions and (d) to collect EMG signals.



(a) **(b)**

Figure 2. 5 Design UniSyd e2MG system: (a) Graphic outline sEMG signal into motor unit action capacities [27] and (b)The UniSyd e2MG system

2.3.2 Characteristics of the Electromyography signal

The EMG signal is a random signal that has an amplitude bounded from (0 -1.5 mV) as the root mean square (RMS), i.e. (0 to 10 mV) peak-to-peak (PP). The energy over the electrical noise level ranged the frequency from (0 – 500 Hz) while the energy of the noise ranged from (50-150 Hz) [28]. There are several sources of the generated noise signal, for example, ambient noise, motion artifacts, the inherent instability of the signal, and inherent noise in the electronic components. The energy under the noise level is not acceptable to analyze. Figure 2.6 displays the EMG signal and its power is taken from the Tibialis anterior (TA) muscle during continuous force and equal contraction at 50% of voluntary maximum [29].

Surface EMG (sEMG) signals are filtered digitally at the bandwidth (20–400 Hz) to increase slow transition and instrumentation noise and divided into the epoch of (250 ms), and this signal is supposed to be motionless. For each column of the 2D array, the central innervation zone (IZ) recognized by visual analysis of the signals. Channels near the acromion show the spreading of the signals selected. Each epoch of each signal is utilizing to estimate the EMG variables. Root mean square value (RMS), mean frequency of the power spectrum (MNF) and the fractal dimension (FD) helps to estimate the use of the box count method, muscle fiber conduction velocity (CV). Also, estimation using a multichannel algorithm, values accepting of the average correlation coefficient between adjacent channels is upper more than 0.75 and entropy [30].

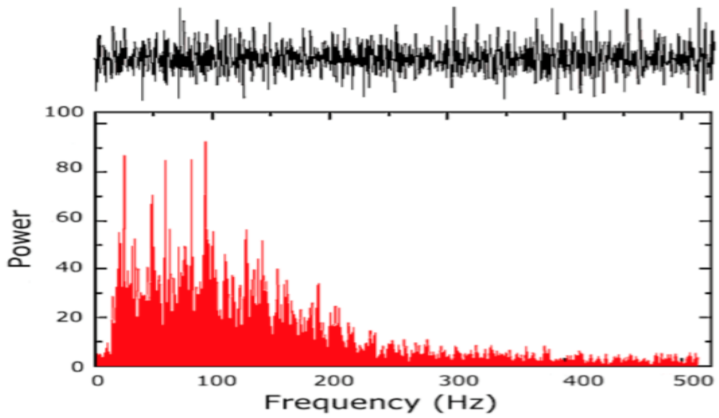


Figure 2. 6 EMG signal and power taken from the TA muscle [29]

2.4. Myoelectric control system

The EMG signal applied as a control system on leg rehabilitation devices. There are two groups of research: clinical and kinesiology. In Clinical research, diseases such as Amyotrophic Lateral Sclerosis, Polymyositis, Carpal Tunnel Syndrome, and Bell's Palsy can be diagnosed using EMG signal; it can also recognize muscle weakness and muscle atrophy and distinguish between neurogenic or myogenic changes. Kinesiology research uses the EMG signal to record the reaction through motivate body, by classifying if the muscles activity is existence or nonexistence, measuring the amount, time and period of muscle activity. The base of biofeedback to patients, find materials relative to muscle fatigue, measure response times on external stimulation, estimate motor control, and to evaluate gait pattern [31].

The EMG signal can use as a control system for the lower limb rehabilitation devices either in EMG-based pattern recognition or non-pattern recognition system [32]. The EMG-based pattern recognition (EMG-PR) or myoelectric pattern recognition (M-PR) consists of a few stages as presented in Figure 2. 77. M-PR is using to classify and recognize the EMG patterns into classes or limb movements. While the EMG-based non-pattern recognition (EMG-non-PR) system could not classify any limb movement, it can use EMG-non-PR signals as a threshold control system or a corresponding control system. The following sections will describe M-PR.

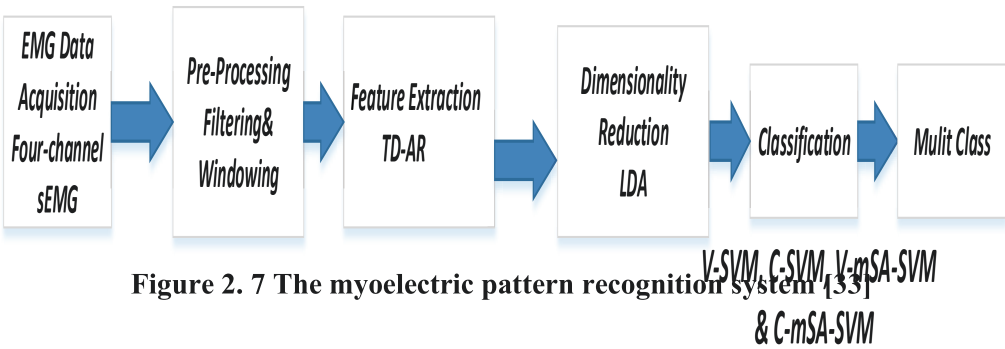


Figure 2. 7 The myoelectric pattern recognition system [33]

2.4.1 Myoelectric pattern recognition (M-PR)

Pattern recognition is a scientific self-control whose object is to classify the data or groups. Pattern recognition is an integral part even in the most fantastic machine intelligence system constructed for decision making. Pattern recognition contains methods which drive the improvement of several applications in different files. These methods can be applied in intelligent artificial applications, such as

Statistical Pattern Recognition, Data Clustering, Syntactic Pattern Recognition, Neural Networks, The application of Fuzzy Sets, Applications of Support Vector Machine (SVM) for Pattern Recognition e.t.c. The technology of PR is applicable in many areas such as artificial intelligence, nerve biology, medicine image analysis, computer engineering, archaeology, geologic reconnoitering, armament technology, space navigation, Computer-aided diagnosis (Medical imaging, EEG, EEG signal analysis), and Bioinformatics (DNA sequences analysis).

This section will address the parts of Myoelectric Pattern Recognition (M-PR) step by step by to clarify each stage of M-PR as shown in figure 2.7

2.4.2 Pre-processing

2.4.2.1 Filtering

To minimize and distract the noise signal from the original signal, we used and adjusted Filters like Band-Pass Filter (BPF) between (15-500 Hz). The power noise signal between (50 Hz and 60 Hz) can be withdrawn by applying a notch filter of (50 or 60 Hz). However, this filter is likely to remove the original signal of the EMG, hence leading to the application of differential amplification.

2.4.2.2 Windowing

The EMG signal can be windowing into two procedures: Disjoint or Overlapped windowing [34, 35], seen in Figure 2.8. Disjoint windowing is related and adjusts the window length. Whereas, overlapped windowing is related and adjusts the window length and window increment. The window increment is the band between two sequential windows. Typically, the Disjoint windowing is the same Overlapped windowing in the case where the window increment is the same as the window length.

Windowing can be defined as a procedure to extract essential data from the myoelectric signals. This information data named as feature data. Applying this feature helps to evaluate the performance of the recognition system [36]. Once the feature is extracted in a data sequence, limited in its time-space or window, then all parts of the recognition system are performed.

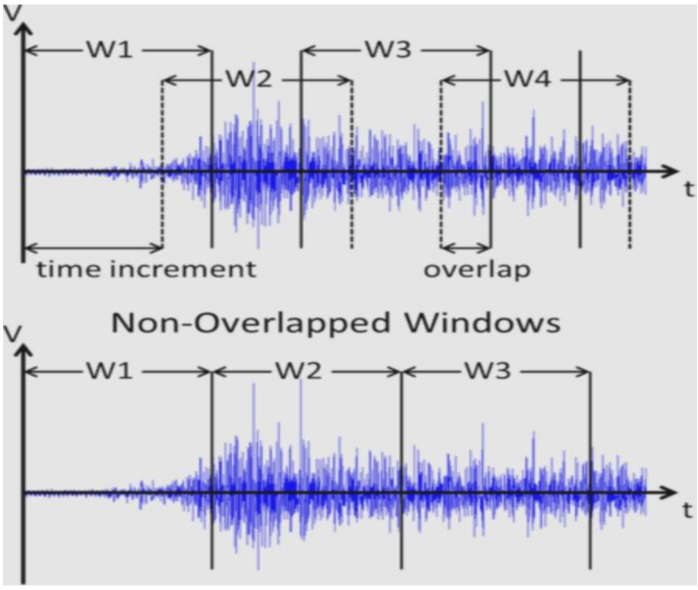


Figure 2. 8 Signal processing: time window. Myoelectric signal windowing by time windowing including overlap or non-overlapped windows. [37]

2.4.3 Feature Extraction

Feature extraction is the next stage using at Myoelectric Pattern Recognition. The feature extraction is a process that transforms patterns into features. For EMG signals, transferring the pattern of EMG signals (windows) to a set of features that covers significant features of the signals [25]. Feature extraction methods address to find the most compatible and informative sets of features and of which precisely define the input signal data in an abbreviated character. Khushaba et al. [38] appropriation of the feature set in EMG-based control, that it must have a feature space last class separate. It also must be robust in a noisy environment and should be less complicated in calculations. Usually, feature extraction with EMG signal contains three stages: Time Domain (TD) features, Frequency Domain (FD) features and Time-Frequency Domain (TFD) or Time-Scale Domain (TSD) [39]. The next section will clarify the TD feature which applied in this thesis.

2.4.1.3 Timing Domain (TD) features

The eight most commonly recommended TD features with the greatest effectiveness for EMG pattern recognition considered in this thesis. These features extracted within an N-sample evaluate time window.

The usefulness of adding the TD features to the procedure was:- easy and quick in running implementation program, less computing complexity, and perfect performance for less noise environment [39]. Conversely, the disadvantage for using TD feature was dealing with non-stationary signals like EMG signals. Some of the most used TD features [40]. Their equations will be described briefly in chapter 3.

2.4.4 Feature Reduction

The results extracted after feature from EMG channels are assumed to be large feature sets. So that, to dominate these sets must be decreased without damage or change the main features. In general, there are two methods of dimensionality reductions: feature selection (FS) and feature projection (FP).

Commonly, the feature projection (FP) with unsupervised methods such as Linear discriminant analysis (LDA) and Principle component analysis (PCA) can be used to extract features from a large data sample. Auto-Encoder is another remarkable process to reach the same purpose in the case of Deep Learning. Recently, the auto-encoder and its deviations have received attention as nonlinear dimensional reduction feature. The predictable Auto-encoder will help to approximate the data to be unique by bushing the output data to be similar to the input data [41].

Frequently, an Auto-encoder adopted as a simple measure to reconstruct a deep structure to help in structural feature learning; limits placed on the parameters for model training. Sparse Auto-Encoder (SAE) was projected to face the average response of each hidden unit to a minimum result. Yu et al. [42] projected a graph normalized auto-encoder, which targets to adopt a graph to control both encoding and decoding. Nevertheless, the irrelevant patterns in the data are still interesting to learn. The existing auto-encoder variants have not been considered in the hidden units and can divide into two stages: the task-relevant and the task-irrelevant.

2.4.5 Class Classification

Classification stage for a diagnosis system is the stage that is responsible for doing the readings near the extracted data, before this in order to produce a diagnosis at the input signal. The Classification requires a classifier to trained and tested, and then this can use to classify a new image. The classification encases two stages: training and testing.

The training stage: a classifier attracts to a trained feature, where the training feature vectors have an assigned label. While at the testing stage, the classifier tends to test the feature vectors that have been kept for the test purpose and estimates the performance of the classifier. Typically, there are two types of classification methods: Class Classification and Label Classification methods. Class Classification can divide into two types depending on how many classes the signal data contains: Single Class Classification (SCC) and Multi-Class Classification (MCC). Label Classification methods can also be divided into two types depending on the set of data associated with the label: Single Label Classification (SLC) and Multi-Label Classification (MLC). In this section of the chapter, we will briefly discuss all these types[43]. Table 2.1 shows the advantages and limitations of MPR system stages.

Table 2. 1 Advantages and limitations for MPR system stages

Methodology	**Advantages**	**limitations**
Filtering BPF	To minimize and distract the noise signal from the original signal	this filter is likely or sometime to remove the original signal of the EMG
Windowing	The EMG signal can be windowed into two procedures: Disjoint or Overlapped windowing [34, 35]	
Feature Extraction	To find the most compatible and informative sets of features and of which precisely define the input signal data in an abbreviated character [38]	it must have a feature space last class separate also, must be robust in a noisy environment and should be less complicated in calculations.
Timing Domain (TD) features	easy and quick in running implementation program, less computing complexity, and perfect performance for less noise environment [39]	dealing with non-stationary signals like EMG signals

Feature Reduction	to dominate these sets must be decreased without damage or change the main features [41].	The results extracted after feature from EMG channels are assumed to be large feature sets
Classifications	a diagnosis stage that is responsible for doing the readings near the extracted data, before this in order to produce a diagnosis at the input signal [43].	The Classification requires a classifier to trained and tested, and then this can use to classify a new signal.

2.4.6 Single-Class Classification (SCC) and Multi-Class Classification (MCC)

This classification method can be distributed into binary (single) methods and multi-class methods. Several classifiers such as Support Vector Machines (SVMs) and Self Advise SVM (SA-SVM) work with the first case and the second case called Multiclass support vector machines (MSVMs). Extreme Learning Machine (ELM) is another method that deals with single and multiclass. This thesis will work with SVM and develop SA-SVM [13, 44] while applying ELM in the experimental part to compare and evaluate the results. In the offline mode, the classification careers the aforementioned optimal parameters.

Other methods have been used by supervised machine learning methodologies to estimate the performance of learning systems such as measure the quality of classification that based on the “confusion matrix” which proceedings correct and incorrectly recognize labels for each class. Estimating the classification performance without any concern for a class is the most common and the most use experiment measure where accuracy did not separate between the numbers of correct labels of different classes.

2.4.6.1 Support Vector Machines (SVMs)

Support vector machines (SVMs) are a machine learning model based on arithmetical learning theory [10]. For a long time, support vector machine (SVM) has widely used in classification applications because of their astounding classification proficiency and its variations. SVM have two keys of learning feature: a first feature where SVM over a nonlinear feature mapping function φ(x), the training sample data first mapped into a higher dimensional feature space. Second feature: the typical optimization technique is applied, to discover the explanation by maximizing the separate margin of two various classes in the feature space, leading to minimizing train errors. Support Vector Method has been used widely to solve regression problems due to their introduction of the epsilon-insensitive loss function [14, 45]. While the simple training algorithm can be a paradigm for linear separators, multiple kernel functions such as Sigmoid, Linear, Polynomial, Radial Basis Function (RBF) comprises of variable degrees of non-linearity and flexibility in the learning system.

As mentioned above the most traditional classifier, the SVMs have some advantages such as decision trees and neural networks. The support vector training primarily comprises optimization of a curved cost purpose. Hence, there is safety to being stuck at a limited minimum in the background propagation neural networks. SVMs constructed on the Structural Risk Minimization (SRM) principle, which reduces the upper bound on the generalization error. As a result, SVMs are less affected by over-fitting, when compared to other algorithms like backpropagation neural networks it applies the empirical risk minimization (ERM) principle.

SVM projects a classification methodology by creating N-dimensional hyperplane, which separates the data perfectly into two or more classes. The SVM model aims to find the perfect hyper-plane [46] which can separate clusters of vectors in a method. In cases where one class of the target variable is on one side of the plane and or with another class are on the other side of the plane. The support vectors (SV) are the vectors near the hyperplane.

Many Bio-medical applications have applied SVMs techniques to classify data or to diagnose diseases. For example by developing an automated diagnosis system for skin cancer [11] and detection and prediction of sleep apnoea [12]. Another area of application is the National Health and Nutrition Examination Survey (NHANES). NHANES used probability data sample of US residents to construct SVM and Logistic Regression classification models for two classification systems: 1) people who have diabetes (diagnosed or undiagnosed) versus people do not have diabetes, 2) people who have undiagnosed diabetes or primer diabetes versus people have not had diabetes. They utilized 14 probable predictors regularly related to diabetes: family history, age, gender, household income, etc. The results appear that the Radial Basis Function (RBF) kernel and Linear kernel operated with the highest efficiency for the classification system 1) and 2), with no difference between Logistic Regression and SVM performed (AUC 0.83 and 0.73 for classification 1) and 2) in two systems. The highest efficiency observed with the SVM model for classification of diabetes in the Pima India Diabetic Dataset from the UCI Machine Learning Laboratory.

Other applications include eight predictions SVM model; laboratory sample data count (plasma glucose concentration, 2 hours' serum insulin) corroborated with 78% accuracy utilized the RBF kernel. SVMs apply across different biomedical classification problems that comprise a patient financial risk model with health statements and clinical encounter data. Also, patient reply to infection consciousness model together applies weighted SVM. The project is comparing different machine learning techniques with Logistic Regression for the prediction of heart disease. It seems to be no significant difference between Logistic Regression and SVM, with the best Linear kernel performing [45, 47]. There are some reasons for using based SVM instead of other methods like Self-Organizing Map (SOM).

The difference between self-organizing map with SVM and SA-SVM that is: Self-Organizing Map (SOM) is an unsupervised classification method based on neural network computing techniques. SOM is base on competitive learning. For competitive learning neuron activation is a utility the space between neuron weight and data input. An activated neuron learns the most and [48] its weights are thus adapted. If a similar pattern founded again, then the same neuron may be activated again. That means that a particular neuron repeatedly wins [49]. While Support Vector Machines (SVM) and Self-advised-SVM supervised classification learning method is self-training. Considering a binary separation problem, the basis of the SVM is mapping the input samples into a high dimensional space and finding the hyperplane that optimally separates the two classes or more classes. The optimal separating hyperplane is chosen following the criteria of maximizing the distance to the closest training instance. Whereas Self-Advised SVM treaties with the neglecting of SVM from the knowledge that can be obtained from misclassified data.

That done by creating advise weights depend on the distance for the, misclassify train data from the correctly classify train data, and through the application of these weights together with decision values of SVM in the test phase — these weights assistance the procedure to reduce outlier data [11].

2.4.6.2 Extreme Learning Machine (ELM)

Huang, Zhou [50] , Presented Extreme Learning Machine (ELM) as a simplification of Single-hidden-Layer Feed-forward Networks (SLFNs) which liberates the iterative tuning by the estimation of weights in each of the hidden and output weights. Furthermore, in a training procedure, the purpose of the ELM is to get the minimum training error and the minimum norm of output weights, which is unlike the original learning algorithm of SLFNs. For these reasons, the training speed is much more rapid when compared to the original SLFNs. Some experiments are conducted to compare classification performances between different types of ELMs, which refers to the accuracy of classification. Furthermore, an evaluation with other renowned classifiers, such as the SVM, SA-SVM, and LDA sees. The SVM and the kernel-based ELM require optimized parameters, where, the RBF kernel involves the cost parameter C and kernel parameter γ. Huang, Zhou [50] Indicated that the of (C,γ) significantly affects the performance of the ELM and SVM

2.4.7 Label Classification methods

The object of this section is to introduce the problem, give a brief overview of its main application and explain its application to the experts. Section 3.5.1, provides a formal definition for Multi-Label Classification. The main areas of application for the Multi-Label Classify are present in Sections 3.5.2. Section 3.5.3 introduces the methods encountered with this issue.

The Multi-label classifying is a predictive data mining application with multiple applications in the world, including automatic labeling of resources like videos, music, images, and texts. The Multi-label data can use as a learning tool and achieved by different methods, such as:

1- Problem transformation method which has three common methods:
  - Binary relevance method (BR).
  - Binary Pairwise classification method (PW)

- Label Combination or label power set method (LC)

2- Adaptation method.
3- Ensembles of classifiers.

In this chapter beginning with defining the Multi-Label classification problem and introduce the mathematical equations that will use in this research thesis. Then it will outline the newest applications that can apply to the class of Multi-Label Classification. The current method of learning for the Multi-Label data is different and include the three methods described above (transformation, adaptation and ensemble classification methods) [51].

2.4.8 Comparison between Multi-class and Multi-label classification

To understand the difference between multi-class classification and multi-label classification, we have to make a comparison:

1- The main differences between Multi-Class classification (traditional) and Multi-Label classification are the output estimated from a trained system. While the traditional classifier will produce one value, the multi-label classifier gives a vector of output values.
2- Traditional single-label classification is interested in learning from a set of data that is related to a single label=1, from a set of separate labels L, $|L| > 1$.
The learning problem is called a binary classification problem (or filtering in the state of textual and web data) when $|L| = 2$, while a known multi-class classification problem when $|L| > 2$ [52].
3- In Multi-Label classification, the data related to a set of labels $Y \subseteq L$. First method for Multi-Label classification involved by the tasks of text categorization and medical diagnosis. For example, in medical diagnosis, the patient might have diabetes and prostate cancer at the same time.[53] Another example of this method shown in the image Figure 2.9 where Figure 2.9a can mark “sea water” and “seabird.”
4- Multi-Label classification deals with data or samples that have multiple class labels at the same time. Each sample is signified by just one single instance. Multi-Label classification has concerned a significant interest from scientists, and it is applied in many areas such as text classification, image marked, video

annotation, bioinformatics, social network and music emotions categorization [54].

5- Multi-Label classification, labels usually links with each other. It has shown to take advantage of label connections. Labels can develop the performance of the classifiers, for example; if one image marked with label "sailing boat," it has a great possibility to assign the label "sea water" as seen in the image of Fig.2.9 (b).

(a) (b)

Figure 2. 9 Two nature's scene image example (a) Seabird Seawater, (b) Sailing and seawater.

2.4.9 Applications of Multi-Label Classification

As indicated in the previous section, the goal of the Multi-Label classifier is to predict a set of appropriate labels for a new instance of data. This section describes numerous application areas that can take advantage of this functionality, and give details of the specific use in areas each fits.

2.4.9.1 Text Classification

Textual documents founded in everywhere; these include all types of agreements and reports, people who have shared their invoices and electronic mail messages. It can also found in our historical medical records, published books and magazines, blog posts, articles in electronic media, and many others. Most of them can classify into more than one category, thus the efficacy of Multi-Label classification to achieve this type of work.

The typical process to transform a set of text documents into Data Mining (DM) depends on the text mining systems. The sourced documents are analyzed, and useless words removed then computed vectors applied with each word incident among the documents. A row describes the last form for each document in the Multi Labelling Dimensionality MLD, and the columns related to words and their frequencies or other types, such as TF/IDF (*Term Frequency/Inverse Document Frequency*).

2.4.9.2 Labelling of Multimedia Resources

While textual documents cover only text, it was used widely previously, images, music, and videos are applied resources in recent times due to the evolution and progression by storage and communication technologies. In 2015, YouTube data assessed 432 000 new videos that uploaded each day. Multi-label classification used to label and identify resources in the clip, recognizing the objects, which appear in sets of images, the emotions created by music clips, or the ideas, which can be derivative from video snips. With this method, a large number of new resources can acceptably classify without human interaction leading to being more cost-effective.

Although the image can represent as a vector of color values, with each pixel as a column, they are pre-proceeded in order to obtain the most convenient features. Pre-proceeded by applying windowing techniques to extract boundaries the image can be difficult (convoluted) to transform the pixel space into an energy space and so on. Similar actions applied in practice with other multimedia resources [55]

2.4.9.3 Genetics/Biology

Lately, Proteomics and Genomics appear as a study field. With this study, huge databases with protein sequences have created. The only a small fraction of these has considered defining their function. Analyzing protein properties and gene expression is an expensive task, but Data Mining (DM) [56] methods can accelerate the procedure and make it inexpensive. The application of Multi-Label classification in this region intends to predict the biological functions of genes. Genes can express more than one function at the same time. Therefore the importance of using Multi Labelling Data (MLD) techniques is to express this problem. The features used as predictors and these are usually the protein's *motifs*, which distinguish their internal structure. Structural motifs point to how the element in the second structural layer linked. Specific patterns, like short chains of amino acids, can be recognized and used as predictive features.

2.4.9.4 Other Application Fields

Other applications for where MLC is applied are the privet and public data-sets. It can be applied to specific applications as shown in this thesis. The most interesting ones are as follows:

1. The study of nonverbal expressions in a speech where the emphasis points to detect how people feel. The aim is to predict several not mutually exclusive emotional states at a given moment. The predictive features recognized by a set of vocal features, an example of which would be intonation and speech rate.
2. The researcher offers a system that automatically classifies patient records. The problem that needs to addresses required for assist of texture documents according to the creative ideologies.
3. The technique presents to improve the progression of the search for related information on Twitter. Many labels are defined to categorize Tweets, including news, opinions, and events.
4. A system that propose to analyze complex events in motion, it combines the track and Multi-label hypergraphs of moving targets in video sequences, detecting the connection between the low-level topographies and high-level impressions which predicted.

Overall, Multi-label classification is appropriate for any situation whatever its type, as long as processed by an MLC system [57].

2. 4.10 Learning from Multi-label Classification

The procedure to achieve an MLC model is similar to that applied to the MC model (traditional), many algorithms based on the initial training phase. It depends on samples that have already labeled to regulate the parameters of the model. When analyzed the model, it can be utilized to expect a label set for new samples.

The Six main approaches that applied are:

Data transformation: where applied to the original Multi Labelling Dimensionality (MLD) and was able to produce one or more binary or multi-class data sets. These can process with traditional classifiers.

1- Method adaptation: aims to adapt existing classification procedures, then they can execute Multi-label data and create several outputs instead of only one.
2- Ensembles classifiers that arise the data transformation.
3- Label correlation information.
4- High-dimensionality problem.
5- Label imbalance.

Label ranking carefully associated with Multi Labelling Dimensionality (MLD), which aims to map each sample in a set data to a set of labels thus creating a total order relationship between them. That is mean; the output of these algorithms is a ranking of labels, where each one assigned to a definite weight. MLC use label ranks to select which labels are significant to the sample [58].

2. 4.10.1 The Data Transformation Methodology

MLC is a more complex technique than traditional classification because the classifier is required to predict multi outputs at once. First step to solve this problem is to break it down into multiple stages making them more straightforward to deal with. The idea behind the transformation is to get Multi Labelling Dimensionality MLD and create datasets that can be processed by binary or multi-class classifiers.

Some methods proposed in the literature are summarized below:

1. **Ignoring multi-label instances**: By dismissing all the samples of data that related to more than one label; this transformation gets a new data set of original multi-class, i.e., just one label per sample. They familiarized it as *Model-I,* where *i is* for *ignore*. The output dataset will have a smaller amount of samples than the original one, except none of the samples required two or more labels. Meanwhile, if there is only one label per sample, then any multi-class classifier can apply [59].
2. **Using the label set as class identifier**: named as a **Label Power Set (LP),** this method clarify that, instead of neglecting labels or sample data, the *Model-n*, where *n* represents for *new class*, it recommends using each of the different combinations of labels (each label set) as an identifier of a new class. The output data set will give

the same number of samples, but only one class. It can address with any multi-class classifier.

3. **Selecting a single label**: The transformation named Models, where *s* represents a **single label**. When a sample data related to a set of labels, one of them selected as a single class. This choice can be random or based on some experiential method. The output by this transformation is a multi-class dataset with the same number of samples than the original one, and each sample has only one class.
4. **Unfolding samples with multiple labels**, named as **Problem Transformation 5 (PT5)**: this technique decomposes each sample into as many samples as the number of that labels it contains, duplicating the input features and allocating it to each sample of the labels. Depending on its distribution on the MLD, there is an option to allocate the weight parameter to each label. The result of this transformation is also a multi-class dataset, and this method offers more samples than the original [60].
5. **Applying banalization techniques**, named as **Binary Relevance (BR),** banalization methods deal with multi-class classification using binary classifiers were select for multi-label classification. The method involves training *k* classifiers, one for each label, taking the samples where the labels appear as confident and all the others as negative. Another application known is as ***cross training*** which trains *k* classifiers, preferably used for samples with multiple labels and always result as a positive.

2. 4.10.2 Adaptation Methodology

Some classification models are primarily intended to resolve binary problems and then expand to solve multi-class problems. SVM is an example of this. Other methods can work with several classes easily such as a kNN classifier which capable of predicting the regular class amongst the neighbors of the new data. Similarly, trees, neural networks, classification rules, and statistical models, have applied to look over both binary and multi-class classifications [61].

The methods learned from adapt of binary classifiers to the multi-class state are voting methods, penalization and divide-and-conquer processes these are valuable while adapting proven algorithms to look over multi-label classification.

A kNN-based system can take the multiple outputs that exist in its neighbors to make an accurate and specific prediction, but the SVM must find the maximum margin boundaries between multiple labels.

2.4.10.3 Ensembles of Classifiers

Classification ensemble is a well-known method designed to develop the performance obtained by individual classifiers. An ensemble increased by a set of classifiers, whose output data commonly combined by weighted or unweighted average. The critical thing in this technique, that the group of weak classifiers, with different biases, is capable of achieving best results than a strong classifier.

The most significant that difference between individual classifiers and the higher expectation of different biases. Variety can be reached through training similar models with different data, for instance with bagging techniques, and or by using mixed classification models. Additionally, ensembles are typically applied to assist with problems encountered such as imbalance. Therefore, it expected that ensemble methods also applied in the multi-label arena.

2. 4.10.4 Label Correlation Information

Numerous amounts of research for the multi-label classification system depends on making the original problem simpler, produce a more natural outcome. This process commonly presumes complete independence between the labels. However, most published research focusing on the significance of label dependency information. These correlations assist in designing the best classification models. The BR transformation method and many algorithms based on it build an individual classifier for each label.

A second mechanism is to combine label dependency information into the classification process indirectly. The LP transformation is another method based on taking complete or partial label sets as class identifiers. Where a group of labels treated like an item and the dependency between them is indirectly embedding into the model at the training process. Chains of classifiers is a similar method, but it depends on binary classifiers instead of multi-class. This technique presents the label predicted by one classifier into a sample dataset, which is an input to the next one [62].

2.4.10.5 High Dimensionality

High dimensionality encountered interrupts multi-label classification more than the traditional classification and causes obstacles. A three-dimensional space theorized:

a- **Instance space:** Multi Labelling Dimensionality (MLD)s contains millions of samples and continues to increase in all contexts, as well as traditional classification, due to the increasing data load.

b- **Input feature space:** two and multi-class data sets with many input features, where MLDs have thousands of features. The dimensionality of this space is made difficult for the learning process.

c- **Output attribute space**: The first method for multi-label classification works which consists of one output feature. By contrast, all MLDs have several outputs, and many of them have thousands, one feature for each label. For this reason, the dimensionality treatment in MLDs is a more complex topic than in a traditional one.

There are non-supervised techniques, such as Principal Component Analysis (PCA) and Latent Semantic Indexing (LSI), which is capable of minimizing the dimensionality by analyzing only the input attributes. These techniques can be utilized to MLDs since there is no need for knowledge about the output variables. The supervised methods characterized by the use of output features to understand which input attributes are more valuable to the information and if it has to be adopted earlier to use in the multi-label context.

2.4.10.6 Label Imbalance

Imbalanced data is another method, which closely linked to multi-label classification. There are many demonstrated techniques to overcome the imbalance in binary classification, and with multi-class datasets. The differences between label distributions in MLDs arise naturally. In Multi Labelling Dimensionality MLD with thousands of labels, for a sample classifying text documents, it is normal that some of them are much more normal. Documents are specifically would have infrequent labels, which will be in the minority against the respective ones. Besides, the transformations applied to multi-label data leads to maximize the imbalance between labels.

The example of BR. Taking into account the positive sample only, where the specific label appears, and all other samples are negative, the original distribution verified affects the symbol of the considered label. The situation is even worse if this label is already a minority label [63].

2.4.11 Post- Processing

Scientists verified that a classifier that integrates the decisions of some classifiers could outperform the best classifier. Integration techniques extend to a wide-range. Ranging from some simple integration rules, such as a majority vote and sum rule, to progressing the oncoming such as dynamic classifier selection and decision templates.

The researchers presented that, the performance of the learning system modified in most cases, regardless if the classifiers existence were integrated whether autonomous or not. The execution of a pure integrator such as the majority vote is easy and straightforward when compared with the other techniques mentioned above. It does not undertake any prior knowledge of individual classifiers' behavior. A study is underway to compare five classifiers integrates and used to assist in the methods of figure recognition: Majority Vote, Logistic Regression, Bayesian, Fuzzy Integrals, and Neural Networks (NNs) when applied together integrate the decisions of seven classifiers. After analyzing the results, the authors indicated that the Majority Vote method is as effective as the other complex systems.

While integrating classifiers output using the majority vote, the data sample is assigned as unclassed that is predicted by a majority of the classifiers. In some applications, the data sample is assigned to classes if more than half of the classifiers permit the identity of the data sample not assigned as a class [64].

2.5 Function Electrical Stimulation (FES) with EMG signal for Rehabilitation devices.

They are having discussed the Myoelectric Pattern Recognition (M-PR) and moving on to consider the use of the FES and EMG devices and its application. New rehabilitation methods concentrate on philosophies of motor learning, which consist of the following: impairment oriented-training (IOT), constraint-induced movement therapy (CIMT), electromyogram (EMG)-triggered neuromuscular stimulation, robotic interactive therapy and virtual reality (VR) [65]. Patients, who suffer from muscle weakness and imbalance movement patterns of the lower limb, show signs of weight loss due to their body weight supported (BWS) ability. The influence of BWS is an essential parameter for gait rehabilitation study. Assisting walk gait devices to improve with body weight supported treadmill training (BWSTT), and body weight supported locomotion training (BWSLT) [66]. Norman Sheall et al.[67] embedded the first neuro augmentative device to relieve severe pain. That was known as dorsal column stimulation. Furthermre development, this technique implemented in patients with movement disorders known as spinal cord stimulation (SCS).

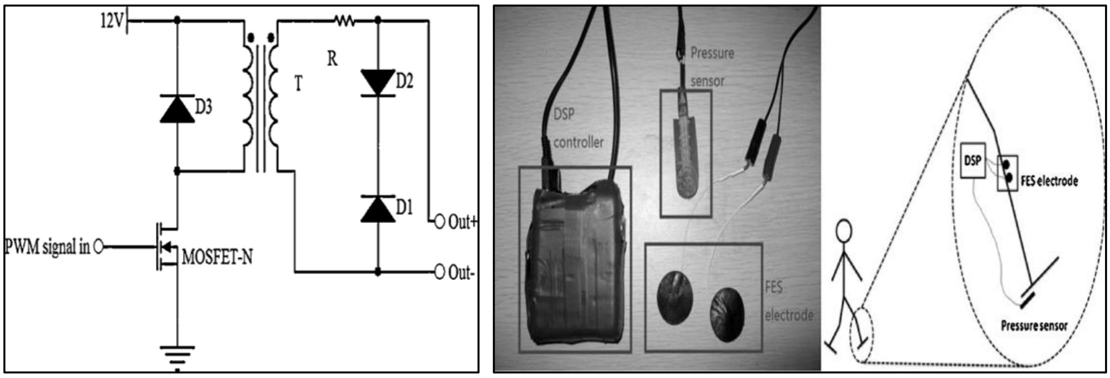
Electrical stimulation (ES) or Neuromuscular Electrical Stimulation [NMES] is a method using to rehabilitate muscle energy, activity, and functionality. When ES applies to the operation of a specific muscle exercise, then it is known as Functional Electrical Stimulation (FES). ES provides a low voltage stimulus to the surface of the patient skin to restore or improve the function of a neurological disorder or injury. ES works by triggering the peripheral nerves and generating a chemical reaction in the sodium-potassium pump that leads to the contraction of the muscle. Therefore the peripheral nerve must be undamaged in order to stimulate the SCI effectively [68].

In 1978, it ascertained that multi-channel FES was applied for 10 to 60 min., three times per week for one month would improve the ability to walk for hemiplegic patients. A study in 1989 applied Multi-channel FES to motivate the lower limb muscles of 20 chronic hemiplegic patients. After daily treatment ranging from one to three weeks, patients who are incapable of walking were walking again. In the 1990s, FES applied to treat the lower limb for extreme stroke patients. This treatment is given weekly for 7 to 21 days. The results showed that most patients were capable of walking and living independently after the use of FES. Several articles concluded that motor practice after brain injury assists in a physiological restructuring which is present appears in the undamaged tissues. The use of Positron Emission Tomography observed similar brain stimulus patterns in stroke patients during either active or passive movements [69-71].

Function neuromuscular stimulation (FNS) is a technique created to implement the use of rehabilitation devices for patients with different neuromuscular disorders.

The mechanism of this technique is to provide an electrical stimulus that activates the neurons leading to the flexion and extension of paralyzed muscles. Problems uncounted with this technique are the quality of control movement and position attained by the patient.

The FNS control system aims to find a suitable design for the stimulation factors to be carried to a set of muscles to perform a specified mission. Many FNS control systems work in an open-loop method when the output of the control ultimately depends on its inputs from the user [72].



(a) **(b)**

Figure 2. 10 Function Electrical Stimulation Electronic: (a) circuit for FES, (b) Whole system for FES

The circuit in Figure 2.10a for FES shows Pulse Width Modulation (PWM) signal drives to the gate of the Metal Oxide Semiconductor Field Effect Transistor (MOSFET) which controls the transformer output current [73]. While Figure 2.10b describes the whole FES system that consists of Digital Signal Processing (DSP) (TMS320F2812, Taxes Instrument Inc., USA) to control the electrical pulse generator, a pair of FES electrodes and a heel pressure sensor sizes [73].

Using the posterior trim or stable ankle AFOs for walking reduces the extreme hip and knee flexion and ankle plantar flexion at the swing phase. In contrast, there is little variation during the stance phase (in the first phase). The patient is not satisfy wearing AFOs; they prefer the posterior trim rather than the stable ankle AFOs because it fits better with their shoes [73, 74] as shown in Figure 2.11.

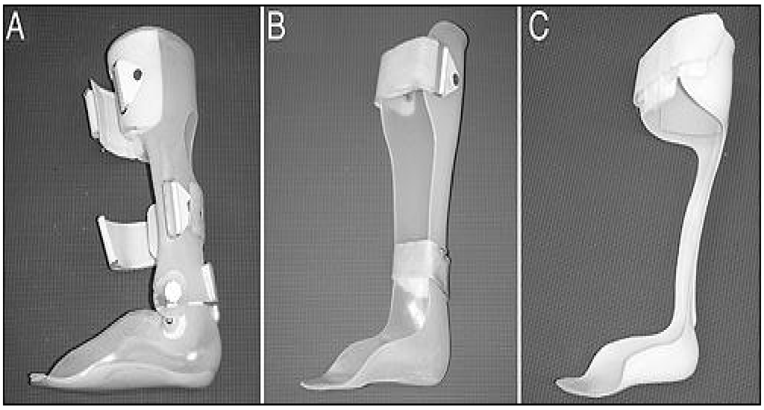


Figure 2. 11 Ankle Foot Orthotic: Solid ankle AFOs for walking (A), the posterior trim (B, C)

The pattern of a walk with FES for patients with foot-drop was more regular and stable than for those wearing the AFO, which needs to be custom-made and specific for each patient. The table below describes the development of research on the work on foot drop devices and analyses the objective, subjects, outcomes for their investigations, which will help patients to lift their foot and walk in stable gait as shown in Table 2.2:

Table 2. 2 The development of researchers works on foot drop devices

Device	**1- Od stock FDS**	**2- Hanger Walk Aid (WA)**
Objective	This FES device motivates the common peroneal nerve at the swing phase of gait.	This device effects speed & energy in the walking gait.
subjects	151 patients with a moderate level of DF caused by upper motor neuron injury.	24 subjects with a severe level of DF from neurological disorders with healthy common peroneal (CP) nerve and its muscular innervation
Outcome	P.Taylor et al.[75] trained patients to use the simulator to improve gait. Patients with stroke confirmed a short period “carry-over” effect[76]. The measurement changed the walk with and without motivation statistically which is essential for stroke and multiple sclerosis patients.	Full acceptance of the WA. The wide range of prices that people pay increases in addition to a universal health care system. People are accustomed to paying for medical care. Even where some devices carry a charge, it is frequently subsidized, such as a quarter of the whole cost[77]
Device	**3- NESS L300Plus, Bioness**	**4- Otto bock my walk**
Objective	To performs the effects of using FES daily on the peroneal and thigh muscle also on the temporal aspects of gait performance with hemiparesis gait.	Some facts about My Gait, a new peripheral stimulator from Otto-bock compensated for shortage control on ankle joint which affected by drop

	Moreover, to demonstrate ankle and knee dysfunction [6].	foot.
subjects	84 subjects with hemiparesis are recruited, diagnosed with: upper motor neuron injury, Hamstrings or quadriceps strength less is than 4/5. Foot drop presents with toe drag while walking. Lower limb muscle tone is between 0–3 rendered to the adapted Ashworth scale, and able to walk independently or with an aid device (cane, walker, etc), capable of tracking multiple-step orders and suitable to respond to electrical stimulation,	
outcome	Assumed to improve the results with suitable control groups and slow the periods of use. Also, Kinetic and kinematic research can be beneficial to recognize the primary mechanism of the dual-channel of FES [78].	My Gait gives an electrical impulse that sent to the nerve; the nerve, in turn, enthuses the muscle to move. Advantages of My Gait is, additional to the dorsiflexion muscles; other muscle groups can stimulate. That led to provide extra support when walking. Also, My Gait: improve walk pattern that gives self-confidence and speed for drop foot.
Device	**5- Otto Bock Acti Gait**	
Objective	To estimate a selected implantable drop-foot stimulator (Acti Gait) which effects walk safely.	
subjects	15 Volunteers implanted in this test. People who suffered a stroke event 6 months prior as a minimum and had a drop-foot affected walk recruited from 3 rehabilitation centers.	
outcome	Change of ankle dorsiflexion with the stimulus increases walk speed and distance walking in (4) minutes. Nerve conducting velocity and different events evaluate the overhaul.	
conclusion	The experiment assessed the safety and performance of the device that was most accepted by patients [79].	

Three types of FDA use transcutaneous peroneal nerve stimulation. The Od Stock Dropped Foot Stimulator as pictured in figure 2.12(on left top) and the wireless L300 in figure 2.12 (on left-bottom) both of which use a heel switch to trigger ankle dorsiflexion. While, the Walk Aide shown in Figure 2.12 (on the right) use a tilting sensor to trigger ankle dorsiflexion [80-82].

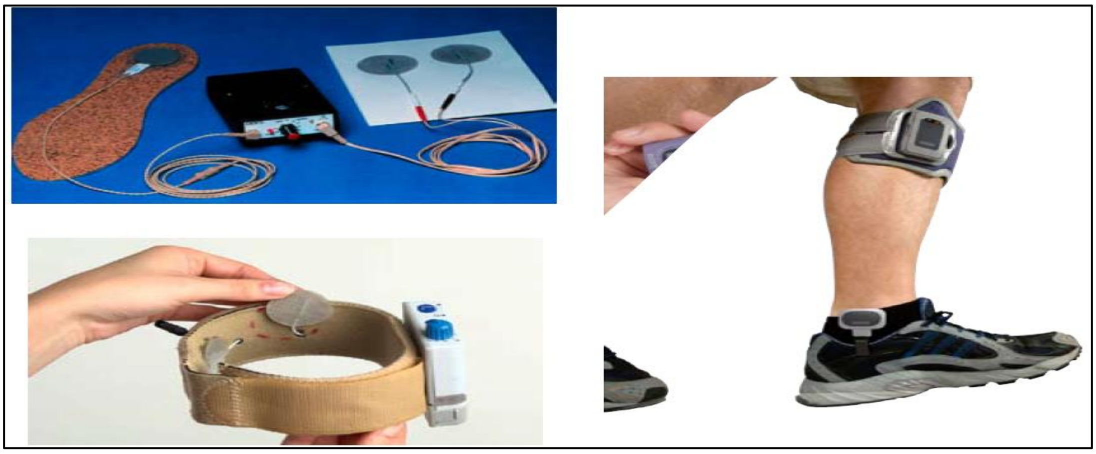


Figure 2. 12 FD Rehabilitation devices (on left-top) Od Stock Dropped Foot, (on left-bottom) Walk Aide and (on the right) L300

The main purpose of these applications is to empower SCI to achieve daily living activities that could not complete. Surface neuron-prosthetic applications are select in this study to restore function in the pediatric SCI patient. There are prosthetic neurons and rehabilitation devices that provide functional electrical stimulation (FES) for people with disorders of the central nervous system. The Bioness H200 Hand Rehabilitation System, L300 Foot Drop system as shown in Figure 2.13, and L300 plus System are an example of such devices [6, 68].

Bioness L300 is a device used to treat the lower limb for DF. This device contains an ES system, where a leg cuff activates and contracts exact leg muscles to enable dorsiflexion with small versions of the foot. It shows patients the difference between wearing Ankle Foot Orthotics (AFO) that permits a regular walking pattern with increased gait speed. Moreover, the L300 Plus device adds a thigh cuff to allow the patients with thigh weakness better control over twisting and straightening the knee.

The FDA has approved the use of Bioness H200 for people with upper limb parlays due to effects SCI or stroke. There is no approval to paediatrics, as it has not identified a lower age limit. Children not listed in the contraindications for use [68].

Figure 2. 13, Bioness L300 Footdrop device Courtesy of Bioness Inc, Valencia, CA [78]

C. Chesler & K. Durfee [83] declared that the electromyogram (EMG) signal was able to indicate motivated muscle fatigue when functional electrical stimulation (FES) used on an FD patient. Surface EMG measurement has the trouble of stimulus artifact during FES. They designed and built an adapted surface stimulation and EMG detection device to reduce this artifact, and it allows the detection of an electrical signal that is created by the muscle during the contraction state.

They implemented an Isometric fatigue experiment by motivating the quadriceps muscle for 20 healthy subjects (12.5 experience) and three spinal cord injured (18 experience) subjects. Peroneal nerve functional electrical stimulation (peroneal FES) has developed the gait during the walk for patients with chronic stroke. Kari Dunning et al.[84] describe the variance of walking with and without a neuroprosthesis the first few weeks after a stroke, offering a clinical viewpoint to consider the use of peroneal FES to rehabilitate neuroprostheses during the severe phases of stroke recovery.

P. Burler et al.[85] explained the Physiological Cost Index (PCI) as an indicator of energy cost for gait and is calculated as in Equation 2.1.

$$\mathrm{PCI} = \frac{(\text{Walking pulse } - \text{ resting pulse})}{\text{gait speed}} \quad (\text{beats/meter}) \qquad (2.1)$$

While the measuring in beats per meter to reflect the power of walking.

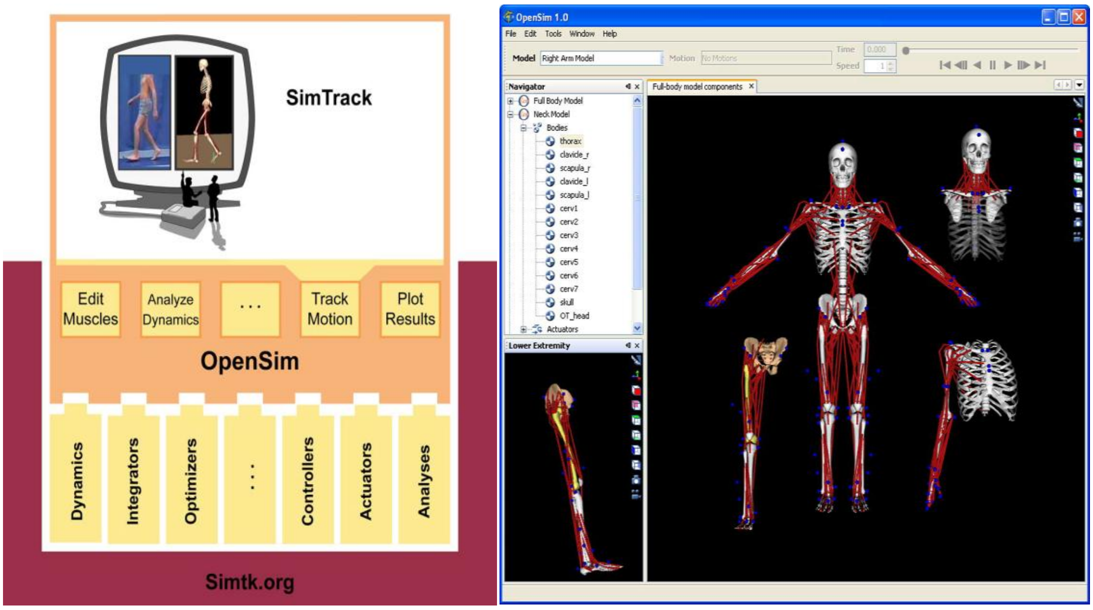
A minimum magnitude suggested the energy-efficiency gait. A connection between higher oxygen consumption and heart rate through walking reported for children and healthy young adults at a different range of speeds. However, this suggests that heart rate monitoring may be more reliable for oxygen rate purpose. It could be used as a precaution in older adults because of the age-related changes in the cardiopulmonary function. That is perfect to use for older adults with heart disease that take medication that damps down the heart rate response.

2.6 Dynamic Simulation for different movements of the musculoskeletal system.

Humans have an extraordinary capability to undertake multifaceted movements, which needs swiftness, timing, and To overcome the problem and to improve the reliability of the diagnosis process, precision. Disability, disease and overall aging are driving factors that lead to the loss of muscle strength, which then leads to limitations in the performance of motor tasks. To fully understanding these movements, Dynamic simulations of movement contributes to understanding the neuromuscular harmonization better; by analyzing athletic performance, evaluating center loading of the musculoskeletal system and categorizing the sources of pathological movement and a scientific base is created for treatment strategy.

In the 1990s, Delp and Loan introduced a musculoskeletal modeling environment, named SImulation of Movement in Musculoskeletal systems (SIMM). SIMM helps operators to generate, modify, and estimate models of different musculoskeletal configurations. SIMM can use by a biomechanics researcher to make computer models of musculoskeletal configurations to mimic movements like walking, stair climbing, cycling, and running. In contrast, it does not support programs to give computation of muscle stimulation that proceeds to synchronize movement and has insufficient tools for the analysis of the outcomes of dynamic simulations. Besides that, SIMM and many commercial packages, for example, Anybody (Anybody Technology), Visual 3D (C-Motion Inc.) or Adams (MSC Software Corp.) do not offer complete access to the source code. Over the last few years, novel software engineering methods have led to develop of more extensible software systems and to improve the simulation stage that occupies a broad spectrum of the biomechanics community.

OpenSim is a software simulation ssystem which permits operatives to programme the model of the musculoskeletal structure and to generate dynamic simulations of different movements such as pathological gait and explore the biomechanical effects of treatments. They provide this Open-Source Simulation environment to increase the progress and sharing of simulation technology and to integrate dynamic simulations into the arena of movement science as shown in Figure 2.14a.



(a) (b)

Figure 2.14 Open Sim System: (a) Diagram of OpenSim, (b) low level computational tools.

OpenSim is essential in the computational component as it permits to derive equations of motion for dynamical systems, implement numerical integration, resolve controlled non-linear optimization problems and offer access to figured muscle control and analyses muscle accelerations. With the available graphical user interface (GUI), the user can access a group of high-level tools to view models as shown in Figure 2.14b, edit muscles, plot results and so on. The Sim-Track allows fixed muscle-driven simulations to generated which characterize the dynamics of separate subjects. OpenSim was established and preserved on Simtk.org and is freely accessible for all with software inclusions. The software is written in ANSI C++, while GUI is written in Java, permitting OpenSim to compile and run on common functioning systems.

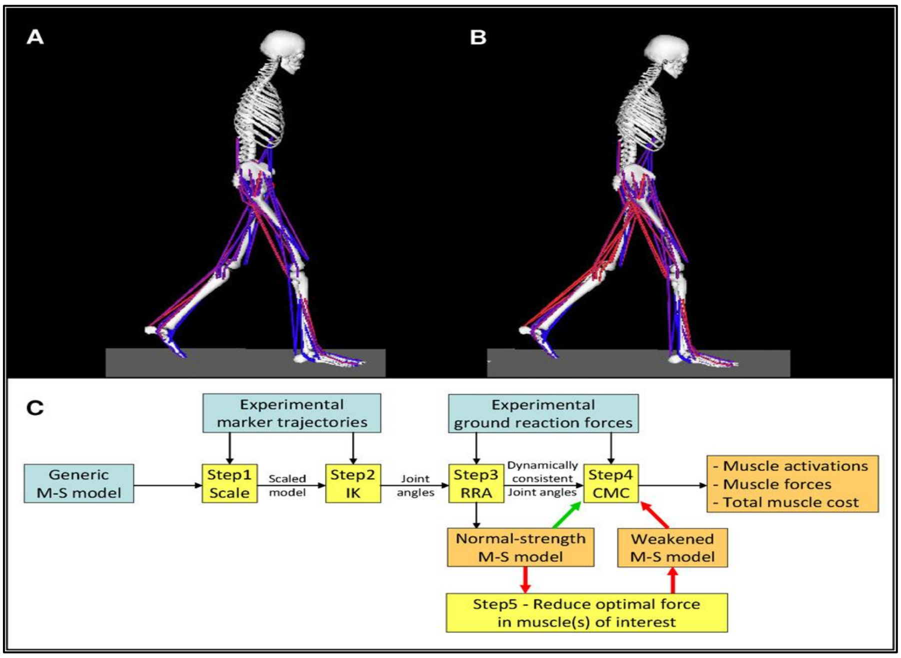


Figure 2.15 Steps for generating a muscle-driven simulation of a subject's motion with SimTrack.

Figure 2.15 shows the input which can be a dynamic musculoskeletal model, experimental kinematics like joint centers, x-y-z trajectories of marker data and joint angles, experiment Ground Reaction Forces and moments attained from the subject.

OpenSim also includes a software development kit (SDK) which contains application programming interfaces (APIs) and consistent libraries. OpenSim is programmable and only needs simple skills of C++ programming and information of the OpenSim API allowing greater competence in the use and development of such programs.[86]

Marjolein et al. [87] tested a range of lower limb muscles weakness before regular walking is affected. They improved simulations for a muscle-driven usual walk and debilitated all main muscle groups so that they could calculate how much weakness will accepted before the implementation of normal gait became difficult.

Ming Xiao et al. [88] compared muscle excitements and assistance to Centre Of Mass (COM) sustenance in 2D and 3D simulations of a healthy walk. They made 2D and 3D models in OpenSim and used them to imitate the same data of a usual healthy walk. They point out that the simulation results for both models were competent to imitate the similar joint kinematics and kinetics in the sagittal plane. While Steele, K.M. et al.[89] made the first simulations for muscle actuated of subjects during crouch gait.

This simulation offers a vision for how separate muscles contribute to the center of mass and joint angular accelerations at a single limb stance. Analysis of these simulations shows that when the crouch gait considered has superior muscle force when compared to the undamaged gait to support body weight and lift the body forward during single limb stance. Also, Liu, M.Q. et al.[90] recognized that the central muscle control is to modulate vertical support and forward movement during walking speeds in unimpaired children. Muscle assistance for support and progress improved with walking speed or during slow as shown in Figure 2.16 and free walking. When the walk slowed, a squarer limb required in initial stance, and muscles provide a majority of support against gravity. These outcomes demonstrate similarities and dissimilarities between complex and straightforward dynamic walking models.

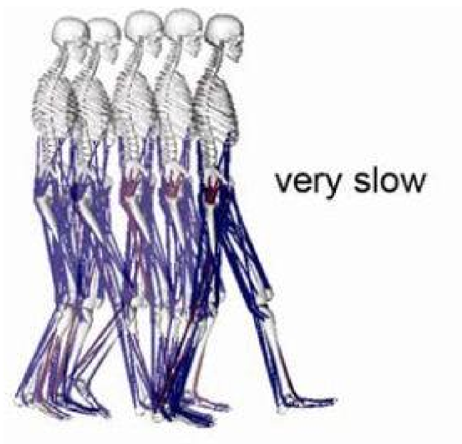


Figure 2.16 Simulink for prolonged walking for unimpaired children

2.7 Summary

The basic theory of the physiology of the lower limb, the muscles that control the leg and the electric signal pathways generated by the muscle activities are present in detail in this chapter. This chapter addresses different techniques that are useful for the use of bio-signals for muscle stimulation such as electromyography or myoelectric signals in addition to the use of different leg rehabilitation devices.

Myoelectric Pattern Recognition (M-PR) methods applied to intelligent artificial applications. The system contains multi-steps. First, the sEMG data is filtering from the noisy signal and then applying to extract and reduce features to minimize the large sample data. The data can now classified into multi-classes. Single Class Classification (SCC) and Multi-Class Classification (MCC) are discussed to identify the difference between them. There is a brief explanation of the details in the next section for Single Label Classification (SLC) and Multi-Label Classification (MLC) with their importance in the application.

Function Electrical Stimulation (FES) and EMG signal for Rehabilitation devices are identify briefly as well. Function Electrical Stimulation (FES) and EMG signal for Rehabilitation devices are identify briefly as well. This section describes fleetingly the most Rehabilitation devices utilized like Od stock FDS, Hanger Walk Aid, NESS L300Plus, Bioness, Otto bock My walk, Otto bock Acti Gait. The most essential and improvement devices are NESS L300Plus, Bioness. The research showed in this thesis will consider the validity of issues with M-PR. The limitations of myoelectric control in the exoskeleton lower limb will discuss further. This thesis proposes different methods to improve the performance of M-PR by proposing classifiers and a feature prediction method that is better suited to M-PR. To fully understand the movements of FD gait; Dynamic Simulations of movement contributes to analyze athletic performance, evaluate center loading of the musculoskeletal system and create a scientific base for treatment. Chapter 4 will discuss first and second contributions for this research with all concepts and experimental results.

Chapter 3

Methodological Review

3.1 Introduction

In the previous sections, Foot Drop and its related information such as Lower Limb Anatomy, diseases, and problems that affect the lower limb, EMG, FES, and other therapies introduced and the scope of the thesis reviewed. This chapter will cover popular methods used to advance and improve the machine-learning methodology, including their application to Foot Drop.

Computer-based Foot Drop diagnosis systems significantly contribute in the early recognition of the stage or level of severity be it Severe, Moderate or Mild. This Chapter will review the most recent developments of these systems and observe the most up to date practices. Identify problems encountered, and recognize and study best practice for data acquisition, pre-processing feature extraction and reduction, and classification of Foot Drop data.

This chapter reviews the background of Myoelectric Pattern Recognition (M-PR) methods applied to intelligent artificial applications. The system contains multi-steps. First, the sEMG data is filtering from the noisy signal and then applying to extract and reduce features to minimize the large sample data. In addition, methodologies such as SVM, SA-SVM and Label Classification and their concepts that will use to binging with research contributions for this thesis.

3.2 Myoelectric pattern recognition (M-PR)

This section will address the parts of Myoelectric Pattern Recognition (M-PR) step by step by to clarify each stage of M-PR and then describe their equations. as shown in figure 3.1

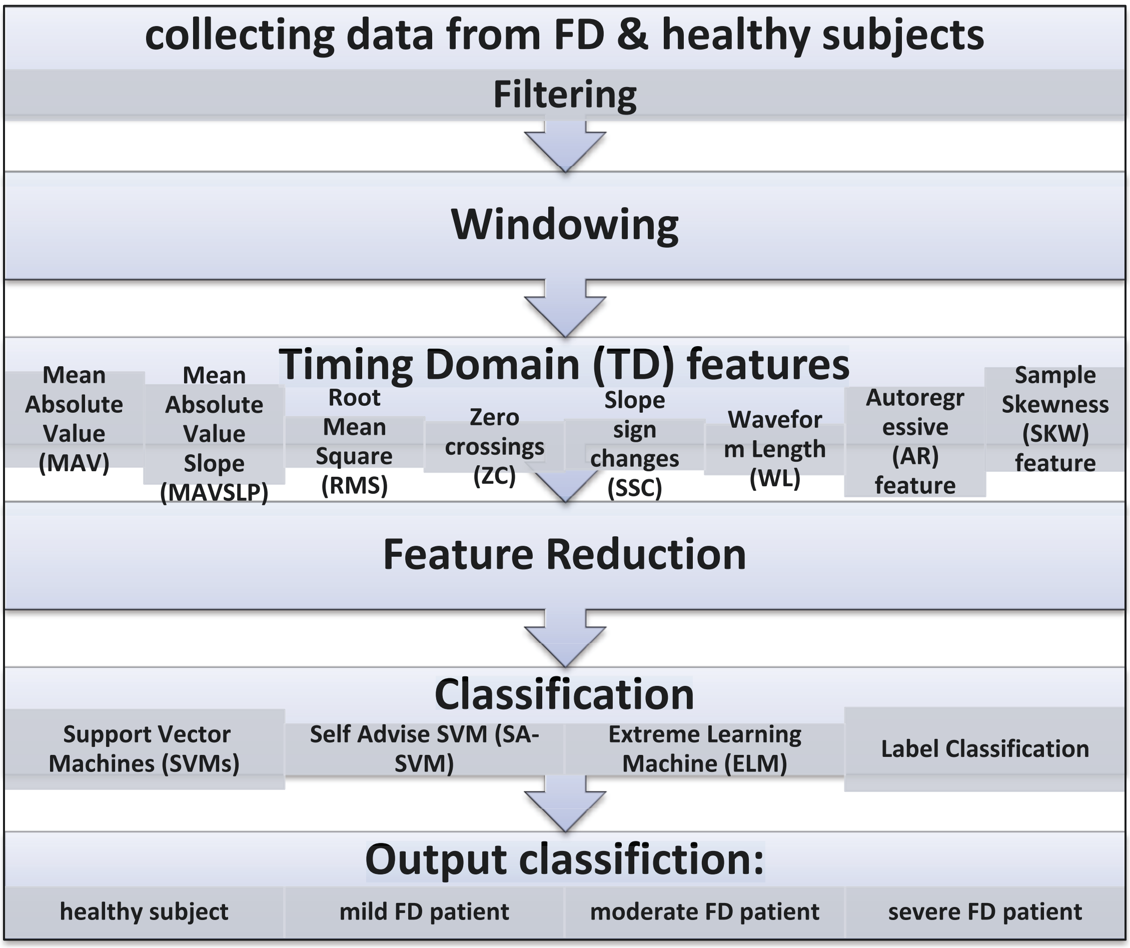


Figure 3. 1 Flowcharts for each stage of M-PR

The part for collecting data will be described with details in chapter 4 section 4.2 “Data Acquisition”, while filtering and windowing presented in chapter 2 section 2.4.2.

3.2.1 Feature Extraction: Timing Domain (TD)

The eight most commonly recommended TD features with the greatest effectiveness for EMG pattern recognition considered in this thesis. These features extracted within an N-sample evaluate time window.

1- **Mean Absolute Value (MAV)**

Mean Absolute Value (MAV) can measured by moving an average of a full-wave rectified the EMG signal. Alternatively, calculate by taking the average of the absolute value of the sEMG (myoelectric) signal. It is a simple method to detect muscle contraction levels and is an interactive feature use in a myoelectric control application. It is defined as shown in Eq.3.1:

$$\mathrm{MAV} = \frac{1}{\mathrm{N}}\sum_{\mathrm{n}=1}^{\mathrm{N}} |\mathrm{X_n}| \tag{3.1}$$

Where N represents the length of the signal and X_n denotes the sEMG windows of this signal [33].

2- Mean Absolute Value Slope (MAVSLP)

Mean Absolute Value Slope (MAVSLP) is an improved version of MAV. The modifications between the MAVs of adjacent windows are determined. Eq.3.2 can be defined as:

$$\mathrm{MAVSLP_i} = \mathrm{MAV_{i+1}} - \mathrm{MAV_i} \tag{3.2}$$

3- Root Mean Square (RMS)

Root Mean Square (RMS) is a format of amplitude modulated the Gaussian random procedure that it is RMS linked to the constant force with non-fatiguing contraction. It is related to the standard deviation. Eq.3.3 expresses it as:

$$\mathrm{RMS} = \sqrt{\frac{1}{\mathrm{N}}\sum_{\mathrm{n}=1}^{\mathrm{N}} \mathrm{x_n^2}} \tag{3.3}$$

4- Zero crossings (ZC)

Zero Crossing (ZC): Is the number of times that the amplitude value of the sEMG signal crosses the zero of the y-axis. In EMG feature, the threshold condition is applied to refrain from the background noise. This feature offers a nearest estimation of frequency domain properties. Eq.3.4 formulates it as:

$$\mathrm{ZC} = \sum_{\mathrm{n}=1}^{\mathrm{N}-1} \left[\mathrm{sgn}(\mathrm{x_n} \times \mathrm{x_{n+1}}) \bigcap |\mathrm{x_n} - \mathrm{x_{n+1}}| \geq \mathrm{threshold}\right] \tag{3.4}$$

Where: $\mathrm{sgn}(x) = \begin{cases} 1, & \text{if } x \geq \text{threshold} \\ 0, & \text{otherwise} \end{cases}$

5- Slope sign changes (SSC)

Slope Sign Change (SSC) is comparable to ZC. It denotes the frequency information of the sEMG signal. Change of numbers amongst a positive and a negative slope. Three consecutive windows are executed with the threshold function to avoid the interference in the sEMG signal. The scheming is defined as shown in Eq.3.5 [91]:

$$SSC = \sum_{n=2}^{N-1} \left[f[(x_n - x_{n-1}) \times (x_n - x_{n+1})] \right] \qquad (3.5)$$

$$\mathrm{f(x)} = \begin{cases} 1, & \text{if x } \geq \text{threshold} \\ 0, & \text{otherwise} \end{cases}$$

6- Waveform Length (WL)

This feature gives a calculation of the complexity of the x signal. It is a feature of the cumulative length of the EMG signal over a window, as appear in Eq.3.6:

$$\mathrm{WL} = \sum_{k=1}^{N} |x_k - x_{k-1}| \qquad (3.6)$$

7- Autoregressive (AR) feature

The Myoelectric signal is a non-stationary signal. However, it can be regarded as a stationary Gaussian process in a short time or a window [92]. Therefore, the EMG signal can model as shown in Eq.3.7:

$$x_k = \sum_{i=1}^{m} a_k x_{k-i} \qquad (3.7)$$

Where $\mathrm{AR} = a_k$ and m^{the} is the order AR model

8- Sample Skewness (SKW) feature

Skewness is a feature which describes a scale of a sample data sets symmetry or no-symmetry. An ideal symmetrical sample data set will have a skewness of 0. The standard supply has a skewness of 0, as shown in Eq.3.8.

$$\text{Skewness} = \frac{n}{(n-1)(n-2)} \sum \frac{(x_i - \bar{x})^3}{s^3} \qquad (3.8)$$

Where n is the sample size, i is the ith value, is the average and s is the sample standard deviation. the exponent in the summation part is (3), i.e. ,t he skewness is the (3th standardized central instant for the prospect model) [41].

3.2.2 Feature Reduction

The classical method for auto-encoder is to symbolize the data signal during a nonlinear encoder to a hidden layer and to utilize the hidden units as the new feature symbols [93] as shown in Equation 3.9.

$$h_i = \sigma\left(W_{1x_i} + b_1\right); \quad \hat{x_i} = \sigma(W_2 h_i + b_2) \qquad (3.9)$$

Where hi $\in R^z$ is the hidden symbol and $\hat{x_i} \in R^d$ is understood as a rebuilding of normalized input xi $\in R^d$. The parameter contains a weight matrices $W_1 \in R^{z*d}$, $W_2 \in R^{d*z}$, and offset vectors $b_1 \in R^z$, $b_2 \in R^d$ with dimension z and d. σ is a non-linear activation function. The auto-encoder with single hidden layer is normally a neural network (NN) with equally input and target output, as shown in Eq. 3.10

$$\underset{W_1,W_2,b_1,b_2}{\text{Min}} = \frac{1}{2n}\sum_{i=1}^{n} \|x_i - \hat{x_i}\|_2^2 \qquad (3.10)$$

where n is a sample size of the data, $\hat{x_i}$ is the rebuilt output while xi is the target output. Two methods of dimensional reduction were applied in this thesis: Linear discriminant analysis (LDA) and Sparse Auto-Encoder (SAE). As mentioned in section2.4.4, the feature set used in this thesis contains MAV, MAVSLP, RMS, ZC, SSC, WL, AR, and SKW. It gives = 22 features for each channel, which means: the feature set produces 88 (22x4) features for the four channels of sEMG signals. The dimensionality reduction methods reduce these features into 12 features for each channel.

3.2.3 Classification

Several classifiers such as Support Vector Machines (SVMs) and Self Advise SVM (SA-SVM) work with the first case and the second case called Multiclass support vector machines (MSVMs). Extreme Learning Machine (ELM) is another method that deals with single and multiclass.

The classification process was verified using fold cross-validation in all experiments. The accuracy expressed in Equation 3.11 uses to ration the performance of the classifier to recognize the leg movements. Additionally, the analysis of variance test like one-way (ANOVA) with a significant level set at p = 0.05 conducted on the classification accuracy for all subjects used to statistical test the importance of the proposed system.

$$\text{Accuracy} = \frac{\text{Number of correct samples}}{\text{Total number of testing samples}} \text{ x } 100\ \% \qquad 3.11$$

3.2.3.1 Support Vector Machine

Vapnik [94] proposed that The Support Vector Machine (SVM) is a vital classification method. Various forms of SVM presented in literature and applied in several different applications. SVM categorizes class problems for two classes and multi-classes. The SVM gives an ideal decision for the formation of a boundary between the two or multi-classes. The margin created that separates the classes, and decision boundary maximized. To theorize the binary classification by a training set of N samples, we must consider a vector for the input data which depends on the ith sample; labeled relate to its class.

The purpose of SVM is to separate the binary labeled training data within the hyperplane that is at the maximum distance from them. That is called a maximum margin hyperplane and it shown in Figure 3.2 [95] the basic idea of the SVM, which simplifies the pair (w, b) described in the hyperplane with the equation $< \mathrm{w}, \mathrm{x} > +\mathrm{b} = 0$. Theorize a binary classification by training sets of N samples$(\mathrm{x}_1, \mathrm{y}_1), \ldots, (\mathrm{x}_\mathrm{i}, \mathrm{y}_\mathrm{i}), \ldots, (\mathrm{x}_\mathrm{N}, \mathrm{y}_\mathrm{N}) \in \mathcal{R}^\mathrm{n} \times \{\pm 1\}$, considering x_i as a vector for input data depending on the ith sample, which labeled as y_i that related to its class.

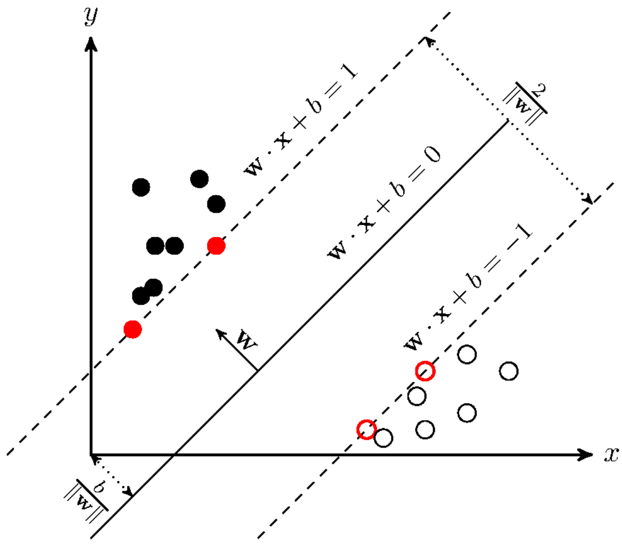


Figure 3. 2 Principle of Support Vector Machine

The hyperplane can separate the train data linearly if the distance for every train data from the hyperplane is specified by the Equation 3.12 and Equation 3.13:

$$y_i(w.x_i + b) \geq 1\,, \qquad i = 1, \ldots, N \qquad (3.12)$$

$$d_i = \frac{w.x_i + b}{||w||}, \qquad (3.13)$$

The merging applies for the two equations above. Results are shown in Equation 3.14:

$$y_i d_i \geq \frac{1}{||w||}. \qquad (3.14)$$

Thus, $\frac{1}{||w||}$ the distance lower bound during the separating hyperplane and the training data, x_i.

The maximum margin of the hyperplane measured to solve the problem of maximizing the $\frac{1}{||w||}$ subject to the restriction 3.12, or by Equation 3.15:

$$\text{Minimize } z = \frac{1}{2} w.w \qquad (3.15)$$

Suppose$(\alpha_1, \alpha_2, \ldots, \alpha_N)$Is N non-negative Lagrange multipliers related to equation 3.1. After a series of steps, the function decision will become in Equation 3.16:

$$f(x) = \text{sign}\left(\sum_{\alpha_i > 0} y_i \alpha_i < x, x_i > +b\right) \qquad (3.16)$$

While the Non-zero for Equation 3.1, is satisfied with an equal sign. Most of α_iThe data is typically zero. The vector w is a linear combination of a comparatively small ratio of the training data x_i. These plot points are known as Support Vectors (SV) because it is closest to the separating hyperplane and these points are necessary to evaluate the hyperplane. The Support Vector is a training pattern that depends on the boundaries of the margin. SVM uses only a small subset of the training samples SVs for the classification. Another category of SV consists of the training data that is beyond their lateral margins. These support vectors viewed as misclassified data [96, 97]. When train data are separate none linearly, it is then difficult to find a separate hyper-plane when there is no separating hyperplane to start on. Luckily, we can solve this problem by introducing N non-negative variables $(\xi_1, \xi_2, \dots, \xi_N)$ as in Equation 3.17:

$$y_i(w.x_i + b) \geq 1 - \xi_i\,, \qquad i = 1, \dots, N. \tag{3.17}$$

The variables ξ_iFound to allow for a small number of misclassified data. That illustrated in the next two steps:

1- If the data x_i satisfies difference in Eq.3.12, then, ξ_i is zero and Eq.3.17 reduces to Eq.3.12
2- If the data x_i does not satisfy the difference in Eq.3.12, the extra term -ξ_i has to add to the right side of Eq.3.12 to get the difference Eq.3.17.

Essentially, some of the training data are ignored with the application of the tolerance parameter to get a linearly separating hyperplane. The separate hyper-plane then considered as the explanation to equation 3.18:

$$\text{Minimize } z = \frac{1}{2} w.w + C\sum_{i=1}^{N} \xi_i \tag{3.18}$$

$$\text{s.t.} \qquad y_i(w.x_i + b) \geq 1 - \xi_i, \qquad i = 1, \dots, N.$$

$C\sum_{i=1}^{N} \xi_i$, This parameter keeps many misclassified points under control. In addition to this, the parameter leads to a stronger solution. The penalty parameter C considered as a regularisation parameter. Preview problem leads to one of these three steps:

1- Maximize the minimum distance 1/w for small C.
2- Minimize the number of misclassified points for large C.

3- Middle values of C is the solution to the problem in Equation 3.18 and is trades its errors for a more considerable margin. For this step, the decision function is given by Equation 3.19:

$$f(x) = \text{sign}\left(\sum_{\alpha_i>0} y_i\alpha_i < x, x_i > +b\right) \qquad (3.19)$$

$$0 \leq \alpha_i \leq C, \qquad i = 1, \dots, N$$

SVM can utilize to produce a non-linear decision function, by projecting the training data to a higher dimensional internal product space F, known as feature space, by applying a non-linear map$\phi(x): \mathcal{R}^n \rightarrow \mathcal{R}^d$. Although the optimal linear hyper-plane calculate in the feature space, by applying kernels, it is capable of making necessary processes in the input space using $k(x_i, x_j) =< \phi(x_i), \phi(x_j) >$it as which is an internal product in the feature space. In terms of these kernels, the decision function can write as following Equation 3.20:

$$f(x) = \text{sign}\left(\sum_{\alpha_i>0} y_i\alpha_i k(x, x_i) + b\right) \qquad (3.20)$$

The decision value for each x of the test sets have either a negative or a positive value dependent on the situation of x and the hyperplane, and this clarified as Equation 3.21:

$$h(x) = \sum_{\alpha_i>0} y_i\alpha_i k(x, x_i) + b \qquad (3.21)$$

There are three mutual kernel functions in SVM:

1. RBF kernel : $K(x_i, x_j) = e^{-\gamma|x_i - x_j|^2}$
2. Polynomial kernel: $K(x_i, x_j) = (x_i x_j + 1)^q$
3. Sigmoid kernel: $K(x_i, x_j) = \tanh(\gamma x_i^T x_j + c)$

Here q, γ, c are kernel parameters.

3.2.3.2 Self-Advised Support Vector Machine (SA-SVM)

Traditional SVM neglects the train data that has not been separate linearly by the kernels through the training phase. That happens during the overview of the tolerance parameters in the objective function and constraints. Therefore, it will be classified incorrect if the data that is identical or similar to this misclassified data appears in the test set. The reason being, the data, which is closest to the misclassified data, is uncertain. This misclassification is not convincing and can control if the Traditional SVM procedure does not neglect the existing data in the training phase.

Non-repeating self-advising method for SVM is adapted, which extracts consequent knowledge from the training phase without the addition of extra parameters. The misclassified data is supply from two prospective sources, first one from outliers and the second one from the data that has not been separate correctly. Advised Weighted-SVM treats the neglect of SVM from the knowledge that can obtained from misclassified data. By the creation of advise weights, which are depending on the distance for the misclassify train data from the correctly classify train data. Also, they are applying these weights together with decision values of SVM in the test phase. These weights assist the procedure in reducing the outlier data[11].

The details of the Self-Advised SVM procedure demonstrated in the following steps:

1- Classify the hyper-plane founded by applying the decision function in Equation 3.22:

$$f(x) = \mathrm{sign}\left(\sum_{\alpha_I > 0} y_i \alpha_i \, k(x, x_i) + b\right) \qquad (3.22)$$

Where, the input vector for samples that labeled with y_i Dependent on its class and αi is the non-negative Lagrange multiplier that contradicts the standard SVM training. Using Radial Basis Function kernel (RBF-SVM) to produce non-linear decision functions as the data includes non-linear separate cases, where all required operations are in the input space. Besides, the RBF kernel is not similar to the linear kernel; it can occur in cases between the relative class labels and attributes that is non-linear.

The sigmoid kernel performs like RBF for specific parameters; it also has some hyper-parameters that can inspire the difficulty of the model. Whereas, the polynomial kernel includes extra hyper-parameters than the RBF kernel and adds more complexity of model when its compared to increasing the classification accuracy.

2- Misclassified data that samples in the first train phase are recognized. The misclassify datasets (MD) in the training phase is calculated by Equation 3.23:

$$MD = \bigcup_{i=1}^{N} x_i \; | y_i \neq \mathrm{sign}\left(\sum_{\alpha_j > 0} y_i \, \alpha_j \, k(x_i, x_j) + b\right) \qquad (3.23)$$

The MD set may be null, but the practical outcomes appear when the existence of misclassified data in the training phase is a communal existence. It should recognized that any technique in the attempt to benefit from misclassified data should have control to affect the outlier data. When the misclassified data is included to resemble samples, the use of misclassified data improved the classification accuracy [98].

3- The algorithm indicates: If MD is null then go to the testing phase, or else compute neighborhood length (NL) for each X_i of MD. Equation 3.24 defined NL.

$$NL(x_i) = \text{minimum}_{x_j} \left(\|x_i - x_j\| \mid y_i \neq y_j \right) \tag{3.24}$$

Where X_j, $j=1 \ldots N$ is the training data that does not belong to the MD set. If the training data is a map as a higher dimension, the distance between xi and xj can evaluated in Equation 3.25 concerning the related RBF kernel.

$$\|\theta(x_i) - \theta(x_j)\| = (k(x_i, x_i) + k(x_j, x_j) - 2k((x_i, x_j))^{0.5} \tag{3.25}$$

4- The calculation for Advised Weight AW (x_k), for each sample x_k from the test, set utilized by Equation 3.26. These AWs represent the closest test data to the misclassified data.

$$\begin{cases} 0 & \forall x_i \in MD, \|x_k - x_i\| > NL(x_i) \text{alternatively}, MD = NUL, \\ \sum 1 - \dfrac{\sum_{x_i} \|x_k - x_i\|}{\sum_{x_i} NL(x_i)} & x_i \in MD, \|x_k - x_i\| \leq NL(x_i) \end{cases} \tag{3.26}$$

5- The absolute value of the SVM decision, for each x_k for the test set considered and scaled in to [0, 1].

6- Finally, for each xk from the test set:

If AW (x_k) < decision value (x_k) then $y_k = \text{sign}\left(\sum_{\alpha_j > 0} y_j \alpha_j k(x_k, x_j) + b\right)$, which identified with standard SVM.

Otherwise $y_k = y_i \mid (\|x_k - x_i\| \leq NL(x_i) \text{ and } x_i \in MD)$

3.2.3.3 Extreme Learning Machine (ELM)

The basic concept of ELM is a single feed-forward network as shown in Figure 3.3.

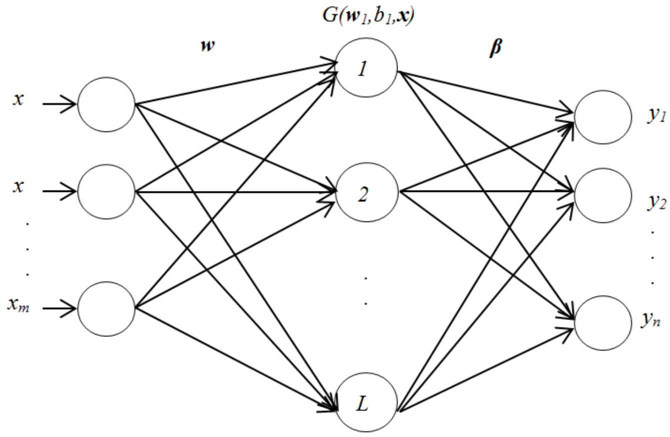


Figure 3. 3 Single feed-forward networks for extreme learning machine

The figure describes that, if N is a sample (x_j, y_j), then the input $x_j = [x_{j1}, x_{j2},\ldots, x_{jm}]^T \in R^m$ and target $y_j= [y_{j1}, y_{j2},\ldots, y_{jn}]^T \in R^n$ while the output of SLFN with L hidden neurons shown in Equation 3.27

$$f_j(x) = \sum_{i=1}^{L} \beta_j g(w_i.x_j + b_i) = \boldsymbol{G\beta} \quad j = 1, \ldots, N \tag{3.27}$$

Where $w_i = [w_{i1}, w_{i2}\ldots, w_{im}]^T$ symbolizes as the vector of the weight relating the *ith* hidden neuron with the input neurons, $\beta_j = [\beta_{j1}, \beta_{j2}, \ldots, \beta_{jL}]^T$ describes the vector of the weight with *i*th hidden neuron with output neuron, b_i is the threshold of the *ith* hidden neuron while g(x) is the activation function of the hidden node. The right side for the equation is a compact form of the SLFN output where G is the hidden layer output matrix as shown in Equation 3.16 and β in Equation 3.28:

$$\boldsymbol{G} = \begin{bmatrix} g(w_1.x_1 + b_1) & \cdots & g(w_L.x_1 + b_L) \\ \vdots & \ddots & \vdots \\ g(w_1.x_N + b_1) & \cdots & g(w_L.x_N + b_L) \end{bmatrix}_{NxL} \tag{3.28}$$

And

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_1^T \end{bmatrix}$$

ELM objective is to reduce train error and norm output weight, so that shown in Equation 3.29:

$$\min_{w_i, b_i, \beta} : ||G(w_1, \dots, w_L, b_1, \dots, b_L)\beta - \boldsymbol{T}||^2 \tag{3.29}$$

Where T is the target that calculates in Equation 3.30:

$$\boldsymbol{T} = \begin{bmatrix} y_1^T \\ \vdots \\ y_N^T \end{bmatrix}_{Nxn} \tag{3.30}$$

The input weights w_i and biases b_i allocated randomly. So that Equation 3.31 can write as shown:

$$\min_{\beta} : ||\mathrm{G}\beta - \mathrm{T}||^2 \tag{3.31}$$

The least-square explanation of Equation 3.20 with least norm is write in Equation 3.32:

$$\beta = \mathrm{G}^{\dagger}\mathrm{T} \tag{3.32}$$

While $\mathbf{G}^{\dagger}$ is the Moore–Penrose generalized inverse of the matrix **G**.

We can formulate the optimization of ELM training using Equation 3.33.

$$\begin{aligned} &Minimize: L_{P_{ELM}} = \frac{1}{2}||\boldsymbol{\beta}||^2 + C\frac{1}{2}\sum_{i=1}^{N}||\boldsymbol{\xi_i}||^2 \\ &Subject\ to: \mathrm{g}(\mathbf{x}_i)\boldsymbol{\beta} = \mathbf{t}_i^T - \boldsymbol{\xi}_i^T \qquad i = 1, \dots., N \end{aligned} \tag{3.33}$$

Where $\boldsymbol{\xi}_i^T = [\xi_{i,1}, \dots, \xi_{i,n}]$ is the vector of output error *m* related to input $\mathbf{x}_i$.

Based on the Karush-Khun-Tucker (KKT) theorem [99], the training of ELM is the answer to the dual optimization problematic as calculating in Equation 3.34.

$$\begin{aligned} \mathrm{L}_{\mathrm{D_{ELM}}} = \frac{1}{2}||\beta||^2 + \mathrm{C}\frac{1}{2}\sum_{i=1}^{N}||\xi_i||^2 \\ - \sum_{i=1}^{N}\sum_{j=1}^{n}\alpha_{i,j}\left(\mathrm{g}(x_i)\beta_j - t_{i,j} + \xi_{i,j}\right) \end{aligned} \tag{3.34}$$

Where β_j is the output weight relating the hidden layer and the j*th* output node and $\beta_j = [\beta_1, \dots, \beta_n]$. With differentiate, we have Equations 3.35. Equation 3.36 and Equation 3.37:

$$\frac{\partial L_{D_{ELM}}}{\partial \beta_j} = 0 \rightarrow \beta_j = \sum_{j=1}^{N} \alpha_{i,j} g(x_i)^T \rightarrow \beta = G^T\alpha \quad (3.35)$$ **0**

$$\frac{\partial L_{D_{ELM}}}{\partial \xi_i} = 0 \rightarrow \boldsymbol{\alpha}_i = C\xi_i \quad (3.36)$$

$$\frac{\partial L_{D_{ELM}}}{\partial \alpha_i} = 0 \rightarrow g(x_i)\beta - t_i^T + \xi_i^T = 0 \quad (3.37)$$

Where $\boldsymbol{\alpha}_i = [\alpha_{i,1}, \dots, \alpha_{i,n}]$ and, $\boldsymbol{\alpha} = [\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_N]$.

By substituting Equation 3.36 and Equation 3.37 to get Equation 3.38:

$$\left(\frac{1}{C} + GG^T\right)\alpha = T \quad (3.38)$$

Where T can write as shown in Equation 3.39:

$$T = \begin{bmatrix} t_1^T \\ \vdots \\ t_N^T \end{bmatrix} = \begin{bmatrix} t_{11} & \cdots & t_{1n} \\ \vdots & \vdots & \vdots \\ t_{N1} & \cdots & t_{No} \end{bmatrix} \quad (3.39)$$

From Equation 3.32 and Equation 3. 38, we have Equation 3.40:

$$\beta = G^T\left(\frac{1}{C} + GG^T\right)T \quad (3.40)$$

Where C is a user-specified parameter. Eventually, the output functions of SLFN could be change to be Equation 3.41:

$$f(x) = g(x)\beta = g(x)G^T\left(\frac{1}{C} + GG^T\right)^{-1} T \quad (3.41)$$

If the training data is a big data, then the output of ELM will be as shown in Equation 3.42[50]:

$$f(x) = g(x)\beta = g(x)G^T\left(\frac{1}{C} + GG^T\right)^{-1} G^T T \quad (3.42)$$

Where g (x) is a feature mapping, (hidden layer output vector) can write like Equation 3.43:

$$g(x) = [Q(w_1, b_1, x) \quad \cdots \quad Q(w, b_1, x)] \quad (3.43)$$

Where, Q is a non-linear piecewise continuous function such as a sigmoid, hard limit, Gaussian, and multi-quadratic functions. The mathematic equations are shown in Appendix C

3.2.3.4 Label Classification

Multi-Label Classification (MLC) is the supervised learning problem where sample data connected to multiple labels. It is different from the original object of single label classification (SLC) such as multi-class or binary where every sample only connected to a single class label. Applications that use the MLC has increased in different fields. For example, text classification, scene and video classification, bioinformatics and Biomedical Text Data [100].

Generally, Binary Relevance (BR) is a method used for Multi-Label Classification (MLC). MLC deliberates every label as an independent binary problem, and its work depends on the lack of appearance for the non-direct modeling label correlations. Most of the existing methods contribute to the complexity of model inter-dependencies between the labels. Another method used in Multi-Label Classification is to perform problem transformation, where a multi-label problem transformed into one or more single-labels like binary or multi-class problems. That activates single-label classifiers and is transformed by their single-label predictions into multi-label predictions. Problem transformation is used to describe each of flexibility and scalability. They apply Support Vector Machines, Naive Bayes, k Nearest Neighbour methods and Perceptron [101].

Where $X^d \subset R$the input sample domain, the sample feature formed as a vector **d**

The sample input forms a vector of d feature$X = [x_1, \dots\dots, x_d]$, while labethe l output domain is$L = \{1, \dots\dots, L\}$. Each sample (x) related to a subset of these labels, which forms as L vector$y = [y_1, \dots\dots, y_L]$. Where y1 =1, if and only if label j is related with sample x, and 0 otherwise.

They assume a set of training data D of N labeled patterns as$D = \{(x^i, y^i) | i = 1, \dots, N\}$. So that the researcher can write multi-label accuracy equation for a set of N test as follow in Equation 3.44:

$$\text{Accuracy} = \frac{1}{N} \sum_{i=1}^{N} \frac{|y^i \wedge \hat{y}^i|}{|y^i \vee \hat{y}^i|} \qquad (3.44)$$

Ensemble Classifier Chains (ECC) signifies a vector of absolute outputs$\widehat{W} = [\widehat{w}_1, \dots, \widehat{w}_L] \in R^L$, while $\widehat{w}_j$ signifying the absolute for the jth label. For prediction vectors $\hat{y}_{1,\ldots,}\hat{y}_{m}$ from repeating 1... m, The absolute equation evaluate as in Equation 3.45:

$$\hat{w}_j = \frac{1}{m}\sum_{k=1}^{m} \hat{y}^i k \qquad (3.45)$$

Also, threshold function *f* apply $\hat{w}_j$ to get a bipartition of appropriate and inappropriate labels: $\hat{y} = f_t(\hat{w})$

Softmax functions also offer single-label classification. Practically, it applies to the multi-label scenario by problem transformation. Softmax loss function can modify the multilabel scenario as shown in Equation 3.46 [102]:

$$\boldsymbol{l_{softmax}} = \sum_{y \in Y_1} \log\left(\frac{exp\left(f_y(x_i)\right)}{\sum_{j \in y} xp\left(f_y(x_i)\right)}\right) \qquad (3.46)$$

3.3 Summary

The purpose of this chapter was to give a background of the research contributions and methodologies that will explain in the next chapters of the thesis. A review of the Support Vector Machine, Advise Weight SVM, Extreme Learning Machine and Label Classification Methods are provided to develop auto control and a diagnostic system for Foot Drop levels. Some of the conditions that affect the technique and performance are reviewed and reported. C**hapter 4** and **chapter 5** will discuss the experimental results presented by this thesis.

Chapter 4

Novel mSA-SVM and LSA-SVM Based Classification for Foot Drop

4.1 Introduction

This chapter will present and discuss the first and second theoretical and practical contributions to this thesis. Before that, **section 4.2** discusses the data collecting from Foot Drop Patients in Rehabilitation Hospital in Sydney, Australia using Ethical Approval (UTS HREC NO.ETH15-0152).

The first contribution as described in **sections 4.3** proposes contribution 1 (Cont.1) for the use of multi-classes using Self- Advised Support Vector Machine SA-SVM to detect and classify leg movement. **Section 4.4,** covers the second contribution 2 (Cont.2) that explains Label SA-SVM for two-class classification employed by the proposed algorithms in the first section. Figure 4.1 shows the general process for detection and classification for leg movement, first part in this flowchart describe the data signal supply from three groups of data collected from different kind of people (old, young, meal, female, and children ect.), the second part is pre-processing (filtering and windowing). Last part is a classification system which, making the contributions (Cont.) in four stages and discusses them briefly in their sections. These data size (three groups of datasets for each experimental test) acceptable in many research area especially in clinical applications and as an example Wu,, . and H. Sakai are woks in the same criteria [15, 103]

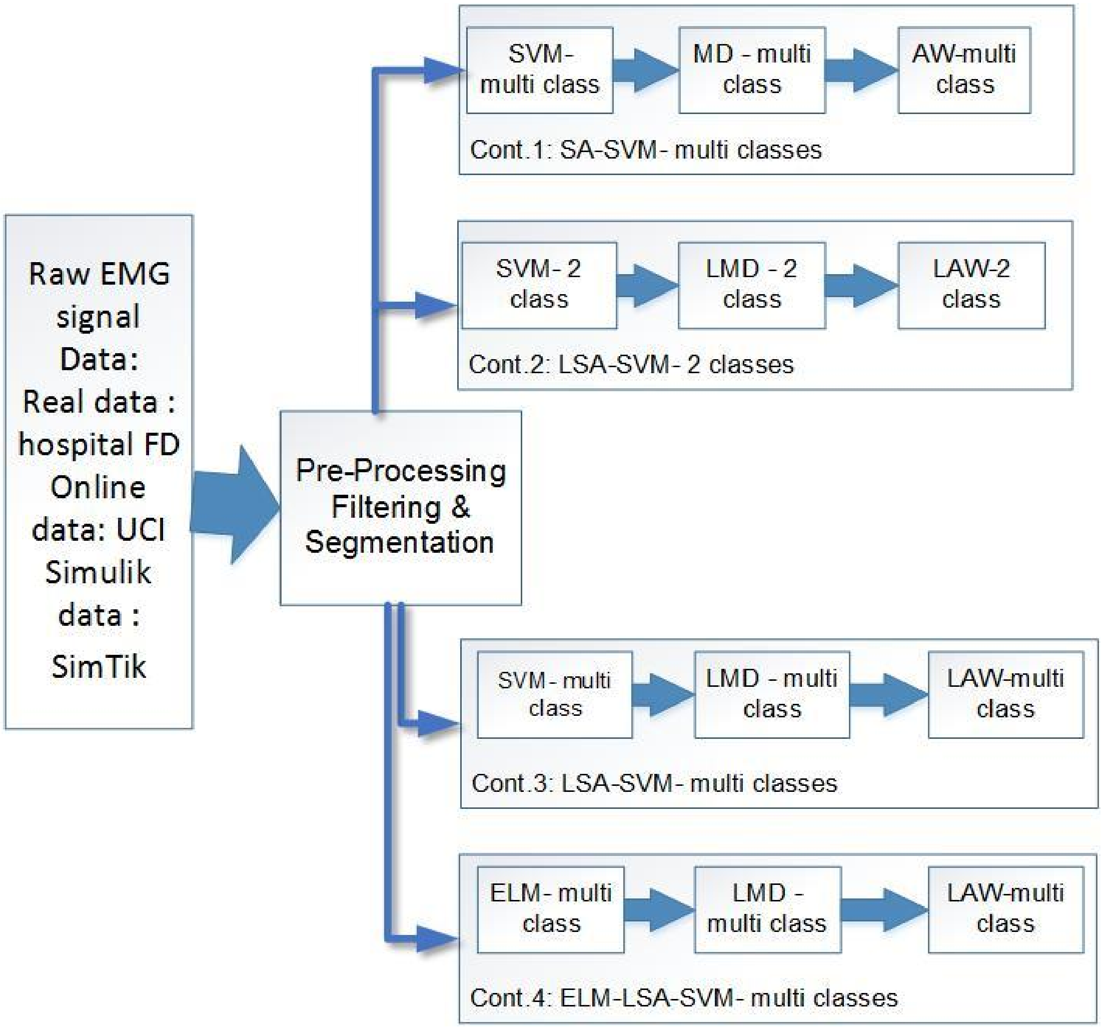


Figure 4. 1 Process of the Foot Drop detection and classification

4.2 Data Acquisition

4.2.1 Introduction

The Myoelectric Pattern Recognition system consists of two main parts: the software and hardware, further elaborated in the phases of the real-time M-PR system, as visible in Figure 4.2 and Figure 4.3. Overall, all phases of real-time pattern recognition based sEMG displayed in Table 4.1 for the experimental stages; the system collects the data from healthy and unhealthy subjects. The collected data is then applied to train the system. Resulting in the output of the classification is the trained classifier and the OpenSim prediction and simulation for gait level.

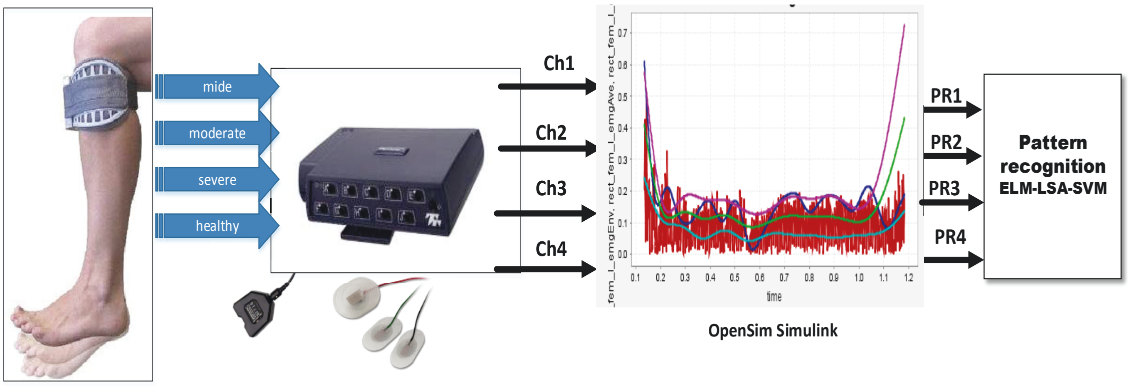


Figure 4. 2 Stages for an Offline myoelectric pattern recognition system

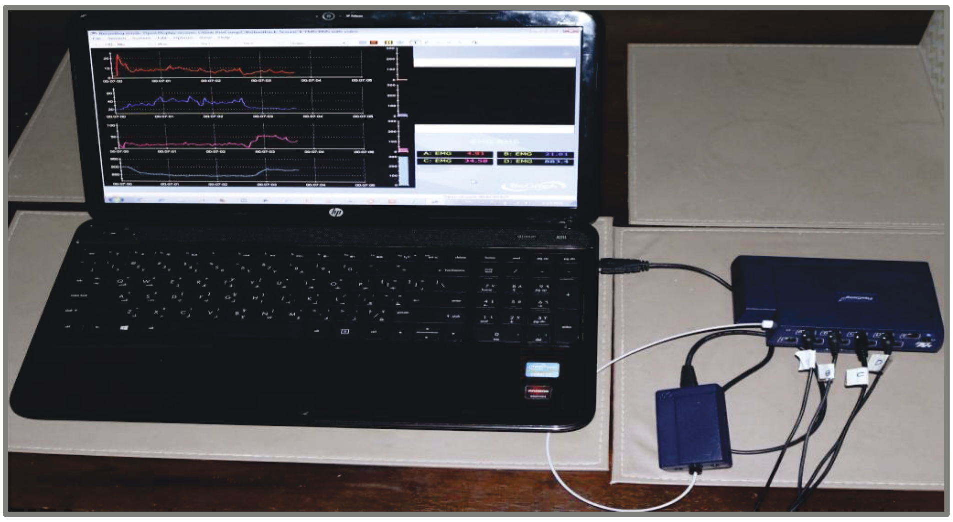


Figure 4. 3 The hardware and software needed for the off-line application

Table 4. 1 The hardware and software needed for the off-line application

Component	Description	Picture
Hardware	Personal Computer Intel Core i7, 2.8 GHz with 16 GB RAM equipped with Windows 10 operating system	
	EMG acquisition device, the FlexComp Infiniti™ System from Thought Technology with frequency sampling 2000 Hz.	

	Four EMG sensors: MyoScan™ T9503M Sensors from Though technology	
	Four electrodes	
	OneProCom Infiniti USB Adapter – TT-USB	
	One small piece of Fiber Optic Cable 15ft. - SA9480	
Software	Matlab R2016b	
	API library from Though Technology connecting the Flexcomp to Matlab	
	OneProCom Infiniti USB Adapter – TT-USB	

The collected EMG signals processed on a Personal Computer Intel Core i7, 2.8 GHz with 16 GB RAM and equipped with a Windows 10 operating system. A band-pass filter used to filter the signals in the frequency band (25 - 550 Hz). A notch filter is applied to remove the 50 Hz line nosiness. The EMG signals are downsized to 1000 Hz to minimize the size. To estimate the recommended *m*SA-SVM, LSA-SVM and ELM-LSA-SVM, the experiment embrace 13 datasets from the UCI machine learning, repository [15, 16]. These databases nominated from the most public benchmarks for classification and diagnosis. The diversity of these databases supports the authentication in this study. The number of instances and the attributes of each database shown in the Tables in each chapter used them. It should be noted that for a dataset with multi-class we used 6 datasets while for two classes were selected 7 datasets [17].

4.2.2 Procedure for collecting sEMG signal data

The data collected from the hospital based on the design procedure that got Ethical Approval to collect data from FD patients in Metro Rehabilitation Hospital in Sydney, Australia using Ethical Approval (UTS HREC NO.ETH15-0152). For the experiment to collecting data, there are 13 subjects involved in the offline experiment. They consisted of six females and seven males aged between 18 to 84 years the average age is 51 years. Foot Drop affected 10 of the subjects. The other three subjects were healthy with no muscle disorder. During the experiment, the subjects are the seat and at a fixed position for the knee as shown in Figure 4.4 and Figure 4.5, to avoid the influence of position movement on EMG signals. During data collection, a few digital filters are applying. They were a band pass filter between 25 and 550 Hz and a notch filter to remove the 50 Hz line nosiness. The EMG signals downsample to 1080 Hz.

Figure 4. 4 The placement of the Four Surface Electrodes

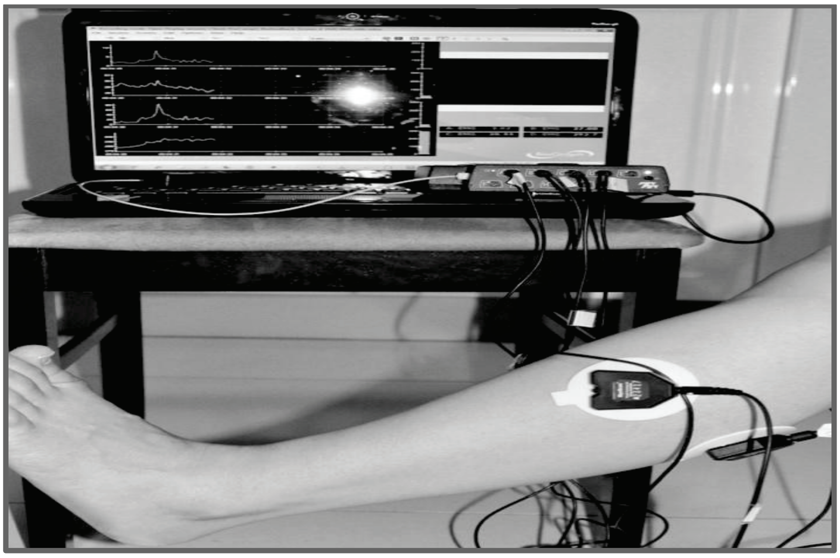


Figure 4. 5 Collect data signals from the leg using sEMG device

The classification was dependent on the sample data from which it acquired. The data collected under the Supervision of the Stroke Coordinator at the Metro-Rehab Hospital. The data collected at three phases at each posture (flexion or extension) of the knee joint. The healthy and unhealthy subject asked to:

1. First trial: move (or get help to move if required) his/her lower limb at the knee joint to flexion and extension (bend and straighten) from the rest position while he/she was sitting on a chair. Each trial takes 3 seconds, and they had a rest of 5 seconds between the two trials. (We take two trials for this case) as shown in Figure 4.6 and Figure 4.7.

2. Second trial: move (or get help to move if required) his/her foot up and down as much as he/she can from rest position while he/she was sitting on the chair. Each trial takes 3 seconds, and there had a rest of 5 seconds between the two trials. (We take five trials for this case) as shown in Figure 4.8 and Figure 4.9.

3. Third trial: move (or getting help to move if required) his/her lower limb (foot and leg) at the knee joint to flex or extend (bend and straighten) with Extension Plantar flexion and Flexion Dorsiflexion from the rest position while he/she was sitting on the chair. Each trial takes 3 seconds, and they had a rest of 5 seconds between two trials. (We take two trials for this case) as shown in Figure 4.10 and Figure 4.11.

Therefore, 12 seconds of data collected for each of the subject trials and 156 seconds for all subject's repetitions. With the offline classification, we conducted 3-fold cross-Validation. To measure the Dorsal Plantar Flexion range for the leg, we used the Goniometric Measurement such as Protractor (Angle Finder and Bevel Square Head) as shown in Figure 4.12. Table 4.2 presents the characteristics for sick and healthy subjects collected from Metro Rehabilitation Hospital.

Table 4. 2 presents the characteristics for sick and healthy subjects collected for Hospital

Simulation ID, for sick & healthy subjects	Gender	Age (Years)	Height (cm)	Weight (kg)	Min KFA* (deg)	Speed (m/s)*	Time since onset of FD (months)	Use electrical stim in rehab. (No/Yes) Bioness Plus L300
Mild 1	F	45	155.7	54.9	15	-------	48	Yes
Mild 2	F	52	160	54.7	22	-------	18	Yes
Mild 3	M	61	176	108	18	-------	15	No
Mild 4	M	64	162	102	15	-------	36	No
Moderate 1	M	84	147	78.2	30o	0.175	1	Yes
Moderate 2	F	68	165	89.4	45	------	3	No
Moderate 3	M	82	172	70	38	------	3	No
Severe 1	M	22	134.6	47	50	0.454	24	Yes
Severe 2	M	68	153	64.7	50	-------	36	Yes
Severe 3	F	60	154.5	81	65	-------	30	Yes
Unimpaired 1	M	60	165	65	0	1.20		No
Unimpaired 2	F	45	160	71	0	1.02		No
Unimpaired 3	F	18	163	64	0	1.17		No

Figure 4. 6 Knee Flexion with Flexion Dorsiflexion

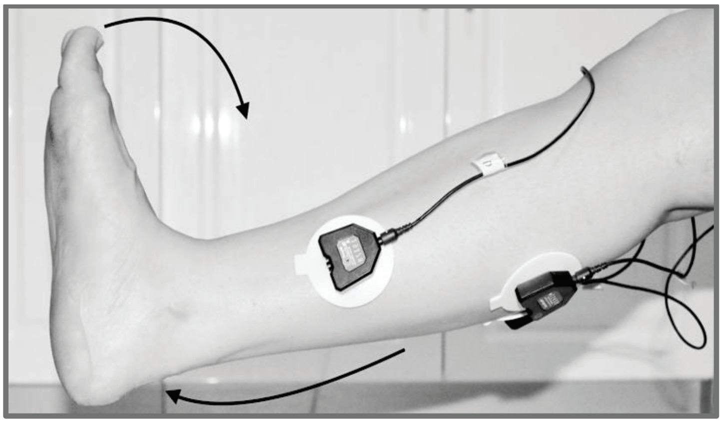


Figure 4. 7 Knee Extension with Flexion Dorsiflexion

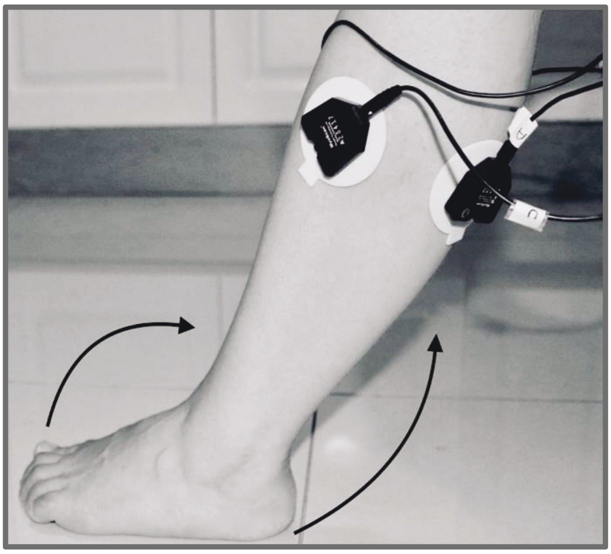


Figure 4. 8 Knee Flexion with Flexion Dorsiflexion

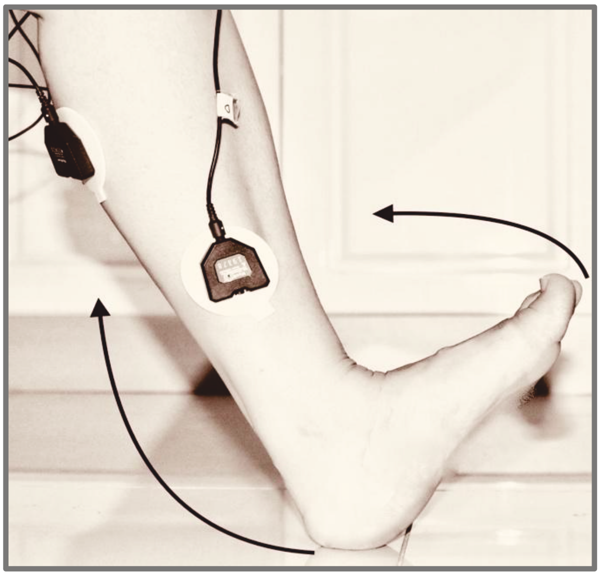


Figure 4. 9 Knee Flexion with Flexion Dorsiflexion

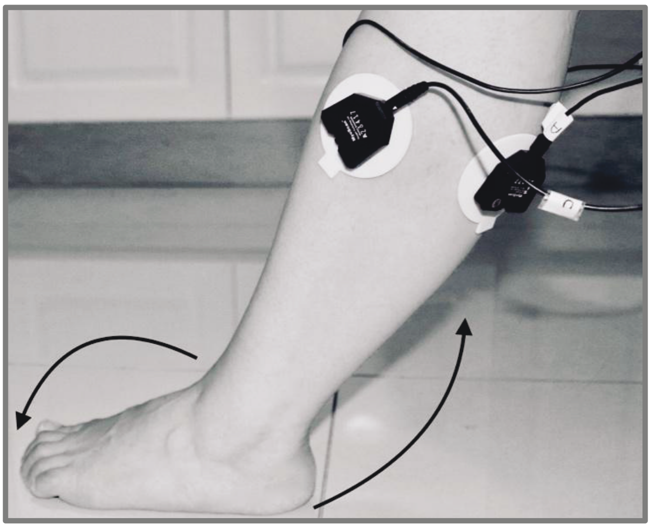


Figure 4. 10 Knee Flexion with Extension Planter Dorsiflexion

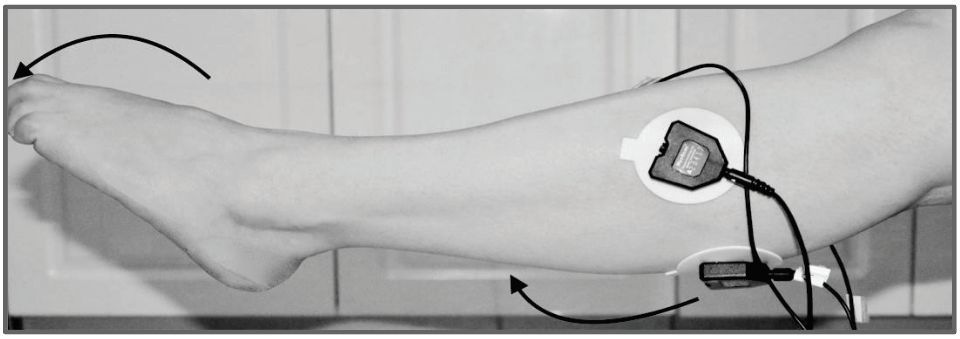


Figure 4. 11 Knee Extension with Extension Plantar-flexion

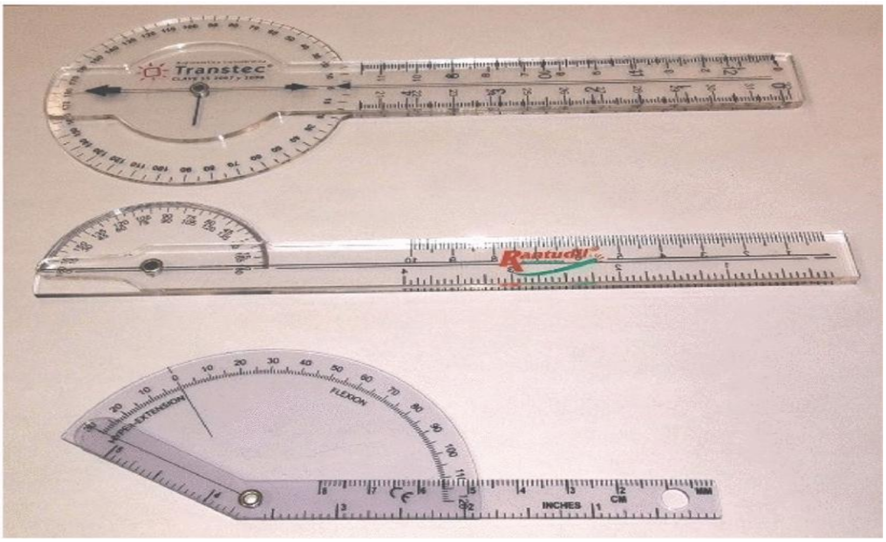


Figure 4. 12   Measure Dorsal Plantar Flexion rang for the leg [104]

For Metro- Hospital dataset, we collected the Surface electromyography (EMG) signals from 13 subjects, from Rectus Femurs (RF), Gastrocnemius (Gas), Soule (Sol) and Tibias Anterior (TA). The OpenSim dataset for CG, provides Surface electromyography (EMG) signals which recorded from ten subjects. sEMG collects signals from the Medial Hamstrings (mH), Biceps Femurs long head (BF), Rectus Femurs (RF), Gastrocnemius (Gas), and Tibias Anterior (TA).

Therefore, there are 5sec x 2(repeated) = 10sec each trial, then 10sec x 3 trial = 30sec for each subject, then 30sec x 1000 = 30,000 data for each class or in total, the number of data is 30,000 x 4 classes = 120,000/channel. There are four channels so that the whole number of sample data is 480,000. The data collected are divide into train data and test data using 3-fold cross-validation.

4.3 Thesis contribution-1(Cont.1): Multi Self-Advised Support Vector Machine (m SA-SVM)

4.3.1 Background

Traditional SVM neglects the train data that not separated linearly by the kernels through the training phase. That happens during the overview of the tolerance parameters in the objective function and constraints. Therefore, it will be classified incorrect if data that is identical or similar to the misclassified data appears in the test

set. The reason being, the data closest to the misclassify data is uncertain. This misclassification is not conclusive and can control if the existing data in the training phase not neglected by the Traditional SVM procedure.
The Non-repeating self-advising method for SVM is adapted and extracts consequent knowledge from the training phase without adding more parameters. The misclassified data supplied from two prospective sources. The first source from outliers and the second source from the data that have not separate correctly. Self-Advised SVM processes the neglect data of SVM from the knowledge is obtained from misclassified data. By creating advice, weights depend on the distance for the misclassify train data from the correctly classify train data, and through the application of these weights together with the decision values of SVM in the test phase. These weights assist the procedure in reducing the outlier data[11].

The one-against-all method is the underlying implementation for SVM multi-class classification. It theories *k* SVM models where *k* is the classes number. The *m*th SVM concept is to train each sample in the *m*th class with positive labels, and each other sample has negative labels. Masood and Maali [13, 44], mentioned in their publication that the multi-classification problem should be considered.

The performance of SA-SVM for a single class (2 classes) needs further development for better application in the general use for several numbers of classes in the classification systems. The multi Advised Weight SVM algorithm is innovated to attain greater flexibility with multi-classification experiments. Modification done for SA-SVM which is totally different from traditional SVM, then the modification for *muti*-SA-SVM represents an important contribution for machine learning methods.

4.3.2 The proposed multi SA-SVM

This section clarifies the proposed multi SA-SVM as illustrates in these steps, which partially published in [33].

1- One versus the Rest (OvR) is one of the simple methods to get M-class classifiers and create a set of binary classifiers f^1, f^2... fro^m all of them trained to discrete one class from rest. Then engage them to get a multi-class SVM classification [95].

Equation 4.1 shown that the decision function is:

$$f^m(x) = sign\,(\sum_{\alpha_{i>0}} y_i \alpha_i^m k(x, x_i) + b^m) \qquad (4.1)$$

For N training data $(x_1, y_1), \dots, (x_i, y_i), \dots, (x_N, y_N)$ where $x_i \in \mathcal{R}^n, i = 1, \dots\dots, N$ and $y_i \in \{1, \dots\dots, k\}$ the class for x_i and α_i is the non-negative Lagrange multiplier that conflicts with normal SVM training.

Where the input vector xi corresponds to the ith labeled by yi depending on the number of classes m.

2- Misclassified data samples that are missing in the primary training stage are recognized. The misclassified datasets (MD) in the training stage can be evaluated by Equation 4.2:

$$\mathrm{MD} = \bigcup_{i=1}^{N} x_i \mid y_i \neq \mathrm{sign}\left(\sum_{\alpha_j > 0} y_i\, \alpha_j^m\, k(x_i, x_j) + b^m\right) \qquad (4.2)$$

3- When MD is empty, then the program loop will enter the testing phase, or calculate neighborhood length (NL) for all member label of MD. NL can be calculated in Equation 4.3:

$$\mathrm{NL}(x_i) = \mathrm{minimum}_{x_j}\left(\|x_i - x_j\| \mid y_i \neq y_j\right) \qquad (4.3)$$

Where x_j, j=1…, N are the train sample data which not related to the MD set. At this time, the train data mapped to a larger dimension.

4- The distance among x_i and x_j calculated with Equation 4.4 mention to the related RBF kernel.

$$\|\theta(x_i) - \theta(x_j)\| = (k(x_i, x_i) + k(x_j, x_j) - 2k((x_i, x_j))^{0.5} \qquad (4.4)$$

For all data sample x_k from the test set, the Advised Weight $AW(x_k)$ can be calculated using Equation 4.5, which clarify how close the test sample data is to the misclassified data.

$$\begin{cases} 0 & \forall x_i \in \mathrm{MD}, \|x_k - x_i\| > \mathrm{NL}(x_i) \text{ alternatively}, \mathrm{MD} = \mathrm{NUL}, \\ \sum 1 - \frac{\sum_{x_i} \|x_k - x_i\|}{\sum_{x_i} \mathrm{NL}(x_i)} & x_i \in \mathrm{MD}, \|x_k - x_i\| \leq \mathrm{NL}(x_i) \end{cases} \qquad (4.5)$$

5- The absolute value of the SVM decision for all x_k from the test set are evaluated and scaled to [0, 1, 2, 3… m] dependent on which class the data tests have.

6- For all x_k from the test set, If (AW (x_k) < decision value (x_k) then the output equal to: $y_k = sign\ (\sum_{\alpha_i>0} y_i \alpha_i^m k(x, x_i) + b^m)$ Which is compatible with normal SVM labelling, otherwise will apply Equation 4.6.

$$y_k = y_i \mid (\|x_k - x_i\| \leq NL\ (x_i) \text{and } x_i \in MD) \tag{4.6}$$

4.3.3 Experimental Setup

4.3.3.1 Pre-Processing data

The experiments are executed to check the performance of mSA-SVM as directed through this state of the art of pattern recognition system shown in Figure 4.13:

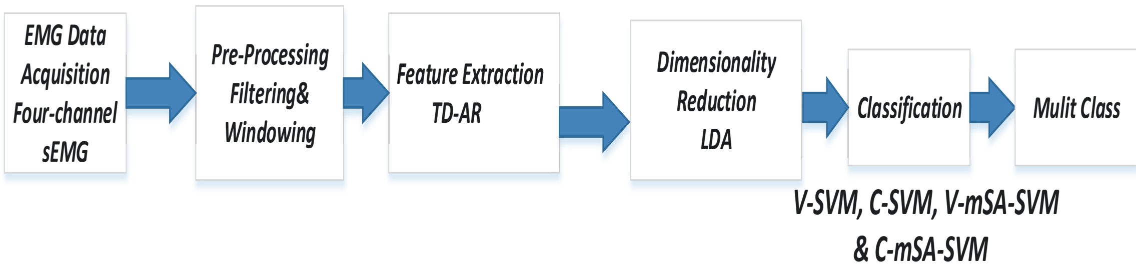


Figure 4. 13 Myoelectric Leg classification using mSA-SVM and other classifiers

Then Signals filtered and window to acquire input signals before the extraction of the data. The Time Domain feature set and Autoregressive model (AR) used. To decrease the dimension of the previews features, LDA activated. The theory behind LDA explains in **Chapter 3 Section 3.2.2**. For the classification stage, the performance of mSA-SVM will be compared with the other: V-mSVM, C-mSVM. The experimental results of mSA-SVM with sEMG signals describe in **Section 4.3.4**. Moreover, the experiments set a benchmark for data sets like the UCI data sets, which are performed to examine the performance of mSA-SVM in different ranges of data sets. A discussion about the findings of mSA-SVM will finalize the section.

4.3.3.2 Feature Extraction

In this experiment, we worked within a feature set that gives the best accuracy than any other set. In **chapter 3 section 3.2.1** the feature set consists of Absolute Value (MAV), Mean Absolute Value Slope (MAVSLP), Root Mean Square (RMS), number of Zero Crossing (ZC), Waveform Length (WL), Sample Skewness (SKW) and Auto Regressive (AR) model Parameters. Moreover, some parameters like an Auto-Encoder as a learning feature and Variance (VAR) included as applied in [2, 105]. This feature set produces 25 x 4 channels = 100 features; all features extracted by using Myolectric toolbox [106] and BioSig toolbox [107].

These parameters verify that the AR model parameters are stable and robust, and resilient to changes in the electrodeposition and signal level [108]. The TD feature listed above was windowed using a disjoint window as a replacement for the sliding window to maintain a low computational cost. A 150-ms window and a 50-ms increment used to form a system, which is appropriate for real data experiment.

4.3.3.3 Dimensionality Reduction

Chapter 3 section 3.2.2 discussed the different categories of the dimensionality reduction technique that tested. For the experiment, we will apply the method based on the speed. However, all these techniques are equitable for real-time application; we preferred to use the fastest one which is Linear discriminant analysis (LDA).

4.3.3.4 Classification

The developing of Myoelectric pattern recognition for lower limb (leg) movement recognition and utilize various components. We use *m*SA-SVM to classify subject level if he/she is mild, moderate, severe or healthy. This section described in details at section 3.2.3.

4.3.4 Experiments and Results

In the experiments, we tested the *v*-*m*SVM, *c*-*m*SVM and *m*SA-SVM to three groups of data sets. **The first data set** is obtained from healthy and unhealthy subjects at the Metro Rehabilitation Hospital as shown in Table 4.2. **The second data set** provided by the Open Sim software and available at the website: www.simtk.org. Data is obtained from this site to test LSA-SVM on children [33].

Table 4.3 represents the characteristics for ten children participated, five of them with cerebral palsy disease walking with a crouched gait and other five are healthy subject. ID labels represent simulation for one gait cycle. The children with cerebral palsy are carefully chosen from a database of subjects, who had experienced motion analysis at Gillette Children's Specialty Healthcare. Principles selecting for the patients comprised: (a) a spastic diplegic cerebral palsy, (b) a minimum knee flexion angle (KFA) during stance more than 15° and (c) a tibial and femoral torsion deformity less than 30°. They requisites from patients not to wear any assistance device while gesture analysis and had not less than two repeated force plate strikes.

The patients with crouch gait CG classified to three cases depend on minimum knee flexion angle during stance. For knee flexion between (15–30)° was organized as Mild crouch gait where knee flexion between (30–50) ° classified as Moderate crouch gait, and knee flexion equal and larger 50° was classified as Severe crouch gait [109] as shown in table 4-3. While for Foot Drop patients, we almost classify their cases depending on minimum knee flexion angle during sitting. For KFA between (15–30)° was classified as Mild FD whereas for KFA between (30° –50°) was classified as Moderate FD and for KFA equal and larger 50° was classify as Severe FD as shown in table 4-2 [89].

The third data set is provided from the UCI machine learning repository data set. The parameter adjustments selected by irregular grid-search technique [110]. Typically, the range of the parameters C and γwas divergent between one classifier to another. For the SVM and SA-SVM, the C and γ ranges were nearly the same that is (2^{-9}, 2^{-8}… 2^{9},2^{10}).

Table 4. 3 Characteristics for sick patients having a crouch gait CG and Healthy subject from SimTik website.

Simulation ID, for sick & healthy subjects	Gender	Age (Years)	Height (cm)	Weight (kg)	Min KFA* (deg.)	Speed (m/s)*
MI01	M	9.4	131.0	28.2	15.5	0.94
MI03	M	9.1	127.6	23.1	21.1	0.67
MO04	F	11.0	143.0	28.7	32.6	1.2

SE02	M	13.2	144.0	35.9	60.3	0.8
SE05	M	16.3	160.0	50.8	86.0	0.7
GIL01(healthy)	F	10.2	0.77	41.1	00.0	1.01
GIL02(healthy)	F	14.6	0.90	66.0	00.0	1.21
GIL04(healthy)	F	11.3	0.72	32.4	00.0	1.15
GIL06(healthy)	F	14.5	0.94	61.9	00.0	1.12
GIL08(healthy)	M	7.0	0.66	26.1	00.0	1.15

4.3.5 Experiments on Myoelectric Leg motion classification

The experiment applied different methodologies to test the performance of *m*SA-SVM. The comparison of *m*SA-SVM and other renowned classifiers such as *v*-*m*SVM, *c*-*m*SVM and *m*SA-SVM completed. Some analysis will be given in each experiment. We should mention that results (average accuracy and training time) for the third data set, benchmark dataset, and UCI will be discussed within the third contribution section where other classifiers discussed.

Tested the performance of *m*SA-SVM and the other contributions in previous sections and compared the results with other well-known classifiers to detect leg movements (healthy or non-healthy) for a single classification. Multi-classification will use for individuals with Foot Drop at various stages and crouch gait cases.

We have windowed the data by the adjoining window using (200 ms) window size and (50ms) increment. While for Open Sim data set for the Crouch Gait patients CG, EMG sampled at 1080 Hz, band-pass filtered within 20 to 400 Hz, rectified, and low-pass filtered at 10 Hz. EMG for all muscles normalized from zero to one [89, 109].

To obtain the best classification performance for the parameters of the classifier were determined using a grid-search method. Table 4.4 offers the optimal parameters for all classifiers. I am utilizing **5-fold** cross-validation to validate the classification performance The Statistical demography for Genders (subjects) shows that 52.17% of Male participate in this experiment while 47.83% of Female whereas statistical demography for age clarifies that this experiment engaged 43.48% of children, 21.17%

of adult and 34.78% of seniors. As we can see from statistical demography for the gender that the ratio for male and female nearly equal while for age we choose to deal with children as half of the participant and another half of adult and senior. In addition to that, this thesis aims for EMG based generalization system that works independent from the gender and other demography factor. This experiment involved four classifiers: v-SVM, c-SVM, c-*m*SA-SVM, v-*m*SA-SVM as can be notice from Table 4.5 and Table 4.6 with plotting results shown in Figures 4.14 and Figure 4.15.

Table 4. 4 Optimal parameters for all classifiers

Dataset	C-SVM		ν-SVM		Feat. type	Win. Size, Win. Inc.
	C	γ	C	γ		
FD-Hospital	2	8	214	0.125	14	200,50
crouch gait	100	8	214	8	14	20,5

Table 4. 5 The average classification accuracy of mSA-SVM across two datasets using five-fold cross-validation compared with C-SVM and V-SVM

Dataset	Accuracy (%)			
	SVM C	SVM ν	SA-SVM C	SA-SVM ν
FD-Hospital	88.39	86.24	90.20	**90.24**
crouch gait CG	67.00	66.12	**70.58**	70.24

Table 4. 6 The training Time consumption for four types of classifier

Dataset	Training Time (ms)			
	SVM C	SVM ν	SA-SVM C	SA-SVM ν
FD-Hospital	**221.8**	276.9	223.6	256.1
crouch gait	400.4	503.8	**250.8**	259.2

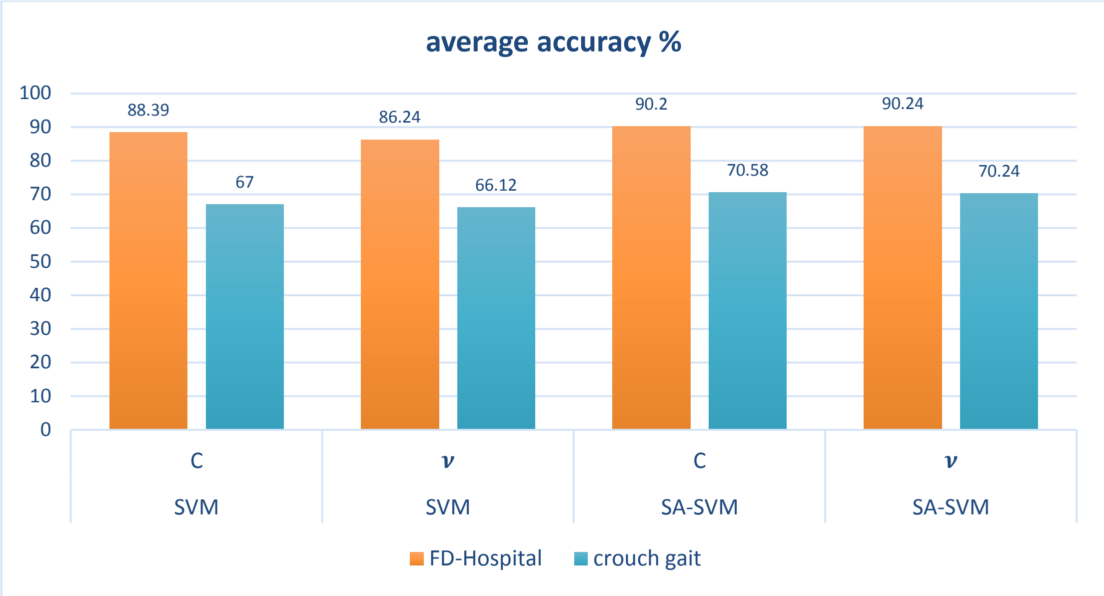


Figure 4. 14 The average classification accuracy of mSA-SVM across two datasets using five-fold cross-validation compared with C-SVM and V-SVM

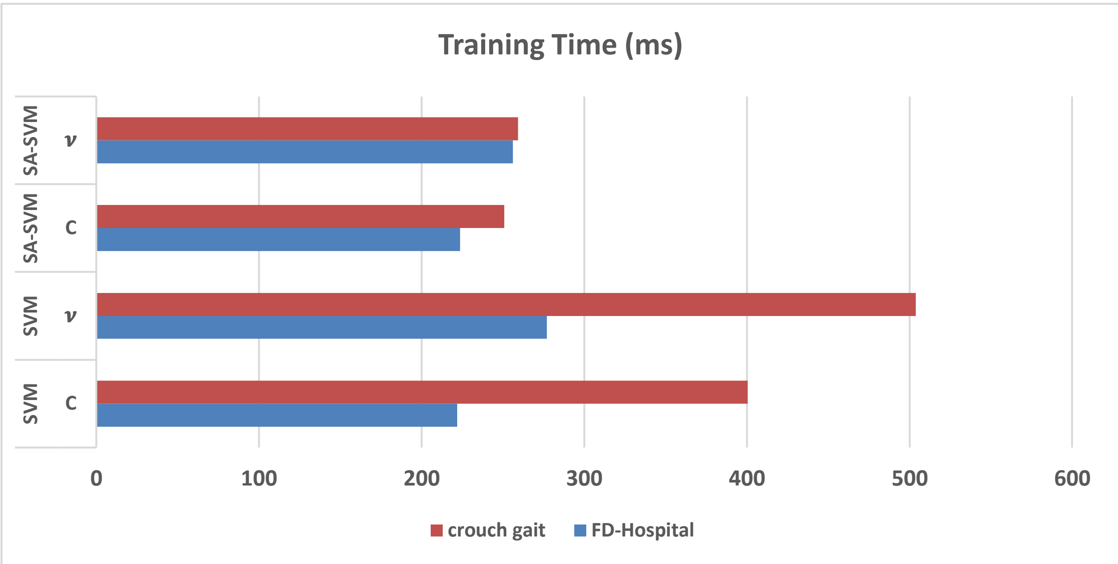


Figure 4. 15 The training Time consumption for four types of classifier

4.4 Thesis contribution-2 (Cont.2): Label Self - Advised Support Vector Machine (LSA-SVM).

One of the most general machine learning algorithms for classification and regression is the Support Vector Machine (SVM). SVM shows exceptional performance when applied in various applications. However, different methods had projected to advance their performance in general or in individual cases.

This thesis proposes a new method to improve SVM performance in general. This method applies to each type of SVMs, which have different kernel types.

The main purpose of recommending this method is to transfer more data from the training phase to the testing phase. This data learned from the misclassified data of the training phase, which ignored the traditional SVM. Experimental results from three groups of data sets that we work on as part of the first contribution in section 4.2 confirm that this methodology can improve the classification performances of C-SVM, V-SVM, and SA-SVM without any additional parameters to the learner algorithm.

This section continues with the background of SA-SVM and elaborates on the simple concepts of SA-SVM. The experiments on Myoelectric pattern recognition for leg movement classification were tables and displayed in a flowchart. The Myoelectric signal collected from Hospital, the OpenSim CG signal and UCI machine learning data sets for the single class was executed to examine the performance of LSA-SVM in more extensive applications. We will conclude with a discussion of LSA-SVM.

4.4.1 Background

The standard SVM disregards the train data that is not separate linearly by the kernels through the training stage, which occurs during the outline of the tolerance parameters in the impartial function and restrictions. For this reason, it will be classified incorrectly if the data is similar or matching the misclassified data that appears in the test set. Ti is happened when the data that is close to the misclassified data is unspecified. Thus resulting in a misclassification that is not sensible and not controlled [13]. Many researchers adapted versions of SVM which are present in chapter 3, targeting to rise classification efficiency and performance for some applications. **Use of the label classification on expert knowledge is one of the methods that was used without increased cost since there will be no addition of an extra parameter.**

For example, Mohammed, A.A propose novel methodologies for texture analysis to improve the single-label classification of the facial features[111]. Masood, A. et al. [13] suggested to engage label classification with class classification techniques; in addition, they dealt with the problem of limited labeled data available, especially for histopathological images. They proposed a novel learning model created on deep belief neural net and semi advised SVM to make effective use of labeled data along with unlabelled data for the training phase. It displayed improved performance when matched with different state of the art approaches for skin cancer diagnosis. The proposed model tested for skin [11]. Several applications in research work with the use of label classification methods LCM to improve: the progression of the search for related information on Twitter, so that five different labels are defined to categorize tweets, including news. The system proposed to analyze complex motion in events; it combines the track and multilabel hypergraphs of moving targets in video sequences. The Adaptation Methodologies: state some classification models were primarily intended to resolve binary problems and then were expanded to solve multi-class problems which SVM as an example of this classifiers. By contrast, other methods can efficiently work with several classes. A kNN classifier is capable of predicting the various class amongst the neighbors of the new data. Similarly, trees, neural networks, classification rules, and statistical models, etc., have been applying to look over both binary and multi-class classifications.

Non-supervised methods, such as Principal Component Analysis (PCA) and Latent Semantic Indexing (LSI) are capable of minimizing the dimensionality by analyzing only the input attributes. These techniques can apply to any multi-label data set (MLDs) because there is no need knowledge about the output variables.

We briefly discussed and explained in **chapter 2 section 2.4.7** about Label Classification methods, and we gave examples for some methodology that apply LC in **chapter 3 section 3.2.3.4**

As clarified in the previous section, the goal of the multi-label classifier is to predict a set of appropriate labels for a new instance of data. This section described numerous application areas, which can take advantage of this functionality while providing details of its specific uses in cases, which fit each of these areas.

4.4.2 Concepts of LSA-SVM

The following concepts describe the LSA-SVM for single classification method, which is partial reported in the submitted Journal paper:

1- Applying the decision function as in Equation 4.7 to classify hyperplane:

$$f(x) = \operatorname{sign}(\sum_{\alpha_{I>0}} y_i \alpha_i \, k(x, x_i) + b \qquad (4.7)$$

Considering x_i as a vector for input data depending on the ith sample, which labeled as y_i , relating to its class.

The pair (w, b) described in the hyperplane with the equation $< w, x > +b = 0$

N samples$(x_1, y_1), \dots, (x_i, y_i), \dots, (x_N, y_N) \in \mathcal{R}^n \times \{\pm 1\}$,

$(\alpha_1, \alpha_2, \dots, \alpha_N)$ Is N non-negative Lagrange multipliers

N Non-negative variables $(\xi_1, \xi_2, \dots, \xi_N)$

Where $x_{i,}$, is the input vector for the ith sample that labeled with y_i related to its class.

2- Misclassified data samples in the first train phase are recognized. The misclassified datasets (MD) in the training phase is calculated by Equation 4.8

$$MD = \cup_{i=1}^{N} x_i \; |y_i \; \neq \operatorname{sign}\left(\sum_{\alpha_j>0} y_i \, \alpha_j \, k(x_i, x_j) + b\right) \qquad (4.8)$$

3- The algorithm indicates that: If MD is empty, then go to the testing phase, or else calculate neighborhood length (NL) for each y_i of label MD. Equation 4.9 defined NL.

$$NL(y_i) = \; \text{minimum}_{y_j} \left(\|y_i - y_j\| \mid x_i \; \neq \; x_j\right) \qquad (4.9)$$

Where y_j, j=1, …., N is the label of training data that do not belong to the label of MD set. The label of training data is mapping to a higher dimension, the distance between y_{ou} and y_j computed according to the following of Equation 4.10 and Equation 4.11 regarding the related RBF kernel.

$$\|\emptyset(y_i) - \; \emptyset(y_j)\| = (k(y_i, y_i) + \; k(y_j, y_j) - 2k((y_i, y_j))^{0.5} \qquad (4.10)$$

i.e., RBF will be: $K(y_{i,}y_j) = e^{-\gamma|y_i - y_j|^2}$ (4.11)

1- For each label y_k from test data, the Lab Advised Weight LAW (y_k) figures out as Equation 4.12. These LAWs represent how close the label test data are to the label of misclassified data.

$$\text{LAW (yk)} =$$

$$= \begin{cases} 0 & \forall x_i \in MD, \|y_k - y_i\| > NL(y_i) \text{ or } MD = NUL, \\ \sum 1 - \frac{\sum_{y_i} \|y_k - y_i\|}{\sum_{y_i} NL(y_i)} & x_i \in MD, \|y_k - y_i\| \leq NL(y_i) \end{cases} \quad (4.12)$$

The absolute value of the SVM decision for each x_k from the test set is calculated and scaled to [0, 1]. Figure 4.16, explain the flow chart the steps above.

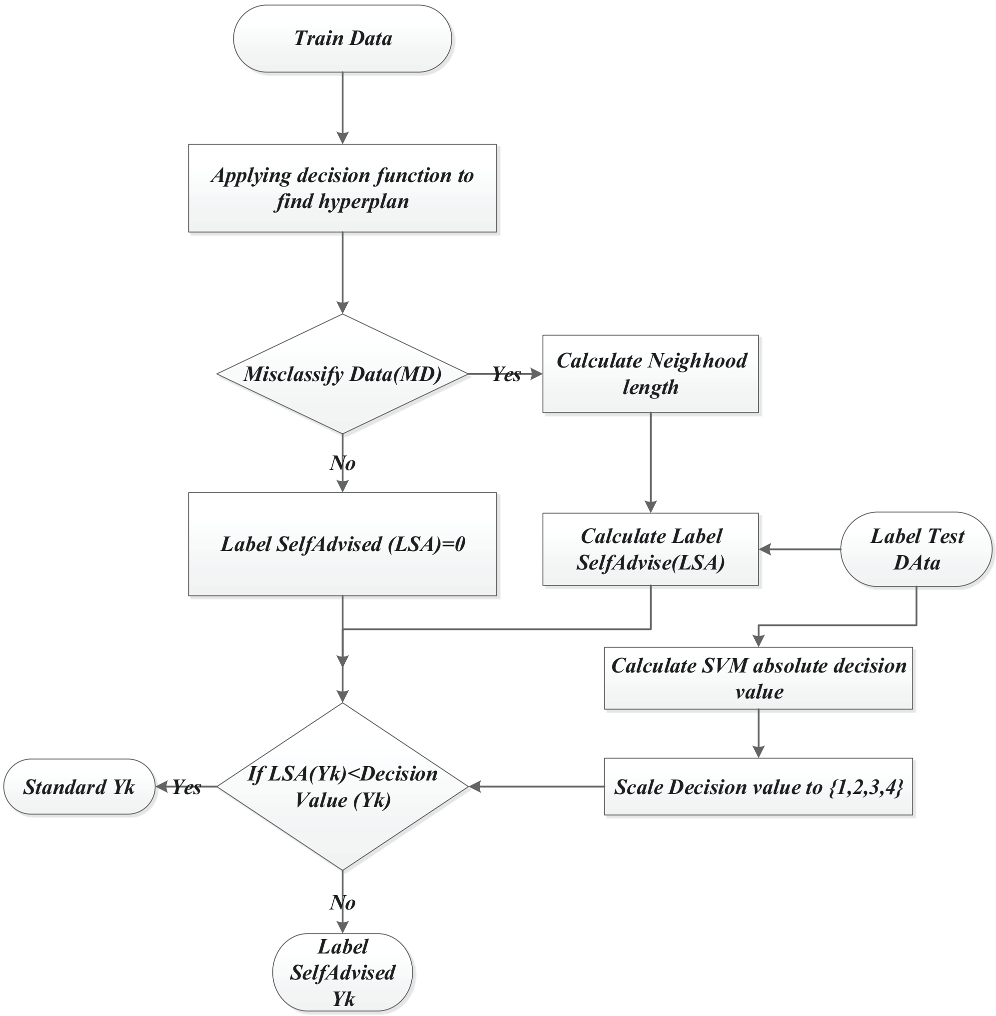


Figure 4. 16 Flowchart for the LSA-SVM method.

2- For each y_k from the label of the test set, If (LAW (y_k) < decision value (y_k) then $y_k = \text{sign}\left(\sum_{\alpha_j>0} y_j \alpha_j\, k(x_k, x_j) + b\right)$ which is compatible with normal SVM labeling, otherwise $y_k = y_i \mid (\|y_k - y_i\| \leq NL(y_i)$ and $x_i \in MD)$

4.4.3 Experiments and results

The experiments check the performance of LSA-SVM for a single class as conducted through the state of art of pattern recognition system shown in Figure 4.17 with the only change to the classifications stage.

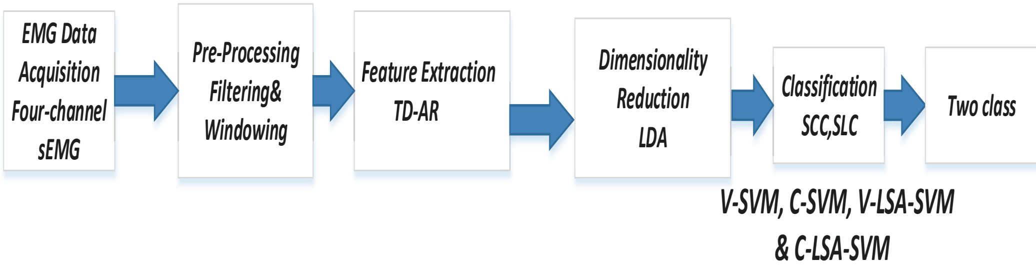


Figure 4. 17 Myoelectric Leg classification using mLSA-SVM and other classifiers

Each stage shown in Figure 4.17, discussed in detail in **section 4.2.3**. The experiments applied v-SVM, c-SVM and LSA-SVM into three different groups of data sets were explained briefly in **section 4.2.4**. **First data set** collected from healthy and unhealthy subjects at Metro Rehabilitation Hospital. **Second data set**, is provided by the Open Sim website: www.simtk.org [33], while the **Third data set** provided by the UCI machine learning repository data set which had seven groups of data set in different applications.

4.4.3.1 Experiments on Myoelectric Leg movement recognition

Various experiments were done to test the performance of LSA-SVM in Myoelectric pattern recognition. First, we adjusted the parameters of C and γ from the range of (2^{-9}, 2^{-8}… 2^{9},2^{10}), as shown in Table 4.7, then we examined the performance of accuracy while we apply *v*-SVM, *c*-SVM, *v*-SA-SVM, *c*-SA-SVM and LSA-SVM. We compare all the classifiers. Some analysis is given in each experiment, the first experiment we used the **first data set FD** and the **second data set CG** to estimate accuracy for each classifier. The Second experiment applies the **third data sets from UCI.**

Table 4. 7 Optimal parameters for all classifiers

Dataset	C-SVM		ν-SVM		Feat. type	Win. Size, Win. Inc.
	C	γ	C	γ		

Hospital	100	0.003	100	0.003	14	50,15
crouch gait	100	0.003	100	0.003	14	20,5

Figure 4.18 and table 4.8 specify that the accuracy of LSA-SVM was higher than SA-SVM and other classifiers in all types of data set. The average accuracy for *v*LSA-SVM as compared with all other five classifiers, is little higher equal to **99.06%,** while in *v*SA-SVM equal to **98.75% for** FD Hospital dataset. However, for CG OpenSim dataset *v*LSA-SVM gives **82.01%** and *v*SA-SVM equal to **80.01%** with five-fold cross-validation training. Table 4.9 and Figure 4.19 presents that training time of *v*SA-SVM in two groups of the dataset is much faster than *v*LSA-SVM. In that case, LSA-SVM method was not achieved the best performance, whereas time consumption equal to **69.4msec** for *v*LSA-SVM and **46.3msec**. For *v*SA-SVM, but it is still reaching a real-time which is less than 300msec as standard real-time measuring.

Overall, in most cases, the adaptation of the Label function used in LSA-SVM can improve the performance of the standard support vector machine and SA-SVM.

Table 4. 8 The average classification accuracy of LSA-SVM across two datasets using five-fold cross-validation compared with other classifiers

Dataset	Accuracy (%)					
	SVM C	SVM $\boldsymbol{\nu}$	SA-SVM C	SA-SVM $\boldsymbol{\nu}$	LSA-SVM C	LSA-SVM $\boldsymbol{\nu}$
Hospital	92.18	93.26	98.86	98.75	98.99	**99.06**
crouch gait	76.20	70.10	81.73	80.01	81.97	**82.01**

Table 4. 9 The training Time consumption for each type of classifier

Dataset	Training Time (ms)					
	SVM C	SVM $\boldsymbol{\nu}$	SA-SVM C	SA-SVM $\boldsymbol{\nu}$	LSA-SVM C	LSA-SVM $\boldsymbol{\nu}$
Hospital	55.3	54.3	107.9	**46.3**	70.7	69.4
crouch gait	48.5	55.1	36.3	**38.0**	48.9	48.4

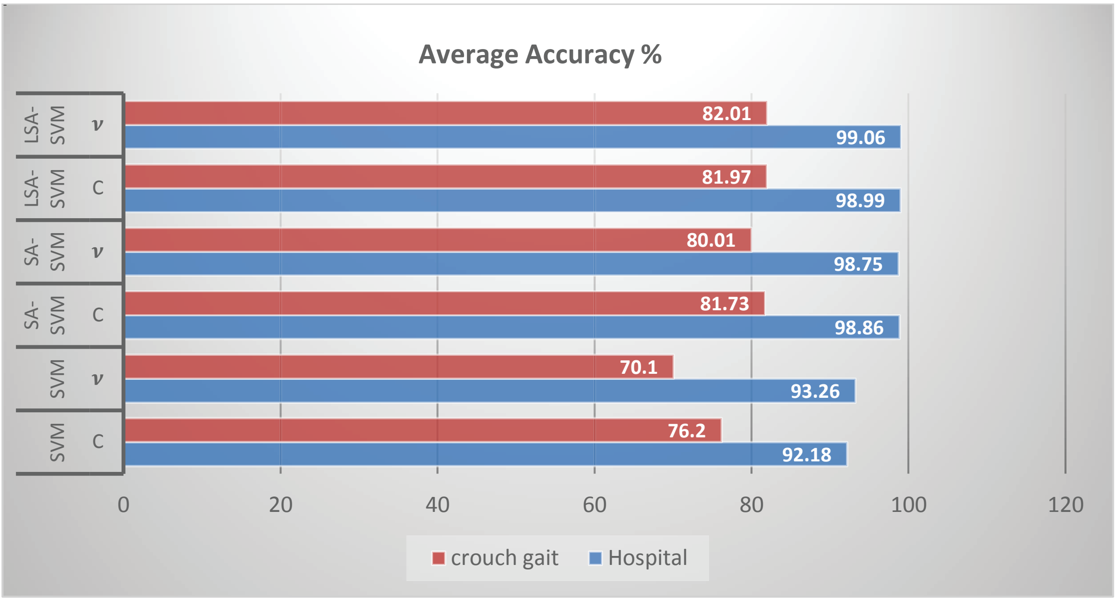


Figure 4. 18 The average classification accuracy of LSA-SVM across two datasets using five-fold cross-validation compared with other classifiers

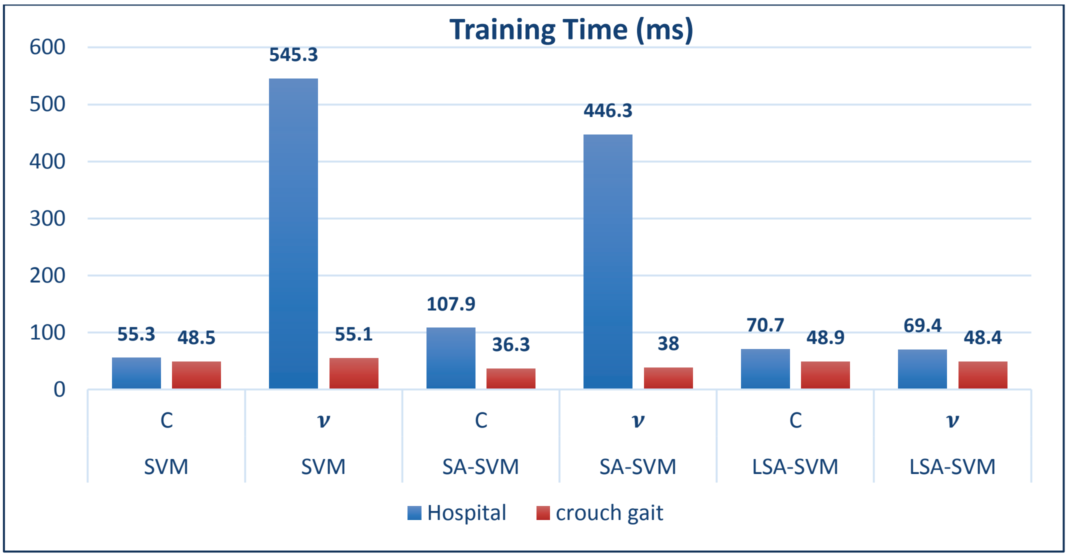


Figure 4. 19 The training Time consumption for each type of classifier

4.4.3.2 Experiments on UCI datasets

LSA-SVM proved that it is capable of classifying four classes of leg movements consisting of a combination of three classes of the Unhealthy leg (Mild, Moderate and Severe patients) and the fourth is Healthy class. This section investigates the performance of LSA-SVM on the benchmark dataset that is accessible online on the UCI machine learning website. The experiment depends on the size of the data. We implement 3-fold cross-validation on the Larger sized data, while we executed 5-fold cross validation for the small and medium-sized data. Table 4.10 shows the data specification for a benchmark for 2-Classes data sets types.

Table 4. 10 Data specification for benchmarking for 2-Classes datasets

Dataset	Group	# Data	# Attributes
Iris	**Small size**	100	4
Parkinson		195	22
Australian Credit Approval (Statlog)	**Medium size**	689	4
Breast Cancer		699	9
Pima Indians diabetes		768	8
Spambase	**Large size**	4601	57
Skin Segmentation		245057	3

This experiment involved six classifiers: v-SVM, c-SVM, c-SA-SVM, v-SA-SVM, c-LSA-SVM, v-LSA-SVM. The optimal parameters are established and noted its effects on the accuracy and performance of the classifier. Table 4.11 explain the parameters that will used in each classifiers

Table 4. 11 The optimal parameters used by each classifier in the UCI dataset experiments for 2 - classes

Dataset	C-SVM		ν-SVM		Feat. type	Win. Size, Win. Inc.
	C	γ	C	γ		
Iris	1	1	1	0.25	11	20, 5
Parkinson	300	0.125	3000	0.003	14	50,15
Australian Credit Approval (Statlog)	0.5	0.125	2^{14}	0.003	14	20,5
Breast Cancer	8200	0.003	2^{14}	0.003	13	50,7
Spambase	3000	0.003	2050	0.005	14	500,120
Pima Indians diabetes	8200	0.0078	2050	0.002	14	50,15
Skin Segmentation	100	0.003	1	2	14	500,150

Table 4.12 and Figure 4.20 show that LSA-SVM has performed reasonably through a seven different data sets with 2 classes. The comparison of LSA-SVM and SA-SVM shows that average accuracy is quite similar in some dataset. As for LSA-SVM, the accuracy of v-LSA-SVM is significantly better than c-LSA-SVM only in "**Breast**

Cancer" equal to 92.09 % dataset. From the results, recognized that LSA-SVM is the most accurate classifier across seven datasets equal to 96.42 in "**Australian Credit (Statelog)**" except "**Pima Indians diabetes**" and "**Spambase**" datasets that is equal to **83.31%** and **62.52%** respectively.

The processing time (Time consumption) of classifiers also calculated. Table 4.13 provides the training time. Figure 4.21 clarifies that the training time of LSA-SVM is one of the slowest classifiers, equated to other classifiers in overall data sets.

Its performance worsens when *v*LSA-SVM works on big data like "Skin Segmentation" dataset equal to **133.1msec**. The *c*LSA-SVM is the slowest classifier take around **267ms** to learn "**Skin Segmentation**" datasets while *v*SA-SVM shows the faster time equal to **0.537msec.**

Table 4. 12 The accuracy of seven classifiers on various data using 5-fold cross validation for small and medium-size data and 3-fold cross validation for large size data for 2 classes

Dataset	Accuracy (%)					
	SVM C	SVM ν	SA-SVM C	SA-SVM ν	LSA-SVM C	LSA-SVM ν
Iris	92.85	90.12	85.71	85.84	92.85	**92.86**
Parkinson	72.32	70.0	71.43	64.29	71.43	**74.10**
Australian Credit (Statelog)	81.6	80.8	88.07	88.53	91.28	**96.42**
Breast Cancer	89.88	88.76	90.96	90.33	**92.09**	91.01
Pima Indians diabetes	**82.5**	80	78.4	77.5	80.95	78.3
Spambase	61.17	**62.52**	61.18	62.53	62.19	61.26
Skin Segmentation	83.31	72.26	60.50	67.47	84.23	**84.43**

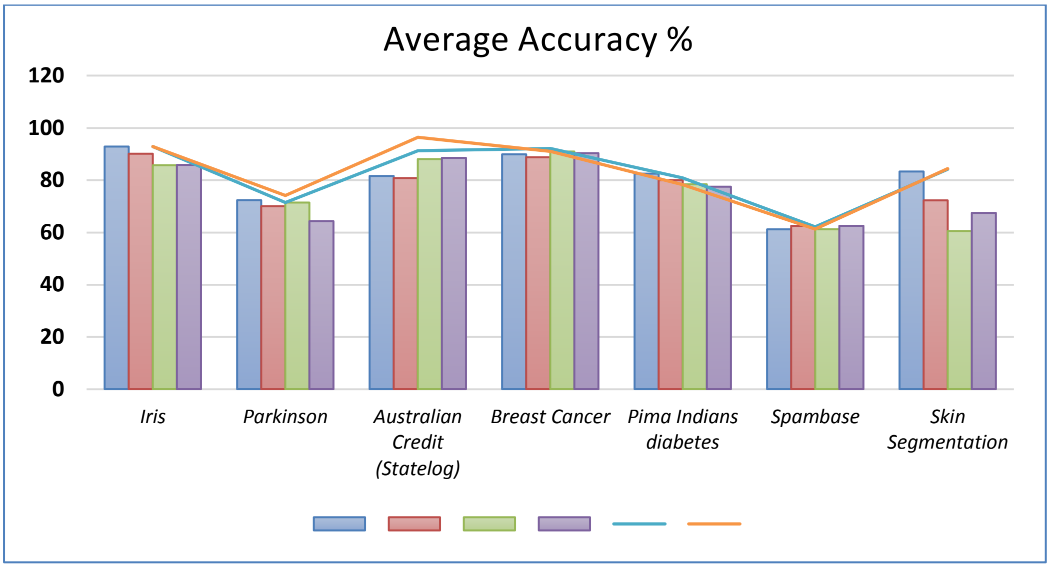


Figure 4. 20 The accuracy of seven classifiers on different data using 5-fold cross validation for small and medium-size data and 3-fold cross validation for large size data for 2 classes

Table 4. 13 The training Time consumption for each type of classifier

Dataset	Training Time (ms)			
	SA-SVM C	**SA-SVM ν**	**LSA-SVM C**	**LSA-SVM ν**
Iris	0.837	0.701	3.6	4.1
Parkinson	0.630	**0.537**	4.8	4.3
Australian Credit (Statelog)	1.10	1.60	4.6	5.2
Breast Cancer	3.20	3.3	7.20	8.70
Pima Indians diabetes	0.901	0.836	5.0	4.7
Spambase	25.8	16.1	34.8	33.4
Skin Segmentation	121.4	86.3	**267.2**	133.1

As shown in table 4.14, which addresses the statistical analyses (standard deviation) for Foot Drop patient and healthy participants

Table 4.14 statistical analyses (standard deviation) for Foot Drop patient and healthy participants

	No. of subjects (with Bioness L300)	No. of subjects (without Bioness L300)	Age(years)	Height (cm)	Weight (kg)
Unimpaired	--	3	17.38	3.56	5.35
Mild Foot Drop	2	2	7.5	7.56	50.38
Moderate Foot Drop	1	3	7.12	11.40	13.77
Severe Foot Drop	3	---	12.94	15.67	24.05

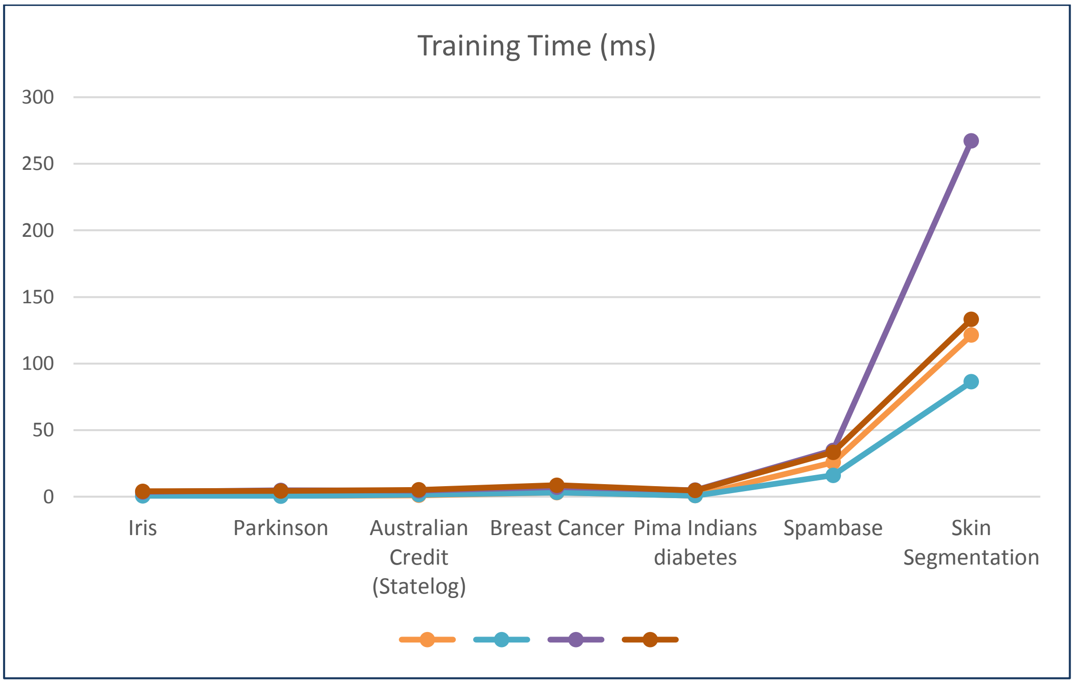


Figure 4. 21 The training Time consumption for each type of classifier

4.5 Summary

The implementation of mSA-SVM and LSA-SVM for Myoelectric leg motion recognition is **one of the main contributions of the thesis**. The other contribution offered in this section describing LSA-SVM. The advantage of LSA-SVM over SA-SVM exists on working with Label data instead of data value itself. Therefore, this is **the second contribution of LSA-SVM**. In addition to the Myoelectric leg motion classification, LSA-SVM has been apply in an extensive collection of classification problems using UCI machine learning data set. The experimental results show that LSA-SVM performs on a wide range of data set size, (small to large). Overall, LSA-SVM is a promising classifier for several classification applications, particularly for Myoelectric pattern recognition.

Implemented the use of Label Self-Advised Support Vector Machine (LSA-SVM) and projected the Self-Advised Support Vector Machine (SA-SVM) for leg motion recognition using sEMG signals. Overall, LSA-SVM could classify four leg movements with an accuracy of 99.06 %. They are deeming its comparably to renowned classifiers such as SA-SVM, SVM. Therefore, LSA-SVM could improve the performance of the advised-based SVM. While comparison training test was executed to compare results of LSA-SVM with all other algorithms, the *p*-value correlated to the ANOVA test by a level of impact $\alpha = 0.05$ was 0.038, which shows the big statistical difference between these groups. Therefore, it can be concluded that LSA-SVM achieved good results than these algorithms. Besides, 68.80 % sensitivity and 76.5 % specificity achieved to classify the testing data for Hospital. A full information and equations of these methods ANOVA test, Sensitivity and Specificity are present in Appendix C.

The tables clarified the short length of data helps classify the severity of foot drop; this appears clear when learning machine training the data with minimum time training and perfect accuracy more than the long data.

Chapter 5

Novel *m*LSA-SVM and ELM-LSA-SVM Based Classification for Foot Drop

5.1 Introduction

This chapter presents the last two contributions to this thesis, incorporating the theoretical and practical systems. **Section 5.2** discusses the third contribution (Cont.3) that focus on developing Cont.2 to the multi-classes. The four contributions will be considered in **section 5.3**, which describes the performance of the engaging ELM method with LSA-SVM.

5.2 Thesis contribution-3 (Cont.3): Multi-Label Self -Advised Support Vector Machine (*m*LSA-SVM).

This section proposes a new multi-class classifier using Support Vector Machine named multi Label Self-Advised Support Vector Machine (*m*LSA-SVM). *m*LSA-SVM is used to classify more than two class in pattern recognition.

We begin with a brief background then followed by the basic concepts of *m*LSA-SVM. The experiments on Myoelectric pattern recognition for leg motion classification will be described later. Not limited to the Myoelectric signal, the experiments on UCI machine learning data sets were able to study the performance of *m*LSA-SVM in large applications. The results and applications of *m*LSA-SVM will conclude the discussion.

5.2.1 Background

Model label correlations have a place in multi-label classification specifically for small dimensions data sets such as CM-based methods etc.[112]. While for a wide range of data sets and evaluation measures, ECC models correlations use a methodology that is effective and not expanded to over-fit. Another investigates technique improves current pairwise rank created in multi-label image classification: the new pairwise rank loss function directly optimizes the label decision model to control which labels are collected in the output [102]. Multi-label classification has been far more effective in data mining applications. Many of the earliest methods found in a standard group setting. Conversely, a few methods found in the stream setting include iSOUP-Tree method; which lead to the development of a streaming multi-target regression [113]. Text annotation is another essential and relevant application; the project consists of metadata conversion to documents, which need time and strength when conducted by humans. For that reason, the ranges of text mining methods are utilizing to automate the procedure. Mainly, built on a keyword extraction or word counts when applying a keyword as a text classification feature, they discovered that: the number of the training sample is less than the number of extracted features, and this complication disturbs the text classification performance. Other issues are the projection of multiple, non-exclusive labels to the documents (multi-label classification). This complication brands the text classification to compute and complicate the outcome when compared with single-label classification [114].

5.2.2 Concepts of mLSA-SVM

The method Multi-Label Self-Advised Support Vector Machine algorithm is partially reported in the submitted Journal paper. It can be illustrated in the 6 steps below:

1- For the first part, to get the decision function for multi-class SVM classification, Equation 5.1 is applied:

$$f^m(x) = sign\left(\sum_{\alpha_{i>0}} y_i \alpha_i^m k(x, x_i) + b^m\right) \tag{5.1}$$

Considering x_i as a vector for input data depending on the ith sample, which labeled as y_i that related to its class.

The pair$(x, x_i) + b^m$ described in the hyper plane with the equation $< (x, x_i) + b^m = 0$. $(\alpha_1, \alpha_2, \dots, i)$ Is i non-negative Lagrange multipliers

Where $x_{i,}$, is the input vector for the ith sample that labeled with y_i related to its class.

For N training data:

$(x_1, y_1), \ldots, (x_i, y_i), \ldots, (x_N, y_N)$, where $x_i \in \mathcal{R}^n, i = 1, \ldots\ldots, N$ and $y_i \in \{1, \ldots\ldots, k\}$ is the class of x_i, and αi is the nonnegative Lagrange multiplier that conflicts with normal SVM training.

Where input vector xi corresponds to the *i*th labeled by yi, depending on the number of classes m.

2- Misclassified data samples that are missing them in the primary training stage will be recognized in this step. The misclassified datasets (MD) in the training stage can be evaluated by considering Equation 5.2:

$$\mathrm{MD} = \bigcup_{i=1}^{N} x_i \ | y_i \ \neq \mathrm{sign}\left(\sum_{\alpha_j > 0} y_i \, \alpha_j^m \, k(x_i, x_j) + b^m \right) \qquad (5.2)$$

3- The algorithm indicates that: If MD is empty, then go to the testing phase, or else calculate neighborhood length (NL) for each y_i of label MD. Equation 5.3 defined NL.

$$\mathrm{NL}(y_i) = \ \mathrm{minimum}_{y_j} \left(\|y_i - y_j\| \mid x_i \ \neq \ x_j \right) \qquad (5.3)$$

Where y_j, j=1,…, N is the label of training data that does not belong to the label of MD set. The label of training data mapped to a higher dimension, the distance between y_i and y_j is total according to the following Equation 5.4 concerning the related RBF kernel.

$$\|\emptyset(y_i) - \ \emptyset(y_j)\| = (k(y_i, y_i) + \ k(y_j, y_j) - 2k((y_i, y_j))^{0.5} \qquad (5.4)$$

4- For each label y_k from test data, the multi-Labe Advised Weight mLAW(y_k) is figured out as Equation 5.5. These *m* LAWs represent how close the label test data is to the label of misclassified data.

$$m\mathrm{LAW(yk)} = \\ = \begin{cases} 0 & \forall x_i \ \in \mathrm{MD}, \|y_k - \ y_i\| > \mathrm{NL}\ (y_i\ ) \mathrm{or\ MD = NUL}, \\ \sum 1 - \dfrac{\sum_{y_i} \|y_k - y_i\|}{\sum_{y_i} \mathrm{NL}(y_i)} & x_i \ \in \mathrm{MD}\ , \|y_k - \ y_i\| \leq \mathrm{NL}\ (y_i\ ) \end{cases} \qquad (5.5)$$

5- The absolute value of the SVM decision for each x_k from the test set is calculated and scaled to [1, 2,3,4,……m].

6- For each y_k from the label of the test set, If (mLAW (y_k) < decision value (y_k) then

$y_k = sign\ (\sum_{\alpha_i>0} y_i \alpha_i^m k(x, x_i) + b^m)$ This is compatible with normal SVM labelling, otherwise: $y_k = y_i \mid (\|y_k - y_i\| \leq NL\ (y_i)$ and $x_i \in MD$

5.2.3 Experiments and results

The experiments check the performance of mLSA-SVM for multi-classes as conducted through state of the art for the pattern recognition system shown in Figure 5.1.

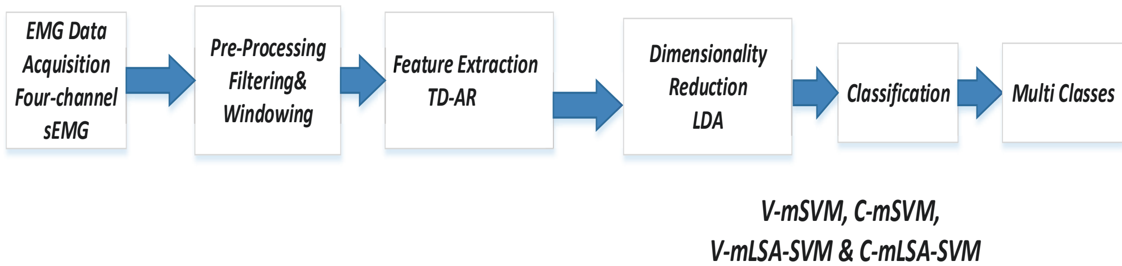


Figure 5. 1 Myoelectric Leg classification using *m*LSA-SVM and other classifiers

Each stage of procedures in Figure 5.1 discuss in detail in **chapter 2 section 2.4.2**. In this experiment, we will apply v-mSVM, c-mSVM, v-mLSA-SVM and c-mLSA-SVM into three different groups of data sets as we explained briefly in **section 5.2.3.1**. The **First data set** collected from healthy and unhealthy subjects at Metro Rehabilitation Hospital. The **Second data set** is provided by the Open Sim website: www.simtk.org [33], while the **Third data set** provided from UCI machine learning repository dataset with seven groups of the dataset in different applications.

5.2.3.1 Experiments on Myoelectric Leg movement recognition

Various experiments were prepared to experiment with the performance of mLSA-SVM in Myoelectric pattern recognition. First, we adjusted the parameters of C and γ from the rang of (2-9, 2-8,…, 29,210) as shown in table 5.1, then we examined the performance of accuracy while we are applying v-mSVM, c-mSVM , v-mSA-SVM, c-mSA-SVM, v-mLSA-SVM and c-mLSA-SVM.

We then compared this for all the classifiers. Some analysis is given in each experiment. The first experiment used the first data set FD and the second data set CG to estimate accuracy for each classifier. The Second experiment applied the third datasets from UCI.

Table 5. 1 The optimal parameters for all classifiers

Dataset	C-SVM		ν-SVM		Feat. type	Win. Size, Win. Inc.
	C	γ	C	γ		
Hospital	2	8	2^{14}	0.125	14	200,50
crouch gait	100	8	2^{14}	8	14	20,5

Table 5.1 and Figure 5.2 specify that the accuracy of *m*LSA-SVM was higher than SA-SVM and other classifiers in all types of dataset. It shows that average accuracy for *mv*LSA-SVM equal to **92.28%** for FD hospital dataset and **71.23%** for *mc*LSA-SVM in CG dataset. In contrast, Table 5.3 and Figure 5.3 indicate that the training time consumption for *m*LSA-SVM is faster than *m*SA-SVM and other classifiers in all types of data set. So that, *mc*LSA-SVM has a time consumption equal to **95msec** in FD dataset which is decidedly faster than all other classifiers, while *mc*SA-SVM has time **250.8msec** in CG dataset which is a little faster than *m*LSA-SVM. For all LSA-SVM, it is still reaching a real-time, which is less than 300msec as standard real-time measuring.

Overall, in most cases, the adaptation of the Label function using it in *m*LSA-SVM can improve the performance of the standard Support Vector Machine and SA-SVM.

Table 5. 2 The average classification accuracy of mLSA-SVM across two datasets using 5-fold cross validation compared with other classifiers

Dataset	Accuracy (%)					
	*m*SVM C	*m*SVM ν	*m*SA-SVM C	*m*SA-SVM ν	*m*LSA-SVM C	*m*LSA-SVM ν
Hospital	88.39	86.24	90.20	90.24	92.24	**92.28**
crouch gait	67.00	66.12	70.58	70.24	**71.23**	71.0

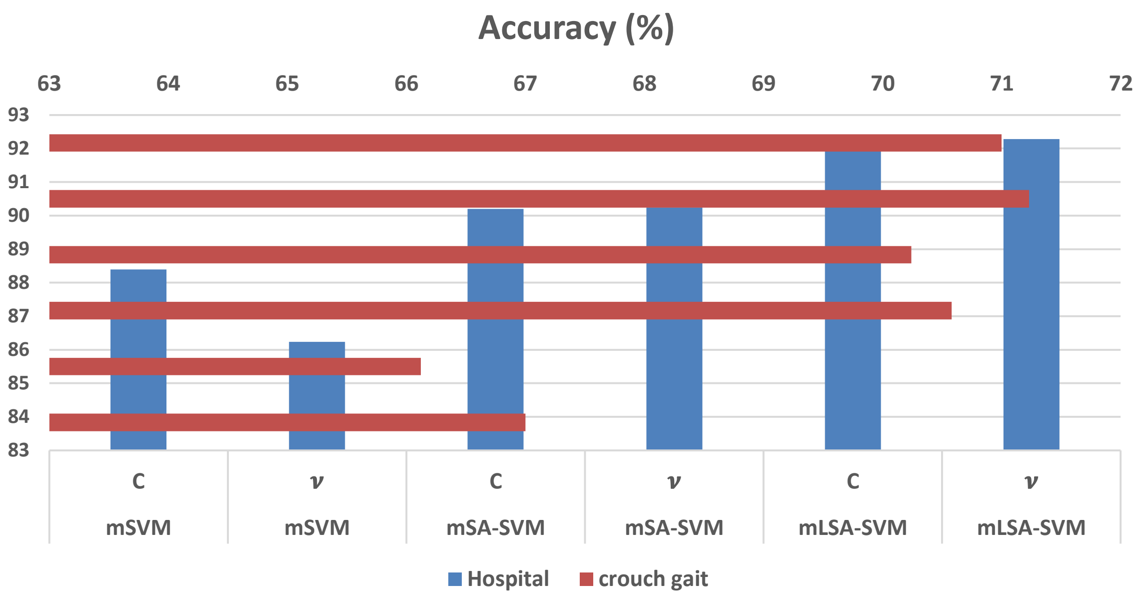


Figure 5. 2 The average classification accuracy of mLSA-SVM across two datasets using 5-fold cross validation compared with other classifiers

Table 5. 3 The training Time consumption for each type of classifier

Dataset	Training Time (ms)					
	*m*SVM C	*m*SVM *ν*	*m*SA-SVM C	*m*SA-SVM *ν*	*m*LSA-SVM C	*m*LSA-SVM *ν*
Hospital	221.8	276.9	223.6	256.1	**95.0**	121.1
crouch gait	400.4	503.8	**250.8**	259.2	258.0	251.0

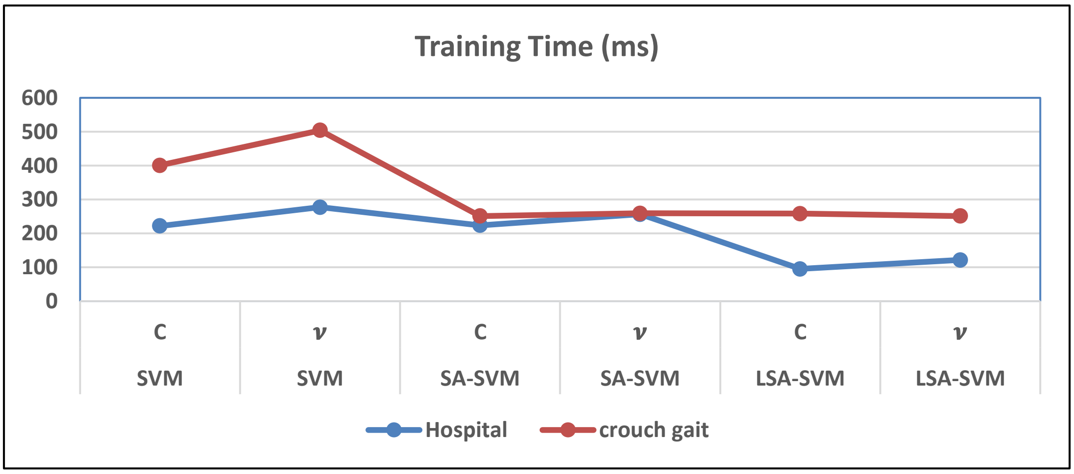


Figure 5. 3 The training Time consumption for each type of classifier

5.2.3.2 Experiments on UCI datasets

*m*LSA-SVM has proven that it has been adept classify four classes of leg movements consisting of combining three classes of the Unhealthy leg (Mild, Moderate and Severe patients) and fourth is Healthy patient class. We investigate further the performance of *m*LSA-SVM on the benchmark dataset that is accessible online on the UCI machine-learning website with multi-classes. The experiment depends on the size of the data that we implemented with 3-fold cross-validation on the Larger sized data while we executed 5-fold cross validation for the small and medium-sized data. Table 5.4 shows the data specification for a benchmark for Multi-Classes data sets types.

Table 5. 4 Data specification for benchmarking for Multi-Classes

Dataset	Group	Data	Attributes	classes
Zoo	**Small size**	101	17	7
Iris		150	4	3
Vehicle	**Medium size**	846	18	4
Vowel		990	10	11
weight lifting exercise	**Large size**	4024	34	5
Sensorless drive		58509	48	11

Table 5. 5 The optimal parameters used by each classifier in the UCI dataset experiments for multi-classes

Dataset	C-SVM		ν-SVM		Feat. No.	Win. Size, Win. Inc.
	C	γ	C	γ		
Zoo	3000	2	2^{14}	4	14	20,5
Iris	2050	0.005	2^{14}	0.5	1	20,5
Vehicle	5000	8	3000	8	13	100,25
Vowel	100	8	100	8	14	100,25
weight lifting exercise	8200	8	100	8	9	200,50

These experiments also involved six classifiers: v-*m*SVM, c-*m*SVM, c-*m*SA-SVM, v-*m*SA-SVM, c-*m*LSA-SVM, v-*m*LSA-SVM. First, we must discover the optimal parameters that affect the accuracy performance of the classifier. Table 5.5 provides all parameters utilized in this experiment.

Table 5.6 and Figure 5.4 shows that *m*LSA-SVM had reasonable performance over the six diverse data sets with multi-classes which gives accuracy equal to **99.26%** in **"Zoo"** dataset. While in **"Vehicle"** dataset the traditional cSVM gives accuracy higher than other equal to **82.16%**. The comparison of *m*LSA-SVM and *m*SA-SVM shows that both classifiers are little similar in the same data set.

Table 5. 6 The accuracy for six classifiers for 6 datasets using 5-fold cross validation for multi-classes

Dataset	Accuracy (%)					
	SVM C	**SVM ν**	**SA-SVM C**	**SA-SVM ν**	**LSA-SVM C**	**LSA-SVM ν**
Zoo	96.52	76.47	97.01	98.50	98.51	**99.26**
Iris	85.71	83.33	92.85	91.66	92.86	**92.86**
Vehicle	**82.16**	68.46	70.53	62.24	70.54	70.54
Vowel	97.43	81.41	96.15	98.71	98.71	**98.72**
weight lifting exercise	70.53	61.63	64.56	64.56	64.66	**70.67**
Sensor less drive	38	27	40.3	**55.21**	50	54.9

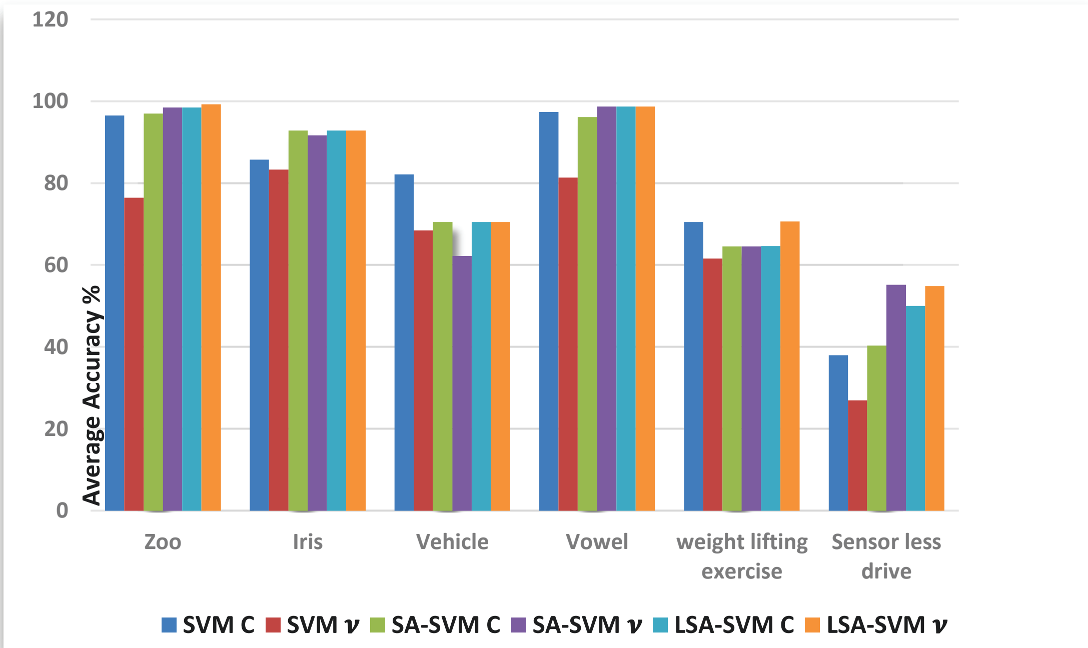


Figure 5. 4 The accuracy for six classifiers for 6 datasets using 5-fold cross validation for multi-classes

The processing time (Time consumption) of classifiers is noted for discussion, where Table 5.7 provides the training time for the four classifiers. Figure 5.5 clarifies that the training time of *m*LSA-SVM is faster than *m*SA-SVM classifier in most datasets, which gives minimum time in **"Iris"** dataset equal to **4.8msec**. *m*LSA-SVM changes to be worse or slower in "**Sensorless drive**" datasets equal to **250.8msec**, while *m*SA-SVM is faster classifier to learn "**Sensorless drive**" datasets equal to **190.5msec.**

Table 5. 7 The training time of four classifiers on various data using 5-fold cross validation for multi-classes.

Dataset	Training Time (ms)			
	SA-SVM C	SA-SVM ν	LSA-SVM C	LSA-SVM ν
Zoo	7.8	14.7	7.4	**6.5**
Iris	10.6	10.7	5.4	**4.8**
Vehicle	8.6	6.1	6.3	**5.2**
Vowel	**12.8**	14.3	14.2	13.4
weight lifting exercise	78.3	78.7	74.8	**74.3**
Sensor less drive	200.2	**190.5**	250.8	200

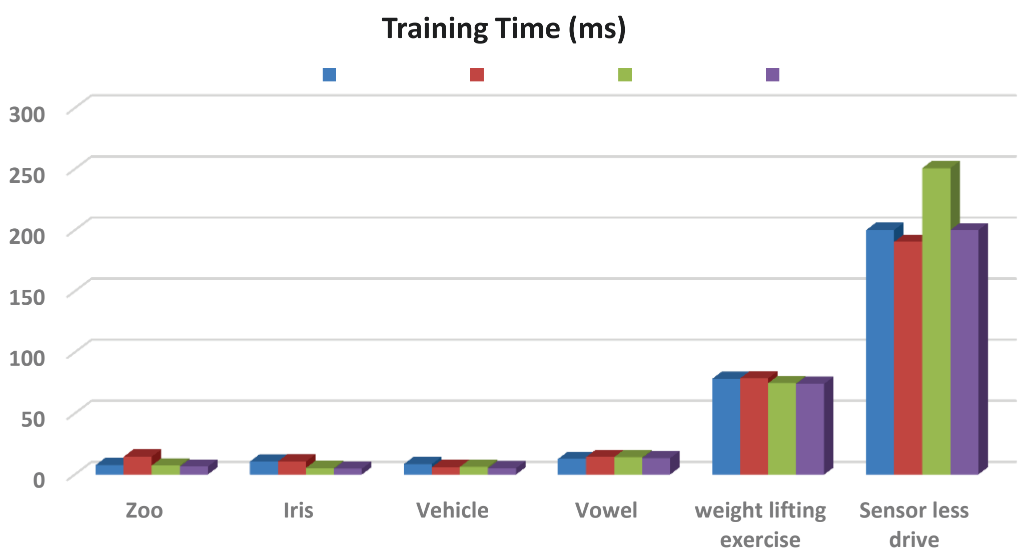


Figure 5. 5 The training time of four classifiers on various data using 5-fold cross validation for multi-classes.

5.3 Thesis Contribution-4 (Cont.4): Extreme Learning Machine-Label Self -Advised Support Vector Machine (ELM-LSA-SVM) for Multi-classification.

This section introduces a new multi-class classifier using Extreme Learning Machine and Support Vector Machine. This combination named Extreme Learning Machine - Label Self-Advised Support Vector Machine (ELM-LSA-SVM). ELM-LSA-SVM is using to classify multi-classes in pattern recognition.

A brief background of ELM-LSA-SVM provides then followed by the basic concepts of ELM-LSA-SVM. The experiments on Myoelectric pattern recognition for leg motion classification will be discussed afterward. Not limited to the Myoelectric signal, the experiments on UCI machine learning data sets can analyze the performance of ELM-LSA-SVM in large applications. A discussion of the application of ELM-LSA-SVM will conclude this section.

5.3.1 Background

ELM has two advantages; firstly, it extends the general single-hidden layer that feeds forward networks, the neuron in the hidden layer may not be identical. Secondly, the hidden node parameters are randomly generating without tuning. Huang et al. [50, 115] show that with the typical optimization method, ELM was able to extend linearly to SVM with minimum optimization limits, and complete the procedure of SVM at ease. Figure 5.6 describes the relationship between SVM and ELM methods. Essentially, SVM can be considered as a generalized form of Single-hidden Layer Feedforward Networks (SLFNs) with N_i as hidden nodes. While each of SVM and ELM used for SLFNs, the hidden layers for each are not tuned. The learning machines for both can be combined using two methodologies: the first method is the traditional ELM, which is diverse from SVM and can use with random kernels. The second method is similar to SVM but more diverse than the least square result of ELM; it can adapt as the traditional optimized method to get the result of ELM [116].

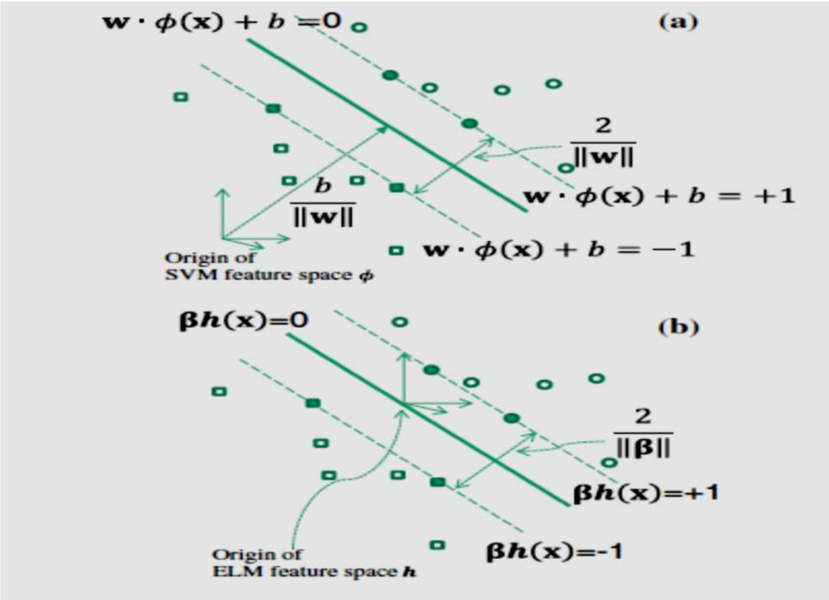


Figure 5. 6 Describe the relationship between SVM and ELM methods[117]

5.3.2 Concepts of ELM-LSA-SVM

In this thesis, we used the first methodology, ELM Kernel to transfer ELM to combine with LSA-SVM, which was partially published in paper conference [105] as we can consider in these steps:

1- The Kernel function of ELM is evaluated using Equation 5.6:

$$\mathrm{K}_{ELM}(\mathrm{x}_i, \mathrm{x}_j) = h(\mathrm{x}_i).h(\mathrm{x}_j)$$
$$= [\mathrm{G}(\mathrm{a}_1, \mathrm{b}_1, \mathrm{x}_i), \ldots, \mathrm{G}(\mathrm{a}_L, \mathrm{b}_L, \mathrm{x}_i)]^T.[\mathrm{G}(\mathrm{a}_1, \mathrm{b}_1, \mathrm{x}_j), \ldots, \mathrm{G}(\mathrm{a}_L, \mathrm{b}_L, \mathrm{x}_j)]^T \qquad (5.6)$$

Where $\mathbf{G}(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{x})$ is a nonlinear piecewise continuous function that satisfies the general estimate ability for ELM theorems, while $\{(a_i, b_i)\}_{i=1}^{L}$ arbitrarily created depending on continuous probability supply as shown in Equation 5.7 [118]:

$$L_{D_{ELM}} = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} y_i y_j \mathrm{K}_{ELM}(\mathbf{x}_i, \mathbf{x}_j)\alpha_i\alpha_j - \sum_{i=1}^{N}\alpha_i \qquad (5.7)$$

For the boundary: $0 \leq \alpha_i \leq C, \quad i = 1, \ldots \ldots, N$

So, the decision function for ELM will be calculated. Equation 5.8:

$$f(x) = sign\left(\sum_{s=1}^{N_S} y_s \alpha_s \mathrm{K}_{ELM}(\boldsymbol{x}, \boldsymbol{x_s})\right) \qquad (5.8)$$

As the $K(x, x_s)$ is the output function of the ith-hidden node of support vector network, $y_i\alpha_s$ is the output weight between the ith-hidden nodes and the output node.

On the other hand, if we need to compare: the SVM kernel function $k(x, x_i)$ needs to satisfy Mercer's state, thus for, the decision function of SVM is Equation 5.9[119]:

$$f(x) = sign\left(\sum_{s=1}^{N_s} y_s\alpha_s \mathrm{K}_{\mathrm{ELM}}(\boldsymbol{x}, \boldsymbol{x_s}) + b\right) \tag{5.9}$$

Ns is the number of support vectors x_s.

2- Misclassified data samples that are missing them in the primary training stage will be recognized in this step. The misclassified datasets (MD) in the training stage can be evaluated as Equation 5.10:

$$\mathrm{MD} = \bigcup_{i=1}^{N} x_s \ |y_s \neq \mathrm{sign}\left(\sum_{s=1}^{N_s} y_s\alpha_s \mathrm{K}_{\mathrm{ELM}}(\mathbf{x}, \mathbf{x_s})\right) \tag{5.10}$$

3- The algorithm indicates that: If MD is empty, then go to the testing phase, or else calculate neighborhood length (NL) for each y_i of label MD. Equation 5.11 defined NL.

$$\mathrm{NL}(y_i) = \mathrm{minimum}_{y_j}\left(\|y_i - y_j\| \mid x_i \neq x_j\right) \tag{5.11}$$

Where y_j, j=1, …., N is the label of training data that does not belong to the label of MD set. The label of training data mapped to a higher dimension, the distance between y_i and y_j is total according to Equation 5.12 concerning the related RBF kernel.

$$\|\emptyset(y_i) - \emptyset(y_j)\| = (k(y_i, y_i) + k(y_j, y_j) - 2k((y_i, y_j))^{0.5} \tag{5.12}$$

4- For each label y_k from test data, the multi-Labe Advised Weight mLAW(y_k) is figuring out as Equation 5.13. These m LAWs represent how much the label test data are closing to the label of misclassified data.

$$m\mathrm{LAW(yk)} = \\ = \begin{cases} 0 & \forall x_i \in \mathrm{MD}, \|y_k - y_i\| > \mathrm{NL}(y_i) \text{or } \mathrm{MD} = \mathrm{NUL}, \\ \sum 1 - \frac{\sum_{y_i}\|y_k - y_i\|}{\sum_{y_i}\mathrm{NL}(y_i)} & x_i \in \mathrm{MD}, \|y_k - y_i\| \leq \mathrm{NL}(y_i) \end{cases} \tag{5.13}$$

5- The absolute value of the SVM decision for each x_k from the test set is calculated and scaled to [1, 2,3,4,……m].

6- For each y_k from the label of the test set, If (mLAW (y_k) < decision value (y_k) then

$y_k = sign\left(\sum_{s=1}^{N_s} y_s \alpha_s \mathrm{K}_{ELM}(\boldsymbol{x}, \boldsymbol{x_s})\right)$

Otherwise compatible with standard ELM labeling $y_k = y_i \mid (\|y_k - y_i\| \leq$ NL (y_i) and $x_i \in$ MD

5.3.3 Experiments and results

The experiments are doing to check the performance of *ELM*-LSA-SVM for multi-classes as conducted through state of the art for the pattern recognition system shown in Figure 5.7.

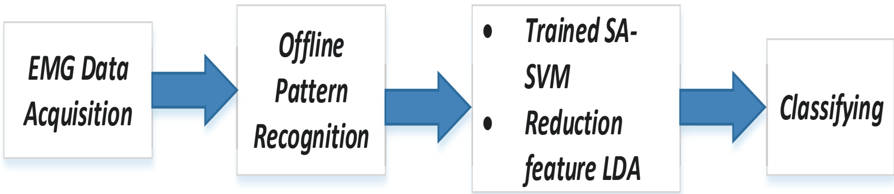


Figure 5. 7 pattern recognition system using ELM-SA-SVM

Section 5.2.3 established the stages of the procedure; it discussed in detail in **section 4.2.3** including its application in this experiment. The First experiment applied TriBas-ELM, Sine-ELM, HarddLim-ELM, RadBas-ELM, and Sig-ELM to select a classifier that will result in the best performance accuracy. The Second experiment, used the classifier identified in experiment one with ELM-SA-SVM and ELM-LSA-SVM with two different groups of data sets as explained briefly in **section 4.2.4**. **The first dataset** collected from healthy and unhealthy subjects at Metro Rehabilitation Hospital. The **Second data set**, provided by the Open Sim website: www.simtk.org.

5.3.3.1 Experiments ELM and other node based ELM

The experiment was prepared to evaluate the performance of ELM-LSA-SVM in Myoelectric pattern recognition. First we adjusted the parameters, where the parameter ranged for kernel-based ELMs for the polynomial kernel: C = (2^{-4}, 2^{-3}, … , 2^{4},2^{5}) and d = (2^{-5}, 2^{-4}, … , 2^{8},2^{9}), and the RBF kernel C = (2^{-9}, 2^{-8}, … , 2^{9},2^{10}) and γ = (2^{-19}, 2^{-18}, …

, 2^7,2^8). The ELM has a hidden layer of parameters which randomly created dependent on the regular distribution. However, we need to select C, and the number of hidden layers L. Fortunately, the combination of (L, C) does not affect the general performance as long as the number of hidden nodes L is large enough [50]. In this study, L gives a value of 2000. As for the range parameter of C, it was varied among (2^{-9}, 2^{-8},…, 2^{29},2^{30}), and it was selected using the coarse grid search method. Some analysis is given in each experiment as shown in Table 5.8. The **first data set FD** used to estimate the accuracy for each classifier.

Table 5. 8 Optimal parameters for all classifiers

Classifier	Hosp. C	CG C	Hosp. γ	CG γ	Feat. type	Hosp. Win. Size, Win. Inc.	CG Win. Size, Win. Inc.	Act.Fun.	L
TriBas-ELM	-	-	-	-	14	200,50	20,5	TriBas	200
ELM-SA-SVM	2^{10}	2^0	2^3	2^3	14	200,50	20,5	TriBas	200
ELM- LSA-SVM	2^{10}	2^0	2^3	2^3	14	200,50	20,5	TriBas	200

This experiment tested the five types of ELM with non-Kernel based ELMs: TriBas-ELM, Sine-ELM, HarddLim-ELM, RadBas-ELM and Sigmoid extreme learning machine (Sig-ELM). The number of hidden nodes changes from 50 to 550. Four-fold cross-validation has applied to confirm the validity of the classification outcome. During windowing, an overlapping window applied with a window length of 250ms along with a 35ms increment. Table 5.9 and Figure 5.8 are shows the classification outcomes. From these outcomes, the best type of node-ELM, which has the most accurate performance (TriBas-ELM) at the node number (200), can select.

Table 5. 9 The average classification accuracy of all node-ELM across Hospital dataset using 3-fold cross-validation.

# hidden Node	Accuracy (%)				
	TriBas-ELM	**Sine-ELM**	**HarddLim-ELM**	**RadBas-ELM**	**Sig-ELM**
50	98.75	98.98	93.90	98.98	98.83
100	98.67	98.90	97.18	98.98	98.83
200	**99.14**	98.90	98.51	98.75	98.83

300	<u>**99.06**</u>	98.90	98.28	98.83	98.98
400	98.98	98.90	98.90	98.75	98.98
500	98.98	98.90	98.83	98.83	98.98
550	98.98	98.90	98.98	98.90	<u>**99.06**</u>

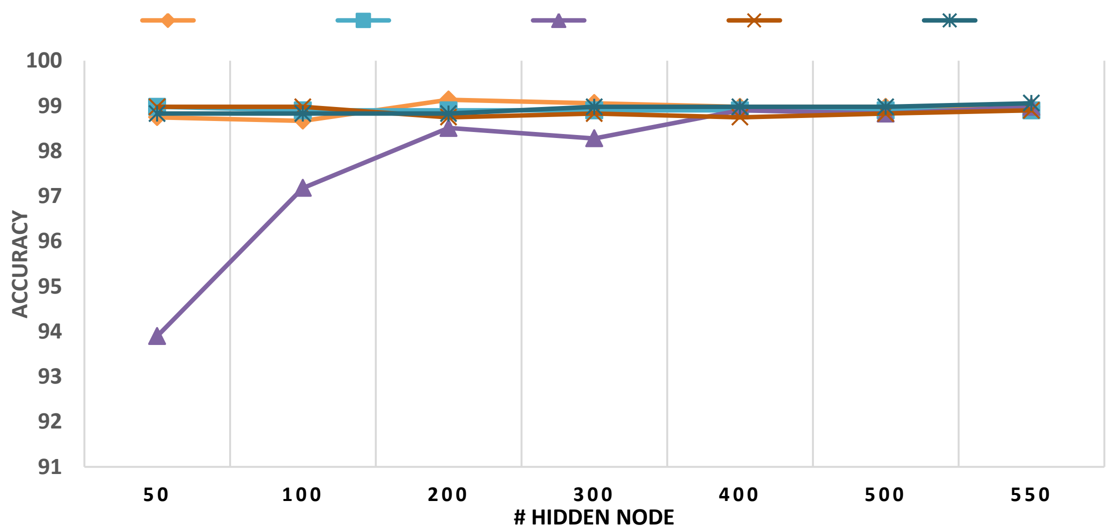


Figure 5. 8 The average classification accuracy of all node-ELM across Hospital dataset using 3-fold cross-validation.

5.3.3.2 Experiments ELM and other node based ELM

Several experiments were done to test the performance of ELM-LSA-SVM in Myoelectric pattern recognition. We examined the performance accuracy of the five classifiers: **TriBas-ELM, *v*-ELM-SA-SVM, *c*-ELM-SA-SVM, *v*-ELM-LSA-SVM and *c*-ELM-LSA-SVM**. An analysis of the experiment has given; the **Hospital data set** and the **CG data set** was used to estimate the accuracy for each classifier.

Table 5.10 and Figure 5.9 specifies that the average accuracy of ELM-LSA-SVM was higher than ELM-SA-SVM and other classifiers when using the CG dataset and it is equal 87.37%, while the accuracy for traditional TriBa-ELM has higher accuracy in FD Hospital dataset equal to 99.14%. While for training time uses Table 5.11 and Figure 5.10 shows a four classifier across Hospital dataset and CG dataset with 3-fold cross-validation.

Figure 5.10 describes that time training for ELM-LSA-SVM methodology is less than other methods 90.9msec and 101.98 msec for both hospital and CG.

Table 5. 10 The average classification accuracy for ELM-LSA-SVM across Hospital dataset and CG dataset at 3-fold cross-validation.

Dataset	Accuracy (%)				
	TriBas -ELM	ELM-SA-SVM C	ELM- SA-SVM ν	ELM-LSA-SVM C	ELM- LSA-SVM ν
Hospital	**99.14**	96.32	79.06	96.56	77.57
crouch gait	78.07	87.20	87.11	**87.37**	86.22

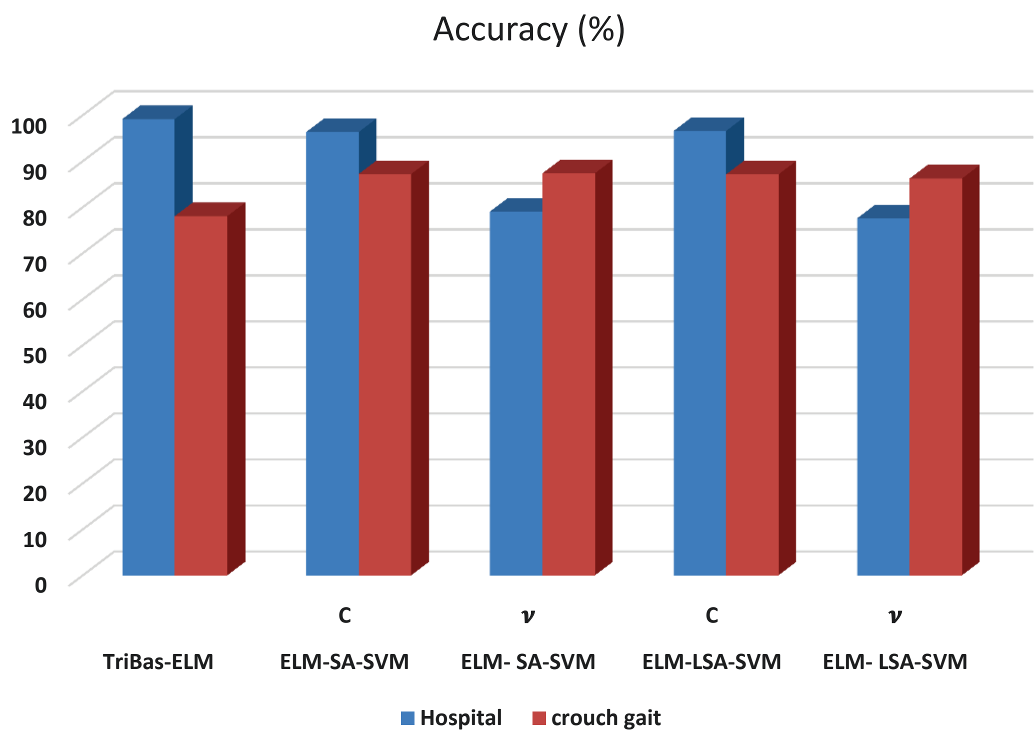
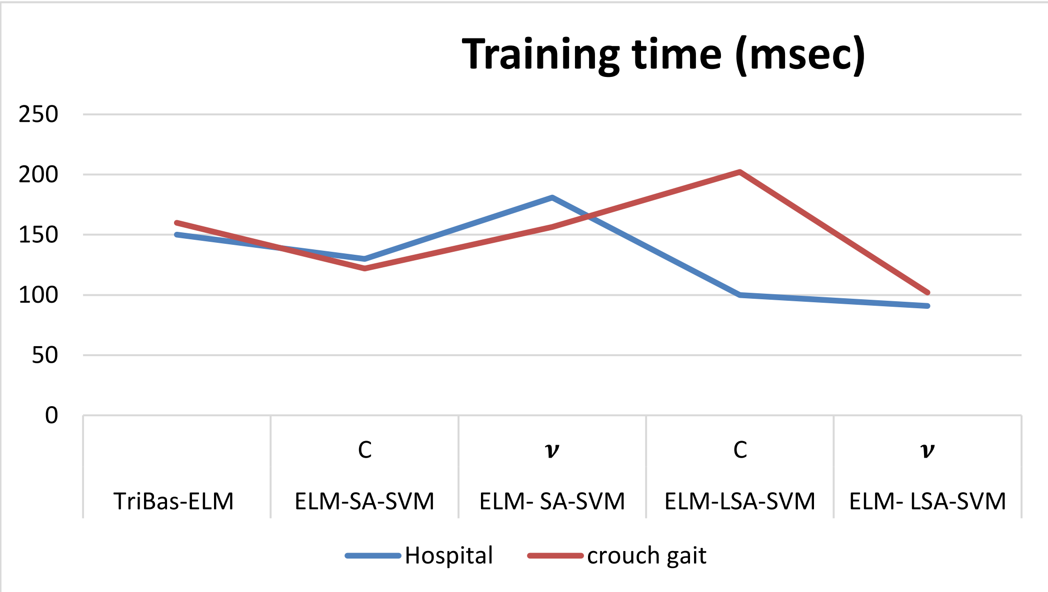


Figure 5. 9 The average classification accuracy across Hospital dataset and CG dataset at 3-fold cross-validation.

Table 5. 11 the training time uses a four classifier across Hospital dataset and CG dataset with 3-fold cross-validation.

	Training time (msec)				
	TriBas-ELM	**ELM-SA-SVM C**	**ELM- SA-SVM V**	**ELM-LSA-SVM C**	**ELM- LSA-SVM V**
Hospital	150.19	130.02	180.9	100	**90.9**
crouch gait	160	122	156.45	202.08	**101.98**

Figure 5. 10 The training time uses four classifiers across Hospital dataset and CG dataset with 3-fold cross-validation.



5.4 Summary

The third and fourth contributions addressed, and it is develop noted in this thesis chapter. LSA-SVM modified from a single class to Multiclass, and with this development, more applications will use this methodology. Then adopted and engaged The fourth methodology matched the base ELM with LSA-SVM to obtain a new method with a higher speed named ELM-LSA-SVM.

For the real data experiment, we applied the latter faster method in real-time application of Myoelectric pattern recognition (M-PR) for leg movement detection. The experimental results on the EMG dataset and benchmark datasets exhibit its benefits. Furthermore, the experimental results with the UCI data set indicate that ELM-LSA-SVM achieves the best performance when applied in CG dataset.

We explain the proposed and implemented of Multi-Label Self-Advised Support Vector Machine (*m*LSA-SVM) and projected Multi-Self-Advised Support Vector Machine (*m*SA-SVM) for leg motion recognition using sEMG signals. Generally, *m*LSA-SVM could classify four leg movements with an accuracy of **92.28 %.** It is comparable to the well-known classifiers such as *m*SA-SVM, *m*SVM. Therefore, *m*LSA-SVM could improve the performance of the advised-based SVM.

The implementation of *m*SA-SVM and *m*LSA-SVM for Myoelectric leg motion recognition is **one of the thesis contributions**. Another contribution that projected in this section is *m*LSA-SVM. The advantage of *m*LSA-SVM over *m*SA-SVM exists with working with Label data instead of data value itself. Therefore, this is **the third contribution of *m*LSA-SVM**. In addition to the Myoelectric leg motion classification, *m*LSA-SVM has applied in an extensive collection of classification problems using the UCI machine learning data set. The experimental results show that *m*LSA-SVM could apply to a wide range of data ranging from small to large-sized data. The *m*LSA-SVM is a promising classifier for several classification applications, particularly for Myoelectric pattern recognition.

To have a wide-ranging comparison training test was executed to compare results of mLSA-SVM and SA-SVM with all other methodologies. The *p*-value correlated to the ANOVA test by a level of impact $\alpha = 0.05$ was 0.04, which shows the big statistical difference between these groups, which shows that *m*LSA-SVM is statistically better than other methodologies. Therefore, it can be concluded that *m*LSA-SVM achieved better results than these algorithms. Besides, 90.03 % sensitivity and 94.50 % specificity achieved to classify the testing data for Hospital, while for Crouch Gate data, 68.03 % sensitivity and 74.2 % specificity achieved. A full information and equations of these methods ANOVA test, Sensitivity and Specificity are present in Appendix C.

The p-value correlated to the ANOVA test by a level of impact $\alpha = 0.05$ was 0.009, which shows the absolute statistical difference between these groups, which shows that **ELM-LSA-SVM** is statistically better than other methodologies.

Therefore, it can conclude that **ELM-LSA-SVM** achieved better results than these algorithms. Also, 96.25 % sensitivity and 99.25 % specificity achieved to classify the testing data for Hospital, while for Crouch Gate data, 84.04 % sensitivity and 90.05 % specificity achieved. A full information and equations of these methods ANOVA test, Sensitivity and Specificity are present in Appendix C.

The tables clarified the short length of data helps classify the severity of foot drop; this appears clear when learning machine training the data with minimum time training and perfect accuracy more than the long data.

Chapter 6

Summary, Conclusion and Future research

This chapter provides a summary of this thesis and major finding of the research conducted in this thesis. Afterward, the conclusion and future works following on from the research in the same field described.

6.1. Thesis summary

The main purpose of this research is to introduce novel classifiers to develop the performance of the M-PR for Foot Drop Rehabilitation devices, which race the problems that occur in the real-time application. This research proposes various methodologies and then examines them with the real-time application. This thesis presented and discusses all contributions, and explains the acquirement briefly for collecting sEMG data signal from Foot Drop Patients in Rehabilitation Hospital in Sydney, Australia using Ethical Approval (UTS HREC NO.ETH15-0152).

Chapter 2 this chapter discussed the Foot Drop problem and its related subjects such as Lower Limb Anatomy, diseases, and complications of the lower limb. Detection and of FD problem using sEMG characteristics electromyography reviewed. Microcontroller system and M-PR stages for leg movement recognition are studied. Furthermore, some conventional approaches and methods use the machine learning system, and their previous applications in Foot Drop are studied and reviewed. Also, functional electrical stimulation (FES) and most common rehabilitation devices used for FD patients studied. At the end of the chapter, dynamic simulation software for different movement of the musculoskeletal system discussed. This simulation named Ope-Sim software, which is simulated and evaluate FD anatomy.

Chapter 3 addresses some Learning Machine methodologies that used in this research like SVM, SA-SVM,ElM, and Label classification methods. For these methods, a review of the Support Vector Machine, Self-Advise SVM, Extreme Learning Machine and Label Classification methods are studies to develop auto-control and a diagnostic system for foot drop.

Chapter 4 describes the data collected from the hospital based on the design procedure that got the ethical approval to collect data from FD patients in Metro Rehabilitation Hospital in Sydney, Australia using Ethical Approval (UTS HREC NO.ETH15-0152). Compliments the novel idea behind Myoelectric Pattern Recognition

In addition to the above **Chapter 4,** proposed a new based SVM classifier method named Label Self-Advised Support Vector Machine (LSA-SVM) and upgraded a Self-Advised Support Vector Machine (SA-SVM) to multi-classes to classify leg motion recognition. Overall, LSA-SVM methods could be classified leg movements if it is healthy or not with an accuracy of 99.06 % comparable to known classifiers such as SA-SVM, SVM. Therefore, LSA-SVM could improve the performance of the advised-based SVM.

The implementation of *m*SA-SVM and LSA-SVM for Myoelectric leg motion recognition is a **contribution-1 of the thesis**. Other contributions that projected in the related section is *L*SA-SVM. The advantage of LSA-SVM over SA-SVM exists on working with Label data instead of data value itself. Therefore, this is **contribution-2** Further to this Myoelectric leg motion classification, LSA-SVM has applied in a vast collection of classification problems using UCI machine learning data set. The experimental results show that LSA-SVM can apply to a wide range of data set from small to large-sized data. Overall, LSA-SVM is a promising classifier for several classification applications, particularly for Myoelectric pattern recognition.

Chapter 5 addresses the third and fourth contributions that had developed in this thesis. Firstly, LSA-SVM has modified from a single class to Multiclass; this development allowed further application in multi-label classification. Secondly, match the base ELM with LSA-SVM to obtain a novel method that was speedier, named ELM-LSA-SVM.

The real-time data experiments apply the faster method in the real-time application of Myoelectric pattern recognition (M-PR) for leg movement detection. The experimental results on the sEMG signal data and benchmark datasets exhibit its benefits as we can see in table 6.1 and 6.2 all kinds of results for all novel methods. Furthermore, the experimental results on UCI datasets indicate that ELM-LSA-SVM achieves the best performance when combined with LSA-SVM. The entire experimental results recorded and display in chapter 5.

Table 6.1 summarizing all results for all contributed methodologies and describing the average classification, while, Table 6.2 clarify the training Time consumption.

Table 6. 1 The average classification accuracy for all contributed methodologies

Accuracy (%)				
	*m*SA-SVM	LSA-SVM	*m*LSA-SVM	ELM-LSA-SVM
FD-Hospital	**90.24**	**99.06**	**92.28**	**99.14**
crouch gait CG	**70.58**	**82.01**	**71.23**	**87.37**
UCI datasets		**96.42 (Statelog)**	**99.26 Zoo**	

Table 6. 2 The training Time consumption for all contributed methodologies

Training Time (ms)				
	*m*SA-SVM	LSA-SVM	*m*LSA-SVM	ELM-LSA-SVM
FD-Hospital	**221.8**	**46.3**	**95.0**	**90.9**
crouch gait CG	**250.8**	**38.0**	**250.8**	**101.98**
UCI datasets		**0.537 Parkinson**	**4.8 Iris**	

The tables of results above describe that the Hospital datasets and UCL datasets give a perfect accuracy when compared to Crouch Gait datasets. While the time consumption for all methodologies identify that short datasets like UCI will process faster than another datasets (Hospital and CG)

6.2 Conclusion

In this thesis, four contributions were proposed adding some new methodologies, novel myoelectric pattern recognition system, and a novel myoelectric controller for exoskeleton leg.

The present chapter briefly summarizes the overall thesis contributions and provides future development application and extension of the methods presented in this thesis.

All experiments conducted in this thesis were presented in various high-quality conferences and has published online. Therefore, the contributions verified globally.

To have a wide-ranging comparison training test was executed to compare results of mLSA-SVM and SA-SVM with all other methodologies. The p-value correlated to the ANOVA test by a level of impact $\alpha = 0.05$ was 0.04, which shows the big statistical difference between these groups, which shows that mLSA-SVM is statistically better than other methodologies. Therefore, it can be concluded that mLSA-SVM achieved better results than these algorithms.

The tables which describe time training in all chapters: clarified that short length of data helps to classify the severity of foot drop; this appears clear when learning machine training the data with minimum time training and perfect accuracy more than the extended data."

We have some limitations (restrictions) in this work:

1- We could not find or get the 25 patients of food Drop (that we decided from the beginning in the same time that we need them), so that we work with only ten patients adding to them three healthy subjects. This restricted the practical work for this thesis by giving less results and comparisons with less dataset (patients).

2- To apply and to transfer the novel methodologies in this thesis, we should buy Bioness L300 foot Drop device to study the electrical circuits then modify our new methods practically, then doing some practical tests instead of simulation part in Open Sim software. Unfortunately, There is no funding to buy Bioness L300 foot Drop device which allows as to build an Auto-control and diagnosis systems. This device cost around 7000$.

3- Another limitation, we could not make an online connection between Open-Sim and Mat-lab software. This restricted the practical test time when I transfer the results offline from mat lab results to simulation part (open sim)

6.3 Recommendation for future research

The following subjects are recommended for future research.

- To implementing the LSA-SVM in FD Rehabilitation devices to give more accuracy and less time for classifying the Foot Drop patients.

- The real-data experiment of Myoelectric pattern recognition (M-PR) involved four sEMG channels and succeeded with an accuracy of 99.06 %. Reducing sEMG channels may be added to develop the performance in the real-data application.
- The consumption time of the Myoelectric controller is approximately 251msec. That time is less than the standard consumption time (300ms). In the future work, changing in Feature extraction and reduction methods will reducing more the time consuming to improve the comfort of the subject for soft human-machine interaction.
- The exoskeleton leg simulation (Open-Sim) that is applied in this study does not have a real-time connection with the Mat-lab program (linked together). For future research, any leg simulation used in the experiments should be linked with Mat-lab and coded to improve the performance and decrease the time delay.
- The implementation of the proposed Myoelectric auto-controller on the leg should be investigated both in real time.
- The feature learning on sEMG signals should be examined to improve the performance of the Myoelectric pattern recognition system.
- Continue with research and try to get faster performance in ELM-LSVM when using FD Hospital dataset.

APPENDIX A: Ethical Approval

That section will provide ethical approval of information regarding the research conducted in this thesis.

HREC Approval Granted - ETH15-0152

Research.Ethics@uts.edu.au

Thu 6/2/2016 11:49 AM

To: Adel Ali Al-Jumaily <Adel.Al-Jumaily@uts.edu.au>; Sahar Adil Abboud <SaharAdil.Abboud-1@student.uts.edu.au>; Research Ethics <research.ethics@uts.edu.au>;

Dear Applicant

Thank you for your response to the Committee's comments for your project titled, "Design Real Time Embedded System for Automated Classification of EMG Signals for Rehabilitation and Devices". Your response satisfactorily addresses the concerns and questions raised by the Committee who agreed that the application now meets the requirements of the NHMRC National Statement on Ethical Conduct in Human Research (2007). I am pleased to inform you that ethics approval is now granted.

Your approval number is UTS HREC REF NO. ETH15-0152.

Approval will be for a period of five (5) years from the date of this correspondence subject to the provision of annual reports.

Please note that the ethical conduct of research is an on-going process. The National Statement on Ethical Conduct in Research Involving Humans requires us to obtain a report about the progress of the research, and in particular about any changes to the research which may have ethical implications. This report form must be completed at least annually, and at the end of the project (if it takes more than a year). The Ethics Secretariat will contact you when it is time to complete your first report.

I also refer you to the AVCC guidelines relating to the storage of data, which require that data be kept for a minimum of 5 years after publication of research. However, in NSW, longer retention requirements are required for research on human subjects with potential long-term effects, research with long-term environmental effects, or research considered of national or international significance, importance, or controversy. If the data from this research project falls into one of these categories, contact University Records for advice on long-term retention.

You should consider this your official letter of approval. If you require a hardcopy please contact Research.Ethics@uts.edu.au.

To access this application, please follow the URLs below:
* if accessing within the UTS network: https://rm.uts.edu.au
* if accessing outside of UTS network: https://remote.uts.edu.au , and click on " RM6 – Research Master Enterprise " after logging in.

We value your feedback on the online ethics process. If you would like to provide feedback please go to: http://surveys.uts.edu.au/surveys/onlineethics/index.cfm

If you have any queries about your ethics approval, or require any amendments to your research in the future, please do not hesitate to contact Research.Ethics@uts.edu.au.

Yours sincerely,

Professor Marion Haas
Chairperson
UTS Human Research Ethics Committee
C/- Research & Innovation Office

University of Technology, Sydney
E: Research.Ethics@uts.edu.au

UTS
UNIVERSITY OF TECHNOLOGY, SYDNEY

CONSENT FORM

I ____________________ ***(participant's name)*** agree to participate in the research project "Automated Classification of EMG Signals for Rehabilitation Application and Devices conducted by **Sahar Adil*,*** *saharadil.abboud-1@student.uts.edu.au* , with student ID: and contact number of the University of Technology, Sydney for her degree PHD.

I understand that the purpose of the study is to investigate the foot drop patient after stroke. This research is based on analysis Electromyography (EMG) muscle signal. EMG will be used to provide indication for foot drop problem. The collected of EMG data is used to classification the types or levels of foot drop problem

I understand that I will be invited to volunteer approximately 2 hours of my time in a single day for data collection. The data collection will involve the placement of an electrode on my lower limb skin to measure an **EMG** (Electromyography) signal. Data will be collected for five trials of each posture (flexion or extension) of knee joint. I will be asked to:

1- **First trial:** moving your lower limb towards the knee joint to flexion and extension from rest position while you sit on chair. Each trial will take 3 seconds and there will be a rest of 5 seconds between two trails. We will take 5 trials for this case.
2- **Second trial:** moving your foot up and down as much as you can from rest position while you sit on chair. Each trial will take 3 seconds and there will be a rest of 5 seconds between two trails. We will take 5 trials for this case.
3- **Third trial:** moving your lower limb (foot and leg) towards knee joint to flexion or extension from rest position while you stand up. Each trial will take 3 seconds and there will be a rest of 5 seconds between two trails. We will take 5 trials for this case.

I understand that I may experience some degree of having some exhaustion (fatigue) in my leg while or after the trails.

I am realize that I can contact Sahar Adil or her supervisor(s) Adel Al Jumaily (Adel.Al-Jumaily@uts.edu.au or 9514 7939) if I have any concerns about the research. I also understand that I am free to withdraw my participation from this research project at any time I wish, without consequences, and without giving a reason.

I agree that Sahar Adil has answered all my questions fully and clearly.

I agree that my medical record to be accessed for this study.

I agree that the research data gathered from this project may be published in a form that does not identify me in any way.

____ ____/____/____
Signature (participant)

____ ____/____/____
Signature (researcher or delegate)

NOTE: This study has been approved by the University of Technology, Sydney Human Research Ethics Committee. If you have any complaints or reservations about any aspect of your participation in this research which you cannot resolve with the researcher, you may contact the Ethics Committee through the Research Ethics Officer (ph: +61 2 9514 9772 Research.Ethics@uts.edu.au) and quote the UTS HREC reference number. Any complaint you make will be treated in confidence and investigated fully and you will be informed of the outcome.

Ethics

INFORMATION SHEET FOR EMG BASED FOOT DROP ANALYSIS AND PREDICTION

WHO IS DOING THE RESEARCH?

My name is Sahar Adil and I am a student Doctorate of Philosophy at UTS. My supervisor is Prof. Adel Al-Jumaily.

WHAT IS THIS RESEARCH ABOUT?

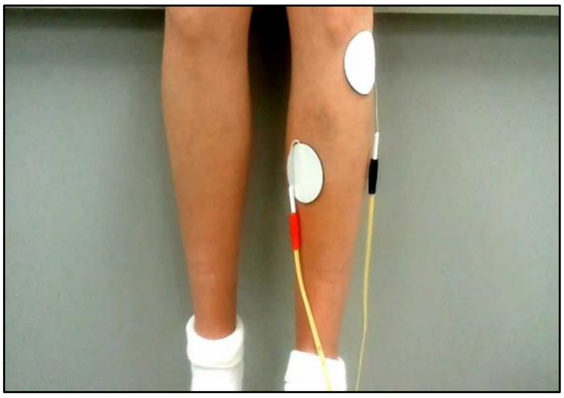
The purpose of the study is to investigate the foot drop in stroke patients. The research aim to find out if it possible to discover the foot drop before it happen. The predication of it before it physically happens will help to deal with it (correct it) and provide the required action that will stop it from occurring. This research is based on analysis Electromyography (EMG) muscle signal. EMG will be used to provide indication for foot drop problem. The collected of EMG data is used to predict foot drop and classify the types or levels of it, thus will help to recommend the correct action to deal with it. The recruitment of research subjects, recording of the signals, observation and monitoring, and expert advice will be take place at Metro Rehab hospital

IF I SAY YES, WHAT WILL IT INVOLVE?

I will invite you to volunteer approximately up to 2 hours of your time in a single day for collector of muscle signal. This data collection will be under the supervision of Anna Barlow (stroke coordinator) will take place in Metro Rehab hospital.

The data collection will involve the placement of an electrode on your lower limb skin to measure an EMG muscle signal, attaching the EMG sensors (electrodes) to your body will not make any extraordinary feeling but just like placing piece of plastic on the skin. Data will be collected for five trials of each posture (flexion or extension). You will be asked to:

1- First trial: moving (or get help to move if required) your lower limb at the knee joint to flexion and extension (bend and straighten) from rest position while you sit on chair. Each trial will take 3 seconds and there will be a rest of 5 seconds between two trials. We will take 5 trials for this case.
2- Second trial: moving (or get help to move if required) your foot up and down as much as you can from rest position while you sit on the chair. Each trial will take 3 seconds and there will be a rest of 5 seconds between two trials. We will take 5 trials for this case.
3- Third trial: moving (or get help to move if required) your lower limb (foot and leg) at the knee joint to flexion or extension (bend and straighten) from rest position while you standing. Each trial will take 3 seconds and there will be a rest of 5 seconds between two trials. We will take 5 trials for this case.



ARE THERE ANY RISKS/INCONVENIENCE?

It may be required to shave a small part of your leg (if needed) where the electrode placed as shown in the figure (in addition to one above the knee). If the shaving required, it will be done by a disposable razor.
You might be having some exhaustion (fatigue in the leg) that may be generated from the trial.

WHY HAVE I BEEN ASKED?

You have a foot drop problem due to a stroke which is what this research about. The outcome of this study is to build a system that can help in dealing with the foot drop by predication before it the physically happens and take action that will correct it. In addition to understanding and correlating the foot drop problem to the muscle signals (EMG) to find a better treatment for it.

DO I HAVE TO SAY YES?

You don't have to say yes. There will be no payment for participation but your participation will contribute to stroke research.

WHAT WILL HAPPEN IF I SAY NO?

Nothing. I will thank you for your time so far and won't contact you about this research again.

IF I SAY YES, CAN I CHANGE MY MIND LATER?

You can change your mind at any time and you don't have to say why. I will thank you for your time so far and won't contact you about this research again.

WHAT IF I HAVE CONCERNS OR A COMPLAINT?

If you have concerns about the research that you can contact my supervisor, Dr Adel Ali Al-Jumaily (Phone 02 95147939, e-mail: Adel.AL-JUMAILY@uts.edu.au) who can help you or please feel free to contact me on .

If you would like to talk to someone who is not connected with the research, you may contact the Research Ethics Officer via Research.Ethics@uts.edu.au, and quote this number *(give UTS HREC Approval Number)*

ACCESS TO MEDICAL RECORD

It may be required to access your medical record to get some extra information regarding your case. This information may be including your diagnosing history, age, weight, height.
None of your information will be accessed by anybody out of the people mentioned above (me, my supervisor, and Metro Rehab hospital staff).
If any of your information used in publications, it will be coded in way that you cannot be identified.

APPENDIX B SOFTWARE CODES

B1. Multi Self- Advised Support Vector Machine (m SA-SVM)

B1.1 *m*SA-SVM code for the training

```
function
[result_tr,NewResult,result,m,otime,ntime,TrainAccu,ER1,ER2]=new_SA_SVMmulC(traindata,trainlabel,testdata,testlabel,SVMtype,C,D,gamma,r,kernel)
% This code is for Multi Self- Advised Support Vector Machine multi-Self Advising SVM
 % Outputs
   % TP1,TN1,FP1,FN1: for original SVM
   % TP2,TN2,FP2,FN2: for self-advising SVM
   %Smdata: numbers of miss data in training
   %Sneighb: number of neighbours in test data
   %training: accuracy of training
   %otime: original time for SVM
   %ntime: new time for self advising SVM
%  inputs
   %SVMtype,C,D,gamma,kernel: parameters of libsvm
best_C=C;best_d=D;best_gamma=gamma;best_kernel=kernel;
param = ['-s ', num2str(SVMtype), ' -t ',num2str(kernel), ' -d ',num2str(D), ' -g ', ...
   num2str(gamma), ' -c ',num2str(C) ' -r ',num2str(r)];
if nargin < 5
   K = 1;
end
ty=0;
if(ty==0)
   start_time=cputime;
   m=svmtrain2(trainlabel,traindata,'-s 0 -t 2 -g 12 -c 8');   %training
   end_time=cputime;
   [TrainPredict, TrainAccu,TrainProb]=svmpredict(trainlabel,traindata,m);  %test
   [TestPredict, TestAccu,TestProb]=svmpredict(testlabel,testdata,m);  %test
else
   cmd='-b 1 -s 0 -t 2 -g 12 -c 8';
   start_time=cputime;
```

```
        m = ovrtrain(trainlabel,traindata,cmd);   %training
        end_time=cputime;
        [TrainPredict, TrainAccu,TrainProb]=ovrpredict(trainlabel,traindata,m); %test
        [TestPredict, TestAccu,TestProb]=ovrpredict(testlabel,testdata,m);  %test
    end
    otime=toc;
    TrainErr = 100 - TrainAccu(1,1)
    TestErr = 100 - TestAccu(1,1)
    Timetraining = end_time-start_time;
    Timetesting=Timetraining;
    result_tr=TrainPredict;
    result=TestPredict;
    PER_tr=TrainProb;
    PER=TestProb;
    NewResult=result;
%% %This function find data that has wrong label from training of SVM
    [mdata,mislabel,mper,t,mIndex,qdata,qlabel]=missdata(traindata,trainlabel,m);
    if ~isempty(qdata)
        training=t(1);  %accuracy of training
        Smdata=0;
        if ~isempty(mdata)
            Smdata=size(mdata,1);
        end;
         o=closetest(qdata,qlabel,testdata,mdata,mislabel,best_C,best_d,best_gamma,best_kernel,mIndex);
        Sneighb=0;
        pos=0;
        neg=0;
        if ~isempty(o)
            [b,b1]=unique(o(:,1));
            Sneighb=length(b);
            for i=1:Sneighb
                bb=find(o(:,1)==b(i));
                S1=0;
                S2=0;
                S3=0;
                S4=0;
                S=0;
                for j=1:length(bb)
                    if o(bb(j),3)==1
```

```
                S1=S1+(abs(mper(o(bb(j),2))))/(1+o(bb(j),4)^2);
            elseif o(bb(j),3)==2
                S2=S2+(abs(mper(o(bb(j),2))))/(1+o(bb(j),4)^2);
            elseif o(bb(j),3)==3
                S3=S3+(abs(mper(o(bb(j),2))))/(1+o(bb(j),4)^2);
            elseif o(bb(j),3)==4
                S4=S4+(abs(mper(o(bb(j),2))))/(1+o(bb(j),4)^2);
            end;
            end;
        if (S1>S2 && S1>S3)&& (S1>S4);
            S=S1;
        elseif (S2>S1&& S2>S3) && (S2>S4);
            S=S2;
        elseif (S3>S1&& S3>S2) && (S3>S4);
            S=S3;
        elseif (S4>S1&& S4>S2)&& (S4>S3);
            S=S4;
        end;
        if S>=(abs(PER(o(bb(j),1))))
            NewResult(o(bb(j),1))=o(bb(j),3);
        end;
    end;
  end;
  ntime=toc;
  [TP1, TN1, FP1, FN1,TP2, TN2, FP2, FN2,ER1,ER2] =measures1(result,NewResult,testlabel);
else
  TP1=0;
  TN1=0;
  FP1=0;
  FN1=0;
  TP2=0;
  TN2=0;
  FP2=0;
  Smdata=0;
  Sneighb=0;
  training=0;
  otime=0;
  ntime=0;
  FN2=0;
```

```
end
end
```

B1.2 closing test data code

```
function o=closetest(qdata,qlabel,testdata,mdata,mislabel,best_C,best_d,best_gamma,best_kernel,~)
o=[];
% traindata1=seprate(traindata,ytrain,1);
% traindata2=seprate(traindata,ytrain,2);
% mdata1=seprate(mdata,mislabel,1);
% mdata2=seprate(mdata,mislabel,2);

%%This function return the test data close to the miss data
for i=1:size(mdata,1)
region(i)=InsurenceRegion(mdata(i,:),mislabel(i),qdata,qlabel,best_C,best_d,best_gamma,best_kernel);
end

%% %This function compute Insurence Region
k=1;
for i=1:size(testdata,1)
   for j=1:size(mdata,1)
dis=(kernelmap(mdata(j,:),mdata(j,:),best_C,best_d,best_gamma,best_kernel)+kernelmap(testdata(i,:),
testdata(i,:).....
 ,best_C,best_d,best_gamma,best_kernel)-
2*kernelmap(mdata(j,:),testdata(i,:),best_C,best_d,best_gamma,best_kernel))^0.5;


      if dis<=region(j);
         o(k,1)=i;
         o(k,2)=j;
         o(k,3)=mislabel(j);
         o(k,4)=dis;
         o(k,5)=region(j);
         k=k+1;
      end;
   end;
end;
```

B1.3 InsurenceRegion

```
function out=InsurenceRegion(x,label,data,ydata,best_C,best_d,best_gamma,best_kernel)
dis=[];
for i=1:size(data,1)
```

```
    if ydata(i)~=label

dis(i)=(kernelmap(x,x,best_C,best_d,best_gamma,best_kernel)+kernelmap(data(i,:),data(i,:),best_C,best_d,best_gamma,best_kernel)-...
  2*kernelmap(x,data(i,:),best_C,best_d,best_gamma,best_kernel)).^0.5;
    end;
end;
out=(min(dis(dis>0)))/2;
```

B2. Label Self -Advised Support Vector Machine (LSA-SVM)

```
function[result_tr,NewResult,result,m,otime,ntime,TrainAccu,ER1,ER2]=LSA_SVM2C(
traindata,trainlabel,testdata,testlabel,SVMtype,C,D,gamma,r,kernel)
% Label Self Self Advising SVM
% Outputs
  % TP1,TN1,FP1,FN1: for original SVM
  % TP2,TN2,FP2,FN2: for self-advising SVM
  %Smdata: numbers of miss data in training
  %Sneighb: number of neighbours in test data
  %training: accuracy of training
  %otime: original time for SVM
  %ntime: new time for self advising SVM
%inputs
  %SVMtype,C,D,gamma,kernel: parameters of libsvm
%% use svmclassfiyku, coz it deal with over train data also
best_C=C;best_d=D;best_gamma=gamma;best_kernel=kernel;
param = ['-s ', num2str(SVMtype), ' -t ',num2str(kernel), ' -d ',num2str(D), ' -g ', ...
   num2str(gamma), ' -c ',num2str(C) ' -r ',num2str(r)];
if nargin < 5
   K = 1;
end
ty=0;
if(ty==0)
   start_time=cputime;
   m=svmtrain2(trainlabel,traindata,'-s 0 -t 2 -g 12 -c 8');   %training
   end_time=cputime;
   [TrainPredict, TrainAccu,TrainProb]=svmpredict(trainlabel,traindata,m);  %test
   [TestPredict, TestAccu,TestProb]=svmpredict(testlabel,testdata,m);  %test
else
```

```
    cmd='-b 1 -s 0 -t 2 -g 12 -c 8';
    start_time=cputime;
    m = ovrtrain(trainlabel,traindata,cmd);   %training
    end_time=cputime;
    [TrainPredict, TrainAccu,TrainProb]=ovrpredict(trainlabel,traindata,m);  %test
    [TestPredict, TestAccu,TestProb]=ovrpredict(testlabel,testdata,m);  %test
end
otime=toc;
TrainErr = 100 - TrainAccu(1,1)
TestErr = 100 - TestAccu(1,1)
Timetraining = end_time-start_time;
Timetesting=Timetraining;
result_tr=TrainPredict;
result=TestPredict;
PER_tr=TrainProb;
PER=TestProb;
NewResult=result;
%%
[mdata,mislabel,mper,t,mIndex,qdata,qlabel]=missdata(traindata,trainlabel,m);
if ~isempty(qdata)
    training=t(1);  %accuracy of training
    Smdata=0;
    if ~isempty(mdata)
        Smdata=size(mdata,1);
    end;

o=closetest3(qdata,qlabel,testdata,testlabel,mdata,mislabel,best_C,best_d,best_gamma,best_kernel,mI
ndex);
    Sneighb=0;
    pos=0;
    neg=0;
    if ~isempty(o)
        [b,b1]=unique(o(:,1));
        Sneighb=length(b);
        for i=1:Sneighb
            bb=find(o(:,1)==b(i));
            S1=0;
            S2=0;
            S=0;
          for j=1:length(bb)
```

```
            if o(bb(j),3)==1
              S1=S1+(abs(mper(o(bb(j),2))))/(1+o(bb(j),4)^2);
            elseif o(bb(j),3)==2
              S2=S2+(abs(mper(o(bb(j),2))))/(1+o(bb(j),4)^2);
            end;
          end;
          if S1>S2
            S=S1;
          else
            S=S2;
          end;
          if S>=(abs(PER(o(bb(j),1))))
            NewResult(o(bb(j),1))=o(bb(j),3);
          end;
        end;
      end;
      ntime=toc;
      [TP1, TN1, FP1, FN1,TP2, TN2, FP2, FN2,ER1,ER2] =measures1(result,NewResult,testlabel);
    else
      TP1=0;
      TN1=0;
      FP1=0;
      FN1=0;
      TP2=0;
      TN2=0;
      FP2=0;
      Smdata=0;
      Sneighb=0;
      training=0;
      otime=0;
      ntime=0;
      FN2=0;
    end
    end
```

B2.2 closing test data code

```
function
o=closetest3(qdata,qlabel,testdata,testlabel,mdata,mislabel,best_C,best_d,best_gamma,best_kernel,~)
o=[];
% traindata1=seprate(traindata,ytrain,1);
% traindata2=seprate(traindata,ytrain,2);
% mdata1=seprate(mdata,mislabel,1);
% mdata2=seprate(mdata,mislabel,2);
for i=1:size(mdata,1)

region_L(i)=InsurenceRegion_L(mdata(i,:),mislabel(i),qdata,qlabel,best_C,best_d,best_gamma,best_k
ernel);
end
k=1;
for i=1:size(testdata,1)
   for j=1:size(mdata,1)

dis_L=(kernelmap(mislabel(j),mislabel(j),best_C,best_d,best_gamma,best_kernel)+kernelmap(testlab
el(i),testlabel(i)...
         ,best_C,best_d,best_gamma,best_kernel)-
2*kernelmap(mislabel(j),testlabel(i),best_C,best_d,best_gamma,best_kernel))^0.5;
      if dis_L<=region(j);
         o(k,1)=i;
         o(k,2)=j;
         o(k,3)=mislabel(j);
         o(k,4)=dis_L;
         o(k,5)=region(j);
         k=k+1;
      end;
   end;
end;
```

B2.3 Label InsurenceRegion

```
function out=InsurenceRegion_L(x,label,data,ydata,best_C,best_d,best_gamma,best_kernel)
dis=[];
for i=1:size(ydata',1)
if data(i)~= x
dis(i)_L=(kernelmap(label,label,best_C,best_d,best_gamma,best_kernel)+kernelmap(ydata(i),ydata(i),bes
t_C,best_d,best_gamma,best_kernel)-
2*kernelmap(label,ydata(i),best_C,best_d,best_gamma,best_kernel))^0.5;
   end;
```

```
end;
out=(min(dis(dis>0)))/2;
```

B3. Multi- Label Self -Advised Support Vector Machine (mLSA-SVM)

```
function[result_tr,NewResult,result,m,otime,ntime,TrainAccu,ER1,ER2]=LSA_SVMmulC(
traindata,trainlabel,testdata,testlabel,SVMtype,C,D,gamma,r,kernel)
% multi-Label Self Advising SVM
% Outputs
  % TP1,TN1,FP1,FN1: for original SVM
  % TP2,TN2,FP2,FN2: for self-advising SVM
  %Smdata: numbers of miss data in training
  %Sneighb: number of neighbours in test data
  %training: accuracy of training
  %otime: original time for SVM
  %ntime: new time for self advising SVM
%inputs
  %SVMtype,C,D,gamma,kernel: parameters of libsvm
%% use svmclassfiyku, coz it deal with over train data also
best_C=C;best_d=D;best_gamma=gamma;best_kernel=kernel;
param = ['-s ', num2str(SVMtype), ' -t ',num2str(kernel), ' -d ',num2str(D), ' -g ', ...
   num2str(gamma), ' -c ',num2str(C) ' -r ',num2str(r)];
if nargin < 5
   K = 1;
end
ty=0;
if(ty==0)
   start_time=cputime;
   m=svmtrain2(trainlabel,traindata,'-s 0 -t 2 -g 12 -c 8');   %training
   end_time=cputime;
   [TrainPredict, TrainAccu,TrainProb]=svmpredict(trainlabel,traindata,m);  %test
   [TestPredict, TestAccu,TestProb]=svmpredict(testlabel,testdata,m);  %test
else
   cmd='-b 1 -s 0 -t 2 -g 12 -c 8';
   start_time=cputime;
   m = ovrtrain(trainlabel,traindata,cmd);   %training
```

```
    end_time=cputime;
    [TrainPredict, TrainAccu,TrainProb]=ovrpredict(trainlabel,traindata,m);  %test
    [TestPredict, TestAccu,TestProb]=ovrpredict(testlabel,testdata,m);  %test
end
otime=toc;
TrainErr = 100 - TrainAccu(1,1)
TestErr = 100 - TestAccu(1,1)
Timetraining = end_time-start_time;
Timetesting=Timetraining;

result_tr=TrainPredict;
result=TestPredict;
PER_tr=TrainProb;
PER=TestProb;
NewResult=result;
%%
[mdata,mislabel,mper,t,mIndex,qdata,qlabel]=missdata(traindata,trainlabel,m);
if ~isempty(qdata)
    training=t(1);  %accuracy of training
    Smdata=0;
    if ~isempty(mdata)
        Smdata=size(mdata,1);
    end;

o=closetest3(qdata,qlabel,testdata,testlabel,mdata,mislabel,best_C,best_d,best_gamma,best_kernel,mIndex);
    Sneighb=0;
    pos=0;
    neg=0;
    if ~isempty(o)
        [b,b1]=unique(o(:,1));
        Sneighb=length(b);
        for i=1:Sneighb
            bb=find(o(:,1)==b(i));
            S1=0;
            S2=0;
            S3=0;
            S4=0;
            S=0;
            for j=1:length(bb)
```

```
            if o(bb(j),3)==1
               S1=S1+(abs(mper(o(bb(j),2))))/(1+o(bb(j),4)^2);
           elseif o(bb(j),3)==2
               S2=S2+(abs(mper(o(bb(j),2))))/(1+o(bb(j),4)^2);
            elseif o(bb(j),3)==3
               S3=S3+(abs(mper(o(bb(j),2))))/(1+o(bb(j),4)^2);
            elseif o(bb(j),3)==4
               S4=S4+(abs(mper(o(bb(j),2))))/(1+o(bb(j),4)^2);
            end;
            end;
         if (S1>S2 && S1>S3)&& (S1>S4);
            S=S1;
         elseif (S2>S1&& S2>S3) && (S2>S4);
            S=S2;
         elseif (S3>S1&& S3>S2) && (S3>S4);
            S=S3;
         elseif (S4>S1&& S4>S2)&& (S4>S3);
            S=S4;
         end;
         if S>=(abs(PER(o(bb(j),1))))
            NewResult(o(bb(j),1))=o(bb(j),3);
         end;
      end;
   end;
   ntime=toc;
   [TP1, TN1, FP1, FN1,TP2, TN2, FP2, FN2,ER1,ER2] =measures1(result,NewResult,testlabel);
else
   TP1=0;
   TN1=0;
   FP1=0;
   FN1=0;
   TP2=0;
   TN2=0;
   FP2=0;
   Smdata=0;
   Sneighb=0;
   training=0;
   otime=0;
   ntime=0;
```

```
    FN2=0;
  end
  end
```

B4. Extreme Learning Machine-Label Self - Advised Support Vector Machine (ELM-LSA-SVM)

```
  %        adding this part in the maine program after reduction feature coded
  %% === Classification using ELM ====
    feat_train=[class_training feature_training];
   feat_test=[class_testing feature_testing];
 node=100;
  act_fn='sig';
[error_training1,error_testing1,classification_testing1,timeTrain,timeTest]...
    = elmclassify_node(feat_train,feat_test,2^14,node,act_fn);
```

APPENDIX C: Statistical Performance Equations.

In this Appendix, some Equations from **chapter 2 and 4** associated with the proposed performance of LSA-SVM algorithms are obtainable. Complete statistics about these equations are present in section 4.2.5.

Table C Measures for Evaluating Performance of a Classifier

Evaluation Parameters
$\mathbf{Sensitivity} = \frac{\mathbf{TP}}{\mathbf{TP+FN}}\ \mathbf{x\ 100\%}$
$\mathbf{specificity} = \frac{\mathbf{TN}}{\mathbf{TN+FP}}\ \mathbf{x\ 100\%}$

Where TN (true negative) is the number of the sample not labeled,

TP (true positive) is the number of a sample labeled (correct sampled),

FN (false negative) is the number of sample unknown and not labeled,

FP (false positive) is the number of sample unknown labeled.

For chapter 3 a non-linear piecewise continuous function, such as a sigmoid, hard limit, Gaussian, and multi-quadratic functions. The mathematic equations as shown in below: Gaussian function

$$Q(w, b, x) = \exp(-b\|x - w\|^2)$$

Sigmoid function

$$Q(w, b, x) = \frac{1}{1 + exp\big(-(w.x + b)\big)}$$

Hard-limit function

$$Q(w, b, x) = \begin{cases} 1, & \text{if } (w.x - b) \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

Multi-quadratic function

$$Q(w, b, x) = (\|x - w\|^2 + b^2)^{1/2}$$

References

1. Westhout, F.D., L.S. Paré, and M.E. Linskey, *Central Causes of Foot Drop: Rare and Underappreciated Differential Diagnoses.* J Spinal Cord Med, 2007. **30**(1): p. 62-6.
2. ., G.E.o.N.D., *Foot Drop.* Retrieved March 16, 2018 from
3. Bajd, T., et al., *Symmetry of FES responses in the lower extremities of paraplegic patients.* Journal of Biomedical Engineering, 1990. **12**(5): p. 415-418.
4. Minassian, K., et al., *Neuromodulation of lower limb motor control in restorative neurology.* Clinical Neurology and Neurosurgery, 2012. **114**(5): p. 489-497.
5. Stein, R.B., et al., *A multicenter trial of a footdrop stimulator controlled by a tilt sensor.* Neurorehabilitation and neural repair, 2006. **20**(3): p. 371-379.
6. Kim, J.-H., et al., *Functional electrical stimulation applied to gluteus medius and tibialis anterior corresponding gait cycle for stroke.* Gait & Posture, 2012. **36**(1): p. 65-67.
7. Springer, S., et al., *Effects of Dual-Channel Functional Electrical Stimulation on Gait Performance in Patients with Hemiparesis.* The Scientific World Journal, 2012. **2012**: p. 8.
8. Roche, A.D., et al., *Prosthetic Myoelectric Control Strategies: A Clinical Perspective.* Current Surgery Reports, 2014. **2**(3): p. 1-11.
9. Young, A.J., L.J. Hargrove, and T.A. Kuiken, *Improving Myoelectric Pattern Recognition Robustness to Electrode Shift by Changing Interelectrode Distance and Electrode Configuration.* Biomedical Engineering, IEEE Transactions on, 2012. **59**(3): p. 645-652.
10. V., V., *Statistical learning theory*. 1998, New york: Wiley.
11. Masood, A., *Developing improved algorithms for detection and analysis of skin cancer.* University of Technology Sydney, Faculty of Engineering and Information Technology, July, 2016.
12. Maali, Y., *An Automatic Sleep Apnea Analysis withSoft Computing Approaches.* thesis, 2014.
13. Masood, A. and A. Al-Jumaily. *SA-SVM based automated diagnostic System for Skin Cancer*. in *Sixth International Conference on Graphic and Image Processing (ICGIP 2014)*. 2015. International Society for Optics and Photonics.
14. Huang, G.B., et al., *Extreme Learning Machine for Regression and Multiclass Classification.* IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 2012. **42**(2): p. 513-529.
15. Sakai, H., C. Liu, and M. Nakata. *Information Dilution: Granule-Based Information Hiding in Table Data - A Case of Lenses Data Set in UCI Machine*

Learning Repository. in *2016 Third International Conference on Computing Measurement Control and Sensor Network (CMCSN)*. 2016.

16. Dua, D.a.G., Casey, *{UCI} Machine Learning Repository.* University of California, Irvine, School of Information and Computer Sciences, 2017.

17. Ashok, P. and G.M.K. Nawaz. *Detecting outliers on UCI repository datasets by Adaptive Rough Fuzzy clustering method*. in *2016 Online International Conference on Green Engineering and Technologies (IC-GET)*. 2016.

18. Cinnamon L. VanPutte, J.L.R., Andrew F. Russo, *Seeley's Essentials of Anatomy & Physiology*. Vol. 8e. 2013: McGraw-Hill Education Learning Technology Specialist.

19. Faez, M., *ANATOMY OF THE LEG.* 2010: p. 39.

20. Udo.Savalli.us, *The muscles shape and places* 2010.

21. Ottobock, ★*MyGait Functional Electrical Stimulation Device for Drop Foot.* Aug 22, 2013

22. GmbH, M.O.B.H.P., *On the way to enhanced quality of life*. p. 4-8.

23. Fatone, S. and A.H. Hansen, *Effect of ankle-foot orthosis on roll-over shape in adults with hemiplegia.* Journal of Rehabilitation Research & Development, 2007. **44**(1): p. 11-20.

24. De Luca, C., *Electromyography*, in *Encyclopedia of Medical Devices and Instrumentation*. 2006, John Wiley & Sons, Inc.

25. Khushaba, R., *Application of Biosignal-Driven Intelligent Systems for Multifunction Prosthesis Control*. 2010, University of Technology, Sydney.

26. Estigoni EH 1, F.C., Song T 2 and Davis GM, *A novel system for evoked EMG analysis: Design, construction, and validation.* Annual Conference of the International FES Society, 2009: p. 1-3.

27. De Luca, C.J., et al., *Decomposition of Surface EMG Signals.* Journal of Neurophysiology, 2006. **96**(3): p. 1646-1657.

28. De Luca, C.J., *Surface electromyography: Detection and recording.* DelSys Incorporated, 2002. **10**: p. 2011.

29. en.wikipedia.org, *EMG signal and TA muscle* 2010.

30. Troiano, A., et al., *Assessment of force and fatigue in isometric contractions of the upper trapezius muscle by surface EMG signal and perceived exertion scale.* Gait & Posture, 2008. **28**(2): p. 179-186.

31. Kai, S. and K. Nakabayashi, *Evoked EMG Makes Measurement of Muscle Tone Possible by Analysis of the H/M Ratio.* 2013.

32. Gopura, R., et al., *Recent Trends in EMG-Based Control Methods for Assistive Robots.* 2013.

33. Adil, S., A.A. Jumaily, and K. Anam. *AW-ELM-based Crouch Gait recognition after ischemic stroke*. in *2016 5th International Conference on Electronic Devices, Systems and Applications (ICEDSA)*. 2016.

34. Englehart, K. and B. Hudgins, *A robust, real-time control scheme for multifunction myoelectric control.* IEEE Trans. Biomed. Eng., 2003. **50**(7): p. 848-854.

35. Oskoei, M.A. and H. Huosheng, *Support Vector Machine-Based Classification Scheme for Myoelectric Control Applied to Upper Limb.* IEEE Trans. Biomed. Eng., 2008. **55**(8): p. 1956-1965.

36. Hargrove, L.J., K. Englehart, and B. Hudgins, *A comparison of surface and intramuscular myoelectric signal classification.* IEEE Trans. Biomed. Eng,, 2007. **54**(5): p. 847-853.

37. Ortiz-Catalan, M., R. Brånemark, and B. Håkansson, *BioPatRec: A modular research platform for the control of artificial limbs based on pattern recognition algorithms.* Source Code Biol Med, 2013. **8**: p. 11.

38. Khushaba, R.N., L. Shi, and S. Kodagoda. *Time-dependent spectral features for limb position invariant myoelectric pattern recognition*. in *Communications and Information Technologies (ISCIT), 2012 International Symposium on*. 2012. IEEE.

39. Phinyomark, A., P. Phukpattaranont, and C. Limsakul, *Feature reduction and selection for EMG signal classification.* Expert Systems with Applications, 2012. **39**(8): p. 7420-7431.

40. Tkach, D., H. Huang, and T.A. Kuiken, *Study of stability of time-domain features for electromyographic pattern recognition.* Journal of NeuroEngineering and Rehabilitation, 2010. **7**(1): p. 21.

41. Wheeler, D., *Are the Skewness and Kurtosis Useful Statistics.* April 2008.

42. Yu, W., et al., *Embedding with Autoencoder Regularization*, in *Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2013, Prague, Czech Republic, September 23-27, 2013, Proceedings, Part III*, H. Blockeel, et al., Editors. 2013, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 208-223.

43. Vinita Dutt, V.C., Imran Khan, *Pattern Recognition: an Overview.* American Journal of Intelligent Systems 2012: p. 23-27.

44. Maali, Y. and A. Al-Jumaily, *Self-advising support vector machine.* Knowledge-Based Systems, 2013. **52**(Supplement C): p. 214-222.

45. Rudd, J.M., *Application of Support Vector Machine Modeling and Graph Theory Metrics for Disease Classification.* 07/2017.

46. Xiaojing, Y., et al. *SVM-based Texture Classification and Application to Early Melanoma Detection*. in *28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2006. EMBS '06.* . 2006. New york, USA.

47. He, X., et al., *A simplified multi-class support vector machine with reduced dual optimization.* Pattern Recognition Letters, 2012. **33**(1): p. 71-82.

48. Chaudhary, V., R.S. Bhatia, and A.K. Ahlawat, *An efficient self-organizing map (E-SOM) learning algorithm using group of neurons*. International Journal of Computational Intelligence Systems, 2014. **7**(5): p. 963-972.

49. Hsu, S.-H., et al., *A two-stage architecture for stock price forecasting by integrating self-organizing map and support vector regression.* Expert Systems with Applications, 2009. **36**(4): p. 7947-7951.

50. Huang, G.B., et al., *Extreme learning machine for regression and multiclass classification.* IEEE Trans. Syst. Man Cybern. B, Cybern., 2012. **42**(2): p. 513-529.

51. Huang, J., et al., *Learning Label-Specific Features and Class-Dependent Labels for Multi-Label Classification.* IEEE Transactions on Knowledge and Data Engineering, 2016. **28**(12): p. 3309-3323.

52. Wu, Q., et al., *ML-FOREST: A Multi-Label Tree Ensemble Method for Multi-Label Classification.* IEEE Transactions on Knowledge and Data Engineering, 2016. **28**(10): p. 2665-2680.

53. Grigorios, T. and K. Ioannis, *Multi-Label Classification: An Overview*, in *Database Technologies: Concepts, Methodologies, Tools, and Applications*, E. John, Editor. 2009, IGI Global: Hershey, PA, USA. p. 309-319.

54. Huang, J., et al. *Group sensitive Classifier Chains for multi-label classification*. in *2015 IEEE International Conference on Multimedia and Expo (ICME)*. 2015.

55. Réjichi, S. and F. Chaabane. *Satellite image time series classification and analysis using an adapted graph labeling*. in *2015 8th International Workshop on the Analysis of Multitemporal Remote Sensing Images (Multi-Temp)*. 2015.

56. Dsouza, K.J. and Z.A. Ansari. *A Novel Data Mining Approach for Multi Variant Text Classification*. in *2015 IEEE International Conference on Cloud Computing in Emerging Markets (CCEM)*. 2015.

57. Chandran, S.A. and J.R. Panicker. *An efficient multi-label classification system using ensemble of classifiers*. in *2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT)*. 2017.

58. Jindal, R. and S. Taneja. *Ranking in multi label classification of text documents using quantifiers*. in *2015 IEEE International Conference on Control System, Computing and Engineering (ICCSCE)*. 2015.

59. Sihao, L., et al. *Multi-label extreme learning machine based on label matrix factorization*. in *2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA)(*2017.

60. Spolaôr, N., et al. *Label Construction for Multi-label Feature Selection*. in *2014 Brazilian Conference on Intelligent Systems*. 2014.

61. Sapozhnikova, E.P. *Multi-label Classification with ART Neural Networks*. in *2009 Second International Workshop on Knowledge Discovery and Data Mining*. 2009.

62. Ye, C., et al. *Multi-label active learning with label correlation for image classification*. in *2015 IEEE International Conference on Image Processing (ICIP)*. 2015.

63. F. Herrera et al., *Multilabel Classification,*. Springer International Publishing Switzerland 2016.

64. Sevakula, R.K. and N.K. Verma, *Assessing Generalization Ability of Majority Vote Point Classifiers.* IEEE Transactions on Neural Networks and Learning Systems, 2017. **28**(12): p. 2985-2997.

65. Krakauer, J.W., *Motor learning: its relevance to stroke recovery and neurorehabilitation.* Current Opinion in Neurology, 2006. **19**(1).

66. Ping, W., et al. *Modulation of weight off-loading level over body-weight supported locomotion training*. in *Rehabilitation Robotics (ICORR), 2011 IEEE International Conference on*. 2011.

67. SHEALY, C.N., J.T. MORTIMER, and J.B. RESWICK, *Electrical Inhibition of Pain by Stimulation of the Dorsal Columns: Preliminary Clinical Report.* Anesthesia & Analgesia, 1967. **46**(4): p. 489-491.

68. Bryden, A.M., et al., *Technology for spinal cord injury rehabilitation and its application to youth.* Journal of Pediatric Rehabilitation Medicine, 2012. **5**(4): p. 287-299.

69. Graupe, D., K.H. Kohn, and S. Basseas, *Above- and below-lesion EMG pattern mapping for controlling electrical stimulation of paraplegics to facilitate unbraced walker-assisted walking.* Journal of Biomedical Engineering, 1988. **10**(4): p. 305-311.

70. Monaghan, C.C., et al., *The effect of FES of the tibial nerve on physiological activation of leg muscles during gait.* Medical Engineering & Physics, 2010. **32**(4): p. 332-338.

71. Yan, T., C.W.Y. Hui-Chan, and L.S.W. Li, *Functional Electrical Stimulation Improves Motor Recovery of the Lower Extremity and Walking Ability of Subjects With First Acute Stroke: A Randomized Placebo-Controlled Trial.* Stroke, 2005. **36**(1): p. 80-85.

72. Abbas, J.J. and R.J. Triolo, *Experimental evaluation of an adaptive feedforward controller for use in functional neuromuscular stimulation systems.* Rehabilitation Engineering, IEEE Transactions on, 1997. **5**(1): p. 12-22.

73. Chen, M., et al., *A self-adaptive foot-drop corrector using functional electrical stimulation (FES) modulated by tibialis anterior electromyography (EMG) dataset.* Medical Engineering & Physics, 2013. **35**(2): p. 195-204.

74. Burdett, R.G. and G. Hassell, *Effects of Three Types of Ankle-Foot Orthoses on the Gait and Bicycling of a Patient with Charcot-Marie-Tooth Disease.* JPO: Journal of Prosthetics and Orthotics, 2004. **16**(1): p. 25-30.

75. Taylor, P.N., et al., *Clinical use of the odstock dropped foot stimulator: Its effect on the speed and effort of walking.* Archives of Physical Medicine and Rehabilitation, 1999. **80**(12): p. 1577-1583.

76. Burridge, J., et al., *Experience of Clinical Use of the Odstock DroppedFoot Stimulator.* Artificial Organs, 1997. **21**(3): p. 254-260.

77. Stein, R.B., et al., *A Multicenter Trial of a Footdrop Stimulator Controlled by a Tilt Sensor.* Neurorehabilitation and Neural Repair, 2006. **20**(3): p. 371-379.

78. Sheffler, L.R. and J. Chae, *Technological Advances in Interventions to Enhance Poststroke Gait.* Physical Medicine and Rehabilitation Clinics of North America, 2013. **24**(2): p. 305-323.

79. Burridge, J., et al., *Phase II trial to evaluate the ActiGait implanted drop-foot stimulator in established hemiplegia.* Journal of Rehabilitation Medicine, 2007. **39**(3): p. 212-218.

80. Chae, J., L. Sheffler, and J. Knutson, *Neuromuscular Electrical Stimulation for Motor Restoration in Hemiplegia.* Topics in Stroke Rehabilitation, 2008. **15**(5): p. 412-426.

81. Foundation, S.N., *Odstock Medical FES.* April 2006: p. 1-16.

82. Therapy, J.o.O.S.P., *new products.* J Orthop Sports Phys Ther, 2006. **36**: p. 625-625.

83. Chesler, N.C. and W.K. Durfee, *Surface EMG as a fatigue indicator during FES-induced isometric muscle contractions.* Journal of Electromyography and Kinesiology, 1997. **7**(1): p. 27-37.

84. Dunning, K., et al., *Neuroprosthesis Peroneal Functional Electrical Stimulation in the Acute Inpatient Rehabilitation Setting: A Case Series.* Physical Therapy, 2009. **89**(5): p. 499-506.

85. Butler, P., et al., *PHYSIOLOGICAL COST INDEX OF WALKING FOR NORMAL CHILDREN AND ITS USE AS AN INDICATOR OF PHYSICAL HANDICAP.* Developmental Medicine & Child Neurology, 1984. **26**(5): p. 607-612.

86. Seth, A., et al., *OpenSim: a musculoskeletal modeling and simulation framework for in silico investigations and exchange.* Procedia Iutam, 2011. **2**: p. 212-232.

87. van der Krogt, M.M., S.L. Delp, and M.H. Schwartz, *How robust is human gait to muscle weakness?* Gait & Posture, 2012. **36**(1): p. 113-119.

88. Xiao, M. and J.S. Higginson, *Muscle function may depend on model selection in forward simulation of normal walking.* Journal of Biomechanics, 2008. **41**(15): p. 3236-3242.

89. Steele, K.M., et al., *Muscle contributions to support and progression during single-limb stance in crouch gait.* Journal of biomechanics, 2010. **43**(11): p. 2099-2105.

90. Liu, M.Q., et al., *Muscle contributions to support and progression over a range of walking speeds.* Journal of biomechanics, 2008. **41**(15): p. 3243-3252.

91. Phinyomark, A., C. Limsakul, and P. Phukpattaranont, *A novel feature extraction for robust EMG pattern recognition.* arXiv preprint arXiv:0912.3973, 2009.

92. Zecca, M., et al., *Control of multifunctional prosthetic hands by processing the electromyographic signal.* Critical Reviews in Biomedical Engineering, 2002. **30**(4-6): p. 459.

93. Shuyang Wang, Z.D., Yun Fu, *Feature Selection Guided Auto-Encoder.* Conference on Artificial Intelligence (AAAI-17), 2017.

94. Vapnik, V., *Statistical learning theory.* Wiley, 1998.

95. Wang, Z. and X. Xue, *Multi-Class Support Vector Machine*, in *Support Vector Machines Applications*, Y. Ma and G. Guo, Editors. 2014, Springer International Publishing: Cham. p. 23-48.

96. Zhan, Y. and D. Shen, *An adaptive error penalization method for training an efficient and generalized SVM.* Pattern Recognition, 2006. **39**(3): p. 342-350.

97. Lau, H.-y., K.-y. Tong, and H. Zhu, *Support vector machine for classification of walking conditions of persons after stroke with dropped foot.* Human movement science, 2009. **28**(4): p. 504-514.

98. Masood, A., A.A.-. Jumaily, and K. Anam. *Self-supervised learning model for skin cancer diagnosis*. in *2015 7th International IEEE/EMBS Conference on Neural Engineering (NER)*. 2015.

99. Fletcher, R., *Practical methods of optimization*. 2013: John Wiley & Sons.

100. Kolesov, A., et al., *On Multilabel Classification Methods of Incompletely Labeled Biomedical Text Data.* Computational and Mathematical Methods in Medicine, 2014. **2014**: p. 11.

101. Read, J., et al., *Classifier chains for multi-label classification.* Machine Learning, 2011. **85**(3): p. 333.

102. Li, Y., Y. Song, and J. Luo. *Improving Pairwise Ranking for Multi-label Image Classification*. in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2017.

103. Wu, X., D. He, and G. Zhou. *The Uniformization and the Feature Selection about the Inconsistent Classification Data Set*. in *2009 WRI Global Congress on Intelligent Systems*. 2009.

104. Trejo, R.L., et al. *Hand goniometric measurements using leap motion*. in *2017 14th IEEE Annual Consumer Communications & Networking Conference (CCNC)*. 2017.

105. Adil, S., et al. *Extreme learning machine based sEMG for drop-foot after stroke detection*. in *2016 Sixth International Conference on Information Science and Technology (ICIST)*. 2016.

106. Chan, A.D.C. and G.C. Green. *Myoelectric control development toolbox*. in *Proc. 30th Conf.Canadian Med. Bio. Eng. Soc.* 2007.

107. Schlogl, A. and C. Brunner, *BioSig: a free and open source software library for BCI research.* Computer, 2008. **41**(10): p. 44-50.

108. Tkach, D., H. Huang, and T.A. Kuiken, *Study of stability of time-domain features for electromyographic pattern recognition.* J Neuro Eng. Rehabil., 2010. **7**(1): p. 21.

109. Steele, K.M., et al., *Muscle contributions to vertical and fore-aft accelerations are altered in subjects with crouch gait.* Gait & posture, 2013. **38**(1): p. 86-91.

110. Chang, C.-C. and C.-J. Lin, *LIBSVM: a library for support vector machines.* ACM Trans. Int. Syst. Tech. (TIST), 2011. **2**(3): p. 27.

111. Mohammed, A.A. and A. Sajjanhar. *Robust single-label classification of facial attributes*. in *2017 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*. 2017.

112. Read, J., et al., *Classifier Chains for Multi-label Classification*, in *Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2009, Bled, Slovenia, September 7-11, 2009, Proceedings, Part II*,

W. Buntine, et al., Editors. 2009, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 254-269.

113. Osojnik, A., P. Panov, and S. Džeroski, *Multi-label classification via multi-target regression on data streams.* Machine Learning, 2017. **106**(6): p. 745-770.

114. Ren, D., et al. *Online biomedical publication classification using Multi-Instance Multi-Label algorithms with feature reduction*. in *2015 IEEE 14th International Conference on Cognitive Informatics & Cognitive Computing (ICCI*CC)*. 2015.

115. Huang, G.-B., X. Ding, and H. Zhou, *Optimization method based extreme learning machine for classification.* Neurocomputing, 2010. **74**(1): p. 155-163.

116. Anam, K., A. Al Jumaily, and Y. Maali, *Index finger motion recognition using self-advise support vector machine*. Vol. 7. 2014. 644-657.

117. Huang, G.-B., *An Insight into Extreme Learning Machines: Random Neurons, Random Features and Kernels.* Cognitive Computation, 2014. **6**(3): p. 376-390.

118. Chi, C., T. Wee Peng, and G.B. Huang. *Extreme learning machines for intrusion detection*. in *The 2012 International Joint Conference on Neural Networks (IJCNN)*. 2012.

119. Wang, G.-G., et al., *Self-adaptive extreme learning machine.* Neural Computing and Applications, 2016. **27**(2): p. 291-303.