

DEVELOPMENT OF A DEEP LEARNING ALGORITHM USING
ELECTROMYOGRAPHY (EMG) AND ACCELERATION TO MONITOR
UPPER EXTREMITY BEHAVIOR WITH APPLICATION TO INDIVIDUALS
POST-STROKE

A Thesis
presented to
the Faculty of California Polytechnic State University,
San Luis Obispo

In Partial Fulfillment
of the Requirements for the Degree
Master of Science in Mechanical Engineering

by
Nathan Dodd

June 2024

© 2024

Nathan Dodd

ALL RIGHTS RESERVED

COMMITTEE MEMBERSHIP

TITLE: Development of a Deep Learning Algorithm using Electromyography (EMG) and Acceleration to Monitor Upper Extremity Behavior with Application to Individuals Post-Stroke

AUTHOR: Nathan Dodd

DATE SUBMITTED: June 2024

COMMITTEE CHAIR: Eric Espinoza-Wade, Ph.D.
Professor of Mechanical Engineering

COMMITTEE MEMBER: Stephen Klisch, Ph.D.
Professor of Mechanical Engineering

COMMITTEE MEMBER: Jonathan Ventura, Ph.D.
Associate Professor of Computer Science

ABSTRACT

Development of a Deep Learning Algorithm using Electromyography (EMG) and Acceleration to Monitor Upper Extremity Behavior with Application to Individuals Post-Stroke

Nathan Dodd

Stroke is a chronic illness which often impairs survivors for extended periods of time, leaving the individual limited in motor function. The ability to perform daily activities (ADL) is closely linked to motor recovery following a stroke. The objective of this work is to employ surface electromyography (sEMG) gathered through a novel, wearable armband sensor to monitor and quantify ADL performance. The first contribution of this work seeks to develop a relationship between sEMG and grip aperture, a metric tied to the success of post-stroke individuals' functional independence. The second contribution of this work aims to develop a deep learning model to classify RTG movements in the home setting using continuous EMG and acceleration data. In contribution one, ten non-disabled participants (10M, 22.5 0.5 years) were recruited. We performed a correlation analysis between aperture and peak EMG value, as well as a one-way non parametric analysis to determine cylinder diameter effect on aperture. In contribution two, one non-disabled participant is instructed to wash a set of dishes. The EMG and acceleration data collected is input into a recurrent neural network (RNN) machine learning model to classify movement patterns. The first contribution's analysis demonstrated a strong positive correlation between aperture and peak EMG value, as well as a statistically significant effect of diameter ($p < 0.001$). The RNN model built in contribution two demonstrated high capability at classifying movement at 94% accuracy and an F1-score of 86%. These results demonstrate promising feasibility for long-term, in-home classification of daily tasks. Future applications of this approach should consider extending the procedure to in-

clude post-stroke individuals, as this could offer valuable insight into motor recovery within the home setting.

Note: All data and code has been made publicly available (<https://github.com/doddfatherr/Thesis-2024>).

TABLE OF CONTENTS

	Page
LIST OF TABLES	ix
LIST OF FIGURES	x
CHAPTER	
1 Introduction	1
2 Background	4
2.1 Stroke	4
2.2 Stroke Care and Recovery	5
2.3 Current Literature Approach to EMG Sensor Use	6
2.4 Research Objective	8
3 Predicting Grip Aperture using Forearm Muscle Activation Data	9
3.1 Introduction	9
3.2 Background	9
3.3 Methods	10
3.3.1 Experimental Design	10
3.3.1.1 Test Bench Setup	10
3.3.1.2 Hardware	11
3.3.1.3 Software	14
3.3.2 Procedure	15
3.3.3 Data Processing and Analysis	16
3.3.4 Statistical Analysis	16
3.3.4.1 Obtaining Ground-Truth Finger Aperture	16
3.3.4.2 Determining Effects of Diameter and Repetition	17

3.4	Results	18
3.4.1	Individual Participant Data	18
3.4.2	Group Data	18
3.5	Discussion	20
3.5.1	Individual Performance	20
3.5.2	Group Performance	21
3.5.3	Next Steps	21
4	Predicting Reach-to-Grasp in the Home Setting using Deep Learning . . .	23
4.1	Introduction	23
4.2	Background	23
4.3	Methods	24
4.3.1	Experimental Design	24
4.3.2	Procedure	25
4.3.3	Segmentation, Labeling, and Feature Extraction	26
4.3.4	Model Architecture	29
4.4	Results	32
4.4.1	Evaluation Method 1: K-fold Cross Validation	35
4.4.2	Evaluation Method 2: Training using limited Data	37
4.5	Discussion	40
4.5.1	Interpretation of Results	40
4.5.2	Future Work	42
4.5.2.1	Model Improvements	42
4.5.2.2	Data Pre-processing Improvements	44
5	Conclusion	46
6	Appendix	48

6.1 Contribution 1 Procedure	48
6.2 Contribution 2 Procedure	50
BIBLIOGRAPHY	53

LIST OF TABLES

Table	Page
3.1 Experimental setup description	13
3.2 Relationships between ground truth aperture and diameter. Note: Corrected p -value = 0.005	19
3.3 Group statistical performance. Note: Corrected p -value = 0.005 . .	19
4.1 Figure 4.1 item descriptions	24
4.2 Extracted time-domain features	27
4.3 Model Hyperparameters	32
4.4 Averaged k-fold performance	36
4.5 Train/Validation Split Model Performance	37

LIST OF FIGURES

Figure	Page
2.1 MindRove armband (left) used during procedure. The armband consists of 8+2 (bias + reference) radially spaced EMG sensors. Example of the raw EMG signal collect is shown (right).	7
3.1 Experimental setup used to capture grip aperture. Setup includes laptop for data collection, document camera, cylinder to simulate a grasping motion, black markers placed on the thumb and forefinger for aperture measurement, a black scaling block, and the MindRove armband located on the forearm.	11
3.2 Placement of Mindrove armband, with IMU rotated counterclockwise 90° of the palm in a supine positioning.	12
3.3 Drawing of experimental setup used to capture grip aperture. . . .	13
3.4 Determination of grip aperture using computer vision algorithm. . .	15
3.5 Data processing workflow; raw sEMG signal (top) is used to calculate RMS_{EMG} (middle). The <code>envelope</code> function provided by MATLAB is utilized to reduce noise, after which <code>findpeaks</code> is employed to locate the peak values (bottom).	17
3.6 Box and whisker plots for participants 1 - 10, each indicating sEMG amplitude across the 5 apertures.	18
3.7 Summary data for all participants across diameters.	19
4.1 Drawing of experimental setup used to simulate Dish Washing. . . .	25
4.2 MATLAB Video Processing Pipeline. Note class on the right, with 1 indicating an RTG instance, 0 indicating non-RTG.	27
4.3 Windowing technique to extract features from raw data.	28
4.4 Pseudocode to determine window label	29
4.5 LSTM cell architecture, where h_{t-1} and C_{t-1} represent the previous cell output and state, respectively [66].	30

4.6	Model architecture.	31
4.7	Model accuracy over 60 training epochs for each fold	36
4.8	Model loss over 60 training epochs for each fold.	36
4.9	Accuracy and loss with respect to data quantity.	38
4.10	Precision, recall, and F1-score with respect to data quantity.	38
4.11	Accuracy curves with respect to data split.	39
4.12	Model loss curves with respect to data split.	39

Chapter 1

INTRODUCTION

Stroke is a leading cause of long-term disability, with many survivors experiencing cognitive and motor deficits [1]. A significant proportion of individuals who have suffered a stroke experience impairment, or even full paralysis in their affected limbs. For those who have experienced a hemiparetic stroke (i.e., a stroke affecting one of the cerebral hemispheres), the contralateral side of the body is typically more affected. This manifests as weaker, more impaired limbs. When evaluating the trajectory of recovery in this population, increased recovery outcomes are associated with increased use of the affected arm [2]. However, when stroke patients are discharged from inpatient therapy, many use their affected arm less frequently in the home setting, and may even begin to voluntarily suppress the use of the limb. This reduced limb use results in difficulty performing activities of daily living (ADLs) such as cleaning, self-care, and environmental interaction. This difficulty is due to limited distal, tactile capacity. As a result, the ability to perform ballistic reach-to-grasp (RTG) movements represents a measure of functional independence, a common metric by which stroke patient recovery is quantified. The ability to monitor the occurrence of RTG movements in the home may provide a quantitative metric of post-stroke recovery; with better monitoring of individuals post-stroke patients in this setting, we can better relate recovery to intervention, and provide evidence for intervention efficacy.

The goal of this thesis is to further our understanding of the relationship between RTG movement, surface electromyography (EMG), and limb acceleration, with the long-term goal of implementing a real-time monitoring system to aid in stroke patient recovery. Our first research question focuses on relating known behaviors to surface

EMG and acceleration data. We initially establish a relationship between grip aperture (a predictor of RTG success) and forearm peak EMG value. Given that grip aperture and RTG are inextricably linked, our initial investigation focuses on pilot validation in a population of unimpaired individuals. We found that in this population an increase in grip aperture corresponds to a statistically significant increase in root-mean-square (RMS) of EMG ($p < 0.001$), suggesting a direct relationship between the two. We found that with an increase in aperture, an increase in EMG variability between participants exists, highlighting the differences in muscle activation patterns. We present our study methodology, results, and relevance with respect to the literature in Chapter 3

This pilot study lays the foundation for our second research question, where we seek to classify reach-to-grasp from continuous, longitudinal, in-home data collection. We simulate an ecologically valid task which requires the participant to make multiple RTG movements of varying directions and speed. Using the collected EMG and acceleration data gathered from this procedure, we train a novel deep learning model to identify RTG movements from continuous data. We developed a model with 94% accuracy and an F1 score of 86% These results suggest that deep learning models, when paired with lightweight, wearable sensors, can detect composite motion in a real-world, in-home setting. It is suggested that future work expands on this approach by extracting other basic, everyday motions to further aid in stroke recovery monitoring.

Overall, we have developed a comprehensive framework for real-time RTG monitoring with application to monitoring upper-extremity behavior. This framework not only provides insights into muscle activation and motion patterns, but also demonstrates the feasibility of using advanced machine learning techniques to support stroke patient recovery. Future work will focus on refining the system's accuracy, exploring addi-

tional applications in rehabilitation, validation with the post-stroke population, and integrating these insights into clinical practice to improve stroke patient outcomes.

Chapter 2

BACKGROUND

2.1 Stroke

An ischemic stroke can occur when the blood supply to a part of the brain is interrupted. This deprives the brain tissue of oxygen which can lead to permanent brain damage, resulting in impaired cognitive abilities, speech, and movement [3]. Each year, about 795,000 people experience a stroke in the United States [1], making it one of the leading causes morbidity and mortality. Cost of illness studies are employed to quantify both the economic cost of inpatient and outpatient care, as well as the indirect costs due to loss of productivity from patients experiencing rehabilitation, physical impairments, and death. The estimated cost of stroke per year in the US is approximately \$103.5 billion, with the average cost per patient per year of \$59,000 [4]. This high cost encompasses the continuum of stroke care that is offered, from the cost of initial hospitalization, to inpatient therapy where a stroke patient is continuously monitored, then to outpatient therapy, where the patient is moved back in to the home setting and comes in for therapy weekly [5]. It is during this time period of outpatient therapy that stroke patients are not typically monitored in the home setting, where much of a patients' time and rehabilitation occurs following a stroke [6, 7]. In addition, outpatient therapy is also limited in time and medical insurance coverage, as providers typically reduce the number of sessions after 1 month, and cease coverage altogether when individuals stop making progress, which can be as early as 6 months to 1 year [8, 9, 10]. If a technique could be developed to better

monitor individuals post-stroke in the home setting with a long-term goal in mind, an improved standard of care could be introduced to achieve better rates of recovery.

2.2 Stroke Care and Recovery

Stroke survivors often experience hemiparesis, a condition characterized by weakness or partial paralysis of one side of the body [11]. Impairment of stroke patient upper extremity occurs in approximately 80% of cases [12]. Of this population, only 5% regain full arm use, and 20% have no functional arm usage [13]. This can affect the functional independence of their everyday life and their ability to perform basic tasks [2]. A common metric used to measure the functional independence and recovery of stroke patients are activities of daily living (ADL). ADLs are routine tasks that individuals typically perform on a daily basis to take care of themselves and maintain independence, such as cooking a meal, moving from sitting to standing, and dressing oneself [14, 15, 16]. ADL performance is often limited for post-stroke patients, particularly movements which require distal dexterity. An ADL that stroke patients often struggle to perform is the reach-to-grasp movement (RTG), a movement one may use when reaching towards a water bottle [16, 17]. This movement, because of the distal dexterity required, often forces stroke patients to use compensatory movements, such as movement of the trunk, dragging of the fingers over the object before fully grasping, and a using weakened grasp. The current process by which the performance of RTG movements is assessed, as well as other ADLs, is typically qualitative, not quantitative, using a general grading assessment to grade a participant on their performance [18, 19]. However, a quantitative measurement by which RTG success is linked to is grip aperture, the euclidean distance measurement between thumb and forefinger [20]. The moment of maximum grip aperture has been observed to be timed precisely with hand transport, indicating that both grasping and reaching are linked

temporally [21]. Measurement systems for aperture are typically goniometer based, not allowing for ease of in-home measurement. A quantitative measure of ADL performance in the home setting still remains to be widely employed in the stroke care continuum.

Many strategies exist for aiding in individual's recovery in the home setting, such as telerehabilitation, robotic devices, games, virtual reality, and sensors [22, 23, 24]. These modalities have shown to engage stroke patients to use their afflicted limbs at a higher frequency and increase activity levels. However, the monitoring of these activity levels remains a challenge. In a review conducted by Chen et al. [25], the researchers concluded that many challenges remain with using sensors in the home setting, such as implementing an accurate, unobtrusive sensor. Implementing a novel monitoring system could give quantitative data on stroke patient recovery, as well as deeper insight into rehabilitation in the home setting.

2.3 Current Literature Approach to EMG Sensor Use

Electromyography (EMG) is a technique used to evaluate the electrical activity of skeletal muscles during contraction. It can be measured using needle-based EMG (nEMG) or non-invasive, surface-based EMG (sEMG). The sEMG approach has been demonstrated to be an effective tool to monitor distal movements, a behavior that has been historically difficult to track with solely inertial measurement unit (IMU) data [26, 27, 28]. Many EMG sensor modalities, such as nEMG or cyberglove approaches, have proven too invasive to practically monitor individuals for extended periods of time [29, 30]. Novel, easily place-able armbands integrating sEMG electrodes provide a robust solution to monitor arm usage through sEMG readings (hereafter referred to simply as EMG).

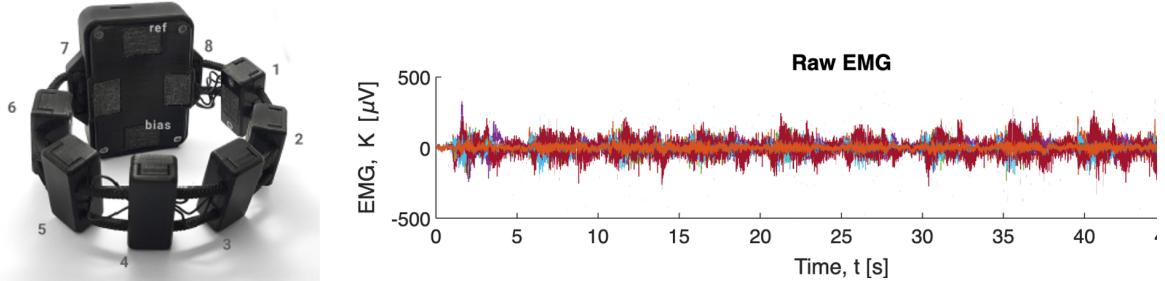


Figure 2.1: MindRove armband (left) used during procedure. The armband consists of 8+2 (bias + reference) radially spaced EMG sensors. Example of the raw EMG signal collect is shown (right).

Current literature aimed at improving stroke care with the use of biosensors focuses on using these armbands to monitor arm usage by post-stroke individuals [31]. One approach to monitor arm usage is by classifying ADL movements from sEMG readings using novel machine learning (ML) techniques [32, 33, 34, 35]. ML models such as linear discriminant analysis (LDA), random forest (RF), hidden markov models (HMM), and artificial neural networks (ANN) have been used to analyze and classify different types of movement from EMG signatures. However, with recent advancements in artificial intelligence (AI) and deep learning (DL), the field has shifted towards more complex models capable of determining hidden, underlying patterns in EMG data. Models such as convolutional neural networks (CNN), recurrent neural networks (RNN), and transformer models have proven to be useful tools in analyzing EMG data [36, 37, 38]. Using these novel DL techniques to classify gestures can allow for the remote monitoring of activity levels of post-stroke individuals.

These techniques often show positive results in a lab setting in detecting and classifying gestures, with many research studies classifying participant movements at rates nearing 100% accuracy. However, studies have shown that when taken out of a research lab environment, the accuracy of machine learning models deteriorates [39]. This is due to the "lab-to-field" phenomena that occurs when trying to replicate out-

side environments in a research lab setting, underscoring the need for further research to bridge this gap [29].

2.4 Research Objective

The contribution of this study is twofold. The first contribution aims to examine the relationship between grip aperture and EMG levels. Grip aperture, a measurement between thumb and forefinger, has been shown to be a valid indicator of RTG success. Doing so can help to quantify RTG movements and provide an insight into ADL performance by stroke patients. The approach is a pilot validation study, conducted on un-impaired individuals, with the long-term goal of applying the procedure to a population of impaired adults to better understand stroke recovery. We expect this study to lay a foundation for future work to further study the relationship and continue to develop real-time monitoring and assessment tools.

The second contribution of this work aims to develop a procedure and algorithm to extract RTG movements from long-term, continuous in-home data. This procedure seeks to simulate an environment which is most-representative of the end goal of developing a non-invasive, real-time monitoring system to monitor stroke patient activity. The EMG and IMU data collected from the in-home procedure will then be used to train a Recurrent Neural Network (RNN) machine learning model that aims to detect these RTG movements from a stream of continuous data. We expect this study to lay groundwork for future studies, with an end goal of fully integrating a deep learning algorithm and EMG sensors to remotely monitor the movements of post-stroke individuals in the home setting.

Chapter 3

PREDICTING GRIP APERTURE USING FOREARM MUSCLE ACTIVATION DATA

3.1 Introduction

Our initial investigation focuses on pilot validation of the approach in a population of un-impaired adults. In this population we predict that increased grip aperture will correspond to increased root-mean-square (RMS) of sEMG data. Additionally, we predict that there exists a linear relationship between aperture and peak RMS of sEMG. We expect this study to provide a foundation for further investigation of the difference in this relationship between healthy and neuromotor-impaired individuals.

3.2 Background

The purpose of our approach is to determine a relationship between forearm sEMG and grip aperture. Doing so can help quantify RTG movements and provide better insight into stroke patient ADL recovery, with the longterm goal of developing real-time monitoring and assessment tools.

A metric commonly used to measure the functional independence and quality of life of stroke survivors are activities of daily living (ADLs) [40]. A common ADL which post-stroke patients often particularly struggle with is the reach-to-grasp (RTG), given the movement necessitates upper-extremity dexterity, a trait often lost after suffering a stroke. A metric closely tied to the success of RTG movements is grip aperture, often quantified as the euclidean distance between thumb and forefinger.

Current approaches to measuring this distance are goniometer-based, limiting their robustness outside of a laboratory setting [41]. Additional information regarding RTG and functional independence, and thus thumb-forefinger aperture, can lead to improved stroke patient care.

3.3 Methods

3.3.1 Experimental Design

Ten non-disabled participants (10M, 22.5 ± 0.5 years) were recruited from a convenience sample at California Polytechnic State University, San Luis Obispo according to the standards of the California State University Institutional Review Board (IRB). Potential participants were approached and provided with detailed information about the study objectives, procedures, and potential risks involved. Informed consent was obtained from each participant prior to their participation in the study. Ethical considerations were carefully addressed throughout the study. Confidentiality of participant information was maintained at all times, and participants were assured that their data would be anonymized and used solely for research purposes.

3.3.1.1 Test Bench Setup

A test bench (depicted in Figure 3.1) was utilized to accurately record both EMG and grip aperture. Novel hardware and software techniques were employed to capture the necessary data. Relevant components found in Figure 3.1 will be discussed in further detail and can be found in Table 3.1.



Figure 3.1: Experimental setup used to capture grip aperture. Setup includes laptop for data collection, document camera, cylinder to simulate a grasping motion, black markers placed on the thumb and forefinger for aperture measurement, a black scaling block, and the MindRove armband located on the forearm.

3.3.1.2 Hardware

The MindRove armband (www.mindrove.com) is a commercially available armband designed to detect muscle activation and movement (Figure 3.1). The armband utilizes 8 surface electromyography (sEMG) nodes and a 3-axis accelerometer ($f_s = 500, 50$ Hz respectively). It has been used in similar studies using EMG to determine upper extremity movement [42, 43, 44]. In this study, the armband was placed approximately 2 inches distal to the olecranon, with the inertial measurement unit (IMU) rotated counterclockwise 90° of the palm in a supine position (Figure 3.2).

The positioning of the armband was chosen to monitor the extensor muscle group, the muscle grouping expecting to have the most activation during grasping movements; this includes the extensor digitorum, extensor pollicis brevis, and pollicis longus.

A limitation which was unaddressed was the phenomena of EMG cross-talk. This is the signal detected over a non-active muscle which is generated from a nearby muscle [45]. In the work conducted by Mogk, J. P. et al., the researcheres analyzed forearm EMG readings during gripping tasks [46]. They noted that the magnitude of common signal between neighboring electrodes decreased as distance between electrodes increased. They additionally note that when looking at flexor and extensor node pairs, there is less than 2% common signal. Because we utilize a commercially available armband with rigid electrode placement, in conjunction with its use in many previous studies, we determined the effect of cross talk to be negligible. However, this effect should be expanded upon in future studies.

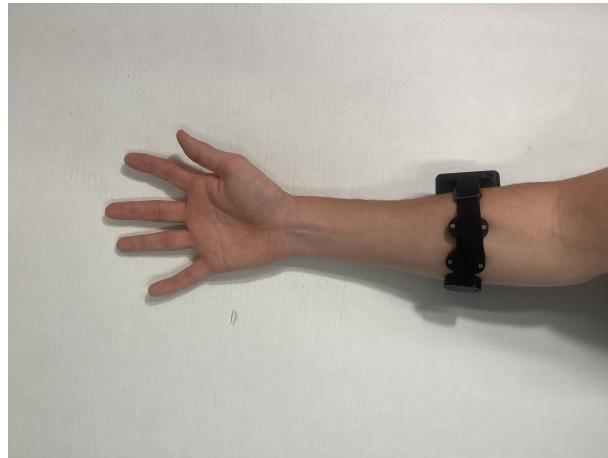


Figure 3.2: Placement of Mindrove armband, with IMU rotated counter-clockwise 90° of the palm in a supine positioning.

The document camera used in this procedure is the IPEVO VZ-R HDMI/USB 8 MP document camera (www.ipevo.com), a commercially available document camera allowing for standardization of camera positioning.

A laptop is connected to both the armband and document camera; before each data collection, the camera is turned on and a recording is started via IPEVO software, then the armband is turned on and data recording is started via mindrove software.

In order to simulate a grasping motion, 3D printed cylinders of varying diameter were placed in front of participants. Participants were instructed to place their hand next to the cylinder and open their hand in a grasping motion, however not actually grasp the item.

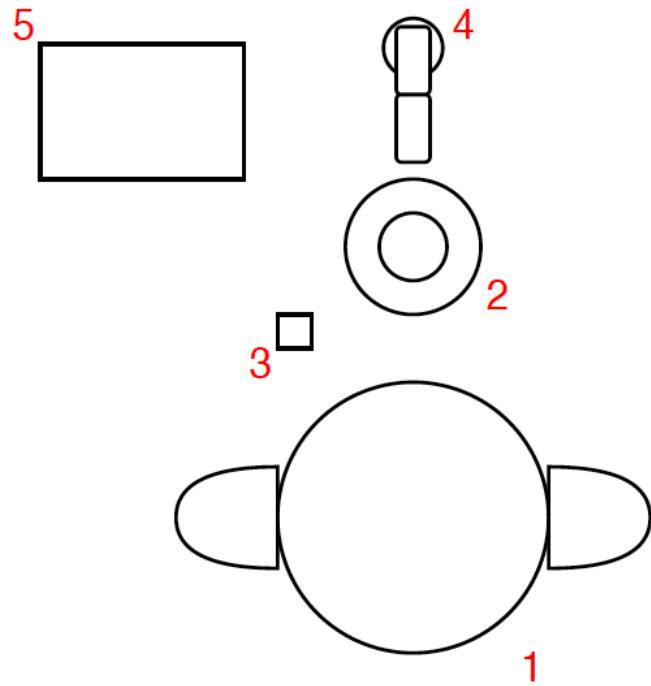


Figure 3.3: Drawing of experimental setup used to capture grip aperture.

Table 3.1: Experimental setup description

Item #	Description
1	Participant
2	Cylinder simulating grasp
3	Ruler
4	Document camera
5	Laptop collecting data

3.3.1.3 Software

Grip aperture was measured using a novel, researcher-developed computer vision system (Figure 3.4). The algorithm was initially validated in fulfillment of a computer vision upper division course, demonstrating promising initial functionality. To allow for ease of measurement and the algorithm to locate finger positioning, remove-able black markers were placed on the insides of the thumb and forefinger. The system utilizes the OpenCV module [47] to calculate grip aperture in 2 distinct steps. The script first locates the ruler in the bottom left of the frame and calculates a camera pixel to distance ratio. The ruler is placed on a 3D printed prism at a height approximately level with the height of the black markers (tests were carried out to determine variability between prism and hand height, determining this variation is insignificant). The algorithm then locates the black markers (shown in Figure 3.4 with green bounding boxes), calculates the distance (in pixels) between markers, and utilizes the previously measured pixel-to-distance conversion factor to calculate and display that distance in the bottom left of the screen.

It is of note that the computer vision algorithm utilizes a binary thresholding technique to locate the markers on a participants' hand. In our procedure, all participants included were of lighter skin tone, allowing the vision application to easily determine the location of the black aperture markers. Due to this, our vision application, like other computer vision systems, can lend itself to potential bias towards participants of fairer skin. Adequate steps would need to be taken to ensure that future iterations of this python script do not induce potential bias.

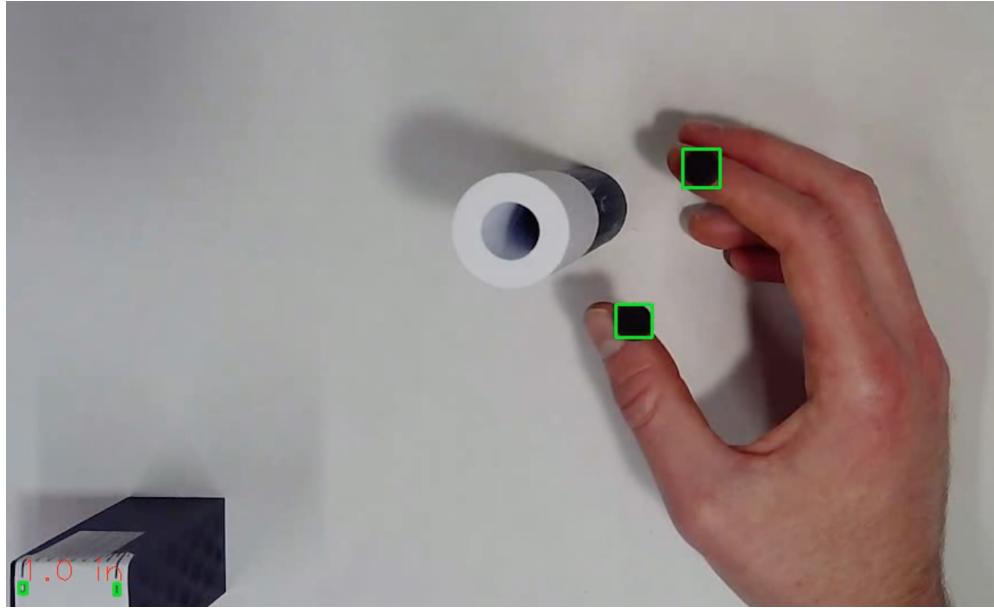


Figure 3.4: Determination of grip aperture using computer vision algorithm.

3.3.2 Procedure

Following informed consent, participants were seated in a relaxed position at a table, where the armband was placed on their forearm (Figure 3.2. Subsequently, a hand dynamometer was utilized to conduct a maximum grip strength test, aiming to standardize participants' surface electromyography (sEMG) data. When the participants were ready to proceed, five cylinders with increasing radii (0.5", 1", 1.5", 2", 2.5") were positioned in front of the participants. These cylinders were denoted as apertures A1, A2, A3, A4, and A5, respectively, for clarity. Acting as visual cues, these cylinders indicated the required grip aperture to the participants.

Participants were instructed to maintain a static hand position while adjusting their hand to an aperture which would allow them to grasp the cylinder. Each grasping motion was repeated 10 times for each cylinder radius to ensure task consistency.

3.3.3 Data Processing and Analysis

Following data collection, raw EMG data from the MindRove armband's eight channels were processed using MATLAB 2023 (www.mathworks.com, Natick, Massachusetts, United States). We computed the root mean square of muscle activation data (RMS_{EMG}) across all eight sEMG channels using the formula:

$$f_{RMS} = \sqrt{\frac{1}{n}(x_1^2 + x_2^2 + \dots + x_n^2)} \quad (3.1)$$

Subsequently, the data underwent additional processing utilizing the MATLAB Signal Processing Toolbox's `envelope` command. This effectively reduced noise in the RMS_{EMG} signal while accurately capturing peak signal values, as illustrated in Figure 3.5. Following envelope processing, the `findpeaks` function was employed to detect peak locations and amplitudes. These identified values were then utilized in subsequent analyses.

3.3.4 Statistical Analysis

3.3.4.1 Obtaining Ground-Truth Finger Aperture

We aimed to establish a relationship between the true (ground-truth) aperture, output from our custom computer vision algorithm, and the maximum electromyography (EMG) for each repetition. However, given the inherent variability in participants' musculoskeletal anatomy, we anticipated that this relationship might differ across individuals. Thus, we conducted correlation analyses within each participant to assess the association between aperture and EMG amplitude. The non-parametric Spearman's rank correlation was employed to evaluate this relationship for each participant.

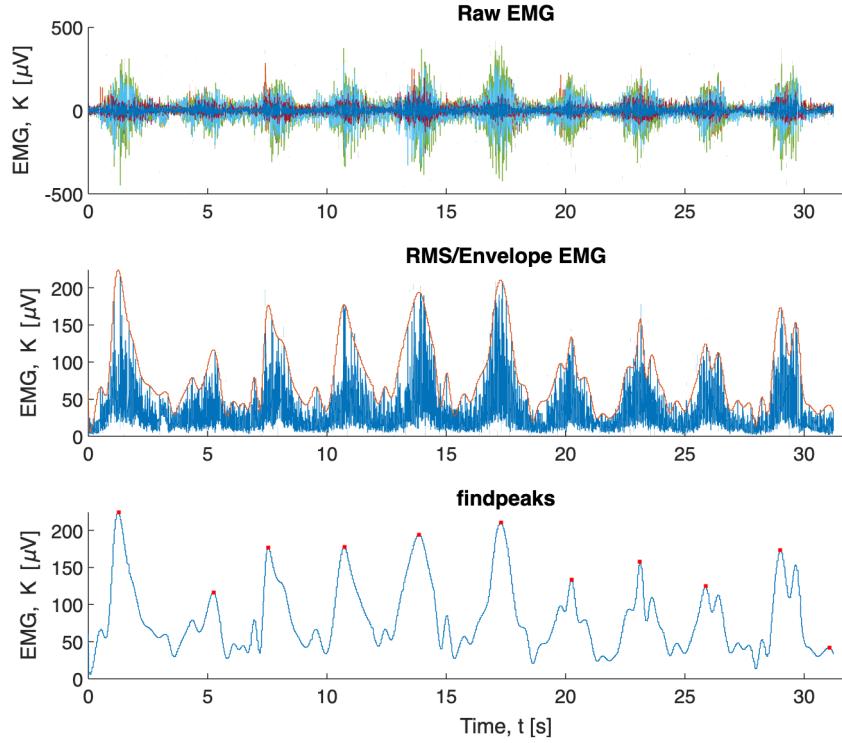


Figure 3.5: Data processing workflow; raw sEMG signal (top) is used to calculate RMS_{EMG} (middle). The envelope function provided by MATLAB is utilized to reduce noise, after which `findpeaks` is employed to locate the peak values (bottom).

3.3.4.2 Determining Effects of Diameter and Repetition

The study was conducted as a one way (cylinder diameter) analysis, with five levels of diameter (D_1, D_2, D_3, D_4 , and D_5). A test for normality of peak EMG readings for each aperture was conducted, and following the result, non-parametric analysis was selected. The Friedman Test was selected to test between cylinder diameters for statistical significance in EMG values with a significance level of $\alpha = 0.05$. In our follow up univariate analysis to determine which aperture groups varied from each other, the Wilcoxon signed-rank test was utilized using a Bonferroni correction. All statistical tests were performed using SPSS (IBM SPSS Statistics for Windows, Version 29.0. Armonk, NY).

3.4 Results

3.4.1 Individual Participant Data

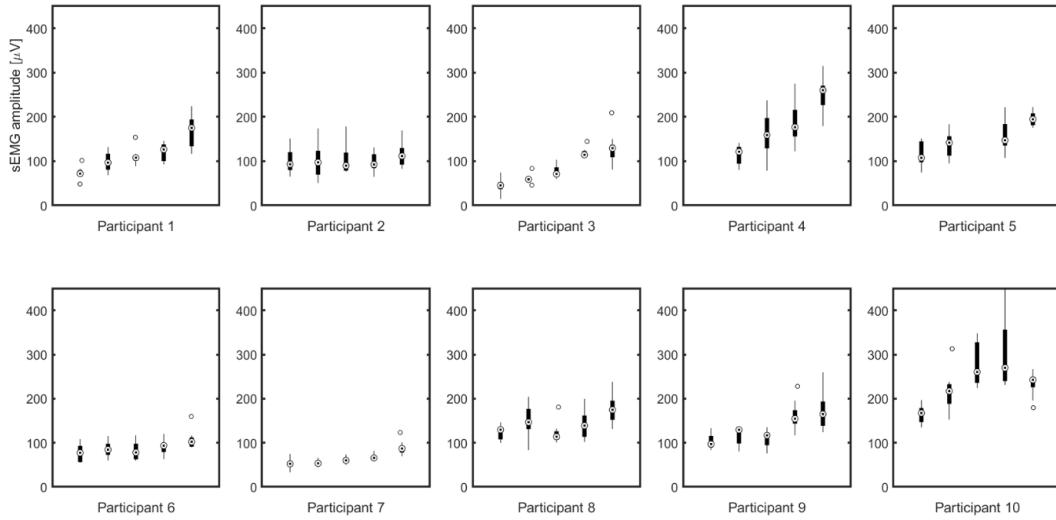


Figure 3.6: Box and whisker plots for participants 1 - 10, each indicating sEMG amplitude across the 5 apertures.

We first visually explored individual participant data (Figure 3.6). From these data, we see a general trend of increasing median sEMG with increasing aperture. To evaluate the relationship between ground truth aperture and sEMG, we evaluated Spearman's rank coefficient within-subjects. Eight of the ten participants demonstrated significant, positive correlation.(Table 3.2).

3.4.2 Group Data

We performed a Friedman test on peak sEMG values to determine the effect of diameter. It was found that diameter had a statistically significant ($p < 0.001$) effect on sEMG readings. Post-hoc analyses using Wilcoxon signed-rank test indicated statistically significant differences between all apertures ($p < 0.001$) except between apertures 2 and 3 (with $p = 0.018$). (Figure 3.7).

Table 3.2: Relationships between ground truth aperture and diameter. Note: Corrected p -value = 0.005

Participant ID	Spearman's ρ	p -value
1	0.758	<0.001
2	0.208	0.152
3	0.883	<0.001
4	0.660	<0.001
5	0.700	<0.001
6	0.408	0.004
7	0.801	<0.001
8	0.304	0.033
9	0.746	<0.001
10	0.604	<0.001

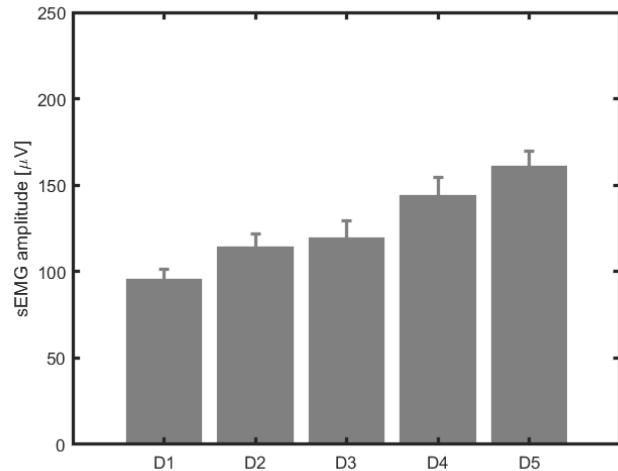


Figure 3.7: Summary data for all participants across diameters.

Table 3.3: Group statistical performance. Note: Corrected p -value = 0.005

Diameter Relationship	Test Statistic, Z	p -value
D1 - D2	-3.563	<0.001
D1 - D3	-3.582	<0.001
D1 - D4	-5.722	<0.001
D1 - D5	-7.052	<0.001
D2 - D3	-2.361	0.018
D2 - D4	-5.111	<0.001
D2 - D5	-6.839	<0.001
D3 - D4	-3.884	<0.001
D3 - D5	-5.914	<0.001
D4 - D5	-4.326	<0.001

3.5 Discussion

3.5.1 Individual Performance

Individual performance analyses indicated a consistent trend suggesting a positive correlation between cylinder diameter and EMG signal amplitude. Within each subject, an increase in cylinder diameter corresponded to heightened EMG signal levels, indicating a direct relationship between grip aperture and muscle activation. Furthermore, within identical diameter groups, the inter-individual variability in EMG signals was lower compared to the variability observed across different diameters. This underscores similar muscle activation patterns across participants when interacting with objects of uniform size, whereas more pronounced differences in muscle activation emerge when comparing across varying cylinder diameters. It is of note that the raw sEMG values for each diameter exhibited variation across participants, which is expected due to the varying musculoskeletal characteristics among individuals. Even for identical movements, participants may engage varying volumes of muscle spindles, contributing to the observed variability in EMG responses.

A further limitation of the study is the inter-participant and intra-session reliability of the EMG data collected. When electrodes are removed and replaced during forearm EMG data collection, the intra-participant correlation deteriorates, underscoring the reliance upon electrode placement [48]. Although the armband was placed at approximately the same location for all participants, variability in this location persists. Due to this, and other confounding factors, the variability of the EMG readings inter-participant were present, however did not significantly affect our statistical results.

3.5.2 Group Performance

The selection of diameters aimed to provide a range of values where sEMG might vary. However, owing to limited prior research on the sensitivity of diameter to sEMG, it's plausible that equal increments may not yield equal changes in sEMG, thus indicating there is not a linear relationship between aperture and sEMG. Our findings supported this notion, indicating differences between all diameters besides D_2 and D_3 . It is suggested that future studies use D_1 , D_2 , D_4 , and D_5 , or potentially a greater final diameter.

Acknowledging potential effects of the protocol on our findings is important. Previous studies on hand posture and forearm sEMG have highlighted various factors influencing study outcomes. For instance, Qing et al. concluded that fatigue and acquisition time can effect the performance of machine learning classifier models trained on sEMG [49]. While our study involved relatively low-stress tasks and short data acquisition periods per participant, the impact of repetition rate on model accuracy warrants consideration. Wang et al. found that as repetition rate increased, model accuracy decreased in certain cases [50]. Although we anticipate minimal influence on our results, recognizing these factors can provide context for interpreting current and future findings.

3.5.3 Next Steps

In future work, there is potential to refine participant instructions to promote more consistent aperture responses. Certain participants reached apertures not consistent with the given cylinder's dimensions. Additionally, the study's inclusion of a small set of male-only participants provides a limitation. Future work should work to include a wider demographic of individuals in the study. Furthermore, while the analyses

conducted offer valuable insights into features sensitive to thumb-forefinger aperture, practical long-term monitoring necessitates the development of models capable of extracting such features from continuous, longitudinal data. Future applications of this approach should consider extending the procedure to include post-stroke individuals, as this could offer valuable insights into motor recovery within the home setting.

Chapter 4

PREDICTING REACH-TO-GRASP IN THE HOME SETTING USING DEEP LEARNING

4.1 Introduction

Our second research question focuses on the classification of reach-to-grasp(RTG) movements from a longitudinal data collection. This study is designed as a proxy for our long-term goal of continuous in-home monitoring, where an individual who has had a stroke can use the armband in the home-setting, and relevant information about their activity levels can be remotely monitored using a novel deep learning algorithm. We predict that our algorithm is capable of extracting RTG movements from longitudinal data, however the armband is likely to incorrectly classify non-RTG movements as such. To our knowledge, this study is the first of its kind to attempt to classify RTG movement from longitudinal, in-home data using EMG and acceleration data.

4.2 Background

Extensive literature exists using deep learning to classify gestures using EMG data [51, 52, 53]; various studies have achieved classification rates above 90%. These studies provide a solid foundation, demonstrating that forearm EMG data alone can be used to gain an understanding of an individual’s movements. However, these studies are generally conducted using static hand poses in a lab environment. This controlled research setting is not representative of the home setting in which this technology is

to be eventually deployed. The end research goal is the ability to give a wearable arm-band to an individual post-stroke as they leave inpatient therapy with the capability of monitoring the movements performed throughout their everyday life. This setting involves many composite motions, involving full body and arm movement, dynamic upper extremity motion, and many external factors which are difficult to predict and model in a research setting. This work aims to simulate the home setting in which a longitudinal, relatively un-controlled data collection occurs for a task selected to gauge the functional independence of the participant. It is our hope that this work will be further built upon to better post-stroke care options.

4.3 Methods

4.3.1 Experimental Design

In order to simulate a real-world environment in which RTG movements would be necessary, an at-home testing procedure was conducted. One non-disabled participant (1M, 23 years) performed the outlined procedure. Our procedure took place in the participant's home, where we set up a dish-washing task using the test bench depicted in Figure 4.1. This activity was selected due to the amount of RTG movements which the activity incurs the participant to make.

Table 4.1: Figure 4.1 item descriptions

Item #	Description
1	Participant
2	Drink glasses before washing
3	Plates before washing
4	Sponge
5	Dish soap
6	Glass & plates after washing

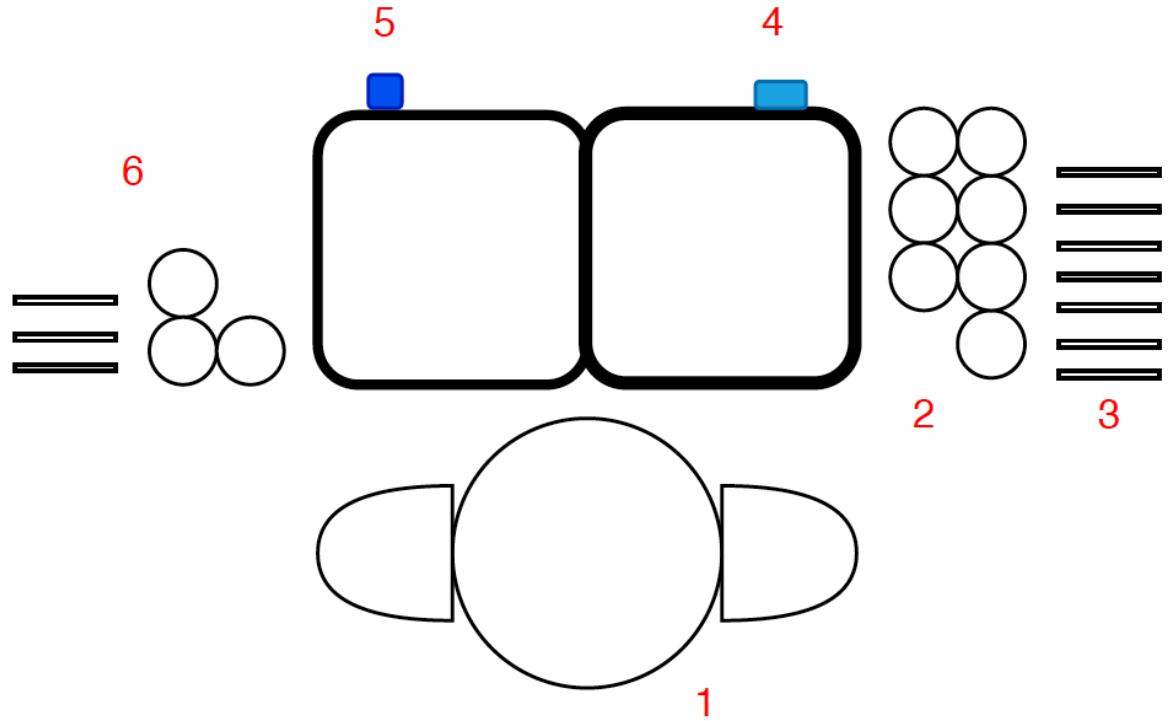


Figure 4.1: Drawing of experimental setup used to simulate Dish Washing.

The commercially available GoPro Hero 9 was used in order to accurately capture ground-truth RTG movements. The camera was fixtured to the participant's chest using a standard GoPro chest mount. The camera was framed so as to capture relevant upper extremity movements. The MindRove armband was once again utilized to collect EMG and inertial measurement unit (IMU) data, with the same forearm positioning as previously indicated in Figure 3.2. These data were streamed to a laptop and appropriately saved following the conclusion of the procedure.

4.3.2 Procedure

The procedure goal is to wash 10 plates and 10 drinking glasses (all of similar sizes) utilizing a sink, dish rack, sponge, and dish soap. The “dirty” dishes are placed on the dominant, armband side of the participant. The participant was instructed to

wash the dishes on the dirty side of the sink using the given sponge and dish soap, and once each dish is clean, to place on the opposite side of the sink. The designed procedure allows for the participant to vary the order in which the task is carried out, without the researcher intervening during the data collection, providing an accurate representation of an in-home setting.

After ensuring both the camera and armband have adequate battery power, all items used in the procedure were in their proper locations, the video recording was started, ensuring that the MindRove user interface is in-screen so as to sync the video and EMG data. Each dish incurred approximately 4 RTG movements: an outward reach to the dirty dish area, an across-body forward reach to the soap, a similar-sided forward reach to the sponge, and a return, proximal dish grasp following use of the sponge. Each data collection was roughly 5 minutes, with an average of 74 RTG movements captured each procedure; a total of 8 procedures were conducted. Following the completion of each procedure, the MindRove program saved the relevant data to a .csv file for further analysis.

4.3.3 Segmentation, Labeling, and Feature Extraction

We used MATLAB 2023 and Apple’s built-in video analysis QuickTime Player for subsequent analysis of raw EMG and acceleration data. The GoPro video recordings were analyzed to extract the timestamps in which an RTG concludes (denoted by the moment in which the hand touches the desired item). Each timestamp was recorded in MATLAB, where a script labels the preceding 375 samples (corresponding to 0.75 seconds) as an RTG movement (class label of 1). It was noted that each RTG task took approximately 0.75 seconds to perform (this value was validated empirically, with no significant change in model accuracy when choosing gesture lengths of 0.5, 1 second). Once this process was concluded, seen in Figure 4.2, the corresponding

.txt file was exported to Python, including the raw EMG, acceleration, and appended class labels.

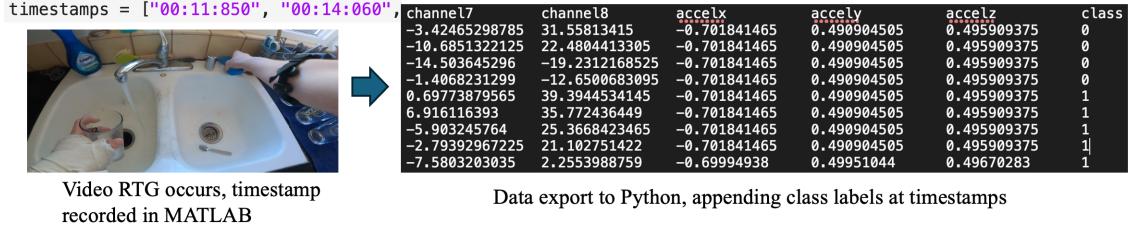


Figure 4.2: MATLAB Video Processing Pipeline. Note class on the right, with 1 indicating an RTG instance, 0 indicating non-RTG.

In order to achieve a robust model, the input should be carefully selected. There are two primary methods for model inputs: collected raw data can be fed as input to a deep learning model in what is referred to as end-to-end learning [65]. Alternatively, manually selected features can be extracted from the raw data to input into the model. Upon conducting a literature review and assessing processing capabilities available, both raw input data and hand-picked features were selected as the input method. The selected time-domain features to be extracted from EMG and acceleration data can be found in Table 4.2. The selected features are widely used in deep learning applications to classify gestures, including EMG and acceleration data [84, 69].

Table 4.2: Extracted time-domain features

Feature	Formula	Description
Mean Absolute Value (<i>MAV</i>)	$MAV = \frac{1}{N} \sum_{i=1}^N x(i) $	Single metric by which overall amplitude can be measured
Root Mean Square (<i>RMS</i>)	$RMS = \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}$	Magnitude or power of a signal
Waveform Length (<i>WL</i>)	$WL = \sum_{i=1}^{N-1} x(i+1) - x(i) $	Represents the total variation or complexity of a signal
Standard Deviation (<i>STD</i>)	$STD = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2}$	Indication of a signal noise and variability
Zero Crossing (<i>ZC</i>)	$ZC = \sum_{i=1}^{N-1} ((x_i \cdot x_{i+1}) < 0)$	Count of times a signal changes its sign, indicating frequency content
Sign Slope Change (<i>SSC</i>)	$SSC = \sum_{i=2}^N ((x_{i-1} - x_{i-2}) \cdot (x_i - x_{i-1}) < 0)$	Count of times a signal changes its slope direction, indicating signal complexity

For both EMG and acceleration raw data, a sliding window approach with window size of 375 samples (750 ms) and an overlap of 20 samples (4 ms) demonstrated the highest model accuracy. Given previous literature and empirical testing, these

parameters were tuned to provide the model with an input size capable of extracting relevant information from the EMG and acceleration data. These 12 features (6 EMG, 6 acceleration) are extracted from each window, normalized, concatenated, and appended to the input feature vector as input to the model(Figure 4.3).

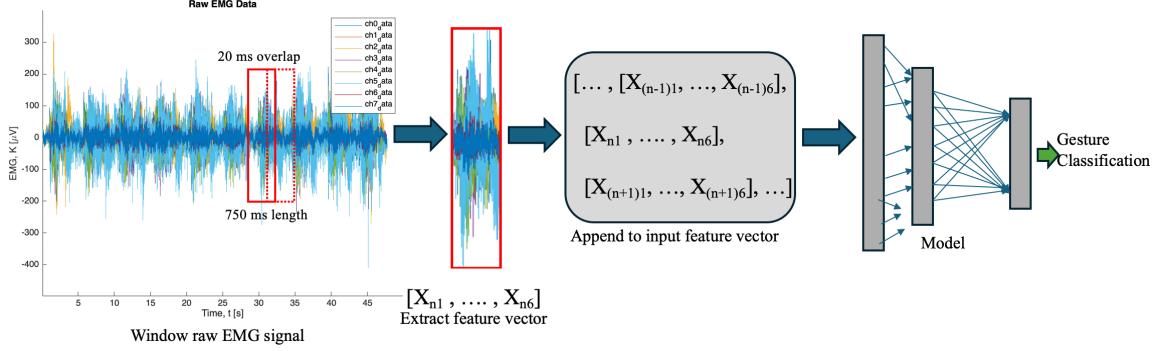


Figure 4.3: Windowing technique to extract features from raw data.

For each single window feature vector, a class label of 0 (non-RTG movement) or 1 (RTG movement) was assigned. Within each 375-sample sliding window, the percentages of each class instance are calculated to determine if an RTG movement is occurring. If the percentage of RTG class instances exceeds a 25% threshold and is the beginning or middle of the RTG movement, the class label for the window is set to 1, indicating an RTG event. If the percentage of RTG instances exceeds a 25% threshold but is not the beginning of the RTG movement, the threshold is increased to 50%.This adjustment helps to avoid false positives during transitions between movements. These values were empirically determined to optimize the model precision. The following pseudocode outlines the logic for determining class labels for each window:

```

function get_window_label(window_samples, threshold_start, threshold_end):
    # Calculate the percentage of RTG class instances in the window
    positive_proportion = count_positive(window_samples) / length(window_samples)

    # Check if the percentage exceeds the 25% threshold
    if positive_proportion >= threshold_start:
        # Determine if it's the beginning, middle, or end of an RTG movement
        if is_beginning(window_samples):
            return 1 # Label for RTG movement (beginning)
        elif is_middle(window_samples):
            return 1 # Label for RTG movement (middle)
        elif is_end(window_samples):
            # If it's the end, increase the threshold to 50%
            if positive_proportion >= threshold_end:
                return 1 # Label for RTG movement (end)
            else:
                return 0 # Label for non-RTG (transition)
        else:
            return 0 # Label for non-RTG movement
    else:
        return 0 # Label for non-RTG movement

```

Figure 4.4: Pseudocode to determine window label

4.3.4 Model Architecture

Given the effectiveness of current deep learning approaches for gesture classification, a recurrent neural network (RNN) architecture was selected. RNNs are a specific type of neural network capable of analyzing sequences of data, where the current prediction is related to the previously predicted values [54]. These have gained popularity via their use in voice recognition assistants, such as Apple’s Siri, Google’s Alexa, as well as predictive text and autocomplete [55]. RNNs have had success in analyzing time-series data, however find difficulty when managing long-term inputs, due to inherent vanishing and exploding gradient issues found among Neural Networks [56]. However, the Long Short Term Memory (LSTM) model, a subset of RNNs, is capable of capturing time dependencies present in long inputs of timeseries data [57, 58]. It does so by utilizing memory cells; each cell has 3 different gates to change the state of the cell: an input, forget, and output gate. The input gate determines which part of the input should be written into the memory cell, typically controlled by a sigmoid activation function. This activation function outputs values between 0 and 1, determining the

proportion of the input to be written to memory. The forget gate determines how much of the current memory cell should be “forgotten”, or discarded, similarly using a sigmoid activation function. The output gate determines the proportion of information to output to the hidden state, once again utilizing sigmoid activation. Utilizing this cell architecture allows for the system to not only look at each timeseries input, but also determine which information of this input is important during training of the model [59].

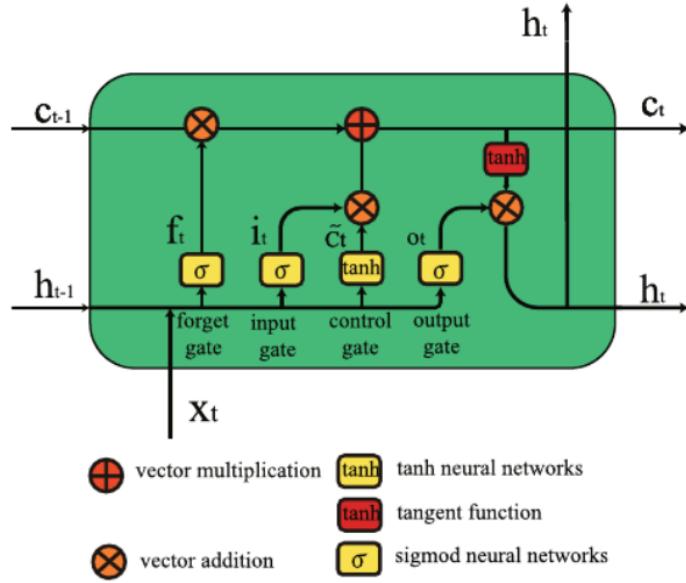


Figure 4.5: LSTM cell architecture, where h_{t-1} and C_{t-1} represent the previous cell output and state, respectively [66].

The Bi-directional LSTM (BiLSTM) model is a specific type of LSTM model that has proven successful in classification tasks [60, 61]. This model architecture analyzes the data in both the forward and backwards directions, allowing the model to gain a deeper understanding of the input sequences. Utilizing the easy interface of the Keras module [62] in python, as well as an existing github repository applying a BiLSTM model for a similar EMG classification problem [63], the model was built in Google’s Colab [64] environment.

For simplicity of inputs, the model is broken into 6 branches: The raw EMG data, raw acceleration data, the zero-crossing feature vector, sign slope feature vector, 1-dimensional extracted EMG features, and 1-dimensional extracted acceleration features. Each branch has its own processing path, tailored to the characteristics of each input. The first 4 branches are first input into their own BiLSTM layer to help capture the temporal dependencies in the data. Then, a dropout layer is utilized to prevent over fitting, outputting to a fully connected (dense) layer. The final two 1-dimensional feature branches are simply flattened and input into a fully connected layer. The outputs of these 6 branches are then concatenated into a dense, 4 layer network of node sizes 100-dropout-12-2. All dense layers use a commonly used Rectified Linear Unit (ReLU) activation functions, except the final layer which has a softmax activation.

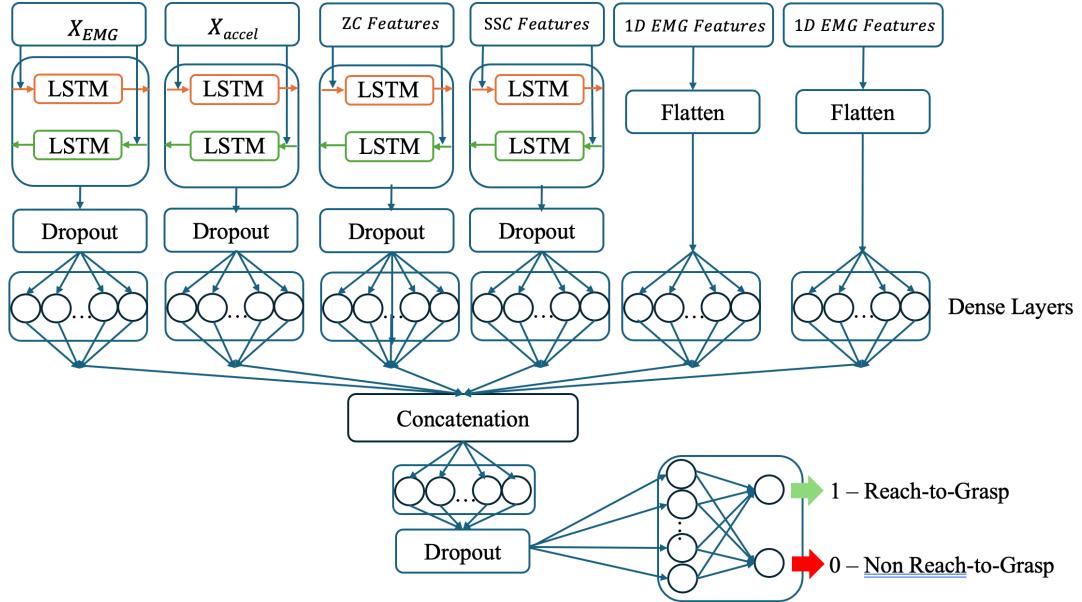


Figure 4.6: Model architecture.

Both empirical testing and a thorough literature review allowed for optimal hyper-parameter selection. K-fold cross validation was utilized to provide a more balanced evaluation of the model and decrease bias in the training dataset. A table of hyper-parameters can be found below.

Table 4.3: Model Hyperparameters

Branch	Layer Parameters	Value	Model Parameter	Value
X _{EMG} X _{acc} ZC, SSC 1-D EMG, Accel Features Concatenated	Bi-LSTM Cells	100	Optimizer	Adam
	Dense Nodes	100	k-fold cross-val	n-folds = 5
	Dropout	0.50	Learning Rate	0.001
	Bi-LSTM Cells	100	Class Weights	Balanced
	Dense Nodes	100	Activation Functions	Relu
	Dropout	0.30		Softmax
Dense Nodes Dense 1 Nodes Dense 2 Nodes Dense 3 Nodes	Dense Nodes	100	Epochs	60
	Dense 1 Nodes	100	Batch Size	256
	Dense 2 Nodes	12		
	Dense 3 Nodes	2		
	Dropout	0.3		

4.4 Results

When evaluating model performance, it is important to keep in mind the metric selection. Model accuracy and loss are the most fundamental metrics in deep learning used to evaluate model performance. Accuracy measures the proportion of correctly classified samples and can be given using the following equation:

$$A = \frac{1}{N} \sum_{i=1}^N f(\hat{y}_i = y_i) \quad (4.1)$$

- N is the total number of samples in the dataset or batch.
- \hat{y}_i is the predicted class label for the i th sample.
- y_i is the true class label for the i th sample.
- $f(\hat{y}_i = y_i)$ is an indicator function that returns 1 if \hat{y}_i and y_i are the same (correct prediction), and 0 otherwise (incorrect prediction).
- A represents the accuracy, calculated as the proportion of correctly predicted samples to the total samples.

Loss quantifies the error made by the model in predicting the target values. It's a measure of how well the model's label predictions match the actual labels, with a lower

loss value indicating a better fit between predictions and true values. The specific loss function varies given the problem-type, however in classification problems, it is common to use categorical cross-entropy as the loss function, given by the following equation:

$$L = - \sum_{i=1}^N \sum_{c=1}^C y_{i,c} \cdot \log(p_{i,c}) \quad (4.2)$$

- N is the number of samples in the dataset or batch.
- C is the number of classes.
- $y_{i,c}$ is the true class label for the i th sample in the dataset for class c .
- $p_{i,c}$ is the predicted probability for the i th sample belonging to class c , typically from the softmax output of the final neural network dense layer.

These two metrics are reliable when analyzing balanced class problems, where the number of instances of each class are all equal. However, in our dataset, there is a large class imbalance: of our 2 class labels, 16% are 1 (RTG movement), 84% are 0 (non-RTG movement). When this is the case, one must rely on more advanced metrics to accurately assess model performance (ex. A model can theoretically predict 0 for every label in our dataset and achieve 84% model accuracy).

Three related metrics to gauge model efficacy in class imbalanced problems are precision, recall, and F1 score. Precision calculates the proportion of true positive predictions among all positive predictions made by the model. This gives a measurement of, given the model has predicted an RTG, the probability that the prediction is correct. It is given by the following formula:

$$P = \frac{TP}{TP + FP} \quad (4.3)$$

- TP stands for true positives, the count of correctly predicted positive samples.
- FP stands for false positives, the count of negative samples incorrectly predicted as positive.
- P represents the precision, calculated as the ratio of true positives to the sum of true and false positives.

Recall, also known as sensitivity or true positive rate, measures the proportion of true positive predictions among all actual positive samples. It is a percentage that given an RTG has occurred, what is the probability it is classified as such. Recall is given by the following equation:

$$R = \frac{TP}{TP + FN} \quad (4.4)$$

- TP stands for true positives, the count of correctly predicted positive samples.
- FN stands for false negatives, the count of positive samples incorrectly predicted as negative.
- R represents recall, calculated as the ratio of true positives to the sum of true positives and false negatives.

The F1 score combines precision and recall into a single value, giving the harmonic mean between the two metrics. It provides a single value to judge model performance that is independent of class balancing. It is given by the following equation:

$$F1 = 2 \cdot \frac{P \cdot R}{P + R} \quad (4.5)$$

- P represents precision, calculated as the ratio of true positives to the sum of true positives and false positives.
- R represents recall, calculated as the ratio of true positives to the sum of true positives and false negatives.
- The F1 score is the harmonic mean of precision and recall, providing a balance between the two metrics.

Metric	Formula	Description
Accuracy	$A = \frac{1}{N} \sum_{i=1}^N f(\hat{y}_i = y_i)$	Proportion of correctly classified samples
Loss	$L = - \sum_{i=1}^N \sum_{c=1}^C y_{i,c} \cdot \log(p_{i,c})$	Quantifies error made by the model
Precision	$P = \frac{TP}{TP+FP}$	Proportion of true positives to all predicted positives
Recall	$R = \frac{TP}{TP+FN}$	Proportion of true positive among all actual positives
F1-Score	$F1 = 2 \cdot \frac{P \cdot R}{P+R}$	Single metric to assess performance by

4.4.1 Evaluation Method 1: K-fold Cross Validation

K-fold cross validation is a common process for assessing model performance and generalizability [81]. It ensures that a model does not rely on a specific subset of the data during the training process. During training, the dataset is split into k equal “folds.” The model is then trained on $k-1$ folds and evaluated on the final fold. This process is repeated k times, changing which fold is the test set each time. After k -iterations of model training and evaluation, the metrics for all k -iterations are then averaged together; doing so provides a more reliable estimate of model performance and generalizability.

Using $k=5$ and the given hyperparameters in Table 4.3, allowing training for 60 epochs (Figures 4.7, 4.8), the model reached an average evaluation accuracy of 93.6%, loss of 18.9%, recall of 91.7%, precision of 79.4%, and F1-score of 85.1%, given in Table 4.4.1.

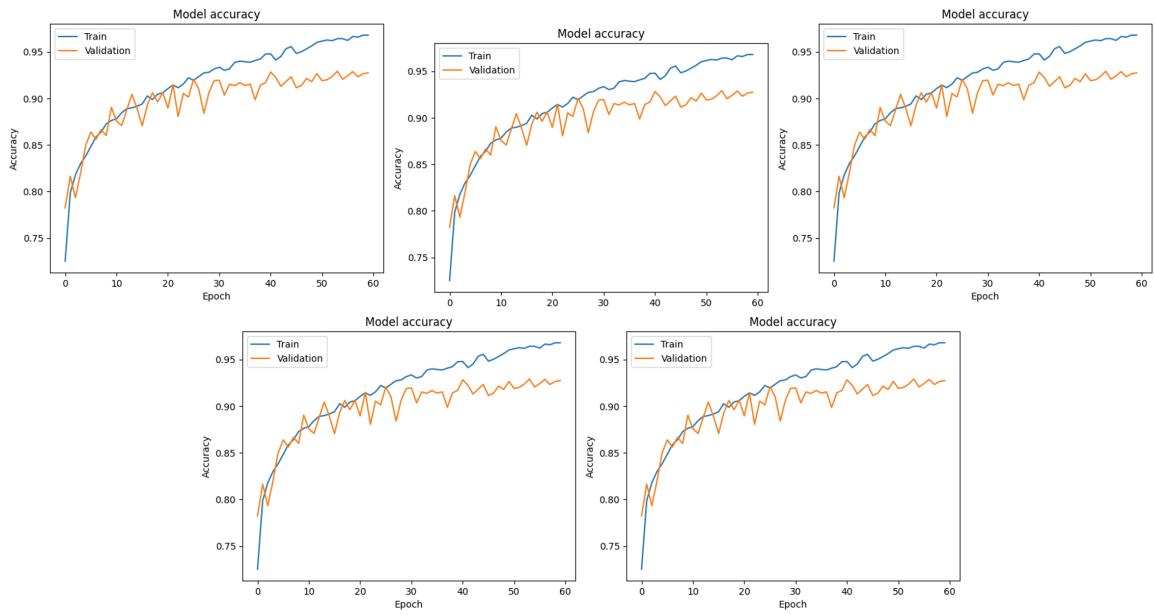


Figure 4.7: Model accuracy over 60 training epochs for each fold

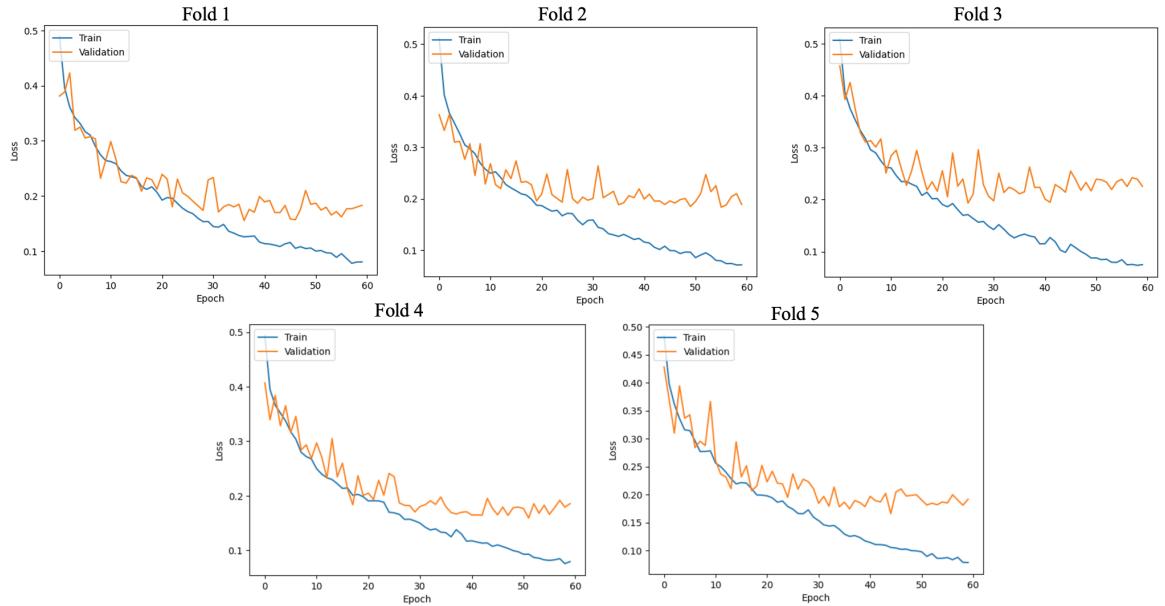


Figure 4.8: Model loss over 60 training epochs for each fold.

Table 4.4: Averaged k-fold performance

Accuracy	Loss	Recall	Precision	F1 Score
0.936	0.189	0.917	0.794	0.851

4.4.2 Evaluation Method 2: Training using limited Data

In a secondary analysis, we decreased the amount of training data iteratively and evaluated the performance of the model. The motivation behind this analysis was to measure the model’s classification performance in low-data settings. In the end use-case of this armband in the home setting, it would be ideal to have a general, pre-trained model available to classify different arm movements of individuals post-stroke. However, based on our findings in contribution 1 and similar studies conducted, model classification accuracy deteriorates when classifying inter-subject. Because of this, eventual application of this armband might entail individual user training and calibration and thus, knowing a model’s capabilities on limited data is of interest.

With this in mind, we begin analysis by first taking 50% of the total data for training and use the remaining 50% for evaluation. After model training and evaluation, the training set is reduced to 30% and the remaining 70% is used as validation. The following trials analyze model performance with 20/80, 10/90, and 5/95 training/-validation data split for a total of 5 trials. The model hyperparameters all remain constant to determine only the effect of training/validation split on model performance. The results of the study can be seen in Table 4.5, Figures 4.9 - 4.12.

Table 4.5: Train/Validation Split Model Performance

Data Split	Accuracy	Loss	Recall	Precision	F1 Score
0.8	0.939	0.191	0.920	0.793	0.852
0.5	0.921	0.247	0.860	0.767	0.811
0.3	0.905	0.339	0.814	0.732	0.771
0.2	0.890	0.397	0.772	0.700	0.734
0.1	0.850	0.511	0.740	0.598	0.661
0.05	0.839	0.664	0.634	0.585	0.608

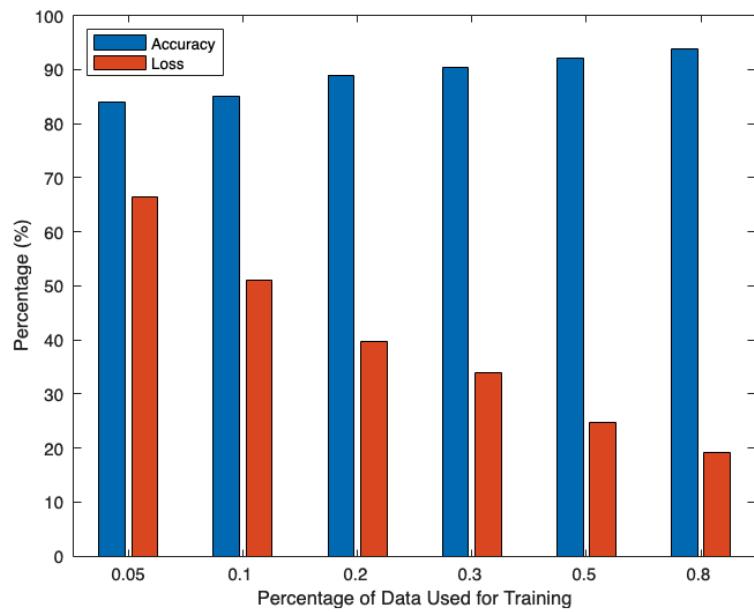


Figure 4.9: Accuracy and loss with respect to data quantity.

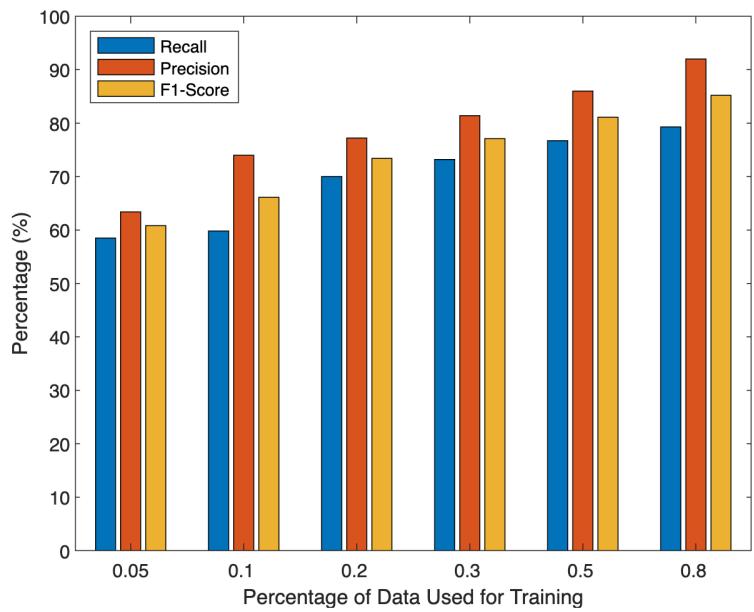


Figure 4.10: Precision, recall, and F1-score with respect to data quantity.

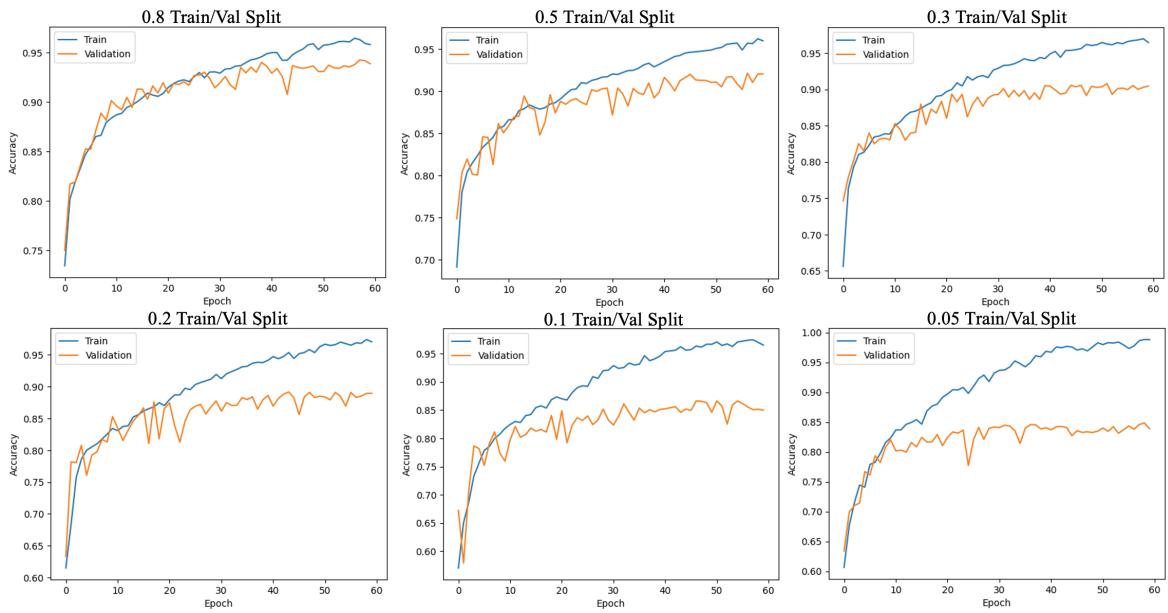


Figure 4.11: Accuracy curves with respect to data split.

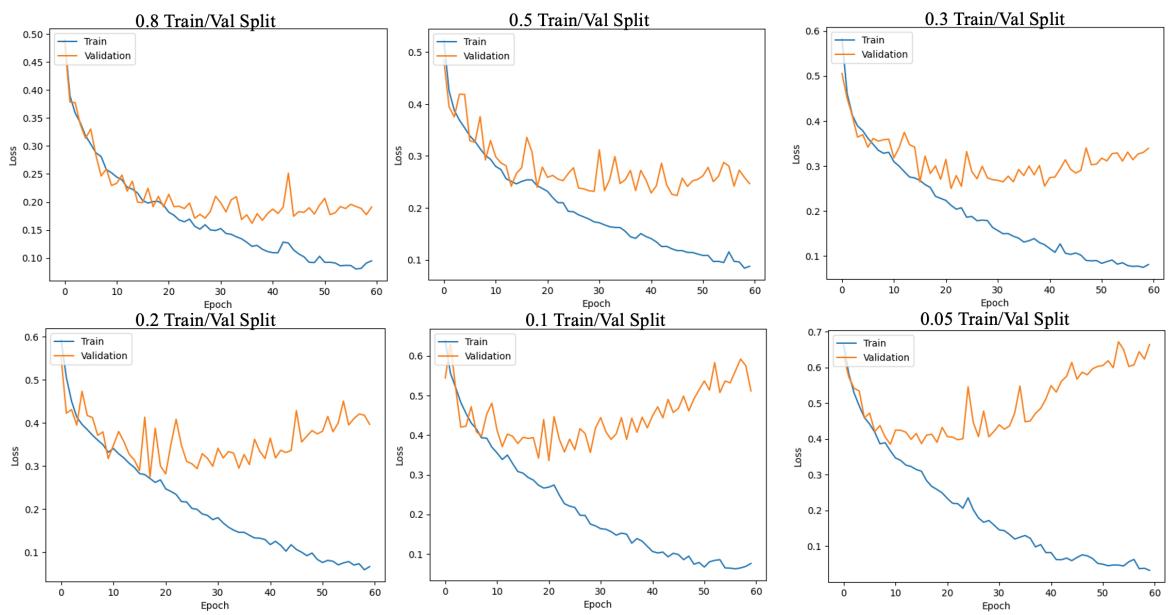


Figure 4.12: Model loss curves with respect to data split.

4.5 Discussion

4.5.1 Interpretation of Results

Although the results are from a novel research methodology, further contextualization is needed to understand their place in the field. When looking at studies using forearm EMG to classify hand movement, much of the current literature focuses on static hand poses and gestures. Work done in this realm often have model accuracy and F1 scores above 0.90. While this is promising, model performance is often lower when classifying dynamic human motion involving both upper extremity control and full arm movement [67, 82], often referred to as human activity recognition (HAR). Pienaar and Malekin’s used an LSTM model to classify dynamic movements such as sitting, walking, and jogging, achieving a model accuracy of 94% and loss below 30%, results comparable to our own [58]. In a study conducted by Xu et. al, although a different RNN is employed using many sensors placed around the body, the classification accuracy for a state of the art model on complex dynamic movements remains at 94.5% [83], underscoring the difficulty in classifying dynamic movements using biosensors.

K-fold cross validation provided the best model results, producing the highest F1 score, lowest loss, and highest accuracy; additionally, the precision and recall were the highest of any trial. This is expected, as the model was trained on approximately 80% of the dataset, whereas models trained on lower proportions of the dataset were incapable of learning the characteristics of the RTG.

When looking at the training cases using limited data (Table 4.5), it can be seen that during the beginning of training, validation loss decreases as accuracy increases. Then, around 30 epochs, the validation loss increases as accuracy remains constant.

This is indicative of overfit in the model, suggesting that more training data is recommended to create a model better capable of generalizing to unseen data. A limitation of this evaluation method, however, is two-fold. The test set between test/train cases was different for each trial: the 80/20 train/test split was only tested on 20% of the data, while the 5/95 split was tested on 95% of the data, a much more difficult task. Because of this, model performance comparison remains difficult. Additionally, the Keras command "modelcheckpoint" was not utilized; due to this, the final model in every trial was used on the test dataset, not the best performing model. For the trials where models experienced overfit (training on 20%, 10%, 5% of dataset), the reported performance metrics were worse than they would have been if the best performing model had been selected, as overfitting likely caused the final model to perform poorly on the test dataset compared to a checkpointed model that captured the best validation performance.

The model struggles to achieve a robust precision, while maintaining a high recall. This indicates that given an RTG has occurred, our model is very capable of classifying it; however, our lower precision states our model as over-predicting these RTG movements. Given the experimental setup, this result is expected, as many movements occur during data collection that are not RTG but have a similar morphology to an RTG movement. The model typically over-predicted an RTG event when the subject performed a reach-to-place (ex: placing of the sponge back in its original location after use) or when two RTG events occurred in quick succession; the model had difficulty in distinguishing these as two separate instances.

4.5.2 Future Work

4.5.2.1 Model Improvements

Although the model is capable of detecting a RTG movement from the given data, there are aspects of the model by which improvements can be made. Although the BiLSTM model has demonstrated its efficacy to correctly classify movements from EMG signals, a similar model architecture with potential benefits is the Convolutional Neural Network LSTM (CNN-LSTM) hybrid network. The CNN has become a pillar of deep learning approaches to EMG analysis, selected for its ability to determine spatial characteristics of the hand, even during composite hand motion [67]. Approaches using a hybrid CNN-LSTM state the benefits of encoding spatial information using the CNN, as well as the temporal characteristics using the LSTM architecture [61, 68]. Implementing this hybrid architecture in future work could prove beneficial in detecting composite motions.

It would also be useful to compare our model performance to that of other existing machine learning models on the same dataset. Although the model results are positive, a performance comparison could prove useful, as the dataset is novel and not a common benchmark dataset that is frequently used. Alternatively, the model can be tested on commonly used benchmark EMG datasets (Ninapro, UCI, BioPatRec) DB2 [72, 73, 74] to allow for comparison with other model architectures. These techniques could give insight into the shortcomings of the procedure, the model architecture, or both.

When working with computationally expensive datasets and models, floating point operations per second (FLOPS) are used to quantify the efficiency of data processing. Given the high signal density of EMG data and the sampling rate of most EMG

sensors, research has been conducted into the efficiency of deep learning models when analyzing EMG data [75]. It has been found that computational complexity in models does not always correlate to computational success, as lightweight models have become increasingly popular [76]. Due to long model long training times, as well as the final, in-home application likely having limited processing power, future work with this model architecture should be conducted into mitigating the computational complexity. Doing so can increase convergence speed and the model hyperparameter tuning process, often a time-intensive labor. Additionally, an ablation study could have been performed in order to determine which features most impacted model performance. Doing so could allow for relevant feature selection and a more lightweight, robust model.

Although there are many hyperparameters, the three that affected model performance the most were number of nodes, number of epochs, and sliding window size. Using previous literature results to select number of layer nodes, when increasing the number of nodes from 30 to 100, model F1-score increased by 13%. When selecting number of epochs to train the model on, a balance between an increase in validation accuracy and loss must be selected. The number of epochs were selected to be when the validation loss began to increase while training loss continued to decrease. This often indicates overfitting in a model, making it at classifying new data. When selecting optimal window size, background research was first conducted. In many studies, optimal model performance is stated to occur when classifying EMG data using a sliding window of 200 ms [61, 65]. However in our analysis, when using the same hyperparameter settings and a window size of 200 ms, model F1-score decreased 10%; highest model performance occurred when a sliding window of 750 ms was used. This can be a function of the data labeling technique, majority class function described previously 4.4, or other unmodeled dynamics present in our data.

4.5.2.2 Data Pre-processing Improvements

In addition to improving the model architecture, improvements can be made upon the data pre-processing methodology. Many studies have shown to have benefited from conducting feature extraction from the frequency domain, including mean frequency, median frequency, and modified mean frequency most commonly [69, 70]. Although a significant proportion of the literature surrounding EMG analysis uses time-domain characteristics, implementation of frequency domain analysis could prove useful in classifying RTG movements.

Another potential drawback to the existing data processing pipeline is the labeling methodology. The problem is simplified to a simple binary classification problem, where the only classes are 1 (RTG is occurring) and 0 (non-RTG movement). This simplification results in a loss of information because the 0 class contains several movements. The approach overlooks the diversity of non-RTG activities, and as well as not accounting for differences within the RTG class. Additionally, when a particular movement contains characteristics between the RTG and non-RTG classes, the model is likely to label the instance as an RTG, increasing the false positive rate. A binary approach may not be the best approach when analyzing dynamic bodily movements. An increase in data labeling, such as adding multiple classes, is likely to improve model performance.

Additionally, the current data processing technique does not adequately address the class imbalance issues. Currently, to address this issue, the built-in Keras tool of *Class Weights* is utilized; this parameter increases the loss for samples from underrepresented classes, while the loss from overrepresented samples is decreased, allowing the model to pay more attention to the minority classes. Although this certainly boosts model performance, further steps can be taken to address the class imbalance,

such as data augmentation techniques (undersampling, oversampling), initializing the bias on the output layer to be the class frequency [71], and implementation of regularization techniques. These can all minimize class imbalance issues and thus improve overall model performance.

Chapter 5

CONCLUSION

Quality of life is drastically reduced for many stroke survivors. We remain unable to effectively monitor the behavior of these individuals in the home setting. Current therapeutic interventions are aimed at regaining motor function, yet do not extend beyond inpatient therapy environments. The contributions of this thesis are aimed at advancing the current standard of stroke care in the home setting via wearable sensors. The ultimate research goal is the distribution of such an armband to a post-stroke individual to wear at home and to remotely monitor activity levels.

Our first research contribution developed a relationship between grip aperture and EMG levels. We developed a novel grip aperture measuring system using a researcher developed computer vision program. We found that forearm sEMG readings were related to the amplitude of thumb-forefinger aperture, a feature critical to performing RTG tasks.

Our second research contribution serves as a framework for future deep learning applications which attempt to classify dynamic human activity in the home setting. The RTG task is one example of a dynamic movement that can be selected to classify stroke patient activity levels. Future work is recommended to expand the number of activity classes, deepen the network architecture, or add a real-time classification component to the system. The system can also be expanded to include several participants from an impaired population, giving a representative sample of future use cases.

Although these contributions are distinct, we anticipate future work to extract grip aperture from in-home RTG motions. In which case, the first contribution provides future research with a knowledge base of EMG signal and grip aperture, while the second contribution can be utilized to train a neural network to capture this aperture from continuous data collection.

Noting both contributions, this work is aimed at bridging the gap between in-home and clinical therapy interventions. Our work focuses on real-world applications using wearable sensors in hopes of improving stroke patient monitoring and thus recovery outcomes. From these contributions, a more robust approach to stroke patient recovery procedure can be formulated and improve the lives of many.

Chapter 6

APPENDIX

6.1 Contribution 1 Procedure

Goal: To develop a relationship between EMG and finger aperture.

Background: Finger aperture (also referred to as grip aperture) is a measurement between forefinger and thumb in the opening and closing of the hand. This movement is integral in reach-to-grasp movements, a movement particularly difficult for post-stroke individuals. If a relationship between finger aperture and EMG can be developed, this can lead to more insight into stroke patient recovery.

Method: Use the Mindrove armband to collect forearm EMG data. Participants will repeat a series of movements, extending the hand in a grasping motion for 10 repetitions. These 10 repetitions will be repeated in 5 sets, with each set simulating a grasp towards an object of varying diameter.

Procedure:

Outline

The procedure will consist of 5 sets of repeated exercises. Each exercise will be 10 repetitions of a grasping motion oriented towards an object of set diameter. Between each set of 10, a larger diameter object will be used to simulate an increase in grip aperture.

Steps

1. Setup:

- When participant is ready to begin, recite background to participant.
- Give armband to participant, ensuring it is on, and properly collecting data (without saving data yet).
- Fasten armband approximately 1" distal from elbow joint. Ensure IMU block is vertical (0° rotation from the palm of participants hand in a supinated position) and fastened securely. Proper positioning can be seen Figure 1.

2. Dynamometer Standardization:

- Hand participant dynamometer and instruct on its operation.
- Begin recording EMG signals and allow for participant to perform maximum grasp.
- Instruct to hold for 3 seconds.
- After participant signals that they are finished, stop recording, and save data file.

3. Grip Aperture:

- Show participant proper hand positioning (Figure 2). Instruct to keep hand resting on paper at all times during trial.
- Have participant sit in a resting position, with dominant hand positioned over paper with printed ruler (Figures 3).

- Allow for participant to perform practice movements, extending hand in a grasping motion towards objects of varying diameter (Figure 4).
- Instruct participant that they will be performing 10 repetitions of each movement. For each movement, pause at maximum extension for approximately 0.5 seconds. Pause in between movements for approximately 1 second.
- When participant states that they are ready to proceed, turn on secondary video recording device (phone/gopro), then turn on document camera.
- Allow participant to proceed.
- Following the conclusion of the 10th repetition, finish the EMG recording, stop the secondary camera recording, then stop document camera recording, and allow time for rest. Repeat grip aperture procedure until 5 sets are completed.

6.2 Contribution 2 Procedure

This procedure provides a simulation of the performance of everyday tasks. Each task is selected with the goal of having the participant perform activities of daily living (ADLs), a metric by which post-stroke individuals' rehabilitation and functional independence is graded upon. Specifically, the ADL selected to be captured is a reach-to-grasp (RTG) movement, an action particularly difficult for post-stroke patients to perform due to the fine finger dexterity required.

The armband in use is the MindRove (www.mindrove.com) armband, a novel armband capable of measuring both muscle activation, via electromyography (EMG) sensors, and full arm movement, via a 3 axis accelerometer. The data recorded will eventually be used as input to a machine learning model to attempt to detect when an individual performs the previously mentioned RTG movements.

The goal of this study is to improve the rehabilitation of stroke patients by monitoring activity levels in the home setting. The armband could be given to a patient as they leave inpatient therapy, and given brief instruction and a pretrained machine learning model, the daily actions of the individual can be remotely sensed and classified. This can improve knowledge of stroke patient recovery and better address the stroke care continuum.

Procedure: Dish Washing

This procedure allows for the performance of many RTG movements during the common household task of washing dishes. The precise order in which dishes are washed is left up to the participant, however the constraints of the procedure are outlined as follows:

Materials

1. Mindrove Armband
2. GoPro (with chest mount)
3. Dish soap
4. Sponge
5. 10 plates of varying size
6. 10 pint sized glasses
7. Drying rack
8. Towel

Setup

1. Fasten Mindrove armband to recommended positioning, with IMU rotated approximately 90° clockwise from palm in a supinated position.
2. Connect armband to Mindrove application using wireless network.
3. Turn on GoPro camera, ensuring it has SD card inserted with adequate charging.
4. Fasten GoPro camera to chest mount; put on the chest mount and tighten around midsection. Wave arms in front of camera to ensure proper framing so as any grasping performed can be captured on camera.
5. Place 10 plates, 10 glasses to the left of sink in the designated locations. Ensure the dish soap, sponge are located appropriately on top of coasters.
6. Ensure the drying rack is empty of any existing dishes.

Experiment

1. When ready to proceed, begin the GoPro recording, then the armband recording, with proper experiment name and number.
2. Walk over to sink and begin grabbing dishes to clean. Wash each dish for approximately 5-10 seconds. The exact order in which dishes are grabbed does not matter, nor does how long/clean the dishes become.
3. When all dishes and glasses are washed and placed onto drying rack, turn off armband recording first, then GoPro recording, saving the file with the relevant name in proper file location.

BIBLIOGRAPHY

- [1] N. M. Salbach, N. E. Mayo, S. Robichaud-Ekstrand, J. A. Hanley, C. L. Richards, and S. Wood-Dauphinee, “Balance self-efficacy and its relevance to physical function and perceived health status after stroke,” *Arch Phys Med Rehabil*, vol. 87, pp. 364–370, Mar 2006.
- [2] D. Chumney, K. Nollinger, K. Shesko, K. Skop, M. Spencer, and R. A. Newton, “Ability of Functional Independence Measure to accurately predict functional outcome of stroke-specific population: systematic review,” *J Rehabil Res Dev*, vol. 47, no. 1, pp. 17–29, 2010.
- [3] V. L. Roger, A. S. Go, D. M. Lloyd-Jones, R. J. Adams, J. D. Berry, T. M. Brown, M. R. Carnethon, S. Dai, G. de Simone, E. S. Ford, C. S. Fox, H. J. Fullerton, C. Gillespie, K. J. Greenlund, S. M. Hailpern, J. A. Heit, P. M. Ho, V. J. Howard, B. M. Kissela, S. J. Kittner, D. T. Lackland, J. H. Lichtman, L. D. Lisabeth, D. M. Makuc, G. M. Marcus, A. Marelli, D. B. Matchar, M. M. McDermott, J. B. Meigs, C. S. Moy, D. Mozaffarian, M. E. Mussolini, G. Nichol, N. P. Paynter, W. D. Rosamond, P. D. Sorlie, R. S. Stafford, T. N. Turan, M. B. Turner, N. D. Wong, and J. Wylie-Rosett, “Heart disease and stroke statistics—2011 update: a report from the American Heart Association,” *Circulation*, vol. 123, pp. e18–e209, Feb 2011.
- [4] T. N. Rochmah, I. T. Rahmawati, M. Dahlui, W. Budiarto, and N. Bilqis, “Economic Burden of Stroke Disease: A Systematic Review,” *Int J Environ Res Public Health*, vol. 18, Jul 2021.
- [5] C. J. Weinstein, J. Stein, R. Arena, B. Bates, L. R. Cherney, S. C. Cramer, F. Deruyter, J. J. Eng, B. Fisher, R. L. Harvey, C. E. Lang,

- M. MacKay-Lyons, K. J. Ottenbacher, S. Pugh, M. J. Reeves, L. G. Richards, W. Stiers, and R. D. Zorowitz, “Guidelines for Adult Stroke Rehabilitation and Recovery: A Guideline for Healthcare Professionals From the American Heart Association/American Stroke Association,” *Stroke*, vol. 47, pp. e98–e169, Jun 2016.
- [6] R. P. Van Peppen, G. Kwakkel, S. Wood-Dauphinee, H. J. Hendriks, P. J. Van der Wees, and J. Dekker, “The impact of physical therapy on functional outcomes after stroke: what’s the evidence?,” *Clin Rehabil*, vol. 18, pp. 833–862, Dec 2004.
- [7] R. A. Werner and S. Kessler, “Effectiveness of an intensive outpatient rehabilitation program for postacute stroke patients,” *Am J Phys Med Rehabil*, vol. 75, no. 2, pp. 114–120, 1996.
- [8] S. R. Chang, N. Hofland, Z. Chen, H. Kovelman, G. F. Wittenberg, and J. Naft, “Improved Disabilities of the Arm, Shoulder and Hand scores after myoelectric arm orthosis use at home in chronic stroke: A retrospective study,” *Prosthet Orthot Int*, Mar 2024.
- [9] Cigna, “Cigna medical coverage policies - physical therapy services.” https://static.cigna.com/assets/chcp/pdf/coveragePolicies/medical/cpg135_physical_therapy.pdf.
- [10] J. K. Freburger, D. Li, A. M. Johnson, and E. P. Fraher, “Physical and Occupational Therapy From the Acute to Community Setting After Stroke: Predictors of Use, Continuity of Care, and Timeliness of Care,” *Arch Phys Med Rehabil*, vol. 99, pp. 1077–1089, Jun 2018.

- [11] S. Wist, J. Clivaz, and M. Sattelmayer, “Muscle strengthening for hemiparesis after stroke: A meta-analysis,” *Ann Phys Rehabil Med*, vol. 59, pp. 114–124, Apr 2016.
- [12] “Physical effects of stroke - hemiparesis.”
<https://www.stroke.org/en/about-stroke/effects-of-stroke/physical-effects-of-stroke/physical-impact/hemiparesis>.
- [13] J. Whitall, S. McCombe Waller, K. H. Silver, and R. F. Macko, “Repetitive bilateral arm training with rhythmic auditory cueing improves motor function in chronic hemiparetic stroke,” *Stroke*, vol. 31, pp. 2390–2395, Oct 2000.
- [14] N. B. Lincoln and J. R. Gladman, “The Extended Activities of Daily Living scale: a further validation,” *Disabil Rehabil*, vol. 14, no. 1, pp. 41–43, 1992.
- [15] J. M. Veerbeek, G. Kwakkel, E. E. van Wegen, J. C. Ket, and M. W. Heymans, “Early prediction of outcome of activities of daily living after stroke: a systematic review,” *Stroke*, vol. 42, pp. 1482–1488, May 2011.
- [16] D. K. Chong, “Measurement of instrumental activities of daily living in stroke,” *Stroke*, vol. 26, pp. 1119–1122, Jun 1995.
- [17] A. J. Turton, P. Cunningham, F. van Wijck, H. Smartt, C. A. Rogers, C. M. Sackley, S. Jowett, S. L. Wolf, K. Wheatley, and P. van Vliet, “Home-based Reach-to-Grasp training for people after stroke is feasible: a pilot randomised controlled trial,” *Clin Rehabil*, vol. 31, pp. 891–903, Jul 2017.
- [18] M. Pashmdarfard and A. Azad, “Assessment tools to evaluate Activities of Daily Living (ADL) and Instrumental Activities of Daily Living (IADL) in older adults: A systematic review,” *Med J Islam Repub Iran*, vol. 34, p. 33, 2020.

- [19] M. E. Mlinac and M. C. Feng, “Assessment of Activities of Daily Living, Self-Care, and Independence,” *Arch Clin Neuropsychol*, vol. 31, pp. 506–516, Sep 2016.
- [20] S. M. Michaelsen, E. C. Magdalon, and M. F. Levin, “Grip aperture scaling to object size in chronic stroke,” *Motor Control*, vol. 13, pp. 197–217, Apr 2009.
- [21] M. K. Rand, A. L. Smiley-Oyen, Y. P. Shimansky, J. R. Bloedel, and G. E. Stelmach, “Control of aperture closure during reach-to-grasp movements in Parkinson’s disease,” *Exp Brain Res*, vol. 168, pp. 131–142, Jan 2006.
- [22] R. Fujii, T. Miki, Y. Nonaka, and S. Tanaka, “Effectiveness of telerehabilitation based on real-time intervention between therapist and participants for improving physical function, activities of daily living and quality of life in people with stroke: A systematic review protocol,” *PLoS One*, vol. 19, no. 4, p. e0297649, 2024.
- [23] G. Devittori, D. Dinacci, D. Romiti, A. Califfi, C. Petrillo, P. Rossi, R. Ranzani, R. Gassert, and O. Lamberty, “Unsupervised robot-assisted rehabilitation after stroke: feasibility, effect on therapy dose, and user experience,” *J Neuroeng Rehabil*, vol. 21, p. 52, Apr 2024.
- [24] D. Rajashekhar, A. Boyer, K. A. Larkin-Kaiser, and S. P. Dukelow, “Technological Advances in Stroke Rehabilitation: Robotics and Virtual Reality,” *Phys Med Rehabil Clin N Am*, vol. 35, pp. 383–398, May 2024.
- [25] Y. Chen, K. T. Abel, J. T. Janecek, Y. Chen, K. Zheng, and S. C. Cramer, “Home-based technologies for stroke rehabilitation: A systematic review,” *Int J Med Inform*, vol. 123, pp. 11–22, Mar 2019.

- [26] D. Ovadia, A. Segal, and N. Rabin, “Classification of hand and wrist movements via surface electromyogram using the random convolutional kernels transform,” *Sci Rep*, vol. 14, p. 4134, Feb 2024.
- [27] M. Pourmokhtari and B. Beigzadeh, “Simple recognition of hand gestures using single-channel EMG signals,” *Proc Inst Mech Eng H*, vol. 238, pp. 372–380, Mar 2024.
- [28] C. Hong, S. Park, and K. Kim, “sEMG-Based Gesture Recognition Using Temporal History,” *IEEE Trans Biomed Eng*, vol. 70, pp. 2655–2666, Sep 2023.
- [29] A. Roda-Sales, N. J. Jarque-Bou, V. Bayarri-Porcar, V. ez, J. L. Sancho-Bru, and M. Vergara, “Electromyography and kinematics data of the hand in activities of daily living with special interest for ergonomics,” *Sci Data*, vol. 10, p. 814, Nov 2023.
- [30] M. S. Totty and E. Wade, “Muscle Activation and Inertial Motion Data for Non-Invasive Classification of Activities of Daily Living,” *IEEE Trans Biomed Eng*, Aug 2017.
- [31] S. n, M. D. a Fogeda, and R. Cano-de-la Cuerda, “An sEMG-Controlled Forearm Bracelet for Assessing and Training Manual Dexterity in Rehabilitation: A Systematic Review,” *J Clin Med*, vol. 11, May 2022.
- [32] A. nez, M. E. zar, and E. Mena-Maldonado, “Real-Time Hand Gesture Recognition Using Surface Electromyography and Machine Learning: A Systematic Literature Review,” *Sensors (Basel)*, vol. 20, Apr 2020.
- [33] N. J. Seo, K. Coupland, C. Finetto, and G. Scronce, “Wearable Sensor to Monitor Quality of Upper Limb Task Practice for Stroke Survivors at Home,” *Sensors (Basel)*, vol. 24, Jan 2024.

- [34] M. Li, G. Scronce, C. Finetto, K. Coupland, M. Zhong, M. E. Lambert, A. Baker, F. Luo, and N. J. Seo, “Application of Deep Learning Algorithm to Monitor Upper Extremity Task Practice,” *Sensors (Basel)*, vol. 23, Jul 2023.
- [35] I. Boukhenoufa, X. Zhai, V. Utti, J. Jackson, and K. D. McDonald-Maier, “A comprehensive evaluation of state-of-the-art time-series deep learning models for activity-recognition in post-stroke rehabilitation assessment,” *Annu Int Conf IEEE Eng Med Biol Soc*, vol. 2021, pp. 2242–2247, Nov 2021.
- [36] M. Emimal, W. J. Hans, T. M. Inbamalar, and N. M. Lindsay, “Classification of EMG signals with CNN features and voting ensemble classifier,” *Comput Methods Biomech Biomed Engin*, pp. 1–15, Feb 2024.
- [37] M. Aviles, J. M. Alvarez-Alvarado, J. B. Robles-Ocampo, P. Y. Sevilla-Camacho, and J. ndiz, “Optimizing RNNs for EMG Signal Classification: A Novel Strategy Using Grey Wolf Optimization,” *Bioengineering (Basel)*, vol. 11, Jan 2024.
- [38] T. Xu, K. Zhao, Y. Hu, L. Li, W. Wang, F. Wang, Y. Zhou, and J. Li, “Transferable non-invasive modal fusion-transformer (NIMFT) for end-to-end hand gesture recognition,” *J Neural Eng*, vol. 21, Apr 2024.
- [39] B. Wang, L. Hargrove, X. Bao, and E. N. Kamavuako, “Surface EMG Statistical and Performance Analysis of Targeted-Muscle-Reinnervated (TMR) Transhumeral Prosthesis Users in Home and Laboratory Settings,” *Sensors (Basel)*, vol. 22, Dec 2022.

- [40] K. J. Ottenbacher, Y. Hsu, C. V. Granger, and R. C. Fiedler, “The reliability of the functional independence measure: A quantitative review,” *Archives of Physical Medicine and Rehabilitation*, vol. 77, p. 1226–1232, Dec 1996.
- [41] C. Gonzalez, T. Ganel, R. Whitwell, B. Morrissey, and M. Goodale, “Practice makes perfect, but only with the right hand: Sensitivity to perceptual illusions with awkward grasps decreases with practice in the right but not the left hand,” *Neuropsychologia*, vol. 46, no. 2, p. 624–631, 2008.
- [42] E. rez, H. G. Gonzalez-Hernandez, and J. A. o, “Physical Exertion Recognition Using Surface Electromyography and Inertial Measurements for Occupational Ergonomics,” *Sensors (Basel)*, vol. 23, Nov 2023.
- [43] C. M. Kollod, N. J. Eftimiu, G. Marton, and I. Ulbert, “Classification of semi-automated labeled mindrove armband recorded emg data,” *2022 IEEE 22nd International Symposium on Computational Intelligence and Informatics and 8th IEEE International Conference on Recent Achievements in Mechatronics, Automation, Computer Science and Robotics (CINTI-MACRo)*, Nov 2022.
- [44] I. Loi, A. Grammatikaki, P. Tsinganos, E. Bozkir, D. Ampeliotis, K. Moustakas, E. Kasneci, and A. Skodras, “Proportional myoelectric control in a virtual reality environment,” *2022 IEEE 14th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP)*, Jun 2022.
- [45] D. Farina, R. Merletti, B. Indino, and T. Graven-Nielsen, “Surface EMG crosstalk evaluated from experimental recordings and simulated signals. Reflections on crosstalk interpretation, quantification and reduction,” *Methods Inf Med*, vol. 43, no. 1, pp. 30–35, 2004.

- [46] J. P. Mogk and P. J. Keir, “Crosstalk in surface electromyography of the proximal forearm during gripping tasks,” *J Electromyogr Kinesiol*, vol. 13, pp. 63–71, Feb 2003.
- [47] “Opencv python module.” <https://docs.opencv.org>.
- [48] A. Hashemi Oskouei, M. G. Paulin, and A. B. Carman, “Intra-session and inter-day reliability of forearm surface EMG during varying hand grip forces,” *J Electromyogr Kinesiol*, vol. 23, pp. 216–222, Feb 2013.
- [49] Z. Qing, Z. Lu, Y. Cai, and J. Wang, “Elements influencing semg-based gesture decoding: Muscle fatigue, forearm angle and acquisition time,” *Sensors*, vol. 21, no. 22, 2021.
- [50] M. Wang, C. Zhao, A. Barr, H. Fan, S. Yu, J. Kapellusch, and C. H. Adamson, “Hand posture and force estimation using surface electromyography and an artificial neural network,” *Human Factors*, vol. 65, no. 3, pp. 382–402, 2023. PMID: 34006135.
- [51] D. R. Hax, P. Penava, S. Krodel, L. Razova, and R. Buettner, “A novel hybrid deep learning architecture for dynamic hand gesture recognition,” *IEEE Access*, vol. 12, p. 28761–28774, 2024.
- [52] R. Tchantchane, H. Zhou, S. Zhang, and G. Alici, “A review of hand gesture recognition systems based on noninvasive wearable sensors,” *Advanced Intelligent Systems*, vol. 5, Jul 2023.
- [53] A. Mujahid, M. J. Awan, A. Yasin, M. A. Mohammed, R. Damaševičius, R. Maskeliūnas, and K. H. Abdulkareem, “Real-time hand gesture recognition based on deep learning yolov3 model,” May 2021.
<https://www.mdpi.com/2076-3417/11/9/4164>.

- [54] Y. Khalifa, D. Mandic, and E. Sejdić, “A review of hidden markov models and recurrent neural networks for event detection and localization in biomedical signals,” *Information Fusion*, vol. 69, p. 52–72, May 2021.
- [55] Z. Liu, Y. Feng, and Z. Chen, “Dialtest: Automated testing for recurrent-neural-network-driven dialogue systems,” *Proceedings of the 30th ACM SIGSOFT International Symposium on Software Testing and Analysis*, Jul 2021.
- [56] R. Pascanu, T. Mikolov, and Y. Bengio, “On the difficulty of training recurrent neural networks,” *Proceedings of Machine Learning Research*, vol. 28, 2013.
- [57] Y. Yu, X. Si, C. Hu, and J. Zhang, “A review of recurrent neural networks: Lstm cells and network architectures,” *Neural Computation*, vol. 31, p. 1235–1270, Jul 2019.
- [58] S. W. Pienaar and R. Malekian, “Human activity recognition using lstm-rnn deep neural network architecture,” *2019 IEEE 2nd Wireless Africa Conference (WAC)*, Aug 2019.
- [59] F. Gers, J. Schmidhuber, and F. Cummins, “Learning to forget: Continual prediction with lstm,” *9th International Conference on Artificial Neural Networks: ICANN '99*, 1999.
- [60] A. Graves, N. Jaitly, and A.-r. Mohamed, “Hybrid speech recognition with deep bidirectional lstm,” *2013 IEEE Workshop on Automatic Speech Recognition and Understanding*, Dec 2013.
- [61] N. K. Karnam, S. R. Dubey, A. C. Turlapaty, and B. Gokaraju, “Emghandnet: A hybrid cnn and bi-lstm architecture for hand activity classification using surface emg signals,” *Biocybernetics and Biomedical Engineering*, vol. 42, p. 325–340, Jan 2022.

- [62] K. Team, “Keras documentation: Lstm layer.”
https://keras.io/api/layers/recurrent_layers/lstm/.
- [63] S. Yuan, “Sherry-s-yuan/gesture classification: A project for neurotech uoft. the model classify 6 different hand gestures from emg signals (8 channels) using bidirectional-lstm, achiving an accuracy of 91 https://github.com/sherry-s-yuan/gesture_classification/tree/master.
- [64] “Google colab.” <https://colab.google/>.
- [65] W. Li, P. Shi, and H. Yu, “Gesture Recognition Using Surface Electromyography and Deep Learning for Prostheses Hand: State-of-the-Art, Challenges, and Future,” *Front Neurosci*, vol. 15, p. 621885, 2021.
- [66] J. Luo, Z. Zhang, Y. Fu, and F. Rao, “Time series prediction of COVID-19 transmission in America using LSTM and XGBoost algorithms,” *Results Phys*, vol. 27, p. 104462, Aug 2021.
- [67] S. Qi, X. Wu, W.-H. Chen, J. Liu, J. Zhang, and J. Wang, “Semg-based recognition of composite motion with convolutional neural network,” *Sensors and Actuators A: Physical*, vol. 311, p. 112046, Aug 2020.
- [68] D. Huang and B. Chen, “Surface emg decoding for hand gestures based on spectrogram and cnn-lstm,” *2019 2nd China Symposium on Cognitive Computing and Hybrid Intelligence (CCHI)*, Sep 2019.
- [69] A. Ben Haj Amor, O. El Ghoul, and M. Jemni, “Sign language recognition using the electromyographic signal: A systematic literature review,” *Sensors*, vol. 23, p. 8343, Oct 2023.

- [70] A. Phinoyomark, S. Thongpanja, H. Hu, P. Phukpattaranont, and C. Limsakul, *The Usefulness of Mean and Median frequencies in Electromyography Analysis*, p. 195–220. 2012.
- [71] A. Karpathy, “A recipe for training neural networks,” Apr 2019.
<https://karpathy.github.io/2019/04/25/recipe/#2-set-up-the-end-to-end-training-evaluation-skeleton--get-dumb-baselines>.
- [72] “Ninapro emg database.” <https://nинапро.хевс.чх/>.
- [73] N. Krilova, I. Kastalskiy, V. Makarov, V. Kazantsev, and S. Lobov, “Emg data for gestures,” Jan 2019.
<https://archive.ics.uci.edu/dataset/481/emg+data+for+gestures>.
- [74] M. Ortiz-Catalan, R. nemark, and B. kansson, “BioPatRec: A modular research platform for the control of artificial limbs based on pattern recognition algorithms,” *Source Code Biol Med*, vol. 8, p. 11, Apr 2013.
- [75] A. E. Olsson, A. rkman, and C. Antfolk, “Automatic discovery of resource-restricted Convolutional Neural Network topologies for myoelectric pattern recognition,” *Comput Biol Med*, vol. 120, p. 103723, May 2020.
- [76] B. Leelakittisin, M. Trakulruangroj, S. Sangnark, T. Wilaiprasitporn, and T. Sudhawiyangkul, “Enhanced lightweight cnn using joint classification with averaging probability for semg-based subject-independent hand gesture recognition,” *IEEE Sensors Journal*, vol. 23, p. 20348–20356, Sep 2023.
- [77] J. Charmant, “Kinovea video annotation resource.”
<https://www.kinovea.org/>.
- [78] “Wrnch: An ai computer vision service.” <https://wrnch.ai/>.

- [79] CMU-Perceptual-Computing-Lab, “Cmu-perceptual-computing-lab/openpose: Openpose: Real-time multi-person keypoint detection library for body, face, hands, and foot estimation.”
<https://github.com/CMU-Perceptual-Computing-Lab/openpose>.
- [80] Microsoft, “Azure kinect body tracking sdk download.” <https://learn.microsoft.com/en-us/azure/kinect-dk/body-sdk-download>.
- [81] “Scikit learn database: K-fold cross validation.” https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.KFold.html.
- [82] G. J. Rani, M. F. Hashmi, and A. Gupta, “Surface electromyography and artificial intelligence for human activity recognition—a systematic review on methods, emerging trends applications, challenges, and future implementation,” *IEEE Access*, vol. 11, p. 105140–105169, Sep 2023.
- [83] C. Xu, D. Chai, J. He, X. Zhang, and S. Duan, “Innohar: A deep neural network for complex human activity recognition,” *IEEE Access*, vol. 7, p. 9893–9902, 2019.
- [84] R. E. Nogales and M. E. Benalcázar, “Analysis and evaluation of feature selection and feature extraction methods,” *International Journal of Computational Intelligence Systems*, vol. 16, Sep 2023.
- [85] N. Dodd and E. Espinoza-Wade, “Predicting grip aperture using muscle activation data,” *IEEE Engineering in Medicine and Biology Society (EMBS) 46th Annual Conference*, Jul 2024.