DEVELOPMENT OF A REAL-TIME GRASPING PATTERN CLASSIFICATION SYSTEM BY FUSING EMG-VISION FOR HAND PROSTHESES

Gamage Dulanjana Manoj Perera

(198115E)

Degree of Master of Science

Department of Mechanical Engineering

University of Moratuwa

Sri Lanka

August 2021

Development of a Real-time Grasping Pattern Classification System by Fusing EMG-Vision for Hand Prostheses

Gamage Dulanjana Manoj Perera

(198115E)

Thesis submitted in partial fulfillment of the requirements for the degree Master of Science in Mechanical Engineering

Department of Mechanical Engineering

University of Moratuwa

Sri Lanka

August 2021

DECLARATION

I declare that this is my own work and this thesis does not incorporate without acknowledgment any material previously submitted for a Degree or Diploma in any other University or institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgment is made in the text.

Also, I hereby grant to University of Moratuwa the non-exclusive right to reproduce and distribute my thesis, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Signature: Date:

The above candidate has carried out research for the Masters thesis under our supervision.

Dr. H. K. G. Punchihewa
Head/Senior Lecturer,
Department of Mechanical Engineering,
University of Moratuwa, Sri Lanka.

Late, Dr. D. G. K. Madusanka
Lecturer,
Department of Mechanical Engineering,
University of Moratuwa, Sri Lanka.

Abstract

The Electromyography (EMG) based trans-radial prostheses have revolutionized the prosthetic industry due to their ability to control the robotic hand using human intention. Although recently developed EMG-based prosthetic hands can classify a significant number of wrist motions, classifying grasping patterns in real-time is challenging. However, the wrist motions alone cannot facilitate a prosthetic hand to grasp objects properly without performing appropriate grasping pattern. The collaboration of EMG and vision has addressed this problem to a certain extent. However they have not been able to achieve significant performance in real-time.

This study proposed a vision-EMG fusion method that can improve the real-time prediction accuracy of the EMG classification system by merging a probability matrix that represents the usage of the six grasping patterns for the targeted object. The You Only Look Once (YOLO) object detection algorithm was utilized to retrieve the probability matrix of the identified object, and it was used to correct the classification error in the EMG classification system by applying Bayesian fusion. Experiments were carried out to collect EMG data from six muscles of 15 subjects during the grasping action for classifier development. In addition, an online survey was conducted to collect data to calculate the respective conditional probability matrix for selected objects. Finally, the five optimized supervised learning EMG classifiers; Artificial Neural Network (ANN), K-nearest neighbor (KNN), Linear Discriminant Analysis (LDA), Naive Bayes (NB), and Decision Tree (DT) were compared to select the best classifier for fusion.

The real-time experiment results revealed that the ANN outperformed other selected classifiers by achieving the highest mean True Positive Rate (mTPR) of $M = 72.86\%$ ($SD = 17.89\%$) for all six grasping patterns. Furthermore, the feature set identified at the experiment (Age, Gender, and Handedness of the user) proved that their influence increases the mTPR of ANN by $M = 16.05\%$ ($SD = 2.70\%$). The proposed system takes $M = 393.89\ ms$ ($SD = 178.23\ ms$) to produce a prediction. Therefore, the user did not feel a delay between intention and execution. Furthermore, proposed system facilitated the user to use suitable multiple grasping patterns for a single object as in real life. In future research works, the functionalities of the system should be expanded to include wrist motions and evaluate the system on amputees.

Keywords **-Surface Electromyography, Real-time Classification, vision feedback, Grasping Pattern, Sensor Fusion**

DEDICATION

In memory of late Dr. Kanishka Madusanka

and

to my loving family who keeps lifting me and inspiring me

in every second of my life.

ACKNOWLEDGMENTS

I received invaluable support and guidance from many people to complete this research work successfully throughout this intense period. I would like to express my sincere gratitude towards all these people who were there for me during my ups and downs.

I am indebted to my thesis supervisor, Late Dr. Kanishka Madusanka, for the encouragement and insightful guidance that he gave me during this intensive period. He was a teacher and a supportive friend who was always there for me when I was down. He allowed this thesis to be my own work, but steered me in the right direction with valuable suggestions. Furthermore, I am incredibly grateful to my co-supervisor, Dr. Himan Punchihewa, for his precious support and suggestions rendered at a crucial stage of the research. I owe my deepest gratitude to him for agreeing to be my supervisor after Dr. Kanishka and spiritually uplifting me to complete the research.

I would like to express my sincere gratitude to Professor Ruwan Gopura for the insightful feedback and endless support he gave me during hard times. Without his administrative support, this thesis would not have been possible. I would also like to extend my gratitude towards Dr. Damith Chathuranga for providing supportive comments and suggestions at the progress reviews. His suggestions have made this thesis more professional and valuable.

I would be amiss if I did not mention Mr. Pubudu Ranaweera and my lab mates, Mr. Achintha Abayasiri, Mr. Sanka Chandrasiri, and Mr. Achintha Iroshan, who were always with me during my ups and downs. I am also thankful

to all my fellow batchmates from the 14^{th} batch especially, Mr. Lakshitha De Silva, Mr. Charuka Lihini, and Mr. Rawisha Serasinghe, for their friendly and insightful feedback. A special token of appreciation is also extended towards Miss. Supipi Fernando who was there with me from the beginning and encouraged me to do my best. I greatly appreciate her caring support.

I would also like to thank Dr. Nalaka Samaraweera for the administrative guidance and support he gave me during final stages of the thesis work.

Last but not least, I am very grateful to my parents and my brother Mr. Kanushka Perera for their continuous support and encouragement.

Dulanjana Perera,
dulanjana.perera@ieee.org

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

AC	Alternative Current
ADL	Activities of Daily Living
ANN	Artificial Neural Network
AR	Autoregressive Coefficient
BDE	Binary Differential Evolution
BF	Bayesian Fusion
BPSO	Binary Particle Swarm Optimization
CART	Classification and Regression Trees
CNN	Convolutional Neural Network
COG	Center Of Gravity
CVS	Cognitive Vision System
DNN	Deep Neural Network
DSOD	Deeply Supervised Object Detectors
DSSD	Deconvolutional Single Shot Detector
DT	Decision Tree
DV	Dependent Variables
ECG	Electroencephalography
EEG	Electroencephalography
EMG	Electromyography
FCNN	Fuzzy Clustering Neural Network
FD	Frequency Domain
F-RCNN	Faster Region-based Convolutional Neural Networks
GLM	General Linear Model
HD-EMG	High Density Electromyography
HL	Hidden Layer

HMM	Hidden Markov Model
HSV	Hue Saturation Value
iEMG	intramuscular Electromyography
IV	Independent Variables
IMU	Inertia Measurement Unit
KNN	K-Nearest Neighbor
LDA	Linear Discriminant Analysis
LR	Learning Rate
mAP	mean Average Precision
MAV	Mean Absolute Value
MBTGA	Modified Binary Tree Growth Algorithm
MC	Metacarpals
MC	Momentum Constant
MLP	Multi-layer Perceptron
MMG	Mechanomyography
mTPR	mean True Positive Rate
MV	Majority Vote
NB	Naive Bayes
NCS	Nerve Conduction Study
NOD	Normalized Onset Distance
non-PR	non-Pattern Recognition
PBPSO	Personal Best Guide Binary Particle Swarm Optimization
PCA	Principle Component Analysis
PR	Pattern Recognition
PSO	Particle Swarm Optimization
RGB	Red Green Blue
RMS	Root Mean Square
ROI	Region Of Interest
RTG	Reach-To-Grasp
SD	Standard Deviation
SFS	Sequential Forward Selection
SOM	Self-Organizing Map

SSC	Sign Slope Change
SVM	Support Vector Machine
TD	Time Domain
TDAR	Time Domain Autoregressive Coefficients
TD-AR	Time Domain-Autoregression
TFD	Time-Frequency Domain
WAMP	Willison Amplitude
WL	Waveform length
YOLO	You Only Look Once
ZC	Zero Crossing

Chapter 1

INTRODUCTION

There are 57.7 million people suffering from amputation and approximately 22.3 million of them are upper limb amputees [13]. South Asia alone has a prevalence of traumatic amputation of around 9.7 million. Therefore, the demand for prosthetic devices is significant. Prosthetic limb is an artificial device that restores the function of a missing limb [13]. Some amputees desire cosmetic devices that provide a lifelike appearance, and others demand active prosthetic devices that can perform movements or the primary functions of the limb (see Figure 1.2). The most common upper limb amputation is *transradial amputation* [13]. According to the Cody *et al.*'s study, the leading traumatic cause for this amputation was falling from a higher level, and the second leading cause was road injuries [13]. The trans-radial amputation results in the loss of the entire hand and most of the lower part of the forearm [14]. Hence, a prosthetic hand should be able to provide the functions of the hand and the wrist in order to perform activities of daily living (ADL) to a certain extent. Robotic active prosthetic hands are popular because of their futuristic appearance and ability to provide primary functions of the lost limb, such as basic hand gestures and finger motions [15]. Moreover, a robotic prosthesis has higher controllability over other active (body-powered) or passive prosthesis because it uses electric actuators to provide complex motions [15]. However, prostheses cannot provide the dexterity of the natural limb because artificial devices fail to achieve the flexibility of the natural limb and sophisticated controlling.

Researchers and scientists have been trying to develop functional prosthetic

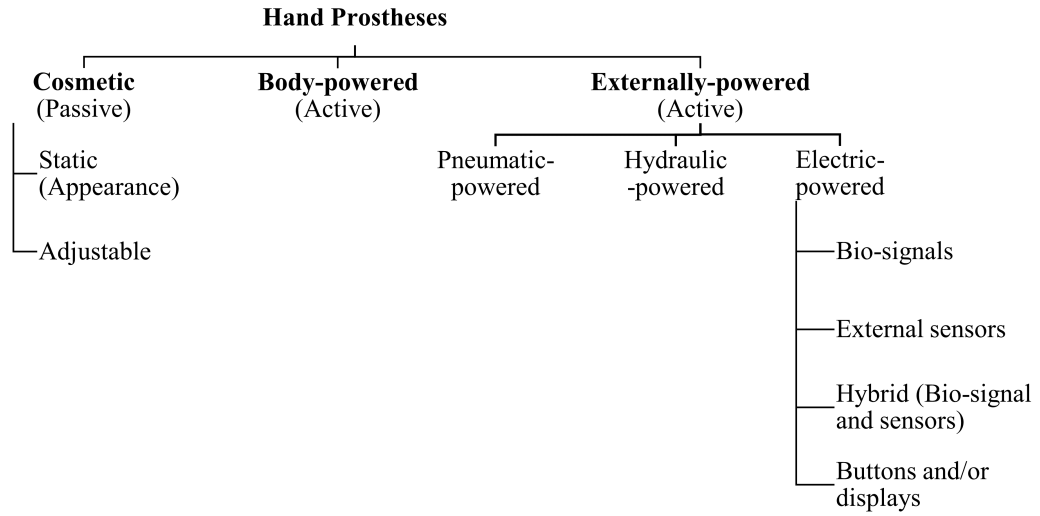


Figure 1.1: Basic categorization of hand prostheses

hands since the early 20^{th} century [3]. G. Schlesinger's development of an electric prosthetic hand in 1919 was a major milestone of prosthetic industry [16]. Before that, body-powered and gas-powered prosthesis dominated the industry [3]. The issue with body/gas-powered prosthetic hands was their limited motion capability and controlling the prosthetic hand using body movements was a grueling task. Therefore, amputees tend to demand electric hands. However, greater advancements started after Daniel Graupe *et al.'s* experiment on Electromyography (EMG) and pattern recognition for prosthetic control in 1978 [17]. Thereafter, the EMG-based prosthetic limb design has been receiving significant attention from the research community.

The Electromyography (EMG) signal is a biological signal that generates in the muscles (in the forearm or leg) during activities. Since it is an electrical signal and foundational theories for signals had been introduced at that time, detecting and controlling it was possible. Even though EMG signal was a reliable and practical source for controlling the prosthetic hand, other biological signal sources such as Mechanomyography (MMG) and Electroencephalography (EEG) were also investigated to study their performance in controlling a prosthetic hand [12, 18–20]. The results have indicated that MMG is far less reliable than EEG, and EEG has shown significant improvement in controlling the prosthetic hand.

(a) (b) (c)

Figure 1.2: (a) Cosmetic leg developed by Ottobock [1]. (b) Body-powered upper limb developed by Ottobock [1]. (c) i-Limb developed by Ossure [2]

However, EEG-based prosthetic limbs are difficult to practically implement, since the EEG acquisition is a more complex task than EMG acquisition.

Controlling of the hand or leg prostheses using biological signals can be achieved in two methods: Proportional control and pattern-recognition-based control methods. The proportional control methods are suitable to control the hand kinematics and dynamics (force and torque), but it fails to identify the user's intention. Pattern-recognition methods can predict the user's intention and control the hand accordingly [17]. It is known as pattern classification. However, classification of grasping patterns using EMG signals is a challenging task because of its stochastic nature, and the primary muscles involved in finger motion (grasping patterns) are located on the missing hand (intrinsic muscles) [21]. Furthermore, the EMG signal change over time and the physiological state (arm reaching speed, arm position, and external weight) of the stump arm [22–25]. Since it directly influences the performance of the control systems, most of the research works are limited to hand gestures (wrist motions and nonverbal communication signals). In contrast, some investigations have classified few grasping patterns in offline conditions when the hand is in a static position.

The hand gestures classification is not challenging as grasping patterns because the remaining muscles have rich motor information of wrist movement than grasp-

ing patterns [26]. Although, several studies have classified 3-4 grasping patterns during the reaching phase in real-time conditions with 80-90% accuracy [27]. The real-time classification is notably difficult than off-line classification due to the stochastic nature of the signal. Therefore, the system needs supplementary aid to predict the user's intention in addition to the sEMG signal to classify a higher number of grasping patterns in real-time [15]. This is the major problem that prosthetic engineers make strenuous efforts to address.

To address the above issue to a certain extent, studies have been carried out for other sensory inputs such as (visual) perception (i.e., camera), for prosthetic hand controllers [28]. The literature has proposed the Eye-In-Hand approach to collaborate with other sensors such as depth sensors, accelerometers, and laser pointers [9, 29–32]. The results of these investigations confirmed that external sensors provide substantial assistance to the prosthetic controllers. The major drawbacks of these methods are that the system entirely depended on the object detection system. The object classification techniques have been reaching the pinnacle of the computer vision field. Even the current high-end mobile devices can execute complex object detection algorithms that can classify around 50 objects with significant accuracy [33]. Association of the surface Electromyography (sEMG) with perception is the latest trend in the prosthetic field and is called the hybrid system [31, 34, 35]. However, previous methods used sEMG only for the triggering action, such as executing the grasping pattern suggested by the system or selecting the grasping patterns from a list [28,34]. This approach is better than pressing buttons on the prosthetic hand, though it cannot use multiple grasping patterns for a single object in real-time.

In real life, humans use several grasping patterns to move and/or operate an object. The previously proposed systems have assigned one specific grasping pattern for a specific object [34]. Even though few studies have defined multiple grasping patterns for an object, the user must remember the order of the grasping patterns for each object and the triggering scheme [36]. Hence, the system would

not be feasible when a large number of grasping patterns are defined in the system. This is a major drawback in vision-based prosthetic hands, in addition to vision-related issues.

1.1 Motivation

Classifying multiple grasping patterns with high accuracy only using sEMG signals is a challenging task. The previous research suggests that additional input can improve classification performance [15]. Even humans use the visual feedback of the current state of the object to decide the grasping pattern in addition to the past experience regarding the object [37]. Then, the brain sends appropriate motor information as EMG signals to the hand to execute the grasping action (see Figure 1.3). The eye-hand cooperation of a person is essential for efficient grasping. Without visual feedback, prehension movements are affected, especially the finger aperture (Hand opening size) [38, 39]. This biological sensor fusion is a proven effective method for accurate prehension. Hence, the implementation of visual feedback with sEMG classification could lead to a new generation of prosthetic hands and increase the predictable number of grasping patterns in real-time.

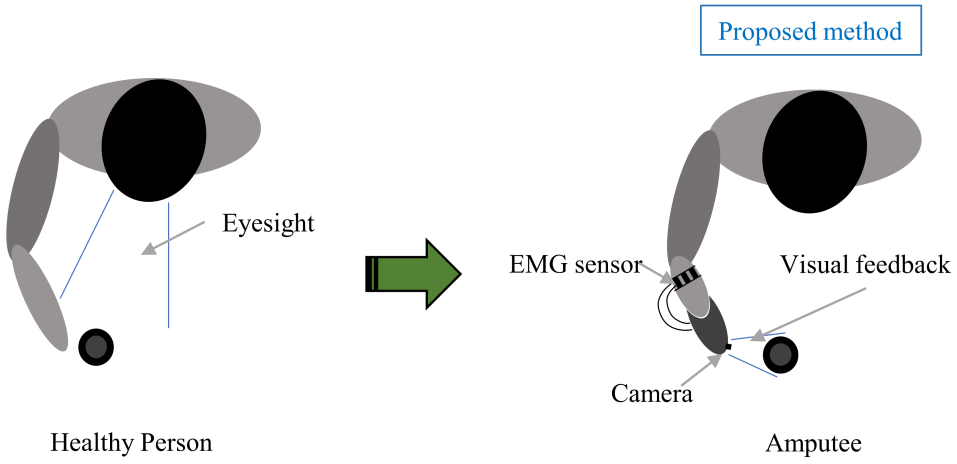


Figure 1.3: Adaptation of visuomotor system of a human to the EMG-based prosthetic hand.

1.2 Aim and objectives of the research

The aim of this research was to develop a classification system that can classify mostly used grasping patterns using vision and sEMG signals.

Objectives

- To develop an algorithm for vision-based grasping pattern classification
- To develop a pattern recognition system to classify the sEMG signals by utilizing predicted data from the developed vision system.
- To validate the system in real-time conditions

To complete the aforementioned objectives, the following methodology was followed. First, a comprehensive literature study was carried out to identify the research gap and the merits of each approach followed so far. Based on the information derived from the literature, experiments and analysis methods were determined to address the research gap. Then, experiments were conducted to collect data, and analyzed to identify opportunities to improve the current methods. sEMG classification systems and object detection systems were developed based on the results and collected data. Sensor fusion method was derived from the existing theories that complement the sEMG and vision information to predict the user intention (grasping pattern) in real-time. Finally, experiments were carried out to validate the proposed system in real-time.

1.3 Contribution to knowledge

This thesis proposes a sensor fusion method that utilizes both sEMG and objects classification results to predict the most suitable grasping patterns for

the identified object. A conditional probability matrix of six grasping patterns for given objects has been proposed. This conditional probability represents the usage of a grasping pattern when user is given corresponding object. The vision system detects the object and retrieves the probability matrix correspond to the identified object. Then, that probability matrix was used to fuse with sEMG classification results to reduce EMG classification error. The real-time experiments conducted to validate the system and results confirm that the fusion improves the current sEMG classification performance while allowing the user to use multiple grasping patterns for a single object. The key contributions to the current knowledge are outlined below.

1. Identified the relationship between demographic information of the user with sEMG signals.
2. Proposed new features for the grasping pattern classification systems and evaluated their performances.
3. Proposed a Bayesian fusion method to improve sEMG classification accuracy using vision feedback (conditional probability matrix).

1.4 Thesis overview

The thesis is composed of eight Chapters. The second Chapter presents the history of the prosthetic devices and the recent developments in the trans-radial prosthetic control systems. The literature review discusses both sEMG-based and vision-based hand prosthetic controllers and the hybrid systems which are used both vision and other external sensors. Chapter 3 reports the experiments conducted to collect sEMG data, and the analysis that was carried out to identify potential factors associated with sEMG data that benefited in sEMG classification. Chapters 4 present the development procedure of the sEMG classification system and the vision system. The classifier development and optimization is

discussed and the selection of the best classifier for the real-time classification is presented in the initial section of the Chapter 4. The implementation of the object detection algorithm in the vision system is discusses in the second section in Chapter 4. The proposed hybrid system is discussed in Chapter 5 with the derivation of the Bayes theorem for the sensor fusion. Chapter 6 presents the overall system validation in real-time conditions with and without the proposed fusion method. The thesis is concluded in Chapter 7.

Chapter 2

LITERATURE REVIEW

A comprehensive literature study was carried out to identify the research gap and the current methods to develop grasping pattern classification systems. The research articles were searched in Scopus and other scholarly databases. The following keywords were used to search the articles. The terms 'electromyography', 'myoelectric', 'prosthetic hand', 'grasping pattern', 'real-time' and 'classification' were used to search the articles related to EMG-based grasping pattern classification system. In order to search articles related to vision-based grasping pattern classification, instead of the terms 'electromyography' and 'myoelectric', the terms 'vision', 'vision-based' and 'perception' were used. The articles were first screened by reading the abstract. Then, the articles related to grasping or hand gesture classification were selected. In the second screening process, articles without proper evaluation criteria were rejected. For that, the results section of the articles was examined. Then articles were grouped according to their evaluation mode, which was real-time and offline. Finally, the data extractions were conducted. Several specific articles which were cited in screened articles were obtained to examine and extract useful information for this research. The information on the history of the prosthetic devices was also extracted from articles and reputed websites.

The initial sections of the review is about the evolution of the prostheses, and the latter parts of the review are more focused on the modern upper limb prosthetic controllers (proportional controllers and classification controllers), especially the Trans-radial prosthetic controllers. The literature review is more

focused on sEMG-based grasping pattern classification systems and vision-based prosthetic hands. Section 2.3 in the literature review discusses only the real-time and offline pattern classifications systems that classify at least two or more grasping patterns with hand gestures.

2.1 History of the prosthetic devices

The prosthetic devices are used to restore the function of the missing limb. Early devices only provided the appearance. However, with advancements in engineering fields such as mechanical and electrical, devices with humanoid motions were introduced. This section discusses a brief history of prosthetic devices and the evolution of EMG-based prosthetic devices, especially prosthetic hands. The section presents the background information and literature of developments until the 20^{th} century that led to the state-of-the-art prosthetic hands developed today. This section was outlined according to the basic categorization of prosthetic devices illustrated in the Figure 1.1.

2.1.1 Cosmetic prosthetic devices

The cosmetic prosthetic is the fundamental device in prostheses that only provide the appearance. The usage of cosmetic prostheses has been mentioned in manuscripts that had been written since the beginning of Egypt, Greek, Indian and Chinese civilizations [3]. Archaeologists have discovered an artificial toe made in 950 B.C. in Egypt, which is the oldest known prosthetic device so far (see Figure 2.1). It was found that the artificial toe is made from wood, and a leather cover was used to mount the toe to the feet [3]. This assembly helped the amputee to maintain balance when walking. A cartonnage toe which is believed to be made in 600 B.C., was found in Egypt, which, however, is used only for cosmetic purposes because of its low flexibility does not provide any facility during walk [40]. Similar inventions can be seen in Greece: Wooden and metal noses

(a) (b) (c)

Figure 2.1: (a) First known prosthetic toe [3]. (b) Roman leg [3]. (c) Artificial iron arm [3].

and Rome: Iron hands, feet, and carved ears. These pre-20th-century prostheses were used as cosmetic devices that substitute for missing limbs or body parts.

2.1.2 Body-powered prosthetic devices

In the 19th-century, inventors implemented various functionalities to these artificial devices so that the user could utilize them for more than just aesthetic appearance [41]. In order to actuate the prosthetic device, cables and pulley systems were integrated to transmit the body motions to the motions in the prosthetic device. Hence, these are known as body-powered prostheses. The literature provides evidence of a prosthetic leg that has a mechanism to control the knee was made from Duralumin alloy in 1928 [3]. Another famous functioning prosthetic device was an artificial arm with a hammer, made for metalworkers. The device consists of several attachments, and the user can change the attachments according to the requirement.

Since then, prosthetic developments were more focused on device functionalities and easy usage. Earlier, functioning prostheses were powered by cable-driven or hand-driven mechanisms, which were commonly known as body-powered pros-

theses [42]. A German dentist developed a body powered prosthetic hand 1818 which has leather straps mount on a shoulder girdle of the amputee [43]. Amputee have to move the intact muscle at arm and shoulder to operate the prosthetic hand. Similar design was introduced in France in 1860 for the wounded soldiers [44]. It can perform open-close the hand and flexion-extension of the fingers. In 1948, Bowden proposed the cable driven body-powered prosthetic hand system instead of leather straps which reduce the bulkiness [45]. These body-powered prostheses have not changed significantly since 1950 to late 90s [46]. Moreover, they were uncomfortable, restrictive and caused unnecessary pressure on the body. Therefore, alternative power source was needed to operate the hand and overcome the aforementioned drawbacks.

2.1.3 Externally-powered prosthetic devices

The externally-powered devices are powered by three basic mediums; hydraulic, pneumatic and electric. This section present the early developments of these three types of powered prosthetic devices. However, review more focused on the electric-powered prosthetic devices, since this research has aimed to develop a control system for electric prosthetic hands.

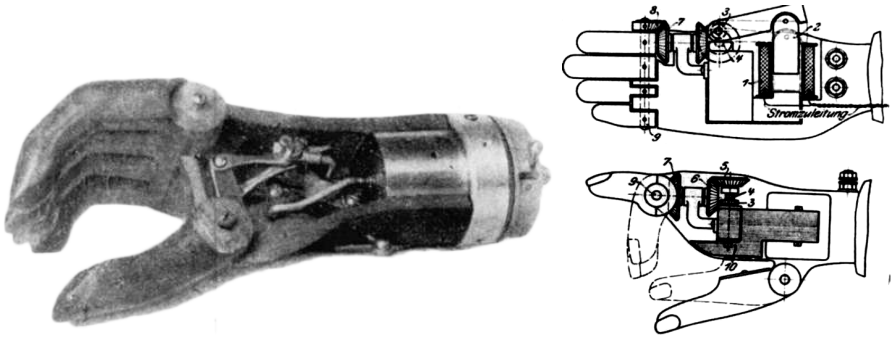
The hydraulic-powered prosthetic foot was introduced by IBM in 1958 [47]. Since then *'Helping Hand'* by Heather *et al.* 1960 [48] and Lewis *et al.* from University of Virginia [49] developed several hydraulic-powered prosthetic devices. In the early 20th century, inventors started to move towards automatic actuation rather than manual operations. As a result of that, gas-powered (CO_2) devices were developed in 1915 in Germany [16]. Even so, in 1963 and 1973, Hugh Steeper developed artificial upper limbs for children who have lost (or have substantially short upper limbs (Phocomelia)) both hands due to Thalidomide [50, 51] which is more advance than German design. Similar technology was used to develop an artificial hand prosthetic that can rotate the wrist using CO_2 gas [52].

These prosthetic devices incorporate several drawbacks, such as the high weight of the system and the limitations in humanoid motion replication. The main reason for these problems is that the mechanisms are significantly bulky. The cable-driven or gas-powered mechanisms required a pully system or pneumatic system to actuate the device. The hydraulic-powered devices have heavy compression system and the system was prone to leak. In addition to bulkiness, achieving precise motion and a human-like motion was challenging.

Electric prosthetic devices

Because of the above mentioned drawbacks, electrical components were implemented in prosthetic devices to improve the quality of the product and its functions. However, the intuition of integrating electrical components in prosthesis was first originated in 1919 in Germany [5]. Alderson collaborated with IBM in 1954 to develop electric-powered limbs [47]. However, they were not widely used since CO2-powered limbs had more functionalities than electric prostheses. Advancement of electrically powered prostheses was begun after Piper's study on recording the electrical potential of surface muscles in 1912 and the discovery of nerve firing frequency [53]. Fundamental principles of sEMG acquisition technology which we use today, were developed by Emil Du Bois-Reymond [54]. He improved Oersted's galvanometer [55] by increasing the number of coils (24,000 coils) to sense the small electrical impulses which are propagating on the muscle during the contraction. This led to significant discoveries such as Piper's Nerve Conduction Study (NCS) about nerve firing frequency [53], Adrian and Bronk's motor unit potential recording device [56], Fritz Buchthal's development of micro-electrode for single muscle fiber in 1934 [57] and development of first commercial EMG machine in 1948 by Jasper, Golseth and Fizzel [58].

Then Reinhold Reiter developed the first myoelectric prosthesis in 1948 [4]. The design was adapted from the Hufner hand and was modified to control electro-magnet. Since it was controlled by a vacuum tube amplifier and powered



(a) (b)

Figure 2.2: (a) First EMG-controlled hand by Reinhold Reiter [4]. (b) First electric hand [5].

externally, the system was bulky and impossible to move. Although Reinhold develops the idea of myoelectric control, several other researchers have found different approaches to control the prosthesis using the myoelectric signal. Professor Norbert Weiner (1947), Berger & Hupper (1952) [59] and Battye *et al.* (1955) [60] were the pioneers in developing myoelectric control prostheses. Transistors were not used in these developments, although they were invented in 1948. The *Russian Hand* which Soviet scientist developed was the first known prosthesis which incorporated transistors [61]. It was used in clinical experiments, and it was sold on a license basis in Canada and the UK.

Since then, significant improvements and groundbreaking milestones were made in the myoelectric-prosthetic field, such as Vaduz hand (1953), also known as French hand [62], English hand (1965) [63], Belgrade Hand (1969) [64], SVEN Hand (1970) [65], Boston Arm (1981) [66] and Otto Bock hands. Until the mid-20th century, researchers and scientists focused on mechanical design, signal acquisition techniques, and system mobility. In the absence of high computational power and methods to analyze complex patterns in signals, EMG signals were limited to proportional control methods. Therefore, the functions of the prosthetic hand were limited. This inspired the researchers to investigate more sophisticated controlling methods using sEMG, such as pattern recognition control mechanisms.

In 1972, Keinosuke Fukunaga proposed a method to identify patterns in a set of data using statistical methods, which revolutionized the machine learning field and significantly affected pattern recognition (PR) applications [17]. In 1978 Daniel Graupe *et al.* proposed a method which adapted from [17], to control a prosthetic arm with several degrees of freedom by discriminating time series data of EMG signal [67]. This was the first robotic prosthetic hand that can automatically determine the user intention using EMG signals. The proposed classifier (Linear discriminant analysis (LDA)) achieved 85% accuracy for classifying four wrist motions in real-time. Several years after, George Saridis *et. al* proposed a different approach to classifying the EMG signal [68]. In this experiment, they considered three signal features which are zero-crossing, variance, and higher-order moments. However, they noticed that motion information was lower in the higher-order moments than the other two features.

These are the major milestone in the evolution of EMG-based prosthetic devices. The following sections review the EMG-based and other methods that are used to control the trans-radial prosthetic hands. Although the review is more focused on the grasping pattern prediction and the hand gestures, more weight of the review is given to the grasping pattern prediction systems.

2.2 State-of-the-art electric prosthetic control systems

This section discusses the recent methods that have been used to control the prosthetic devices. The EMG signal is not the only biological signal that used for the prosthetic control. The EEG and MMG were also used recently for prosthetic control [12, 18–20, 69]. However, those signals are difficult to extract and sensitive to noises which makes them challenging to use for the control system [18]. The EEG extraction from the head, needs specific cap and wearing it all the day to control a prosthetic hand is not practical as well as a tedious task.

Furthermore, recent developments have employed other external sensors such as depth sensors, laser beams, Inertial Measurement Units (IMU) and RGB vision sensors for the control system of the prosthetic device [29–31]. The visual feedback was mainly used to identify the object and determine the corresponding grasping type for the object [9]. The other sensory inputs were used to aid the vision-based control systems. These approaches were also have achieved significant performance although the functionalities were limited. Among these approaches, EMG-based and external sensory inputs such as vision-based approaches have shown opportunities to improve.

2.2.1 EMG and vision sensory inputs for prosthetic control

The pattern recognition (PR) and non-pattern recognition (non-PR) based myoelectric control systems were introduced from time to time. Furthermore, non-PR EMG methods are limited in interpreting the data and are mainly used for low latency and low computational power required applications such as proportional control movements (on/off control) [70]. On the other hand, PR based control system is highly efficient in interpreting the complex movements of the residual arm, such as grasping patterns, hand gestures, and individual finger motions, and it can identify the simultaneous movements [27]. The reason for that is the PR method collects information from the EMG signal in various aspects and finds distinct patterns in them to distinguished the different motions. Unlike EEG, EMG signal can be easily extract using a small electrodes which can be implemented in the prosthetic hand. The collaboration of vision and EMG has significant performance but the EMG was used only for the triggering action [34]. Only few has used EMG for classification but it did not classify the users actual intention [36]. However, studying these approaches have given better perspective of the current methods and their drawbacks. Therefore, the following section discusses only the EMG-based (PR-based) and vision-based control systems of trans-radial prostheses.

2.3 EMG-based prosthetic hand control systems

Recently developed EMG-based PR approaches follow six main steps to classify the EMG patterns from the signal source and control the hand (see Figure 2.3). The review is presented in this order to discuss the merits of the techniques recently used in each section. In this section, EMG-based control methods are discussed, but the post-processing methods are discussed at the final section since, those methods are used for both EMG and vision-based control systems.

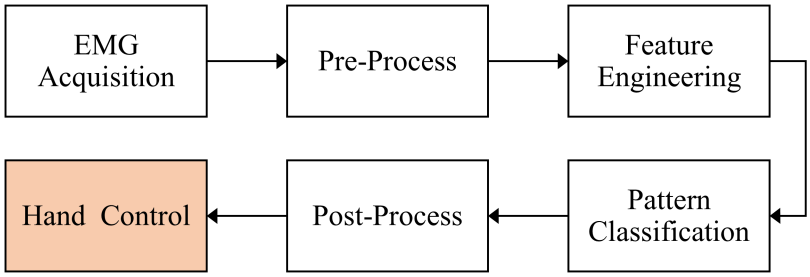


Figure 2.3: Pattern recognition (PR) control scheme for EMG-based prosthetic hand

2.3.1 EMG acquisition

The first step in the EMG classification is to capture the EMG signal from the user. Since the review discusses only the hand prostheses, the EMG signal of the muscles at the arm (hand, forearm, and shoulder) is considered. The surface EMG (sEMG) acquisition is the widely used method because it is practical and straightforward. On the other hand, intramuscular EMG (iEMG) is a complex but reliable method. In order to capture the iEMG, electrodes have to be contacted with muscle belly physically. It has to be surgically achieved by a skilled physician. It has been found that iEMG has significantly less cross talk (i.e., less correlation between electrodes) compared to sEMG. However, as the Figure 2.4 depicts, iEMG classification results have not outperformed the classification accuracy of sEMG significantly [6]. Hargrove *et al.* have analyzed the data with paired t-test and identified that difference between the mean classification accu-

racy of both methods is not significant ($p > 0.100$). Therefore, Hargrove *et al.* suggest that, instead of iEMG, sEMG can be used to classify the signals without a significant information loss. This is a remarkable finding because implementing an iEMG prosthetic control system is a challenging task. Furthermore, it would raise numerous issues such as the inability to detach the device and the complications during the iEMG acquisition. Although it facilitates acquiring EMG signals from deep muscles, performance is similar when certain surface muscles are used.

The most notable finding in this experiment was optimum electrode combination for higher classification accuracy. They claimed that the optimum number of electrodes for both selected muscle arrangement and the symmetrical placing arrangement is four (see Figure 2.6) [6]. An insignificant accuracy increment was observed when electrode count was increased from four to eight in a symmetrical arrangement. A recent study conducted for both intact arm and amputated arm has confirmed Hargrove *et al.'s* conclusion [21]. Moreover, they have stated that the number of channels (muscles) should be within 4-6 to have optimum classification performance for grasping patterns. However, this applies when a single electrode is used for a single muscle. Several studies have used High-Density EMG (HD-EMG) strips with an array of electrodes (usually 32-256 electrodes) to cover the part of the forearm. The recently proposed electrode count for the HD-EMG acquisition is 100 [71, 72]. However, this number depends on where the strip is placed and the other parameters such as window length and the number of patterns classify.

It is important to mention that, with the advancement of the technology miniaturization of sensors, dry electrodes that work as transducers for muscular inputs have replaced traditional gel EMG electrodes and have improved performance [73]. The electrode placement should be done by following a special skin preparation in order to achieve good signal acquisition without any artifacts and interference (i.e., noises) [74]. In the Hermens *at el.* recommendation for sEMG acquisition, the procedure starts with rubbing the surface of the hand

with slightly abrasive material so that, stratum corneum layer of the epidermis can be removed safely. Next, an alcohol wipe is used to remove remaining skin cells. After that, the electrodes can be attached firmly to the hand to achieve low skin-electrode impedance.

Since this research work is interested only in trans-radial amputation (below elbow), muscles that are typically loose after the amputation is intrinsic muscles (muscle at hand) [75]. In many amputation cases, these intrinsic muscles are entirely removed with some portion of extrinsic muscle at the forearm (the stump arm). However, there are eight muscles at the forearm whose tendons are

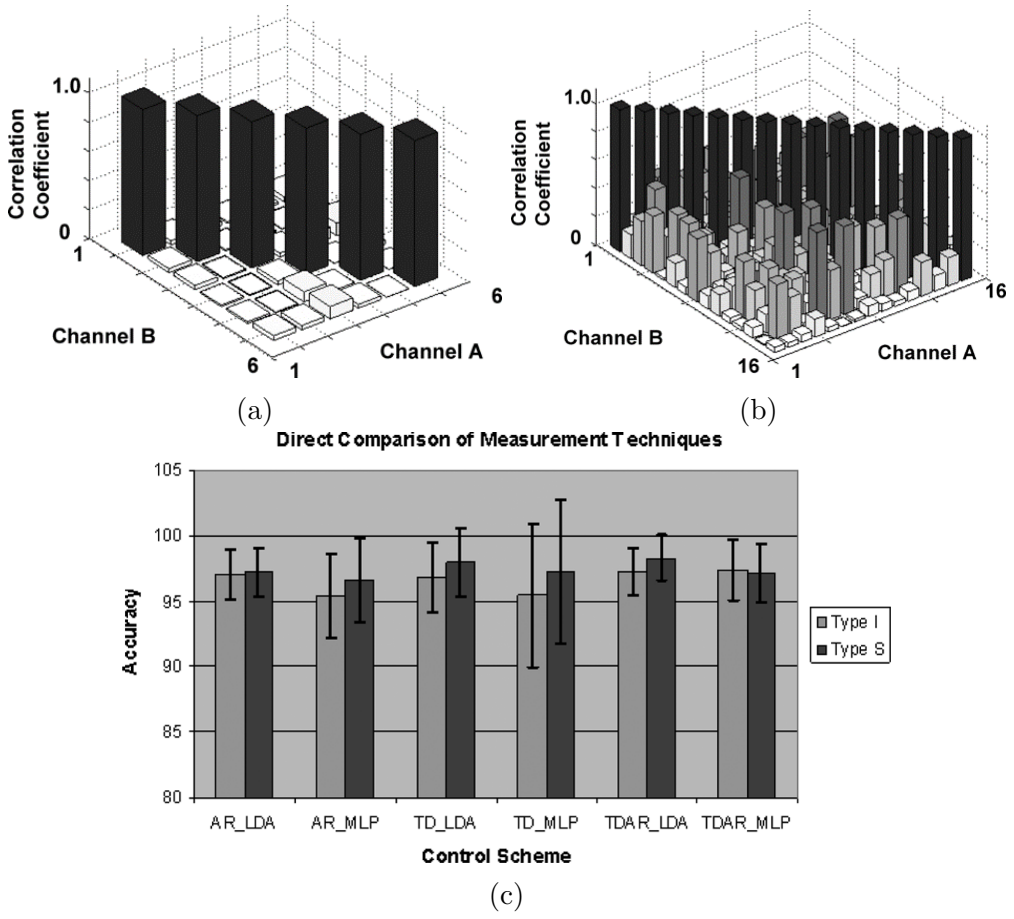
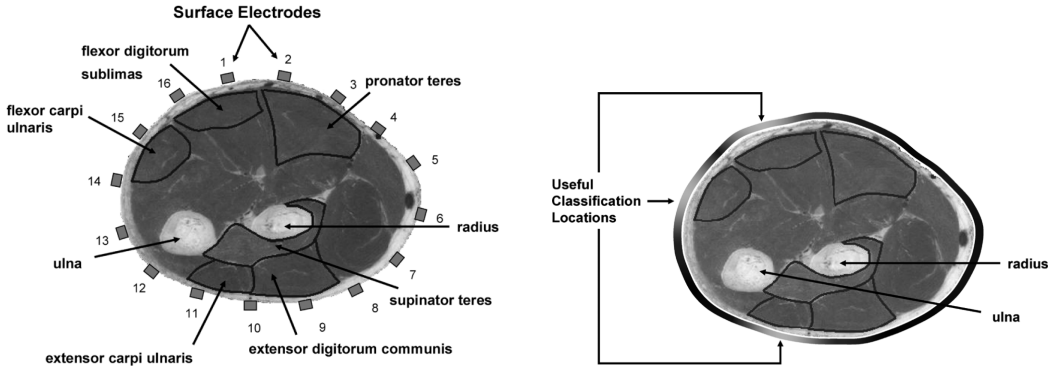


Figure 2.4: (a) iEMG crosstalk among 6 electrodes. (b) sEMG Crosstalk among 16 electrodes. (c) Classification results of Linear Discriminant Analysis (LDA) and Multilayer perceptron for iEMG (Type 1) and sEMG (Type 2). AR- Autoregressive coefficient; TD- Time Domain features; TDAR- Time Domain Autoregressive coefficients [6]



(a) (b)

Figure 2.5: Optimum number of electrodes for symmetrical arrangement and optimum place arrangement [6]

attached to the fingers. Thus, these muscles have motor information of grasping actions (finger motion). However, only four of them can be accessed from the surface of the forearm, and the other four are located at intermediate and deep muscle layers. These four surface muscles are Palmaris Longus, Extensor Digitorum, Extensor Digiti minimi, and Flexor Digitorum Superficialis. In addition to the above muscles, studies have suggested few other muscles at the forearm, such as Flexor Carpi Ulnaris and Extensor Carpi Ulnaris, are suitable for hand gesture motion classification [6, 76–79]. For the grasping pattern classification, Extensor Carpiradialis and Extensor Carpi Ulnaris are preferable because tendons of these muscles are attached to the index finger, and the little finger [76]. Despite that, some studies have suggested muscles control the wrist movements, such as Extensor Carpi Radialis Longus and Extensor Carpi Ulnaris have significant impact on the classification performance [80]. It is because specific grasping patterns have identical wrist formations. Hence, their motor information (sEMG) provides additional aid to the classifiers to further classify user intention. Few investigations have suggested shoulder muscles for the prosthetic control such as Triceps, Biceps, Anterior Deltoid, Pectoralis Major, and Infraspinatus [80, 81]. However, studies have not investigated the effect of shoulder muscles. When considering the anatomical aspects, these shoulder muscles would not contribute significantly to the classification performance.

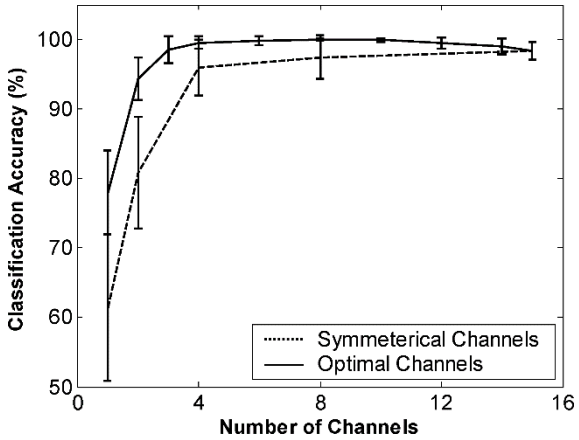


Figure 2.6: Optimum number of electrodes for symmetrical arrangement and optimum place arrangement [6]

2.3.2 Pre-processing of EMG signal

Pre-processing is a method that removes unwanted signal information and abnormalities in the data. Signal conditioning prior to the classification is important because sEMG contains numerous noises that are produced by the equipment, electromagnetic interference, and motion. Moreover, sEMG consists of Electroencephalography (ECG) signal, which is not required for the grasping pattern classification [82]. It affects the sEMG pattern adversely. These noises have to be filtered out prior to feature extraction. Otherwise, classification results may not be accurate as expected when the new data is given. It is impossible to remove noise by just performing skin preparation because electrical components (i.e., resistors, capacitors, ICs, etc.) in amplifier and electromagnetic interference of the power line affect the signal [83]. Generally, high-quality electrical components are used to develop the amplifiers and other electrical devices (i.e., silver chloride for electrode material with low impedance) so that inherent noises in the EMG acquisition system can be minimized. Other common noises are the ambient noise and the motion artifacts. Ambient noise, also known as Power-Line Interference (PLI), can be removed using 50 Hz (or 60 Hz) notch filters. However, motion artifacts such as the relative motion of electrode with respect to the skin during hand movements can induce (1-10) Hz range interference, which has to be elim-

inated by using a high pass filter [84]. Previous investigations have confirmed that useful sEMG signal frequencies are distributed within a range of (20-500) Hz [71, 72, 79, 82, 85, 86]. Hence, to extract that particular range from the raw signal, Butterworth bandpass filter is used [86]. It is important to mention that sampling frequency affects the classification performance [86]. Previous studies reported that low sampling rates decrease the classification accuracy [87, 88]. In Angkoon *et. al.* study, they proved that the 1000 Hz sampling rate produces significantly higher classification accuracy than using 200 Hz (sampling rate at wearable EMG acquisition systems) [86]. The higher sample rate has higher signal information than lower sampling rates (i.e., data points per second). Therefore, a sample rate that is higher than 200 Hz is appropriate for better classification performance [86].

2.3.3 Segmentation of EMG data

Segmentation is a process that divides the signal into equal partitions. Since EMG is an analog signal, the computer has to read it as a digital signal. The digital signal is a sequence of data points that proximate the analog signal pattern with equal time intervals. However, the individual digital signal point has no valuable meaning. Therefore the signal is divide into segments and considers all the signal points in a segment. In that case, that particular segment represents a certain range of the original signal (a portion of a signal pattern). Thus the computer can identify any patterns in the segmentation. In the classification problem, this segmentation is known as a window. This window is used to calculate different properties of the signal (i.e., features). Since the signal is a time-variant, the size of the window length is defined using time typically in *milliseconds*. This particular time period is called a window length. There are two types of windows: overlapping windows and non-overlapping (adjacent) windows. Most of the studies have suggested that an overlapping window is suitable because it minimizes the effect of abnormalities (i.e., spikes) in the signal to the classification [89]. The

Farina *et al.* have pointed out that lower window length (less than 128ms) results in drastically lower classification accuracy [90]. Englehart and Hudgins confirmed this in their study [91]. However, the overall time for the final output should be less than 300ms in order to reach the real-time response [7, 81]. Recent studies have suggested that the window length for the real-time application should be between 100 ms – 256 ms so that users can experience the lag-free motion with high-reliability [78, 81, 89]. Furthermore, the non-overlapping window has to be used for the real-time classification since the overlapping window method uses a particular portion of the previous window and the other portion from the current window. That resulted in lag since it takes three windows to cover the current window data. Hence, for real-time classification, disjoint windows are used.

2.3.4 Feature extraction

Feature engineering techniques are used in almost every machine learning fields, such as image processing, classification, and regression. The raw signal has no valuable meaning apart from the raw signal pattern. Hence, features are calculated to extract useful information. However, in feature engineering, the respective domain knowledge (analog electrical signal knowledge) is used to extract useful information from the raw signal. The different theories in the respective domain are applied to extract information from the raw signal. For example, the Root Mean Square (RMS) of a signal is widely used in sEMG classification, which represents a rectified mean amplitude of the signal (or some cases, power dissipated at a resistor). Similarly, numerous features have been introduced in different applications, but not all features are suitable for every application. Therefore, feature selection methods are used to determine the best features for the problem.

The appropriate features produce better classification performance in terms of accuracy and computation time. Generally, EMG features are extracted in the form of time-domain (TD), frequency-domain (FD), and time-frequency domain

(TFD). The time-domain signal represents the amplitude variation with respect to time [92] and the frequency-domain generally represents power with respect to the frequency spectrum.

Each of these domains has unique features which have different properties (see Table 3.1). Therefore, features must be carefully selected as the classifier can easily identify the discrepancies in classes while consuming minimum computation time. Hudgins has proposed the unique time-domain feature set, which is still used because of its compatibility with a range of classifiers [93]. They are Mean Absolute Value (MAV), Waveform length (WL), Zero Crossing (ZC), and Sign Slope Change (SSC). Since the simple feature set was used, the overall classification delay was 250ms. In addition, Willison Amplitude (WAMP) [94], Root Mean Square (RMS) [95], Autoregressive (AR) model parameters [96] and time domain-autoregression (TD-AR) were also used as features. RMS is the most used feature in pattern recognition applications because it is most appropriate for non-fatigue situations [96]. However, AR and TD-AR feature calculations take time but demonstrated good classification accuracy. Jie Liu has tested a combination of AR (six orders) and RMS, and archived 97.5% accuracy [97]. L. Hargrove et al. [6] indicated that the combination of the time-domain features and the autoregressive (AR) features had an outstanding performance in EMG pattern recognition. However, these results were based on their data set, and it may not be suitable for every problem. D. Zhang *et al.* investigated the effect of different domain features (i.e., TD, FD, TFD, Nonlinear entropy domain, and fractal domain) on classification accuracy [76]. Investigation proved that time-domain features produced the highest accuracy (97.4%) within 0.51ms, whereas entropy domain and fractal domain cost significant time for the calculations (320.1ms). Some studies have augmented the current features to new features and claim those new features outperform the conventional features [98]. The merits of the most common features are presented in Table 3.1. These empirical values give an idea about the suitability of certain features in certain applications. However, the feature selection for the interested data domain is essential to get the best

results.

The features should have good class separability, among others. The class separability can be identified and improved by performing the Principal Component Analysis (PCA) or Linear Discriminant Analysis (LDA) on features [76]. PCA and LDA reduce the dimensionality of features which improves the classification time and accuracy. However, PCA investigates the direction of maximum variance (with respect to other considered features), and LDA investigates the subspace that maximizes class separability [99]. The results of Daohui Zang's experiment and proves that LDA produces the highest class separability, which eventually resulting in the highest classification accuracy [76]. In addition to the conventional feature selection methods, few studies have investigate the selection power of *Sequential Forward Selection (SFS)* and *Particle Swarm Optimization (PSO)* methods [98]. The results indicated that the features selected by the PSO produce the highest classification accuracy. An extensive investigation has been conducted to evaluate novel edition of PSO (personal best (Pbest) Guide Binary Particle Swarm Optimization PBPSO), Binary Particle Swarm Optimization (BPSO), Genetic Algorithm (GA), Modified Binary Tree Growth Algorithm (MBTGA), and Binary Differential Evolution (BDE) [100]. Their results indicated that the proposed PBPSO could select the best lowest number of features. It selected 14 features out of 480 on average (for each participant), and the average classification accuracy was 88.83%. All the other methods selected above 200 features for the classification. However, this is very effective when the number of features is significantly high. When the number of features is lower, then the greedy method or random selection is better, since they evaluate almost entire feature combinations.

Table 2.1: Merits of three major feature domains [12]

Featuer Domain	**Feature types**	**Advantages**	**Disadvantages**
Time Domain Feature	iEMG; MAV; MMAV; RMS; SSI; VAR, LD; WL; ZC; Average Amplitude Change (AAC); WAMP; SSC; MAVS; MHW; AR; CC; KUR; SKW; ASS	Simple to Compute; Quick; More often used in force- EMG relations	Noise sensitive
Frequency Domain Features	MNF; MdF; Power Spectrum (PS); Mean Power (MP); SM[k]; Frequency Ratio (FR); PSD; Peak Frequency (PF)	More often used in fatigue estimation	Computationally Complex; Information loss due to spectral Leakage
Time-Frequency Features	Wavelet transform; short-time Fourier transform; "Cohen Class", S-transform and BD	Regarded best for information extraction; High dimensional output	Computationally expensive

2.3.5 Classification methods

The classification is the critical part of the grasping pattern prediction system. In the classification process, the classifier decides the output class (category) of the given inputs. In order to perform the decision process, the classifier algorithm has to be mathematically developed (a mathematical model) to comply with the data. This is known as *Training* the classifier with known inputs and outputs (label data). In the training process, the classifier tunes the parameters of the model so that it can map given inputs to the corresponding outputs. If the model map the data properly, it can determine the actual class (output) of new data precisely. In EMG classification, the model is trained with the selected features of the signals and the corresponding grasping pattern. This review primarily focuses on the grasping pattern with hand gesture classification systems.

In 1998, Sang-Hui proposed a method to recognize the patterns in EMG signal using Dempster-Shafer theory and Fuzzy mapping function [94]. The results have

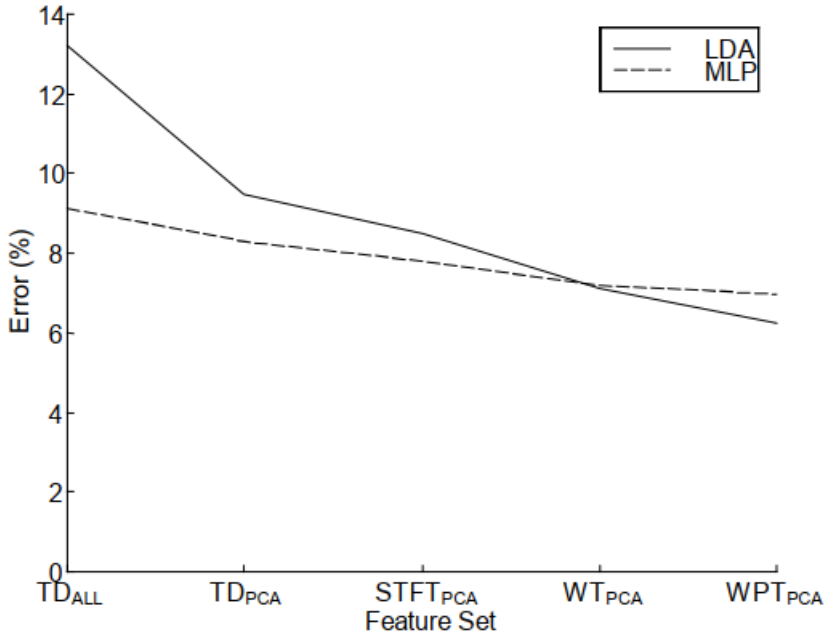


Figure 2.7: Classification error of MLP and LDA with respect to different feature set [7].

suggested that multiple features produce the highest accuracy over previous methods because Dempster-Shafer theory is more suitable for applications with high uncertainties such as EMG signals classification. Englehart K. *et. al.* investigated the performances of classification of EMG signals, between Linear Discriminant Analysis (LDA) and Multi-layer Perceptron (MLP) (see Fig. 2.7) [7]. The investigation has indicated that the LDA performance is increased with time-frequency domain feature set (WPT) while MLP maintains an average 9% error for all feature sets (TD and FD). Another important finding was that PCA dimensionality reduction has a significant impact on LDA classification. In addition to them, random forest classifier [101] and minimum distance classifier [76] were proposed for classifying hand gestures. Although they resulted in considerably high accuracy 87% and 97.7% respectively, it shows poor performance with high variance of data and high-class separation.

The selection of a classifier is important as the selection of features. The most suitable classifier has to be selected according to the application parameters. In 2007, Levi J. *et. al.* investigates the performances of different classifiers with AR features [6]. Results indicate no significant differences in accuracies of different classifiers (MLP, LDA, Gaussian Mixture Model Hidden Markov Model). Hence, for further studies, they implemented the LDA MLP (backpropagation algorithm) classifiers. Many recent studies adopted LDA because it has shown greater accuracy and simple implementation [79, 102–105]. K-nearest neighbor (KNN) is another widely used statistical classifier for pattern recognition [106]. Christodoulou *et al.* has suggested that KNN is not suitable for applications with higher variations in features, though SVM and Self-organizing Map (SOP) classify accurately [107]. However, a recent study has achieved 88.8% classification accuracy using KNN [100]. Christodoulou *et al.* studied the neuromuscular disorders with Amplitude-modulation and Frequency-modulation features, whereas Jingwei *et al.* studied the hand movement with conventional signal features. Their results have indicated that the performance of the classifier depends on the data as well as the features. Hence, multiple classifiers should be evaluated to

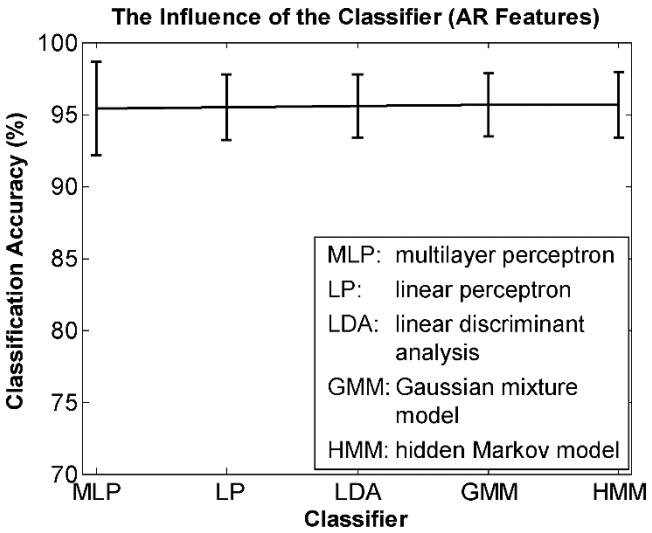


Figure 2.8: Classification result of different classifiers for AR feature [6]

determine the most suitable classifiers for the data domain.

Artificial Neural Network for pattern classification is mainly used for non-linear modeling of EMG signals. The advantage of ANN is that it can be trained for a specific task which can use to identify the linear and non-linear patterns in EMG signals [105,108]. For the real-time classification, several works of literature have suggested that Fuzzy Clustering Neural Network (FCNN) provides significantly low error while less computation time [109]. Karlik *et al.* compared the performances of different classifiers such as MLP, Conic Section Function Neural Network (CSFNN), FCNN-6, and FCNN-2, in order to identify the optimum classifier for their application. Results have proved that FCNN-2 produces the highest classification accuracy, which was 98.3%.

In 2005, Chan and Englehart proposed a hidden Markov model (HMM) based pattern recognition method for prosthesis control [110]. Hence the model follows a probabilistic approach to pattern recognition, and it is resilient to a continuous feed of myosignal variations. In this investigation, HMM method showed higher accuracy than MLP as well as the response time. Xiaolong *et. al.* proposed a self-recalibrating mechanism for MLP so that it can adapt (retrain the network) to a new data set [111]. Since it takes some time to collect data for retraining, the

system is not suitable for real-time prosthetic control. However, Riilloa *et al.* have investigated the performance of LDA, Support Vector Machine (SVM), and ANN against different time windows, and the study concluded that ANN obtained the best real-time classification results [112]. According to the [108], [110] and [112], MLP (or ANN) is better than conventional classifiers (i.e. SVM, LDA, KNN) but, perform less than probabilistic classifiers (i.e. FCNN, HMM). The reason for that is, SVM and KNN classifiers' performances depend on the kernel characteristics, and they exhibit a non-parametric nature [113]. In addition to the ANN, Deep Neural Networks (DNN) such as Convolutional neural network (CNN) has been used for EMG classification [114–116]. The CNN is suitable when the feature space is large. Hence, it can be considered as the grid of features and perform pool calculations. In most cases, CNN outperforms the other conventional classifies such as SVM, LDA, and KNN but not significant.

Table 2.2: Summary of offline and real-time grasping pattern classification. Here only mention the highest performed classifiers only. [HG- Hand Gesture; HC-Hand Close; HO-Hand Open; R-rest; IFP-Index Finger Pinch; MFP-Middle Finger Pinch; KG-Key Grip; CG-Chuck Grip; LG-Lateral Grip; PG-Power Grip; PH-Point; PD-Precision Disk; P2F- Prismatic-2 Finger; P4F- Prismatic-4 Finger, UP-ulnar Pinch] [ECOC–Error Correcting Output Codes; NB-Naive Bayes; ESN-Echo state network]

Reference	**Classifier**	**Patterns (grasping + hand gestures)**	**Accuracy (%)**
[76]	PCA+LDA	HC, HO, **IFP, MFP** and 5-HG	97.4
[6]	LDA, MLP	**KG, CG,** HO, R and 6-HG	95-99
[117]	LDA	**LG, PG, IFP** and 9-HG	90-95
[118]	ANN	**FG, PG,** HC	76
[119]	NB	R, HC, HO and 2-HG	98
[80]	ESN	**PD, CG, P2F, P4F, UP**	93
[111]	CNN	HO, **PG, KG, CG, IFG, LG** and 4-HG	74[2]
[120]	ECOC	**LG, KG, IFP**, HO	98[1]
[121]	LDA	HO, HC, **FG, KG, IFP** and 2-HG	67.3[1]
[78]	SVM	**KG, PG, PD, IFP, P2F, P4F**	80 (**PG, P2F**)[1]
[8]	SVM	HC, individual finger movement (9)	90[1]
[122]	LDA	HC, HO, **PD, PH, IFP** and 6 HG	90[1]
[21]	LDA	**CG, KG, PG, IFP, P2F**	69.4[1]

[1] Real-Time classification accuracy

[2] High computational time

In this research work, five supervised learning classifiers were selected for the investigation. As literature has suggested, the performance of the proposed system should be evaluated for several classifiers in order to determine the best classification system for the application. Therefore, the ANN, LDA, KNN, NB, and DT were developed and optimized for real-time classification.

2.3.6 Post-processing for final prediction

Results from the classifier may not be accurate as expected. The classification results may fluctuate among classes. Therefore, post-processing methods have been proposed to overcome the limitations in the classifier. In the post-processing stages, errors or spurious misclassifications resulted during the prediction, are corrected by applying post-process techniques. To minimize this classification error, Simon *et al.* [123] introduced the practice of decision-based velocity ramp functions as a post-processing method. It reduces the speed of action when the classifier decision is changed. That helps to smooth the transition motion from current action to new action. Huang *et al.* proposed Majority Vote (MV) tech-

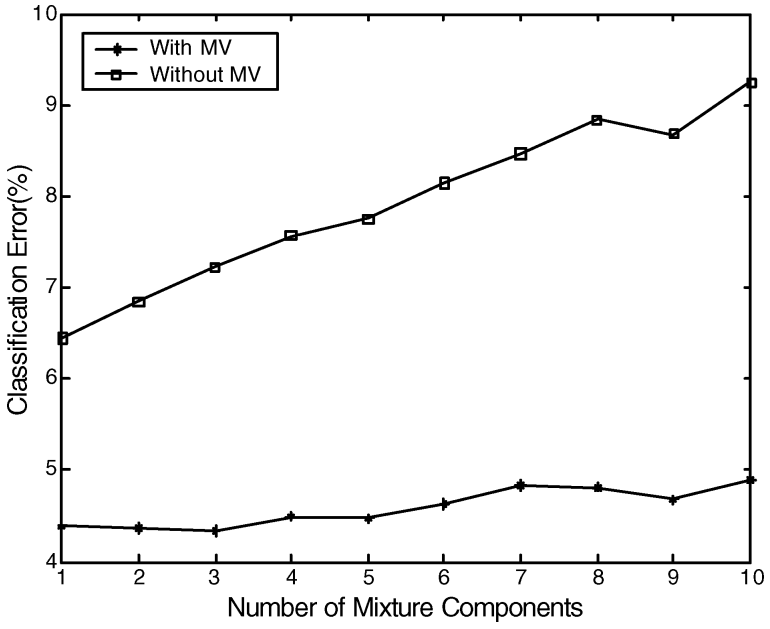


Figure 2.9: Effect of majority vote of classification results [6]

nique to reduce the classification error (see Fig. 2.9) [124]. This technique uses certain previous outputs and current output to smooth the output or correct the output [7, 80, 125]. Similar to the MV technique, [111] proposed a re-calibration mechanism that uses previously predicted labels (predictions errors) to retrain the network. The Bayesian Fusion (BF) is another commonly used post-processing technique to minimize the fluctuations in the output [8]. Similar to the MV, Bf uses previous outputs to correct current output. [8] has investigated the performance of both MV and BF and has shown that the on average, BF produces the accurate results than MV. However, the difference is not statistically significant. The similar study has been conducted by Li H. *et al.* and prove the [8] conclusion [126]. The implementation and computation cost is less MV and BF than others. Therefore, MV and BF were used in this research and compare the performance of both to determine the best post-processing method for this application.

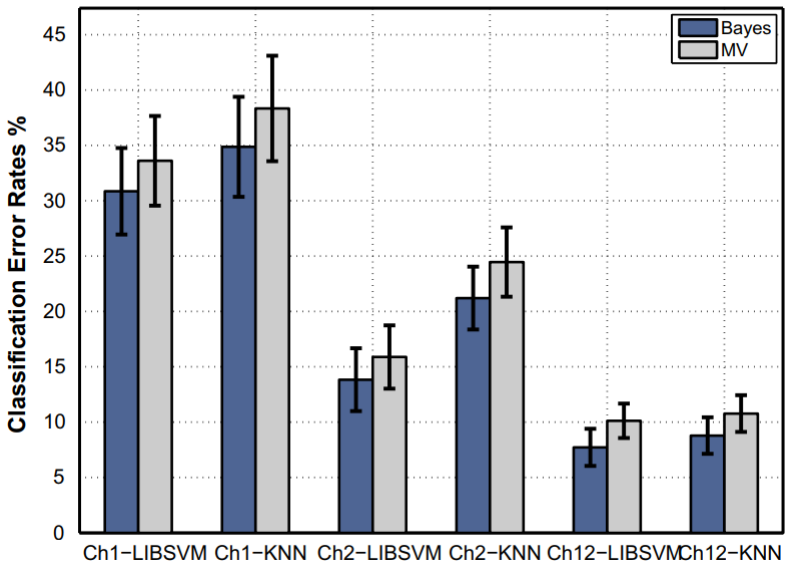


Figure 2.10: Comparison of classification error rates of KNN and SVM when MV and BF are used for post-processing [8]

2.4 Vision-based prosthetic hand control systems

The human visuomotor function inspires Vision-based grasping pattern classification. As mentions in the introduction, humans use visual feedback to determine the grasping pattern and navigate the hand towards the object. A similar approach was used recently to determine the grasping pattern in both industrial applications and the prosthetic hand field. Object detection is primarily achieved by employing a specific CNN model. There are highly advance real-time object detection algorithms available at the moment, such as You Only Look Once (YOLO) [10], Faster Region-based Convolutional Neural Networks (Faster R-CNN) [127], Deconvolutional Single Shot Detector (DSSD) [128], and Deeply Supervised Object Detectors (DSOD) [129]. Most of these models are pre-trained for more than 50 domestic objects and available freely on the official websites. The [130] has evaluated several object detection algorithms, including the models mentioned above. The results indicated that the YOLO model exhibited a lower computation time (0.03 seconds/image) while achieving 78.6 mean Average Precision (mAP(%)). The highest mAP was achieved by PVANET+, which is 83.8%, and it takes 0.05 seconds for an image. Hence it takes significant memory, it is not suitable for embedded systems. Since YOLO has shown significant accuracy (mAP) and a higher frame rate (45 FPS), it suitable to implement on a prosthetic hand. The implementation of the YOLO in this research work is discussed in Section 4.2 and a brief description of the algorithm is given in Appendix J. Furthermore, integrating a high-speed and high-resolution camera on the prosthetic hand is possible since most modern smartphones have high-quality visual sensors. Because of these reasons, the vision was used for prosthetic control systems as an external sensor. The following paragraphs discuss the merits of previous vision-based grasping pattern classification systems and the different approaches in vision-based classification.

The main objective of using a camera for grasping is to identify the object and

its properties, such as size and shape. Therefore, researchers have investigated the vision-based controlling method for prostheses because these are achievable objectives. The vision-based controlling method was initially studied to investigate the behavior of a robot hand during object manipulation [131]. Ramnath proposed that a camera mounted on the hand provides more significant advantages than mounted externally [132]. The main advantage is that this does not need camera calibration for the manipulator coordinate system since both systems are in the same reference plane. However, some drawbacks are incorporated with the eye-in-hand method, such as lighting condition changes and unable to measure depth. These can be overcome by implementing external light sources and depth sensors such as sonar [30]. Despite this, [29] suggested that multiple cameras can be employed to overcome depth measuring difficulty and object identification, although the system has to be calibrated. However, stereovision object detection and depth measuring is a computationally expensive process. Therefore, practical implementation of the stereovision raises problems in addition to grasping pattern classification issues. The incorporation of an additional sensor to measure the distance, such as sonar or sharp sensor [30], can improve the performance of the system by providing a facility to measure the object dimensions [31, 32]. Strbac *et. al.* have used a stereo camera with a laser beam to measure the depth of the selected object, which the user chooses by pointing the laser dot onto the object [29]. Beneficial of object dimension measuring is that system has the opportunity to determine the precise grasping type as well as finger and wrist configuration based on the object dimensions [31, 32]. This approach allows the system to determine the grasping pattern based on defined rules. The number of grasping patterns depends on how decision rules are developed and the depth of the decision trees. The major drawback in this dimension-based approach is that the computation time and the low precision in distance calculation [32]. The distance to the object is used to determine the object dimensions, and if errors occur during the measuring, it affects the grasping decisions as well. Therefore, researches were focused on object detection-based pattern classification methods.

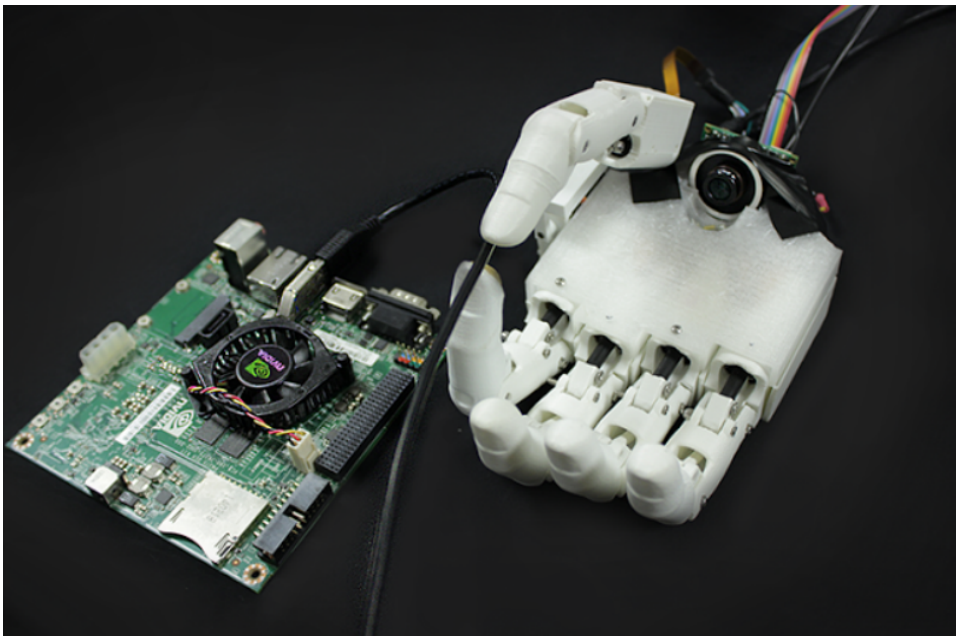


Figure 2.11: The propose Eye-in-Hand method for vision-based grasping pattern classification system by Joseph *et al.* [9].

Object identification is the primary objective in vision-based prostheses. This has been accomplished in different methods such as color segmentation [131], laser point approach [29–31], CNN classification [9,34] and edge detection [133]. Once the object is identified, properties such as orientation, dimension, and center of gravity (CoG) can be estimated by using image process techniques such as particle analysis (for the axis of the object, orientation) [32, 133] and moments of image (for CoG) [134]. Then cognitive vision system (CVS) can use this information to decide the grasping pattern (from a list/dictionary), aperture size, and finger pose, and ultimately control the hand appropriately [9]. These systems were developed to identify the grasping type entirely based on image information [9, 29, 30, 32]. The main drawback of using the dictionary approach is that list has to be developed for each object. Besides, the number of grasping patterns assigned for each object was limited to 1. When multiple grasping patterns are assigned, the user has to remember the order of each object and select it by other means (using sEMG triggering). This approach was used investigated by several teams, and they employed EMG or MMG to trig the action or select the pattern from a list/dictionary suggested by the vision system [31,34,35,135]. The

bio-signal was used only for the selection. The [35] proposed a method to teach the system by keeping the image of the object and corresponding grasp type, so then the system will use this information in the future for the decision. However, it needs a large memory space, and retrieving image data from the database is time-consuming. Chunyuan *et al.* have proposed a triggering scheme (wrist flexion and extension) to select grasping pattern [36]. They further proposed a method to use RGB-depth images to determine the grasping pattern. A List of objects with corresponding grasping patterns is feed along with the depth images to the CNN model. Once the model detects an object, it outputs respective grasping patterns. A fusion method was used to combine vision results and the EMG triggering scheme results to produce the final output. This method gives the user to freedom to select the multiple grasping patterns for a single object. However, the user has to remember the triggering scheme.

The recent studies have introduced novel frameworks to determine the targeted object when other objects surround it [135, 136]. This new approach facilitates the decision system to determine the grasping pattern (from database) or hand orientation without any external sensor. This approach is practically important since, in real-life, the object of interest is surrounded by other objects. Hence system should be able to identify the target object correctly to execute efficient grasping action. However, this research employed only the object detection techniques and the proposed dictionary approach to classify grasping patterns. Since EMG classification is also simultaneously execute, the hybrid system requires light-weight algorithms to produce the final prediction within 500ms.

Table 2.3: summary of Vision-based Prostheses. (OD-Object Detection; GP-Grasp Prediction; TA-Triggering action; FC-Finger Configuration; WC- Writs Configuration; P-Pinch; PG-Power Grip; TG-Tool Grip; KG-Key Grip; 2JC- 2Jaw chuck; 3JC- 3 Jaw chuck; LG-Lateral Grip; HG-Hook Grip; PT-Point Grip) (DA- Dictionary Approach; ODA- Object Dimension and Area)

Reference	**Sensor**	**Objective**	**Grasp Types**	**Method**	**Accuracy (%)**
[34]	Camera, Myo band	OD, GP, TA (EMG)	12 grasping types	DA	-
[133]	Camera, MMG sensor	OD, GP, TA (MMG)	P, PG	ODA	84.4
[9]	Camera	OD, GP	P, PG, 3JC, TG, KG	DA	93.2
[31]	Sonar, Camera, Laser point, EMG sensor	OD, GP, FC, TA (EMG)	KG, 2JC, 3JC, LG	ODA	84
[30]	Sonar, Camera, Laser point, Accelerometer	OD, GP	P, KG, PG	ODA	-
[29]	Camera (stereo), Laser point	OD, GP	P, KG, PG	ODA	90
[32]	Camera, Sonar	OD, GP, WC	P, KG, PG	ODA	-
[35]	Camera, Laser point, LCD display, Myo band	OD, GP, TA	P, HG, KG, PG	DA	89.60
[36]	RGB-Depth Camera(external), EMG	OD, GP, TA	P, HG, 3JC, PT, PG, LG, KG	DA	89.8

Chapter 3

PRELIMINARY STUDY ON CLASSIFIER DEVELOPMENT

This chapter discusses details of the study that was carried out to investigate the behavior of the hand and the sEMG signal during Reach-To-Grasp (RTG) motion. In addition to that, sEMG signals of six muscles when executing six grasping patterns were collected for classifier development. The identified relationships and the key findings are discussed at the end of the section.

3.1 EMG and onset distance data collection

When grasping an object, the upper limb performs a series of motions at the shoulder, elbow, and hand to reach that object and grasp it. It is known as Reach-To-Grasp motion. As mentioned in the Chapter 2, developing a classifier to predict grasping patterns when the hand is in motion rather than in a static posture is a more practical approach for real-time classification [137]. Therefore, further studies have to be conducted to investigate the RTG motion. Since the previous RTG investigations focused on hand speed and the hand aperture relationship with the kinematics of other limbs, this experiment was designed to investigate the sEMG signal (Root Mean Square (RMS) of EMG signal) change with a person's demographic factor such as Age, Gender, and the Handedness of the participant. A similar procedure was followed to examine the variation of hand onset distance of each grasping pattern. However, the previous studies have

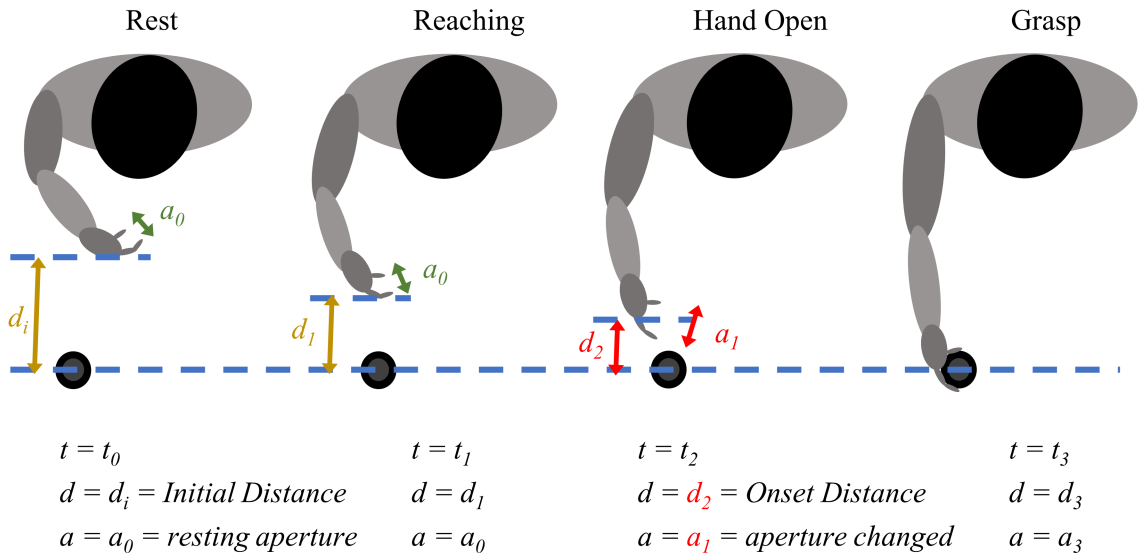


Figure 3.1: Overview of hand onset experiment during the RTG motion.

not discussed the effect of Age, Gender, and the Handedness of the participant (or the amputee) on the RTG motion and the grasping patterns. Therefore, this experiment was designed to examine this gap and collect EMG data to develop an EMG classifier. In addition to the experiment details, descriptive and inferential statistics of the collected data are discussed. The aim of the experiment was to identify the variation of hand onset distance and sEMG signal with selected grasping patterns. The hand onset distance is the distance from the object to the hand when the hand initiates the grasping formation (from resting posture to the grasping pattern posture). Figure 3.1 illustrates the RTG motion and the hand onset. Here, the d_i and a are the distance to the hand from the object and hand aperture size. The t indicates the time instance at each moment. The d_2 and a_1 are the hand onset distance and the hand aperture size at that instance. The sEMG of the RTG motion was used for EMG analysis and the classifier development. In addition to that, the effect of demographic factors such as Age, Gender, and Handedness on the grasping pattern was investigated. A Systematic research method was followed to conduct the experiment. The sEMG signals of six selected muscles were collected from 15 healthy subjects to analyze and develop an EMG classification system. The following sections describe the selection criteria

of the muscles, experiment protocol, analysis methods, and the key findings of the experiments.

3.1.1 Muscle selection

As discussed in Section 2.3.1, six muscles were selected for the study. Among the six muscles, four of them are superficial muscles (surface muscles) whose tendons are attached to the fingers (Phalanges) and palm (Metacarpals). The other two muscles from the superficial layer of the forearm muscles, though their primary responsibility is to performed wrist motions. However, according to the [78], EMG information of these wrist motion muscles has a significant impact on the EMG classification. Moreover, as suggested by the [6, 21] in their investigations, the optimum electrode count (number of muscles) should be within 4-6. Therefore, the number of muscles was limited to six, and Figure 3.2 illustrates the selected muscle anatomy on the forearm. Table 3.1 shows the anatomical information of selected muscles.

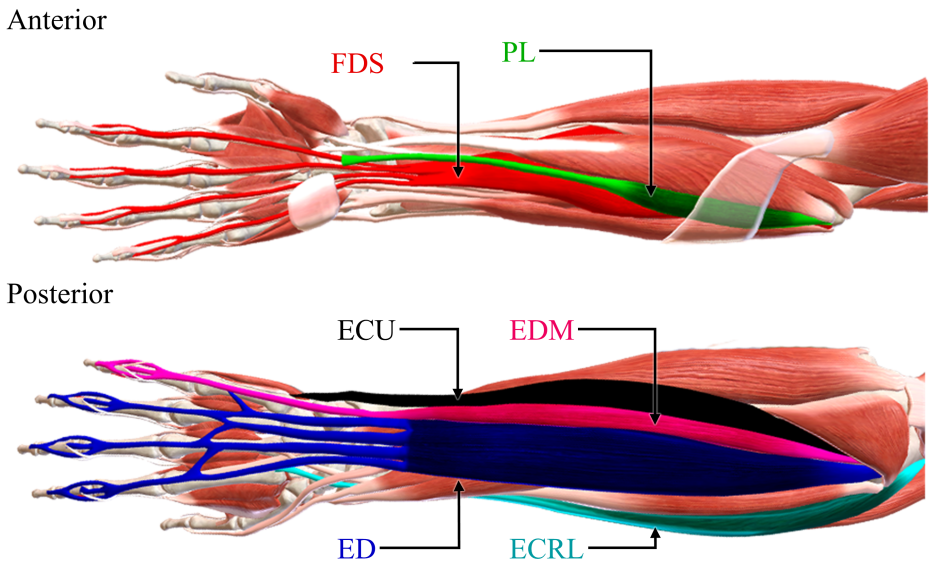


Figure 3.2: Anatomy of the selected muscle at forearm. FDS - Flexor Digitorum Superficialis; PL- Palmaris Longus; ECU - Extensor Carpi Ulnaris; EDM - Extensor Digiti minimi; ED - Extensor Digitorum; ECRL - Extensor Carpi Radialis Longus. (image is adapted from BioDigital.com)

Table 3.1: Selected Muscle for sEMG Extraction

Muscle	**Compartment**	**Attached region**	**Function**
PL	Anterior	Proximal phalages of all fingers except thumb	Flexes writ joint & assit gripping force
ED	Posterior	Middle & Distal phalages of all fingers except thumb	Extends fingers
EDM	Posterior	Middle phalages of little finger	Extends little finger
FDS	Anterior	Middle phalages of all fingers except thumb	Flexes fingers
ECRL	Posterior	Dosal surface of MC[1]II	Extends & abducts the wrist
ECU	Posterior	Medial side of MC[1]V	Extends & abducts the wrist

[1] Metacarpals

3.1.2 Sample size selection

The power analysis is used to estimate the sample size (number of participants) for given for *standard deviation (SD)* of the sample and expected difference *(ED)* between two samples. The *ED* represent the difference of the considered parameter (distance or the RMS value) of two grasping patterns that this experiment expects to observe. A pilot test was conducted to estimate the *SD* of each grasping pattern. Since the experiment was designed to investigate two aspects of the RTG motion (onset distance and RMS), the individual sample size was estimated for 80% power, and the largest sample size was selected [138].

A pilot test was carried out for five participants for a single trial and single attempt. A *SDs* were calculated for EMG experiment (i.e. RMS study) and distance experiment (i.e. Onset study). Since the experiment considered the difference between two samples (2-sample t test), the *pooled standard deviation* had to be calculated. The equation 3.1 gives the *pooled SD* of two *SD*s of two patterns (5 participants). The *pooled SD* was calculated for every pair of grasping patterns, and the highest *pooled SD* was selected to estimate sample size. The

reason for that is that the highest *pooled SD* required the largest sample size. The expected difference of the mean was defined as $RMS_{difference} = 0.03$ and the $NOD_{differece} = 25\%$ as the difference of every grasping patterns pairs (in both experiment) exhibited the maximum difference of $RMS_{difference} = 0.035$ and the $NOD_{differece} = 26.094\%$. The sample size was calculated for each grasping pattern in each experiment (for 95% Confidence Interval). The results suggested that at least 14 participants should be required for the RMS study and 13 participants for the onset distance study. Therefore, 15 participants were selected for the experiment. The details of the sample size calculation are reported in Appendix B.

$$SD_{pooled} = \sqrt{\frac{(SD_1^2 + SD_2^2)}{2}} \tag{3.1}$$

Participants

The number of participants (or the sample size) for the quantitative research on RTG motion study is typically between 6-15 [38, 78, 80, 139, 140]. The power analysis of the pilot test indicated that the number of participants (sample size) should be at least 14 to have 80% statistical power for the given expected difference. Therefore, the 15 participants were selected randomly but ensured that the *Palmaris Longus* muscle is present (or the tendon is visible from the surface). Out of 15 participants, 9 of them were males, and the rest 6 participants were females. The participants were divided into 3 Age categories for less complicated analysis. The following Table 3.2 presents the summary of participants' details.

Table 3.2: Details of the participants.

Category Type	**Age group (yrs)**	**Male/Female**
1	20 – 25	3/3
2	26 – 30	4/1
3	31 – 35	2/2

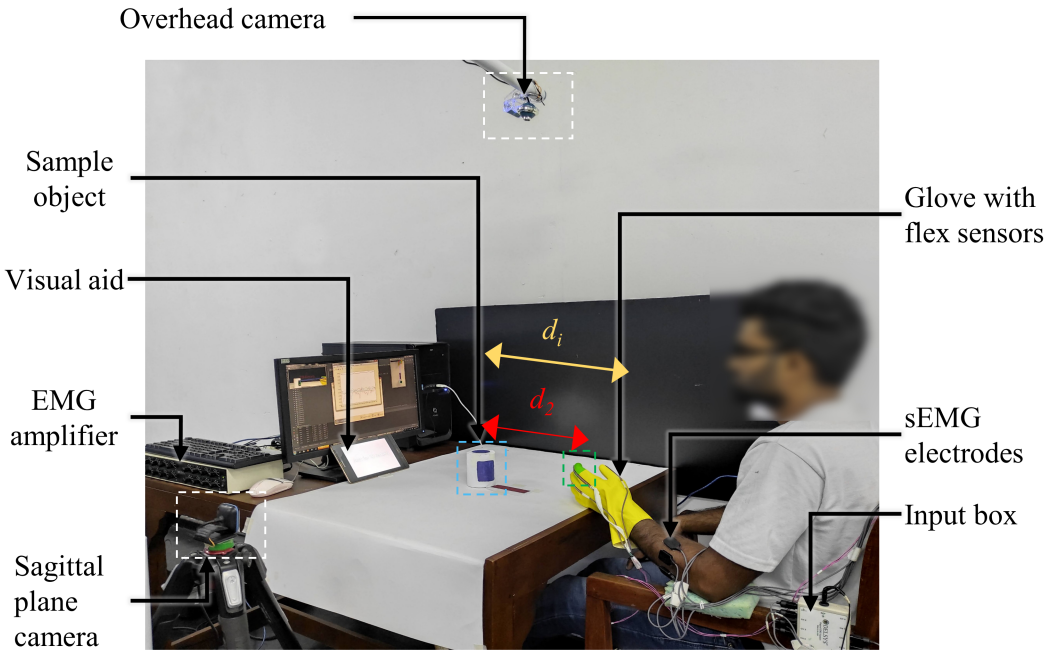


Figure 3.3: The flex sensor with glove is marked with green color and the object (cylindrical, lateral power grasp) is marked with blue color for the distance calculation. d_i is the initial distance and d_2 is the hand onset distance.

3.1.3 Experimental setup

In order to identify the hand onset distance, cameras and flex sensors were used to measure and detect the hand position and motions. Two cameras at perpendicular plane were placed to have uninterrupted distance calculation (see Figure 3.3). An image processing-based contact-less distance calculation method was used to measure the distance to the hand from the object. One camera was placed at the transverse plane (overhead position, at 103cm above the table), whereas the second camera was mounted at a plane parallel to the sagittal plane of the participants (left side of the subject, 83cm away from the object). It facilitated a comfortable environment for the participants and promoted the natural grasping motions of the participants.

A python script carried out the contact-less distance measurement. The continuous video feeds from both cameras (30 frames per second) were fed to the python program to detect specific colors and measure the distance in real-time.

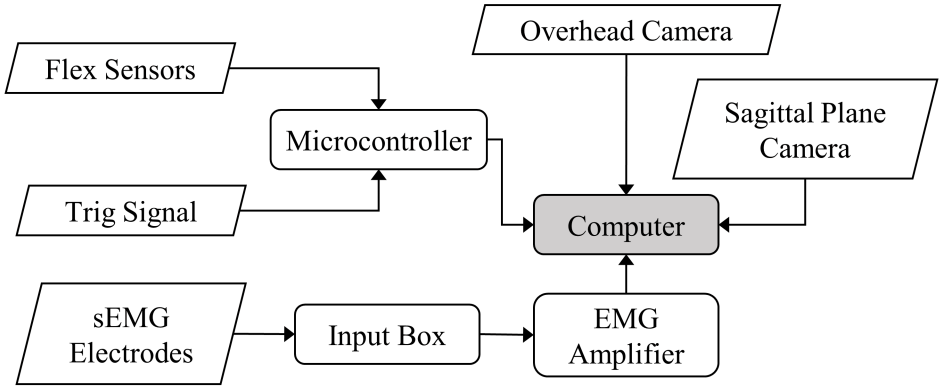


Figure 3.4: Schematic diagram of the experiment setup.

The corresponding object and the glove were marked with blue and green colors, respectively. The algorithm was developed to detect any color as the different lighting conditions change the Hue and Saturation of color, making it difficult to detect. The algorithm captures the image from the video feed and converts the color space from RGB (Red Green Blue) to HSV (Hue Saturation Value). Thus, it facilitated the program to detect the entire dynamic range (light and dark areas) of a color. The corresponding HSV values were obtained using another algorithm, and these values can be calculated at any time of the experiment. Furthermore, the experiment program facilitates storing the video feeds separately, pause the program when needed, and the ability to select Region of Interest (ROI) of each video feed in order to reduce the computation burden and to avoid false detection.

The hand onset was identified using the flex sensors attached to a glove (see Figure 3.5). The resistance values of the flex sensors are read by an Arduino microcontroller (ATmega328, ATmel) and stored in the computer for post-analysis. A recent study reported that most of the grasping patterns, including the pattern suggested in this research, the Thumb, Index finger, and Middle finger are essential to forming the grasping pattern [139]. Therefore, only these three fingers were selected to reduce the complexity of the experiment and to maintain the quality of measuring the hand aperture (onset). The sEMG signals from 6 mus-

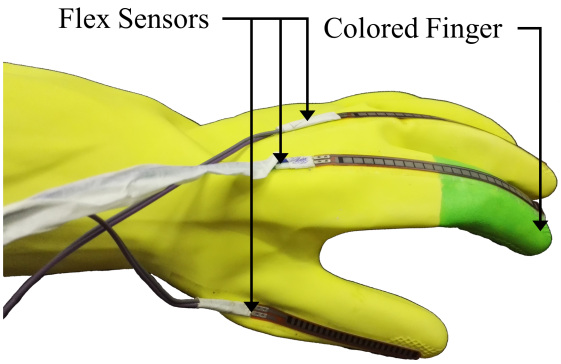


Figure 3.5: Three flex sensors were connected to the Arduino Uno microcontroller. The green color mark is identified by the distance calculation python program.

cles were simultaneously recorded using the Delsys Bagnoli-16 EMG acquisition system. The sEMG machine was interfaced with a personal computer through a data acquisition card (PCI-6220, National Instruments). The sample rate was set to 4000 Hz, and the amplification gain was set to 1k. All the data (sEMG, distance, and resistance values) was recorded simultaneously with a timestamp for the post process.

3.1.4 Experimental protocol

The participants were given verbal and visual instructions regarding the objectives of the experiment and the procedure of the experiment. Then the electrodes were placed on the hand according to the guidelines published by Hermie J. *at el.* [74]. Before the experiment, participants were informed to carry out test runs to familiarize themselves with the grasping pattern and the timing. A visual guide was given to follow the time plan (see Figure 3.3). The participant should perform two trials for each grasping pattern, and a single trial contains 5 attempts. Hence the single grasping pattern has 10 observations (attempts or samples). Thus a single person performed 60 attempts for all the grasping patterns. The resting time between two attempts was 13 seconds, whereas the 60 seconds for two trials. The resting time was extended at the request of participants. The

grasping patterns were determined using the statistics provided by Zheng *et al.* in their investigation [141]. The *Disk, Key Lateral Power, Pinch, Point* and *Power* grasps were the 6 most frequently used patterns in a domestic environment. The hand formation and the corresponding object are depicted in the following figure.

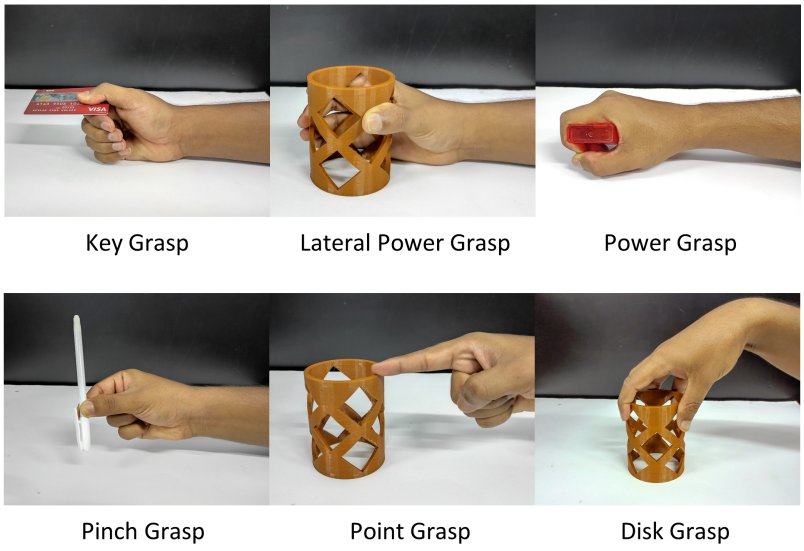


Figure 3.6: The selected grasping patterns and corresponding object used for the experiment.

The real objects were used to promote the natural grasping motions. The reaching and grasping time was set as 5 seconds. The reaching and retracting time was approximately 1.5 seconds, and the grasping time was set as 2 seconds. The reaching and retracting speeds had to restrict since the distance calculation was achieved from the image process. The lower speeds produce precise distance calculations. The cameras were calibrated every time before the experiment starts to minimize the errors. A red rectangle (see Figure 3.3) on the table was used to calibrate the top camera, and a similar red rectangle was used to calibrate the side camera. The object was placed at approximately $300mm$ distance from the initial position (d_i). The participants were instructed to maintain a constant initial distance throughout the experiment and the same resting hand posture (to detect hand onset). Although a variation of the d_i was observed ($mean = 303.62 \pm 11.42mm$). Hence the normalized distance was considered, and the details are discussed in Section 3.1.5. At the end of each trial, collected data was

verified manually by plotting the data appropriately (sEMG signal, distance, and resistance).

3.1.5 EMG and distance data pre-processing

The Root Mean Square (RMS) of a signal is the most widely used feature to represent the signal since it loosely represents the signal power. Hence, the RMS was used to represent the muscle activities of the 6 muscles. The principal component analysis, PCA (deeply discussed in Section 4.1.2) of the muscles indicated that 4^{th} muscle (Flexor Digitorum Superficialis) accounts for the highest variation of both first 2 principal components. Therefore, the data from 4^{th} channel (i.e., Flexor Digitorum Superficialis) was used for the RMS study (sEMG signal study). The additional details of the PCA are provided in Appendix A.

The RMS of each grasping attempt was calculated, as mentioned in Section 4.1.3 however, the window length was considered as the entire 2 seconds of the grasping attempt. The signal filtering and the other pre-processing were carried out as mentioned in Section 4.1.1. The hand onset was detected manually. Figure 3.7 illustrates the hand distance variation and the resistance variation of the 3 flex sensors at the fingers during a single trial (5 attempts). The y-axis represents both *Ohms*, Ω and the distance in *millimeters*. When the flex sensors bend with the finger flexion and extension (motion), the resistance is changed accordingly. The decrease in sensor value indicates the finger extension. This phenomenon was used to identify the finger onset (hand onset).

The time instance of leading finger extension was used to identify the hand onset distance on the distance graphs. The distance value at that time instance is considered as the Onset distance. The initial hand (resting) distance (horizontal line at the top of the distance graph) was extracted from the graph for normalization. The initial values and the onset values were extracted and stored for normalization. The normalized values were used to determine the sample size

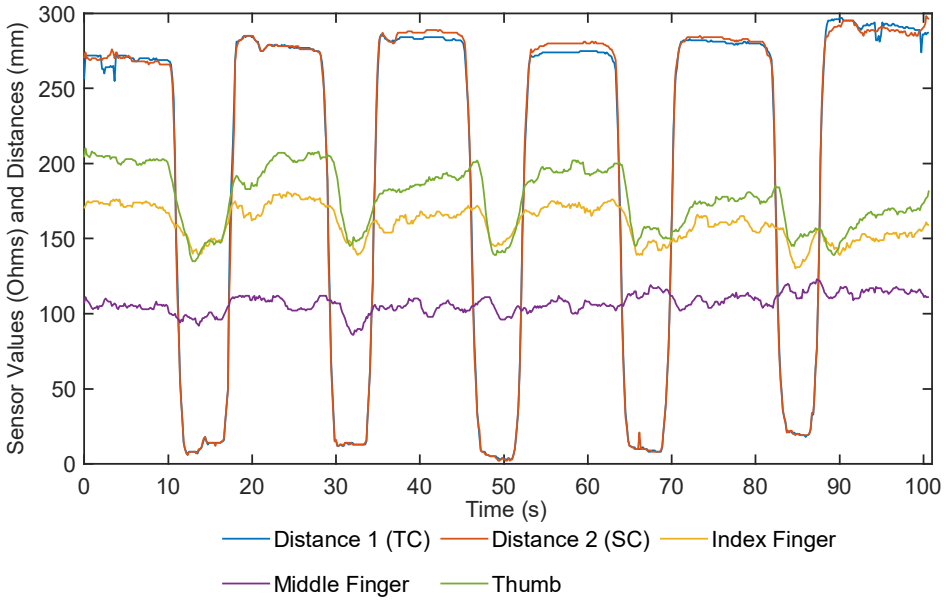


Figure 3.7: The finger motion variation and the hand distance (d_i in the Figure 3.3) variation. TC- Top camera and SC- Side camera. (Key Grasp-Participant 9)

and other statistical parameters. The following equation was used to calculate the normalized Onset distance (NOD). The d_2 and the d_i were the onset distance and initial distance (in the Figure 3.3 and 3.1).

$$NOD = \frac{d_2}{d_i} \times 100\% \tag{3.2}$$

3.2 Statistical analysis

The statistical analysis was carried out using Minitab 19. All the analysis was carried out for 95% Confidence Interval ($\alpha = 0.05$). The total observations (including all the trials) of the onset study was 900, and the RMS study was 525. The Ryan-Joiner normality test (similar to Shapiro-Wilk) was conducted to identify the data distribution of both RMS and Onset distance study since it has the most power among other normality tests [142]. The results indicated that the

data in both studies do not follow the normal distribution.

1. RMS test - $p < 0.010$ for $n = 525$.
2. Onset Distance - $p < 0.010$ for $n = 900$

However, the studies have shown that the parametric test can be applied when the samples are > 30 even when the data distribution is not normal [143, 144]. Moreover, the studies have shown that statistical power is more significant when applying a parametric test on the non-normal data than a non-parametric test. Therefore following sections discuss the parametric and non-parametric test results of both studies. The parametric tests were used to investigate the deep relationships of the data with certain factors.

3.2.1 RMS study of EMG data

The main objective of the experiment is to identify the relationship between signal and the grasping patterns as well as the participants' demographic information. The RMS of the signal was used for the analysis, and the selection procedure is mentioned in Section 3.1.5. As the discussion in Section 4.1.1, few grasping attempts of some participants have to be removed due to saturation. Hence the total sample size (total attempts) for the RMS study was 525.

The Kruskal-Wallis test (non-parametric test) was applied to the total sample (all 6 grasping pattern attempts) to determine a significant median difference between grasping patterns. The results indicated that at least one median difference is statistically significant from others ($DF = 5, \;\; H = 127.12, \;\; p = 0.0000$). Figure 3.8 illustrates the mean, median, and confidence interval of RMS for each grasping pattern.

The relationship of demographic factors with the grasping pattern was studied using *General Linear Model (GLM)*. Since the sample size is 525, a parametric

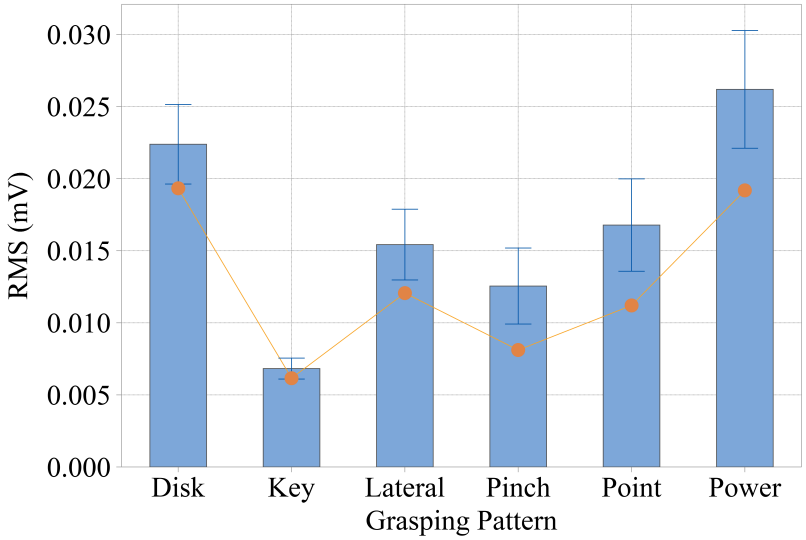


Figure 3.8: The interval plot of mean RMS of each grasping pattern. Orange dot indicate the median whereas the blue line indicates the 95% CI.

test can be applied to non-normal data. The GLM explores the linear relationship between given factors (independent variables (IV)) and the response (dependent variable (DV)). In this investigation, demographic factors and the grasping patterns are the IVs, and RMS is the DV. Furthermore, it carefully evaluates and visualizes multi-variables in a single graph. Thus more meaningful interpretations can be produced. Besides, the GLM does not consider the distribution form of the error term of the linear regress, and thereby it is more suitable for real-world applications.

The ANOVA GLM was computed, and the results indicate that the *gender* has no significant relationship with the RMS of the data ($F = 14.61$, $p = 0.110$). However, results suggest that all the other factors have a strong relationship with the RMS ($p < 0.05$). Therefore, the individual factors affect the sEMG signals of the grasping patterns. Figure 3.9 illustrates the main effect of each factor on the RMS and summarizes the ANOVA results. The detailed ANOVA table is reported in the Appendix C.

The relationship of interaction of each factor with the RMS revealed interesting results (see Figure 3.10). Although *gender* has no significant main effect on the

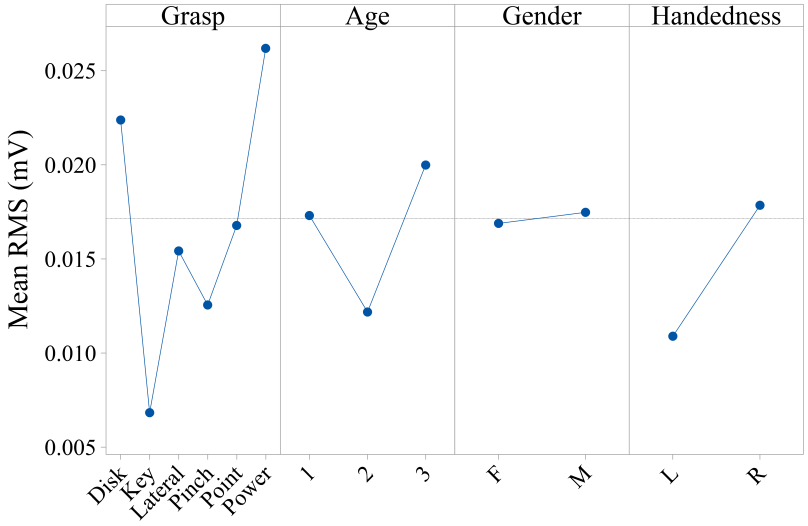


Figure 3.9: The summary of the ANOVA GLM of RMS study. The main effects illustrates how different factors affect the RMS of the signal.

RMS, the interaction of *gender* and *grasping patterns* has significant effect on the RMS. The interaction plots revealed a combination of different demographic factors that affect the RMS. For example, a person from age category 2 exhibits lower amplitude in the sEMG signal than the other two age categories. The statistical details are tabulated in the Appendix C. The following figures depict the summary of the interaction analysis.

3.2.2 Normalized onset distance study

A similar analysis was conducted for the Normalized onset distance (NOD) study. The results of the Kruskal-Wallis test showed that at least one median of the grasping patterns (normalized distance) is different from others ($DF = 5 \quad H = 164.53, \quad p = 0.0000$). The sample size of the Onset study was 900. Hence, the parametric tests are valid for the analysis. The Figure 3.11 illustrates the mean normalized onset distance.

The ANOVA GLM results indicate that all the factors and interactions strongly correlate with the normalized onset distance ($p < 0.05$). Therefore, hand onset

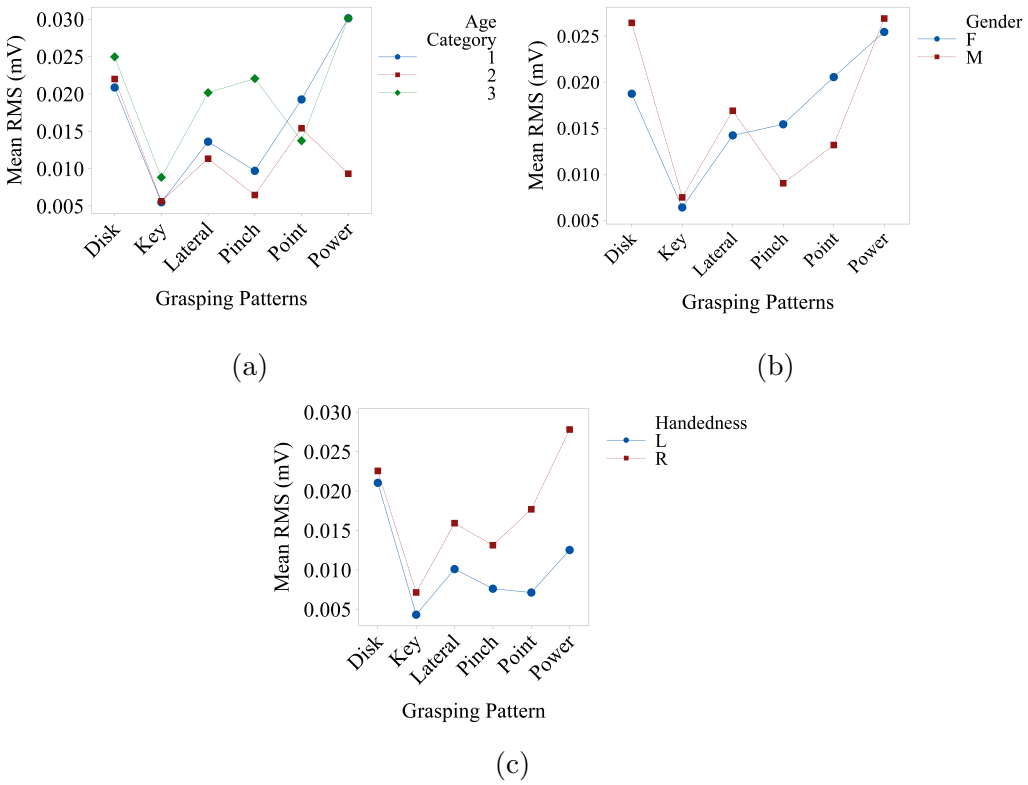


Figure 3.10: Interaction analysis of *grasping pattern*, *gender*, *age* and *handedness* on RMS of the signal.

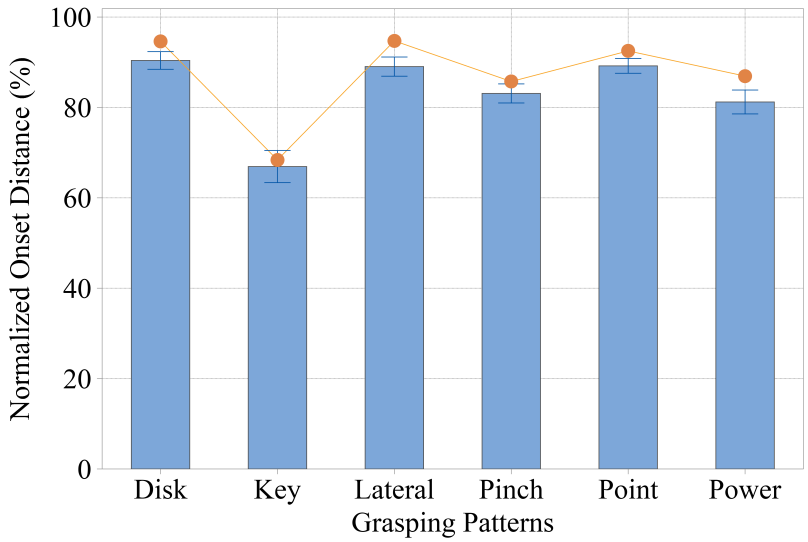


Figure 3.11: The interval plot of mean/median normalized onset distance of each grasping pattern. Orange dot indicate the median whereas the blue line indicates the 95% CI.

depends on the participants' demographic factors. Figure 3.12 shows that onset distance highly depends on the grasping pattern. However, *age* and the *handedness* has less effect on the onset distance than *gender*. The details are reported at the Appendix D.

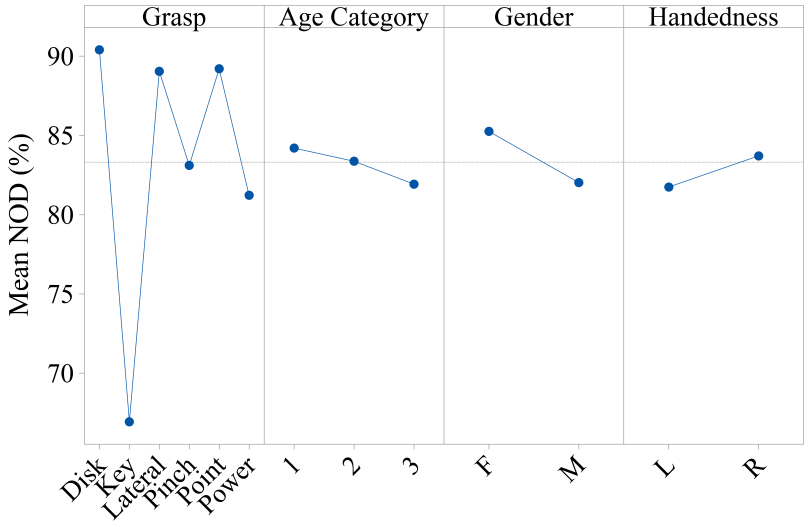


Figure 3.12: The summary of the ANOVA GLM of Onset study. The main effects illustrates how different factors affect the normalized onset distance of the signal.

The interactions of each demographic factor with grasping patterns have a significant effect on the normalized onset distance. The Figures 3.13 depicts that the demographic factors affect the hand onset distance significantly. For example, a right-handed person tends to form a hand closer to the object than a left-handed person.

3.3 Summary

The statistical analysis of the RMS of sEMG signal, Normalized onset distance, and the participants' demographic information have revealed significant insights regarding mutual relationships of each other. The Ryan-Joiner test indicated that both RMS and normalized distances do not follow normal distributions. Since the higher number of samples in each experiment, General linear model

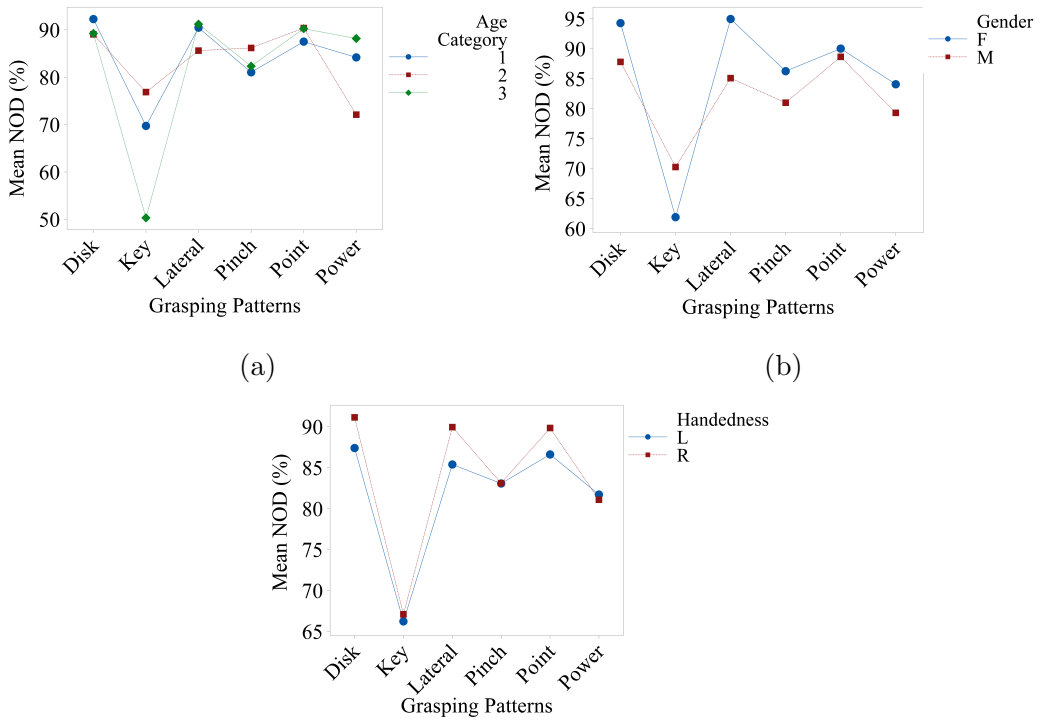


Figure 3.13: Interaction analysis of *grasping pattern*, *gender*, *age* and *handedness* on Normalized onset distance.

with an ANOVA test is applicable for the relationship analysis (effect and the interactions) although samples are not normal. The Kruskal-Wallis test revealed that at least one median is significantly different from others, and the interval plot (see Figure 3.8) illustrates the practical significance. The median RMS of the Disk, Key, Lateral, and Power grasps indicate a difference, whereas Pinch and Point have similar median RMS. The reason for that might be due to the similar finger arrangement in each grasping pattern.

Furthermore, studies revealed that the participants' demographic factors; Age, Gender, and handedness, have a distinct effect on the RMS. The participants in Age category 2 (26-30 years) showed the lowest mean RMS for all the grasping patterns, and it is statistically significant (no overlapping of CI with other categories) (see Appendix C). However, the gender-wise difference is insignificant. The interaction of each factor with grasping patterns revealed the most crucial

information needs to develop the sEMG classifiers. The results indicated that the RMS of the sEMG signal depends on the participant's age, gender, and handedness. Therefore, each participant has a distinct RMS of sEMG. Hence, the association of demographic factors for discriminating the grasping pattern could have a positive effect. Therefore, this research work investigated the effect of Age, Gender, and Handedness of the participants when employing them as features for the sEMG classification.

The normalized hand onset distances exhibited similar mean distances for most of the grasping patterns. However, the Key grasping pattern showed the lowest distance among others. The reason for that is that the finger arrangement of the Key grasp is closer to the Resting-state. Hence, participants start to form the hand for Key grasp when the hand is closer to the object. Even though the GLM ANOVA results revealed that the effect of each factor is statistically significant, Age and Handedness effect is comparatively smaller than others (see Figure 3.12 and Appendix D). Thus, the interaction of each factor with grasping pattern showed a significant effect on the Normalized onset distance as the RMS study. However, this information is valuable if the mean normalized onset distances are significantly different from each other. Hence, the research work did not employ the object distance for grasping pattern classification.

The key findings of this analysis were used for the EMG classifier development. The demographic factors of the user (age, gender and handedness) were used as input features for the classifiers. following chapter discusses the development process and illustrates the positive effect of identified demographic features on the classification accuracy.

Chapter 4

DEVELOPMENT OF GRASPING PATTERN CLASSIFICATION SYSTEMS

Predicting the grasping intention of the user is one of the main objectives of this research. This chapter discusses the development process of 5 sEMG classifiers and development of vision classification system to predict 6 grasping pattern. The initial section discusses the optimization and performance of five supervised learning algorithms (classifiers) for the collected sEMG signals. The second section presents the implementation of vision system and the process of determining the conditional probability matrix required for the vision system for the grasping pattern predictions.

4.1 Development of EMG classification system

The mathematical predictive models have been developed to forecast the results from new data. These models approximately map the input variables to output variables; therefore, the prediction is approximated. Furthermore, as discussed in Section 2.1.3, pattern recognition (PR) was used to predict the grasping intention of the user, and it is known as the classification prediction model. The PR techniques find the likelihoods of the output categories for given input data. In other words, the classifier calculates the probability of categorizing the given data into the given output categories.

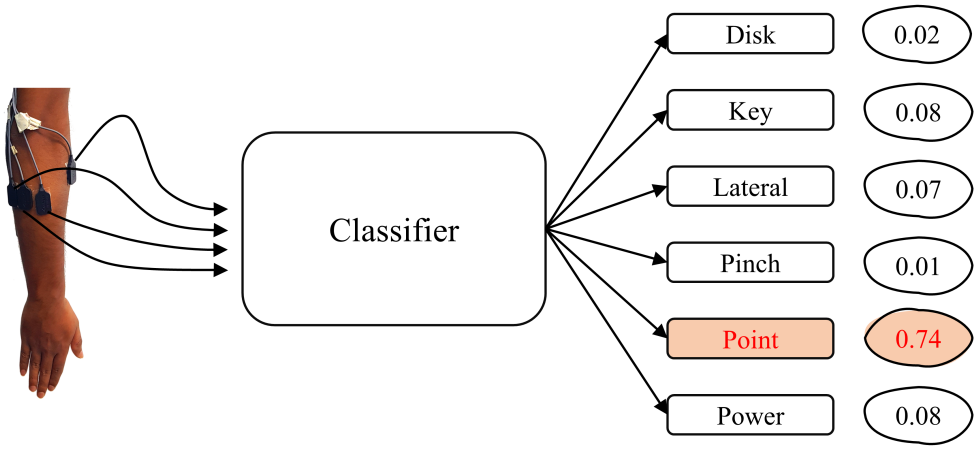


Figure 4.1: Overview of a classifier. The grasping pattern with highest likelihood (probability) is considered as predicted user intention.

This section presents the developments of Artificial Neural Network (ANN), Gaussian Naive Bayes (NB), K-nearest neighbor (Knn), Decision Tree (DT), and Linear Discriminant Analysis (LDA) classifiers for the sEMG classification. The sEMG signals of 6 muscles were filtered with specific filter banks and calculated the features. Then the features were fed into the classifiers to train the models as a set of groups. The effect of different feature groups was identified and, using grid search and Bayesian search methods, networks, and other classifiers were optimized.

4.1.1 Preprocessing of EMG data

Data preprocessing is a crucial step in machine learning. Since the classifier is trained using the collected data, the performance of the classifier depends on the quality of the data. Therefore, data should be cleaned and properly arranged to eliminate the bias effect and to achieve higher prediction accuracy.

Since the experiment gathered approximately 15 million data points from 15 participants, data might contain outliers, missing values, or saturated signals. Furthermore, since the sEMG signals have mV amplitudes, they are susceptible

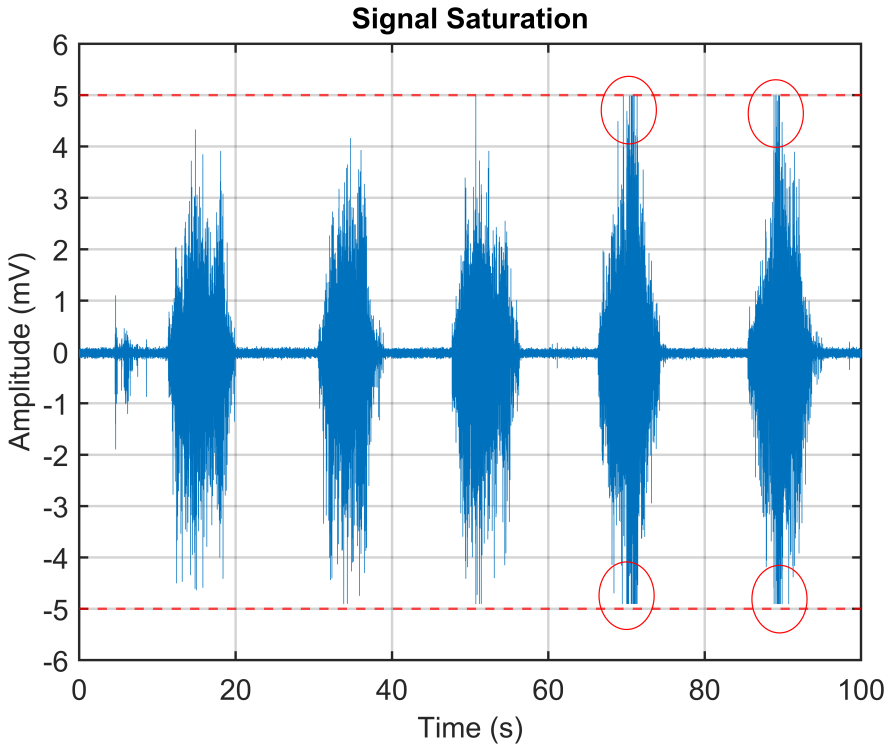


Figure 4.2: Saturated raw signal of Extensor Digitorum muscle of Subject 2, power grasp - trial 2. Subject is weighted 92Kg. Red ellipses indicate the saturated signal regions (4th and 5th attempts).

to signal saturation due to external artifacts. Therefore, a manual data screen process was carried out to investigate the outliers and other artifacts. Since this research interests only the sEMG signals of the grasping actions (during the RTG motion), the signals related to grasping action were extracted from the raw signal. Therefore, a manual extraction method was followed instead of the onset detection method since the manual method is more reliable and flexible. Besides, it helped the developer identify the behavior of the data domain and helpful in the filter development step.

As suspected, few sEMG signals of the experimental trials had the saturation effect due to high contraction force (see the Figure 4.2). The Delsys EMG acquisition system can amplify sEMG signals up to 10mV [-5mV, 5mV]. However, some participants have unintentionally exceeded this limit when performing some grasping patterns such as Power grasp and Lateral Power grasp. This effect

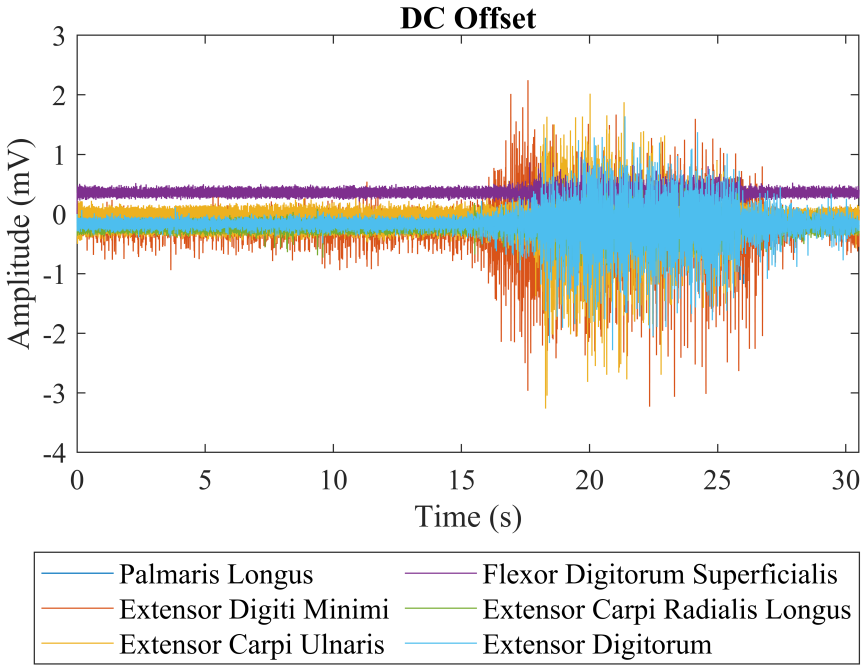


Figure 4.3: Illustration of positive and negative DC offset of the EMG channels.

was observed only on the EMG signals of male participants with high weights. Since their muscle contraction power is considerably higher than other regular participants, the amplified signal exceeds the 10mV amplitude and is saturated. Therefore, it deformed the typical signal pattern of the grasping action.

During the manual investigation of signals, a DC offset (or signal shift) was identified in most of the channels (muscles). Figure 4.3 illustrates the initial 30 seconds of raw sEMG signals of all muscles. It can be caused due to a defect of the amplifier and electrodes. However, the DC offset can be easily removed using a high-pass filter because the DC offset is the 0 Hz frequency signal.

After these anomalies were identified, the most suitable sEMG signals were selected for filtering and feature calculations. According to the literature, the best frequency range for the sEMG classification is (20-500) Hz [86]. Therefore, signals were filtered using a bandpass filter. Furthermore, several spikes were revealed in DFT (Discrete Fourier Transform) analysis. It is due to the harmonics

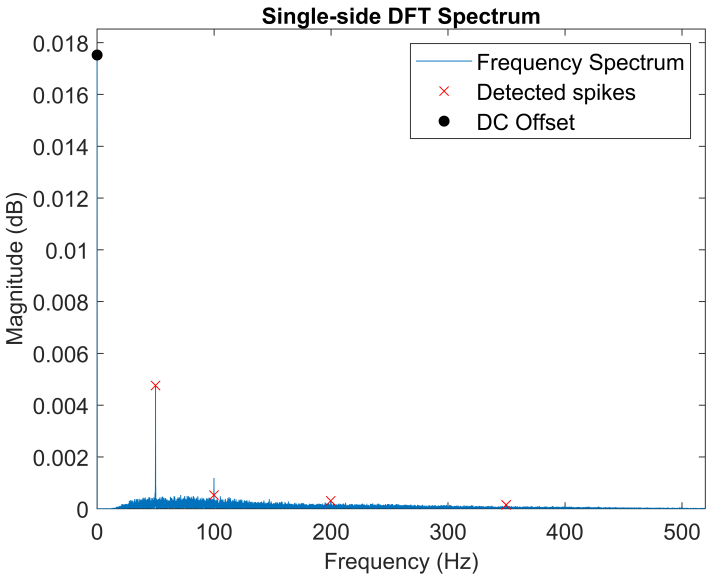


Figure 4.4: Illustration of initial stage of spikes detection and DC offset. DC offset is represented by the 0 Hz frequency and it is denoted by black dot on the graph.

of the Powerline Interference (PLI), which is 50 Hz. The PLI harmonics occur due to the non-linear loads connected to the same power grid where the experiment equipment is connected. Since these harmonics randomly occur on channels, an algorithm was developed to detect the spikes automatically and perform the filter operations.

The analysis of the DFT revealed that there are erratic spikes are occurred at PLI harmonic spikes. The PLI harmonics typically arise at 50 Hz, 100 Hz, 150 Hz, 250 Hz, and 300 Hz. They are denoted as F_a. The 50 Hz is F_1 and others were denoted using the same acronym. However, erratic spikes (49.1 Hz,50.4 Hz, etc.) may arise because of the fluctuation of AC frequency (see Figure 4.6). The amplitude of the PLI harmonic spikes and the erratic spikes decreases with the frequency increases (see Figure 4.4).

Based on this information, the algorithm was developed to detect the spikes and apply filters properly. The general overview of the algorithm is illustrated in Figure 4.5. Since harmonic spikes occur around specific frequencies, a range (r_i) was defined to calculate a threshold t_i. The range was defined as [F_a −

10, $F_a + 10]$ The reason to calculate the threshold is that not every erratic spike has significant amplitude compare to the surrounding frequencies and frequencies with amplitudes higher than the threshold is considered as undesired spike (f_{ij}). The threshold is defined as $1.8a_i$, and a_i is the amplitude of the median frequency of the range r_i. The f_{ij} represents the j^{th} spike at i^{th} range. There can be several spikes in a single range i.

$$R = \{r_1, r_2, r_3, ...r_i\} \quad \text{and} \quad i = 1, 2, ..., 7 \tag{4.1}$$

where,

$$r_1 = [40, 60], r_2 = [90, 110], .., r_7 = [390, 410]$$

$$f_{ij} = \{f_{11}, f_{12}, f_{21}, ..., f_{ij}\} \quad \text{and} \quad i = 1, 2, ..., 7$$
$$j = number\ of\ spikes\ at\ r_i$$

However, in the next step of the algorithm, these frequencies were screened to minimize the filtering process. The two-stage filtering process was followed since single-notch filtering is insufficient to remove all the spikes in the signal. The investigation revealed that there were some erratic spikes remain after the first stage. Hence, the second stage was defined with lower bandwidth to keep the critical information of the signal. In the first stage, the algorithm selects the frequencies divisible by 50 and sets the selected frequency values as the notch frequency of the notch filter. Hence, several notch filters were applied in the first stage, and the number of notch filters equals the number of spikes selected in the screening process. This stage removes the PLI harmonic spikes and most erratic spikes. A similar procedure was followed in the second screening process except, the frequencies that are not divisible by 50 were filtered. This was followed to

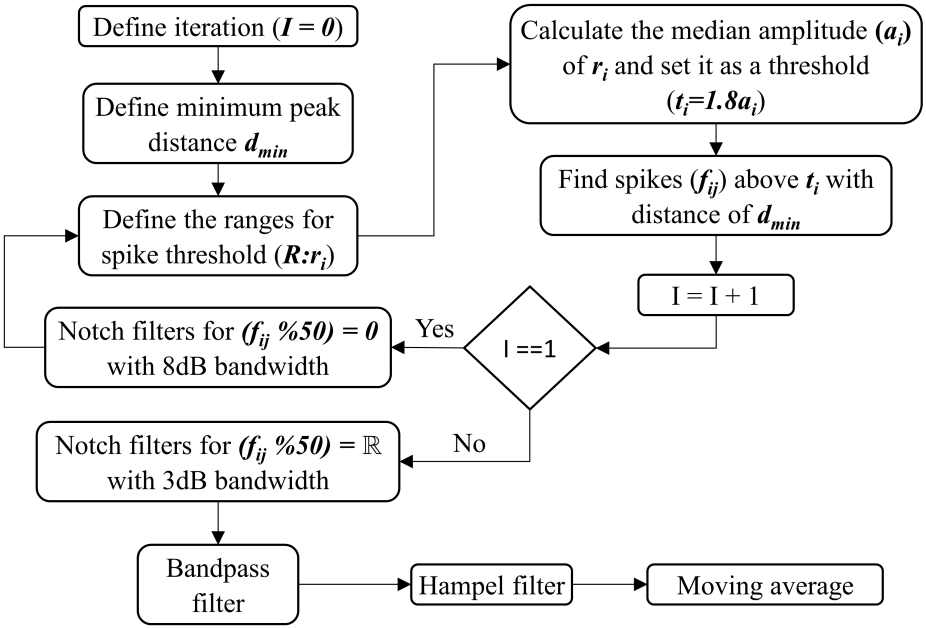


Figure 4.5: Overview of the algorithm of the filtering system.

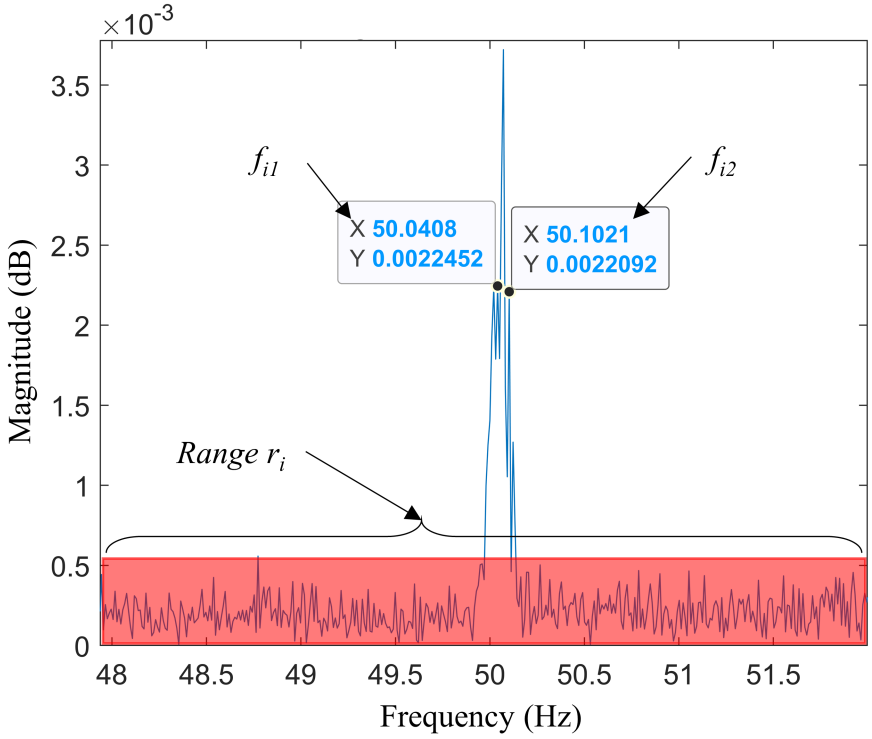


Figure 4.6: The erratic (surrounding) frequencies. This is due to the fluctuation of utility frequency.

remove the remaining erratic spikes. The bandwidth of the first stage notch filters was 8 dB, whereas 3 dB for the second stage notch filter. The intention of using 8 dB in the first stage is to remove both spikes as much as possible.

After the Notch filtering process is finished, (20-500) Hz IIR Butterworth band-pass filter is applied to remove the motion artifacts, ECG artifacts, DC offset, and unnecessary high frequencies. Even though a thorough filtering process is applied, outliers cannot be removed using conventional filtering methods. Therefore, a Hampel filter was used to remove the outliers within the *3 standard deviations*. 400 data points were considered to calculate the *standard deviations* as the real-time process collects 400 data point at a time. Finally, a moving average with $2ms$ window length was applied to smooth the fluctuations of the signal.

4.1.2 Feature selection

The prediction models are developed using the features calculated from the filtered signals. It is essential to select the suitable features for the model because the performance (the accuracy of the prediction) is dependent on how well the model maps the input (features) to the output (classes). There is a large number

Table 4.1: Feature groups selected for the investigation.

Time Domain (TD)	**Autoregressive Coefficient (AR)[1]**	**Frequency Domain (FD)**	**Nonlinear Features (Non)**
Root Mean Squared (RMS)	Coefficient 1	Average Energy	Sample Entropy
Zero Crossing	Coefficient 2	Mean Frequency	Skewness
Wavelength	Coefficient 3	Median Frequency	
Sign Slope Change	Coefficient 4		
Mean			
Variance			
Willison Amplitude			
Integral EMG			

[1] AR is frequency domain features.

of signal features available, and many of them have proved their effectiveness [93, 95]. However, when the number of features (inputs) increases, the model training time increases significantly, and the system memory required for the training process increases [100]. Therefore, it can cause to decrease in the accuracy of the models by overfitting the data. Hence, the features should be simple enough to calculate but significantly represent the signal information.

Based on the literature, Time-domain features proved the highest effectiveness among other domains [6]. However, this investigation studied the effectiveness of the other feature domains on the classifier to ensure that models are well optimized for the application. The following Table 4.1 represents the selected feature groups and the respective features which are selected from the previous studies [12, 102, 145, 146].

The features can be selected using different methods. The selection methods generally fall into two categories: 'Supervised' and 'Unsupervised' methods. The supervised selection methods need to know the output of features, whereas unsupervised methods decide the output using specific algorithms [147]. Most unsupervised methods are statistical methods, but some supervised methods use statistics to rank the features. However, some selection methods cost high computation time, such as wrapper methods (Greedy search, Forward selection), whereas many statistical methods use less time, such as filter methods (Fisher score, Correlation-based feature selection, PCA).

In this research, features were primarily selected wrapper method, which uses a set of features and evaluates the model performance. The Principle Component Analysis was conducted for the initial identification of the most significant features. The following sections describe the procedure and the results obtained from PCA.

Principal component analysis of features

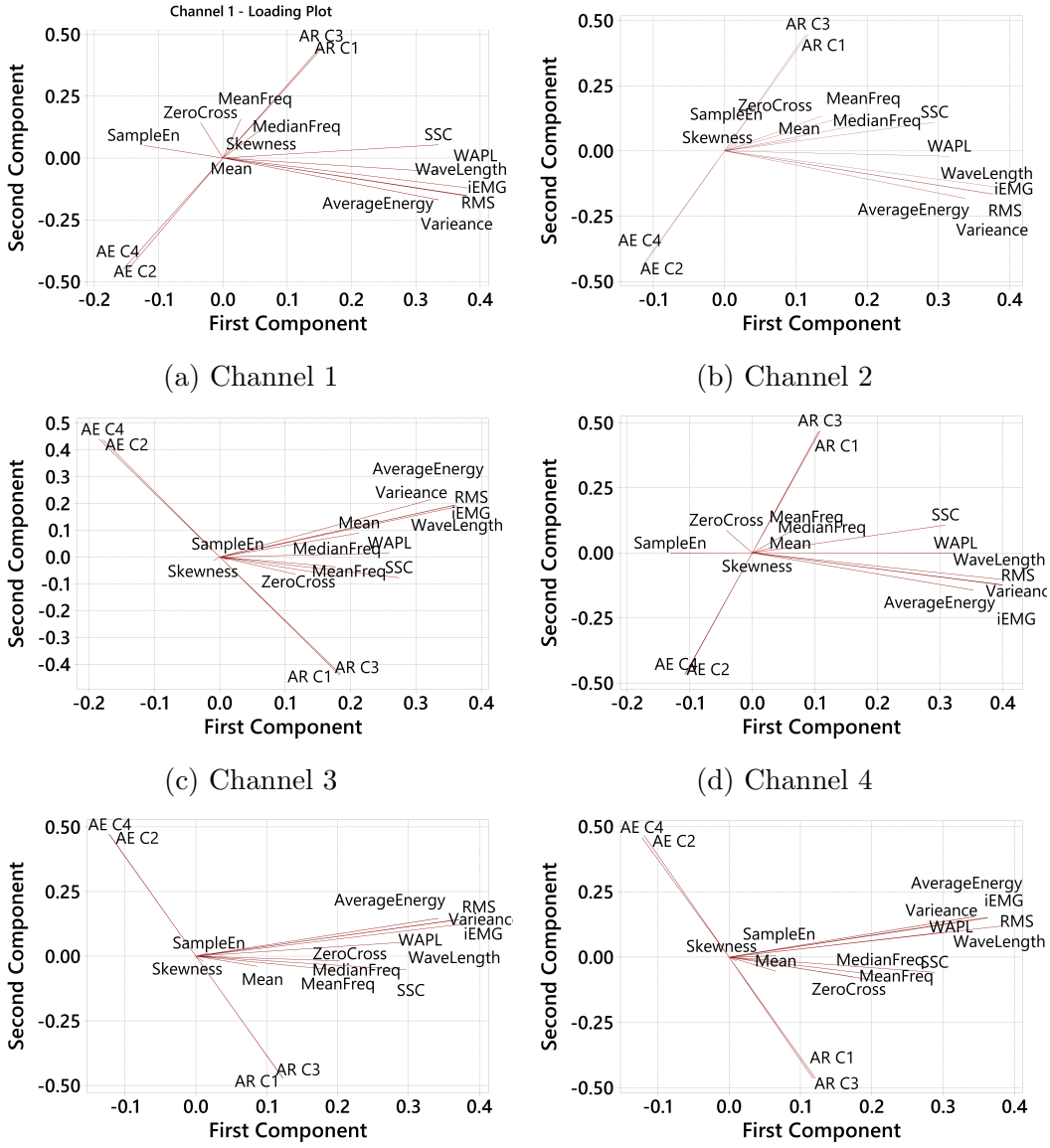
As discussed in Section 2.3.4 the Principal Component Analysis is a widely used tool for dimensionality reduction. However, it can further be used for the feature selection [148]. The PCA calculates several principal components (PCs) representing all the features given (represent the signal variation). These few PCs are used instead of all the features. It is known as dimensionality reduction. The cumulative contribution of the signal variation of the selected PCs should be at least 85% to have a considerable performance of the model. In this project, models were trained using the first five PCs. The results are discussed in the 4.3. However, in this section, the selection criteria of the features are discussed.

The loading plot illustrates how much of a weight particular feature has on the PC. A higher weight (positive or negative) implies that the particular feature has higher contribution for that PC in terms of the variation. The Figure 4.7 illustrates the loading plots of each channels (muscles).

The figures show that TD features positively correlate with each other and a negative correlation among AR features. The understanding of the correlation of features is useful when feature selection. In general, positive correlate features are more suitable than non-correlated or negatively correlate features. Because positively correlate inputs help the model to map the input and outputs more precisely. Although it not always the best selection criteria, it provides a reasonable perspective about the features. Hence, features were grouped as Table 4.1 and followed the wrapper method to select the features.

Feature Grouping

As discussed in the previous section, PCA revealed that TD features significantly influence the signal. The previous researches further confirm it. However, these feature sets were used for the initial model training and investigated their



(a) Channel 1

(b) Channel 2

(c) Channel 3

(d) Channel 4

(e) Channel 5

(f) Channel 6

Figure 4.7: TD features have strong positive influence to the PC1 which has approximately 30% variation of the signal in all channels. AR features have both positive and negative influence to the PC2.

effect on the classification performance. Following is the group of features investigated.

1. TD features

2. AR features

3. TD and AR features

4. FNon features (Frequency and Nonlinear)

5. All features

6. 5 Principal Component

The performance of the feature groups was evaluated by calculating the mean True Positive Rate (mTPR) of all output classes (six grasping patterns). The mTPR was calculated for the testing data set, and the testing set consists of signals of six grasping patterns. The investigation process is similar to the wrapper method, but the final feature group selection is conducted through manual trend analysis. Each feature group is used to train several ANN models with selected neurons in the first hidden layer. Then, the feature group with a positive trend is used for further ANN optimization and the other conventional model train. The following figure presents the feature group selection process.

4.1.3 Development protocol of classifiers

The development of the sEMG classification system is one of the objectives of the research. The previous section introduced the classifiers that were used in this project. Although the introductions were brief, the following sections will

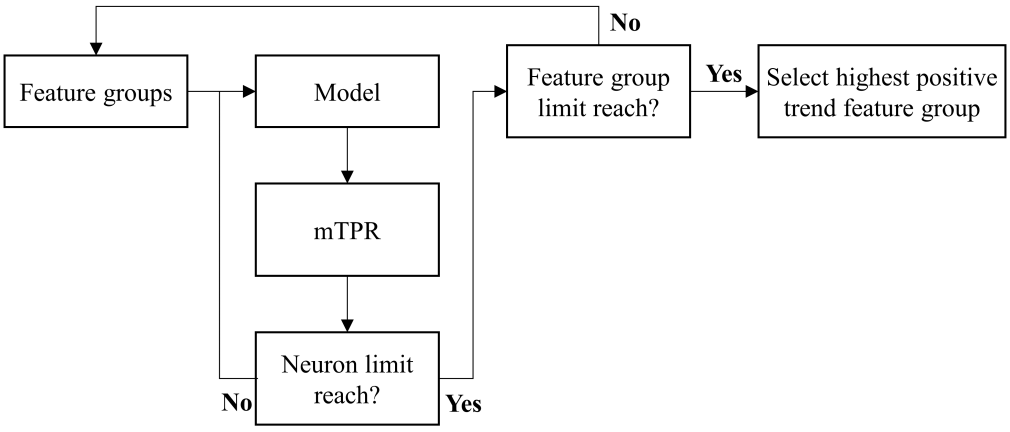
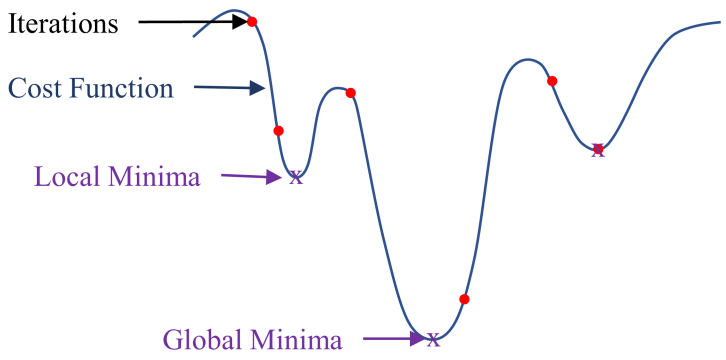
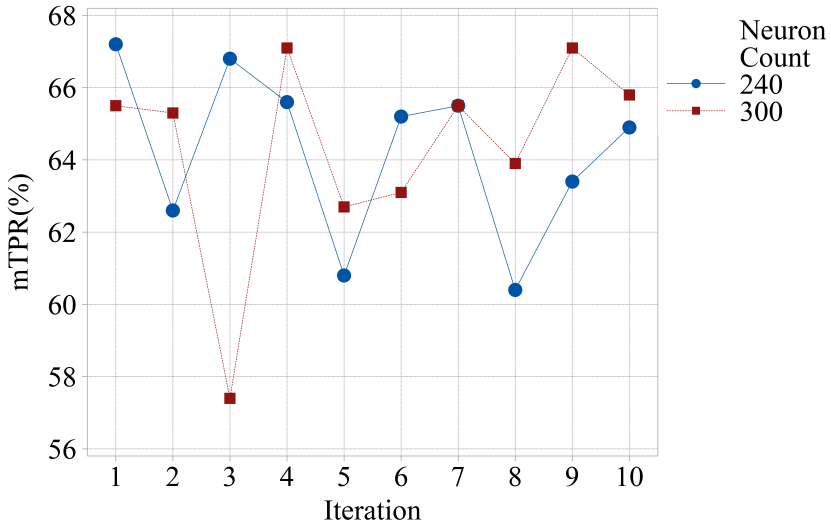


Figure 4.8: Development procedure of the classification models.



(a) Different cost values at different iteration in same ANN model



(b) Variation of the model performance during 10 iterations

Figure 4.9: Variation of model performances (mTRP) when the training process is iterated ten times.

discuss the model parameters more comprehensively. This section discusses the protocol that followed to develop the classifiers. First, the window length analysis was performed at the initial stage of the development. Then data was split into two subsets: Training Data and Testing Data. Finally, feature selection and parameter optimization was carried out. The results illustrated in this report are from the Testing Data set.

Highest performed model selection

The model developments and evaluations were executed in Matlab 2019 platform. When the ANN model was developed, weights initializing was achieved using *Nguyen-Windrow* weight initializing algorithm. Since the initial weight assignment is random, each initializing places the model in a different place on the cost-function (see Figure 4.9a). This can be considered as an optimization because by training the model several times with different initial weights, different minimum cost values can be achieved. Then the model with the lowest cost value (i.e., the highest mean True Positive Rate (mTPR)) is selected for the prediction. Figure 4.9b shows the variation of mTPR of 10 iterations of two ANN models of 240-hidden layer neurons and 300-hidden layer neurons. All the models were trained 10 times and selected the highest mTPR model iteration for further analysis.

Representation of results

As describe in the previous section, mTPR was used to evaluate the performance of the models. The following equation shows calculation of mTPR. Here, TP_i is the True Positive count of i^{th} class and FN_i is the False Negative count of i^{th} class. N represent the number of classes (grasping patterns). In this case $N = 6$.

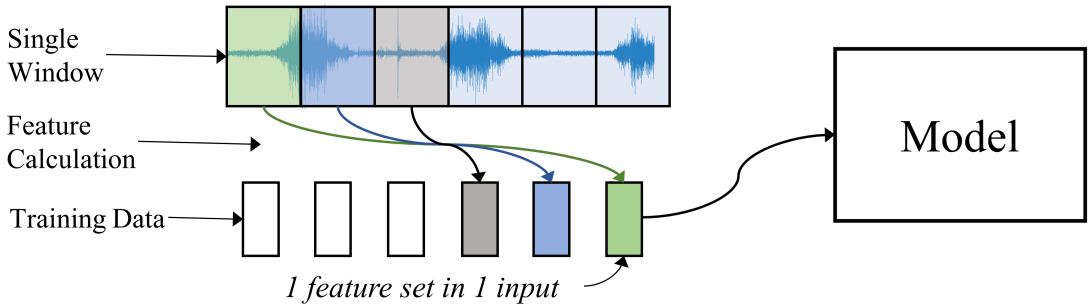
$$mTPR = \frac{1}{N} \sum_{i=1}^{N} \frac{TP_i}{TP_i + FN_i} \times 100\% \tag{4.2}$$

Window length analysis

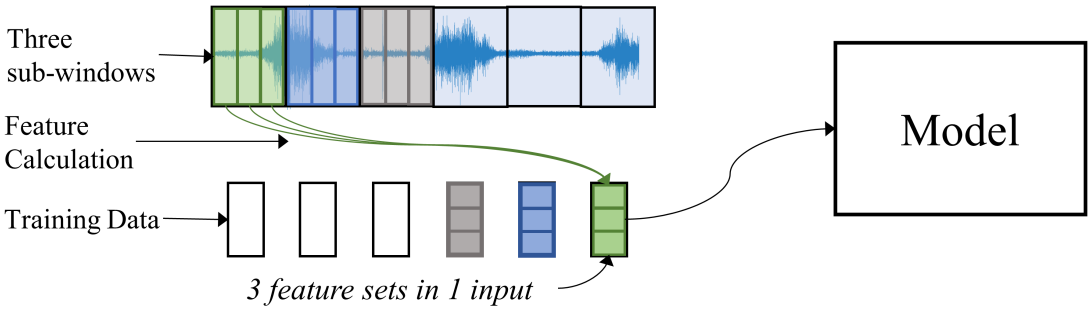
Section 2 discussed the importance and the limitation of the window length. Since the real-time should produce the final prediction within $300ms$, the window

length was limited to $100ms$, and the remaining $200ms$ was used for the processing. The EMG acquisition system can process 4000 data points per second (4000 Hz sample rate). Hence, in a $100ms$ time length, it processes 400 data points that are more than sufficient to calculate the features. Since there are no studies about multiple windows classification, the effect of the sub-windows of a feature window on the classifier was investigated. Since this was a simple quantitative study, only the TD features were selected to check the model performance because of their proven effectiveness [76].

In general, all the data points in a window are used to calculate a single TD feature set (i.e., 8 TD feature). This study examined the effect of multiple feature sets calculation in a single window (i.e. 2 or more TD feature sets). Instead of reducing the window length, dividing the $100ms$ window length into sub-windows and calculate feature sets for each sub-windows. Then the calculated feature sets were used to train the model. Figure 4.10 illustrates the process of training data.



(a) Conventional Feature calculation and model training



(b) Investigated method

Figure 4.10: Illustration of sub-windowing of the main window.

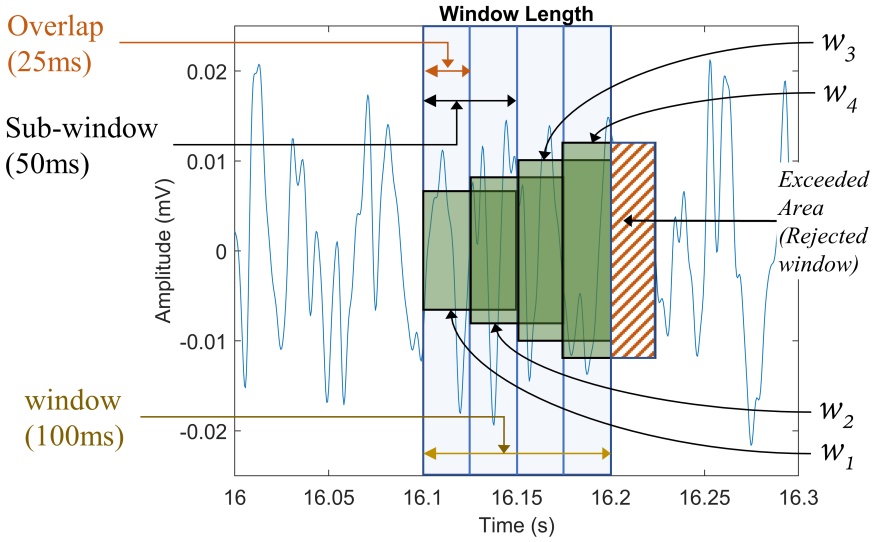


Figure 4.11: Illustration of sub-window of 50ms length.

In that example, window length is divided into three. It is practically impossible as the $\frac{100}{3}ms$ is a rational fraction. However, this is only for demonstration purposes.

There are eight TD features in a single channel and 48 features for all six channels in a single sub-window. This number is changed according to the number of sub-windows or the dividing factor. Table 4.2 shows the dividing factors, sub-window length, and respective total number of features. The overlap was set to 50% of the sub-window length. For example, when the window is divided into 2 (dividing factor), then the sub-window length is $50ms$ as illustrates in the

Table 4.2: Sub-window Analysis

Diving Factor	**Length of a sub-window (ms)**	**No. Sub-windows in a Window**	**No. Features without exceeded window[1]**	**No. Features in a Window [2]**
1	100	1	1	48
2	50	4	3	144
5	20	10	9	432
8	12.5	16	15	720

[1] Last sub-window exceed the window length.

[2] each sub window has 6 channels, and each channel has 8 features.

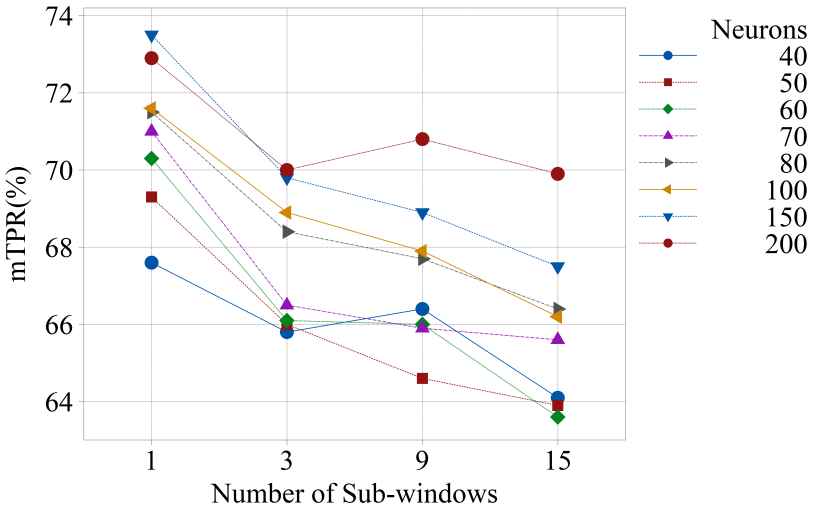


Figure 4.12: Training analysis of the sub-windows. mTPR is presented in the y-axis. This figure shows the training accuracy of each model.

Figure 4.12 and overlap is $25ms$. Since sub-windows are calculated using the sliding window method, 4 sub-windows can be calculated. However, the last sub-window exceed the window length, and it was rejected when the features are sort for the training.

As the example illustrated in Figure 4.11, the last window (w_4) is rejected as it exceeds the window length. That red area is not practically achievable in a real-time situation since that data is in the next window, which is not acquired at the moment. The investigation was performed for ANN with different neuron counts for 1 hidden layer. The inputs of the ANN were varied according to the number of sub-windows, and the outputs were the 6 grasping patterns. As discussed in previous sections, maximum mTPRs of each 10 iterations were plotted for the analysis (see Figure 4.12). The figure shows that the single window model yields the highest mTPR in each neuron count. This may be due to high information density within a window. Multiple feature sets were not effective as a single feature set because of the high input nodes [149]. Therefore, the conventional single window was selected for the feature calculation.

Data splitting

The extracted signals were combined and were shuffled in order to minimize the bias effect. The data is then split into two main groups: The training group and the Testing group. The 1.6 million data points were allocated for the Training group, and 1 million data points were allocated for the Testing group. The feature calculation reduced the Training group to 22,000 observations and the Testing group to 15,700 observations. However, during the ANN training process, Training observations were divided into 3 sub-sets: Training (60%), Validation (20%), and Testing (20%). This was followed to avoid overfitting and bias. The Testing observation (15,000 observations) was used to evaluate the performance of the model for the new data. All the results mentioned in this report are from the Testing observations unless otherwise specified.

4.1.4 Optimization of classifiers

Optimization of the classifier is an important step. It searches for the best parameters and the hyperparameters of the model, which gives the highest performance. This section discussed the Grid search and the Bayesian optimization method that followed to optimize the classifiers. The conventional classifier (NB, DT, KNN, LDA) was optimized using the Bayesian method, whereas the ANN was optimized using both methods. The reason for that is that ANN takes considerable time to train, and the number of variables that need to be optimized higher than other classifiers. Hence, the initial optimizations were carried out using the Grid search, and later part of the optimization was done by the Bayesian optimization.

Table 4.3: Parameter sets of the ANN structure for the Grid search method. HL-Hidden Layer

Parameter	Values
Feature Groups	TD, AR, TDAR, FNon, ALL, PCA[1]
No. neurons in HL 1 [2]	40, 50, 60, 70, 80, 100, 150, 200, 240, 300, 350, (400, 500, 600, 700)
No. neurons in HL 2	40, 50, 60, 70, 80, 100, 150, 200, 240, 300, 350, 400, 450

[1] PCA feature set was used only at the initial stage

[2] For specific cases this set was extended to 700

Table 4.4: Hyperparameter of the Learning rate for the Grid search method

Hyperparameter	Values
Learning rate	0.0001, 0.0004, 0.0008, 0.001, 0.005, 0.007, 0.01, 0.04, 0.08, 0.1

Grid search optimization

The Grid search method is a traditional way of performing both parameter and hyperparameter optimization. It is also known as *parameter sweep* which is a *exhaustive searching* method. The parameters are defined as sets and perform training for each combination. This research focused on the number of hidden layers, neurons on each hidden layer, and the feature groups. The initial *Learning rate* selection was an independent analysis. However, at the end of the optimization, both the learning rate and momentum constant are optimized. The Grid search method is compliant with the parallel computation, though process cost time. The following Table 4.3 and 4.4 report the parameter sets that were used for the grid search optimization.

Grid search optimization analysis and results

The grid search method was used only for the ANN optimization. The initial search was about *learning rate* because it has a direct effect on the neural network.

It is impossible to analytically calculate the learning rate for a given model [150]. Only the trial and error method has to be followed to identify the best value. Typically it is between 1 and 10^{-6} [151]. One simple approach that literature has suggested the grid search method, which has the range of $[1, 10^{-5}]$. Hence, as the Table 4.4 reports, learning rates were evaluated for TD features. Three different hidden layer neuron configurations were selected for the evaluation. The network has only one hidden layer and 48 inputs. The comprehensive details of neural network parameters and hyperparameters used for the training are reported in Appendix H.1.

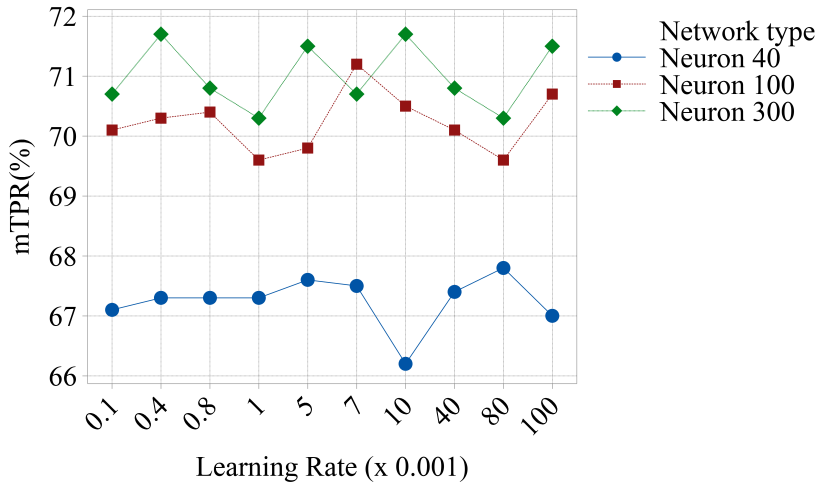


Figure 4.13: mTPR comparison of three networks with different learning rates

Figure 4.13 reveals that the 0.007 produces the highest mTPR for all the configurations. Hence it was selected for further analysis. However, the final optimization was carried out to find the best *learning rate* and the *momentum constant* for each layer. The parameters and hyperparameters of the following analysis are as same as Appendix H.1 unless otherwise specified.

After selecting the proper learning rate, core analysis was started. In the beginning, the performance of a single hidden layer was evaluated for all the feature categories.

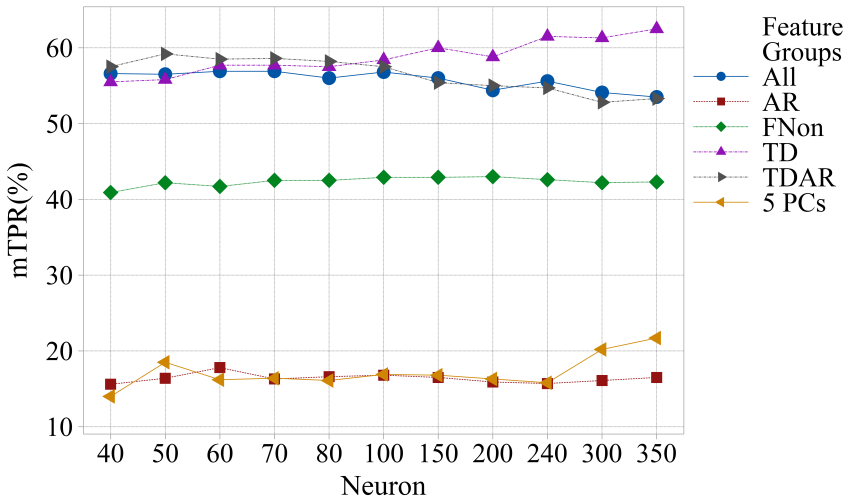


Figure 4.14: mTPR comparison of different feature groups (without the demographic features) and different hidden layer neurons

According to Figure 4.14, TD features showed a positive uptrend than others, whereas the *5-Principle components* did not show expected performance. It was outperformed by the ALL features (17 features) by a considerable amount. However, the investigation was conducted further to evaluate the identified demographic features. The same network configurations except the inputs are increased by 3 (the Age, Gender, and Handedness).

A significant improvement was noticed when the *demographic* information is added as features. As the statistical analysis has suggested, these features have a distinctive effect on signals of each grasping pattern. Therefore, the features improved the performance of the classifier when the network tries to find the patterns of the signals. Figure 4.16 shows the performance increment percentage of each feature group at different neuron configurations. The increment percentage of a group was calculated as follows.

$$Increment = \frac{mTPR_{with\ Demog} - mTPR_{without\ Demog}}{mTPR_{without\ Demog}} \times 100\% \qquad (4.3)$$

The TD, FNon and All, feature groups were improved by 16%-21% approx-

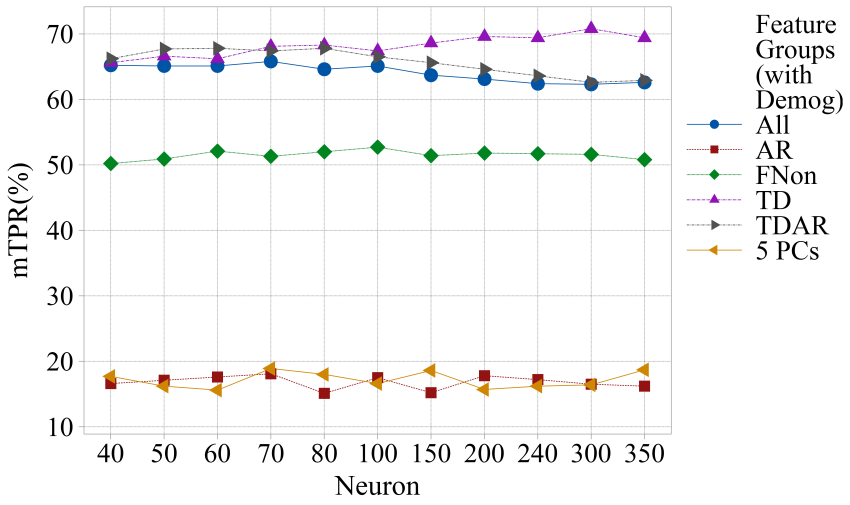


Figure 4.15: mTPR comparison of different feature groups with demographic features and different hidden layer neurons

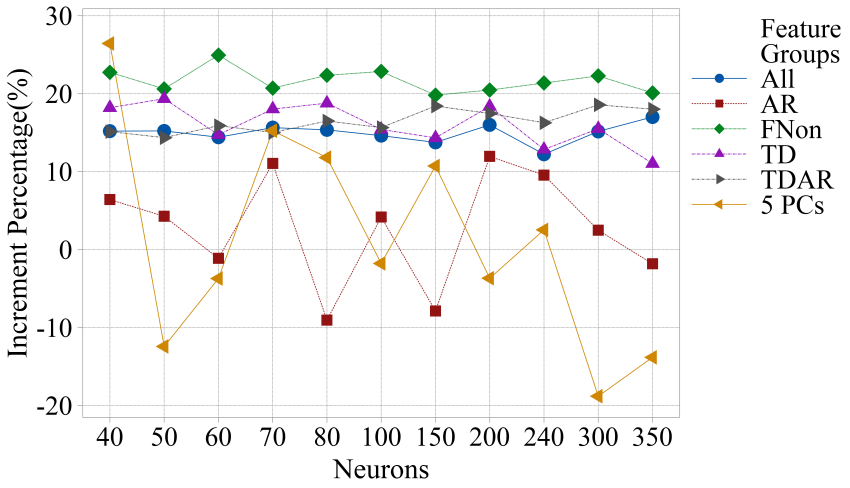


Figure 4.16: Performance improvement of different feature groups and different neuron configurations.

imately, although other groups did not show significant improvement. The AR and PC feature groups' performances were decreased in some neuron configurations. However, these analysis results indicate that the demographic features have a significant positive effect on the performances of the classifiers. The based on the results illustrated in Figure 4.15, the TD group showed an uptrend. Hence,

the TD group was selected for further analysis (40 - 700 neuron range). The trend analysis of the TD group showed that the number of neurons in the 1^{st} hidden layer is limited to 500 as the higher neurons do not improve the performance of the classifier. The 300-neuron model exhibited the highest mTPR.

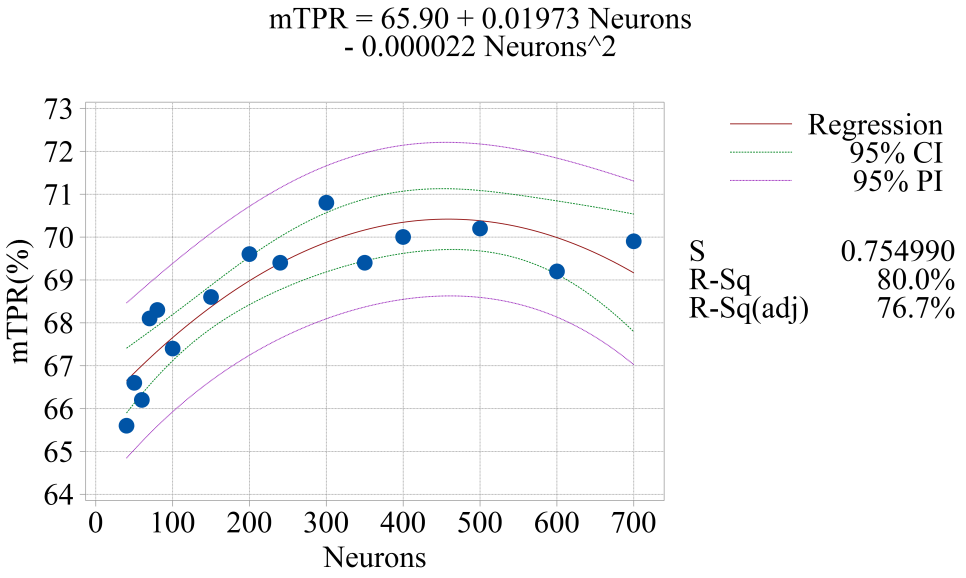


Figure 4.17: The trend analysis of the TD group. The regression equation is noted at the top of the figure. Other necessary statistical information are mentioned within the figure.

After the 1^{st} hidden layer optimization was finished, 2^{nd} hidden layer optimization was started. The configurations of neuron count in hidden layers are tabulated in Table 4.5.

Table 4.5: Number of Neurons in each hidden layer

Hidden Layer	**Neuron count**
HL 1	70, 80, 100, 150, 200, 240, 300
HL 2	40, 50, 60, 70, 80, 100, 150, 200, 240, 260, 300, 350, 400, 450

Each of the neuron combination in Table 4.5 was evaluated. Figure 4.18 revealed that some of the combinations have an uptrend. However, the graphs were not clearly defined the trends. The [80, 400] and [300, 450] neuron configurations

produced the higher mTPR. A trend analysis of each line was carried out to identify the best neuron configuration for further analysis.

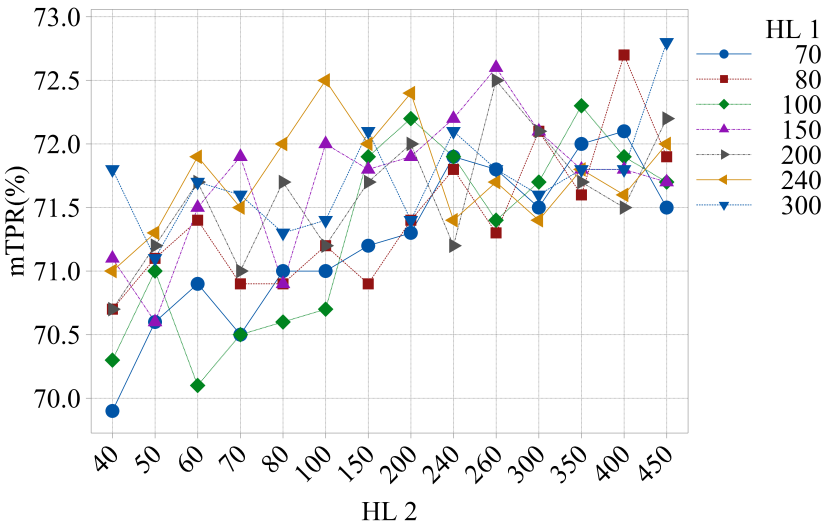


Figure 4.18: Double HL analysis of TD features.

The following Table 4.6 tabulates the summary of the trend analysis. The trend lines were defined as the number of neurons at HL 1. The descriptive figures are illustrated in the Appendix H.2. A positive (+) coefficient for both first-order and second-order terms indicates an uptrend. Moreover, a positive (+) second-order and negative (-) first-order term indicates the uptrend. All the other combinations indicate downtrends.

Table 4.6: summary of the trend analysis

HL 1	**Constant**	**first order**	**second order**	**trend**
70	69.94	0.0114	-0.000017	down
80	70.78	0.0032	0.0000	up
100	69.79	0.0139	-0.00002	down
150	70.65	0.0115	-0.00002	down
200	70.88	0.0064	-0.00001	down
240	71.52	0.0028	-0.00001	down
300	71.59	-0.0012	0.00001	up

According to the table, HL1-80 and HL1-300 had the uptrend and were selected for further analysis. However, training time for the HL2 for a higher number of neurons (< 500) is a time-consuming study. The third layer analysis more time-consuming than HL2 analysis. Hence, Bayesian optimization was used for the third hidden layer as well as other hyperparameter tunings.

Bayesian optimization

The Bayesian Optimization algorithm is a frequently used optimizer for non-convex, nonlinear, and high dimensional problems [152]. It is a practical approach to search for the maximum value of computationally expensive functions such as Neural Networks. The algorithm predicts the maximum point at the function without evaluating the expensive function but evaluating an approximated statistical model of the expensive function. This research used Bayesian Optimization to tune the hyperparameters of the ANN and other conventional classifiers such as Hidden layers, number of neurons, or number of neighbors in kNN. The evaluation function of the ANN is the mTPR of the test data. Therefore, the Bayesian optimizer searches the highest performed model for new data (Test data). The conventional classifier optimization was achieved by evaluating the cross-entropy performance function. The comprehensive details are discussed in the following section, and additional information is reported in the Appendices.

Bayesian optimization of ANN

The optimizable parameters of these evaluations are the second hidden layer neurons and the third hidden layer neurons. The parameters of the Bayesian optimization are listed in Table 4.7. According to Table 4.6, HL1-80 and HL1-300 showed the uptrend, and they were already evaluated up to 450 neurons for the HL2. Therefore, the second hidden layer neuron range for the Bayesian optimization was set to [500, 2000]-integers. However, for the third hidden layer,

Table 4.7: The parameters of the Bayesian optimizer

Parameter	**value**
Acquisition function	Expected Improvement per second plus
Exploration Rate	0.2
No. of seeds	10
Evaluations	5

[40, 2000]-integers as it was never evaluated before.

The results of the optimization are tabulated in Table 4.8. The bold numbers indicate the optimized value. The extended analysis of the second hidden layer results indicated that the [80 623]-model has the highest mTPR. Hence it was used as the first and second hidden layer neuron count for the third layer analysis. On the other hand, the [300 500]-model needed significant time to train because of the large number of weights and biases. Besides, [80 623]-model took 22.5 hrs to evaluate the third hidden layer neuron count during the optimization (50 iterations).

Table 4.8: Optimized hidden layer neuron count. The bold numbers represents the optimized variable

HL 1	**HL 2**	**HL 3**	**Max. mTPR(%)**	**Time elapsed (hr)**
80	**623**	-	73	2.04
300	**500**	-	72.3	6.48
80	623	**500**	72.6	22.51

Since the model is supposed to be implemented on real-time application, a less complex [80 623]-model is selected. However, the hyperparameters of the selected model were again optimized. Since there are two layers, the learning rate (LR) and the momentum constant (MC) of each layer should be selected for the best performance. The Table 4.9 reports the Bayesian results of the hyperparameter optimization of [80 623]-model. The combinations of four variables were searched for the highest mTPR.

Table 4.9: Optimized hyperparameters

mTPR(%)	LR - HL 1	MC - HL 2	LR - HL 2	MC - HL 1
73.1	0.01278	0.6705	0.04467	0.1136

The following Figure 4.19 illustrates the Bayesian search of two-variable optimization (LR and MC of a single layer). First, the algorithm approximates the objective function (red surface - surrogate mean) and searches the next minimum point (-mTPR) using an acquisition function. Then evaluates that point on the objective function and updates the surrogate function. Again search for the new minimum point on the updated surrogate function (red surface). This iterative process was performed for 50 iterations and output the lowest (-)mTPR with the corresponding hyperparameters (MC and LR).

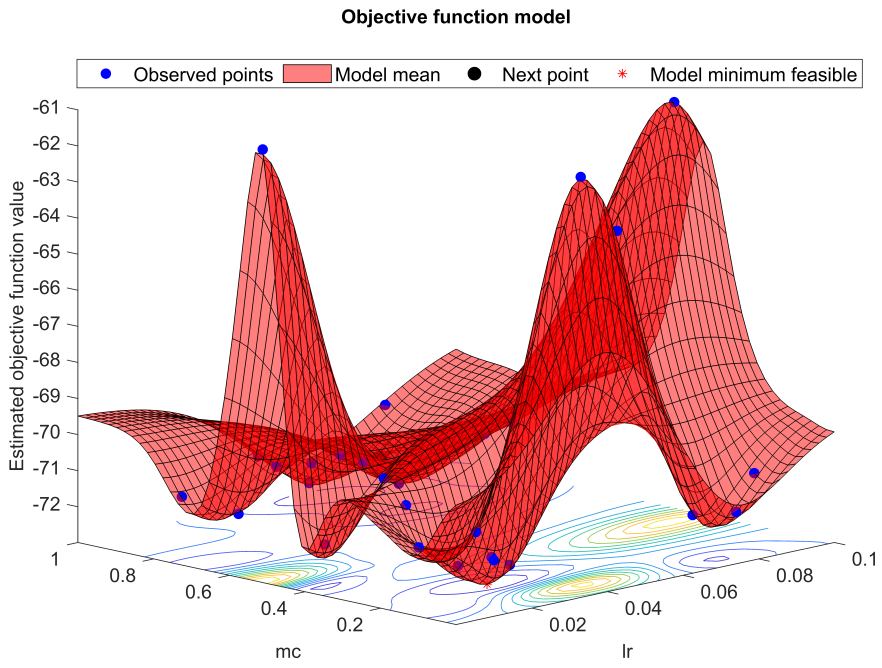


Figure 4.19: The surrogate function of two variables (LR and MC). The algorithm is searching for the minimum point on the surrogate mean surface (red surface).

The final architecture of the optimized ANN is depicted in the Figure 4.20 and corresponding parameters and the hyper-parameters are reported in the Table 4.10.

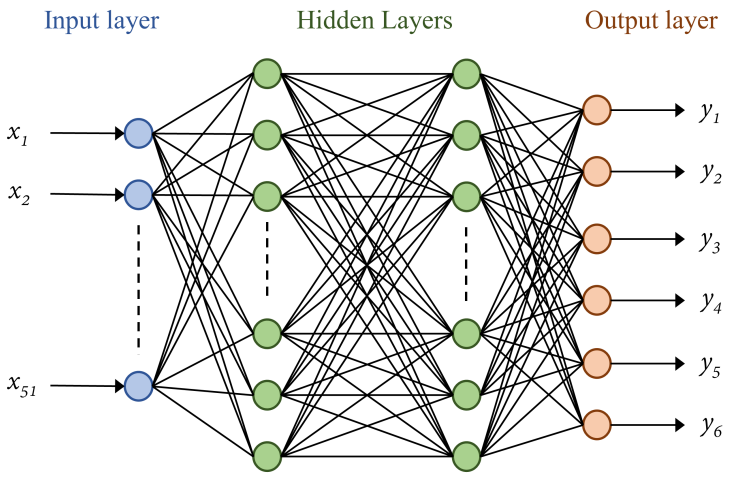


Figure 4.20: Optimized neural network architecture.

Table 4.10: Optimized hyperparameters of the TD feature-models without demographic features. TD-model without demographic features produced the highest mTPR.

Layer	**Node/Neurons**	**LR**	**MC**
Input	51	-	-
Hidden 1	80	0.01278	0.6705
Hidden 2	623	0.04467	0.1136
Output	6	-	-

Bayesian optimization of conventional classifiers

A similar approach was followed to find the optimum parameters for the KNN, LDA, NB, and DT. The optimized parameters were tabulated in the following Table 4.11.

Table 4.11: Optimized hyperparameters for the conventional classifiers.

Classifier	**Parameter 1**	**Parameter2**
LDA	Delta	Gamma
DT	Min. Leaf Size	Max. Number of Splits
KNN	No. of Neighbors	Distance
NB	Distribution	Width

The optimized results of the conventional classifiers are illustrated in the fol-

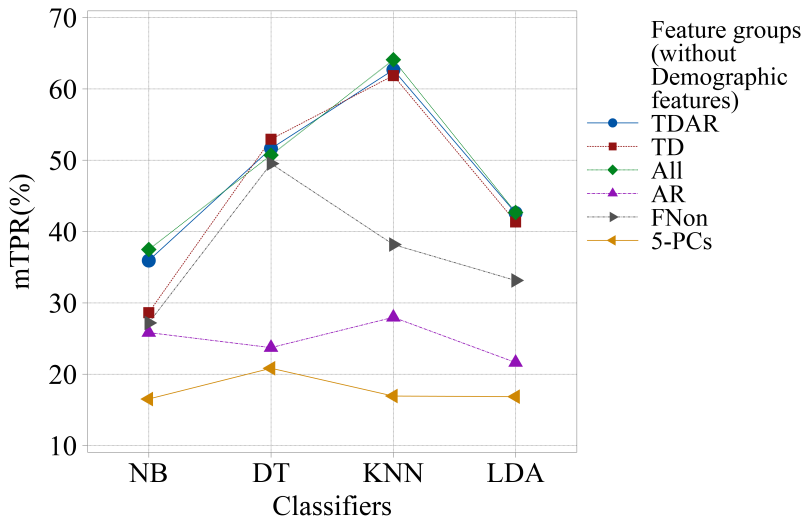


Figure 4.21: The feature groups performances on the conventional classifiers. The demographic features are not considered.

lowing Figure 4.21. The performance of the classifiers was evaluated for the selected feature groups without the demographic features. However, DT and KNN showed notable performances. The other classifiers exhibited poor classification accuracy. The reason might be that the data is not normally distributed, and models unable to generalized for the given data.

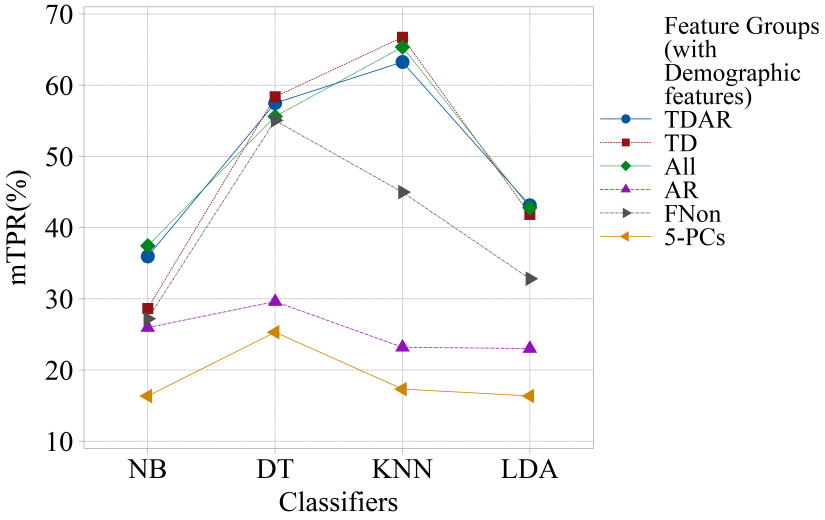


Figure 4.22: The performance of the feature groups with demographic features on the conventional classifiers.

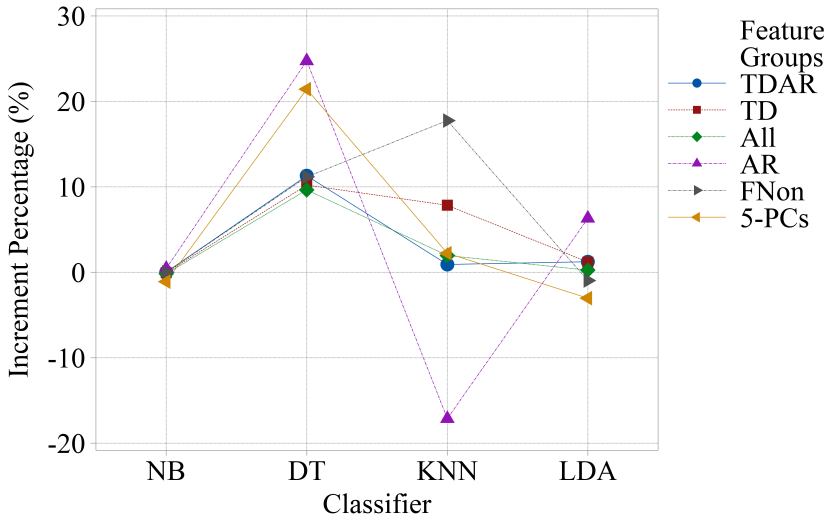


Figure 4.23: The mTPR increment due to the demographic features.

However, the demographic features have a considerable effect on several classifiers but not every feature group. The performances of the classifiers with demographic information are illustrated in the following Figure **??**.

According to Appendix I.1, the highest mTPR was observed on the KNN classifier when the feature group is TD with demographic features. The overall accuracy was increased with demographic factors but not a significant amount. Besides, these models did not outperform the optimized ANN model, which exhibited *73.1% mTPR*. Therefore, the ANN with two hidden layers was used for the real-time EMG classification. The optimized parameters of the conventional classifiers are tabulated in the Table 4.12.

Table 4.12: Optimized hyperparameters of the TD feature-models with demographic features. TD-model with demographic features produced the highest mTPR.

Classifier	**Parameter 1**	**Parameter2**
LDA	0.010046	0.000367
DT	7	1068
KNN	15	seuclidean
NB	kernel	1.0335

4.2 Development of vision-based classification system

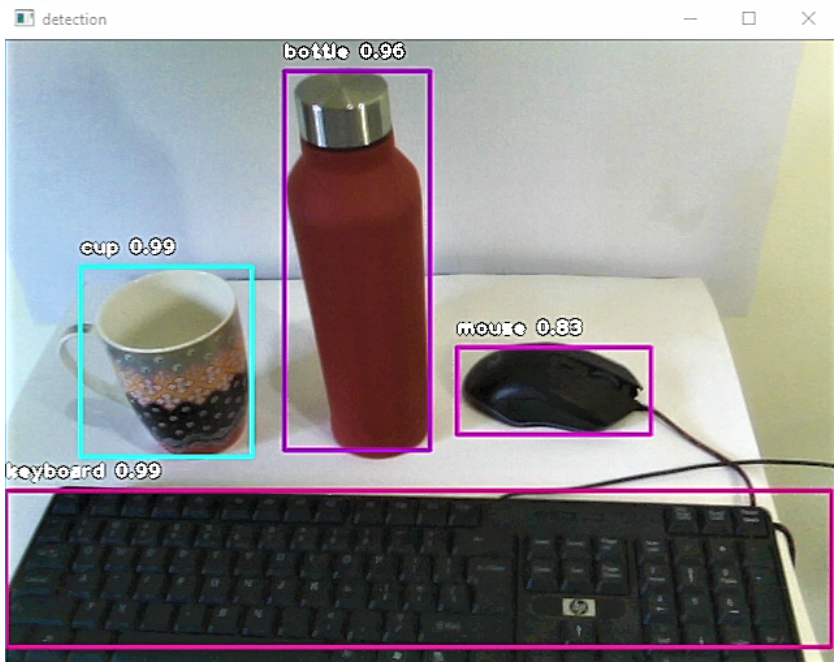


Figure 4.24: The real-time object detection. The respective confidence score is also presented.

This section discusses the development of vision-based grasping pattern classification system and the derivation of conditional probability matrix. The vision system used the YOLO to identify the object. Then the system retrieves the corresponding conditional probability matrix for the object. This matrix represent the probability of grasping pattern for the detected object similar to the EMG classification result matrix. The highest probability pattern can be considered as the predicted grasping pattern for the object. However, the algorithm only calculate the probability matrix and send it to the hybrid system to predict the final grasping pattern using the EMG probability matrix. Furthermore, the algorithm was restricted to detect one object in an image and images (or video feed) usually contain single object to avoid the multiple detection. Figure 4.24 illustrates the real-time detection with corresponding confidence score for the detection.

The vision system was developed on the Python 3.6 platform. The CUDA 10 libraries were used alone with the Tensorflow 1.30 and Keras 2.0.8. The

object detection was performed by YOLOv3 using the NVIDIA Geforce GTX 750 (4GB). The pre-trained YOLOv3 network was used for better performances, and COCO data set was used to train that model. Since the pre-trained model can classify 80 object types, including domestic objects like spoons, cups, bottles, keyboards, and laptops, no additional training was required. However, to reduce the computation time, this object list was reduced to 49 classes to detect only the domestic objects. The following sections present the derivation of conditional probability matrix and vision algorithm that used YOLO to detect objects.

4.2.1 Proposed probability matrix

A survey was conducted to calculate the grasping pattern usage of selected objects. i.e., the probability of a grasping pattern uses to grasp an object. The survey aims to collect the frequency of the selected grasping pattern used to grasp a given object. The grasping patterns selected for the survey were *Disk, Key, Lateral, Pinch, Point, and Power*. The number of objects was limited to 11. The objects are *Scissor, Cup, Bottle, Mouse, Fork, Laptop, Keyboard, Knife, Remote, Phone, and Spoon*. The participants were asked to select the frequency of the given grasping pattern use to grasp the given object. For example, *How frequently you use the* Disk *grasp to operate, hold or move the Bottle?*. The user was given a Likert scale to select the frequency, and the frequency was based on his/her experience/knowledge on grasping the given object.

A detailed description with visual aid was given to comprehend the questions easily, and the survey was designed using both English and Sinhala language. The participants were selected randomly with different demographic factors such as age, gender, occupation, and handedness. The 5-point Likert scale was used to collect the opinion of the participants regarding the grasping pattern usage. As usual, intervals between each Likert item are assumed equal. The items and the corresponding scores are given in the following table.

Table 4.13: 5-point Likert scale parameters

Item	Never	Rarely	Sometimes	Often	Always
Score	1	2	3	4	5

The pilot test was carried out for 6 participants to determine the sample size and to get the external opinion on the questionnaire. The questionnaire consists of 22 questions. Initial questions are regarding the demographic information, and few questions were given to familiarize with the concept of grasping. Each main-questions (11 main-questions for 11 objects) consist of 6 sub-questions (6 grasping patterns), and each sub-question has the Likert scale. The *Standard Deviation (SD)* and the *Margin of Error (MOE)* were calculated for each main question to determine the sample size for each question. Although the sample size of every main-questions was 9 for the given *SD* and *MOE*, the actual survey sample size was increased to have higher statistical power. The statistical details of the pilot test are presented in Table F.1. Hence the 35 participants were selected, but only 31 responded to the survey (88.57% response rate). There were 17 males and 14 females who participated in the survey, and the $31.23yrs(14.1)$ mean age group was considered for the frequency calculation. The *Cronbach's Alpha*, $\alpha = 0.8242$ indicated that the questions have adequate reliability and consistency.

The score of the each sub-questions (grasping pattern) was calculated as the following equation. The L_j is the Likert score of item j^{th} (as the Table 4.13) and f_{ji} is the number of occurrences of j^{th} Likert item in i^{th} grasping pattern. K is the Number of Likert items ($K = 5$). $Score_i$ is the cumulative score of the i^{th} grasping pattern (i.e. grasping class).

$$Score_i = \sum_{j=1}^{K} L_j f_{ji} \tag{4.4}$$

The scores for each grasping pattern in the main-question (with respect to the object) were calculated as equation 4.4. Then the ratio of each j^{th} score with

respect to the total score of the main-question was calculated as equation 4.5. The $p(c_i|O)$ was considered as the conditional probability of i^{th} grasping pattern used to grasp the given object O. The complete probability matrix is listed at the Appendix F.

$$p(c_i|O) = \frac{Score_i}{\sum_{i=1}^{6} Score_i} \tag{4.5}$$

This $p(c_i|O)$ is used for the propose fusion method to reduce sEMG classification error (i.e. reduce the False Positive).

4.2.2 Vision algorithm

The developed vision system retrieves the probability matrix each time when the detection take place. The YOLO detects objects in the given image and calculates the respective confidence score. The confidence score represent the probability that image (or selected region) contains the detected object. If the confidence score is higher than the 80%, algorithm finds the respective matrix of the detected object and send the matrix to the Matlab script which the EMG classification executes. Since it takes around 30 seconds to start detecting objects in the image at the initial execution, EMG classification and other executions were paused until the first detection take place. Once it is finished, algorithm waits for the response from the Matlab script which classify the EMG and fuse the two results. The following algorithm illustrates the procedure to determine the respective grasping pattern probability matrix for the fusion system.

Algorithm 1: Algorithm of vision system

$flag_{initial} \leftarrow 0$

$flag_{matlab} \leftarrow 0$

while *'x' key is not pressed* **do**

  $img \leftarrow$ RGB camera image

  convert img to python array object

  if $flag_{initial} == 0$ **then**

    $detected_{copy}, detection \leftarrow$ **YOLO(img)**

  end

  $flag_{initial} \leftarrow 1$

  send $flag_{initial}$ to matlab

  while $flag_{matlab} \neq 1$ **do**

    read $flag_{matlab}$ `// waiting for matlab response`

  end

  $detected_{copy}, detection \leftarrow$ **YOLO(img)**

  for *all objects in detection* **do**

    $item \leftarrow detection['objectname']$

    $confidence_{score} \leftarrow detection['confidence']$

  end

  remove duplicates in $item$ and $confidence$ arrays

  if $confidence >= 80$ **then**

    $Prob_{matrix} \leftarrow$ **ConditionalProb**$(item)$

    send $Prob_{matrix}$ to hybrid system algorithm

  end

end

Function *ConditionalProb(item)*:

  `// search the conditional probability matrix of` $item$

  $Prob_{matrix} \leftarrow matrix(item)$

return $Prob_{matrix}$

4.3 Summary

The sEMG classification was carried out for five classifiers, and each classifier showed different performances. In addition, the effect of conventional signal features and the introduced new features were investigated. The results revealed that the TD features with demographic factors produced the highest classification performances for ANN, whereas other conventional classifiers showed diverse behaviors. The summary of the optimized classifiers testing results is tabulated in Table 4.14. As the statistical analysis suggested, the demographic features improve the class separability in classification, and as a result, the system can identify the pattern of new data efficiently and accurately.

Table 4.14: Optimum testing results of ANN, LDA, Knn, NB and DT for each grasping pattern. The best feature group was TD with demographic factors

Grasp	**ANN**	**KNN**	**DT**	**LDA**	**NB**
Disk	72.2	64.7	58.6	37.4	13.3
Key	80.7	80.6	63.6	57.8	81.0
Lateral	73.2	66.4	57.5	41.7	19.0
Pinch	62.3	56.2	47.3	32.4	20.2
Point	78.2	72.4	64.4	44.7	11.9
Power	73.3	60.2	58.7	37.3	26.5
mean & Std	73.31±6.34	66.74±8.74	58.37±6.13	41.86±8.86	28.6±26.2

The vision algorithm can detect multiple objects with higher confidence. The system efficiently retrieves the conditional probability matrix and send it to the hybrid system real-time. Although images have low resolution (640x480) YOLO detects object accurately with higher confidence.

Chapter 5

EMG-VISION HYBRID SYSTEM

This chapter presents the derivation of the proposed hybrid system. The EMG classification results and the probability matrix from the vision system are fused to produce the final improved prediction. The fusion method is based on the Bayes theorem. The final results are further improved by employing post-processing methods (see Figure 5.1).

Post-processing of the results is essential when the output is not at an acceptable level. The proposed fusion method acts as a post-processing technique as it improves the output of the classification system. In addition to the fusion method, two post-processing methods were investigated. The Majority Vote (MV) and the Bayesian fusion (BF) were evaluated to identify the best technique for this classification system. The following sections present the proposed fusion method and brief introductions to MV and BF.

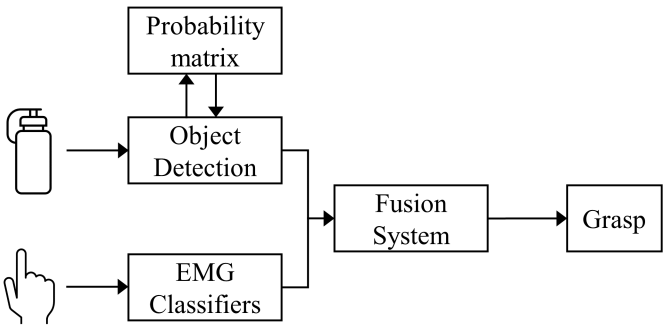


Figure 5.1: Hybrid system overview.

5.1 Data fusion using Bayes theorem

As describe in the section 4.2, visual input was used to enhance the classification results. Bayesian fusion (Not the post-processing technique) was used to fuse both classification output($p(c_i|I)$) and the object detection system output (probability matrix), $p(c_i|O)$. The general Bayes theorem calculates the conditional probabilities, but it needs an extensive database. It assumes that all the inputs are dependent on each other. In practice, it is intractable to solve conditional probabilities for some cases. However, in many Bayes theorem applications, the aforementioned assumption is neglected [153]. They assume that all the variables are independent of each other. The studies have shown that the Bayes theorem approach can succeed even the independence assumption is violated [153].

The classification system outputs the six conditional probabilities $p(c_i|I)$, with $i \in [1, 6]$, corresponds to the probability of i^{th} grasping pattern when given the *Intention I* (EMG signals). Note that the summation of the output probabilities are equal to one, i.e., $\sum_{i=1}^{6} p(c_i|I) = 1$. The vision system probability matrix represent the conditional probabilities, $p(c_i|O)$, where $i \in [1, 6]$ and O is the detected object. The variable I and the event O are independent events as the user intention is not change with the object. For example, whether it is a *cup* or *bottle*, use intention is to grasp the object. However, grasping pattern is changed according to the intention. Then we can compute the conditional probability of a *grasping pattern* when both the *object* and the *intention* is given which is $p(c_i|O, I)$. The equations 5.1 presents the Bayes approach for this conditional probability.

$$p(c_i|O, I) = \frac{p(O|c_i, I)p(c_i|I)}{p(O|I)} \tag{5.1}$$

Since the variable I and O are independent, $p(O|c_i, I) = p(O|c_i)$ and $p(O|I) = p(O)$ [8]. Update the equation 5.1 with the assumptions,

$$p(c_i|O,I) = \frac{p(O|c_i)p(c_i|I)}{p(O)} \tag{5.2}$$

Apply the Bayes theorem to $p(O|c_i)$,

$$p(O|c_i) = \frac{p(c_i|O)p(O)}{p(c_i)} \tag{5.3}$$

Substitute the 5.3 to the 5.2,

$$p(c_i|O,I) = \frac{p(c_i|O)p(O)p(c_i|I)}{p(O)p(c_i)} \tag{5.4}$$

The $p(O)$ at numerator and denominator cancel out. then the $p(c_i|O,I)$ simplified to,

$$p(c_i|O,I) = \frac{p(c_i|O)p(c_i|I)}{p(c_i)} \tag{5.5}$$

Here, $p(c_i)$ is a prior probability of the class i and it is known as *normalization constant*. Since $0 \leq p(c_i|O), p(c_i|I) \leq 1$, the product of the both terms is between [0, 1]. Hence the *normalization constant* was set to 1. The 5.5 simplified to following.

$$p(c_i|O,I) = p(c_i|O)p(c_i|I) \tag{5.6}$$

The $p(c_i|O,I)$ is the final predicted probability of the i^{th} grasping pattern. Then the final output of the classification system is a 6×1 matrix,

$$C = \begin{bmatrix} P(c_1|O,I) \\ P(c_2|O,I) \\ \vdots \\ P(c_6|O,I) \end{bmatrix}$$

The C matrix is fed to the final post-processing algorithm (Majority Vote or Bayesian Fusion) to get the steady output.

5.2 Post-processing of fused data using majority vote and Bayesian fusion

The MV algorithm is a simple data post-processing method that selects the highest voted element within a defined range (windows) as the output of that range. It is further used to smooth the outputs when spurious misclassifications are present [91]. To calculate the highest voted class, MV needs the recently produced final outputs of the classification system. However, for this investigation, votes of the classes were not considered. Instead of that, the mean of each class within the 4 windows was calculated, and the highest mean class was selected for the current output. For the simplicity lets take $P(c_{ji}|O,I) = P(c_i|w_n)$ where the w_n is the n^{th} window considered.

$$G_{final} = \underset{i\in[1,6]}{\operatorname{argmax}} \left(\frac{1}{4} \sum_{n=1}^{4} P(c_i|w_n) \right) \tag{5.7}$$

Here, j is the class number, $P(c_i|w_n)$ is the i^{th} class probability at n^{th} window. G_{final} is the best classification class among other classes. The mean vector, $\left(\frac{1}{4}\sum_{n=1}^{4} P(c_i|w_n)\right)$ is assigned as the current output.

In the section 5.1, Bayes theorem was also used to fuse the two outputs. However, it is further used as a post-processing method to fuse recent windows to form final results [8]. A similar approach was used as the MV post-processing technique to produce the final output from the BF. The Equation 5.8 represent the BF approach [8].

$$G_{final} = \underset{i \in [1,6]}{\operatorname{argmax}} \left(\frac{1}{4} \prod_{n=1}^{4} p(c_i|w_n) \right) \tag{5.8}$$

5.3 Algorithm of the proposed hybrid system

The both EMG classification system and the vision-based classification system execute simultaneously on Matlab and Python, respectively. The 100ms window sEMG data is collected and processed by the ANN EMG classification system, while vision system detect objects and retrieves the probability matrix. The hybrid system fuse both results to produce the final improved result. The vision system start initially and execute the algorithm 1. Once the vision system send the $YOLO_{start}$ flag to the hybrid system, it starts to process the EMG signals while sending another flag to the vision system indicating that hybrid algorithm (Matlab script) has started without errors. Then vision system starts to continuously detect objects and sends corresponding probability matrices, $Prob_{matrix}$ to the hybrid system. Once the hybrid system receives the first $Prob_{matrix}$, system start to fuse the two results from EMG and vision system. The hybrid algorithm performs a binary classification prior to the EMG classification to differentiate *resting state* and *grasping state*. If the mean absolute amplitude of the EMG signal within 100 ms window is less than 0.06 mV, it is considered as *resting state* EMG signal. The other windows with higher mean absolute amplitude are used for the classification. The Majority Vote (or Bayesian fusion) is used to determine the best result using the results of previous four windows. The following algorithm illustrates the basic procedure of the hybrid system.

Algorithm 2: Algorithm of hybrid system

```
set demographic factors of the user
windows ← 4        // set the number of majority (bayesian) windows to 4
set EMG DAQ
load filters for EMG signal
startDetection ← 0
flag_matlab ← 0
ProbMat ← 0
while YOLO_end ≠ 1 do
    if YOLO_end = 1 then
        stop EMG DAQ                    // stoping flag from vision system
    end
    if YOLO_start == 1 then
        flag_matlab ← 1
        send flag_matlab to vision algorithm
        start EMG DAQ
    end
    if Prob_matrix is recieved then
        ProbMat = Prob_matrix        // probability matrix from the vision
         system
        startDetection = 1
    end
    if startDetection = 1 then
        collect EMG data of 100ms
        emg ← data                      // assign sEMG data to emg variable
        prediction ←FusionClassification(emg, ProbMat, windows)
    end
    Display the prediction                          // or send to simulation
end
Function FusionClassification(emg, ProbMat, windows):
    if EMG amplitude < 0.06 then
        emg_prediction ← Rest                                  // Hand resting
    else
        emg_prediction ← EMGClassification(emg)          // predict EMG
         pattern
    end
    Prob_matrix ← matrix(item)
    fusion_result = emg_prediction × Prob_matrix                // Sensor Fusion
    prediction ← MajorityVote(fusion_result, windows)
    // Bayesian fusion is also checked during the validation
return prediction
```

Chapter 6

REAL-TIME VALIDATION OF HYBRID SYSTEM

This chapter discusses the validation process that was carried out to evaluate the performance of proposed grasping pattern classification system in real-time. The system validation was carried out in two stages: EMG classification system validation and combine system validation. Since previous studies have already validated YOLO, this research did not carry out any experiments to validate the vision system. Furthermore, since the system was developed for the real-time application, all the validations were performed in a real-time environment.

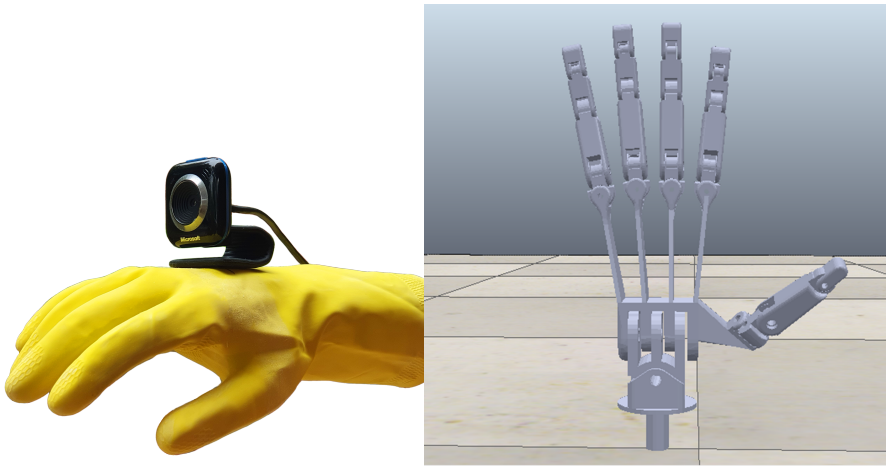
For the experiment. participants were given a latex glove with a camera mounted between knuckles and the wrist (see Figure 6.1a). When the user's hand reaches to the object to grasp, camera detects the object and produces the probability matrix. The sEMG signals of six muscles were collected and classify during the experiment. The experiment algorithm is same as the algorithm discussed in the section 5.3 but system keeps the EMG classification results to compare with hybrid system results. Furthermore, experiment algorithm calculates the Bayesian fusion in addition to majority vote to compare their performances.

In order to visually validate the performances, a computer simulation was used. A prosthetic hand capable of finger flexion-extension and abduction adduction was designed in the SolidWork 2017 platform and used to simulate the grasping patterns in the V-rep environment. The following sections discuss the simulation details and the validation process of each subsystem.

6.1 Grasping pattern simulation

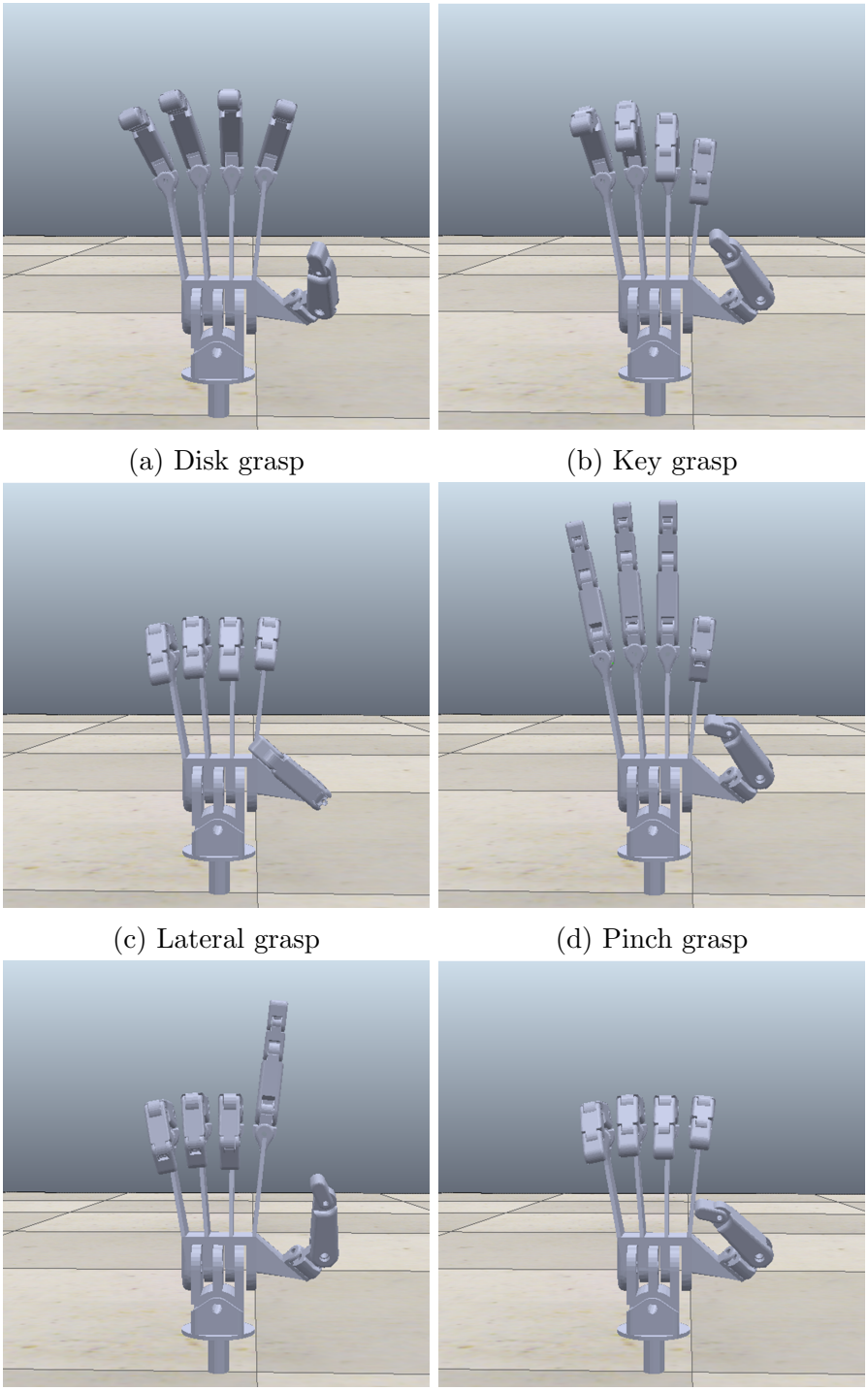
Due to the lack of resources, the physical demonstration was not possible. Instead of that, real-time computer simulation was performed to verify the system visually. It further helps the participants when performing grasping patterns for the validations. The V-rep 3.6.2 PRO Education version (Coppelia Robotics AG, Switzerland) was used since it has the proper libraries to connect with Matlab 2019. The finger joint arrangements of 6 grasping patterns were pre-programmed. Once the predicted grasping patterns are received from the Matlab program, V-rep simulates the hand accordingly.

The finger arrangement is not anatomically accurate, but it gives the idea of the grasp. Note that the finger arrangement of the *Lateral* and *Power* grasps is similar but position of the thumb is different (see the Figure 6.2c and Figure 6.2f). It is difficult to arrange the anatomically accurate grasping poses using the designed hand, but they are noticeable when the real-time simulation happens. Figure 6.1b illustrates the resting hand pose. When the user does not perform any grasping pattern, the system identifies it as a resting position.



(a) (b)

Figure 6.1: (a) Eye-in-hand camera setup for real-time experiments. (b) Normal or the rest position of the simulated hand.



(a) Disk grasp (b) Key grasp

(c) Lateral grasp (d) Pinch grasp

(e) Point grasp (f) Power grasp

Figure 6.2: Simulated grasping pattern.

6.2 Validation of the EMG system

As mentioned previously, all the validations were carried out in a real-time environment. The validation of the EMG system was used to compare the performance of the fusion system. These validation results were used as a benchmark for the complete system validation. A simulation environment was used to visualize the predicted grasping pattern. The participant was prepared as per the discussed in the experiment section 3.1.4 and instructed to grasp the objects using given grasping patterns. Since the participant performed only given grasping pattern (True Positive) at a time, the true positive rate was calculated. As long as the participant executes the grasping action, the system tries to predict it. Figure 6.3 depicts the experiment system overview of the real-time validation. The camera fee

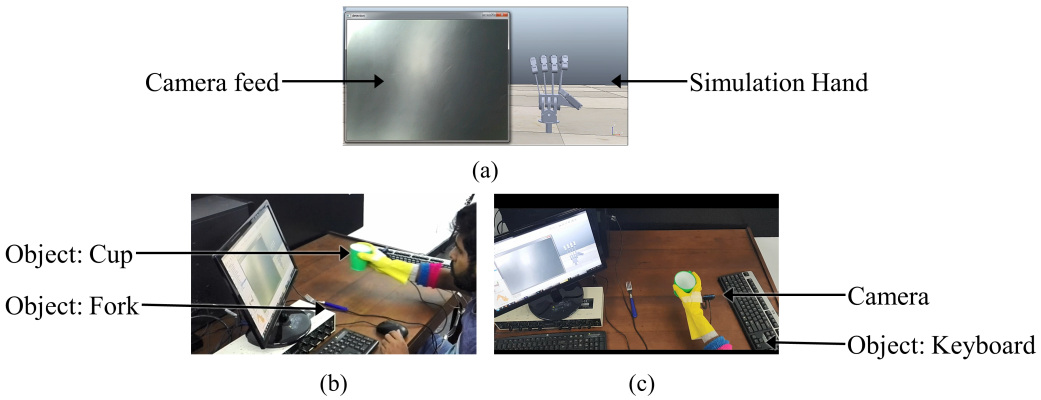


Figure 6.3: V-rep simulates the lateral grasp. (a)v-rep simulation with the camera feedback (left window). (b) The isometric view of the experiment. (c) Plan view of the experiment.

Table 6.1 reports the real-time prediction accuracy of the grasping patterns when no visual feedback is given. A single grasping pattern was executed 6 times (trials), and the approximate time for a single trial was 2-3s. The corresponding prediction count is reported in the table. The *Rest* was excluded from the prediction count.

Table 6.1: Online performance of EMG classification without the fusion system . The mean of the six trials were tabulated with respective *standard deviations*.

Grasping Pattern	**mTPR (%)**	**PP/TP[1]**
Disk	77.49(10.98)	498/637
Key	13.82(8.76)	118/770
Lateral	43.06(19.89)	304/560
Pinch	8.87(14.25)[2]	126/1193
Point	18.78(1.71)	127/530
Power	7.70(9.16)[2]	113/1194

[1] PP-predicted positives, TP-true positives

[2] TPR results were skewed to left

The results in Table 6.1 were produced by the optimized ANN model. The output of the model was not fused with the visual feedback. However, the result is stabled by the MV method. Even though that, most of the results are unacceptable. Only the disk pattern was classified with considerable accuracy. Figure 6.4 illustrates the real-time experiment of the Lateral grasp prediction. According to the illustration, EMG classification alone did not produce acceptable output. The output was varied throughout the classes. This is mainly due to the stochastic behavior of the EMG signals, and the quality of the signal changes with time.

Furthermore, the computation time was evaluated. All the models were considered for evaluation. The results revealed that for a single classification, the model takes less than 12 ms. This time only accounted for the prediction and feature calculation. A 100 ms window is required to collect the data for features. Therefore total time would be less than 120 ms to produce the results.

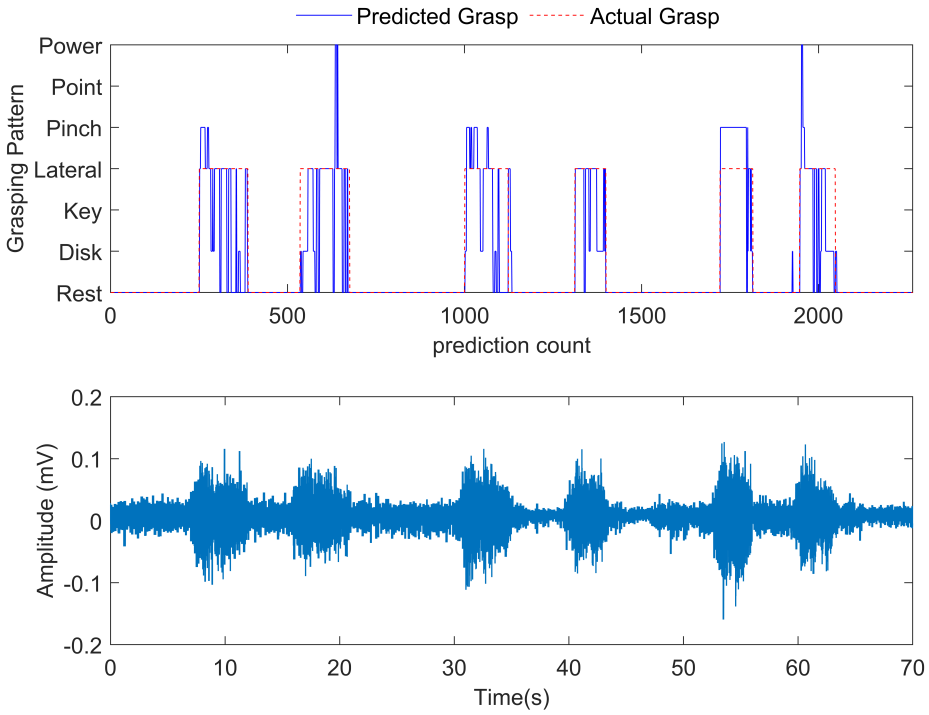


Figure 6.4: Examples of the prediction errors of Lateral grasp when no fusion was utilized. The object used was the cup.

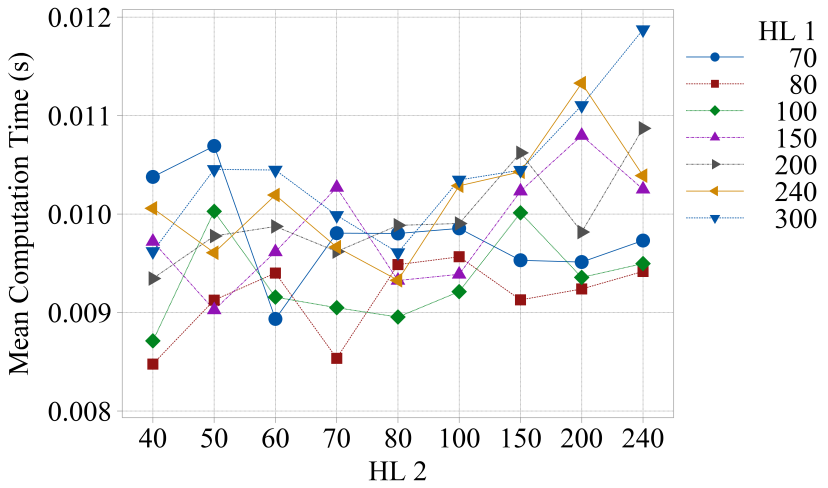


Figure 6.5: Model computation time for different hidden layer configurations

6.3 Validation of the vision system

Several researchers already validate the YOLOv3. The mean Average Precision was the evaluation metrics (see Figure 6.6). The COCO dataset was used to evaluate the model. The YOLOv3 has significantly low computation time but high precision. Although these values depend on the hardware (GPU and memory) used in the evaluated system, the effect on each algorithm is proportional. The average time to detect an object in a single image in the proposed system was M = 138.98 ms (SD=1.527 ms). Thus, other algorithms may take more than $250ms$.

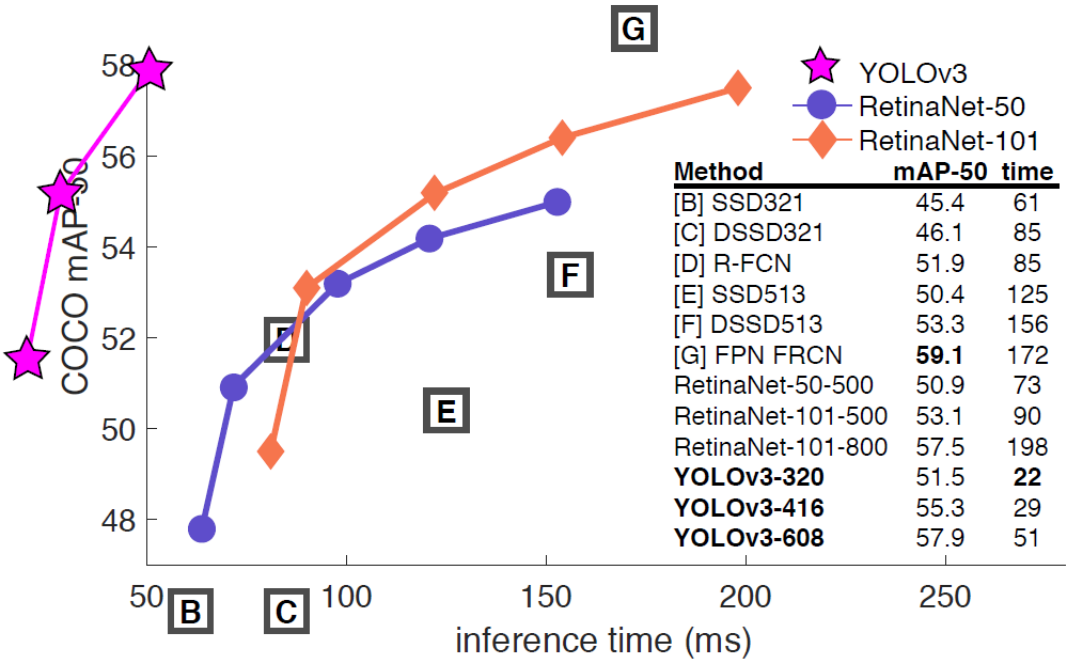


Figure 6.6: Model computation time for different hidden layer configurations [10]

6.4 Validation of proposed hybrid system

The most important validation is the complete system validation. The procedure was similar to the EMG system validation. The Moving Average and the Bayesian fusion post-process were considered during the evaluation. These values were calculated from the same data used for the EMG system validation and the fusion system with MV validation. The Table 6.2 presents the complete system performances.

Table 6.2: Complete fusion system online performances. The mean of the 6 trials were tabulated with respective *standard deviations*.

Grasping Pattern	**mTPR (%) of MV**	**mTPR (%) of BF**
Disk	73.97(13.45)	25.07(7.84)
Key	73.39(5.06)	14.79(4.50)
Lateral	73.94(7.40)	18.03(2.58)
Pinch	62.47(18.52)	17.15(5.23)
Point	87.38(3.15)	21.02(2.84)
Power	71.65(28.41)	16.62(5.54)

The predictions of the fusion system with MV is illustrated in the Figure 6.7. The Bayesian fusion as a post-processing method was not suitable for this application. The results indicate that the performances were drastically reduced when the BF is utilized. The outputs fluctuated even though the 4-window MV was used. However, the fluctuation was comparably reduced by the fusion system. Even most of the outputs were *Rest* state, which can be further minimized by employing slip detection method at the prosthetic hand. The notable improvement was that the reduction in *False Positive* (*wrong classification*).

The complete system takes $M = 393.89$ ms ($SD = 178.23$ ms) to produce output from EMG inputs. This time accounts for the 100 ms window, feature calculation, prediction, object detection, and fusion. However, the computation time can be varied from 125.75 ms to 919.86 ms, depends on the computer processing schedule.

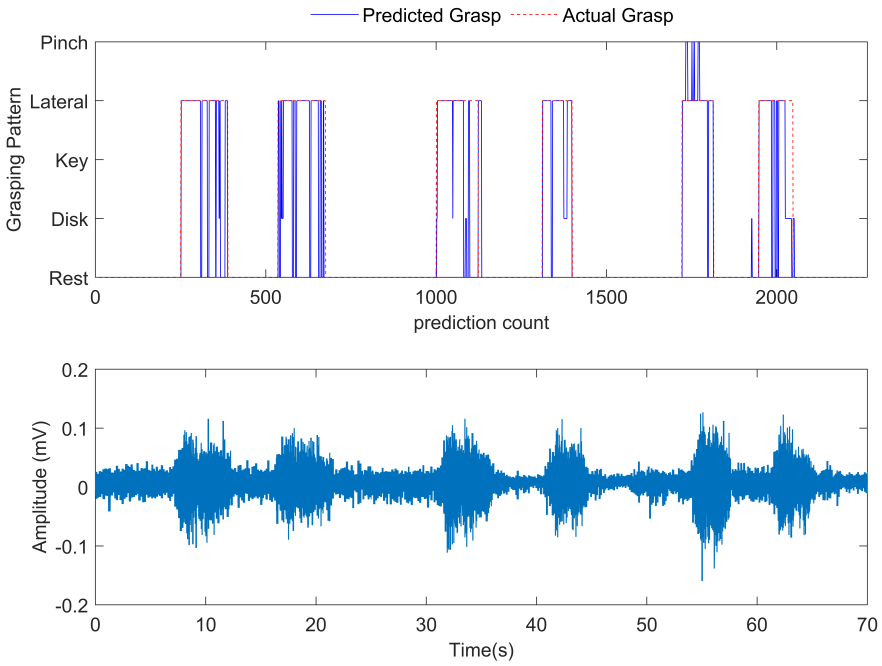


Figure 6.7: Examples of the prediction errors of lateral grasp when no fusion was utilized. The object used was the cup.

6.5 Discussion

The developed sEMG classification system was evaluated in real-time conditions without the fusion system. However, the results indicated that the model could not predict the grasp accurately from new data. The mTPR of the experiment was M = 30.92% (SD = 31.17%). The model exhibited accepted accuracy only for the disk grasp, which is M = 77.49%. The main reason for that would be the stochastic nature of the EMG signal. The model was trained using previously collected data. However, the testing result of the optimized model for the same day collected data was mTPR = 73.1%, which is acceptable. This is the major issue in prosthetic hand controllers.

The proposed fusion system improved the real-time EMG classification performance to M = 72.86% (SD = 17.89%) by fusing the conditional probability matrix with the EMG classification results. The probability matrix corrects the

EMG classification error. This correction can be seen in Figure 6.4 and Figure 6.7. Although false-positive predictions are reduced, few inputs were falsely classified as resting. It is maybe due to the threshold defined in the algorithm. When the EMG signal amplitude drops to less than the threshold during the grasping action, the system identifies it as a resting state. This error can be mitigated by employing additional sensors such as IMU and slip detection method to determine the hand movement and object slipping.

Furthermore, this false resting-state prediction is not continuous. Hence, a mechanical hand cannot perform these sudden transitions because states are changed within 400 ms. It was also noted in the simulation. The torque parameters of the fingers in the simulation were set to resemble the actual hand. Therefore, the abrupt state changes did not simulate. Hence, this false resting prediction will not be a significant issue when an actual prosthetic hand is used.

Chapter 7

CONCLUSION AND FUTURE DIRECTION

7.1 Conclusion

The major problem in the robotic prosthetic hand is the limited functionalities that can perform real-time with an amputee's intention. Electromyography-based and vision-based prosthetic control systems have been introduced to overcome this issue to a certain extent, but only a few could classify the grasping patterns in real-time. Moreover, the number of patterns was limited to five, and most of the other studies focused on hand gestures and wrist movements when the hand is in a static position. However, it is challenging to develop a system that can classify a high number of grasping patterns using only sEMG signals of the remaining muscle at the stump forearm arm. This thesis proposes a vision-EMG fusion method that can classify six frequently used gasping patterns with mTPR of M=72.86% (SD=17.89%) in real-time dynamic conditions. The approach utilizes an optimized sEMG classification system that can classify six frequently used grasping patterns in real-time. The proposed object detection system uses YOLO to detect objects and retrieve a probability matrix containing conditional probabilities of each six grasping patterns for each identifiable object. The Bayesian fusion method (not the post-processing technique) was used to fuse both sEMG and probability matrix outputs from the vision system to determine the final user intention for the detected object. The system continuously predicts the user's intention. When the user does not perform any grasping action, the system automatically identifies it as a *Resting* state. In addition to the classification system,

research work revealed that the sEMG signals depend on the participant's demographic factors such as Age, Gender, and Handedness and can be utilized as features for the classification system.

The approach proposed in this report used the vision to retrieve a probability matrix that is previously defined. Unlike the previous approach, this approach facilitates the user to use multiple grasping patterns for a single object. The validation experiment results indicated that the system could predict the user's intention with considerable accuracy within 400 ms. The results further indicated that the proposed fusion system outperformed the performance of most of previously developed real-time classification systems, which only could classify 4-5 grasping patterns. However, the sEMG classifier was failed to outperform the previous classifiers without the fusion system in real-time conditions. Furthermore, the developed fusion system is robust enough to classify the patterns during the RTG motion. Unlike previous classifiers, the model was trained for dynamic conditions.

7.2 Future direction

This research work is limited to 6 frequently used grasping patterns. Furthermore, the sensory inputs are limited to sEMG and RGB vision. As for the future directions, the number of grasping patterns should be increased with additional wrist movements such as flexion-extension and pronation-supination. In order to achieve that, the classification system may need an ensemble learning method. Furthermore, the fusion system can be further expanded to employ other sensory inputs such as depth sensor, MMG, and IMU. Thus, it could improve the real-time classification performance but should meet the real-time task completion time of 300-400 ms. Finally, the effectiveness of the current systems should be evaluated for the amputees to measure the practical potential of the system.

LIST OF PUBLICATIONS

---

1. **Dulanjana M. Perera** and D. G. K. Madusanka, "Vision-EMG Fusion Method for Real-time Grasping Pattern Classification System," presented at the 7^{th} *Moratuwa Engineering Research Conference (MERCon)*, Moratuwa, Sri Lanka, Jul. 27-28, 2021.

REFERENCES

[1] "bebionic hand," access: 2019-09-14. [Online]. Available: https://www.ottobockus.com/prosthetics/upper-limb-prosthetics/solution-overview/bebionic-hand/

[2] "Össur. life without limitations," access: 2020-03-12. [Online]. Available: https://www.ossur.com/en-us/prosthetics/arms

[3] M. Kazamel and P. P. Warren, "History of electromyography and nerve conduction studies: A tribute to the founding fathers," *Journal of Clinical Neuroscience*, vol. 43, pp. 54–60, 2017.

[4] R. Reiter, "Eine neue elektrokunsthand," *Grenzgebiete der Medizin*, vol. 1, no. 4, pp. 133–135, 1948.

[5] G. Schlesinger, "Der mechanische aufbau der künstlichen glieder," in *Ersatzglieder und Arbeitshilfen*. Springer, 1919, pp. 321–661.

[6] L. J. Hargrove, K. Englehart, and B. Hudgins, "A comparison of surface and intramuscular myoelectric signal classification," *IEEE transactions on biomedical engineering*, vol. 54, no. 5, pp. 847–853, 2007.

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

[8] R. N. Khushaba, S. Kodagoda, M. Takruri, and G. Dissanayake, "Toward improved control of prosthetic fingers using surface electromyogram (emg)

signals," *Expert Systems with Applications*, vol. 39, no. 12, pp. 10 731–10 738, 2012.

[9] J. DeGol, A. Akhtar, B. Manja, and T. Bretl, "Automatic grasp selection using a camera in a hand prosthesis," in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2016, pp. 431–434.

[10] J. Redmon, S. Divvala, R. B. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 779–788, 2016.

[11] L. Weng, "Object Detection Part 4: Fast Detection Models," Dec. 2018. [Online]. Available: https://lilianweng.github.io/2018/12/27/object-detection-part-4.html

[12] S. M. Khan, A. A. Khan, and O. Farooq, "Selection of features and classifiers for emg/eeg based upper limb assistive devices-a review," *IEEE reviews in biomedical engineering*, 2019.

[13] C. L. McDonald, S. Westcott-McCoy, M. R. Weaver, J. Haagsma, and D. Kartin, "Global prevalence of traumatic non-fatal limb amputation," *Prosthetics and orthotics international*, 2020, pMID: 33274665.

[14] L. S. M. Tintle, L. M. F. Baechler, C. G. P. Nanos III, L. J. A. Forsberg, and M. B. K. Potter, "Traumatic and trauma-related amputations: Part ii: Upper extremity and future directions," *JBJS*, vol. 92, no. 18, pp. 2934–2945, 2010.

[15] N. Jiang, S. Dosen, K.-R. Muller, and D. Farina, "Myoelectric control of artificial limbs—is there a need to change focus? [in the spotlight]," *IEEE Signal Processing Magazine*, vol. 29, no. 5, pp. 152–150, 2012.

[16] W. Dahlheim, "Pressluft hand für kreigsbeschädigte industriearbeiter z. komprimierte und flüssige gase," *German Patent*, 1915.

[17] K. Fukunaga, *Introduction to statistical pattern recognition.* Elsevier, 2013.

[18] Y. Jiang, C. Chen, X. Zhang, W. Zhou, C. Chen, and S. Lemos, "Eeg-based hand motion pattern recognition using deep learning network algorithms," in *Proceedings of the 2020 9th International Conference on Computing and Pattern Recognition*, 2020, pp. 73–79.

[19] M.-K. Liu, Y.-T. Lin, Z.-W. Qiu, C.-K. Kuo, and C.-K. Wu, "Hand gesture recognition by a mmg-based wearable device," *IEEE Sensors Journal*, vol. 20, no. 24, pp. 14 703–14 712, 2020.

[20] P. Kaczmarek, T. Mankowski, and J. Tomczynski, "Towards sensor position-invariant hand gesture recognition using a mechanomyographic interface," *2017 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)*, pp. 53–58, 2017.

[21] G. Li, A. E. Schultz, and T. A. Kuiken, "Quantifying pattern recognition—based myoelectric control of multifunctional transradial prostheses," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 18, no. 2, pp. 185–192, 2010.

[22] A. Waris, I. Niazi, M. Jamil, K. Englehart, W. Jensen, and E. Kamavuako, "Multiday evaluation of techniques for emg-based classification of hand motions," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, pp. 1526–1534, 2019.

[23] J. He, D. Zhang, N. Jiang, X. Sheng, D. Farina, and X. Zhu, "User adaptation in long-term, open-loop myoelectric training: implications for emg pattern recognition in prosthesis control." *Journal of neural engineering*, vol. 12 4, p. 046005, 2015.

[24] E. Scheme, K. Biron, and K. Englehart, "Improving myoelectric pattern recognition positional robustness using advanced training protocols," *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 4828–4831, 2011.

[25] C. Cipriani, R. Sassu, M. Controzzi, and M. Carrozza, "Influence of the weight actions of the hand prosthesis on the performance of pattern recognition based myoelectric control: Preliminary study," *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 1620–1623, 2011.

[26] M. Simão, N. Mendes, O. Gibaru, and P. Neto, "A review on electromyography decoding and pattern recognition for human-machine interaction," *IEEE Access*, vol. 7, pp. 39 564–39 582, 2019.

[27] N. Parajuli, N. Sreenivasan, P. Bifulco, M. Cesarelli, S. Savino, V. Niola, D. Esposito, T. J. Hamilton, G. R. Naik, U. Gunawardana *et al.*, "Real-time emg based pattern recognition control for hand prostheses: a review on existing methods, challenges and future implementation," *Sensors*, vol. 19, no. 20, p. 4596, 2019.

[28] D. Madusanka, L. Wijayasingha, R. Gopura, Y. Amarasinghe, and G. Mann, "A review on hybrid myoelectric control systems for upper limb prosthesis," in *2015 Moratuwa Engineering Research Conference (MERCon)*. IEEE, 2015, pp. 136–141.

[29] M. Štrbac and M. Marković, "Stereovision system for estimation of the grasp type for electrotherapy," *Serbian Journal of Electrical Engineering*, vol. 8, no. 1, pp. 17–25, 2011.

[30] o. Klisić, M. Kostić, S. Došen, and D. B. Popović, "Control of prehension for the transradial prosthesis: natural-like image recognition system," *Journal of Automatic Control*, vol. 19, no. 1, pp. 27–31, 2009.

[31] S. Došen, C. Cipriani, M. Kostić, M. Controzzi, M. C. Carrozza, and D. B. Popović, "Cognitive vision system for control of dexterous prosthetic hands: experimental evaluation," *Journal of neuroengineering and rehabilitation*, vol. 7, no. 1, p. 42, 2010.

[32] S. Došen and D. B. Popović, "Transradial prosthesis: artificial vision for control of prehension," *Artificial organs*, vol. 35, no. 1, pp. 37–48, 2011.

[33] Q.-C. Mao, H.-M. Sun, Y.-B. Liu, and R.-S. Jia, "Mini-yolov3: real-time object detector for embedded applications," *IEEE Access*, vol. 7, pp. 133 529–133 538, 2019.

[34] D. T. Andrade, A. Ishikawa, A. D. Munoz, and E. Rohmer, "A hybrid approach for the actuation of upper limb prostheses based on computer vision," in *2017 Latin American Robotics Symposium (LARS) and 2017 Brazilian Symposium on Robotics (SBR)*. IEEE, 2017, pp. 1–6.

[35] J. Fajardo, V. Ferman, A. Muñoz, D. Andrade, A. R. Neto, and E. Rohmer, "User-prosthesis interface for upper limb prosthesis based on object classification," in *2018 Latin American Robotic Symposium, 2018 Brazilian Symposium on Robotics (SBR) and 2018 Workshop on Robotics in Education (WRE)*. IEEE, 2018, pp. 390–395.

[36] C. Shi, D. Yang, J. Zhao, and H. Liu, "Computer vision-based grasp pattern recognition with application to myoelectric control of dexterous hand prosthesis," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, pp. 2090–2099, 2020.

[37] J. Doyon, A. M. Owen, M. Petrides, V. Sziklas, and A. C. Evans, "Functional anatomy of visuomotor skill learning in human subjects examined with positron emission tomography," *European Journal of Neuroscience*, vol. 8, no. 4, pp. 637–648, 1996.

[38] M. K. Rand, M. Lemay, L. M. Squire, Y. P. Shimansky, and G. E. Stelmach, "Role of vision in aperture closure control during reach-to-grasp movements," *Experimental Brain Research*, vol. 181, no. 3, pp. 447–460, 2007.

[39] S. Jackson, G. Jackson, and J. Rosicky, "Are non-relevant objects represented in working memory? the effect of non-target objects on reach and

grasp kinematics," *Experimental Brain Research*, vol. 102, no. 3, pp. 519–530, 1995.

[40] M. Garber, "The perfect, 3,000-year-old toe: A brief history of prosthetic limbs," 2013, accessed: 2020-02-23. [Online]. Available: https://www.theatlantic.com/technology/archive/2013/11/the-perfect-3-000-year-old-toe-a-brief-history-of-prosthetic-limbs/281653/

[41] "Artificial arm with four tool attachments, england, 1945-1955 — science museum group collection," accessed: 2020-02-23. [Online]. Available: https://collection.sciencemuseumgroup.org.uk/objects/ co479542/artificial-arm-with-four-tool-attachments-england-1945-1955-artificial-arm

[42] "Hanger-simpson wooden hand with all fingers articular, 1948," accessed: 2020-02-25. [Online]. Available: https://collection.sciencemuseumgroup.org.uk/objects/co477906/hanger-simpson-wooden-hand-with-all-fingers-articu-prosthetic-hand

[43] R. H. Meier, D. J. Atkins *et al.*, *Functional restoration of adults and children with upper extremity amputation*. Demos Medical Pub., 2004.

[44] D. Masson, G. Grove, J. Morley, and M. Morris, *Macmillan's Magazine*. Macmillan and Company, 1861, vol. 3.

[45] K. J. Zuo and J. L. Olson, "The evolution of functional hand replacement: From iron prostheses to hand transplantation," *Plastic Surgery*, vol. 22, no. 1, pp. 44–51, 2014.

[46] M. LeBlanc, "Current evaluation of hydraulics to replace the cable force transmission system for body-powered upper-limb prostheses," *Assistive Technology*, vol. 2, no. 3, pp. 101–108, 1990.

[47] S. Alderson, "The electric arm," in *Human Limbs and Their Substitutes*, P. William, P. Klopsteg, and Ed, Eds. McGraw-Hill, 1954, ch. 13.

[48] A. Heather and T. A. Smith, "Helping hand, a hydraulically operated mech anical hand," *Orthopedic and Prosthetic Appliance Journal*, vol. 14, 1960.

[49] D. Lewis, "Hydraulics for prosthetic devices," *Orthotics and Prosthetics*, vol. 22, pp. 23–28, 1968.

[50] "Co2 gas-powered artificial arms, roehampton, england, 1963," accessed: 2020-02-24. [Online]. Available: https://collection.sciencemuseumgroup.org.uk/objects/co476753/co2-gas-powered-artificial-arms-roehampton-england-1963-artificial-arm

[51] "Pair of co2 powered upper limbs for a 10 year old," accessed: 2020-02-24. [Online]. Available: https://collection.sciencemuseumgroup.org.uk/objects/co476750/pair-of-co2-powered-upper-limbs-for-a-10-year-old-prosthetic-arm

[52] "Artificial hand prosthetic device powered by carbo," accessed: 2020-02-24. [Online]. Available: https://collection.sciencemuseumgroup.org.uk/objects/co203089/artificial-hand-prosthetic-device-powered-by-carbo

[53] H. Piper, *Elektrophysiologie menschlicher muskeln*. Springer-Verlag, 2013.

[54] E. H. Du Bois-Reymond, *Gesammelte Abhandlungen zur allgemeinen Muskel-und Nervenphysik*. Veit, 1877, vol. 2.

[55] H. C. Ørsted and K. L. Rahbek, *Experimenta circa effectum conflictus electrici in acum magneticam: med oversættelse til dansk fra Hesperus bind III pag. 312-21, udgivet af Knud Lyhne Rahbek, 1820*. Københavns Universitets Fond til Tilvejebringelse af Læremidler, 1970.

[56] E. D. Adrian and D. W. Bronk, "The discharge of impulses in motor nerve fibres: Part ii. the frequency of discharge in reflex and voluntary contractions," *The Journal of physiology*, vol. 67, no. 2, pp. 9–151, 1929.

[57] F. Buchthal and T. Péterfi, "Die potentialdifferenzen einzelner lebender muskelfasern im ruhezustand und während der kontraktion," *Pflüger's Archiv für die gesamte Physiologie des Menschen und der Tiere*, vol. 234, no. 1, pp. 527–543, 1934.

[58] F. J. Bonner Jr and A. B. Devleschoward, "Aaem minimonograph# 45: the early development of electromyography," *Muscle & Nerve: Official Journal of the American Association of Electrodiagnostic Medicine*, vol. 18, no. 8, pp. 825–833, 1995.

[59] N. Berger and C. R. Huppert, "The use of electrical and mechanical muscular forces for the control of an electrical prosthesis." *The American journal of occupational therapy: official publication of the American Occupational Therapy Association*, vol. 6, no. 3, p. 110, 1952.

[60] C. Battye, A. Nightingale, and J. Whillis, "The use of myo-electric currents in the operation of prostheses," *The Journal of bone and joint surgery. British volume*, vol. 37, no. 3, pp. 506–510, 1955.

[61] A. Kobrinskii, S. Bolkhovitin, L. Voskoboinikowa, D. Ioffe, E. Polyan, B. Popov, Y. Slavutski, A. Sysin, and Y. Yakobson, "Problems of bioelectric control, automatic and remote control," in *Proceedings of 1st IFAC International Congress, Moscow*, 1960, pp. 619–623.

[62] E. Wilms and L. Nader, "Die technik der vaduzer hand," *Orthopädie Technik*, vol. 2, no. 7, p. 7, 1951.

[63] A. Bottomley, "Myo-electric control of powered prostheses," *The Journal of bone and joint surgery. British volume*, vol. 47, no. 3, pp. 411–415, 1965.

[64] M. Rikich, "The belgrade hand prosthesis." *The Basic Problems of Prehension, Movement and Control Artificial Limbs*, vol. 183, no. sJ, pp. 60–67, 1969.

[65] P. Herberts and I. Petersen, "Possibilities for control of powered devices by myoelectric signals." *Scandinavian journal of rehabilitation medicine*, vol. 2, no. 4, pp. 164–170, 1970.

[66] R. W. Mann, "Cybernetic limb prosthesis: The alza distinguished lecture," *Annals of Biomedical Engineering*, vol. 9, no. 1, pp. 1–43, 1981.

[67] D. Graupe, J. Magnussen, and A. Beex, "A microprocessor system for multifunctional control of upper-limb prostheses via myoelectric signal identification," *IEEE Transactions on automatic control*, vol. 23, no. 4, pp. 538–544, 1978.

[68] G. N. Saridis and T. P. Gootee, "Emg pattern analysis and classification for a prosthetic arm," *IEEE Transactions on Biomedical Engineering*, vol. 6, pp. 403–412, 1982.

[69] C. J. Perera, I. Naotunna, C. Sadaruwan, R. A. R. C. Gopura, and T. D. Lalitharatne, "Ssvep based bmi for a meal assistance robot," in *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2016, pp. 002 295–002 300.

[70] M. A. Oskoei and H. Hu, "Myoelectric control systems—a survey," *Biomedical signal processing and control*, vol. 2, no. 4, pp. 275–294, 2007.

[71] J. Lara, N. Paskaranandavadivel, and L. K. Cheng, "Hd-emg electrode count and feature selection influence on pattern-based movement classification accuracy," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2020, pp. 4787–4790.

[72] B. Yu, X. Zhang, L. Wu, X. Chen, and X. Chen, "A novel postprocessing method for robust myoelectric pattern-recognition control through movement pattern transition detection," *IEEE Transactions on Human-Machine Systems*, vol. 50, pp. 32–41, 2020.

[73] G. D. Gargiulo, P. Bifulco, M. Cesarelli, A. Fratini, and M. Romano, "Problems in assessment of novel biopotential front-end with dry electrode: A brief review," *Machines*, vol. 2, no. 1, pp. 87–98, 2014.

[74] H. J. Hermens, B. Freriks, C. Disselhorst-Klug, and G. Rau, "Development of recommendations for semg sensors and sensor placement procedures," *Journal of electromyography and Kinesiology*, vol. 10, no. 5, pp. 361–374, 2000.

[75] R. L. Drake, W. Vogl, A. W. M. Mitchell, and H. Gray, *Gray's anatomy for students*. Philadelphia: Elsevier/Churchill Livingstone, 2005.

[76] D. Zhang, X. Zhao, J. Han, and Y. Zhao, "A comparative study on pca and lda based emg pattern recognition for anthropomorphic robotic hand," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2014, pp. 4850–4855.

[77] G. Hägg and E. Milerad, "Forearm extensor and flexor muscle exertion during simulated gripping work—an electromyographic study," *Clinical biomechanics*, vol. 12, no. 1, pp. 39–43, 1997.

[78] I. Batzianoulis, N. E. Krausz, A. M. Simon, L. Hargrove, and A. Billard, "Decoding the grasping intention from electromyography during reaching motions," *Journal of neuroengineering and rehabilitation*, vol. 15, no. 1, p. 57, 2018.

[79] S. Pancholi and A. Joshi, "Improved classification scheme using fused wavelet packet transform based features for intelligent myoelectric prostheses," *IEEE Transactions on Industrial Electronics*, vol. 67, pp. 8517–8525, 2020.

[80] I. Batzianoulis, S. El-Khoury, E. Pirondini, M. Coscia, S. Micera, and A. Billard, "Emg-based decoding of grasp gestures in reaching-to-grasping motions," *Robotics and Autonomous Systems*, vol. 91, pp. 59–70, 2017.

[81] A. yuan Zhang, Q. Li, N. Gao, L. Wang, and Y. Wu, "Influence of different feature selection methods on emg pattern recognition," *2019 IEEE International Conference on Mechatronics and Automation (ICMA)*, pp. 880–885, 2019.

[82] S. Ma, B. Lv, C. Lin, X. Sheng, and X. Zhu, "Emg signal filtering based on variational mode decomposition and sub-band thresholding," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, pp. 47–58, 2021.

[83] A. Phinyomark, E. Campbell, and E. Scheme, "Surface electromyography (emg) signal processing, classification, and practical considerations," 2020, pp. 3–29.

[84] R. H. Chowdhury, M. B. Reaz, M. A. B. M. Ali, A. A. Bakar, K. Chellappan, and T. G. Chang, "Surface electromyography signal processing and classification techniques," *Sensors*, vol. 13, no. 9, pp. 12 431–12 466, 2013.

[85] A. Phinyomark, P. Phukpattaranont, and C. Limsakul, "Feature reduction and selection for emg signal classification," *Expert systems with applications*, vol. 39, no. 8, pp. 7420–7431, 2012.

[86] A. Phinyomark, R. N Khushaba, and E. Scheme, "Feature extraction and selection for myoelectric control based on wearable emg sensors," *Sensors*, vol. 18, no. 5, p. 1615, 2018.

[87] H. Chen, Y. Zhang, Z. Zhang, Y. Fang, H. Liu, and C. Yao, "Exploring the relation between emg sampling frequency and hand motion recognition accuracy," in *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2017, pp. 1139–1144.

[88] G. Li, Y. Li, L. Yu, and Y. Geng, "Conditioning and sampling issues of emg signals in motion recognition of multifunctional myoelectric prostheses," *Annals of biomedical engineering*, vol. 39, no. 6, pp. 1779–1787, 2011.

[89] R. Menon, G. D. Caterina, H. Lakany, L. Petropoulakis, B. Conway, and J. Soraghan, "Study on interaction between temporal and spatial information in classification of emg signals for myoelectric prostheses," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, pp. 1832–1842, 2017.

[90] D. Farina and R. Merletti, "Comparison of algorithms for estimation of emg variables during voluntary isometric contractions," *Journal of Electromyography and Kinesiology*, vol. 10, no. 5, pp. 337–349, 2000.

[91] K. Englehart and B. Hudgins, "A robust, real-time control scheme for multifunction myoelectric control," *IEEE transactions on biomedical engineering*, vol. 50, no. 7, pp. 848–854, 2003.

[92] J. Liu, "Adaptive myoelectric pattern recognition toward improved multifunctional prosthesis control," *Medical engineering & physics*, vol. 37, no. 4, pp. 424–430, 2015.

[93] B. Hudgins, P. Parker, and R. N. Scott, "A new strategy for multifunction myoelectric control," *IEEE Transactions on Biomedical Engineering*, vol. 40, no. 1, pp. 82–94, 1993.

[94] S.-H. Park and S.-P. Lee, "Emg pattern recognition based on artificial intelligence techniques," *IEEE transactions on Rehabilitation Engineering*, vol. 6, no. 4, pp. 400–405, 1998.

[95] M. Tavakolan, Z. G. Xiao, J. Webb, and C. Menon, "Emg processing for classification of hand gestures and regression of wrist torque," in *2012 4th IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob)*. IEEE, 2012, pp. 1770–1775.

[96] A. H. Al-Timemy, G. Bugmann, J. Escudero, and N. Outram, "Classification of finger movements for the dexterous hand prosthesis control with surface electromyography," *IEEE journal of biomedical and health informatics*, vol. 17, no. 3, pp. 608–618, 2013.

[97] J. Liu, "Feature dimensionality reduction for myoelectric pattern recognition: A comparison study of feature selection and feature projection methods," *Medical engineering & physics*, vol. 36, no. 12, pp. 1716–1720, 2014.

[98] A. Turlapaty and B. Gokaraju, "Feature analysis for classification of physical actions using surface emg data," *IEEE Sensors Journal*, vol. 19, pp. 12 196–12 204, 2019.

[99] A. M. Martínez and A. C. Kak, "Pca versus lda," *IEEE transactions on pattern analysis and machine intelligence*, vol. 23, no. 2, pp. 228–233, 2001.

[100] J. Too, A. R. Abdullah, N. Mohd Saad, and W. Tee, "Emg feature selection and classification using a pbest-guide binary particle swarm optimization," *Computation*, vol. 7, no. 1, p. 12, 2019.

[101] C. Meeker, S. Park, L. Bishop, J. Stein, and M. Ciocarlie, "Emg pattern classification to control a hand orthosis for functional grasp assistance after stroke," in *2017 international conference on rehabilitation robotics (ICORR)*. IEEE, 2017, pp. 1203–1210.

[102] S. Ahmad, A. Ishak, and S. Ali, "Review of electromyographic control systems based on pattern recognition," in *5th Kuala Lumpur International Conference on Biomedical Engineering 2011*. Springer, 2011, pp. 556–559.

[103] L. Pan, D. Zhang, J. Liu, X. Sheng, and X. Zhu, "Continuous estimation of finger joint angles under different static wrist motions from surface emg signals," *Biomedical Signal Processing and Control*, vol. 14, pp. 265–271, 2014.

[104] S. Pancholi and A. M. Joshi, "Electromyography-based hand gesture recognition system for upper limb amputees," *IEEE Sensors Letters*, vol. 3, no. 3, pp. 1–4, 2019.

[105] B. Saeed, M. Z. ur Rehman, S. O. Gilani, F. Amin, A. Waris, M. Jamil, and M. Shafique, "Leveraging ann and lda classifiers for characterizing differ-

ent hand movements using emg signals," *Arabian Journal for Science and Engineering*, vol. 46, pp. 1761–1769, 2020.

[106] D. C. Dening, F. G. Gray, and R. M. Haralick, "Prosthesis control using a nearest neighbor electromyographic pattern classifier," *IEEE Transactions on Biomedical Engineering*, no. 6, pp. 356–360, 1983.

[107] C. I. Christodoulou, P. A. Kaplanis, V. Murray, M. S. Pattichis, C. S. Pattichis, and T. Kyriakides, "Multi-scale am–fm analysis for the classification of surface electromyographic signals," *Biomedical Signal Processing and Control*, vol. 7, no. 3, pp. 265–269, 2012.

[108] R. J. Oweis, R. Rihani, and A. Alkhawaja, "Ann-based emg classification for myoelectric control," *International Journal of Medical Engineering and Informatics*, vol. 6, no. 4, pp. 365–380, 2014.

[109] B. Karlik, M. O. Tokhi, and M. Alci, "A fuzzy clustering neural network architecture for multifunction upper-limb prosthesis," *IEEE Transactions on Biomedical Engineering*, vol. 50, no. 11, pp. 1255–1261, 2003.

[110] A. D. Chan and K. B. Englehart, "Continuous myoelectric control for powered prostheses using hidden markov models," *IEEE Transactions on Biomedical Engineering*, vol. 52, no. 1, pp. 121–124, 2004.

[111] X. Zhai, B. Jelfs, R. H. Chan, and C. Tin, "Self-recalibrating surface emg pattern recognition for neuroprosthesis control based on convolutional neural network," *Frontiers in neuroscience*, vol. 11, p. 379, 2017.

[112] F. Riillo, L. R. Quitadamo, F. Cavrini, E. Gruppioni, C. A. Pinto, N. C. Pastò, L. Sbernini, L. Albero, and G. Saggio, "Optimization of emg-based hand gesture recognition: Supervised vs. unsupervised data preprocessing on healthy subjects and transradial amputees," *Biomedical Signal Processing and Control*, vol. 14, pp. 117–125, 2014.

[113] M. Zardoshti-Kermani, B. C. Wheeler, K. Badie, and R. M. Hashemi, "Emg feature evaluation for movement control of upper extremity prostheses," *IEEE Transactions on Rehabilitation Engineering*, vol. 3, no. 4, pp. 324–333, 1995.

[114] A. Ameri, M. Akhaee, E. Scheme, and K. Englehart, "A deep transfer learning approach to reducing the effect of electrode shift in emg pattern recognition-based control," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, pp. 370–379, 2020.

[115] D. Huang and B. Chen, "Surface emg decoding for hand gestures based on spectrogram and cnn-lstm," *2019 2nd China Symposium on Cognitive Computing and Hybrid Intelligence (CCHI)*, pp. 123–126, 2019.

[116] M. Atzori, M. Cognolato, and H. Müller, "Deep learning with convolutional neural networks applied to electromyography data: A resource for the classification of movements for prosthetic hands," *Frontiers in Neurorobotics*, vol. 10, 2016.

[117] T. R. Farrell *et al.*, "A comparison of the effects of electrode implantation and targeting on pattern classification accuracy for prosthesis control," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 9, pp. 2198–2211, 2008.

[118] M. Gandolla, S. Ferrante, G. Ferrigno, D. Baldassini, F. Molteni, E. Guanziroli, M. Cotti Cottini, C. Seneci, and A. Pedrocchi, "Artificial neural network emg classifier for functional hand grasp movements prediction," *Journal of International Medical Research*, vol. 45, no. 6, pp. 1831–1847, 2017.

[119] M. Jochumsen, A. Waris, and E. N. Kamavuako, "The effect of arm position on classification of hand gestures with intramuscular emg," *Biomedical Signal Processing and Control*, vol. 43, pp. 1–8, 2018.

[120] G. Kanitz, C. Cipriani, and B. B. Edin, "Classification of transient myoelectric signals for the control of multi-grasp hand prostheses," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 9, pp. 1756–1764, 2018.

[121] E. Mastinu, J. Ahlberg, E. Lendaro, L. Hermansson, B. Håkansson, and M. Ortiz-Catalan, "An alternative myoelectric pattern recognition approach for the control of hand prostheses: A case study of use in daily life by a dysmelia subject," *IEEE journal of translational engineering in health and medicine*, vol. 6, pp. 1–12, 2018.

[122] K. Xu, W. Guo, L. Hua, X. Sheng, and X. Zhu, "A prosthetic arm based on emg pattern recognition," in *2016 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. IEEE, 2016, pp. 1179–1184.

[123] A. M. Simon, L. J. Hargrove, B. A. Lock, and T. A. Kuiken, "A decision-based velocity ramp for minimizing the effect of misclassifications during real-time pattern recognition control," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 8, pp. 2360–2368, 2011.

[124] Y. Huang, K. B. Englehart, B. Hudgins, and A. D. Chan, "A gaussian mixture model based classification scheme for myoelectric control of powered upper limb prostheses," *IEEE Transactions on Biomedical Engineering*, vol. 52, no. 11, pp. 1801–1811, 2005.

[125] E. J. Scheme, K. B. Englehart, and B. S. Hudgins, "Selective classification for improved robustness of myoelectric control under nonideal conditions," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 6, pp. 1698–1705, 2011.

[126] A. H. Al-Timemy, R. Khushaba, and J. Escudero, "A comparison of post-processing techniques on the performance of emg based pattern recognition system for the transradial amputees," *2016 3rd Middle East Conference on Biomedical Engineering (MECBME)*, pp. 46–49, 2016.

[127] S. Ren, K. He, R. B. Girshick, and J. Sun, "Faster R-CNN: towards real-time object detection with region proposal networks," *CoRR*, vol. abs/1506.01497, 2015. [Online]. Available: http://arxiv.org/abs/1506.01497

[128] C.-Y. Fu, W. Liu, A. Ranga, A. Tyagi, and A. C. Berg, "Dssd: Deconvolutional single shot detector," *arXiv preprint arXiv:1701.06659*, 2017.

[129] Z. Shen, Z. Liu, J. Li, Y.-G. Jiang, Y. Chen, and X. Xue, "Dsod: Learning deeply supervised object detectors from scratch," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 1919–1927.

[130] Z.-Q. Zhao, P. Zheng, S.-t. Xu, and X. Wu, "Object detection with deep learning: A review," *IEEE transactions on neural networks and learning systems*, vol. 30, no. 11, pp. 3212–3232, 2019.

[131] S. Jafari and R. Jarvis, "Eye-to-hand coordination: a coarse-calibrated approach," in *Australasian Conf on Robotics and Automation*, 2004.

[132] K. Ramnath, "A framework for robotic vision-based grasping task," *Project report, The Robotics Institute, Carnegie Mellon University*, 2004.

[133] M. Gardner, R. Woodward, R. Vaidyanathan, E. Bürdet, and B. C. Khoo, "An unobtrusive vision system to reduce the cognitive burden of hand prosthesis control," in *2014 13th International Conference on Control Automation Robotics & Vision (ICARCV)*. IEEE, 2014, pp. 1279–1284.

[134] C. Copot, A. Burlacu, and C. Lazar, "Image moments based visual control algorithm for servoing systems," in *2009 IEEE 5th International Conference on Intelligent Computer Communication and Processing*. IEEE, 2009, pp. 157–160.

[135] Y. He, R. Kubozono, O. Fukuda, N. Yamaguchi, and H. Okumura, "Vision-based assistance for myoelectric hand control," *IEEE Access*, vol. 8, pp. 201 956–201 965, 2020.

[136] Y. He, R. Shima, O. Fukuda, N. Bu, N. Yamaguchi, and H. Okumura, "Development of distributed control system for vision-based myoelectric prosthetic hand," *IEEE Access*, vol. 7, pp. 54 542–54 549, 2019.

[137] J. L. Betthauser, C. Hunt, L. Osborn, M. Masters, G. Levay, R. Kaliki, and N. Thakor, "Limb position tolerant pattern recognition for myoelectric prosthesis control with adaptive sparse representations from extreme learning," *IEEE Transactions on Biomedical Engineering*, vol. 65, pp. 770–778, 2018.

[138] J. P. Ioannidis, "Why most published research findings are false," *PLoS medicine*, vol. 2, no. 8, p. e124, 2005.

[139] S. Betti, G. Zani, S. Guerra, U. Castiello, and L. Sartori, "Reach-to-grasp movements: a multimodal techniques study," *Frontiers in psychology*, vol. 9, p. 990, 2018.

[140] M. K. Rand, Y. Shimansky, A. B. Hossain, and G. E. Stelmach, "Quantitative model of transport-aperture coordination during reach-to-grasp movements," *Experimental brain research*, vol. 188, no. 2, pp. 263–274, 2008.

[141] J. Z. Zheng, S. De La Rosa, and A. M. Dollar, "An investigation of grasp type and frequency in daily household and machine shop tasks," in *2011 IEEE International Conference on Robotics and Automation*. IEEE, 2011, pp. 4169–4175.

[142] A. Ghasemi and S. Zahediasl, "Normality tests for statistical analysis: a guide for non-statisticians," *International journal of endocrinology and metabolism*, vol. 10, no. 2, p. 486, 2012.

[143] D. G. Altman and J. M. Bland, "Statistics notes: the normal distribution," *Bmj*, vol. 310, no. 6975, p. 298, 1995.

[144] A. C. Elliott and W. A. Woodward, *Statistical analysis quick reference guidebook: With SPSS examples*. Sage, 2007.

[145] O. W. Samuel, M. G. Asogbon, Y. Geng, A. H. Al-Timemy, S. Pirbhulal, N. Ji, S. Chen, P. Fang, and G. Li, "Intelligent emg pattern recognition control method for upper-limb multifunctional prostheses: advances, current challenges, and future prospects," *IEEE Access*, vol. 7, pp. 10 150–10 165, 2019.

[146] N. Nazmi, M. A. Abdul Rahman, S.-I. Yamamoto, S. A. Ahmad, H. Zamzuri, and S. A. Mazlan, "A review of classification techniques of emg signals during isotonic and isometric contractions," *Sensors*, vol. 16, no. 8, p. 1304, 2016.

[147] S. Solorio-Fernández, J. A. Carrasco-Ochoa, and J. F. Martínez-Trinidad, "A review of unsupervised feature selection methods," *Artificial Intelligence Review*, vol. 53, no. 2, pp. 907–948, 2020.

[148] F. Song, Z. Guo, and D. Mei, "Feature selection using principal component analysis," in *2010 international conference on system science, engineering design and manufacturing informatization*, vol. 1. IEEE, 2010, pp. 27–30.

[149] G. Zhang and M. Y. Hu, "Neural network forecasting of the british pound/us dollar exchange rate," *Omega*, vol. 26, no. 4, pp. 495–506, 1998.

[150] R. Reed and R. J. MarksII, *Neural smithing: supervised learning in feedforward artificial neural networks*. Mit Press, 1999.

[151] Y. Bengio, "Practical recommendations for gradient-based training of deep architectures," in *Neural networks: Tricks of the trade*. Springer, 2012, pp. 437–478.

[152] M. A. Osborne, "Bayesian gaussian processes for sequential prediction, optimisation and quadrature," Ph.D. dissertation, Oxford University, UK, 2010.

[153] P. Domingos and M. Pazzani, "Beyond independence: Conditions for the optimality of the simple bayesian classi er," in *Proc. 13th Intl. Conf. Machine Learning*. Citeseer, 1996, pp. 105–112.

Appendix A

PCA OF THE CHANNEL SELECTION FOR RMS STUDY

Table A.1: Eigen analysis of the principal components

Parameter	**PC1**	**PC2**	**PC3**	**PC4**	**PC5**	**PC6**
Eigenvalue	3.3925	0.8618	0.5562	0.5353	0.4173	0.2369
Proportion	0.5650	0.1440	0.0930	0.0890	0.0700	0.0390
Cumulative Prop.	0.5650	0.7090	0.8020	0.8910	0.9610	1.0000

Table A.2: Eigenvector correspond to the each muscle in each principal component

Muscle	**PC1**	**PC2**	**PC3**	**PC4**	**PC5**	**PC6**
Muscle 1	0.4450	-0.3210	-0.1830	0.3960	0.3030	-0.6450
Muscle 2	0.4250	-0.0060	0.3090	-0.5580	0.6230	0.1590
Muscle 3	0.3970	0.2550	-0.6680	-0.4630	-0.3260	-0.1020
Muscle 4	*0.4420*	*-0.4110*	-0.1700	0.3210	-0.1120	0.7010
Muscle 5	0.4240	-0.0420	0.6230	-0.0800	-0.6180	-0.2030
Muscle 6	0.2990	0.8130	0.0860	0.4560	0.1380	0.1220

Appendix B

PILOT TEST DATA OF THE EXPERIMENT

B.1 Distance Data

Table B.1: Normalized onset distance data of 5-participants (pilot test).

Participant	Disk	Key	Lateral	Pinch	Point	Power
1	90.29	62.82	91.78	55.41	60.9	56.66
2	98	57.94	98.71	97.09	98.36	97.97
3	96.49	50.8	97.4	90.03	86.22	67.21
4	82.41	92.97	87.27	86.93	92.45	33.9
5	87.74	59.93	67.01	87.54	95.44	92.67
Mean	90.986	64.892	88.434	83.4	86.674	69.682
SD	6.404274	16.31107	12.82111	16.15869	15.09365	26.3851

Table B.2: Pooled SD (every grasping pair) of normalized onset distance data of 5-participants (pilot test).

	Disk	Key	Lateral	Pinch	Point
Key	12.39				
Lateral	10.13	14.67			
Pinch	12.29	16.24	14.59		
Point	11.59	15.71	14.00	15.64	
Power	19.20	21.93	20.74	21.88	21.49

Table B.3: Absolute difference (every mean grasping pairs) of normalized onset distance data of 5-participants (pilot test).

	Disk	Key	Lateral	Pinch	Point
Key	26.094				
Lateral	2.552	23.542			
Pinch	7.586	18.508	5.034		
Point	4.312	21.782	1.76	3.274	
Power	21.304	4.79	18.752	13.718	16.992

B.2 RMS Data

Table B.4: RMS data of 5-participants (pilot test).

Participant	Disk	Key	Lateral	Pinch	Point	Power
1	0.033114	0.0082	0.013994	0.009189	0.025098	0.022843
2	0.02222	0.007756	0.012137	0.006406	0.032553	0.063482
3	0.013402	0.007967	0.015392	0.005142	0.009267	0.031725
4	0.007922	0.006364	0.010889	0.018776	0.015506	0.019201
5	0.014591	0.007493	0.00987	0.011658	0.072273	0.078556
Mean	0.01825	0.007556	0.012456	0.010234	0.030939	0.043162
SD	0.00975	0.000716	0.002249	0.005401	0.024764	0.026379

Table B.5: Pooled SD (every grasping pair) of RMS data of 5-participants (pilot test).

	Disk	Key	Lateral	Pinch	Point
Key	0.006913				
Lateral	0.007075	0.001669			
Pinch	0.007881	0.003852	0.004137		
Point	0.018819	0.017518	0.017583	0.017923	
Power	0.019886	0.018660	0.018721	0.019040	0.025585

Table B.6: Absolute difference (every mean grasping pair) of RMS data of 5-participants (pilot test).

	Disk	Key	Lateral	Pinch	Point
Key	0.010694				
Lateral	0.005794	0.0049			
Pinch	0.008016	0.002678	0.002222		
Point	0.01269	0.023383	0.018483	0.020705	
Power	0.024912	0.035605	0.030705	0.032927	0.012222

Appendix C

ANOVA GLM RESULTS OF THE RMS STUDY

Table C.1: ANOVA GLM table of the RMS study. The *gender* effect has no significant relashionship with the RMS. Hence the a error was produced at the analysis

Source	**DF**	**Adj SS**	**Adj MS**	**F-Value**	**P-Value**
Grasp	5	0.005398	0.001080	6.96	0.000
Age	2	0.004532	0.002266	14.61	0.000
Gender	1	0.000398	0.000398	2.56	*0.110*
Handedness	*1*	*0.001637*	0.001637	10.55	0.001
Grasp*Age	10	0.010867	0.001087	7.01	0.000
Grasp*Gender	5	0.006582	0.001316	8.49	0.000
Grasp*Handedness	5	0.003670	0.000734	4.73	0.000
Error	495	0.076792	0.000155		
Lack-of-Fit	6	0.008513	0.001419	10.16	0.000
Pure Error	489	0.068278	0.000140		
Total	524	0.121283			

Appendix D

ANOVA GLM RESULTS OF THE ONSET STUDY

Table D.1: ANOVA GLM table of the onset study.

Source	**DF**	**Adj SS**	**Adj MS**	**F-Value**	**P-Value**
Age Category	2	1650	825.1	4.39	0.013
Handedness	1	1084	1083.7	5.77	*0.017*
Gender	*1*	*2610*	2609.9	13.89	0.000
Grasp	5	48158	9631.6	51.24	0.000
Grasp*Age Category	10	25445	2544.5	13.54	0.000
Grasp*Handedness	5	4258	851.6	4.53	0.000
Grasp*Gender	5	4147	829.5	4.41	0.001
Error	870	163525	188.0		
Lack-of-Fit	18	17655	980.9	5.73	0.000
Pure Error	852	145869	171.2		
Total	899	259526			

Appendix E

FILTER PARAMETERS

Table E.1: IIR single notch filter parameters

Parameter	**Value**
Sampling frequency	4000 Hz
Notch frequencies	from the screen process
Bandwidth	8 Hz, 3 Hz
Bandwidth attenuation	0.04dB

Table E.2: IIR Butterworth bandpass filter parameters

Parameter	**Value**
Sampling frequency	4000 Hz
First stopband frequency	10 Hz
First passband frequency	20 Hz
Second passband frequency	480 Hz
First stopband frequency	500 Hz
First stopband attenuation	1dB
Passband ripple	1dB
Second stopband attenuation	10dB
Match	Stopband

Appendix F

SURVEY QUESTIONNAIRE

Table F.1: Pilot test results of 6 participants. Sample size was calculated for given Margin of Error (MOE) and Standard Deviation (StDev)

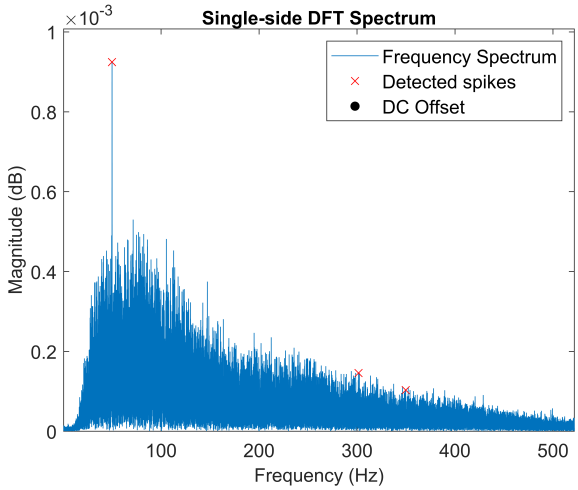
Object	N	Mean Score	SE Mean	MOE	StDev	Sample Size
Scissor	6	68.8	11.1	21.756	27.2	9
Cup	6	76.0	12.9	25.284	31.6	9
Bottle	6	74.3	15.0	29.400	36.8	9
Mouse	6	65.3	10.6	20.776	25.9	9
Fork	6	69.2	11.4	22.344	27.9	9
Laptop	6	70.2	8.44	16.542	20.67	9
Keyboard	6	61.7	12.30	24.108	30.3	9
Knife	6	66.5	13.20	25.872	32.3	9
Remote	6	79.3	10.50	20.580	25.8	9
Phone	6	81.67	9.72	19.050	23.80	9
Spoon	6	72.8	11.30	22.142	27.7	9

Table F.2: Derived conditional probability matrix

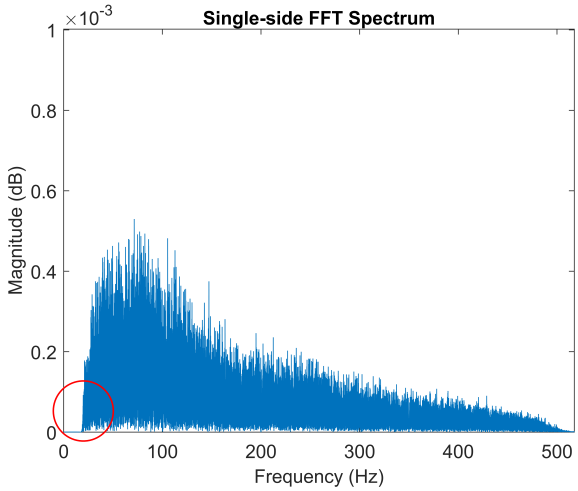
Object	Disk	Key	Lateral	Pinch	Point	Power
Scissor	0.11	0.20	0.14	0.17	0.10	0.28
Cup	0.21	0.12	0.26	0.1	0.09	0.21
Bottle	0.21	0.1	0.3	0.1	0.1	0.19
Mouse	0.18	0.09	0.12	0.15	0.17	0.28
Fork	0.1	0.25	0.12	0.24	0.1	0.18
Laptop	0.1	0.14	0.15	0.16	0.24	0.21
Keyboard	0.11	0.13	0.15	0.12	0.33	0.17
Knife	0.1	0.2	0.14	0.13	0.12	0.32
Remote	0.11	0.19	0.17	0.09	0.23	0.21
Phone	0.11	0.16	0.18	0.11	0.23	0.2
Spoon	0.1	0.25	0.12	0.21	0.11	0.21

Appendix G

SIGNAL FILTERING



(a)



(b)

Figure G.1: (a) Second stage filter. (b) Bandpass filtering.

Appendix H

GRID SEARCH OPTIMIZATION

H.1 Neural network parameters and hyperparameters

Table H.1: Parameters and the hyperparameters of selected neural networks

Hyperparameter	**Values**
Inputs	48 (8 TD × 6 Ch)[1]
Outputs	6 classes
hidden layers	1(40, 100)[2]
Training Function	6 Scaled Conjugate Gradient
Data Splitting	as section 4.1.3 mentioned
Loss function	Cross-entropy
Min. gradient of loss Func.	0.0001
Validation checks	40
Sigma	0.000008
Momentum Constant	0.05
Epochs	1000
Lambda	0.0000005
Regularization	0.05

[1] TD - Time Domain, Ch - Channels(Muscles)

[2] Two networks with 40 and 100 hidden layer neurons

H.2 Trend Analysis of the hidden layers

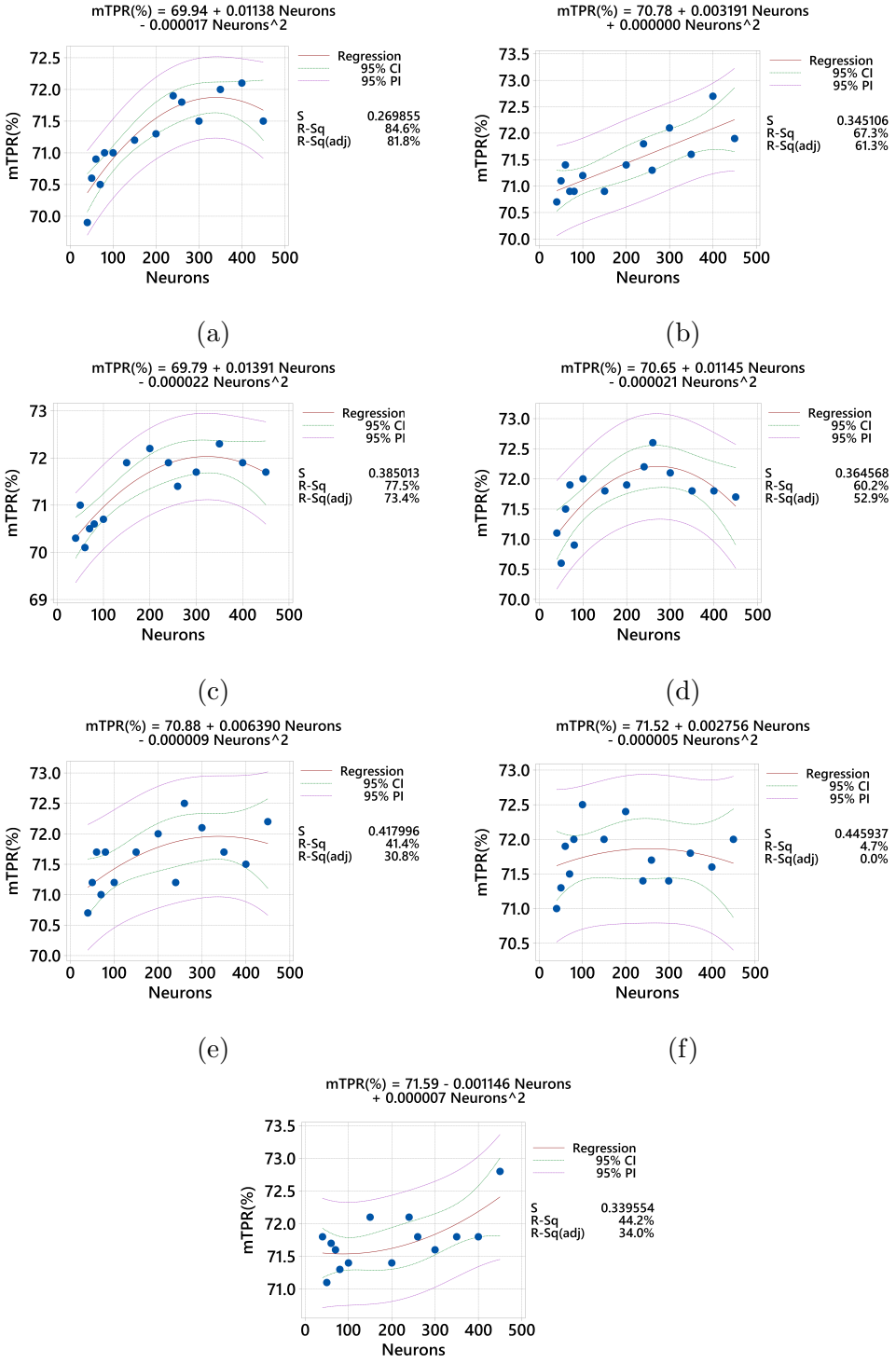


Figure H.1: Trend analysis of HL-2 neuron configuration.

Appendix I

BAYESIAN OPTIMIZATION RESULTS

I.1 Optimized results of the conventional classifiers

Table I.1: Optimized results of the conventional classifiers. Feature groups with the demographic features.

Classifier	TDAR	TD	All	AR	FNon	5-PCs
NB	35.96	28.64	37.44	25.97	27.21	16.36
DT	57.49	58.37	55.61	29.61	55.06	25.32
KNN	63.26	66.74	65.36	23.20	44.98	17.33
LDA	43.14	41.86	42.80	23.02	32.84	16.36

Table I.2: Optimized results of the conventional classifiers. Feature groups without demographic features.

Classifier	TDAR	TD	All	AR	FNon	5-PCs
NB	35.95	28.65	37.50	25.85	27.21	16.54
DT	51.65	52.96	50.72	23.74	49.53	20.85
KNN	62.68	61.88	64.10	27.99	38.19	16.95
LDA	42.61	41.36	42.69	21.65	33.15	16.87

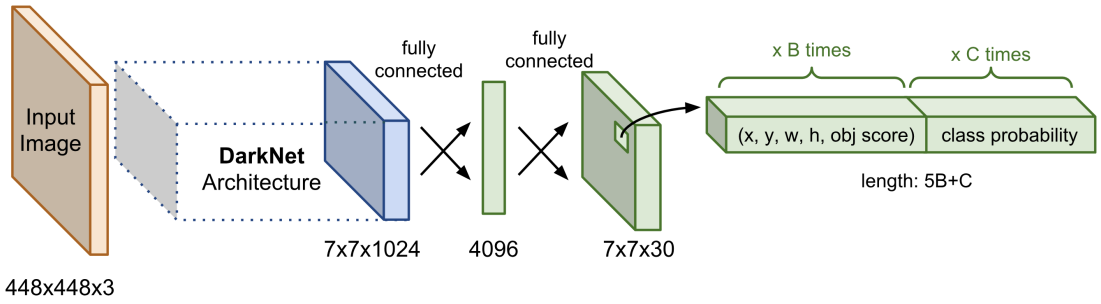
Appendix J

YOLO ALGORITHM

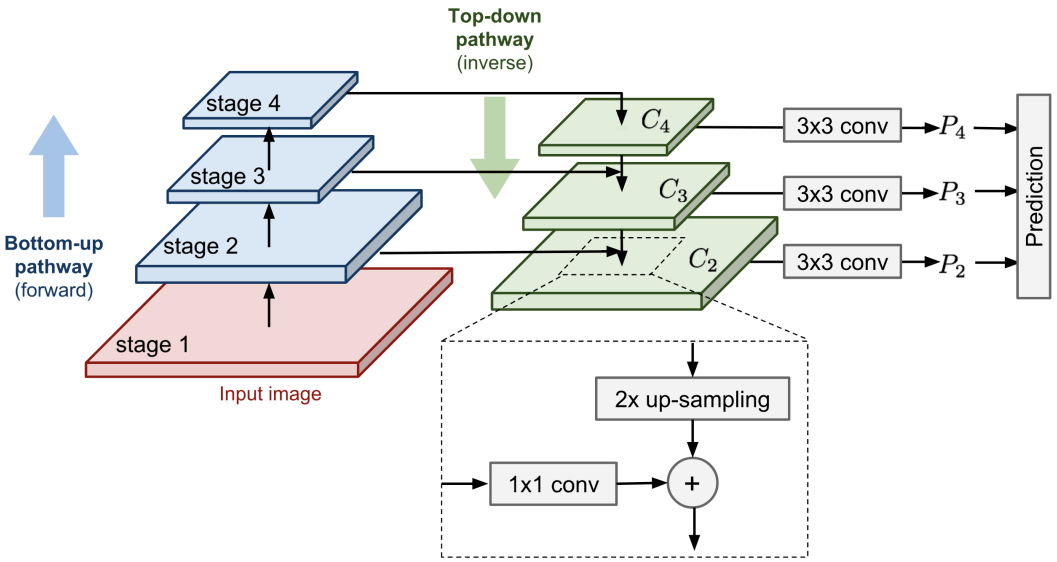
The YOLO is Convolution Neural Network (CNN) based object detection algorithm. Unlike other detection algorithms such as, Region based Convolutional Neural Networks(R-CNN), YOLO uses single-feedforward CNN which is typically use for regression problems. Hence, the detection time is significantly lower than other methods.

Before YOLO was introduce, the detection algorithms use the same image for several times to compute the probability of the particular object presence in the image as well as the object location. However the YOLO architecture uses the image only one time to compute all the above outputs. That is why the YOLO takes comparably less time to detect object and their corresponding locations [10].

The algorithm look on the image only one time and start the predictions. That is why it is name as *You Only Look Once*. Once the image is received, it calculate the features of the entire image using the pre-trained DarkNet (see Figure J.1a). Then algorithm breaks the feature image into a grid ($S \times S$, $S = 7$) as the Figure J.2. Each cell predicts B number of bonding boxes ($B = 2$ and the box attributes x, y, w, h) and confidence scores, CS of each boxes. The confidence score (CS) reflects the presence of a object in that particular bounding box but it doesn't have information about the type of the object (Object class C). In the above Figure J.2, there bounding boxes that has thick outlines. When a confidence score of a particular box is high, then the thickness of that bounding box is high. Simultaneously it calculates the class conditional probability of the particular



(a)



(b)

Figure J.1: (a) The model architecture of YOLO (b) The model architecture of ResNet (featurized image pyramid). (image is adapted from, *Lil'Log* [11])

object ($P(C_i|object)$) in each cell S_i. The FigureJ.2 illustrates the $P(C_i|object)$ as a color map. The following are the predicted outputs of the algorithm.

1. Boxes 4 attributes ($x\ coordinate, y\ coordinate, width, height$)
2. Confidence score of the box
3. Class probability distribution

Then the final output is $S \times S \times (B * 5 + C)$ dimension long. Here, B is the number of boxes in each cell, 5 means the combinations of 4 attributes and confidence score of the box and C is the number of classes.

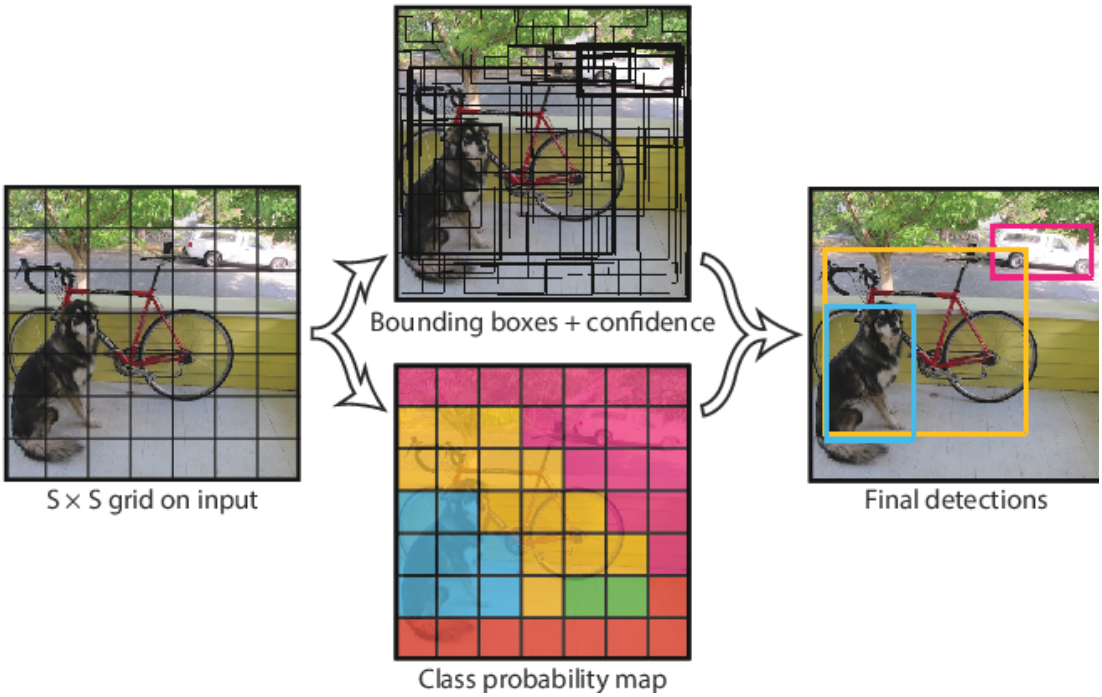


Figure J.2: The overview of the YOLO algorithm. It devide the image into grids and predict the object presence in each grid simultaneously while detecting the locations [10].

At the end of the algorithm the final score of a object in the bounding box is calculated by $P(C_i|object) \times CS$. A class with final score above a threshold value is considered as the detected object class of that particular bounding box. The bounding boxes with lower CS are rejected.