HAND GESTURE RECOGNITION VIA ELECTROMYOGRAPHIC (EMG)

ARMBAND FOR CAD SOFTWARE CONTROL

A Thesis
IN
Electrical Engineering

Presented to the Faculty of the University
of Missouri–Kansas City in partial fulfillment of
the requirements for the degree

MASTER OF SCIENCE

by
ALA-ADDIN NABULSI

B.S., University of Jordan, 2014

Kansas City, Missouri
2018

HAND GESTURE RECOGNITION VIA ELECTROMYOGRAPHIC (EMG)

ARMBAND FOR CAD SOFTWARE CONTROL

Ala-Addin Nabulsi, Candidate for the Master of Science Degree

University of Missouri–Kansas City, 2018

ABSTRACT

In the past, computers - whether personal or at work - required a mouse and keyboard to interact with them, and they are still used to this day. Even for video games a physical tool (controller) is needed to interact with the gaming environment. Previously that was acceptable since that was how these electronic devices were conceived, but with the recent boom in Virtual reality (VR) and Augmented Reality (AR), that reality has already started to change. VR and AR have existed for well over 2 decades [1], yet only in the last 5 years have they started getting closer to reaching their true potential. With this new technology, we can go to virtual worlds, interact with creatures that never previously existed, and visualize information in ways never thought possible before.

With the emergence of VR came the need to change the way we interact with the virtual environment and, with that, the way we interact with technology as a whole. And what better controller for the job than the human hand. If the user can interact with

technology with hand gestures then the whole process becomes intuitive, eliminating any training time, and giving the user a more natural experience. For this, Hand Gesture Recognition (HGR) systems will be needed.

HGR systems recognize the user's hand shape by means of a glove, cameras, or biosignals. One particularly useful biosignal for this task is the forearm Electromyographic (EMG) Signal. This signal reflects the contraction state of the forearm muscles. EMG signals are already being used in prosthetics to help amputees have more natural control over their prosthetic limbs. They can also be used for translating sign language, or just generally in Human-Machine-Interaction (HMI).

This work proposes a method to interact with computers using hand gestures, specifically for a Computer Aided Design (CAD) software known as Solidworks. To achieve this a commercial EMG armband (the Myo - Thalmic Labs) was used to record 8-channel EMG signals from a group of volunteers over the span of 3 visits. The data set was then preprocessed and segmented. The resulting data set consisted of 10 hand gestures performed by 10 subjects, with 162 samples per gesture. A total of 11 feature sets were extracted and applied to 4 different machine learning models.

A 9-fold cross validation and testing was performed and the classifiers over all the feature sets were evaluated and compared. The best model validation performance was achieved by the Linear Discriminant Analysis (LDA) model with an average Area Under Curve (AUC) of 76.35% and an average Equal Error Rate (EER) of 29.73%.

In future work, we propose to use the HGR method developed in this thesis in multiple applications such as mapping certain shortcut commands in Solidworks (and other applications) to hand gestures.

APPROVAL PAGE

The faculty listed below, appointed by the Dean of the School of Computing and Engineering, has examined the thesis titled "Hand Gesture Recognition via Electromyographic (EMG) Armband for CAD Software CONTROL," presented by Ala-Addin Nabulsi, candidate for the Master of Science degree, and certifies that in their opinion it is worthy of acceptance.

<u>Supervisory Committee</u>

Reza Derakhshani, Ph.D., Committee Chair
Department of Computer Science & Electrical Engineering

Ahmed Hassan, Ph.D
Department of Computer Science & Electrical Engineering

Katherine Bloemker, Ph.D.
Department of Computer Science & Electrical Engineering

CONTENTS

Appendix

ILLUSTRATIONS

TABLES

ACRONYMS AND ABBREVIATIONS

AR: Augmented Reality

AUC: Area Under Curve

CAD: Computer Aided Design

CNNs: Convolutional Neural Networks

ECG: Electrocardiography

EEG: Electroencephalography

EER: Equal Error Rate

EMG: Electromyography

HGR: Hang Gesture Recognition

HMD: Head-Mounted-Display

HMI: Human Machine Interaction

IMU: Inertial Measurement Unit

IPC: Intuitive Prosthetic Control

KNN: K-Nearest Neighbor

LDA: Linear Discriminant Analysis

ML: Machine Learning

NIRS: Near-Infrared Spectroscopy

PCA: Principal Component Analysis

ROC: Receiver Operating Characteristic

sEMG: Surface Electromyography

SLI: Sign Language Interpretation

SVMs: Support Vector Machines

VR: Virtual Reality

ACKNOWLEDGMENTS

Firstly, I would like to take this opportunity to express my deep gratitude and thanks to my advisors and mentors Dr. Reza Derakhshani and Dr. Ahmed Hassan for all the guidance they have given me throughout these two long years. If not for them, I probably would not have entered into the field of research and would have just taken the courses and graduated for the sole purpose of obtaining another certificate. I am grateful for the time and resources they've both given me during my time getting my Master's degree.

Secondly, I would like to thank my family for believing in me, for always being there for me to support me in the hard times as well as the easy times.

I would also like to thank my colleagues and friends, Dr. Ajita Rattani, Jesse Lowe, Ahmad Mohammad, Mark (Hoang) Nguyen, Sai Narsi Reddy, and Clayton Kettlewell. A lot of what I know now is from the time we spent together, the intriguing discussions we had and all their support. I am grateful for being a member of UMKC's CIBIT lab. This gave me the unique opportunity to interact with researchers on a regular basis. Every meeting I attended covered amazing topics, thought provoking discussions and unique insights on how research should be done.

CHAPTER 1

INTRODUCTION

The purpose of this paper is to show the feasibility of using hand gestures, via a wearable armband (Myo armband) with biosignal sensors, to control Computer Aided Design (CAD) software such as Solidworks. This would allow for a more intuitive method of control, thus bringing the experience of using such sophisticated software one step closer to the average user [2].

1.1 Background

A biosignal could be defined as any signal in the human body that can be measured and/or monitored. Electric biosignals are possibly the most used biosignals in the medical field. Examples include Electroencephalography (EEG), Electrocardiography (ECG), Mechanomyography (MMG), and Electromyography (EMG). The biosignal of interest in this paper is the EMG signal.

Electromyography (EMG), also referred as Electromyogram, is an electric biosignal that reflects the contraction state of a muscle [3]. It can be acquired through two types of sensors: Intramuscular EMG (Intra-EMG) sensors or Surface EMG (sEMG) sensors. Intra-EMG sensors are needle-shaped electrodes that penetrate the skin and go into the muscles for better conductivity. They are mostly used in diagnosing neurological disorders [4]. The sEMG, on the other hand, is a flat sensor that is placed on top of the skin

region over the muscle of interest. Since sEMG is non-invasive, it is a preferred method for many applications.[1]

EMG signals can, if utilized correctly, tell a great deal about the human body. For example, EMG signals can be recorded from a subject's forearm and fed it into a Machine Learning model that would then be trained to recognize what hand gesture is being made at that given time. Hand Gesture Recognition (HGR) systems can learn from a variety of sensor data, whether the data is in the form of images, audio signals or even biosignals.

1.2   Types of Hand Gesture Recognition Systems

1. Visual Based HGR: (V-HGR)

   V-HGR systems use a variety of image capture devices to track the motion of the hands and what gestures they are making. Such devices include depth-aware cameras (structured light or time-of-flight cameras), single 2D camera, or even stereo cameras. These HGR devices can either be fixed to a table[2] [5] or fixed to a Head-Mounted-Display (HMD)[3].

   [Devices: Leap Motion (figure 1a), Microsoft Kinect 2 (figure 1b), Nimble Sense]

2. Glove Based HGR: (G-HGR)

   These gloves, whether wired or wireless, have inertial measurement units (IMUs) to track the orientation of the hands, and flex sensors along the fingers to track

---

[1]From this point onward, any mention of EMG refers to sEMG unless stated otherwise.
[2]https://www.xbox.com/en-US/xbox-one/accessories/kinect
[3]https://www.leapmotion.com

Figure 1: Devices used in V-HGR: (a) The Leap Motion controller, (b) The Microsoft Kinect 2

individual finger movement; but require an additional Visual-Based device to track position.

An intriguing fact about one of these data gloves in particular (Gloveone)[4] is that it gives haptic feedback (vibrations) to the fingers and the palm of the hand to simulate touch or force. This is highly desirable, whether in Virtual Reality (VR) or Augmented Reality (AR) applications. Such systems are still currently under development by some companies, while others are about to release their data gloves[5]. [Devices: Gloveone, Manus VR Glove (HTC Vive), Cyberglove III, Essential Reality P5 Gaming Glove]

3. Biosignal Based HGR: (B-HGR)

Until a few years ago, the only existing HGR systems for VR were mainly visual-based or glove-based, which were still under development. A third category recently emerged, which can be referred to as Biosignal based HGR system. This paper will focus on the EMG based HGR (EMG-HGR) systems. In EMG-HGR

---

[4]https://www.neurodigital.es/
[5]https://manus-vr.com/

systems, IMUs are used to track position and orientation; but the hand gesture is tracked using the EMG sensors. The EMG sensors are electrodes that sense the biosignal potential of the forearm muscles as they contract and relax to perform different hand gestures. This HGR system[6] is in the form of a wireless armband worn on the forearm known as the Myo armband (figure 2).

Figure 2: The Myo Armband (Thalmic Labs)

A few research groups have come up with their own configurations [6], but so far only one such system is available commercially (the Myo). It has been shown that adding another modality, such as a Near-Infrared Spectroscopy (NIRS) sensor could improve HGR accuracy [7].

As of March 14th, 2017 there are at least 9 VR gloves in development[7]. This shows how HMI is shifting from using controllers towards using hand gestures. That is why there is a need to build better, more reliable, HGR systems. Although the V-HGR and G-HGR systems are more accurate as of now, the EMG-HGR provides an inconspicuous equivalent that also protects the user's privacy (no cameras). Though more development/improvements need to be made to EMG-HGR systems in order to consider them as

[6]https://www.myo.com/
[7]http://virtualrealitytimes.com/2017/03/14/full-list-of-glove-controllers-for-vr/

reliable input methods.

1.3 Applications of Hand Gesture Recognition

HGR systems are already having a great impact on many research fields, including: Human Machine Interaction (HMI), Sign Language Interpretation (SLI), and Intuitive Prosthetic Control (IPC).

- Human-Machine-Interaction (HMI)

  In HMI, a user can control/interact with a computer in a way that is more natural through hand gestures, by means of HGR systems. Combining this with voice recognition, keyboards and mice may become obsolete in the near future.

- Sign Language Interpretation (SLI)

  In SLI, HGR systems can help provide an equivalent to existing voice-based language translation applications for the people who use sign language in their everyday interactions [8].

- Intuitive Prosthetic Control (IPC)

  Prosthetics' physical designs have undergone many changes over the years as well as did their method of control. IPC is a more modern method of prosthetic control that seeks to use the EMG signals of subject's distal muscles on the amputated limb in order to control their prosthetic in a more natural fashion and significantly improve the subject's quality of life [9].

1.4 Computer Aided Design

Computer Aided Design (CAD) is using a computer to create, modify, analyze, or optimize a design [10]. CAD allows manufacturers and designers to bring their ideas into a virtual reality before building them in the real world. This allows them to analyze their designs and simulate how they may run in the real world in order to "debug" them before sending them off to production.

There are many commercially available CAD software, with Autodesk's Auto-CAD and Inventor, and Dassault Systemes's SolidWorks being among the most widely used. These programs started off having command-line based control, which meant everything was drawn using just the keyboard as an input device. This method is efficient for those trained to use it but comes off as complicated for the non-engineering/programming-oriented individuals. Later there was a shift that involved adding the mouse for input. This added more degrees of freedom to the users and hence made it easier, since using the mouse to draw is slightly more intuitive than the keyboard. With these two devices, in addition to the ever-evolving user interfaces and visuals, drawing in CAD has become a lot easier to use than in the early days. With the emergence of environments such as VR and AR, the method of interacting with CAD software will evolve once more to become highly intuitive, to a point where a non-technical individual can learn to use it within minutes. As with other VR/AR applications, the HMI in future CAD will be by means of hand gestures and HGR systems.

1.5 Overview of EMG - Hand Gesture Recognition System

In an EMG-HGR system, the acquired EMG signals undergo some signal processing before going into the Machine Learning Model. The involved signal preprocessing includes scaling, filtering, projecting the data onto new dimensions, or obtaining any of its statistical and frequency information. Once the preprocessing is complete, the signal is then ready to go into the Machine Learning (ML) Model. The model extracts features from the data, which are then used to classify the EMG signal sample. This classification then triggers an action in the Myo-controlled system being used. An illustration of this process is shown in figure 3.
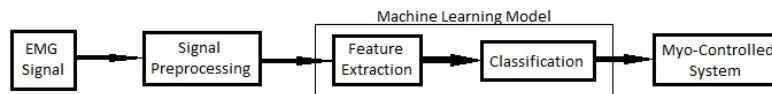
Figure 3: Overview of an EMG-HGR System

The chosen hand gestures can be linked to specific keyboard keys using key bindings. Afterwards, those keys can be used in a software's shortcut menu (e.g. Solidworks) to be mapped to their assigned actions.

CHAPTER 2

RELATED WORK

Previous works on HGR with the Myo armband have used different methods for each aspect of the proposed HGR systems, the aspects being the preprocessing methods (features extracted), the machine learning models used, whether or not other modalities were used alongside the EMG signals, and the different data sets collected and used. The published variations in these aspects are discussed individually below.

2.1 Signal Preprocessing Methods

Once an EMG signal sample is acquired, it needs to be preprocessed in accordance with the classification method to be used. For EMG signals, preprocessing may include: full-wave rectification of the signal, normalization, and filtering (to remove noise from power sources and other biosignals). After this is done, other EMG features can be extracted to be fed into the classifier. Extracted features in previous work include:

1. Mean Absolute Value (MAV)

2. Zero Crossings (ZC)

3. Waveform Length (WL)

4. Root Mean Square (RMS)

5. Wilson Amplitude (WAMP)

6. Maximum (Max)

7. Minimum (Min)

8. Standard Deviation (SD)

9. Histogram (Hist)

10. Short-time Fourier Transform (STFT, A.K.A. Spectrograms)

11. Variance (VAR)

12. Discrete Wavelet Transform (DWT)

13. Amplitude Spectrum (AS)

14. Modified Mean Frequency (MF Mean)

15. Modified Median Frequency (MF Median)

16. Energy Ratio

17. MAV of DWT Coefficients

18. SD of DWT coefficients

19. Ratios between MAV of DWT coefficients

20. Pairwise-Inner-Product (PIP)

21. Teager Energy (TE)

Table 1 shows the various features that have been most commonly extracted before classification. Some of these feature will be explained further in Chapter 5. As seen in Table 1, depending on the selected feature(s), acceptable results could be achieved with using as many as 9 features as an input [2], or using just 1 [11]. A more detailed account of Table 1 will follow.

Kerber et al [2] achieved a classification accuracy of 95% by extracting and combining (at feature level) the MAV, ZC, WAMP, Hist, VAR, AS, MF Mean, MF Median, and the Energy Ratio.

Gieser et al [12] achieved a classification accuracy of 89.95% (KNN), 90.75% (SVM) by extracting and combining (at feature level) the MAV, RMS, Max, Min, and VAR.

Luh et al [13] achieved a classification accuracy of 89.38% by extracting and combining (at feature level) the MAV, SD, DWT, MAV of DWT coefficients, SD of DWT coefficients, and the Ratios between MAV of DWT coefficients.

Lin et al [14] achieved a classification accuracy of 83% by extracting and fusing (at feature level) the MAV, WL, RMS, PIP, and TE.

Allard et al [11] achieved a classification accuracy of 97.9% by extracting the STFT (Spectrogram) feature.

Lobov et al [15] achieved a classification accuracy of 97.03% by extracting the RMS feature.

Amirabdollahian et al [16] achieved classification accuracies of 94.9% (linear), 89.8% (polynomial), 31.1% (RBF) by extracting the Waveform length feature.

Benalcazar et al [17] achieved a classification accuracy of 89.5% without extracting any features. They only fully-rectified the signals then passed them through a low-pass filter (LP(Abs(x))) then fed the samples to the classifier.

2.2 Machine Learning Models

As can be seen in Table 2, the best results were obtained by using Convolutional Neural Networks (CNNs) [11], followed by Multilayer Perceptrons (MLPs) [15], SVM [16], then the remaining statistical methods [2, 12–14]. A more detailed account of table 2 will follow.

Kerber et al [2] achieved a classification accuracy of 95% by means of a Support Vector Machine (SVM) classifier.

Gieser et al [12] compared the k-Nearest Neighbor (KNN) classifier with the SVM. They achieved a classification accuracy of 89.95% with the KNN classifier and 90.75% with the SVM classifier.

Luh et al [13] achieved a classification accuracy of 89.38% by means of a Multilayer Perceptron (MLP), also known as the Artificial Neural Network (ANN).

Lin et al [14] compared the ANN, KNN, Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA), and SVM classifiers. They achieved a classification accuracy of 83% with the ANN classifier, 78% with the KNN classifier, 81% with the QDA classifier, and 81% with the SVM classifier.

Allard et al [11] achieved a classification accuracy of 97.9% by using 2 identical Convolutional Neural Networks (CNNs).

Lobov et al [15] achieved a classification accuracy of 97.03% by using an MLP, with 2 layers (layer size 1 = 9, layer size 2 = 7).

Amirabdollahian et al [16] used an SVM classifier and compared different kernels, namely the Linear, polynomial and Radial Basis Function (RBF) Kernels. They achieved a classification accuracy of 94.9% with a linear kernel, 89.8% with a polynomial kernel, and 31.1% with a Radial Basis Function (RBF) kernel.

Benalcazar et al [17] achieved a classification accuracy of 89.5% by means of a KNN classifier, with Dynamic Time Warping for calculating distance between neighbors.

2.3 Sensor Fusion

Using motion data alongside EMG, the classification performance greatly improves. As can be seen from Table 3, whether adding acceleration [18], or acceleration and orientation data [19–21], the classification accuracy was always above 95%. The reason for this is that the motion data accounts for the hand and/or forearm's motion while performing a hand gesture (in addition to the hand's shape). A more detailed account will follow.

Kerber et al [2] achieved a classification accuracy of 95% by only using Myo's 8 EMG signals.

Gieser et al [12] achieved a maximum classification accuracy of 90.75% using just the 8 EMG signals.

Luh et al [13] were able to achieve a classification accuracy of 89.38% using the 8 EMG channels.

Lin et al [14] achieved an accuracy of 83% using just the 8EMG signals.

Allard et al [11] used the 8 EMG signals to achieve a classification accuracy of 97.9%.

Lobov et al [15] achieved an accuracy of 97.03% using only the 8 EMG signals.

Amirabdollahian et al [16] achieved classification accuracies of 94.9% (linear kernel), 89.8% (polynomial kernel), and 31.1% (RBF kernels) by using just the 8 EMG signals.

Benalcazar et al [17] were able to achieve a classification accuracy of 89.5% using the 8 EMG signals from the Myo armband.

Paudyal et al [19] achieved a classification accuracy of 95.36% after utilizing the Myo's IMU data (9-axis).

Yeh et al [18] achieved a classification accuracy of 97.9% when then added Myo's accelerometer data (3-axis).

Paudyal et al [20] achieved a classification accuracy of 97.7%. This was achieved with data from 2 Myo armbands, utilizing all the signals from both armbands (EMG+IMU).

Kutafina et al [21] used 2 Myo armbands (EMG+IMU) to achieve a classification accuracy of 98.3%.

2.4   Data Sets Used

The data sets collected for EMG classification vary in size as well as their other aspects. They can be seen in Table 4. Their sizes varied from 5 subjects [13] up to 35 [12], as well as their number of samples per class, ranging from 40 [15, 20] up to 4000 [13]. A

more detailed account will follow.

Kerber et al [2] collected EMG data, using 1 Myo armband, from 14 subjects performing 40 hand gestures, 5 times each. This resulted in each gesture class consisting of 70 samples. All study participants in the study were able-bodied. (i.e. none of the participants had amputated arms.)

Gieser et al [12] collected EMG data, using 1 Myo armband, from 35 subjects performing 11 hand gestures, 3 times each. This resulted in each gesture class consisting of 105 samples. It was not clear whether or not there were any amputees in the data set.

Luh et al [13] collected EMG data, using 1 Myo armband, from 5 subjects performing 17 hand gestures, 800 times each. This resulted in each gesture class consisting of 4000 samples. All study participants in the study were able-bodied.

Lin et al [14] collected EMG data, using 1 Myo armband, from 10 subjects performing 9 hand gestures, 5 times each. This resulted in each gesture class consisting of 50 samples. All study participants in the study were able-bodied.

Allard et al [11] collected EMG data, using 1 Myo armband, from 18 subjects performing 7 hand gestures. Each gesture class consisting of 70 samples (through windowing). All study participants in the study were able-bodied.

Lobov et al [15] collected EMG data, using 1 Myo armband, from 10 subjects performing 7 hand gestures, 4 times each. This resulted in each gesture class consisting of 40 samples. All study participants in the study were able-bodied.

Amirabdollahian et al [16] collected EMG data, using 1 Myo armband, from 25 subjects performing 4 hand gestures, 5 times each. This resulted in each gesture class

consisting of 125 samples. All study participants in the study were able-bodied.

Benalcazar et al [17] collected EMG data, using 1 Myo armband, from 10 subjects performing 5 hand gestures, 60 times each. This resulted in each gesture class consisting of 600 samples. All study participants in the study were able-bodied.

Paudyal et al [19] collected EMG and IMU data, using 1 Myo armband, from 9 subjects performing 26 hand gestures, 6 times each. This resulted in each gesture class consisting of 54 samples. All study participants in the study were able-bodied.

Yeh et al [18] collected EMG and Accelerometer data, using 2 Myo armbands, from 11 subjects performing 9 hand gestures, 10 times each. This resulted in each gesture class consisting of 110 samples. All study participants in the study were able-bodied.

Paudyal et al [20] collected EMG and IMU data, using 2 Myo armbands, from 10 subjects performing 20 hand gestures, 4 times each. This resulted in each gesture class consisting of 40 samples. All study participants in the study were able-bodied.

Kutafina et al [21] collected EMG and IMU data, using 2 Myo armbands, from 17 subjects performing 9 hand gestures, 3 times each. This resulted in each gesture class consisting of 51 samples. All study participants in the study were able-bodied.

As it can be seen, there is not a lot of consistency in the data sets regarding the number of test subjects, number of gestures collected, which gestures are collected, the number of times a gesture is repeated, or the type of muscle contractions (isotonic/isometric).

Another issue that arises, due to the data sets being in-house (private), is the inability of researchers to actually compare their work with one another. This is halting any significant, quantifiable, progress to be made in EMG signal classification. A solution to

this would be having an agreed-upon methodology/protocol for EMG data collection regarding the previously mentioned inconsistencies, the type of data collected, and the data storage methods. This data set can incorporate able-bodied individuals, as well as amputees, and include gestures that would account for the different application of EMG Signal Classification. Having such a data set collected and publicly available would provide researchers with a benchmark to compare their work. Only in this way can significant, quantifiable progress be made in EMG classification.

Table 1: Extracted EMG Features

Features	[2]	[12]	[13]	[14]	[11]	[15]	[16]	[17]
MAV	✓	✓	✓	✓				
ZC	✓							
WL				✓			✓	
RMS		✓		✓		✓		
WAMP	✓							
Max		✓						
Min		✓						
SD			✓					
Hist	✓							
STFT					✓			
VAR	✓	✓						
DWT			✓					
AS	✓							
MF Mean	✓							
MF Median	✓							
Energy Ratio	✓							
*			✓					
**			✓					
***			✓					
PIP				✓				
TE				✓				
LPF(Abs(x))								✓
Accuracy	95%	89.95% (KNN), 90.75% (SVM)	89.38%	83%	97.90%	97.03%	94.9% (linear), 89.8% (polynomial), 31.1% (RBF)	89.5%

* MAV of DWT coefficients
** SD of DWT coefficients
*** Ratios between MAV of DWT coefficient

Table 2: Machine Learning Methods Used

Author	Classifier Type	Accuracy
[2]	SVM	95%
[12]	KNN & SVM compared	89.95% (KNN), 90.75% (SVM)
[13]	MLP (compared with fuzzy logic)	89.38%
[14]	ANN, KNN, LDA, QDA, SVM	83%(ANN), 78%(KNN), 80%(LDA), 81%(QDA), 81%(SVM)
[11]	2 cascaded identical CNNs	97.90%
[15]	MLP (2 layers: Ls9+Ls7)	97.03%
[16]	SVM (Linear, Polynomial, and RBF Kernels)	94.9%, 89.8%, 31.1%
[17]	KNN (with Dynamic time warping)	89.5%

Table 3: Sensor Fusion

Author	Other Modalities	Accuracy
[2]	-	95%
[12]	-	89.95% (KNN), 90.75% (SVM)
[13]	-	89.38%
[14]	-	83%
[11]	-	97.90%
[15]	-	97.03%
[16]	-	94.9% (linear), 89.8% (polynomial), 31.1% (RBF)
[17]	-	89.5%
[19]	9-axis IMU data	95.36%
[18]	3 axis Accelerometer	97.90%
[20]	9 IMU channels (x2)	97.7%
[21]	9 IMU channels (x2)	98.30%

Table 4: Collected Data sets

Author	Able-bodied	Subjects	Classes	Samples/ Class	Armbands (Myo)	Other Modalities
[2]	✓	14	40	70	1	-
[12]	N/A	35	11	105	1	-
[13]	✓	5	17	4000	1	-
[14]	✓	10	9	50	1	-
[11]	✓	18	7	70*	1	-
[15]	✓	10	7	40	1	-
[16]	✓	25	4	125	1	-
[17]	✓	10	5	600	1	-
[19]	✓	9	26	54	1	IMU
[18]	✓	11	9	110	1	Accelerometer
[20]	✓	10	20	40	2	IMU
[21]	✓	17	9	51	2	IMU

* using windowing.

CHAPTER 3

CHOOSING HAND GESTURES

3.1 Identifying Possible Actions in Solidworks

Hand gestures help a great deal with human communication. They can have different meanings in different cultures and/or regions. Before choosing the hand gestures for this project, the actions/commands to be used had to first be identified. In CAD software (namely, Solidworks) there are different type of actions that can be done. These can be divided into five categories:

- Manipulation:

  This category involves manipulating an objects drawing in order to get a better look, closer look, or overall look.

  [Actions: Pan, Zoom, and Change Orientation]

- Drawing: (2D)

  This category includes an assortment of 2D objects to select from to start drawing with the mouse.

  [2D Shapes: Lines, Rectangles, Polygons, Circles, Arcs, etc.]

- Editing: (2D)

  This category includes actions that can be done to 2D drawings.

[Actions: Fillet, Trim, Extend, Move, Copy, Delete, Scale, etc.]

- Editing: (3D)

  Includes actions that can be done to 3D objects in addition to 2D drawings.

  [Actions: Fillet, Extrude, Revolve, Sweep, Loft, etc.]

- Executive Actions:

  Includes actions such as Confirm, Cancel, Undo, and Exit Sketch.

After identifying the possible CAD actions, it is possible to assign a hand gesture to each action. This task is easier said than done since the human hand, starting from the wrist to the finger tips, has 27 Degrees of Freedom (DoF). Each of DoF has a continuous range of possible angles, giving an almost infinite number of possible hand gesture variations. Adding other factors, such as two hands/their relative position/other context, it becomes fairly difficult to categorize hand gestures. For this, a Hand Gesture Taxonomy[1] had to be found or developed.

3.2   Hand Gesture Taxonomy

Previous studies in hand gestures focused on using hand motion, as a part of body motion, to describe body language. The objective, for body language, was different in nature than HMI with hand gestures since there was no need to characterize hand gestures with such a detail as would be needed for HMI. This could be due to the fact that, in the

---

[1]A taxonomy is a technique of hierarchical classification. It is used to name defining class characteristics that can be used to distinguish one class from another.

case of body language, the observer is a human that can register and understand a variety of hand movements, no matter how subtle, without any conscious thinking. In HMI, the observer of hand gestures is a computer analyzing images, signals, etc. So in order to define any gestures used in HMI, their characteristics would need to be as detailed (and as clear) as possible. Vafaei's work [22] made this possible. He described a Hand Gesture Taxonomy that adopted previous work and adapted it to be used in HMI. His taxonomy contained a wide variety of distinctive hand features that would make hand gestures easily defined, and hence distinguishable.

His proposed taxonomy can be seen in table 5. A description of it is shown below. The taxonomy is described by 11 dimensions, divided into two groups:

- Group 1: Gesture Mapping

  These dimensions describe how gestures are mapped to tasks, this includes:

  1. Nature:

     Whether the gesture is manipulative, Pantomimic (i.e. imitates an action), Symbolic (i.e. visually depicts a symbol), Pointing, or Abstract.

  2. Form:

     Whether the gesture is static (no motion/change in gesture), dynamic (has motion/change in gesture), or a stroke (consists of tapping motion).

  3. Binding:

     Whether the gesture is bound to a certain object, reference point in a room, or is world independent.

4. Temporal:

   Whether the gesture is continuous (i.e. action performed during gesture) or discrete (i.e. action occurs after completing gesture).

5. Context:

   Whether or not the gesture requires prior context.

- Group 2: Physical Characteristics

  These dimensions describe the physical attributes of a gesture, including:

  1. Dimensionality:

     Whether the gesture occurs in 1 axis (linear translational motion), 2 axes (planar translational motion), 3 axes (translational/rotational motion in 3D space), or 6 axes (gesture both translational and rotational motion in 3D space).

  2. Complexity:

     Whether the gesture consists of a simple gesture (simple) or a sequence of hand shapes (Compound).

  3. Body Part:

     What body part(s) is involved when in performing the gesture.

  4. Handedness:

     Whether the gesture is performed by dominant hand, non-dominant, or both.

  5. Hand Shape:

     Description of hand shape, fingers, etc.

6. Range of Motion (ROM):

Whether joint rotation is small or large, compared to joint's ROM (i.e. whether a gesture is subtle or not).

Table 5: Hand Gesture Taxonomy

Group	Dimension	Type
Gesture Mapping	Nature	Manipulative, Pantomimic, Symbolic, Pointing, Abstract
	Form	Static, Dynamic, Stroke
	Binding	Object-Centric, World-Dependent, World Independent
	Temporal	Continuous, Discrete
	Context	In-Context, No-Context
Physical Characteristics	Dimensionality	Single-Axis, 2-Axes, 3-axes, 6-Axes
	Complexity	Simple, Compound
	Body Part	Hand, Arm, Head, Shoulder, Foot
	Handedness	Dominant, Non-Dominant, Bi-Manual
	Hand Shape	e.g. Flat, Open, Bent, Index Finger, Fist, ASL Shapes, ...
	ROM	Small, Large

3.3 Chosen Hand Gestures and Their Actions

After researching the hand gesture taxonomy, it seemed best to choose well established/recognized hand gestures instead of inventing new gestures. For this reason, 13 hand signs were selected from the American Sign Language (ASL) alphabet, in addition to 4 other hand gestures, adding up to 17 hand gestures in total. Most of the gestures were chosen due to them representing their mapped action, either visually or by other

means. The chosen gesture can be viewed in figure 4 and their mappings can be found in table 6. Then the characteristics of the selected hand gestures, according to the proposed taxonomy, were logged. They can be found in tables A.1 - A.4 in the Appendix.
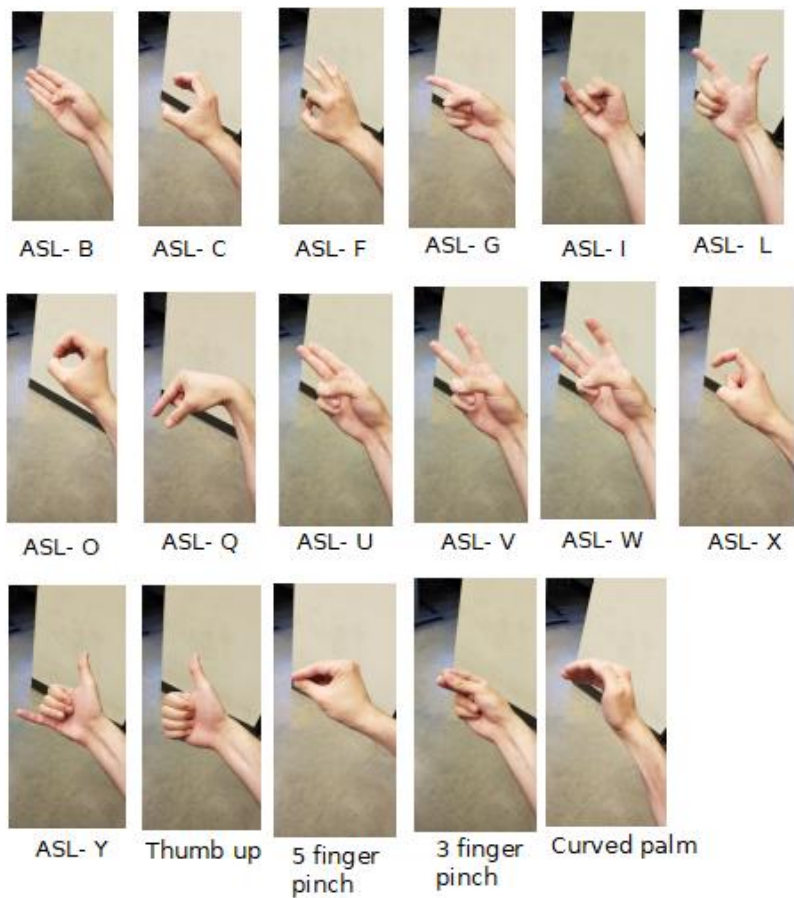
Figure 4: Visual Representation of Chosen Hand Gestures

Table 6: Mapping of Chosen Hand Gestures

Command Category	Task	Hand Shape
Editing	Fillet	ASL - U
	Chamfer	Curved Palm
	Trim	ASL - V
	Extend	ASL - L
	Move	3-finger pinch
	Copy	ASL - C
	Delete	ASL - Q
	Scale	ASL - F
	Smart Dimensions	ASL - I
Manipulate	Change Orientation	5-finger pinch
Actions	Confirm	thumbs up
	Undo	ASL - Y
	Cancel	ASL - W
	Exit Sketch	ASL - B
	Toggle Mouse Movement	ASL - G
	Mouse Click	ASL - X
3D Draw	Revolve	ASL - O

CHAPTER 4

DATA COLLECTION

4.1   IRB Protocol

Before collecting any data, approval was needed from the Institutional Review Board (IRB) at the University of Missouri – Kansas City. To attain IRB approval, all the investigators involved in the study, in addition to any research assistants, had to undergo the Collaborative Institutional Training Initiative (CITI) training titled "Group 1 Biomedical Investigator (UMKC and TMC)". This training assures that the investigators are knowledgeable in how to ensure the collected data's privacy and the safety of the individuals (or animals) that the data is collected from.

The investigative team involved in the research consisted of an associate professor (Principal Investigator), a masters student (Co-Investigator), and a bachelors student (Research assistant). The co-investigator conducted the data collection sessions, while the research assistant assisted with the data preprocessing and segmentation. The IRB approval form can be found in Appendix-B.

Upon approval from the IRB at the University of Missouri – Kansas City, 12 volunteers participated in the study. EMG and IMU data was collected, via the Myo armband, from the 12 subjects over the course of 3 visits. The procedures were explained to the volunteers and they each provided written consent prior to starting data collection. The consent form can be found in Appendix-B.

4.2 Data Collection Process

In the beginning of a data collection session (i.e. visit), the volunteer would wear the Myo armband on their right forearm in the manner recommended by Thalmic Labs, the Thalmic Labs Logo facing upwards (forearm posterior) and the USB port facing the wrist. Then it is allowed to "warm up" for a few minutes. Once warmed up, the Myo connects to the computer used for data collection. A syncing gesture, required whenever the Myo is to be worn, is performed. The syncing gesture consists of a full wrist extension at maximum force. The amplitude of maximal contraction is used to normalize the 8 signals. Once the Myo is synced and connected to the computer, then it is possible to start using the MATLAB data collection script. The script generated a Graphical User Interface (GUI) that allowed the co-investigator to key in the subject number, the gesture number, the set[1] number, and the visit number. It also showed a timer. It will be discussed in the next section.

During the data collection session, the volunteer repeats a gesture 60 times, in sets of 20, while wearing the Myo armband. This is repeated for the 17 hand gestures while taking scheduled breaks in between sets and gestures. Extra breaks were also taken as needed. This process was repeated for 3 visits with the 12 subjects.

The data sessions were also recorded with a camera mounted onto the computer collecting the data. In its field of view were a secondary monitor, that had the GUI in view, and the participant's hand that the data was collected from. The purpose of the video was to review the footage for any mistakes done during the session, so those parts

---

[1]Also referred to as "section".

28

of the EMG data streams could be cut out.

4.3   Data Collection Script

A custom MATLAB script was written to be used in data collection. The script was built upon an open-source MATLAB script, MyoMex[2]. Using MyoMex, a GUI was designed in MATLAB that allowed the researcher to key in the subject number, visit number, gesture number, and section number.

The section number made it possible to collect a gesture's data over several time periods (sets), allowing for rest periods in between. It could also allow for the collection of one sample at a time, eliminating the need for later data segmentation. Although that would make the data collection experience unpleasant for all those involved in the session.

The data collection GUI also contained keys to Start and Stop data recording and a timer to keep track of time. The timer was used as a reference to use when cutting out mistakes from the collected data. The GUI layout can be viewed in figure 5.

This script was meant to be used in order to streamline the data collection process and, once packaged and available to public, would allow for other researchers to also collect data using the Myo armband. This would also allow for standardization of the EMG data collection process and allow other researchers to contribute to expanding this data set into a larger, publicly available, benchmark data set for EMG-based Hand Gesture Recognition.

---

[2]"https://www.mathworks.com/matlabcentral/fileexchange/55817-myo-sdk-matlab-mex-wrapper"

29
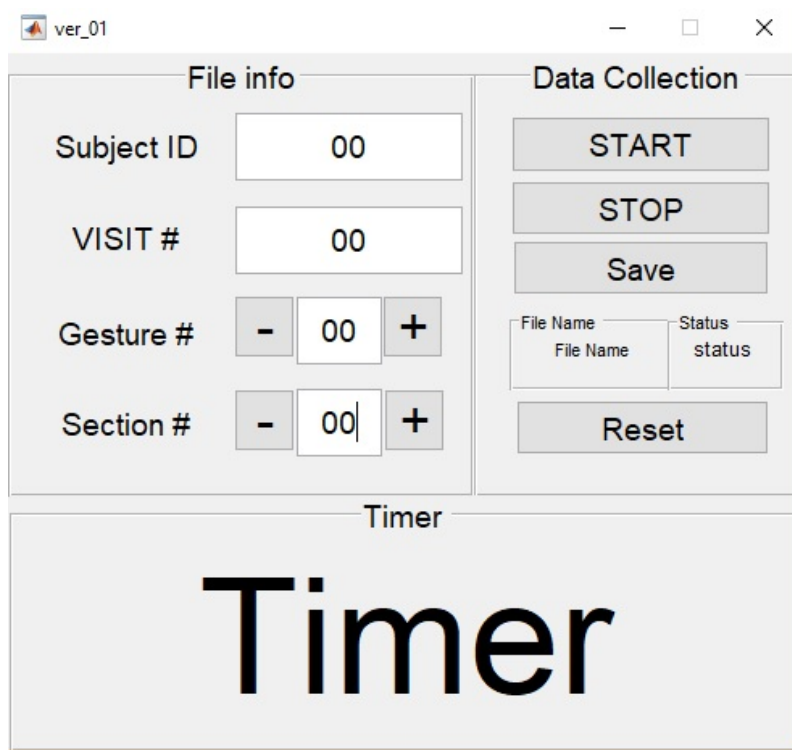
Figure 5: Data Collection GUI

To decrease the burden on the subjects during data collection, data recording continued running through the session, even when mistakes were made. In addition to that, if a sample for a certain gesture was forgotten, that sample would be performed later in the session and a note would be made regarding it. The method of dealing with such circumstances will be described in Chapter 5.

4.4 Collected Dataset

The final collected dataset consisted of 17 hand gestures, each repeated 180 times (from 3 visits) per subject. This resulted in each gesture class having 2160 samples (60 times x 3 visits x 12 subjects), i.e. 36720 samples in total.

CHAPTER 5

SIGNAL PREPROCESSING

Before feeding the EMG data into any of the ML algorithms, the collected data first had to go through processing to have it in a format suitable for the ML scripts. This included editing the collected data, segmenting it into individual samples, and extracting features, which will be discussed in the following sections.

5.1   Sample Editing

After data collection, the first step of the data preprocessing was to edit the data: cropping out the parts where the subject made any mistakes, such as performing the wrong hand gesture or making a mistake in performing the hand gesture. For this, a custom-made MATLAB scripts (Editing Script) was written to perform the Delete, Copy, and Paste operations on the collected data. The Delete operation allowed to delete the parts of the signal where mistakes, as mentioned above, were made. The Copy and Paste operations allowed a forgotten sample (done at the end of the session) to be taken from the end of the sample set and be pasted into its correct place, while ensuring that no sample exists twice in the data set, to keep the number of samples per class per subject relatively consistent.

This script would start by plotting the 8 signals in 8 subplots (figure 6) and ask the user to select which operation to perform: Copy, Delete or Paste. Then the user selects 1 of the 8 signals to perform this operation on by simply clicking on the subplot, directly

from the figure window. Once a signal is selected, it is re-plotted, separately, in a larger window. There the user can select a segment to copy by clicking on beginning and end points straight from the plot. The same process applies for deleting a segment. There is no limit on how many samples can be copied. For segment deletion, the user can keep deleting until there are no data points left.

For pasting a segment, a figure with subplots with all the copied samples from that gesture is shown. The user then selects a copied segment from the list and then selects a point on the original signal plot as a starting point to insert the selected copied segment. These operations were repeated until all the mistakes/notes were taken care of.
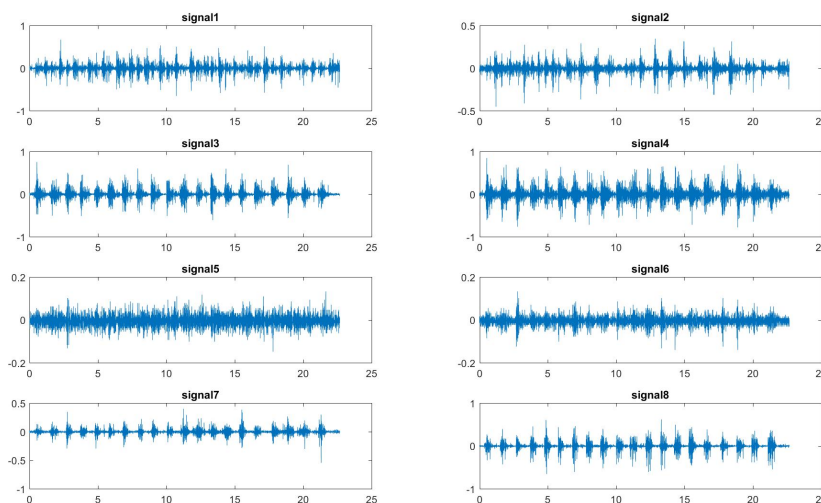
Figure 6: Original EMG Plots (Set of 20 samples)

To demonstrate the script's output, the EMG sample set in figure 6 is used. The first sample (first bi-directional spike, most visible in plot of signal 3) is copied, then the next 19 samples are deleted. The first sample is then pasted 4 times after the first sample.

This results in 5 repetitions of the first sample, as shown in figure 7.
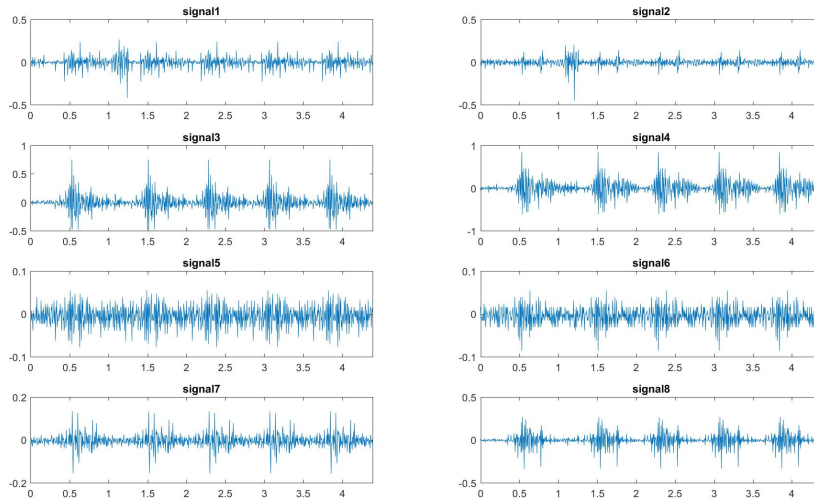
Figure 7: Demo of Edited EMG Plots

5.2   Sample Segmentation

Since the EMG samples were collected in batches (20 at a time), the collected data needed to be segmented into individual samples. An EMG Onset Detection algorithm was implemented to automatically detect the time instances in which a hand gesture had occurred and segment the collected data accordingly. The process was performed by means of a custom-made MATLAB script.

The Segmentation script starts by plotting all 8 signals in 8 subplots – 20 samples per subplot. Overlaying each signal plot is a low-pass filtered (LP) version of the fully-rectified (FR) signal (with a specific starting cutoff frequency). Overlaying those are the thresholds for each signal, which is the Mean Absolute Value (MAV) of the FR-LP signal

multiplied by a certain starting weight. Using this threshold, a rectangular signal (set of windows) was produced for each of the 8 signals. One of the 8 window sets will be used to segment all 8 signals into individual segments. A preview is shown in figure 8. This demonstrates how the example from the previous section (5 copied samples) would have been segmented.
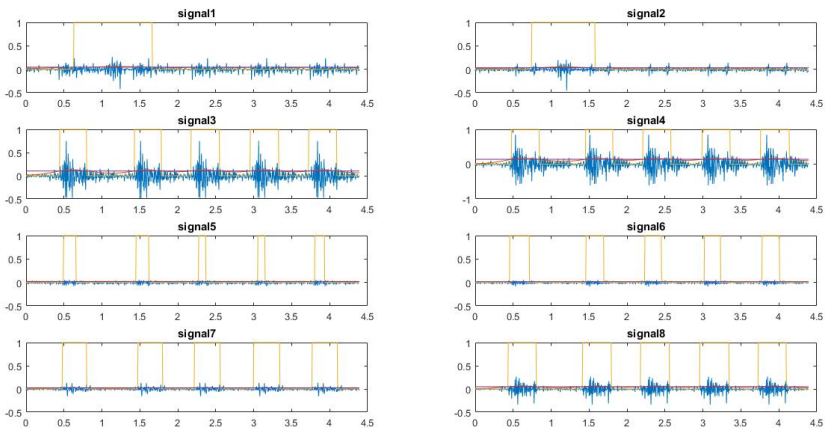
Figure 8: Segmentation of 5 copied samples from previous section: The blue line is the edited signal. The red line is the LPF[ABS(x(t))] signal. The purple is the MAV multiplied by a certain constant. The yellow line is the window set generated by thresholding.

The segmentation script allows the user to change the cutoff frequency and the MAV-weight until at least one of the window sets appropriately divides the signal up into 20 distinct samples. Then the user can select which of the 8 window sets best segments the signals into 20 samples. The selected window set would then be used for segmentation over all 8 channels. After segmentation, the individual samples are saved.

Figure 9 shows how increasing the frequency from 1 Hz (figure 8) to 10 Hz can change the 8 window sets.
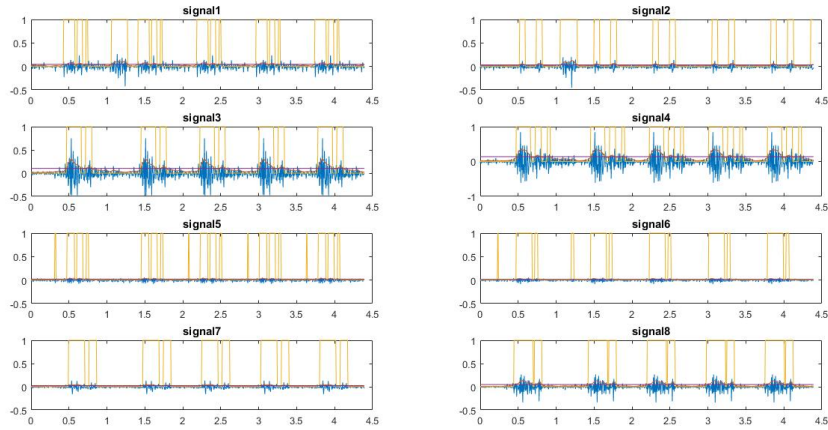
Figure 9: Segmentation of Demo samples using LPF frequency of 10 Hz.

The effect of increasing the MAV weight from 1.1 (figure 8) to 1.3 can be observed in figure 10, the effect of decreasing it to 0.9 is shown in figure 11. It clearly shown how increasing the MAV weight decreases the window widths and decreasing it increases the window widths.

5.3   Preprocessed Data Set

After the preprocessing, some hand gestures and subjects were excluded due to them being fairly difficult to segment using the segmentation script. Segmenting those samples would require knowing the starting and ending points of each sample individually. Although cumbersome, it maybe done in the future to increase the size of the data set.

The final preprocessed data set consists of 10 subjects performing 10 hand gestures, each repeated 162 times (from 3 visits) per subject. This resulted in each gesture
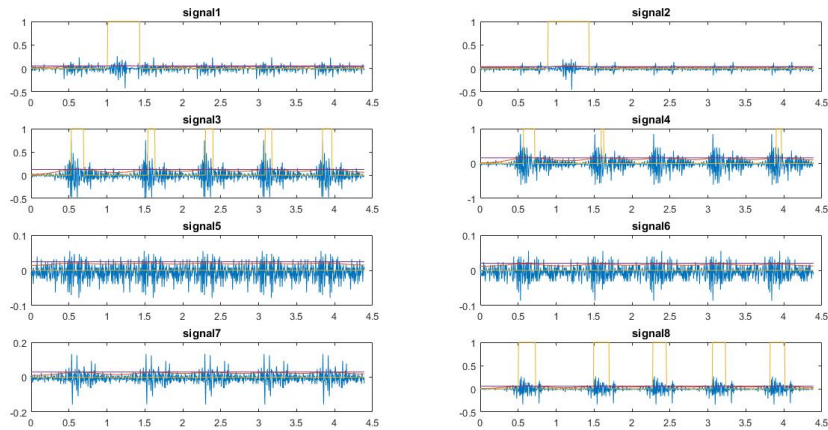
Figure 10: Segmentation of 5 copied samples from previous section (increasing MAV weight from 1.1 to 1.3).

class having 1620 samples (54 times x 3 visits x 10 subjects) per gesture class, i.e. 16200 samples in total. The 10 hand gestures were the 5-finger pinch and the thumbs up gestures, in addition to the ASL letters U, V, L, F, I, Y, B and G.

5.4   Generating the Final Data Set, Targets, and Labels

After the data was segmented it was possible to rearrange the data into the final data set. All the samples were arranged in a 10x10 cell array such that each cell had the samples of 1 subject for 1 gesture for all visits:

(54 samples) x (3 visits) = 162 samples per subject per gesture.

Once the data set cell array was created, 2 similar cell arrays were created with the samples' corresponding targets and labels. This would make it much easier when dividing the data sets into training, validation, and testing sets.

The distribution of the sample lengths was plotted, showing that the sample lengths,

36
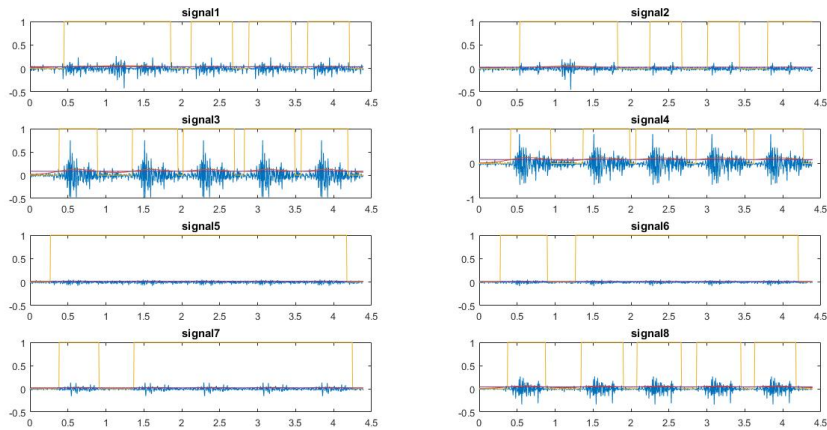
Figure 11: Segmentation of 5 copied samples from previous section (decreasing MAV weight from 1.1 to 0.9).

for all the data, ranged from 62 data points to 462 data points per channel. Their plot can

be seen in figure 12. The ML algorithms used cannot take samples of different lengths, so

all the samples were padded with zeros to make all their lengths equal to 462. The data

set was then normalized from 0 to 1.
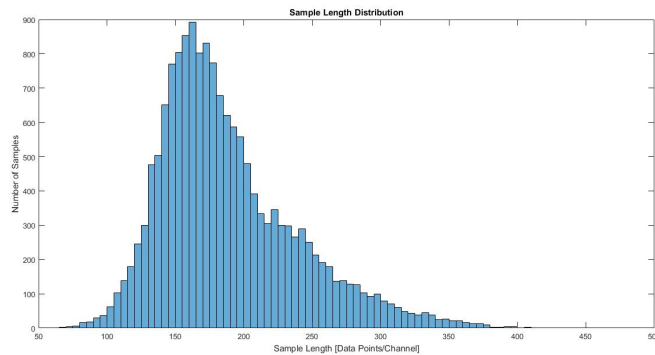
Figure 12: EMG Sample Length Distribution

5.5 EMG Features

Eleven feature sets, used in the literature, were extracted for each of the 8 signals. They are:

1. Root Mean Square (RMS)

   The root mean square for a channel (c) is calculated according to the following equation, resulting in 1 value per channel:

$$RMS\left(c\right) = \sqrt{\frac{1}{w}\sum_{i=1}^{w} x_c\left(i\right)^2} \qquad (5.1)$$

2. Mean Absolute Value (MAV)

   The mean absolute value for a channel (c) is calculated according to the following equation, resulting in 1 value per channel:

$$MAV\left(c\right) = \frac{1}{w}\sum_{i=1}^{w} \left|x_c\left(i\right)\right| \qquad (5.2)$$

3. Variance (Var)

   The Variance for a channel (c) is calculated according to the following equation, resulting in 1 value per channel:

$$Var\left(c\right) = \frac{1}{w}\sum_{i=1}^{w} x_c\left(i\right)^2 \qquad (5.3)$$

4. Mean Frequency (Freq-mean)

   The mean frequency, in terms of the sampling frequency (fs=200 Hz) is calculated for a channel (c), resulting in 1 value per channel.

5. Median Frequency (Freq-med)

   The median frequency, in terms of the sampling frequency (fs=200 Hz) is calculated for a channel (c), resulting in 1 value per channel.

6. Integrated Absolute Value (IAV)

   The Integrated Absolute value for a channel (c) is calculated according to the following equation, resulting in 1 value per channel:

$$IAV(c) = \frac{1}{w} \sum_{i=1}^{w} x_c(i) \tag{5.4}$$

7. Wilson Amplitude (WAMP)

   The Wilson Amplitude for a channel (c) is calculated according to the following equation, resulting in 1 value per channel:

$$WAMP(c) = \sum_{i=1}^{w} f\left(|x_c(i) - x_c(i+1)|\right) \tag{5.5}$$

where

$$f(x_c) = \begin{cases} 1 & x_c < 0 \\ 0 & otherwise \end{cases} \tag{5.6}$$

8. Frequency Amplitude Spectrum

   The amplitude spectrum for a channel (c) is calculated according to the equation below. The frequency bins are grouped into 5 bins of equal size and their average computed, resulting in 5 values for a channel (c).

$$A_c = \sum_{j=1}^{W} |fft_c(j)| \tag{5.7}$$

39

9. Energy of Approximate Wavelet Coefficients

A bi-orthogonal mother wavelet was used to decompose each signal into 8 scales. For each scale, the approximate wavelet coefficients are generated, and their energies calculated, resulting in 8 values per channel. The wavelet transform equation is shown below:

$$F(a, b) = \int_{-\infty}^{\infty} f(x)\psi_{(a,b)}^{*}(x)dx \tag{5.8}$$

where the $*$ is the complex conjugate symbol and function ψ is a mother wavelet function.

10. All feature sets (1 to 9)

This feature set is a concatenation of the feature sets 1 to 9 for each sample, resulting in 20 values per channel.

11. Spectrograms (reduced dimensionality)

This feature set resulted from generating a channel c's spectrogram, using 8 segments and 50% overlap. Then the absolute values of all 8 spectrograms are concatenated, generating an 8-channel image.

A spectrogram is generated using the short-time Fourier Transform, described in the following equation:

$$STFT(X_c(t))(\tau, \omega) \equiv X_c(\tau, \omega) \tag{5.9}$$

$$X_c(\tau,\omega) = \int_{-\infty}^{\infty} x_c(t)\,\omega\,(t - \tau)\,e^{-j\omega t}dt \tag{5.10}$$

40

After generating the spectrograms, Principal Component Analysis (PCA) is used to reduce their dimensionality using the principle components corresponding to 90% of the covariance matrix, resulting in 536 values per channel (compared to the original 1806 values per channel). (i.e. the spectrogram's dimensionality was reduced to less than a third of its original size.)

Upon generating a feature set for a given sample, it is then normalized from 0 to 1 across all channels (individually) to maintain consistency among subjects/gestures. The generated feature sets, and their corresponding targets and labels, were then divided into 80%, 10%, and 10% for training, validation and testing respectfully in accordance with a subject independent scheme. (i.e. using data of different subjects for training, validation and testing.)

CHAPTER 6

MACHINE LEARNING MODELS USED

Machine Learning (ML) models have been used on quite a variety of data, including image data, audio, video and biosignals (e.g. EMG signals). Due to the small size of the data set, the ML models used were statistical-based. These models were used individually and as in committees. In the future, when the data set reaches a sufficient size, other ML models (such as deep learning models) can be used and compared. The models used are described below.

6.1 Statistical-based Methods

Statistical-based methods are, as the name implies, methods that involve using a data set's populational distribution in order to classify it. The methods used are briefly described below:

- Support Vector Machine (SVM):

  SVMs binary classifiers work by projecting the data into a higher dimensional space, in which the data become linearly separable, separating the data, then bringing the boundary back to the main space and identifying each class' edge examples, A.K.A Support Vectors (SVs). Through this method, the margin between the SVs and the boundary is then maximized, which maximizes generalizability. The SVM

model used consists of 10 binary SVMs (in a one-vs-all coding scheme), each having a Gaussian kernel, a kernel scale of 1, and a box constraint of 10.

- K-Nearest Neighbor (KNN):

  The KNN classifier works by taking each sample in the data set and looking at its "K" nearest neighbors. The class with the higher neighbor-subset in the nearest neighbors is selected as the samples class. The KNN model used 3 nearest neighbors, and a maximum of 50 data points in a leaf node and Euclidean distance as a distance metric (in its KD-tree nearest neighbor search method).

- Linear Discriminant Analysis (LDA):

  A Fisher LDA classifier projects the data into a subspace in order to maximize the inter-class distance and intra-class scatter. The LDA model (Fisher) used has a regularized linear discriminant. It uses 1 covariance matrix generated from the whole data set.

- Naïve Bayesian (NB):

  The NB classifier classifies the data (using Bayes' Rule) based on the data's prior probabilities and maximum likelihood. The NB model uses an unbounded Gaussian kernel to model the data. The class prior probabilities are equal for all classes.

A 9-fold cross validation was performed using 8 subjects' data for training and 1 subject's data for validation in each fold. Performance of the 4 models was compared by means of the Area Under Curve (AUC) and Equal Error Rate (EER) metrics obtained

43

from the Receiver Operating Characteristic (ROC) curve on the validation data. Then all models were tested using the testing data set, and the best performing model (overall) was selected based on the best validation performance. The best performing model for the test subject was selected based on the best performing testing model. The results will be discussed in the next section.

At this point, it is worth pointing out that the classification results in Chapter 2 are reported in terms of accuracy, as apposed to that of this work, which are in terms of ROC's AUC and EER. The AUC and EER metrics are more generalized and would still hold true in circumstances where the data set's sample distribution is uneven among all classes. To ease the comparison of this work with others, it is important to note that a model's classification accuracy can be easily obtained by the following formula:

$$Accuracy = 1 - EER \tag{6.1}$$

6.2 Committees

A scheme (and script) were developed to fuse the 4 models together, at score level, to form committees. Although some tests were performed, the results were still inconclusive. More tests on committees will be done in the future to find out whether or not fusing these 4 classifiers would, in fact, yield a greater classification performance than that of the individual models. The proposed score fusion scheme is described below:

This scheme would account for all possible combinations of the 4 models, adding up to 11 possible combinations in total (6 permutations of 2 models + 4 permutations of

3 models + 1 permutation of 4 models).

The combinations can be seen below:

Table 7: All possible combinations of the SVM, KNN, LDA, and NB models

Combination	No. of Classifiers	Classifiers
1	2	SVM, KNN
2	2	SVM, LDA
3	2	SVM, NB
4	2	KNN, LDA
5	2	KNN, NB
6	2	LDA, NB
7	3	SVM, KNN, LDA
8	3	SVM, KNN, NB
9	3	SVM, LDA, NB
10	3	KNN, LDA, NB
11	4	SVM, KNN, LDA, NB

For each possible combination, the scores were fused using the following 8 methods:

1. Weighted sum using EER:

   Score weights are generated using the reciprocal of their ROC's Equal Error Rate (EERs), after rescaling the reciprocals to sum up to 1. Then a weighted sum of the individual model scores is performed to obtain the final fused scores. This method is described in mathematical form below:

   $$S_{w\ sum} = sum\left(w_1 S_1, \ldots, w_n S_n\right) \qquad (6.2)$$

   where

   $$w_n = \frac{1}{EER_n} \qquad (6.3)$$

45

2. Weighted sum using AUC:

Score weights are generated using their ROC's Area Under Curve (AUCs), after rescaling them to sum up to 1. Then a weighted sum of the individual model scores is performed to obtain the final fused scores. This method is described in mathematical form below:

$$S_{w\ sum} = sum\left(w_1 S_1, \ldots , w_n S_n\right) \tag{6.4}$$

where

$$w_n = AUC_n \tag{6.5}$$

3. Weighted sum using EER and AUC:

Score weights are generated using the product of their AUCs and the reciprocal of their EERs and rescaling them to sum up to 1. Then a weighted sum of the individual model scores is performed to obtain the final fused scores. This method is described in mathematical form below:

$$S_{w\ sum} = sum\left(w_1 S_1, \ldots , w_n S_n\right) \tag{6.6}$$

where

$$w_n = \frac{AUC_n}{EER_n} \tag{6.7}$$

4. Product:

Scores from each model are multiplied together, element by element (also referred to as Hadamard Product), to obtain final fused scores. This method is described in mathematical form below:

$$S_{prod} = S_1 \circ \dots \circ S_n \tag{6.8}$$

5. Max:

Scores from each model are compared, element by element, and the maximum from each element is used to obtain the final fused scores. This method is described in mathematical form below:

$$S_{max} = max(S_1, \dots , S_n) \tag{6.9}$$

6. Min:

Scores from each model are compared, element by element, and the minimum from each element is used to obtain the final fused scores. This method is described in mathematical form below:

$$S_{min} = min(S_1, \dots , S_n) \tag{6.10}$$

7. Median

Scores from each model are compared, element by element, and the median from each element is used to obtain the final fused scores. This method is described in

mathematical form below:

$$S_{median} = median(S_1, \dots , S_n) \tag{6.11}$$

8. Mean:

Scores from each model are compared, element by element, and the mean from each element is used to obtain the final fused scores. This method is described in mathematical form below:

$$S_{mean} = mean(S_1, \dots , S_n) \tag{6.12}$$

Through these methods, a total of 88 committees (11 permutations x 8 score fusion methods), would be evaluated using 9-fold cross validation and then their classification performance can then be compared with that of the 4 individual models.

CHAPTER 7

RESULTS AND DISCUSSION

After training the 4 classifiers with the 11 feature sets, and performing 9-fold cross validation, the best performing models were evaluated with the testing data set. The individual results will be discussed in the following sections.

7.1   Validation

The validation AUC and EER results for the 4 individual classifiers can be viewed in tabular format in tables 8 and 9, respectfully. The validation AUCs and EERs can also be viewed in figures A.1 and A.2 in the appendix[1].

1. SVM

   As can be seen in tables 8 and 9, the best performing feature set for the SVM classifier is the spectrogram, with an average AUC 75.8% and EER of 29.87%. It is followed by the feature set 10, with an average AUC of 69.18% and EER of 35.04%. The third, fourth, and fifth best feature sets were the RMS (average AUC=68.4%, EER=36.6%), the MAV (average AUC=68.39%, EER=36.7%), and IAV (average AUC=68.39%, EER=36.7%) respectfully. They are followed by the Var (average AUC=65.6%, EER=38.76%), Frequency Amplitude Spectrum (average AUC=66.04%, EER=38.01%), and the Energy of the Approximate Wavelet

---

[1]All AUC and EER results (for both figures and tables) are reported as a percentage

49

Coefficients (average AUC=64.49%, EER=39.6%).

The Worst performing feature sets for this classifier were the Mean Frequency, the Median Frequency, and the Wilson Amplitude, all three yielding average AUCs under 55% and EERs above 45%.

2. KNN

From observing tables 10 and 11, it can be seen that the KNN classifier is the worst performing classifier among the them all. Its best feature set, the spectrogram, only yielded an average AUC of 63% and an average EER of 37.3%. Following that are the MAV (average AUC=61.46%, EER=38.8%), the IAV (average AUC=61.46%, EER=38.8%), and feature set 10 (average AUC=61.48%, EER=38.86%). Its third best are the RMS (average AUC=60.65%, EER=39.56%) and the Frequency Amplitude Spectrum (average AUC=60.75%, EER=39.48%). Its fourth best are the Var (average AUC=59.88%, EER=40.29%) and the Energy of the Approximate Wavelet Coefficients (average AUC=59.4%, EER=40.94%).

Its worst performing feature sets were also the Mean Frequency, the Median Frequency, and the Wilson Amplitude, all three yielding average AUCs under 55% and EERs above 45%.

3. LDA

As can be seen in tables 10 and 11, it is clear that the LDA classifier is the best

performing classifier overall, with its best performing feature set being the spectrogram. It yielded an average AUC of 76.35% and an average EER of 29.73%. Its second best was the RMS (average AUC=75.68%, EER=30.83%), the MAV (average AUC=75.25%, EER=30.81%), and the IAV (average AUC=75.7%, EER=30.81%). Its third best performing feature sets were the VAR (average AUC=74.66%, EER=31.5%), and the Frequency Amplitude Spectrum (average AUC=74.81%, EER=31.27%). Surprisingly, feature set 10 only yielded an average AUC of 73.68% and an average EER of 31.77%, compared to the feature sets summing up to it. Following that was the Energy of the Energy of Approximate Wavelet Coefficients (average AUC=70.81%, EER=34.41%). Although still in the last 3 feature sets, the Mean Frequency (average AUC=60.61%, EER=42.06%) and the Median Frequency (average AUC=58.95%, EER=43.3%) performed better with the LDA classifier than with the other classifiers.

Its worst performing feature sets the Wilson Amplitude, yielding average AUC of 52.7% and EER of 48.2%.

4. NB

Looking at tables 10 and 11, one can see the best performing feature set for the NB classifier is the RMS, yielding an average AUC of 73.26% and an average EER of 31.6%. Closely following the RMS were the MAV (average AUC=72.75%, EER=32.13%) and the IAV (average AUC=72.75%, EER=32.13%). The third best performing feature sets was the Var (average AUC=71.39%, EER=33.42%) followed by the frequency Amplitude Spectrum (average AUC=68.63%, EER=33.08%).

The Energy of Approximate Wavelet Coefficients yielded an average AUC of 61.99% and an average EER of 38.95%, followed by feature set 10 (average AUC=31.57%, EER=38.58%), the Mean Frequency (average AUC=60.8%, EER=41.77%), and the Median Frequency (average AUC=58.96%, EER=43.15%).

The worst performing feature sets for the NB classifier were the Wilson Amplitude, yielding an average AUC of 53.14% and an average EER of 48.1%, and surprisingly, the spectrogram, yielding an average AUC of 53.19% and an average EER of 47.59%.

Looking at the overall classification performance from all 4 classifiers, it can be observed that the RMS, MAV, VAR and IAV feature sets all had similar performances. This could suggest that these feature sets may be interchangeable, hence using all 4 of them may be redundant.

It can also be observed that the Mean Frequency, the Median Frequency, and the Wilson Amplitude feature sets had the worst overall performance among the 4 classifiers. This poor performance may be due to the fact that the Myo armband collects data at a sampling frequency of 200 Hz, while EMG signals are usually collected at a minimum of about 1000 Hz [23]. This low sampling frequency could be causing some aliasing in the collected data.

To compare classifiers, it can be seen that the best overall performers were the LDA and NB, followed by the SVM classifier. The KNN classifier performed the worst (overall) among the 4 classifiers.

Table 8: Validation AUCs of Individual Classifiers

Feature set	SVM	knn	LDA	NB
RMS	68.44 ±0.62	60.65 ±0.38	75.68 ±1.02	73.26 ±1.16
MAV	68.39 ±0.70	61.46 ±0.40	75.25 ±1.03	72.75 ±1.14
Var	65.60 ±0.60	59.88 ±0.41	74.66 ±0.97	71.39 ±0.79
Freq. (mean)	52.20 ±0.01	52.14 ±0.02	60.61 ±0.09	60.80 ±0.12
Freq. (median)	51.39 ±0.01	51.88 ±0.01	58.95 ±0.07	58.96 ±0.08
IAV	68.39 ±0.70	61.46 ±0.40	75.25 ±1.03	72.75 ±1.14
W. Amplitude	50.51 ±0.01	50.30 ±0.01	52.70 ±0.11	53.14 ±0.13
Freq. Amp. Spect.	66.04 ±0.71	60.75 ±0.42	74.81 ±0.75	68.63 ±0.84
Energy (Wavelets)	64.49 ±0.37	59.40 ±0.18	70.81 ±0.68	61.99 ±0.43
FS1:FS9	69.18 ±0.80	61.48 ±0.26	73.68 ±0.97	61.57 ±0.44
Spectrogram	75.80 ±0.56	63.03 ±0.13	76.35 ±0.52	53.19 ±0.08

7.2 Testing

The testing AUC and EER results for the 4 individual classifiers can be viewed in tabular format in tables 10 and 11, respectfully. The testing AUCs and EERs can also be viewed in figures A.3 and A.4 in the appendix.

1. SVM

   As can be seen in tables 10 and 11, the best performing feature sets from the SVM classifier were the RMS (average AUC=83.56%, EER=24.8%), MAV (average AUC=83.23%, EER=24.37%), and the IAV (average AUC=83.23%, EER=24.37%). Those were followed by the Var (average AUC=91.62%, EER=26.62%), the Frequency Amplitude Spectrum (average AUC=80.47%, EER=25.63%), feature set 10 (average AUC=79.44%, EER=27.18%), and the Energy of Approximate Wavelet Coefficients (average AUC=76.15%, EER=31.1%). The Spectrogram yielded an average AUC of 62.24% and an average EER of 39.88%.

Table 9: Validation EERs of Individual Classifiers

Feature set	SVM	knn	LDA	NB
RMS	36.64 ±0.34	39.56 ±0.36	30.83 ±0.60	31.60 ±0.57
MAV	36.71 ±0.39	38.81 ±0.38	30.81 ±0.55	32.13 ±0.53
Var	38.76 ±0.30	40.29 ±0.38	31.52 ±0.50	33.42 ±0.34
Freq. (mean)	48.42 ±0.01	47.92 ±0.02	42.06 ±0.05	41.77 ±0.07
Freq. (median)	49.00 ±0.01	48.13 ±0.01	43.30 ±0.03	43.15 ±0.04
IAV	36.72 ±0.39	38.81 ±0.38	30.81 ±0.55	32.13 ±0.53
W. Amplitude	49.55 ±0.01	49.74 ±0.01	48.22 ±0.05	48.05 ±0.07
Freq. Amp. Spect.	38.01 ±0.45	39.48 ±0.39	31.27 ±0.40	33.08 ±0.60
Energy (Wavelets)	39.60 ±0.17	40.94 ±0.17	34.41 ±0.34	38.95 ±0.34
FS1:FS9	35.04 ±0.53	38.86 ±0.24	31.77 ±0.55	38.58 ±0.38
Spectrogram	29.87 ±0.33	37.30 ±0.11	29.73 ±0.30	47.59 ±0.06

The Worst performing feature sets for the SVM classifier were the Mean Frequency, the Median Frequency, and the Wilson Amplitude, all three yielding average AUCs under 55% and EERs above 45%.

2. KNN

From tables 10 and 11, it is shown that the KNN its the worst performer, overall, among the 4 classifiers during testing. Its highest performing feature sets were the MAV (average AUC=72.17%, EER=28.66%) and the IAV (average AUC=72.17%, EER=28.66%). Its second-best performers were the RMS (average AUC=70.65%, EER=30.41%) and the Frequency Amplitude Spectrum (average AUC=80.47%, EER=30.5%). Closely following were feature set 10 (average AUC=69.86%, EER=31.22%) and the VAR (average AUC=69.22%, EER=31.67%). The Energy of Approximate Wavelet Coefficients yielded an average AUC of 65.29% and an average EER of 35.37%.

The Worst performing feature sets for the KNN classifier were, as per usual, the Mean Frequency, the Median Frequency, and the Wilson Amplitude, in addition to the spectrogram, all yielding average AUCs under 55% and EERs above 45%.

3. LDA

Referring to tables 10 and 11, The LDA classifier has some reasonably well performing feature sets, in comparison to the previous 2 classifiers. Its best performing feature set is the RMS, yielding an average AUC of 9103% and an average EER of 16.84%. Closely following it were feature set 10 (average AUC=89.64%, EER=18.73%), the IAV (average AUC=89.24%, EER=18.99%), the MAV (average AUC=89.24%, EER=18.99%), and the VAR (average AUC=89%, EER=19.47%). The Frequency Amplitude Spectrum feature set yielded an average AUC of 88.23% and an average EER of 20.3%, and the Energy of Approximate Wavelet Coefficients yielded an average AUC of 84.14% and an average EER of 24%.

The worst feature sets for the LDA classifier were the Mean Frequency (average AUC=63.59%, EER=40.44%), the Wilson Amplitude (average AUC=62.14%, EER=41.62%), the Spectrogram (average AUC=60.1%, EER=41.75%), and then the Median Frequency (average AUC=57.98%, EER=44.43%).

It is worth noting here that even the worst performing feature sets performed better with the LDA classifier than they usually did with the other classifiers.

4. NB

It can be seen from tables 10 and 11 that the NB classifier has one of the best per-forming feature sets, among all the classifiers, for the testing evaluation. This best feature set was the RMS, yielding an average AUC of 92.07% and an average EER of 15.8%. It is closely followed by the Frequency Amplitude Spectrum (average AUC=90.11%, EER=17.38%), the IAV (average AUC=89.46%, EER=18.51%), the MAV (average AUC=89.46%, EER=18.51%), and the Var (average AUC=87.65%, EER=20.51%). Feature set 10 yielded an average AUC of 80.36% and an average EER of 21.16%, followed by the Energy of Approximate Wavelet Coefficients, with an average AUC of 74.51% and an average EER of 28.59%.

The worst performing feature sets for the NB classifier were the Mean Frequency (average AUC=63.7%, EER=40.44%), the Wilson Amplitude (average AUC=62.79%, EER=41.15%), the Median Frequency (average AUC=57.26%, EER=45.46%), and lastly the Spectrogram (average AUC=51.91%, EER=48.47%).

The testing results in this paper only reflect how the classifiers (and feature sets) performed with 1 test subject's data. Due to this fact, any observations here cannot be generalized to their performance in the real world. Their general performance can only really be observed by collecting data from a much larger number of subjects and evaluating based on that.

Keeping this fact in mind, it was observed that (for this test subject) the best per-forming feature sets, overall, were the RMS, the MAV, the VAR, the IAV, and the Frequency Amplitude Spectrum. The overall performance of the Energy of Approximate Wavelet Coefficients and feature set 10 varied among the different classifiers, performing

well for some (compared to other feature sets) while worse in others. The worst overall testing performance resulted from the Mean Frequency, the Median frequency, the Wilson Amplitude, and the Spectrogram feature sets.

So the overall best validation performance was achieved by the LDA classifier, using the Spectrogram feature set, achieving an average AUC of 76.35% and an average EER of 29.73%. Unfortunately, in testing this model, it only achieved an average AUC of 60.1% and average EER of 41.75%.

The best performing model for the test subject used was the NB classifier, with the RMS feature set, achieving an average accuracy of 92.07% and an average EER of 15.81%.

It is worth restating that these testing results are not absolute, as the testing data set consists of only one subject, hence it cannot be used to infer the performance of any of these models in the real world.

Table 10: Testing AUCs of Individual Classifiers

Feature set	SVM	knn	LDA	NB
RMS	83.56 ±0.03	70.65 ±0.01	91.03 ±0.01	92.07 ±0.01
MAV	83.23 ±0.05	72.17 ±0.02	89.24 ±0.02	89.46 ±0.06
Var	81.62 ±0.05	69.22 ±0.01	89.00 ±0.01	87.65 ±0.03
Freq. (mean)	53.49 ±0.00	52.94 ±0.00	63.59 ±0.00	63.70 ±0.00
Freq. (median)	51.31 ±0.00	52.20 ±0.00	57.98 ±0.00	57.26 ±0.00
IAV	83.23 ±0.05	72.17 ±0.02	89.24 ±0.02	89.46 ±0.06
W. Amplitude	53.24 ±0.00	53.09 ±0.00	62.14 ±0.00	62.80 ±0.00
Freq. Amp. Spect.	80.47 ±0.03	70.32 ±0.01	88.23 ±0.02	90.11 ±0.05
Energy (Wavelet)	76.15 ±0.04	65.29 ±0.01	84.14 ±0.01	74.51 ±0.05
FS1:FS9	79.44 ±0.06	69.86 ±0.01	89.64 ±0.01	80.36 ±0.08
Spectrogram	62.24 ±0.01	53.38 ±0.01	60.10 ±0.01	51.91 ±0.08

Table 11: Testing EERs of Individual Classifiers

Feature set	SVM	knn	LDA	NB
RMS	24.81 ±0.04	30.41 ±0.01	16.84 ±0.01	15.81 ±0.01
MAV	24.37 ±0.05	28.66 ±0.02	18.99 ±0.02	18.51 ±0.03
Var	26.62 ±0.05	31.67 ±0.01	19.47 ±0.01	20.51 ±0.05
Freq. (mean)	47.66 ±0.00	47.21 ±0.00	40.44 ±0.00	40.44 ±0.00
Freq. (median)	48.91 ±0.00	47.83 ±0.00	44.43 ±0.00	45.46 ±0.00
IAV	24.37 ±0.05	28.66 ±0.02	18.99 ±0.02	18.51 ±0.03
W. Amplitude	47.98 ±0.00	47.02 ±0.00	41.62 ±0.00	41.15 ±0.00
Freq. Amp. Spect.	25.63 ±0.01	30.50 ±0.01	20.30 ±0.02	17.38 ±0.03
Energy (Wavelet)	31.09 ±0.03	35.37 ±0.01	24.01 ±0.01	28.59 ±0.02
FS1:FS9	27.18 ±0.05	31.22 ±0.01	18.73 ±0.01	21.16 ±0.06
Spectrogram	39.89 ±0.01	46.58 ±0.01	41.75 ±0.00	48.47 ±0.07

CHAPTER 8

CONCLUSION

This work proposed performing hand gesture recognition from forearm EMG signals, using a commercial EMG armband (the Myo armband). Once recognized, the hand gestures can be mapped (using key binding) to certain command shortcuts in CAD software (Solidworks). This would provide an alternative to visual or glove-based HGR systems used in new emerging HMI technologies, such as Virtual Reality (VR) and Augmented Reality (AR).

EMG-based HGR has the advantages of being convenient (more so than glove-based HGR) and non-invasive, in terms of privacy, compared to visual-based HGR. If a high classification performance can be achieved, then the EMG-HGR systems could provide a reliable alternative to the existing HGR systems used in AR and VR.

Several classifiers are evaluated to perform this task, namely the Support vector Machine, the K-Nearest Neighbor, the Linear (Fisher) Discriminant Analysis, and the Naive Bayesian Classifier. These 4 classifiers were trained with and evaluated using 11 different feature sets, namely the Root Mean Square, the Mean Absolute Value, the Variance, the Mean Frequency, the Median frequency, the Integrated Absolute Value, the Wilson Amplitude, the Amplitude Frequency Spectrum, the Energy of the Approximate Wavelet Coefficients, a concatenation of the previous 9 feature sets, and the Spectrogram.

A Hand Gesture Taxonomy was used to describe hand gestures in order to define

the hand gestures to be used in this work. Then a dictionary of 17 hand gestures was selected, consisting of 13 American Sign Language (ASL) letter signs and 4 additional signs. After identifying the possible executable actions in Solidworks, the chosen hand gestures were assigned to suitable actions in Solidworks.

A data set of the selected hand gestures was collected. It consisted of EMG data from 12 subjects performing the 17 hand gestures, 180 times per gesture, over the span of 3 visits. There were no restrictions on the time taken to perform each hand gesture sample. Then the collected data underwent preprocessing and segmentation. The remaining data set consisted of 10 hand gestures from 10 subjects, each containing 162 samples per gesture. The previously mentioned feature sets were then extracted from the preprocessed data set.

The data was divided into training (80%), validation (10%) and testing (10%) in subject independent scheme. (i.e. 8 subjects for training, 1 for validation, and 1 for testing.)

A 9-fold cross validation was performed on all the models (4 classifiers + 11 feature sets). The best validation performance was average AUC of 76.35% and an average EER of 29.73% using the LDA classifier with the Spectrogram feature set. During testing, this model achieved an average AUC of 60.1% and average EER of 41.75% (table 12).

The best performing model for testing was the NB classifier with the RMS feature set, achieving an average accuracy of 92.07% and an average EER of 15.81% (table 13). For its validation this model achieved an average AUC of 73.26% and an average EER of 31.6%.

The reported testing results were only for 1 test subject and cannot be used to infer the model's real-world performance. A lot more data would need to be collected to deduce a statistically-sound evaluation of these models in the real world.

Table 12: Best Model (Based on Validation): LDA (Spectrogram)

Set	AUC	EER
Validation	76.35 $\pm$0.52	29.73 $\pm$0.30
Testing	60.10 $\pm$0.01	41.75 $\pm$0.00

Table 13: Best Model (Based on Testing): NB (RMS)

Set	AUC	EER
Validation	73.26 $\pm$1.16	31.60 $\pm$0.57
Testing	92.07 $\pm$0.01	15.81 $\pm$0.01

Based on the validation performance, either of the SVM, LDA, or NB models could be used for EMG-HGR. It is possible that combining these models, at score level, can yield better classifications results. Different feature set combinations can also be evaluated, in comparison with performance on individual feature sets.

This work proves that it is feasible to use forearm EMG signals for HGR and that it shows potential of becoming an alternate modality for HGR. Although promising, the EMG-HGR models need to achieve much better classification results in order to be considered as a reliable input modality for HMI. To be able to infer anything statistically significant, much more data would need to be collected and used to train and evaluate the HGR models.

CHAPTER 9

POSSIBLE FUTURE DIRECTIONS

As mentioned in 2.4, to have highly effective, intuitive EMG-HGR systems there is a need to accelerate research in ML with EMG signals. An integral factor in this would be to establish a benchmark data set for comparing results in EMG-HGR. Such a benchmark would have a statistically-significant number of subjects as well as an extensive hand gesture vocabulary, all collected by means of agreed upon data collection standards for forearm EMG signals. The data set collected in this study, although small, could serve as a beginning of such a benchmark, or at least a stepping stone towards one.

Future plans include making this data set available to public, which could help other researchers to apply their unique ideas to EMG signals without needing to worry about collecting the data themselves. With some modifications to the data collection script, it can also be uploaded online as a stand-alone package to allow other researchers to contribute to the data set's growth or even to for live EMG data streaming from the Myo armband to evaluate their own ML models in real time.

One significant addition to the data collection script would be to add a motion tracking system (glove/visual/etc.) to track the hand and finger digit movements. The finger kinematics data could be used to ease the segmentation process. It would also enable the possibility of attempting regression on the EMG signals to generate the hand kinematics.

EMG Data can also be collected from amputees in order to analyze it and compare it with able-bodied subjects' data to see how the same ML models perform on both data types. This could help create more IPC models.

In addition to that, other EMG features can be extracted, and different feature combinations can be evaluated through methods of feature subset selection. Other models (e.g. Neural Networks and their variations) can be trained and compared or their scores fused to find the best performing model. In addition to that, different committees can be created from the models and then evaluated to attempt achieving better classification results.

TABLES AND FIGURES

Table A.1: Gesture Mapping and Description - Part1

	Category	Manipulate	3D Draw	
	Task	Orientation	Revolve	Revolved Cut
Group 1: Gesture Mapping	Nature Form Binding	Manipulative Static World-Indep.	Pantomimic Static World-Indep.	Pantomimic Static World-Indep.
	Temporal Context	Discrete n	Discrete n	Discrete y
Group 2: Physical Characteristics	Dimension. Complexity Body Part Handedness	6-axis Simple hand+arm Dom. hand / Bimanual	Simple Hand Dom. hand	Compound Hand Dom. hand
	Hand Shape	5-finger pinch	ASL - O	ASL - O,V

Table A.2: Gesture Mapping and Description - Part2

	Category	Actions					
	Task	Confirm	Undo	Cancel	Exit Sketch	Toggle Mouse	Mouse Click
Group 1: Gesture Mapping	Nature Form Binding Temporal Context	Symbolic Static World-Indep. Discrete n	Abstract Static World-Indep. Discrete n	Abstract Static World-Indep. Discrete n	Abstract Static World-Indep. Discrete n	Pointing Static World-Indep. Discrete n	Pointing Static World-Indep. Discrete n
Group 2: Physical Characteristics	Dimension. Complexity Body Part Handedness	Simple Hand Dom. hand	Simple Hand Dom. hand	Simple Hand Dom. hand	Simple Hand Dom. hand	Simple Hand Dom. hand	Simple Hand Dom. hand
	Hand Shape	thumbs up	ASL - Y	ASL - W	ASL - B	ASL - G	ASL - X

Table A.3: Gesture Mapping and Description - Part3

	Category	Editing (1)			
	Task	Fillet/ Chamfer	Trim	Extend	Move
Group 1: Gesture Mapping	Nature Form Binding Temporal Context	Pantomimic Static World-Indep. Discrete n	Pantomimic Static World-Indep. Discrete n	Pantomimic Static World-Indep. Discrete n	Manipulative Static World-Indep. Discrete n
Group 2: Physical Characteristics	Dimension. Complexity Body Part Handedness	Simple Hand Dominant hand	Simple Hand Dominant hand	Simple Hand Dominant hand	Simple Hand Dominant hand
	Hand Shape	ASL - U	ASL - V	ASL - L	3-finger pinch

Table A.4: Gesture Mapping and Description - Part4

|	Category	Editing (2)			
	Task	Copy	Delete	Scale	Smart Dimensions
Group 1: Gesture Mapping	Nature Form Binding	Symbolic Static World-Indep.	Pantomimic Static World-Indep.	Manipulative Static World-Indep.	Abstract Static World-Indep.
	Temporal Context	Discrete no	Discrete no	Discrete no	Discrete no
Group 2: Physical Characteristics	Dimension. Complexity Body Part Handedness	- Simple Hand Dominant hand	- Simple Hand Dominant hand	1-Axis Simple hand+arm Dominant hand / Bimanual	- Simple Hand Dominant hand
	Hand Shape	ASL - C	ASL - Q	ASL - F	ASL - I
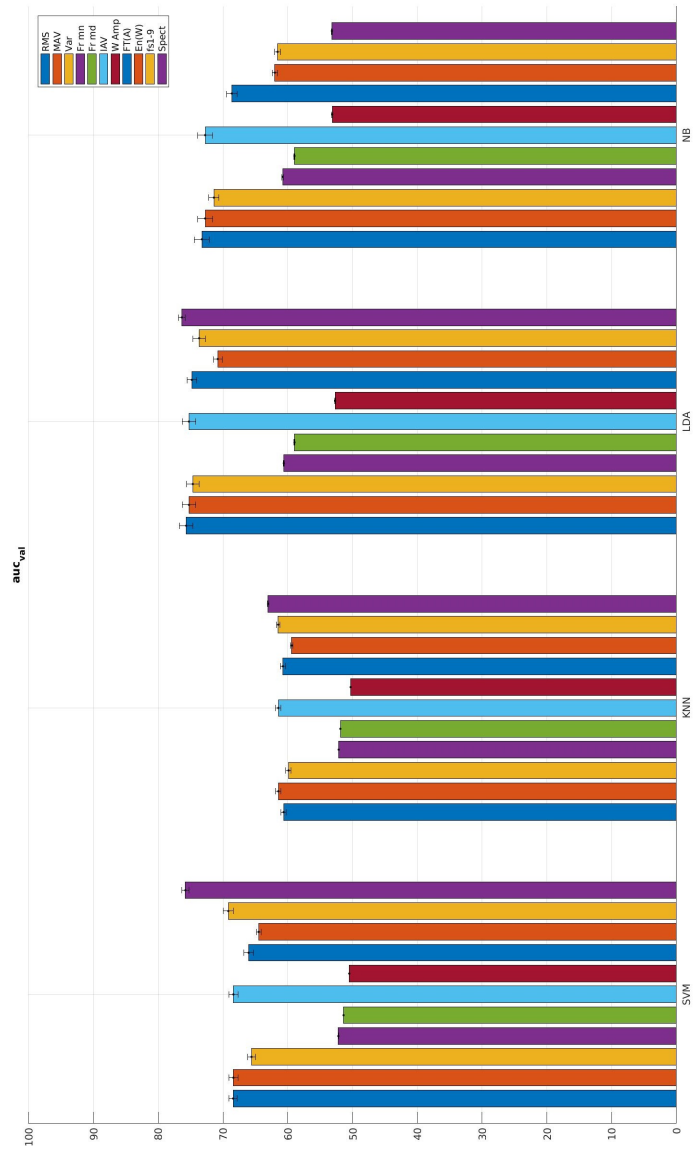
Figure A.1: Validation AUCs of Individual Classifiers
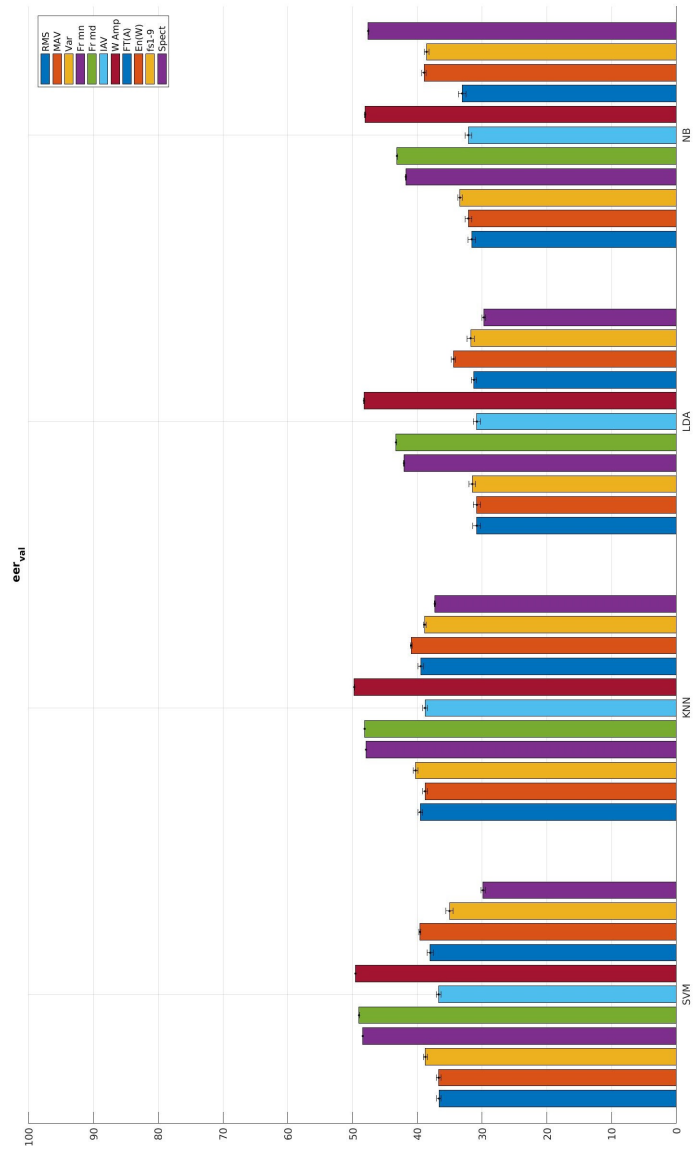
67
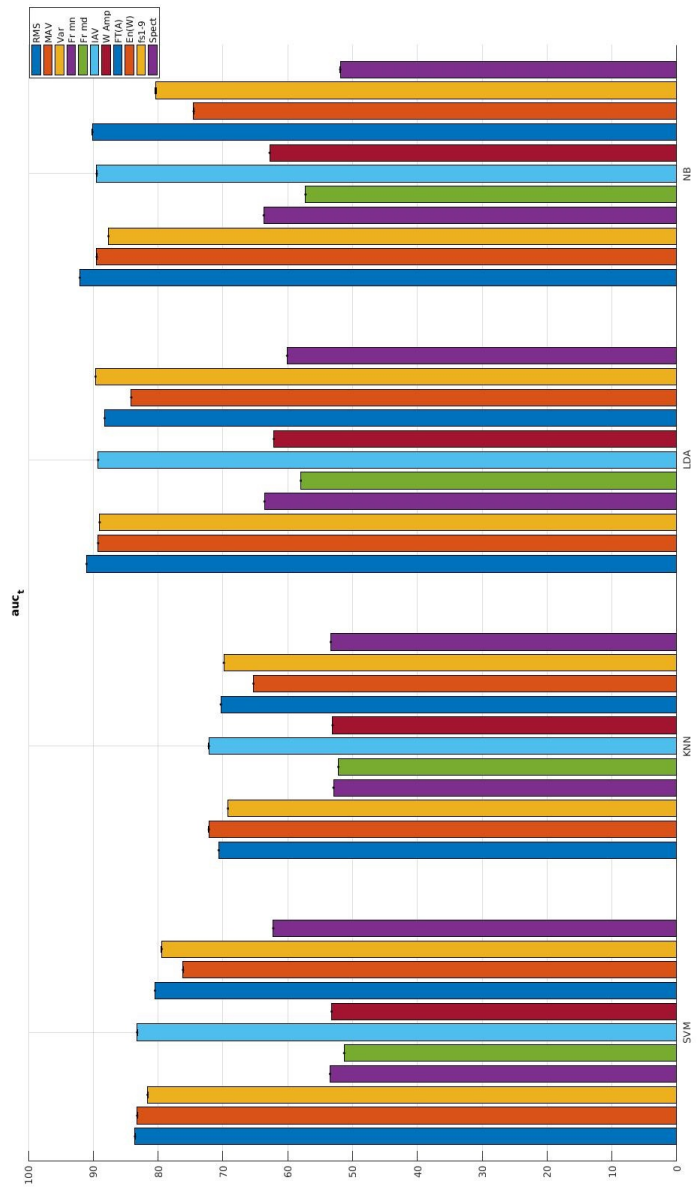
Figure A.2: Validation EERs of Individual Classifiers

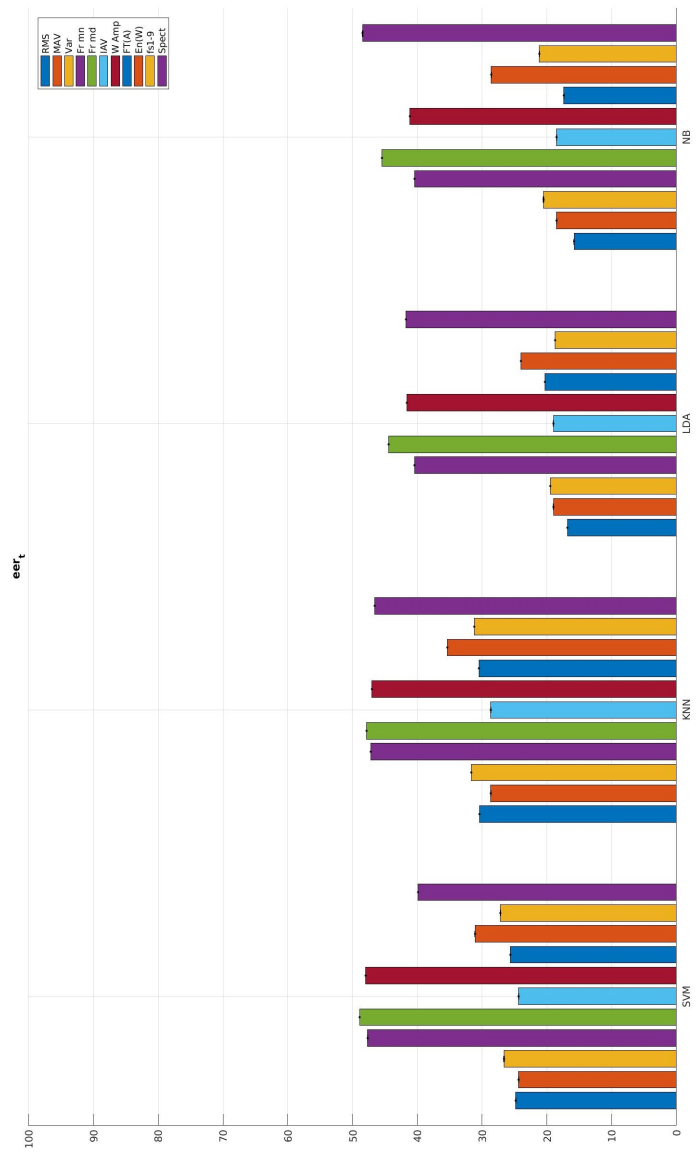Figure A.3: Testing AUCs of Individual Classifiers

Figure A.4: Testing EERs of Individual Classifiers

APPENDIX B

IRB APPROVAL AND CONSENT FORMS

UMKC
5319 Rockhill Road
Kansas City, MO 64110
TEL: (816) 235-5927
FAX: (816) 235-5602

UMKC
Office of Research Services

NOTICE OF NEW APPROVAL

Principal Investigator: Dr. Ahmed Hassan
School of Computing and Engineering
Kansas City, MO 64110

Protocol Number: 16-416
Protocol Title: Collection of Forearm Surface Electromyogram (sEMG) Biosignals for Hand Gesture Recognition
Type of Review: Designated Review
Expedited Category #: 4

Date of Approval: 02/24/2017
Date of Expiration: 02/23/2018

Dear Dr. Hassan,

The above referenced study, and your participation as a principal investigator, was reviewed and approved, under the applicable IRB regulations at 21 CFR 50 and 56 (FDA) or 45 CFR 46 (OHRP), by the UMKC IRB. You are granted permission to conduct your study as described in your application.

Your protocol was approved under Expedited Review Regulatory Criteria at 45 CFR 46.110 or 21 CFT 56.110 under Category #4 as follows:
4. Collection of data through noninvasive procedures (not involving general anesthesia or sedation) routinely employed in clinical practice, excluding procedures involving x-rays or microwaves. Where medical devices are employed, they must be cleared/approved for marketing. (Studies intended to evaluate the safety and effectiveness of the medical device are not generally eligible for expedited review, including studies of cleared medical devices for new indications.)

Examples: (a) physical sensors that are applied either to the surface of the body or at a distance and do not involve input of significant amounts of energy into the subject or an invasion of the subject=s privacy; (b) weighing or testing sensory acuity; (c) magnetic resonance imaging; (d) electrocardiography, electroencephalography, thermography, detection of naturally occurring radioactivity, electroretinography, ultrasound, diagnostic infrared imaging, doppler blood flow, and echocardiography; (e) moderate exercise, muscular strength testing, body composition assessment, and flexibility testing where appropriate given the age, weight, and health of the individual.

This approval includes the following documents:

Attachments
        citiCompletionReportAlaAddin
        Thesis Supervisor Approval Email
        Hassan_Biomedical_Training
        Volunteer Recruiting Script (Informal)
        Volunteer Recruiting Email

If a consent is being used in this research study you may find the stamped version in section 16 of your application.

The ability to conduct this study will expire on or before 02/23/2018 unless a request for continuing review is received and approved. If you intend to continue conduct of this study, it is your responsibility to provide a Continuing Review form prior to the expiration of approval or a final report if you plan to close the study.

This approval is issued under the University of Missouri - Kansas City's Federal Wide Assurance FWA00005427 with the Office for Human Research Protections (OHRP). If you have any questions regarding your obligations under the Board's Assurance, please do not hesitate to contact us.

There are 5 stipulations of approval:

1) No subjects may be involved in any study procedure prior to the IRB approval date or after the expiration date. (PIs and sponsors are responsible for initiating Continuing Review proceedings).
2) All unanticipated or serious adverse events must be reported to the IRB.
3) All protocol modifications must be IRB approved prior to implementation unless they are intended to reduce risk. This includes any change of investigator.
4) All protocol deviations must be reported to the IRB.
5) All recruitment materials and methods must be approved by the IRB prior to being used.

Please contact the Research Compliance Office (email: umkcirb@umkc.edu; phone: (816)235-5927) if you have questions or require further information.

Thank you,

Cynthia Thompson

CONSENT FORM FOR PARTICIPATION IN A RESEARCH STUDY

Forearm Surface Electromyogram (sEMG) Biosignal Dataset for Hand Gesture Recognition

<u>**Introduction**</u>

You are being asked to volunteer for a research study. The data collection will take place at Room 454, Flarsheim Hall, and/or right outside Flarsheim Hall building, on UMKC's Volker Campus. The Investigators in charge of this study Dr. Ahmed Hassan (Assistant Professor, University of Missouri, Kansas City) and Ala-Addin Nabulsi (Graduate Student, University of Missouri, Kansas City).

The study team is asking you to take part in this research study because you are an adult student or employee at UMKC. Research studies only include people who choose to take part. Please read this consent form carefully and take your time making your decision. The study doctor or staff will go over this consent form with you. Ask him/her to explain anything that you do not understand. Think about it and talk it over with your family and friends before you decide if you want to take part in this research study. This consent form explains what to expect: the risks, discomforts, and benefits, if any, if you consent to be in the study.

<u>**Background**</u>
Hand gesture recognition is the study of hand gestures to be used in human-computer interaction. Hand gestures can be observed from multiple perspectives, be it visual (through cameras), biosignals (such as the electric signals in the forearm muscles), or otherwise. For either perspective, patterns can be found in the images or the electric signals which may be used to identify which hand gesture is being made at the time.

<u>**Purpose**</u>
The purpose of this research study is to advance the technology for hand gesture recognition based on unique patterns observed from the forearm muscles also called surface-Electromyography (s-EMG) signals for human-computer-interaction. In order to promote research in this field: (i) we need to collect a adequately-sized sEMG dataset using a sensor-filled-armband and infrared cameras, (ii) the collected dataset will allow us to study and compare the efficacy of various biosignal processing and pattern recognition algorithms designed for hand-gesture recognition in the context of wearable devices, and (iii) further, the collected and anonymized dataset will be publicly available to promote research and facilitate fair comparison and independent replication of the results from different research groups. You will be one of about 30 subjects in the study at Room 454, Flarsheim Hall, and/or right outside Flarsheim Hall building, on UMKC's Volker Campus.

UMKC IRB
Approved
from: 02/27/2017   to: 02/23/2018

Subject Initials _____

Study Visit Procedures and Treatments
The following study visits and procedures will occur:
You will be one of the 30 healthy adult volunteers who will be asked to perform 18 hand-gestures while wearing an armband on your forearm. The sensors in the armband are non-invasive. Hand-gesture data will also be collected with an infrared camera.
During your first visit, the informed consent form will be given and you will be provided with ample time to read, understand, and ask any questions you might have. You will be required to pay 3 visits in the time lapse of 1-2 weeks. At each visit, you will be asked to perform each of the 18 hand gestures 60 times in a row, with short breaks in between each gesture and the next. Each visit may take about 90 minutes. You will always be at the discretion to withdraw at any time.
When you are done taking part in this study, you will *not* have access to the study treatment/procedure after completing the study. This is because this research is not of therapeutic nature.
The forearm electric signals from the collected dataset will be made publicly available, and will not be associated with any identifiable information related to you.

Possible Risks or Side Effects of Taking Part in this Study
To the best of our knowledge, recording data with sEMG sensors, IMUs and cameras (infrared or ordinary) has no adverse side effects. The risks of this study are the possible loss of privacy or breach/loss of confidentiality. We will take measures to minimize this risk by storing your recorded data under folders only labelled with a random study number, and by using password-protected computer systems accessible only by the listed investigators. If you experience any discomfort, please inform the investigators and we will stop the collection process. Only anonymized statistical data will be published.

Possible Benefits for Taking Part in this Study
There are no direct benefits to you for participating in this study. However, your participation may help the advancement of technologies used in human-computer-interaction.

Costs for Taking Part in this Study
There are no anticipated, direct or indirect, costs to you for participating in this study.

Alternatives to Study Participation
The alternative is to not participate.

Confidentiality and Access to your Records
Results of this research may be published for scientific purposes or presented to scientific groups; however, you will not be personally identified. The Institutional Review Board or other regulatory agencies may be given access to research study records that contain your data labelled with a random study number. Recorded information that may possibly identify you and the consent form signed by you will be reviewed to verify the study procedures that were performed. Your records will be kept as confidential as possible under local, state and federal law, but absolute confidentiality cannot be guaranteed. All data will be filed at room 454 Flarsheim Hall, under adequate security and with the

UMKC IRB
Approved
from: 02/27/2017   to: 02/23/2018

Subject Initials _____

accessibility restricted to the research personnel only. An anonymous electronic copy of research data will be provided to the study co-investigator. Further, only raw biosignal data (without any demographic information) will be publicly available for research purposes.

In Case of Injury
The University of Missouri-Kansas City appreciates people who help it gain knowledge by being in research studies. It is not the University's policy to pay for or provide medical treatment for persons who participate in studies. If you think you have been harmed because you were in this study, please call the investigator Dr. Ahmed Hassan at ███████████, or the IRB Administrator of UMKC's Adult Health Sciences Institutional Review Board at 816-235-5927.

Contacts for Questions about the Study
If you have any questions or concerns regarding this study, or if any problems arise, you may call the study Investigator Dr. Ahmed Hassan, at ███████████ or Co-Investigator Ala-Addin Nabulsi, ███████████. To express concerns of pressure about your participation in the study or ask questions you may also contact IRB Administrator of UMKC's Adult Health Sciences Institutional Review Board at 816-235-5927. If you have questions regarding your rights as a research participant please contact the IRB Administrator of UMKC's Adult Health Sciences Institutional Review Board at 816-235-5927.

Voluntary Participation
You volunteer and consent to participate in this research study. A copy of this consent form will be given to you. Your participation in this research is voluntary; you are free to discontinue participation in this study at any time and for any reason; refusal to participate will involve no penalty or loss of care to which you are normally entitled; you may also discontinue participation at any time without penalty or loss of benefits to which you are entitled. You will be removed from the study at any time if it is necessary because of medical reasons. Additionally, you will be informed of any significant findings developed during the course of this research that might adversely affect you. You have read this Consent for Research or it has been read to you. Further, the purpose of the study, risks involved, and procedures which will be performed have been explained to you. You have had the chance to ask questions, and you may ask questions at any time during the course of the study by calling Dr. Ahmed Hassan, at ███████████ or Ala-Addin Nabulsi ███████████.

_____   _____
Signature (Volunteer Subject)   Date   Printed Name (Volunteer Subject)

_____   _____
Signature of Person Obtaining Consent   Date   Printed Name of Person Obtaining Consent

Page 3 of 3 Version Date: 12/14/2016	**UMKC IRB** Approved from: 02/27/2017 to: 02/23/2018	Subject Initials _____

REFERENCE LIST

[1] I. E. Sutherland, "A head-mounted three dimensional display," in *Proceedings of the December 9-11, 1968, fall joint computer conference, part I.* ACM, 1968, pp. 757–764.

[2] F. Kerber, M. Puhl, and A. Krüger, "User-independent real-time hand gesture recognition based on surface electromyography," in *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services.* ACM, 2017, p. 36.

[3] M. C. Garcia and T. Vieira, "Surface electromyography: Why, when and how to use it," *Revista andaluza de medicina del deporte*, vol. 4, no. 1, pp. 17–28, 2011.

[4] J. D. Enderle and J. D. Bronzino, *"Biomedical Sensors" and "Biosignal Processing" Introduction to biomedical engineering*, 2012, pp.610-741.

[5] Z. Zhang, "Microsoft kinect sensor and its effect," *IEEE multimedia*, vol. 19, no. 2, pp. 4–10, 2012.

[6] V. Amatanon, S. Chanhang, P. Naiyanetr, and S. Thongpang, "Sign language-thai alphabet conversion based on electromyogram (emg)," in *Biomedical Engineering International Conference (BMEiCON), 2014 7th.* IEEE, 2014, pp. 1–4.

[7] W. Guo, P. Yao, X. Sheng, H. Liu, and X. Zhu, "A wireless wearable semg and nirs acquisition system for an enhanced human-computer interface," in *Systems, Man*

and Cybernetics (SMC), 2014 IEEE International Conference on. IEEE, 2014, pp. 2192–2197.

[8] M. Mohandes, S. Aliyu, and M. Deriche, "Prototype arabic sign language recognition using multi-sensor data fusion of two leap motion controllers," in *Systems, Signals & Devices (SSD), 2015 12th International Multi-Conference on.* IEEE, 2015, pp. 1–6.

[9] J. Fu, L. Xiong, X. Song, Z. Yan, and Y. Xie, "Identification of finger movements from forearm surface emg using an augmented probabilistic neural network," in *System Integration (SII), 2017 IEEE/SICE International Symposium on.* IEEE, 2017, pp. 547–552.

[10] M. Sarcar, K. M. Rao, and K. L. Narayan, *"Fundametals of CAD" in Computer aided design and manufacturing.* PHI Learning Pvt. Ltd., 2008, pp.3.

[11] U. C. Allard, F. Nougarou, C. L. Fall, P. Giguère, C. Gosselin, F. Laviolette, and B. Gosselin, "A convolutional neural network for robotic arm guidance using semg based frequency-features," in *International Conference on Intelligent Robots and Systems (IROS).* IEEE, 2016, pp. 2464–2470.

[12] S. Gieser, V. Kanal, and F. Makedon, "Evaluation of a low cost emg sensor as a modality for use in virtual reality applications," in *International Conference on Virtual, Augmented and Mixed Reality.* Springer, 2017, pp. 97–110.

[13] G.-C. Luh, Y.-H. Ma, C.-J. Yen, and H.-A. Lin, "Muscle-gesture robot hand control based on semg signals with wavelet transform features and neural network classifier," in *Machine Learning and Cybernetics (ICMLC), 2016 International Conference on*, vol. 2. IEEE, 2016, pp. 627–632.

[14] J. F.-S. Lin, A.-A. Samadani, and D. Kulić, "Segmentation by data point classification applied to forearm surface emg," in *Smart City 360*. Springer, 2016, pp. 153–165.

[15] S. Lobov, V. Mironov, I. Kastalskiy, and V. Kazantsev, "Combined use of command-proportional control of external robotic devices based on electromyography signals." *Medical Technologies in Medicine/Sovremennye Tehnologii v Medicine*, vol. 7, no. 4, 2015.

[16] F. Amirabdollahian and M. Walters, "Application of support vector machines to detect hand and wrist gestures using a myoelectric armband," in *International conference on rehabilitation robotics , London , United Kingdom*, 2017.

[17] M. E. Benalcázar, C. Motoche, J. A. Zea, A. G. Jaramillo, C. E. Anchundia, P. Zambrano, M. Segura, F. B. Palacios, and M. Pérez, "Real-time hand gesture recognition using the myo armband and muscle activity detection," in *Ecuador Technical Chapters Meeting (ETCM), 2017 IEEE*. IEEE, 2017, pp. 1–6.

[18] C.-W. Yeh, T.-Y. Pan, and M.-C. Hu, "A sensor-based official basketball referee signals recognition system using deep belief networks," in *International Conference on Multimedia Modeling*. Springer, 2017, pp. 565–575.

[19] P. Paudyal, J. Lee, A. Banerjee, and S. K. Gupta, "Dyfav: Dynamic feature selection and voting for real-time recognition of fingerspelled alphabet using wearables," in *Proceedings of the 22nd International Conference on Intelligent User Interfaces*. ACM, 2017, pp. 457–467.

[20] P. Paudyal, A. Banerjee, and S. K. Gupta, "Sceptre: a pervasive, non-invasive, and programmable gesture recognition technology," in *Proceedings of the 21st International Conference on Intelligent User Interfaces*. ACM, 2016, pp. 282–293.

[21] E. Kutafina, D. Laukamp, R. Bettermann, U. Schroeder, and S. M. Jonas, "Wearable sensors for elearning of manual tasks: Using forearm emg in hand hygiene training," *Sensors*, vol. 16, no. 8, p. 1221, 2016.

[22] F. Vafaei, "Taxonomy of gestures in human computer interaction," M.S. thesis, North Dakota State University, USA, 2013.

[23] R. Merletti and P. Di Torino, "Standards for reporting emg data," *J Electromyogr Kinesiol*, vol. 9, no. 1, pp. 3–4, 1999.

VITA

Ala-Addin Nabulsi was born in the state of Kansas in 1990. He graduated from the University of Jordan (Amman, Jordan) in august 2014 with an Bachelor's Degree in Mechatronics Engineering. For his senior project, he (with the help of his team) designed and built an electrically powered exoskeletal arm. Their project was nominated for the Jordan Engineering Association Competition.

In 2015 he was accepted into the University of Missouri - Kansas City (UMKC) to start his Master's Degree in Electrical Engineering. There, he pursued his passion of Machine Learning and Biosignal Processing and started to research their application in Human Machine Interaction.

In 2017 he was admitted into UMKC's Interdisciplinary Ph.D. program to continue his research, majoring in Electrical Engineering and Mechanical Engineering as a Co-discipline. He was awarded the Ph.D. Fellowship in Big Data and Analytics in his first year in the Ph.D. program. He spends his days contemplating more intuitive ways to interact with technology and how to make them a reality.