

Integrating Computer Vision with NeuroMuscular Interfacing for the Semi-Autonomous Control of Robotic Limbs



Milia Helena Hasbani

Department of Bioengineering - Imperial College London

Supervised by Professor Dario Farina

2023

Submitted in part fulfillment of the requirements for the degree of Doctor of Philosophy
of Imperial College London and the Diploma of Imperial College London

Abstract

Grasping with an active prosthetic hand is a complex process leading to traditionally un-intuitive control where users cycle through grasp types using buttons, co-contraction, or thresholding. To overcome such limitations of state-of-the-art, we introduce a shared control strategy based on a Human Intention Detection system with assistance from a Computer Vision (CV) system. High-density surface EMG (HDsEMG) is used to control a prosthetic wrist and hand while the CV system analyses visual inputs to select a grasp type which is applied upon hand closure. Kinematic Prediction Models were trained from concurrent HDsEMG and motion capture recordings to predict kinematics of wrist movements and closing and opening of the hand. We systematically investigated the complexity of the training movement sets used to train the models, resulting in successful predictions of kinematics when training with a reduced training set ($r^2 = 0.70$, $RMSE = 0.13$ for a 3-Degree of Freedom (DoF) system). The Kinematic Prediction Models were extended to an online system and validated by individuals with limb loss who were able to control a prosthesis. Participants successfully completed 100% of all target-reaching tasks with a cursor on a screen for 2-DoF movements. In parallel, a CV system was developed and trained to predict one of four grasp types (general, spherical, pinch, tripod) from the image of an object. The novel CV algorithm correctly classified 88% of grasps from unseen images in an offline setting and demonstrated an accuracy of 79% when used in real-time reaching and grasping experiments with an eye-in-hand camera. The two systems were integrated into a real-time shared control framework for the semi-automatic control of a prosthetic hand. The CV system complements the user's direct control and reduces the cognitive effort required to grasp objects, while still allowing the user to feel in natural control of the prosthesis.

Statement of Originality

This Thesis is my own work and reports the results of my original research. Any work that is the result of collaboration or the input of others has been specifically stated within the text. Information derived from the work of others has been acknowledged and referenced.

Milia Helena Hasbani

Copyright Statement

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial-No Derivatives 4.0 International Licence (CC BY-NC-ND).

Under this licence, you may copy and redistribute the material in any medium or format on the condition that; you credit the author, do not use it for commercial purposes and do not distribute modified versions of the work.

When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text.

Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Acknowledgements

Firstly, I would like to thank my supervisor Professor Dario Farina for welcoming me into your group, allowing me the opportunity to work on this exciting research topic, and supporting me throughout the PhD. Thank you also to Dr Deren Barsakcioglu for your support and supervision. I very much appreciate all your insight, advice, and contributions to my research.

I would like to thank the EPSRC for funding my research, and TASKA Prosthetics for allowing us to use their exciting technology. I would like to thank Ruth Nicolson, Dorset Orthopaedic, and Mr Norbert Kang for allowing me to experience clinical settings and place my work into wider context. Thanks also to Carole Maroun for helping me with the cover art design.

Many thanks to all the participants in my studies. I could not have done this research without you. Your insights have been invaluable, and it has been a pleasure working with you and designing technology that I hope will one day have an impact on your everyday lives.

Particularly helpful to me during this time were Moon, Jumpei, Irene, Simone, and Patrick. Thank you for always being so helpful and generous with your time and sharing your knowledge. Thank you to Rachel for collaborating with me, helping me materialise my ideas, and being patient with my supervision. Special thanks to Dr Angela Kedgley for your unwavering guidance and mentorship.

I also had great pleasure working with all my friends, colleagues, and collaborators, in the Neuromechanics and Rehabilitation Technology Group. Thank you to all of you and to all the friends I have made throughout my Imperial PhD journey, whether from the Graduate Students' Union, the Department of Bioengineering, or the Beit Wardening Team, for making my last few years memorable. Special thanks to Julia for your friendship and support as we both navigated our PhD journeys.

I would like to extend my deepest appreciation to James for being so encouraging, patient, and, supportive. Thank you also for all your help as a Python and L^AT_EX wizard.

Finally, I would like to thank my family, especially my parents, for your unparalleled support. Thank you for always believing in me.

Publications and Presentations

Publications

M. H. Hasbani, D. Y. Barsakcioglu, M. K. Jung, and D. Farina, “Simultaneous and proportional control of wrist and hand degrees of freedom with kinematic prediction models from high-density emg”, 2022 44th Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC), pp. 764–767, 7 2022.

B. Grandi Sgambato, **M. H. Hasbani**, D. Y. Barsakcioglu, J. Ibáñez, A. Jakob, M. Fournelle, M. X. Tang, D. Farina, “High Performance Wearable Ultrasound as a Human-Machine Interface for wrist and hand kinematic tracking”, submitted, 2023. TechRxiv version: doi.org/10.36227/techrxiv.22127435

Presentations

Imperial College Department of Bioengineering Department Seminar: PhD Students Showcase, “Integrating computer vision with neural interfacing for semi-autonomous control of robotic limbs”, London (online), 16th June 2021, Oral Presentation, Invited Speaker.

International Society on Electrophysiology and Kinesiology (ISEK) Congress, Quebec City, Canada, June 2022, abstract accepted for an individual oral presentation, presentation cancelled due to COVID.

44th Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC), Glasgow, Scotland July 2022, Poster Presentation.

Contents

0.1	Abbreviations	22
1	Introduction	24
1.1	Motivation	24
1.2	Background	26
1.2.1	Clinical case and current solutions	26
1.2.2	Grasping	31
1.2.3	Computer Vision for grasping objects	35
1.2.4	Computer Vision aided prostheses research	37
1.3	Current challenges and our direction	41
1.3.1	System overview and aims	43
1.4	Thesis structure	44
2	Offline evaluation of myocontrol	46
2.1	Background	46
2.1.1	Machine Learning techniques	48
2.1.2	Electrode set	49
2.1.3	Training time optimisation for myocontrol	50
2.1.4	Experiments	50
2.2	Aim and hypothesis	51
2.3	Methods	52
2.3.1	Experimental protocol	54
2.3.2	Motion capture processing	59

2.3.3	Electromyography processing	60
2.3.4	Kinematic Prediction Models	61
2.3.5	Training set evaluation	63
2.3.6	Statistical methods	63
2.4	Results	64
2.4.1	Regression	64
2.4.2	Classification	70
2.5	Discussion	74
3	Computer Vision Assistance	77
3.1	Introduction	77
3.1.1	Camera selection and placement	79
3.1.2	Grasp type classification	80
3.1.3	Aims	81
3.2	Offline algorithm development and training	82
3.2.1	Image dataset	82
3.2.2	Image pre-processing	83
3.2.3	Algorithm development	84
3.2.4	Results	88
3.3	Computer Vision Real-time System development	89
3.3.1	Hardware	89
3.3.2	System overview	90
3.3.3	Image processing cycle	93
3.3.4	Updating the state machine	94
3.4	Validation of the real-time system	94
3.4.1	Experimental protocol	95
3.4.2	Results	102
3.5	Discussion	104

4	Integration and validation of the shared EMG-Computer Vision system	112
4.1	EMG-CV controller integrated system design	113
4.1.1	TASKA prosthetic hand	114
4.1.2	EMG-RS framework	118
4.1.3	CV-RS integration with EMG-RS framework	122
4.1.4	Socket design	124
4.2	Validation of the EMG-RS with participants with upper limb loss	125
4.2.1	Experiment design and evaluation	127
4.2.2	Experimental protocol	129
4.2.3	Results	136
4.2.4	Discussion	147
4.3	EMG-CV Shared Control System validation	150
4.4	Conclusion	153
5	Conclusion	155
5.1	Thesis contributions and achievements	155
5.2	Future outlook	158
A	EMG feature extraction and evaluation	162
A.1	Variogram	162
A.2	2D cross-correlation	163
A.3	Feature selection	163
B	Permissions for reuse	169

List of Tables

1.1	Comparison of hand prosthetic systems that incorporate computer vision to help grasp objects. Abbreviations: AR: Augmented Reality; CNN: convoluted Neural Networks, EMG: Electromyography; IMU: Inertial Measurement Unit; MMG: Mechanomyography; Ext: Extension; Flex: Flexion; Pro: Pronation; Sup: Supination.	38
2.1	Movement sets performed by the subjects in the different tasks. Movement sets include individual DoF movements as well as sequential wrist and hand movements and simultaneous wrist and hand movements.	58
2.2	Classification of closing and opening accuracy of the 4-Degree of Freedom (DoF) system.	72
2.3	Classification of closing and opening accuracy of the 3-DoF system.	72
3.1	Subject information for CV-RS experiment.	96
3.2	List of objects, their dimensions, and grasping contexts.	99
3.3	Grasp type used by each subject for each grasp context in the both rounds of recording where they were first unrestricted to any grasp type and then limited to the 4 grasp types used in the CV-RS.	102
3.4	Accuracies attained when varying f and W	104
4.1	Participant information for real-time HDsEMG control study.	128
4.2	Properties of the High Density surface Electromyography (HDsEMG) electrodes used in the experiments. Three different types of electrodes were used: GR10MM0808, GR08MM1305, GR08MM1305 (OTBioelettronica, Italy). i.e.d refers to interelectrode distance. The sensor material is Cu with chemical gilding. Electrode images from [157].	128
4.3	High Density surface EMG electrode grid configurations for each participant.	130

4.4	Summary of the different experimental protocols performed with each participant. LL represents participant with Limb Loss.	131
4.5	Complete set of training and testing tasks, with further categorical information for the target reaching tasks.	135

List of Figures

1.1	Autumn is a LEGO [®] Friends character who was born without the lower part of her left arm. Image reproduced from [2], ©2022 The LEGO Group.	25
1.2	Feix's GRASP taxonomy classifies human grasps into 33 categories. ©2016 IEEE, reproduced from [57].	32
1.3	Representation of the selected grasp type options as performed by the TASKA prosthetic hand [59]. ©TASKA Prosthetics 2023.	33
1.4	Overview of proposed shared control for a prosthetic device. Information recorded from the user via HDsEMG and from a camera is used to control a prosthetic hand and wrist.	44
1.5	System diagram for proposed shared controller. Output from a human intention detection system and a computer vision assistance system are used for shared control of a prosthetic hand.	45
2.1	Human wrist degrees of freedom.	48
2.2	Placement positions of the optical markers. US: styloid process of ulna; RS: styloid process of radius; EM: medial epicondyle of elbow; EL; lateral epicondyle of elbow; Mi: head of the metacarpal bone of finger i (i = 1-3); Pi: head of proximal phalanx of finger i (i = 1-3); Di: head of distal phalanx of the thumb (i = 1) and head of middle phalanx of long fingers (i = 2-3). Image from [91] ©2022 IEEE adapted from [110].	54
2.3	Example of 2 repetitions of a movement, here flexion of the wrist.	55
2.4	Grasp types used in Task Set 2.	55
2.5	Example of 1 repetition of a sequential movement, here flexion of the wrist followed by closing the hand.	56

2.6	Example of 1 repetition of a simultaneous movement, here flexion of the wrist combined with closing the hand.	57
2.7	Example of identification of closing and opening events from MCPJ2 and IPJ2 angular velocity. Additionally, any points in the ‘neutral’ (hand fully open) state were also labelled as ‘opening’, while samples in the fully closed state were labelled as ‘closing’.	60
2.8	Overview of Kinematic Prediction Models training and testing.	61
2.8	Examples of the angle data predicted by the linear regressor compared with the true data.	66
2.9	r^2 scores for flexion-extension, pronation-supination, and radial and ulnar deviation, and MCPJ2 angles when different movement sets are used to train and test a linear regressor. Here ‘simple’ refers to movement sets 1 and 2, ‘sequential’ refers to movement set 3, ‘simultaneous’ refers to movement set 4, and ‘all’ refers to all the previously mentioned movement sets.	67

2.10	4-DoF system showing a) r^2 and b) normalised RMSE scores for flexion-extension, pronation-supination, and radial and ulnar deviation wrist angles when different movement sets are used to train a linear regressor, tested on individual wrist movements, sequential wrist and hand movements, and simultaneous wrist and hand movements. Jitter plot shows individual results for each cross-validation of each DoF for each of 8 subjects.	68
2.11	3-DoF system showing a) r^2 and b) RMSE scores for flexion-extension and pronation-supination wrist angles when different movement sets are used to train a linear regressor, tested on individual wrist movements, sequential wrist and hand movements, and simultaneous wrist and hand movements. Jitter plot shows individual results for each cross-validation of each DoF for each of 8 subjects. * shows statistically significant difference with $p < 0.05$	69
2.12	Normalised confusion matrices for the 4-DoF system. a) recall and b) precision for classes predicted when different movement sets are used to train an SVM classifier regressor, tested on individual wrist movements, sequential wrist and hand movements, and simultaneous wrist and hand movements, for 8 subjects.	71
2.13	Normalised confusion matrices for the 3-DoF system. a) recall b) precision for classes predicted when different movement sets are used to train an SVM classifier regressor, tested on individual wrist movements, sequential wrist and hand movements, and simultaneous wrist and hand movements, for 8 subjects.	73
3.1	Eye-in-hand vs eye-to-hand robotic vision configurations.	80
3.2	Adaptation of the Cutkosky grasp taxonomy for the Taska hand.	81
3.3	Overview of CV offline algorithm.	82
3.4	Step 1: Image pre-processing steps.	84

3.5	Distribution of images based on 2 metrics: maximum pixel value in the mapped range 1-255 and the number of nonzero pixels, in the original dataset as well as the balanced dataset.	86
3.6	Elbow plot for the Kmeans clustering model indicating 4 as the most appropriate number of clusters. k is the number of clusters. Inertia is the within-cluster sum-of-square as defined in Equation 3.2.	87
3.7	Examples of K-means clustering (Step 3).	88
3.8	Step 2: Predicted level-sets from SS-CNN with associated RGB image.	88
3.9	Normalised confusion matrix for grasp classification after having labelled a sample of the dataset.	89
3.10	Example of an image processing cycle.	91
3.11	Computer Vision Real-time System Flowchart.	92
3.12	Webcam with attachment and as worn by a participant	96
3.13	Setup of the experiment, showing subject's starting position, webcam wristband, object placement, custom backdrop, and external cameras.	97
3.14	Example outputs from the CV-RS showing the overlay of the object bounding box, object label, and grasp state overlayed over the original input image.	103
3.15	Frequency of class output for the different predicted classes in windows of size 10 frames beginning at the frame number indicated. The 'true' classes associated with each object are indicated as solid lines in each plot, whereas the dashed lines represent incorrect predictions. Data is averaged over 3 trials and both subjects.	105

4.1	EMG-CV shared controller complete system diagram. The user and the camera both observe an object during a reaching task. HDsEMG is recorded from the user and processed to produce instructions for wrist positions and closing/opening the hand per the user's intention. Visual input from the camera is processed to produce Grasp State outputs that are sent to the shared controller, where wrist and hand control mapping takes place. The commands for the prosthetic movement are sent to the local controller to generate movement in the prosthetic hand and wrist. The user then grasps the object via the prosthesis upon their intention to close the hand.	114
4.2	Overview of the equipment used in the EMG-CV shared control system.	115
4.3	TASKA hand and wrist units on the custom stand.	116
4.4	Graphic User Interface with cursor that can be either followed by the user for training or controlled by the user. The cursor is a blue arrow that moves in the x-axis space for flexion-extension, rotates to indicate pronation-supination, and becomes smaller to indicate hand closing.	120
4.5	EMG Real-time System Flowchart. Acquired EMG is processed in real-time to produce positions for flexion-extension, pronation-supination, and hand closing/opening, which are used to control a cursor on the screen.	121
4.6	Initial Socket Design with a clear plastic test socket connected to the QWD for TASKA connection.	125
4.7	Proof-of-concept prototype design. The prosthesis is connected to the Mitt Sleeve via a custom 3-D printed connector.	125
4.8	Experimental Setup for offline validation of EMG-RS. High Density electrodes on the participant's arm are connected to the Quattrocento amplifier. The EMG signals are sent to the laptop where they are processed and used to control a cursor on the screen, as well as the TASKA hand, which is shown here supported on the table.	129
4.9	Example of 1 repetition of a training movement, here flexion and extension of the wrist.	133

4.10	All targets shown in the cursor space.	136
4.9	Heatmaps of the mean EMG activity during held contractions during training tasks for each participant, for both the 3- and 2-DoF control systems. The EMG channels are arranged according to the grid arrangements. Channels with 0mV activity are those that have been manually removed to due noisy signals.	139
4.10	Overview of the observed real-time target-reaching task performance across 6 performance metrics separated by number of target DoFs and the number of DoFs involved in training the system. All metrics besides completion rate are based on successful attempts.	142
4.11	Observed real-time target-reaching task performance across 3 metrics separated by Activation Level (<i>ActLvl</i>) for tasks in the 2-DoF trained system.	143
4.12	Overview of the individual participant performance observed real-time target-reaching tasks in the 2-DoF system across 6 performance metrics separated by number of target DoFs. All metrics besides completion rate are based on successful attempts.	144
4.13	Bar plots of mean values for responses to statements in the user survey for the 2-DoF real-time control system. Responses from Strongly Agree to Strongly Disagree are mapped with values from 1 to 5. Error bars show standard deviation.	146
4.14	Histogram of survey responses. Response map: 1. Strongly Agree 2. Agree 3. Neutral 4. Disagree 5. Strongly Disagree.	147
4.15	User wearing the complete proof-of-concept system with shared control.	152
A.1	2D cross-correlation, as used as a feature for kinematic prediction. Each dot represents one Root Mean Square (RMS) value over a specific window. These values are arranged in grids based on the original EMG channel alignment. X represents cross-correlation between the RMS values of the grids. The three peak values of the resulting cross-correlograms were used as features.	164
A.2	4-DoF system showing r^2 scores across wrist angles when using different features to train the linear regressor with the individual wrist movements	165

A.3	4-DoF system showing normalised Root Mean Square Error (RMSE) scores across wrist angles when using different features to train the linear regressor with the individual wrist movements	166
A.4	3-DoF system showing r^2 scores across wrist angles when using different features to train the linear regressor with the individual wrist movements	167
A.5	3-DoF system showing normalised RMSE scores across wrist angles when using different features to train the linear regressor with the individual wrist movements	168

List of Supplementary Videos

2.1	Experimental setup showing a subject perform a set of extension movements while following the video on the screen. HDsEMG is recorded from the forearm, while motion capture is recorded from markers on the hand and arm. The corresponding tracked motion capture recording is also shown. Video can be viewed at: https://youtu.be/IJTHNLlh_RU	53
3.1	Demonstration of the Computer Vision system being used to automatically control the grasp type of a prosthetic hand. The grasp type changes from target to target: ‘pinch’ for the pen; ‘general’ for the bottle; ‘spherical’ for the ball. Video can be viewed at: https://youtu.be/sMm7I66R-jQ	95
3.2	Video Output from the fixed external cameras during the second round of reaching and grasping tasks. The participant reaches for the pen and grasps it 3 times. Video available at: https://youtu.be/U70Cb6oHigo	98
3.3	Video Output from the eye-in-hand camera during the second round of reaching and grasping tasks. The participant reaches for the pen and grasps it 3 times. The corresponding level-set mask which is output from the CNN is also shown along with the continuously output Grasp States. Video available at: https://youtu.be/Xw4qdvd9KlI	98
4.1	A participant with limb loss at the transradial level controls the TASKA hand via HDsEMG and the Kinematic Prediction Models of the Human Intention Detection System. The intended movements are displayed on the screen, but the participant is following instructions from the researcher resulting in a perceived delay. Video available at: https://youtu.be/xa-9GJdbfBI	145

4.2	Demonstration of the shared control during a pilot experiment with a user with limb loss. The user controls the rotation of the prosthesis as well as closing and opening the prosthetic hand via the Human Intention Detection system. The prosthesis changes grasp type automatically based on the Computer Vision Assistance. The image shows the user wearing the complete proof-of-concept system and reaching towards a bottle.	
	Video available at: https://youtu.be/CGhTeKkKpFw	152

0.1 Abbreviations

AI Artificial Intelligence

AR Augmented Reality

CNN Convolutional Neural Network

CNS Central Nervous System

CV Computer Vision

CV-RS Computer Vision - Real-time System

DoF Degree of Freedom

EL Lateral Epicondyle of elbow

EM Medial Epicondyle of elbow

EMG Electromyography

EMG-RS Electromyography - Real-time System

GS Grasp State

GUI Graphical User Interface

HDsEMG High Density surface Electromyography

IMU Inertial Measurement Unit

IPJ Interphalangeal Joint

LDA Linear Discriminant Analysis

MCPJ Metacarpophalangeal Joint

ML Machine Learning

MMG Mechanomyography

PWM Pulse-Width Modulation

QDW Quick Disconnect Wrist

RGB Red-Green-Blue

RMS Root Mean Square

RMSE Root Mean Square Error

ROI Region of Interest

RS Radial Styloid

SS-CNN Semantic Segmentation-Convolutional Neural Network

SVM Support Vector Machine

TMR Targeted Muscle Reinnervation

UDP User Datagram Protocol

US Ulnar Styloid

VGG Visual Geometry Group

Chapter 1

Introduction

1.1 Motivation

In October 2022, LEGO[®] announced a new generation of LEGO[®] Friends for 2023 [1]. One of these characters, Autumn, has a limb difference – she is missing part of her arm [2], as shown in Figure 1.1. She is not alone. There are an estimated 250,000 upper limb amputees in the UK alone [3] who are lacking the functionality of at least one hand, and whose interaction with their environment and with others has been dramatically influenced by the loss of their limb. The number of patients with an amputation or congenital limb deficiency, both upper and lower limb, attending specialist rehabilitation service centres in the UK is estimated at 55,000–60,000 [4], with NHS England spending approximately £60 million per year on these services [4]. In the United States, there were around 541,000 people living with different levels of upper limb loss in 2005; a number which is expected to at least double by 2050 [5]. Although some people are born with limb loss, approximately 3500 and 5200 upper limb amputations are reported each year in Italy and the UK, respectively [6].

The hand is a marvellous and dexterous tool that humans use to interact with their environment - from grasping and manipulating objects to interacting with others and using touch and pressure sensory information to understand the world around us. We are heavily dependent on our hands for a multitude of tasks; thus, the loss of a hand due to amputation can be traumatic with lasting psychological and physical effects.

Some of this functionality can be restored with the use of a prosthesis, which can act as a partial substitute for the lost limb, thus enhancing the user’s quality of life. Still, most prosthesis users are unsatisfied with their prosthesis, such that only around 45% of prosthesis users use it throughout the day, most of them preferring cosmetic usage over functionality (around 80%) [7]. The reason behind this rejection of prostheses can be traced back to a number of things, including time consuming and complicated training needs, the absence of tactile feedback, poorly developed fine motor control, and difficult and unintuitive control over the device [8–10]. As a result, these devices do not feel like a natural extension of the body [11, 12].

To tackle this challenge, we aim to improve the embodiment of the prosthetic by focusing on the task of grasping an object using a prosthetic hand. We aim to make this as natural and intuitive as possible, alongside a more innate overall control of the wrist function. This will be achieved through first developing a robust electromyographic control of 2 DoFs of the hand and wrist, to detect and follow the user’s intention intuitively. The control of the prosthesis will be shared by the user with a Computer Vision (CV) algorithm. The integrated CV system will aid the user in selecting the grasp type and finger configuration to be used when grasping objects. This will complement the user’s control, and reduce the effort and cognitive load needed by the user to complete grasping tasks.

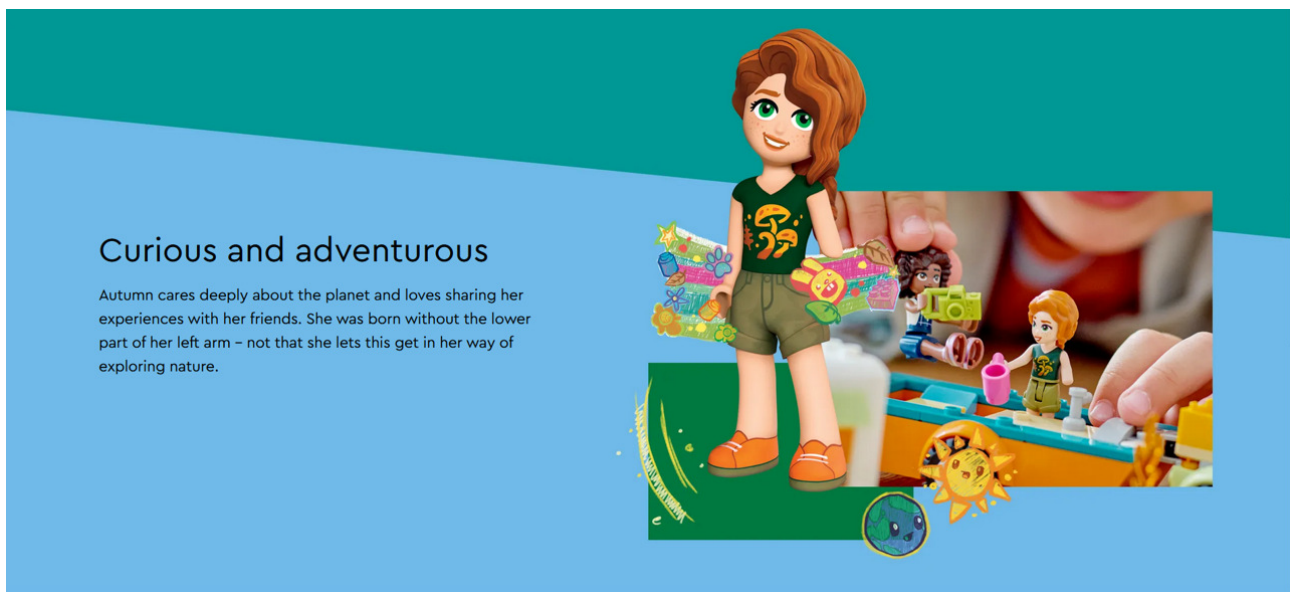


Figure 1.1: Autumn is a LEGO[®] Friends character who was born without the lower part of her left arm. Image reproduced from [2], ©2022 The LEGO Group.

1.2 Background

1.2.1 Clinical case and current solutions

Limb loss can occur for a number of reasons, whether the limb loss is congenital, or acquired. Amputation etiologies include trauma due to combat injury or accidents, infection, dysvascular disease, dysvascular disease with comorbidity of diabetes, cancer [5, 13], or elective amputation for bionic reconstruction e.g. after a brachial plexus injury [14], among other causes. Upper limb loss can take place at different levels of the arm, and can be classified as transcarpal, wrist disarticulation, transradial, elbow disarticulation, transhumeral, shoulder disarticulation and forequarter [6]. The limb loss can also be unilateral or bilateral. In this thesis we are targeting limb loss at the transradial level, which accounts for 12% of the upper limb absence in Italy and the United Kingdom [6].

Use of a hand prosthesis can enhance the lifestyle of a person with limb loss to a certain extent by acting as a partial substitute for the lost limb. Hand and wrist prostheses have undergone numerous improvements throughout their existence, such that there are currently a wide variety of passive and active systems offering diverse functionality [8]. There are three main groups of hand prostheses: passive functional or cosmetic devices, body powered devices, and myoelectric control devices. Each prosthetic group has advantages and limitations for the user, with no prosthesis able to fully reproduce the complexity of the natural hand.

- **Passive functional devices** include aesthetic hands with no intrinsic moving parts. They address cosmetic concerns of limb difference, and can be used for non-grasping tasks such as stabilising, supporting and pushing [15].
- **Body powered devices** use a cord to open the hand with one simple motion, usually pulled by the shoulder or the opposite shoulder [15]. These are usually limited to a single grip, with reduced functional movements. Compensatory movements such as over-use in the contralateral side are thought to lead to bio-mechanical imbalance issues in the user [15]. Although they provide natural sensory feedback to the user [8], body-powered prostheses lack the cosmetic effect of cosmetic prostheses which can have social and psychological impact [15].

- **Externally or electrically powered devices, typically, myoelectric control devices**, are powered by an external source. Remaining muscles activate sensors to control motors [15] to move the hand and wrist. Limitations of myoelectric prosthetic hands, in both a clinical and research setting, include the lack of sensory feedback to the user, the weight of the device, and concerns about durability [15]. The user may also be required to learn multiple coordinated movements to control the device and grip patterns, which can require significant user concentration and training [15].

Myoelectric control of prostheses

Many prosthetic devices use myoelectric control, where surface Electromyography (EMG) signals are recorded from muscle activity from the residual limb [16].

For myoelectric control, the user's intention can be detected in the form of a muscular contraction by placing surface electrodes on the skin over the location of a muscle [17]. The sum of the electrical contributions of the active motor units detected by the electrodes is known as surface, or global, EMG [18]. These myoelectric signals are generated during varying levels of static muscle contraction and dynamic limb motion, and, since they are repeatable and distinguishable, they can be used to control a system – such as a prosthetic device [17, 19].

Amputees are still able to generate myoelectric potentials in residual muscles. After amputation, many patients still perceive the presence of their lost limb, also known as a phantom limb [20]. This missing limb is still represented in the brain, which can still send commands and signals to the muscles, both present and missing. Thus, activating the phantom limb involves contracting the residual muscles in the residual limb, which can be detected by myoelectric potentials on the surface of the skin and used to control a prosthesis [21]. As such, myoelectric control can provide a useful human-machine interface, and this type of control is being explored excessively in relation to its application in prosthetics, as amputees are still able to generate these signals [16, 17, 22]. Of course the degree to which individuals with limb loss can control a prosthesis depends on the anatomy of the residual limb and muscles. There are two main approaches that are currently utilised in myoelectric control of prostheses for

transradial amputee users: direct, or sequential, control and pattern recognition [22].

Electrically powered hand prostheses are typically controlled by direct control using two bipolar EMG electrodes placed over the flexors and extensors of the limb, in a conventional method which has been used commercially and clinically since the 1960s [23]. The two EMG signals are typically used to control the velocity of closing and opening the prosthetic hand [21]. Thus, muscle groups of antagonistic muscles are linked to a single DoF, limiting the control to one DoF at a time. In some cases, the user can still control other DoFs, or switch the grasp type used, by cycling through different DoFs or pre-set hand grasp patterns. This switching is typically done via a co-contraction [16]. Heuristics for switching between DoFs are time-consuming or counter-intuitive and do not allow for simultaneous control of multiple DoFs.

To overcome the limitations of traditional control, and allow for more advanced control strategies, various Machine Learning (ML) techniques have been proposed, mainly in a research setting, while becoming more popular in a clinical setting. These techniques involve extracting information from a greater number of EMG channels, typically 4 to 10 [17, 24, 25]. Pattern recognition is the most common approach. In pattern recognition, classification techniques are used to map features extracted from myoelectric inputs to a discrete set of intended outputs [26–28]. This allows users to control more than one DoF without having to switch methods. However, the DoFs can only be controlled sequentially, in an on/off fashion. The discrete control can be extended by adding proportional activation for each DoF based on EMG amplitude [27, 29]. Alternatively, movements can be simultaneously combined, as in [30], but this may reduce the classification accuracy, and does not extend to independent proportional control of the combined DoFs.

In order to improve simultaneous and proportional control over multiple DoFs, regression based approaches have more recently been explored, mainly limited to a scientific setting [31–36]. Instead of mapping EMG features to a discrete set of tasks, regression maps the EMG features to continuous and proportional movements within the range of that DoF. This allows for independent, proportional, and simultaneous control which can also be performed over combinations of multiple DoFs. For example, a prosthesis user can close the hand while rotating the wrist.

For both classification and regression, the ML models need to be trained. The models are unique to each individual, and may require large amounts of data in order to predict kinematics with a relatively high accuracy. Typically, features, rather than raw signals, are extracted from the data and fed into a classifier or regressor [17]. The most common features include time domain features which are based on signal amplitude [17]. These include features such as: linear envelope, mean absolute value, RMS, and zero crossings among others [37]. Frequency domain features such as power spectral density and moments, short time Fourier transform, and spectral magnitude averages, are also employed [37]. Other features include time-frequency domain and spatial filtering features [37, 38].

The EMG data used for training also needs to be labelled with the corresponding movements. For able-bodied participants the labels are obtained as the user performs the relevant movements while EMG is recorded. However, for participants with limb loss, labels can be generated from the kinematics or forces of the intact limb during bilaterally mirrored movements [33, 39]. Alternatively, labels can be generated from visual cues that are provided for the users to follow [35]. This method can be used for people with bilateral or uni-lateral limb loss, or able bodied subjects.

EMG can be recorded intramuscularly or with electrodes on the skin. Intramuscular EMG allows for detection of activity at the level of single motor units. However, due to the invasiveness of intramuscular recordings, intramuscular EMG is more traditionally used in neurophysiology research [40]. We focus our research on surface EMG, especially for myocontrol where only surface EMG is currently used commercially.

The number of electrodes used to detect surface EMG signals vary in different devices. Most commercial sequential control devices use only 2 EMG electrodes, on antagonistic muscle pairs. For more advanced control mechanisms, some devices use a small array of electrodes, such as the MyoTM band which consists of 8 electrodes arranged in a wearable band [41]. In recent research, high density electrode arrays are being used, [22, 28, 33, 42], although not yet commercially, providing improved performance for myoelectric control. HDsEMG uses multiple closely spaced electrodes overlying a limited area of the skin [43]. HDsEMG allows spatial EMG activity to be recorded, alongside temporal activity, allowing new muscle characteristics to be detected [43]. Use of HDsEMG along with spatial

features has also shown advancements in myoelectric prosthesis control over more traditional feature analysis [28, 42]. In [42], for example, HDsEMG was used to implement a variogram as a feature for classification of 9 movement classes with an accuracy of $\sim 95\%$. HDsEMG also has the advantage of being non-invasive. HDsEMG research continues to evolve with more advanced research techniques, such the use of blind source separation [44] to more closely capture the activity of individual motor unit action potentials, but these are still limited to research settings.

Even though ML approaches have shown promising results in academia for a number of decades, the impact of ML approaches on clinical and commercial applications is still limited [21]. All major prosthetic manufacturers still rely on the traditional two-channel control. Improvements in terms of robustness of ML methods has led to a few companies [45, 46] offering pattern recognition based controllers. Despite these improvements in terms of robustness of ML methods in clinics, several limitations still exist. High performance results under controlled laboratory conditions usually decreased substantially with the introductions of real world factors [21]. These factors include: small electrode shifts [47, 48], changing arm positions [49], variable skin conditions [19], additional load due to the weight of the prosthesis [50], or time between algorithm training and application [51].

Surgical techniques

Several clinical and surgical techniques and solutions have emerged to complement prosthetic use. During amputation, although some muscles are lost, their innervating nerves may remain. Targeted Muscle Reinnervation (TMR) is a surgical intervention that involves redirecting these nerves to other muscle tissues [8]. This serves to amplify the activity of the redirected nerves, allowing their electrical activity to still be captured via HDsEMG recordings and blind source separation [8].

To address the issues that arise from wearing a socket, such as sweat, discomfort, pressure on the limb, lack of natural feedback, an alternative form of biomechanical interface can be provided via Osseointegration [8]. Osseointegration involves insertion of a metal implant to the residual skeletal structures to attach to the prosthesis [8].

1.2.2 Grasping

Alongside providing sensory information through touch, and playing a great role in social interactions and communication, one of the main functions of the hand is interacting with our surroundings, including through grasping objects. This thesis focuses on prehensile movements of the hand, where an object is “seized and held partly or wholly within the compass of the hand” as opposed to non-prehensile movements where objects are simply manipulated by pushing or lifting movements without grasping [52].

Reaching and grasping are complex motions during which the neuromusculoskeletal system, following instructions from the Central Nervous System (CNS), achieves several challenging tasks. The task can be broken down into bringing the hand into the appropriate location in the vicinity of the object for reaching, while concurrently, the fingers form into a grip in anticipation of contact for grasping [53]. The process begins with the visual assessment of the target object to determine its properties, both extrinsic (location and orientation relative to the observer) and intrinsic (size and shape) [53]. This information is used by the brain to plan the motor task; the visual information about the extrinsic properties informs the reaching aspect, while information about the intrinsic properties informs grasping [53]. The motor task is also guided through a closed loop feedback system involving vision, tactile information, proprioception and other prior experience of the object properties. The complex sensory-motor planning and implementation all takes place with incredible ease and seemingly little effort - the person may be concerned with the high level goal of the task, while the actual implementation happens in an almost automated matter at a subconscious level [54].

Grasping itself is musculoskeletally a very complicated task involving up to 35 muscles in the forearm and hand [55]. For example,

“Three phases operate when grasping large objects. In the first the hand opens widely from action of both the long extensors and the intrinsic. Then the object is surrounded largely by intrinsic action. Closure and firm grasp result from the strong action of the extrinsic flexors and the intrinsic of both the thumb and fingers. Finally, slight extension of the wrist tightens the grip even more.” [55]

In terms of the actual posture of the hand during prehensile function, grasping can take place in many

different forms depending on the size, shape, and orientation of the object being grasped as well as the way in which it will be manipulated. Factors influencing the grasp chosen include the shape of the object, size of the object, miscellaneous factors such as weight and texture, as well as intended activity [52]. Despite the variability in grasping postures, Napier [52] groups them into either power or precision grasps, or a combination of both. The taxonomy of grasping is further studied by many, with Cutkosky later [56] extending this into the field of automation such that grasp types can be classified according to rules based on the precision of the task and the size of the object [56]. Further, according to Feix taxonomy [57], human hand grasps can be classified into 33 types, as demonstrated in Figure 1.2.

Opp: VF:	Power						Intermediate		Precision					
	Palm		Pad				Side		Pad				Side	
	3-5	2-5	2	2-3	2-4	2-5	2	3	3-4	2	2-3	2-4	2-5	3
Thumb Abducted	1: Large Diameter	2: Small Diameter	3: Medium Wrap	10: Power Disk	11: Power Sphere									
Thumb Adducted	17: Index Finger Extension	4: Adducted Thumb	5: Light Tool	15: Fixed Hook	30: Palmar		16: Lateral	25: Lateral Tripod					22: Parallel Extension	
							29: Stick							
							32: Ventral							

Figure 1.2: Feix’s GRASP taxonomy classifies human grasps into 33 categories. ©2016 IEEE, reproduced from [57].

Grasping an object with a prosthesis

While deciding on which grasp type to use when interacting with an object may come naturally to most of us, the task is not as effortless for prosthesis users. Trans-radial prosthesis users are typically able to complete the reaching aspect of the reaching and grasping task, but may struggle with grasping itself.

Many prosthesis users may not even have access to multi-grip myoelectric prostheses. NHS England has only recently in August 2022 introduced new policy [15] to recommend the use of multi-grip myoelectric control prosthetic hands. Until then, prosthetic hands available on the NHS consisted of body powered, passive functional, or single grip myoelectric control prostheses [15]. Thus, both body powered and myoelectric prosthesis users would be constrained to a basic general grasp for opening and closing. This may constrain the user to limited tasks, where precision tasks become especially difficult. Relying on a single grip means users may struggle to pick up certain objects and may need compensatory movements such as manipulation of the object with the other hand to facilitate the grip [15]. Some commercial prosthetic devices offer different attachments that can be replaced for different tasks [58].

With increasing mechanical dexterity, multi-grip myoelectric prosthesis users may have access to a range of grasp types, with different commercial prostheses offering different gestures – ranging from 1 all the way to around 20. Examples of 4 grasp types, which can be replicated by the TASKA prosthetic hand are demonstrated in Figure 1.3.



Figure 1.3: Representation of the selected grasp type options as performed by the TASKA prosthetic hand [59]. ©TASKA Prosthetics 2023.

However, even if there are multiple grasp types, prosthesis users have additional steps to go through

in activating the grasping. They have to press a button or co-contract (tense up their forearm muscles) [16], to cycle between different grasp types to choose from. For example with the Hero-arm from Open Bionics, the only DoF that can be activated by the user is closing/opening¹. Activating the flexors closes the hand. Once the hand is closed a co-contraction fixes the hand in the closed position, and another co-contraction releases the hand back to the open position. Activating the extensor muscles allows the user to cycle through a series of pre-set grips. Pressing a button on the hand switches between 2 groups of pre-set grips. Some prosthetic hands have an accompanying mobile phone application which can be used to help calibrate, set thresholds, and choose grip patterns. For example with the TASKA hand [60], 2 sets of preset grips can be determined using the accompanying mobile phone application. These preset grips can then be cycled through by pressing respective buttons on the back of the hand, one button for each set of grips.

The need to cycle through different grasp types via co-contraction, thresholding, or button pressing is due to the complexity of understanding the choice of grasp type from only the EMG. From the extrinsic muscles of the arm, it is not possible to differentiate between precise single finger movements with EMG, and this may be increasingly challenging with varying lengths of residual limbs and muscle availability. This is especially challenging with a limited number of EMG channels, as is the case with most commercial prostheses, as users already have to cycle through different DoFs and different movement options, so changing the grasp type becomes an extension of this.

Relying on this cyclical switching between grasp types though is cumbersome and unintuitive [9]. Pressing a button is also not a viable option for people with bilateral upper limb loss. Thus, one of the challenges in clinical prosthetic control that we wish to address is the lack of intuitive control, especially when it comes to grasping. To develop a more intuitive system for grasping objects, we propose to integrate CV to simulate visual sensory information and complement the information extracted from EMG. As previously discussed, vision plays a major role in assessing the choice of grasp type. As such, in this thesis, we will pursue the notion that sensory information, including visual information, can be collected and input through the robotic system itself, allowing for semi-autonomous control of

¹This is based on conversation with a Hero Arm user, without an active wrist. The Hero Arm also comes with the option of an active wrist system

the prosthesis, such that the user is able to perform fewer input actions when grasping an object in the correct alignment.

1.2.3 Computer Vision for grasping objects

In order to tackle the challenge of grasp type choice and implementation in the field of automated robotics, several robotic systems use sensors, CV, and ML techniques for grasp detection [61]. Cameras and CV have been implemented across a range of fields in robotics to recognise objects and position and control the end-effector's orientation. The camera is typically fixed on the robot (or on the user's head, if being used in a rehabilitation engineering setting) [62], or placed on top of the end-effector (eye-in-hand) (see Figure 3.1). Robots can be programmed to grasp previously known and, thus, recognisable objects or even scripted to carry out series of complex manipulation tasks; however, autonomously grasping a novel and previously unknown object is a separate and challenging task [63]. There are several tools and methods that can be utilised if the 3D model of the object to be grasped is available or if the object is previously known, but, in practice, it is difficult to reconstruct a full 3D model of a novel object from a 2D stereo vision system in order to apply these methods [63]. As a result, work has been done on using vision to locate grasping points on an object and end-effector orientation in real-world settings, though mostly for planar objects, including methods that rely on local visual features, simplifying and classifying the object shape, as well as addressing feasibility and force-closure criteria [63]. When it comes to more common objects typically found in the home or office, identifying where to grasp the object becomes a challenging perception problem, since these objects come in an array of various shapes and sizes, in the midst of a scene cluttered with background noise [63].

In [63], visual features are used to predict the ideal grasping point of an object, applied through a learning approach in which the projection of a grasping point on the image plane is identified over several camera positions. Supervised learning is used to identify patches containing the grasping points based on training data comprising objects with the labelled location of the grasping point in each image [63]. The points identified in each image are then triangulated using a probabilistic

approach to quantify the 3D grasping point [63].

Similarly, [64] also employs CV methods to determine the best contact points for grasping, expanding to multiple contact points rather than a 3D region as in [63]. Applied to a desk cleaning robot, the system in [64] employs depth data as well as intensity image data to create a point cloud from which grasp candidates are chosen and ranked based on their features using a trained Support Vector Machine (SVM) model.

The system in [65] aims to provide integrated vision-guided grasping for service robots. It uses 2D image information to segment objects, extract their boundaries, and compute their orientation. The grasping approach is based on determining the optimal points for grasping, rather than the specific grasp based on the identification and classification of an object, using shape features to preserve grasp stability based on curvature, angular, and distance thresholds [65]. The camera is placed on the end-effector (eye-in-hand configuration) [65].

In [66], CV (as well as eye-tracking) are used for identification and selection of targets in the experimental workspace, using a Kinect sensor to record the 3D point cloud depth data and Red-Green-Blue (RGB) information from the workspace. Points were grouped as objects using a Euclidean clustering extraction algorithm before being classified as either cylinders or spheres, and the geometric properties of these shapes, including the radius and orientation, were used to select the optimal grasp types [66]. The camera was fixed on a mount above the robotic hand, replicating the location of where the head would be, and visual feedback was provided to the user via a Graphical User Interface (GUI) on a monitor [66].

Several other ML and classification techniques have been employed for robotic grasp recognition. Some systems, for example, use a sliding window detection framework which employs a classifier to determine whether small patches of an image constitute good grasps for an object [61]. Instead of applying the classifier to numerous parts of the image, the system in [66] makes use of a large database of objects with labelled grasps and applies a Convolutional Neural Network (CNN) approach, applying a single, albeit large, network once to an image to directly predict grasp coordinates.

The problem addressed in robotic grasping is mainly where to grasp the object as well as how to

perform the reaching task and orient the end-effector, whereas with prostheses, the question is how the object is going to be picked up. The user can move the prosthetic hand into the right area and can choose best location to place the hand prosthesis.

1.2.4 Computer Vision aided prostheses research

Seeing as, biologically, the process of grasping an object relies heavily on sensory information from a number of sources, researchers have attempted to mimic these natural systems through integrating different sensors and techniques within prosthetics to improve their control. Hybrid myoelectric control systems use sensor fusion methods by combining EMG with other bio signals or external sensory inputs to control prostheses [67, 68]. Surface EMG sensors were fused with different types of sensors, such as acceleration and gyrosensors [12], [13] from [69].

As such, CV and ML techniques have been implemented to aid with grasp selection in prosthetic hands in research settings. While visual hand prostheses is a new area of research, some recent work has shown significant promise for eventual clinical impact. A brief comparison of these systems can be found in Table 1.1.

The device in [70, 71] uses a camera, alongside a depth ultrasound sensor and a laser, all incorporated onto the dorsal side of a prosthetic hand. Quantified thresholds of the size of the object are used to determine the grasp type; the tripod and pinch grasp are selected for small and very small objects respectively, while lateral or palmar grasps are selected for larger (and wider) objects [71]. These are extended to incorporate different sizes for each grasp types in [70]. The orientation of the object is used to decide the rotation of the wrist in [71], using external motion capture to automatically rotate the wrist. However, in [70], the artificial hand is attached to a user via a splint, who is able to rotate the hand physically through pronation/supination. 2 EMG electrodes are also used to control opening and closing the hand through extension and flexion respectively. The CV system is activated by activating the finger extensor muscles. The work in [70] was then extended in [54, 72], where a similar rule based principle for grasp detection was used by fitting preset 3D shapes to each object. Cameras were fixed to custom-made glasses, while other sensors were embedded within the prosthesis

Table 1.1: Comparison of hand prosthetic systems that incorporate computer vision to help grasp objects. Abbreviations: AR: Augmented Reality; CNN: convoluted Neural Networks, EMG: Electromyography; IMU: Inertial Measurement Unit; MMG: Mechanomyography; Ext: Extension; Flex: Flexion; Pro: Pronation; Sup: Supination.

System reference(s)	Biological Sensors	Biological control	Vision/other Sensors	Grasping algorithm	No of grasp types	Other info
Dosen et al. 2010 [70, 71]	2 EMG electrodes	Ext to open, Flex to close	Camera, depth ultrasound sensor, laser all on hand	Size threshold	4 (with different sizes)	
Markovic et al. 2014, 2015 [54, 72]	2 EMG electrodes	Sequential manual control of multiple DoFs, switched via co-contractions, triggering the automatic system	Camera on glasses, Inertial Measurement Unit (IMU) on prosthesis	Predefined 3D object shape fitting and thresholds; hand is automatically rotated	2	AR feedback interface
Gardner et al. 2014, 2017 [73, 74]	MMG sensor	Open hand or initiate automatic control	Eye-to-hand camera, 2 IMUs on arm	Edge detection for object identification; reaching motions for grasp choice selection	3	
Degol et al. 2016 [75]	N/A	N/A	Camera in palm	VGG CNN predicts grasp type from image	5	Embedded GPU microprocessor
Andrade et al. 2017 [76]	Myo TM band	Cycle through operations for autonomous system	Camera on hand	Object recognition proposes appropriate grasp types which user can cycle through	14	Simulated prosthetic environment
Ghazaei et al. 2017, 2019 [77, 78]	2 EMG electrodes	Indicate when to take a picture	Camera on hand	CNN to extract features and classify grasp types	4	
N Bu et al., 2017; Shima et al. 2018; He et al. 2019, 2020 [79–82]	2 EMG electrodes or Myo TM band	Operation of automated system	Offline images; or camera on hand	CNN Object classification for preset motion & orientation planning	N/A	3-finger gripper
Fukuda et al. 2021 [69]	Myo TM band	Motion decisions	Camera on end-effector	Bimodal non-linear mapping from EMG and CV to motion planning	10	3-finger gripper
Mouchoux et al. 2021 [83]	8 EMG electrodes	Close the hand when using autonomous grasping system; manipulate object in 2-DoF otherwise	RGB-Depth Camera on glasses, IMU and motion capture on arm	Autonomous system decides on grip type, amount of wrist Flex/Ext, Pro/Sup, & size of hand aperture	2	AR feedback interface

itself to create an artificial view of the scene and analyse the environment [72]. The system also made use of Augmented Reality (AR) on the glasses for feedback which allows the user to provide feedback on grasp size to a feedback loop via EMG. Although this system allows the user to have some control over some aspects of the grasp and interact with the system, providing feedback, this is still in some ways sequential and unintuitive, requiring additional accessories and steps before grasping can take place. The work was more recently further extended in [83], with the use of 8 myoelectric electrodes. They extended the automated system in [72] for use in cluttered scenes along with automated control of the grip type, amount of wrist Flex/Ext, Pro/Sup, and size of hand aperture during reaching and grasping task [83]. However, the user can still rely on myoelectric classification control (4-6 classes / 2-3 DoFs) when manipulating the object / when not in a reaching/grasping task.

In [76], EMG recorded from the Myo™ band is used to operate an automated system which assigns a label to objects in images and matches a choice of grasp type to each object label. The user can then cycle through these grasp type options using the EMG. Instead of using EMG, [73, 74] uses Mechanomyography (MMG), simply to open the hand if closed or start the automatic loop if open. IMUs on the users arm are used to determine if the arm is already moving to reduce noise artifacts in the MMG. A CV system separates images from a camera on the user’s head into blobs and then matches objects to the blobs. A few possible grasp types are assigned to each object. The IMUs are also used to record data during the reaching phase. The extracted intention for interaction is used to select a suitable grasp type from the previously proposed grasp types.

As image recognition and CV applications and techniques have advanced, deep learning methods have also been applied to the control of prosthetics [69, 75, 77, 79–82]. Degol et al.[75] use a trained CNN to directly classify grasp types based on object images; however, they do not integrate any biological sensors. In [77], a webcam was placed directly on a 2-channel EMG controlled prosthetic. Using a large database of images classified with their respective grasp type, a CNN was applied to extract features from the object to classify it into grasp types. The EMG signals are only used to activate the CV algorithm.

The research in [79–82] also use object classification, however they use this to determine the

orientation of a 3-finger gripper. For each classified object, a preset motion for the gripper is calculated based on a lookup table for each object. Shima et al. [80] extend this concept introduced in [79], by adding information from a depth sensor to classify objects into 6 categories of shapes. In [79, 81, 82], EMG signals from either 2 electrodes or a MyoTM band are used to trigger closing/opening or to operate the automated CV system. He et al. [81] also use an IMU to localise the camera's orientation, whereas [82] focuses on the target object selection phase of such a system. Instead of using rule-based EMG for operation of a CV system, [69] attempt to combine the EMG and vision information inputs into one bimodal network. They use nonlinear modelling to map between the two kinds of sensor inputs and a motion output through training a bimodal neural network in an end-to-end manner [69]. They use their framework to classify movements into 5 object categories with 2 grasping motions per object.

All these deep learning approaches are limited to the objects that they train with for classification. However, a potential drawback of relying on object classification is that the number of household and office everyday items is unquantifiable, thus attempting to use object identification for grasp selection becomes impractical, as not every object can be considered [77].

Accordingly, the application of such CV systems is not limited to upper limb prostheses but can be applied to a wide variety of disciplines, whether in rehabilitation engineering or in automated robotics. CV and ML techniques for grasp detection are also being applied to other scenarios in rehabilitation engineering. For example, an exoskeleton that learns the user's reaching and grasping patterns [84]. Another example is outlined in [62] where a robotic system uses CV to detect the target object, understand the relative orientation of the object to be grasped, and set the corresponding orientation of the end-effector for grasping. This is applied in assistive robotic devices such as robotic orthotic gloves and tele-operated robotic hands [62]. The artificial perception system in [85] uses a Kinect RGB-Depth sensor and CV algorithms for the control of functional electrical stimulation (FES) assisted grasping. The system tracks the hand and performs object analysis to identify the position and shape of the object with respect to the position of the hand, and thus select the grasp type [85].

1.3 Current challenges and our direction

From the current solutions described above, we can collate a few challenges that we wish to address in this thesis. Although there have been significant efforts invested to advance control algorithms, resulting in more capable systems [8], the abandonment rates of bionic hands by the users remain high [9, 10, 86, 87]. Almost 30% of the participants in the survey in [86] who have suffered below-elbow amputations, rejected the use of prosthesis due to lack of comfort, weight, and unsatisfactory function. The reason behind this abandonment of prostheses can be traced back to a number of things, including time consuming and complicated training needs, the absence of tactile feedback, poorly developed fine motor control, and difficult and unintuitive control over the device [8–10]. As a result, these devices do not feel like a natural extension of the body [11, 12].

[87] suggests that most researchers seem to agree that electrical activity recorded from the muscle or peripheral nerve is the most appropriate means of controlling a powered prosthetic hand. However, while users desire a highly dexterous powered prosthetic hand that is lightweight and strong, they do not seem to use them when available [87], indicating a need for more robust myoelectric control methods. Additionally, below elbow prosthesis users with motorised grippers, but only passive wrist units have shown greater measures of compensatory movements compared to able-bodied participants [88]. Olsen et al. [89] found that active wrists significantly reduced the subjective workload on the prosthesis users. Thus, one of the main focuses of this research is to ensure a robust control strategy over multiple DoFs. As complicated training times are also a point of contention for complex myoelectric control systems, we will also aim to address this.

Another challenge in clinical prosthetic control that we wish to address is the lack of intuitive control, especially when it comes to grasping. Cyclical switching between grasp types is cumbersome and not always suitable for people with bilateral limb loss if pressing a button is required. Automation can help reduce the number of movements and cognitive load needed to perform grasps in the correct alignment. Specifically, computer vision can simulate the significant role that the physiological sense of vision plays in assessing the choice of grasp type. Thus, to improve the intuitiveness of grasping objects with a prosthesis, we propose to integrate CV to complement the information extracted from

EMG and allow for semi-automatic control of the prosthesis. In addition, it is usually difficult to create a general myoelectric pattern recognition model that translates to multiple users due to the movement patterns differing from person to person, but CV can (largely) be applied to most users [81, 90].

Although the idea of combining an autonomous system with the user control of a prosthesis has been previously explored, and is inevitable considering bio-mimicry of design, there are some limitations in the current research proposals that we aim to address. Firstly, most of the presented systems, except for [83], use very basic EMG to control when to take a picture or when to close/open the hand. In this case the EMG is used to operate an automated system, rather than in an intuitive shared control strategy. This does not allow for any other wrist or prosthetic functionality besides grasping. However, we aim to truly integrate a shared control where the user is still able to control the wrist, as mentioned above, to allow for reduced compensatory movements and greater functionality. Although the systems in [72, 83] allow the user to have some control over some aspects of the grasp and interact with the system, providing feedback, this is still in some ways sequential and unintuitive, requiring additional accessories and steps before grasping can take place. We aim to reduce the cognitive burden and need for additional steps. In these systems as well as in [73], the camera is also mounted on the head or externally to the prosthesis. This requires the use of additional accessories, and although it provides a direction of the user’s gaze, it does not provide the point of view of the hand prosthesis itself.

Moreover, several systems rely on training on very specific objects or tasks [73] or on deep learning for object classification [76, 79–82]. Markovic et al., Markovic et al., and Mouchoux et al. [54, 72, 83] also rely on objects conforming to generalised shapes. These are thus limited to specific applications, and do not all generalise to novel objects, or novel uses of objects. The research in [79–82], while proposing interesting use cases for CV for prostheses, apply these techniques for motion planning for a three finger gripper. This does not take into account multiple grasp types for different objects that may be proposed in more complex prostheses. They also rely on orienting the gripper automatically, along with [83], which may be a suitable approach to reduce the cognitive burden of the user. On the

other hand it reduces some of the control the user has over the system, and thus may affect the user's sense of embodiment. While it is evident that providing automation for grasp selection can simplify the myoelectric control of a prosthetic hand for the user, grasp selection should not only be based on an object's size, shape and orientation, but on the intention of the user [72]. As such, in this thesis we propose a system for shared semi-autonomous control of a prosthetic hand. Since a prosthetic hand is an extension of human body, its control should reflect the user's intention.

1.3.1 System overview and aims

We aim to focus on user experience and intuitive grasping. As such, any sensors should be integrated within the prosthetic so there are no external devices. The EMG should not just be used as a threshold to decide when to take a picture or open/close the hand. Instead, the EMG should be used for continuous control the wrist function, regardless of the use of the fingers. At the same time, the intention to close the hand should come naturally and intuitively from EMG. Our grasp detection models are also innovative, and they generalise to novel objects and different object approaches. We are proposing a shared control strategy with the following specifications, such that the user feels that they are in control of the prosthetic as much as possible:

- Dexterous
- Real-time
- User-control centric
- Intuitive
- Switch between 4 grasp types
- 2-DoF proportional wrist control

These will be realised with the following method:

- Human Intention Detection via HDsEMG

- Simultaneous 2-DoF wrist control
- Hand activation command detection
- CV Assistance
 - Visual input and analysis
 - Grasp type selection
 - Applied upon hand closing

All of this has not been achieved with current commercial or research efforts. The control should come from the user's intentions the majority of the time through EMG – which will provide control for the wrist and the intention to close the hand or open it. The CV Assistance system should be responsible only for selecting the grasp type – this will be applied when the user gives the instruction to close the hand. An overview of the overall system is presented in Figure 1.4 and in Figure 1.5, which also indicates a breakdown of the thesis structure.

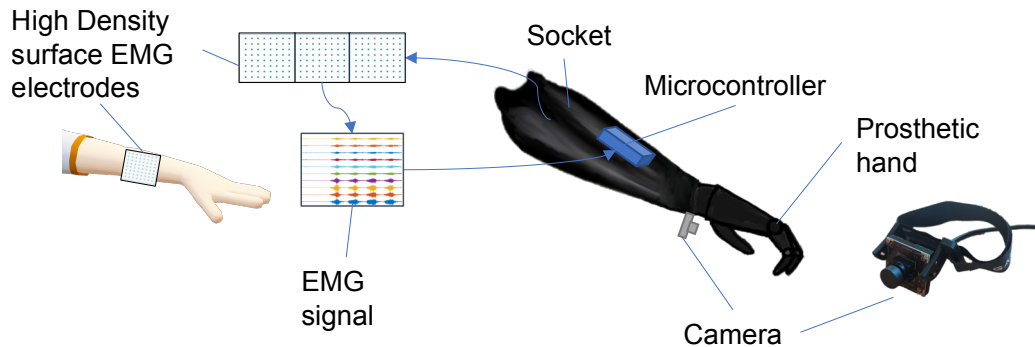


Figure 1.4: Overview of proposed shared control for a prosthetic device. Information recorded from the user via HDsEMG and from a camera is used to control a prosthetic hand and wrist.

1.4 Thesis structure

In Chapter 1, I have placed this thesis in wider context surrounding limb loss and the need for greater intuitiveness over the control of prostheses. I have reviewed the use cases for CV within prosthetic control, and I have highlighted the overall aims of the thesis.

In Chapter 2, I develop an offline EMG control system for the prediction of hand and wrist kinematics in 4 DoF, allowing for simultaneous and proportional control of flexion-extension, pronation-

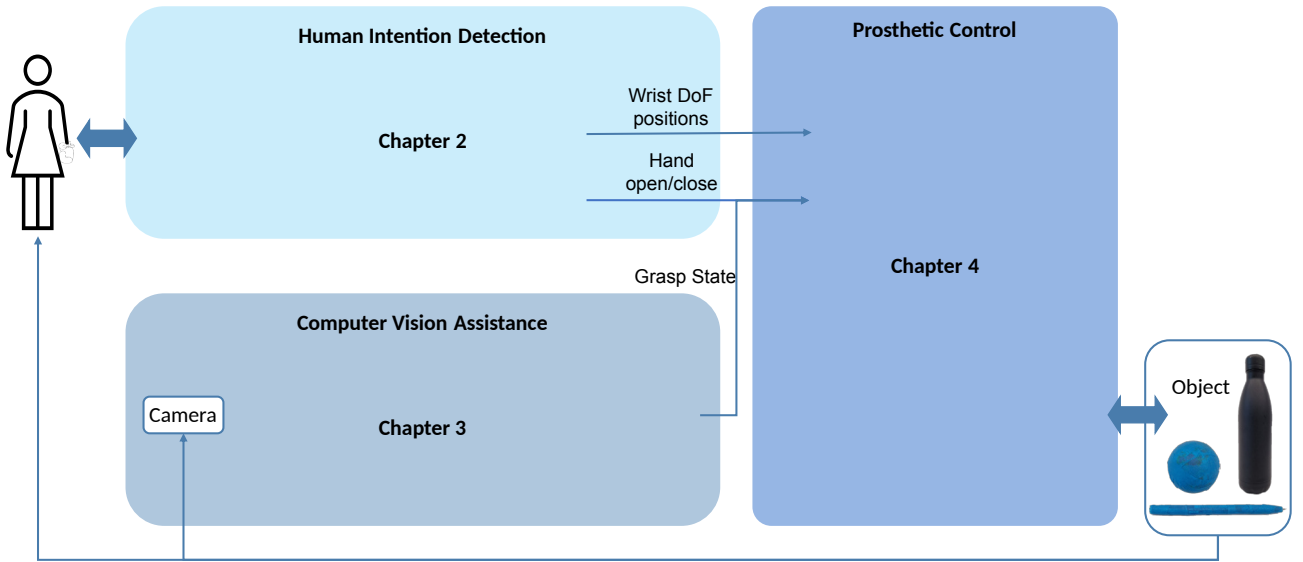


Figure 1.5: System diagram for proposed shared controller. Output from a human intention detection system and a computer vision assistance system are used for shared control of a prosthetic hand.

supination, and radial-ulnar deviation, as well as simultaneous control of hand closing and opening. I also design a study to record EMG and motion capture simultaneously, in which I evaluate the proposed control system in 3 and 4 DoFs and investigate the effect of reducing the movements involved in training the system.

In Chapter 3, I develop a novel CV approach to determine the grasp type of a given object. I train and validate the system offline before adapting it to use in a real-time continuous system that could be integrated with a prosthetic hand. The Computer Vision Assistance system is validated in real-time in a reach and grasp experiment.

In Chapter 4, I integrate the two previously developed systems into one real-time framework (EMG-CV shared controller). I first integrate the Kinematic Prediction Models developed in Chapter 2 into a real-time framework before also integrating the Computer Vision Assistance. The framework is connected to a prosthetic hand and wrist in a proof-of-concept smart prosthetic system. I then validate the Human Intention Detection System with participants with limb loss, before validating the use of the EMG-CV shared controller.

Finally, I review the content of this thesis in Chapter 5 together with a discussion of its contributions and potential future work.

Chapter 2

Offline evaluation of myocontrol

In this chapter¹, we focus on the offline development of the human intention detection controller, which is a part of the overall shared control system. We aim to develop a 3-DoF myoelectric control of the hand and wrist to improve intuitive control and overcome existing limitations of myocontrol, including the inability to control more than one function at a time, complex training requirements, and reduced numbers of DoFs that can be controlled concurrently. We applied regression techniques to achieve independent proportional and simultaneous control for the wrist. At the same time, we complemented this with classification techniques to determine the intention to grasp, regardless of the position or movement of the wrist. We then devised an experimental study, in which we used motion capture to acquire precise kinematics of the hand and wrist, recorded concurrently with HDsEMG, during a series of wrist and hand movements in individual, sequential, and simultaneous DoFs. This allowed us to train our prediction models with various schemes in an offline setting to inform the real-time implementation of the controller.

2.1 Background

Myoelectric prostheses can help restore functionality for users with upper limb loss or deficiencies, by predicting the user intention based on EMG signals recorded from the residuum [8, 92]. However,

¹This chapter is based on [91], in which the study and part of the analysis are described. However, the text has been modified from the original article.

a proportion of people with upper limb loss are not satisfied with their prosthesis (around 33% of surveyed participants rejected their upper limb prosthesis in [93] and [9]). Dissatisfaction tends to be due to time consuming and complicated training needs, as well as non-intuitive control [8, 9]. In [94], 33% of myoelectric prosthesis users reported problems with the functionality of their prosthesis, and 15% used the prosthesis under 8 hours a day. Of the unilateral transradial amputee veterans surveyed in [13], only 18% use a single- or multi-DoF myoelectric prosthesis, with 55% using a body-powered device, 4% using a cosmetic prosthesis, and 23% opting not to use any device.

To tackle such challenges with uptake of myoelectric prostheses, we aim to make the task of grasping an object as natural as possible, alongside a more intuitive overall control of the wrist function. This will be achieved through first developing a robust electromyographic control of 4 DoFs of the hand and wrist, following the user’s intention. The control shall be shared between the user’s intention, detected through surface electromyography, and assisted by CV for grasp control, with the instruction to close the hand acting as the switch between the two systems. Thus, not only does the myoelectric control need to be robust, but the detection of the decoding of the intention to close the hand must be precise too, no matter in which position the hand is when the instruction to close arrives. The 4 DoFs of the system include 3 DoFs of the wrist – flexion-extension, pronation-supination, and radial-ulnar deviation – as well as opening/closing of the hand. These movements are shown in Figure 2.1. It is important to mention that the radial-ulnar deviation is investigated in this work; however, not all commercial prosthetic hands are able to perform this function. In fact, most current commercially available myoelectric control systems lack the possibility to control the wrist at all, which is critical in achieving tasks in a natural way. This can cause a host of issues such as in the shoulder due to compensatory movements that are required to overcome this lack of functionality [95]. Therefore, it is very important to achieve the desired wrist function in 3 DoFs, and one of the main research questions involves how to establish a robust and intuitive simultaneous and proportional paradigm to accomplish this.

To switch between EMG control and CV control in the shared control system, as described in Figure 1.4 and Figure 1.5, our approach is to differentiate between wrist and hand function using the

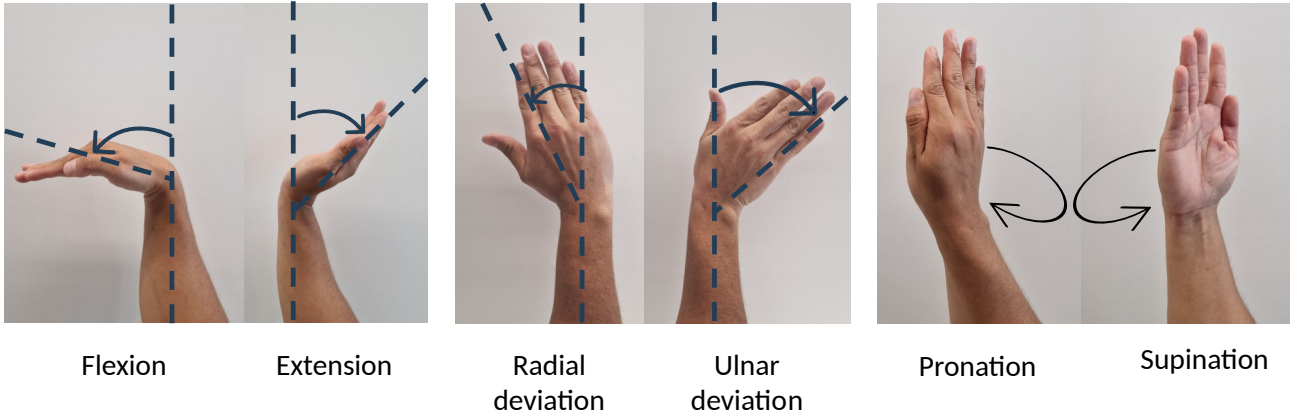


Figure 2.1: Human wrist degrees of freedom.

EMG signals. With this distinction, we should be able to use the classification in a decision tree which would determine the type of control used in the shared control. This involves exploring which EMG features and ML techniques would allow the EMG signals to be classified according to whether they represent wrist function or hand function (opening/closing of the fingers).

2.1.1 Machine Learning techniques

The main focus for upper limb prosthesis control research has been pattern recognition techniques used to classify the movements into different DoFs, most commonly using time-domain or autoregressive features with Linear Discriminant Analysis (LDA), SVM, or hidden Markov models [96]. The main drawback is the ability to control only one DoF at a time (sequential control). Alternatively, regression techniques can be used for simultaneous and proportional control of multiple DoFs [31, 32]. These include linear regression, multi-layer perceptrons, and kernel ridge regression, where continuous outputs are estimated for each DoF, instead of class selection [21, 31, 32, 97–103].

Nevertheless, regression-based algorithms currently fail to provide reliable control over more than 2 DoFs [22, 92, 98]. Recently, various solutions have been proposed to overcome the challenges of simultaneous control from surface EMG, including varying the number and type of sensors [67], or combining sequential and simultaneous control strategies [104]. Here, we aim to develop a 3-DoF myoelectric control of the hand and wrist to improve intuitive control and overcome existing limitations of myoelectric control, including the inability to control more than one function at a time, by combining regression and classification methods. We applied regression techniques to achieve independent proportional

and simultaneous control for the wrist. At the same time, we complemented this with classification techniques to determine the intention to grasp, regardless of the position or movement of the wrist.

2.1.2 Electrode set

While we could have used a small number of channels, as is used traditionally both commercially (2 channels) and in research settings (up to 10 channels), we now have access to higher density systems which provide better spatio-temporal resolution across various muscles of interest. Thus, we decided to use a greater number of channels, specifically with HDsEMG, because 1) we aim to generate a full database of EMG signals from multiple locations as a basis for future research and 2) we assumed that multi-channel recordings would improve the likelihood of good regression over 3 DoFs, which we are aware is challenging. As such, to increase the chances of a successful kinematic prediction, the use of HDsEMG was chosen for the promise it has shown in similar applications [28, 42], since it can provide richer information from muscle activity [22, 33]. The HDsEMG will also enable the investigation of spatial domain features for motion prediction. The electrodes will be placed on the skin, just below the elbow. Since the aim of the research project is to be beneficial for individuals with transradial limb loss as the end user, the electrode matrices were chosen to be placed just under the elbow, as this is the residuum section retained by most transradial amputees. This is also the most common placement for such electrodes when used for prosthesis control, especially when it comes to the wrist control as the bodies of the muscles responsible for moving the wrist are found here, as well as some extrinsic finger muscles. These include Flexor Carpi Ulnaris, Flexor Carpi Radialis, Flexor Digitorum Superficialis, Pronator Teres, Extensor Carpi Radialis Longus, Extensor Carpi Radialis Brevis, Extensor Digitorum (Communis), Extensor Digiti Minimi, Extensor Carpi Ulnaris, Flexor Digitorum Profundus, and Supinator [105].

Using HDsEMG does complicate the myocontrol aspect of the research. However, the offline development underpins the rest of the EMG development for the real-time control system. In that case we would be considering how users with limb loss would be able to control the system. Although a smaller number of electrodes may be simpler for them to control, it would not allow for control of

more than 1 DoF simultaneously. Additionally, we need to accommodate for a variation in individual anatomies in the end user, so using an array of HDsEMG electrodes would increase the ability to control multiple DoFs, which in turn would allow for controlling the wrist as well as initiating closing and opening the hand once the CV element is introduced.

2.1.3 Training time optimisation for myocontrol

Increasing the training set of movements collected from the user can also mitigate some challenges in simultaneous control, although this can be onerous for the patient [8]. Simultaneous movements are also particularly difficult to train since people with limb loss may have difficulties in repeating them consistently. Therefore, it is important to develop control systems that can extrapolate simultaneous movements when trained on only individual or sequential movements. Thus, we also aimed to investigate in a systematic manner which movements are most suitable for training the models for wrist movement and grasping, in order to reduce time and complexity for learning and training when implemented in an online or clinical setting. As a result, in the study, subjects performed a series of wrist and hand movements in individual, sequential, and simultaneous DoFs. This allowed us to train our prediction models with various schemes in an offline setting to inform the real-time implementation of the controller.

2.1.4 Experiments

To train the EMG signals for classification, they first need to be labelled, which can be performed in a straightforward manner when the user is following a strict pattern of movements for training and testing. However, in real-life scenarios, such movements will not be so well-defined. In addition, we would like to increase the freedom and the range of movements performed in a training scenario, to better reflect everyday motions. However, labelling such movements would prove tricky for purely EMG recordings, as the movements and their timings would be difficult to define. Thus, to more accurately label such movements, we incorporate motion capture into the study, alongside the EMG acquisition. This follows the concept that recording the kinematics will allow us to identify movements

in a combination of DoFs, enabling us to correctly label the EMG recordings during the movements. Motion capture can be used to calculate the angles between different sections of the hand and wrist in various DoFs; thus, they can help to identify which motions are taking place at each point in time. By mapping EMG to kinematics we aim to identify the desired movements from the EMG signal.

This chapter will focus on creating a system with simultaneous and proportional control in 3 DoFs, as well as simultaneously with closing/opening control, with the following requirements: minimal training, and robustness over non-stationary conditions. Following from this, the user decides at which point to start grasping the object, which is done when the instruction to close the hand is given through the EMG interface. The user must feel in control of the prosthesis, and the whole process should be as natural as possible.

2.2 Aim and hypothesis

By using motion capture to define and identify the different motions of the hand and wrist, including the intention to close the hand, and employing HDsEMG, which provides high spatio-temporal information, we hypothesise that we can robustly infer the motions of the wrist as combinations of the following DoFs: flexion-extension, pronation-supination and radial-ulnar deviation. We also hypothesise that we can discern the intention to close the hand robustly in a variety of movement combinations, while comparing how these differences affect the HDsEMG signals produced by the subject. In addition, we hypothesise that we are able to discern the intention to close the hand robustly in a variety of movement combinations, while comparing how these differences affect the HDsEMG signals produced by the subject. This study will also enable us to create a dataset of simultaneous HDsEMG and motion capture recordings for movements of the hand and wrist, which provides a good basis for a number of different analyses. This will allow us to test and validate algorithms that can identify wrist movement vs hand movement (closing and opening of the hand) as well as algorithms that may allow simultaneous control of multiple DoFs of the wrist.

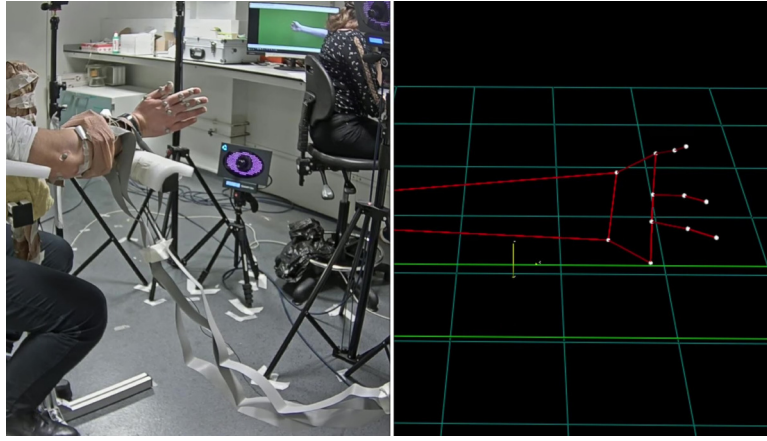
The study addresses the following research questions:

- Can we robustly infer the motions of the wrist as combinations of the following DoFs: flexion-extension, supination-pronation for a 2 DoF wrist; and flexion-extension, radial-ulnar deviation, supination-pronation for a 3 DoF wrist?
 - What is the minimum training data set that we can use, and still retain robust performance?
 - Is the regressor capable of predicting wrist angles while ignoring confounding factors such as the hand movements?
- Can we robustly distinguish between two classes of hand movements: hand closing/closed, and hand opening/open?
 - What is the minimum training data set that we can use, and still retain robust performance?
 - Is the classifier capable of predicting hand movement while ignoring confounding factors such as the wrist movements?

2.3 Methods

HDsEMG and optical marker motion capture were recorded concurrently from the lower arms of 8 able-bodied subjects (mean age 27.13 ± 5.84 ; 6 males, 2 females; all right-handed), performing a series of wrist and hand movements in individual, sequential, and simultaneous DoFs. The experimental procedure was approved by the Imperial College Research Ethics Committee (JRCO number: 18IC4685). During the experiment, the subjects were guided to an adjustable stool surrounded by the motion capture cameras. The subject’s right (dominant) arm was positioned on a metallic cushioned support frame resting at a neutral position (hand relaxed and Ulnar Styloid (US) pointing downwards). A computer screen was positioned outside of the capture area at a comfortable viewing distance for the subject and videos were shown to guide them in performing the required movements. The setup and an example of a subject performing one task with the corresponding motion capture recording are shown in Supplementary Video 2.1.

Three adhesive HDsEMG grids of 8x8 electrodes with 10-mm inter-electrode distance were placed on the subject’s arm below the elbow (except for one subject who only required two grids due to a



Supplementary Video 2.1: Experimental setup showing a subject perform a set of extension movements while following the video on the screen. HDsEMG is recorded from the forearm, while motion capture is recorded from markers on the hand and arm. The corresponding tracked motion capture recording is also shown. Video can be viewed at: https://youtu.be/IJTHNLih_RU

smaller arm circumference), covering the bellies of the wrist extensor and flexor muscles. Conductive paste was used to improve the electrode-skin contact. Prior to electrode placement, the skin above the target muscles was shaved and cleansed with abrasive cream and alcohol. The electrodes were secured with medical tape. The acquisition of the HDsEMG channels was realised through an earlier version of the *myoViewer*, a MATLAB (MATLAB 2019a, Mathworks) programme developed in our research group [106], using the Quattrocento amplifier (OTBioelettronica, Italy). The monopolar HDsEMG signals were recorded, amplified, sampled at 2048 Hz, A/D converted to 16 bits, and digitally band-pass filtered (10-500 Hz). An optical marker motion capture system (SMART DX, BTS Bioengineering) composed of 8 cameras was used to record the positions of 14 markers placed on the hand, wrist and elbow at 200 frames per second. 20mm sized markers were used, and the marker placement is shown in Figure 2.2. The configuration for the marker placement was modified from [107–110] for the hand and [111, 112] for the rest of the arm. The motion capture cameras were positioned to ensure that throughout the duration of the movements each marker would be visible by two or more cameras. This was a carefully managed and developed setup to ensure no markers would be hidden during recordings, and the cameras were calibrated before each recording session. A trigger signal was used to synchronise the EMG and motion capture recordings. An Arduino Uno was used to generate an active low trigger pulse to be sampled both by the EMG amplifier and by the Motion capture system and used to align the recordings.

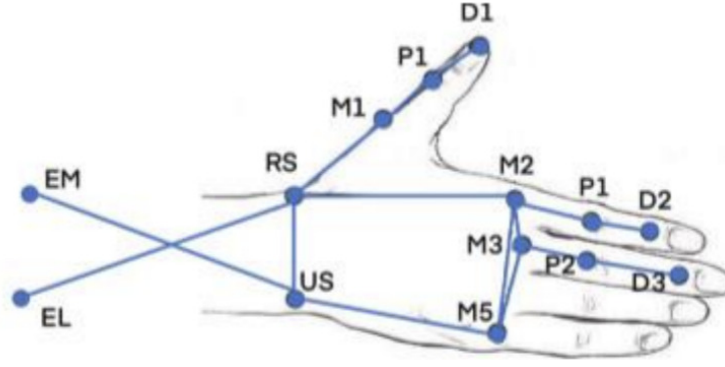


Figure 2.2: Placement positions of the optical markers. US: styloid process of ulna; RS: styloid process of radius; EM: medial epicondyle of elbow; EL: lateral epicondyle of elbow; Mi: head of the metacarpal bone of finger i ($i = 1-3$); Pi: head of proximal phalanx of finger i ($i = 1-3$); Di: head of distal phalanx of the thumb ($i = 1$) and head of middle phalanx of long fingers ($i = 2-3$). Image from [91] ©2022 IEEE adapted from [110].

2.3.1 Experimental protocol

The tasks, listed in Table 2.1, required subjects to move their wrist or hand from neutral to the required position in one or more DoF (2s), hold (for 3s), then return to neutral (2s). Each movement was repeated 5 times, recorded continuously, with 2s of rest between repetitions. For sequential movements, the subject first moved the wrist to the target, which they maintained while closing and opening their hand (transition times of 2s).

Task Set 1 - tasks in 1 DoF

The first task set involves the subject moving their wrist or hand from neutral to the extreme position in one DoF, holding it there, and moving it back to neutral. This is done following a video on a screen with a movement transition time of 2s, and a contraction period of 3s. There were 5 repetitions of each movement, recorded continuously, with 2s of rest between each repetition. The DoFs associated with the wrist included here are flexion of the hand, extension of the hand, pronation of the forearm, supination of the forearm, radial deviation of the wrist, and ulnar deviation of the wrist. There was a short rest period between each recording of 5 trials. The first 2 repetitions of the flexion movement recording are represented in Figure 2.3.

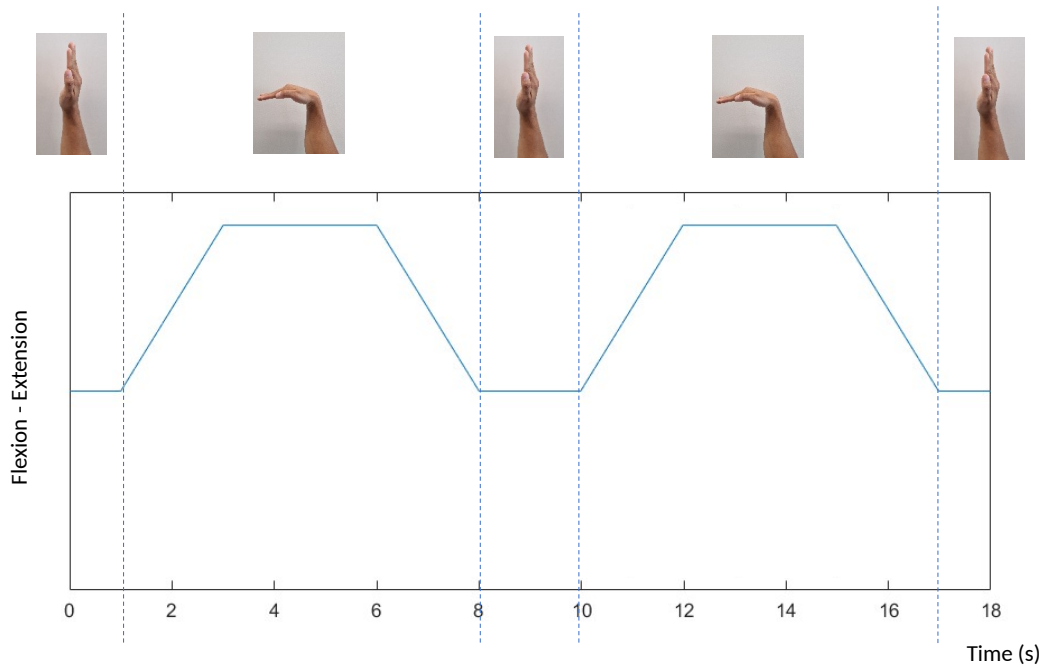


Figure 2.3: Example of 2 repetitions of a movement, here flexion of the wrist.

Task Set 2 - grasping movements

For the second task set, the subject was asked to close and open their fingers in different grasp types. This is done following a video on a screen with a movement transition time of 2s, and a contraction period of 3s. There were 5 repetitions of the movement, recorded continuously, with 2s of rest between each repetition. The grasp types to be included here are: General grasp (power sphere), tripod pinch, pincer precision pinch, and fixed hook grasp. There was a short rest period between recordings for the different grasp types. The different grasp types used are shown in Figure 2.4. As a note, the different grasp types were recorded, but were not addressed in the data analysis.

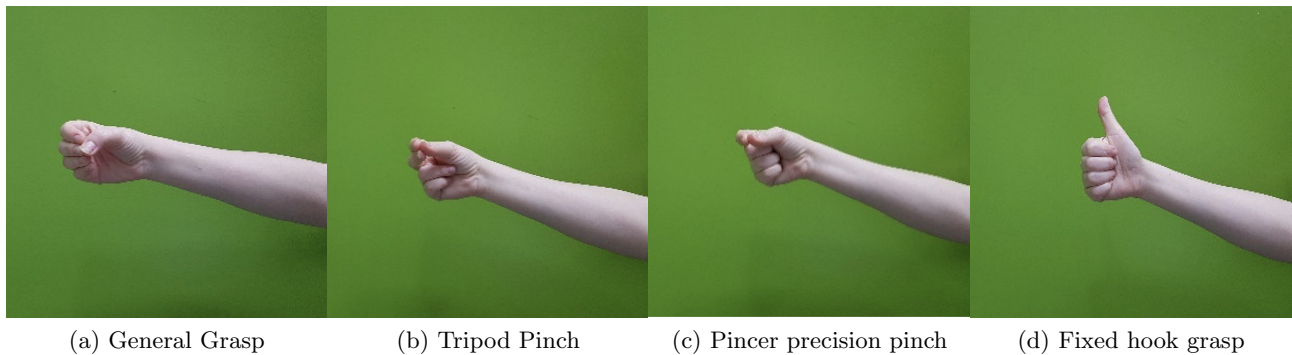


Figure 2.4: Grasp types used in Task Set 2.

Task Set 3 - closing the hand in different positions

For the third task set, the subject was asked to close their fingers (general grasp) while the hand was in different positions. They first moved the wrist to the extreme position of one DoF, for example flexed wrist, and then closed the hand while the wrist is flexed as shown in Figure 2.5. They were asked to hold the closed position, before opening the hand again, all while maintaining the flexed (or other starting) position. This is done following a video on a screen with a movement transition time of 2s to move the wrist, followed by 2s transition for closing the hand and a contraction period of 3s. There will be 5 repetitions of the finger movement, with 2s of rest between each repetition. This was followed by returning the wrist to the neutral position over a period of 2s. The wrist movements for the initial position were the following: flexion of the wrist, extension of the wrist, pronation of the wrist, supination of the wrist, radial deviation of the wrist, and ulnar deviation of the wrist.

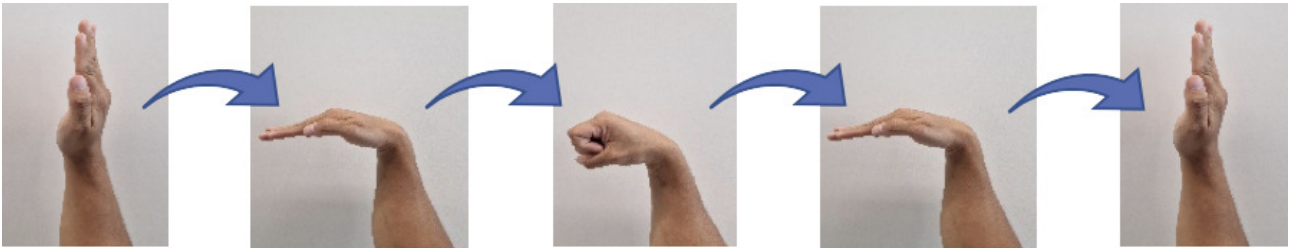


Figure 2.5: Example of 1 repetition of a sequential movement, here flexion of the wrist followed by closing the hand.

Task Set 4 – combined DoFs of the wrist and hand

The fourth task set involved the subject moving the wrist in one DoF while closing the fingers simultaneously, holding the position, and then returning to neutral as shown in Figure 2.6. This is done following a video on a screen with a movement transition time of 2s, and a contraction period of 3s. There were 5 repetitions of each movement, recorded continuously, with 2s of rest between each repetition. The DoFs included here were flexion of the hand while closing the fingers, extension of the hand while closing the fingers, pronation of the forearm while closing the fingers, supination of the forearm while closing the fingers, radial deviation of the wrist while closing the fingers, and ulnar deviation of the wrist while closing the fingers. There was a short rest period between each recording

of 5 trials.

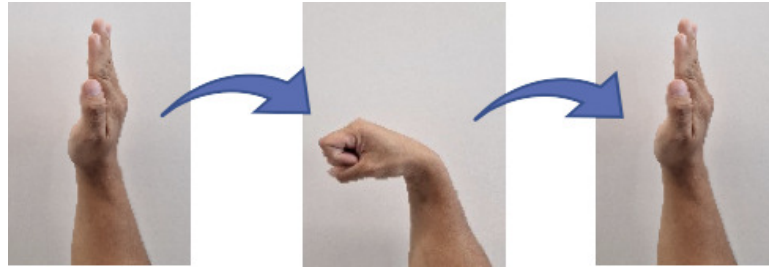


Figure 2.6: Example of 1 repetition of a simultaneous movement, here flexion of the wrist combined with closing the hand.

Task Set 5 – combined DoFs of the wrist

The final task set involved moving the wrist in multiple DoFs simultaneously. The hand started in the neutral position, and then the subject moved through four positions: 1. wrist flexed and radially deviated, 2. wrist extended and radially deviated, 3. wrist extended and ulnarly deviated, and 4. wrist flexed and ulnarly deviated, before returning the hand to the neutral position. This set includes two recordings: 1. The movements outlined were performed in succession, moving steadily from one position to the next over 2s, and holding the position for 2s, over 5 repetitions, with 2s rest in the neutral position. 2. The movements outlined were performed in succession, moving steadily from one position to the next over 2s, over 5 repetitions, with 2s rest in the neutral position (without holding each position distinctly).

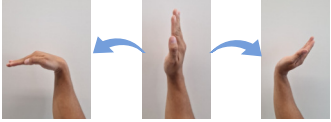
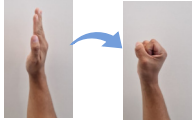
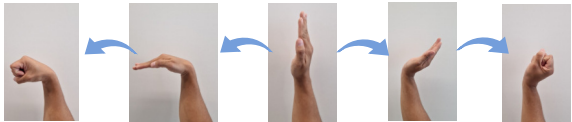
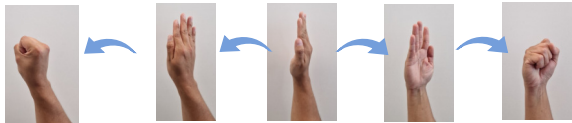
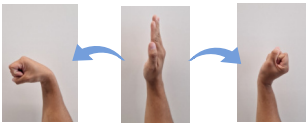
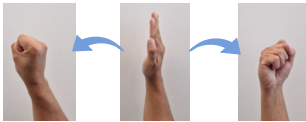
Movement set	Description	Degrees of Freedom
Individual wrist	Separate 1 DoF wrist movements: Wrist Flexion-Extension	
	Wrist Pronation-Supination	
	Wrist Radial-Ulnar Deviation	
Individual hand	1 DoF Hand closing/opening	
Sequential wrist and hand	Opening/closing hand in different wrist positions: Hand Opening-Closing after Wrist Flexion-Extension	
	Hand Opening-Closing after Wrist Pronation-Supination	
	Hand Opening-Closing after Wrist Radial-Ulnar Deviation	
Simultaneous wrist and hand	Opening/closing hand while wrist moves in 1 DoF: Hand Opening-Closing during Wrist Flexion-Extension	
	Hand Opening-Closing during Wrist Pronation-Supination	
	Hand Opening-Closing during Wrist Radial-Ulnar Deviation	

Table 2.1: Movement sets performed by the subjects in the different tasks. Movement sets include individual DoF movements as well as sequential wrist and hand movements and simultaneous wrist and hand movements.

2.3.2 Motion capture processing

The relevant motion capture data was segmented based on the trigger signal. The marker coordinates (x,y,z) in a 3D reference space were interpolated using a 1st order spline function and filtered with a low-pass 2nd order Butterworth filter (Python 3.7.7, Python Software Foundation). The cut-off frequency was determined using residual analysis filtering [113, 114] and was used for all data recorded for one subject. The following joint angles were then calculated: flexion-extension, radial-ulnar deviation, and supination-pronation of the wrist; Metacarpophalangeal Joint (MCPJ)s and proximal Interphalangeal Joint (IPJ)s of the first three fingers. The finger angles were calculated following the methods described in [110].

Wrist angles were calculated from an Euler XZY rotation between the hand reference frame (as in [110]) and the forearm reference frame (adapted from [111]), where the origin is the midpoint between the US and Radial Styloid (RS). The X-axis is a vector from US to RS, and the Z-axis is a cross between the X-axis and a temporary Y-axis (line connecting the midpoint between the Lateral Epicondyle of elbow (EL) and Medial Epicondyle of elbow (EM) to the origin, pointing distally). The angles were then low-pass filtered (5 Hz, 2nd order) and resampled to match the frequency of the EMG features (16 Hz).

The resampled angle data was also used to split each recording into 5 parts, one trial for each of the 5 repetitions in the recording. The segmentation was performed using a semi-automatic cumulative summation algorithm [115] applied to detect abrupt changes in the angle data, followed by a manual check and fine tuning.

Data Labelling

To determine the grasp intention, labels were generated to classify the data into hand ‘opening’ or ‘closing’, using the angles of the index finger, as adapted from [110]. Hand ‘closing’ was defined as samples in which the IPJ2 angular velocity was below 10% of its minimum velocity. Hand ‘opening’ was when the MCPJ2 angular velocity was above 10% of its peak velocity. Peak velocities were calculated from data in movement set 2. Any points in the ‘neutral’ (hand fully open) state were also

labelled as ‘opening’, while samples in the fully closed state were labelled as ‘closing’. An example of the labelling of the angular data is shown in Figure 2.7.

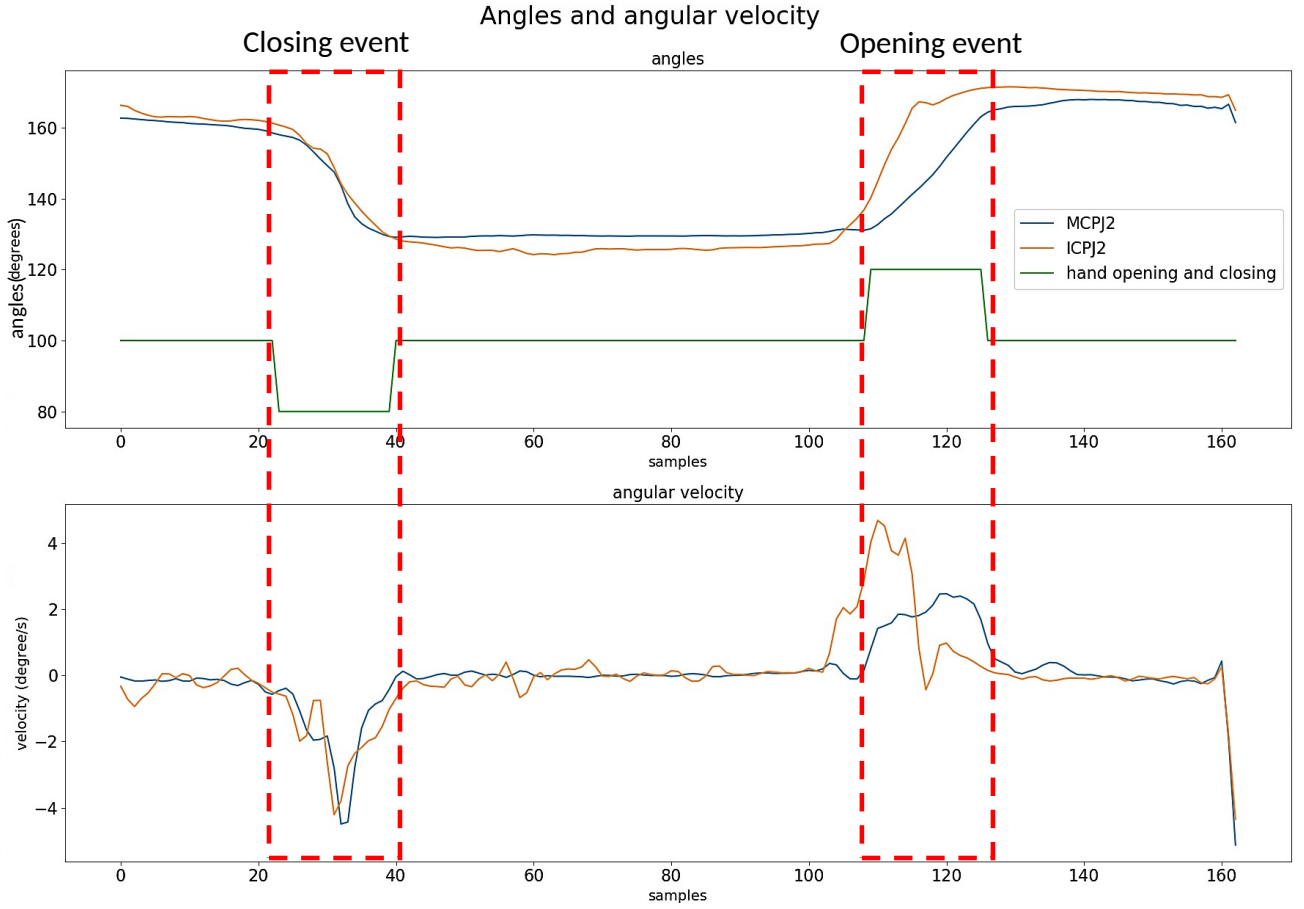


Figure 2.7: Example of identification of closing and opening events from MCPJ2 and IPJ2 angular velocity. Additionally, any points in the ‘neutral’ (hand fully open) state were also labelled as ‘opening’, while samples in the fully closed state were labelled as ‘closing’.

2.3.3 Electromyography processing

The EMG signals were segmented based on the trigger signal, filtered with a 20-500 Hz 5th order Butterworth band-pass and 3rd order band-stops at 50, 100, 150 Hz, and rectified, with noisy or disconnected channels set to 0. No channel-wise normalisation was performed, in order to avoid amplification of channels with low signal-to-noise ratio, which could negatively influence the prediction models. The EMG processing and feature extraction was performed on MATLAB (MATLAB 2019a, Mathworks). The first feature extracted from the EMG signals is the RMS for each channel, representing the amplitude envelope and a reduction in the size of the data. EMG amplitude envelope was selected as it is a robust, commonly used signal feature in myocontrol, which is linearly related

to force [18]. The envelope was extracted from the filtered EMG by estimates of the RMS values in non-overlapping intervals of 62.5ms. After this feature extraction, the data was split into 5 trials per recording, based on the splitting information from the angle data.

Feature Extraction

Since we have access to thorough HDsEMG recordings, preliminary research was done to explore and investigate the effect of using different spatial features, specifically the variogram and 2D-cross-correlation, on the prediction models. This is described in Appendix A and led to the conclusion that the RMS feature is the most suitable for our application.

2.3.4 Kinematic Prediction Models

The extracted features and kinematics were used to train regression and classification models to predict the kinematics of wrist movements and opening and closing of the hand, respectively. Certain parameters were determined in preliminary findings. An overview of the Kinematic Prediction Models is shown in Figure 2.8.

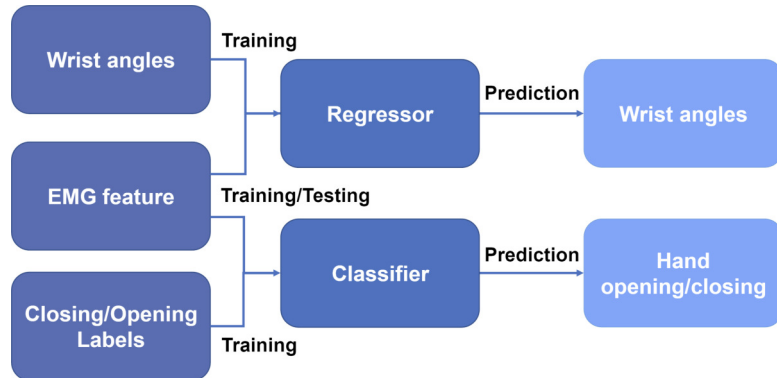


Figure 2.8: Overview of Kinematic Prediction Models training and testing.

Regression for wrist angles

A linear regressor [116] was trained to predict the wrist flexion-extension and pronation-supination angles. One model was created per subject, feature, angle, and training data set. Four trials from each movement in the training set were used for training, and one trial of each movement in the testing set was used for testing in 5-fold cross validation. The training movement trials were randomized

and shuffled. Predicted angles were then low-pass filtered (4th order Butterworth, 2 Hz). To assess the validity of the models, r^2 values for each angle each run of the cross validation were calculated according to [117] where r is the Pearson Correlation coefficient, defined as

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 (y_i - \bar{y})^2}} \quad (2.1)$$

where $\bar{x}$ is the mean of the vector x and $\bar{y}$ is the mean of the vector y , the true angles and the predicted angles, respectively, and i is the index of observation. r is the Pearson Correlation coefficient, which is then squared, such that $0 < r^2 < 1$.

The normalised RMSE is also calculated for each angle in each run of the cross validation according to [116]. The RMSE was normalised over the range of the testing data for that angle. If y_i is the predicted value of the i^{th} sample, and x_i is the corresponding true value, then the RMSE estimated over n samples is defined as

$$RMSE(y, \hat{y}) = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2} \quad (2.2)$$

Classification of grasp intention

An SVM classifier with a Radial Basis Function kernel [116] was trained to predict the classification labels from the features. One model was created per subject, feature, angle, and training data set. Four trials from each movement in the training set were used and one trial of each movement in the testing set was used for testing in 5-fold cross validation. The training trials were randomised and shuffled. Classification smoothing was applied to the predicted classes, such that the most repeated class within a moving window of 3 samples was chosen as the class for the central sample. To assess the validity of the models, confusion matrices were created from the predictions of all the runs and normalised to show the recall and precision.

2.3.5 Training set evaluation

The prediction models were trained as either a 4-DoF or a 3-DoF systems. In the 4-DoF system all the movements listed are used, whereas in the 3-DoF system any movements involving radial and ulnar deviation are not used. This is to evaluate the performance of the prediction models and training sets in an ideal scenario, as well as in a more practical scenario, conforming to our hardware limitations. The kinematic prediction models were trained on different movement sets (Table 2.1) to investigate the influence of the training set on performance. Each movement set includes four separate recordings, except for Set 2 which only includes one. This is because only the general grasp was used in these investigations. Each recording lasted approximately 45s and included 5 repetitions (4 of which were used for training for each fold of the validation). The regression training sets were: the individual wrist and hand sets (5 movements in 3-DoF system and 7 movements in 4-DoF system); individual wrist and sequential wrist and hand movements, and individual wrist and simultaneous wrist and hand movements (8 movements each in 3-DoF system and 12 movements each in 4-DoF system); and a set including all the recorded possible movements relating to the wrist (12 movements in the 3-DoF system or 16 movements in the 4-DoF system). For classification of the hand movements the following sets were used: individual hand movements (1 movement); individual hand and sequential wrist and hand movements, and individual hand and simultaneous wrist and hand movements (5 movements in 3-DoF system and 7 movements in 4-DoF system); and a set including all the recorded possible movements relating to the hand (9 movements in 3-DoF system and 13 movements in 4-DoF system).

2.3.6 Statistical methods

Statistical analysis was performed on the r^2 and RMSE scores for the regression results in the training set evaluation for the 3-DoF system. The datasets of the r^2 and RMSE scores do not follow a normal distribution (Shapiro-Wilk test with $p < 0.05$) and presented non-homogeneous variances for the r^2 results (Levene's test with $p < 0.05$, except for the 'individual wrist' testing set), but homogeneous variances for the RMSE results (Levene's test with $p > 0.05$, except for the 'Simultaneous wrist and hand' testing set). As such, non-parametric Kruskal-Wallis H tests [117] were used, where the factor

is the training movements set. This was followed by Mann-Whitney U-tests for post hoc multiple comparison [117], corrected with a Bonferroni correction. The level of significant statistical difference was set as 0.05 ($p < 0.05$).

2.4 Results

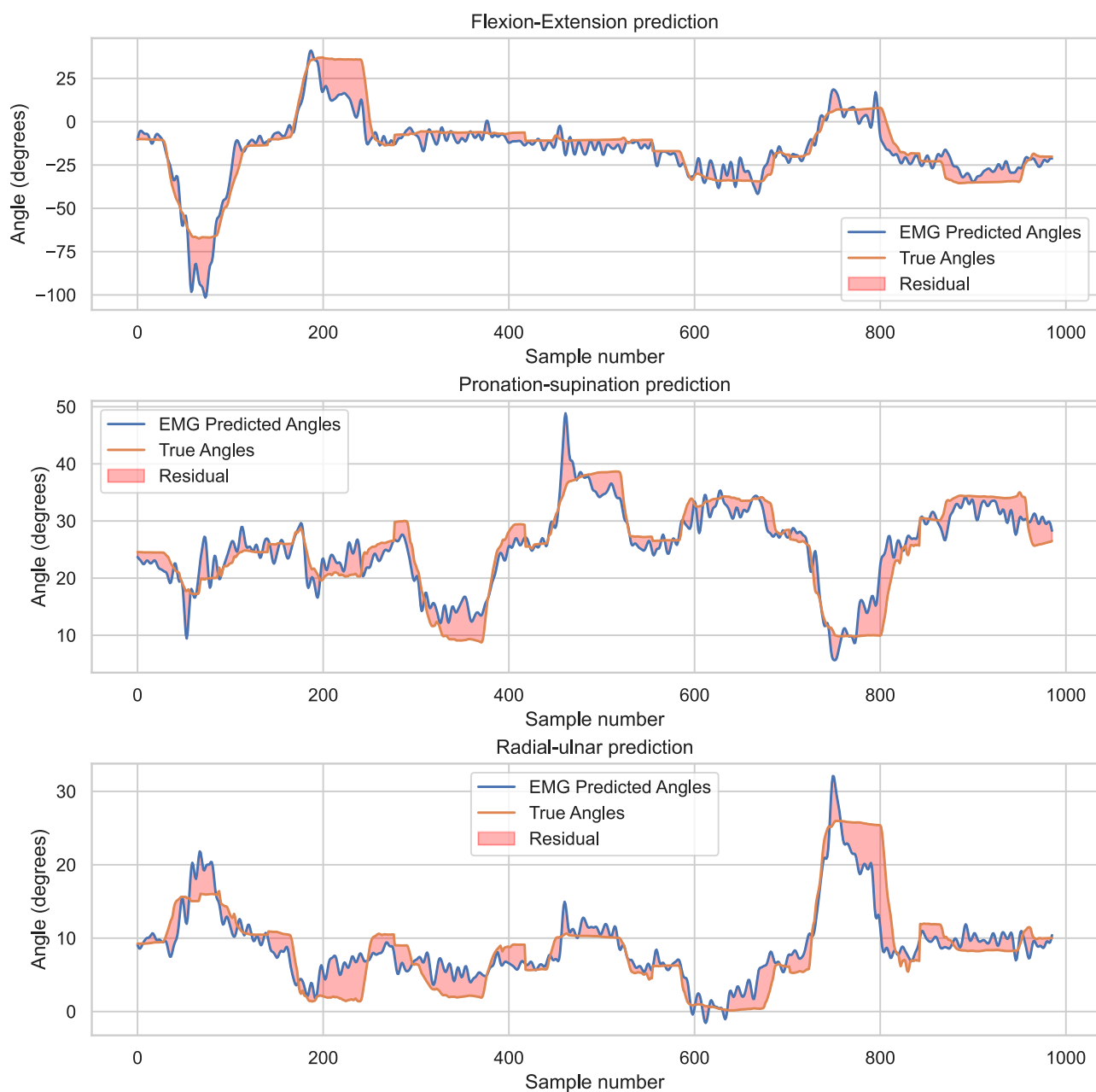
We were able to predict the movements of the wrist and hand using regression and classification, respectively.

2.4.1 Regression

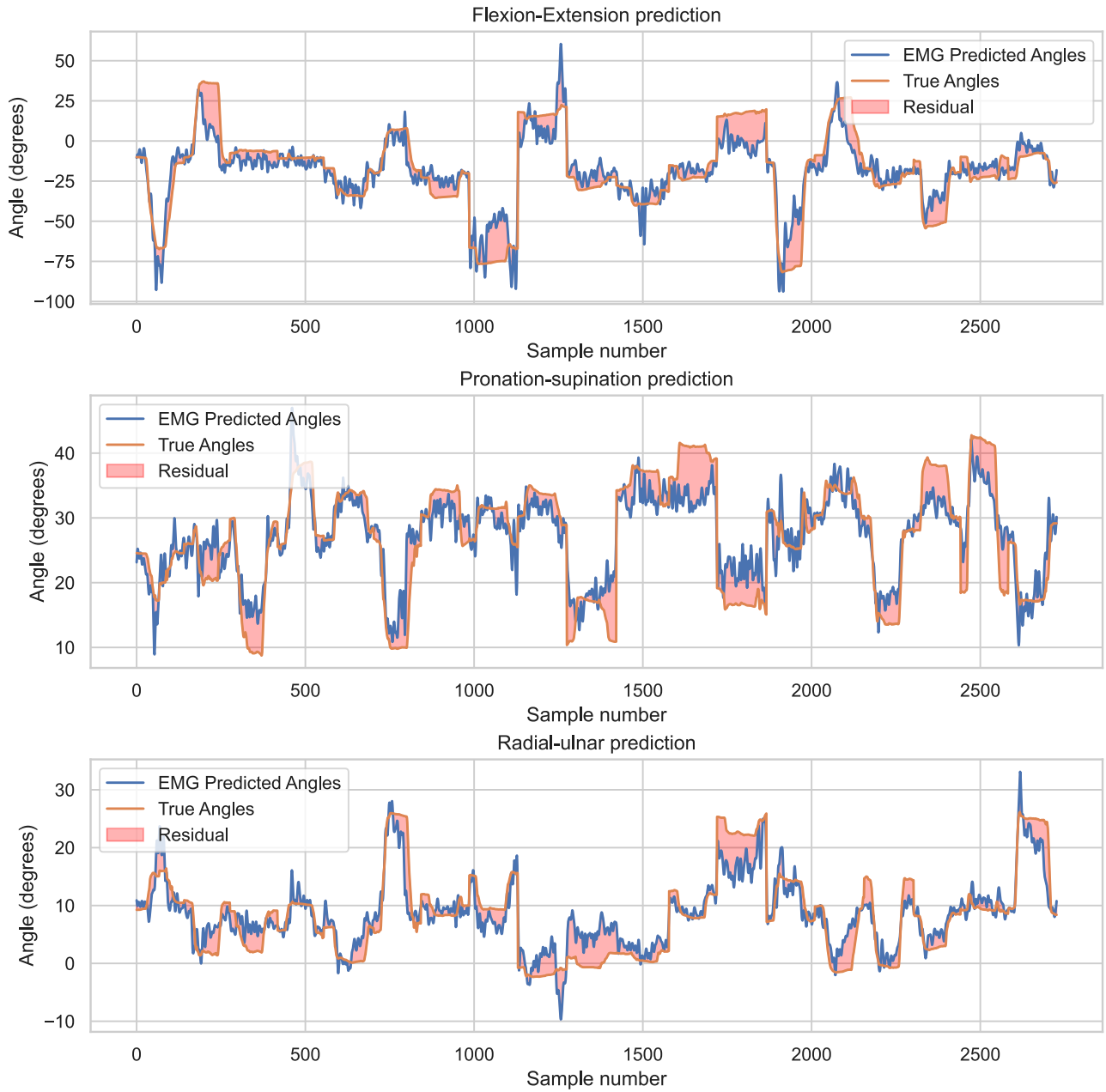
Figure 2.8 shows some examples of the angle data predicted by the linear regressor compared with the true data. The residual shows the difference between the two sets of angles. Figure 2.9 shows the overall performance of the linear regressor when training and testing the system on the same movement set(s), broken down by angle. The performance is most robust for the flexion-extension and radial-ulnar deviation angles, while still providing reasonable performance for the pronation-supination, and MCPJ2 angles.

Next we move on to an evaluation of using different combinations of movement sets to train the regressor, when testing the models on different movements sets involving 4 DoFs (3 DoFs of wrist and 1 DoF of hand). Figure 2.10 shows the r^2 and RMSE scores for the systematic evaluation of using different combinations of movement sets to train the regressor, when testing the models on different movements sets in the 4-DoF system. As expected, training with only the individual wrist and individual hand movements provides the least robust performance for movement combinations between the wrist and hand movements (median $r^2 = 0.43$, median $RMSE = 0.17$ for overall performance). Training with individual wrist movements plus either sequential (median $r^2 = 0.55$, median $RMSE = 0.15$) or simultaneous (median $r^2 = 0.54$, median $RMSE = 0.15$) data yield similar results, but the performance is not as effective as when training with all movement sets available (except for simultaneous wrist) (median $r^2 = 0.59$, median $RMSE = 0.14$).

Figure 2.11a shows the r^2 scores and Figure 2.11b shows the RMSE scores for the systematic



(a) Predictions for a regression model trained and tested on movement sets 1 and 2.



(b) Predictions for models trained and tested on movement sets 1, 2, 3, and 4.

Figure 2.8: Examples of the angle data predicted by the linear regressor compared with the true data.

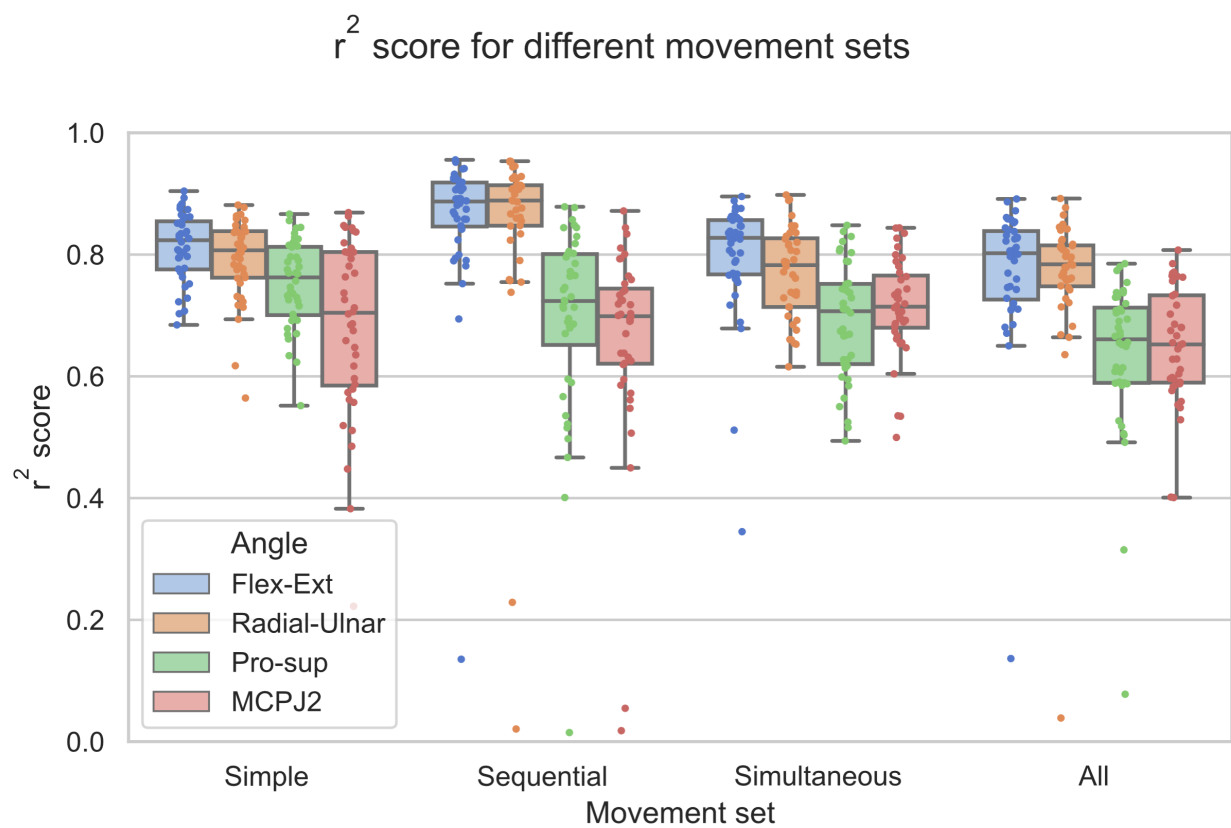
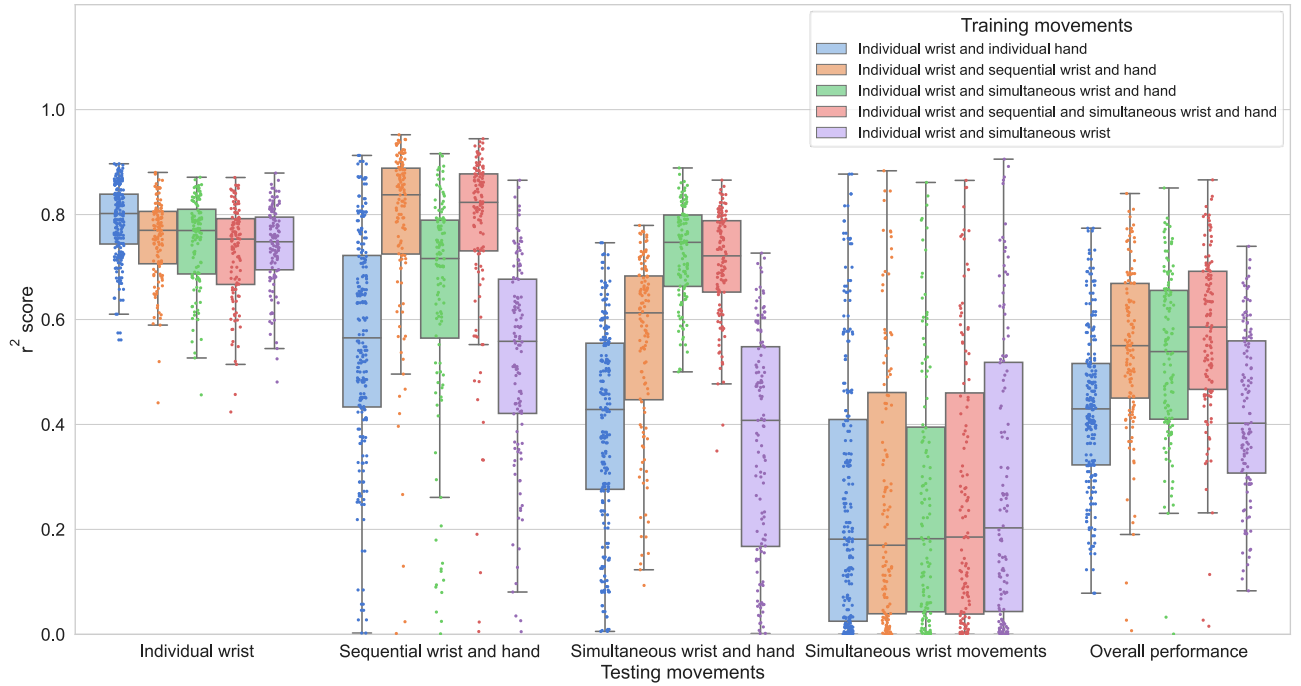
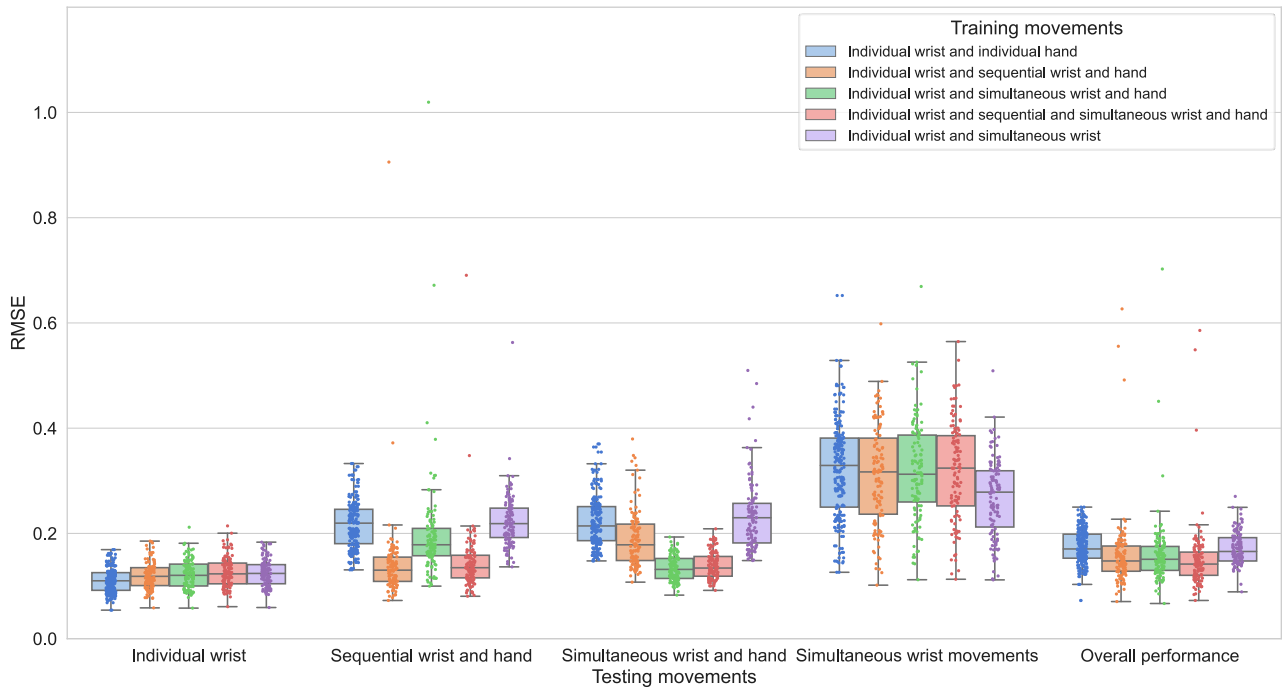


Figure 2.9: r^2 scores for flexion-extension, pronation-supination, and radial and ulnar deviation, and MCPJ2 angles when different movement sets are used to train and test a linear regressor. Here ‘simple’ refers to movement sets 1 and 2, ‘sequential’ refers to movement set 3, ‘simultaneous’ refers to movement set 4, and ‘all’ refers to all the previously mentioned movement sets.



(a) r^2 scores



(b) Normalised RMSE scores

Figure 2.10: 4-DoF system showing a) r^2 and b) normalised RMSE scores for flexion-extension, pronation-supination, and radial and ulnar deviation wrist angles when different movement sets are used to train a linear regressor, tested on individual wrist movements, sequential wrist and hand movements, and simultaneous wrist and hand movements. Jitter plot shows individual results for each cross-validation of each DoF for each of 8 subjects.

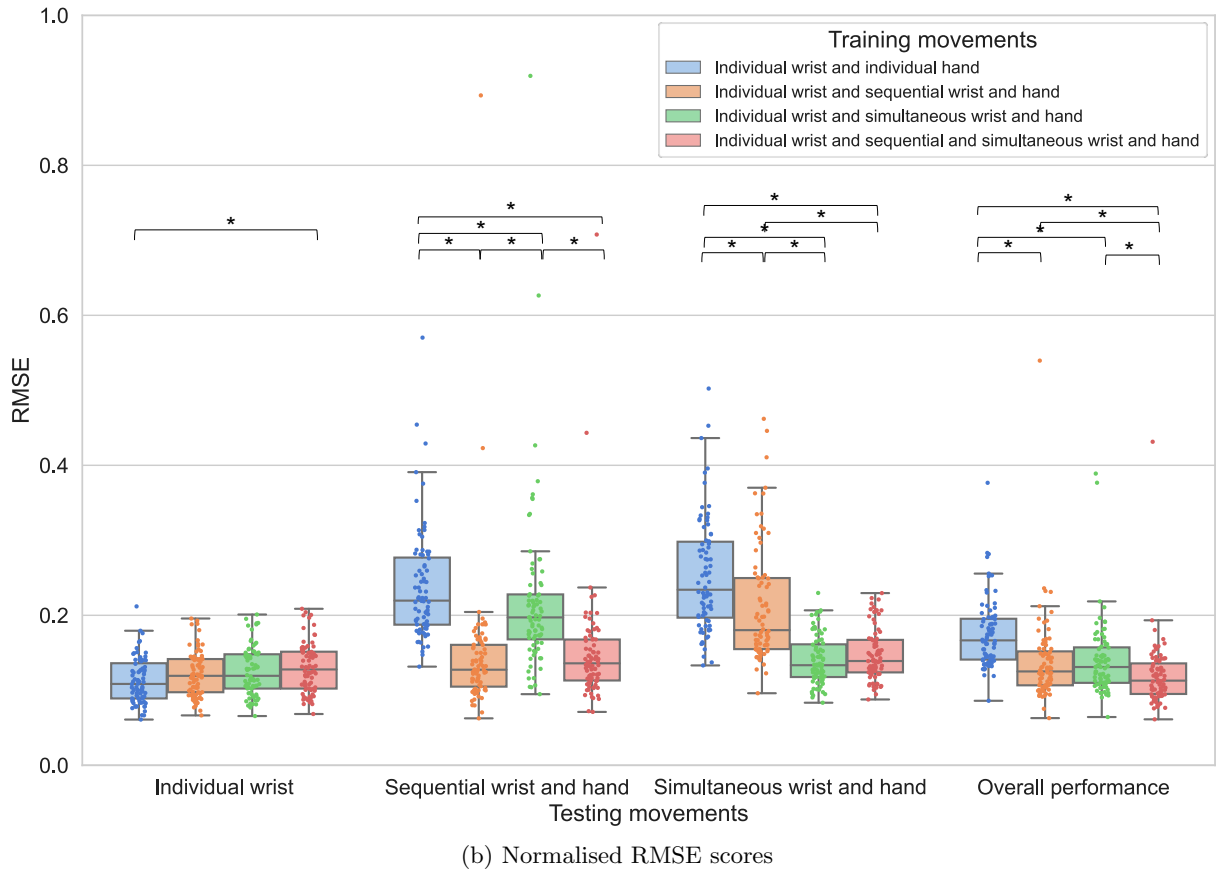
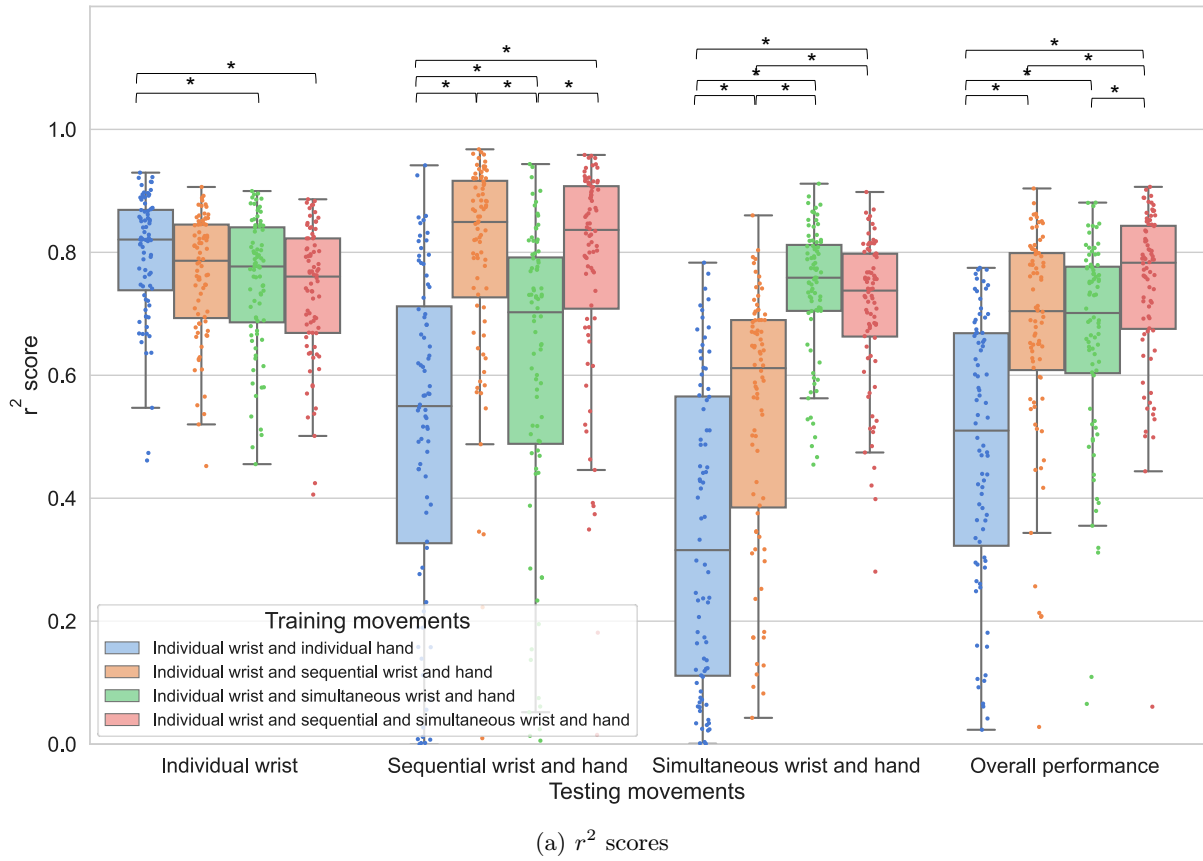


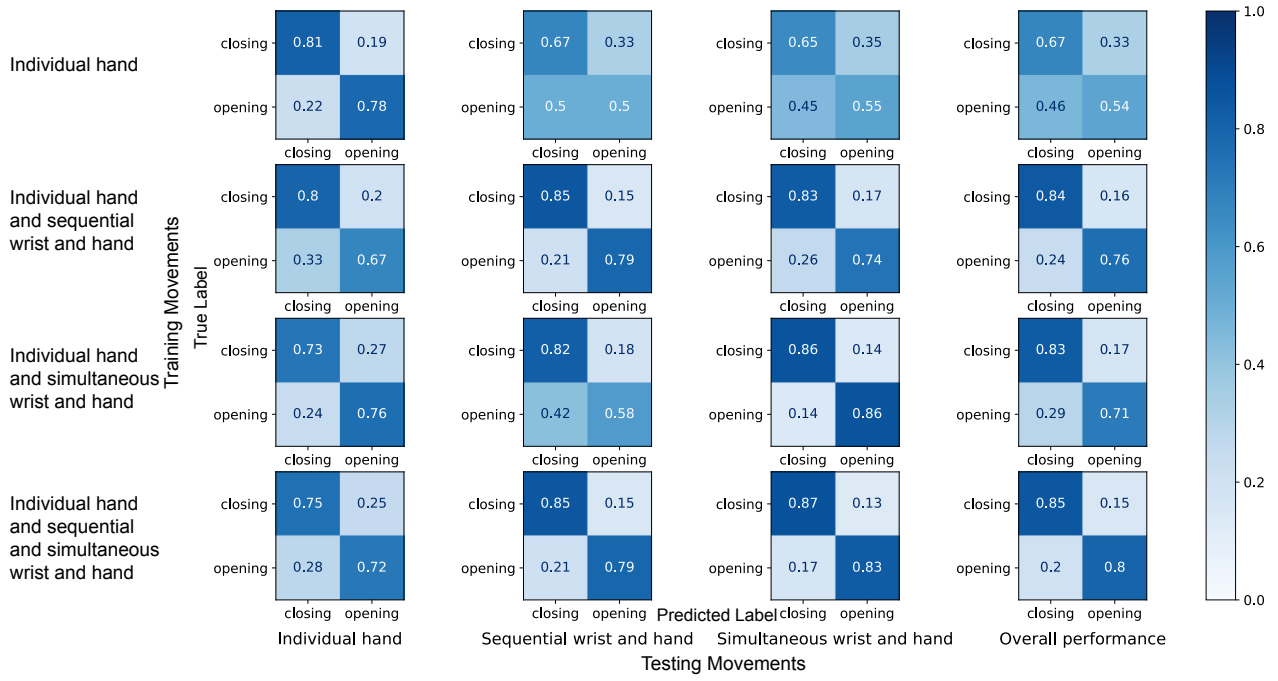
Figure 2.11: 3-DoF system showing a) r^2 and b) RMSE scores for flexion-extension and pronation-supination wrist angles when different movement sets are used to train a linear regressor, tested on individual wrist movements, sequential wrist and hand movements, and simultaneous wrist and hand movements. Jitter plot shows individual results for each cross-validation of each DoF for each of 8 subjects. * shows statistically significant difference with $p < 0.05$.

evaluation of using different combinations of movement sets to train the regressor, when testing the models on different movements sets in the 3-DoF system. The Kruskal-Wallis H tests reveal significant differences between training movement sets for each of the testing sets. The post hoc tests show statistically significant differences between most of the training sets for all testing sets, except for a few cases. As expected, training with only the individual wrist and individual hand movements provides the least robust performance for movement combinations (median $r^2 = 0.51$, median $RMSE = 0.17$, for overall performance). Training with individual wrist and hand movements plus either sequential (median $r^2 = 0.70$, median $RMSE = 0.13$) or simultaneous (median $r^2 = 0.70$, median $RMSE = 0.13$) data yield statistically similar results (for r^2 at least) but the performance is not as effective as when training with all movement sets available (median $r^2 = 0.78$, median $RMSE = 0.11$).

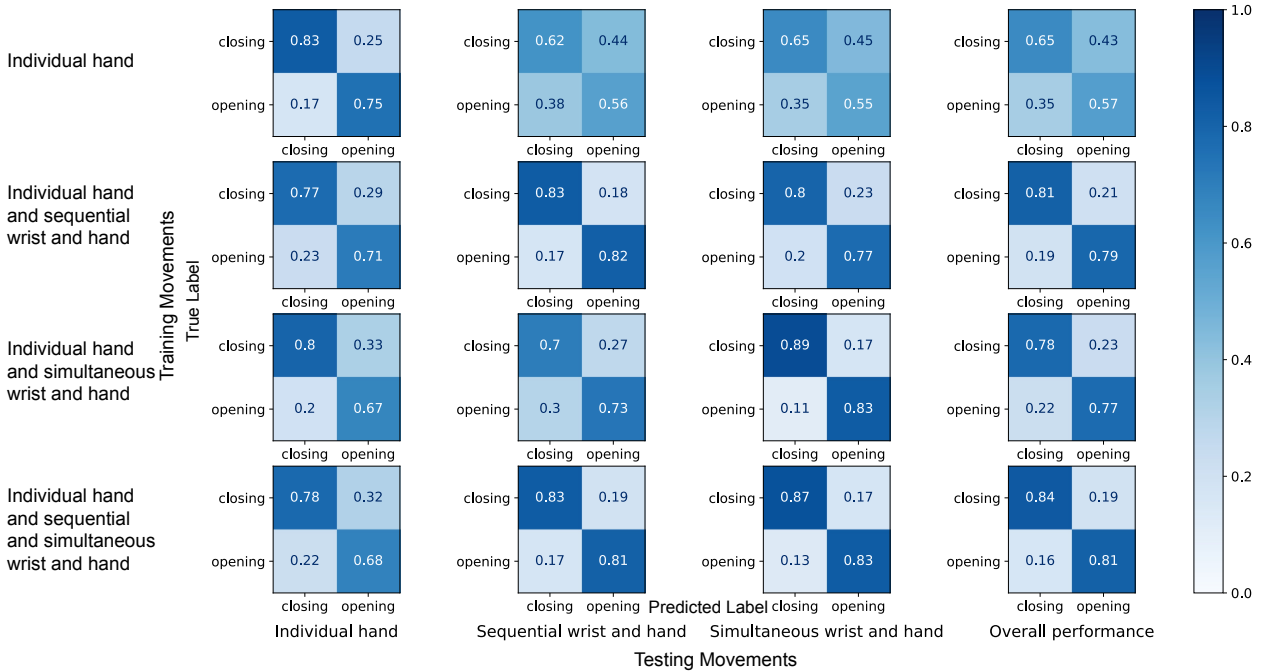
2.4.2 Classification

Figure 2.12 shows the normalised confusion matrices for cross validation classification using different training and testing movements in the 4-DoF system. Training with individual hand movements does not allow for accurate predictions of hand intention when involving movements of the wrist (overall performance ‘closing’ recall: 0.67; ‘opening’ recall: 0.54; ‘closing’ precision: 0.65; ‘opening’ precision: 0.57). However, training with individual hand movements and either sequential or simultaneous hand movements, or both, yield similar robust results when it comes to overall performance (overall performance ‘closing’ recalls: 0.84, 0.83, 0.85; ‘opening’ recalls: 0.76, 0.71, 0.8, respectively; ‘closing’ precision: 0.81, 0.78, 0.84; ‘opening’ precision: 0.79, 0.77, 0.81). The overall classification accuracy for each training and testing set are listed in Table 2.2.

Figure 2.13 shows the normalised confusion matrices for cross validation classification using different training and testing movements in the 3-DoF system. Training with individual hand movements does not allow for accurate predictions of hand intention when involving movements of the wrist (overall performance ‘closing’ recall: 0.67; ‘opening’ recall: 0.54; ‘closing’ precision: 0.65; ‘opening’ precision: 0.56). However, training with individual hand movements and either sequential or simultaneous hand movements, or both, yield similar robust results when it comes to overall performance



(a) Recall



(b) Precision

Figure 2.12: Normalised confusion matrices for the 4-DoF system. a) recall and b) precision for classes predicted when different movement sets are used to train an SVM classifier regressor, tested on individual wrist movements, sequential wrist and hand movements, and simultaneous wrist and hand movements, for 8 subjects.

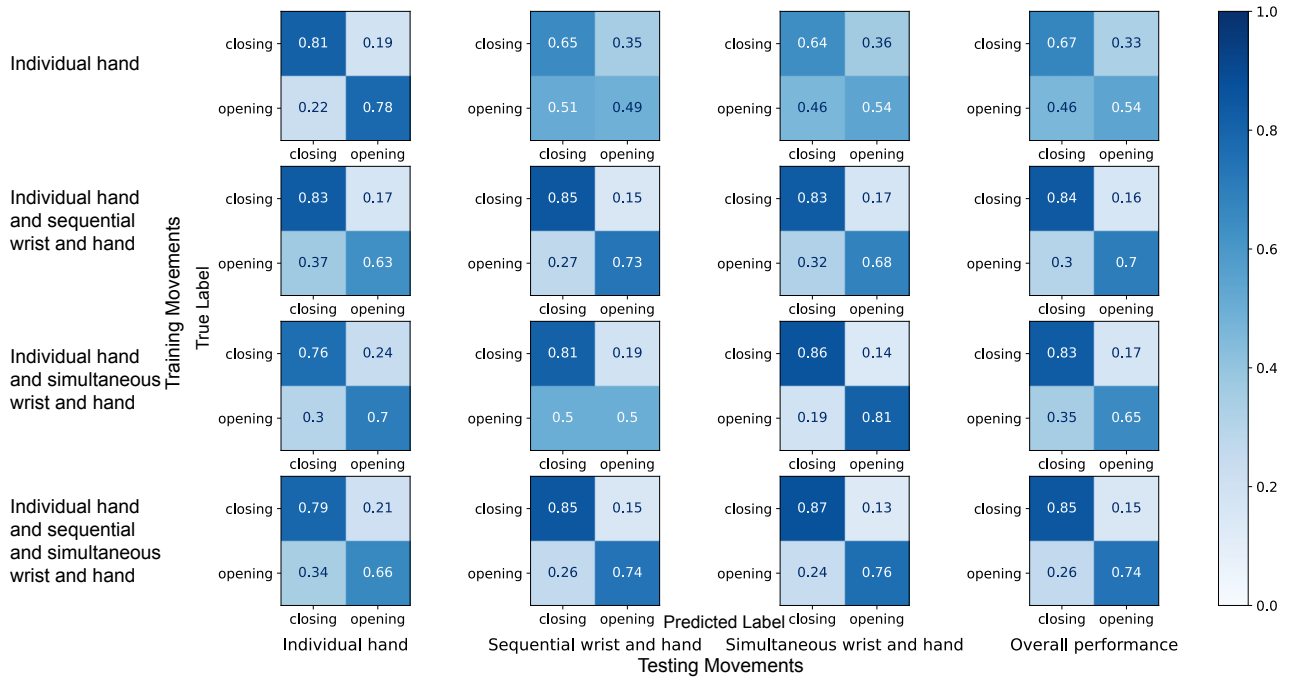
Table 2.2: Classification of closing and opening accuracy of the 4-DoF system.

		Testing Sets			
	Accuracy (%)	Individual hand	Sequential wrist and hand	Simultaneous wrist and hand	Overall performance
Training Sets	Individual hand	79.74	59.36	60.82	61.54
	Individual hand and sequential wrist and hand	74.35	82.46	78.78	80.24
	Individual hand and simultaneous wrist and hand	73.98	71.07	85.97	77.85
	Individual hand and sequential and simultaneous wrist and hand	73.32	82.26	85.29	82.94

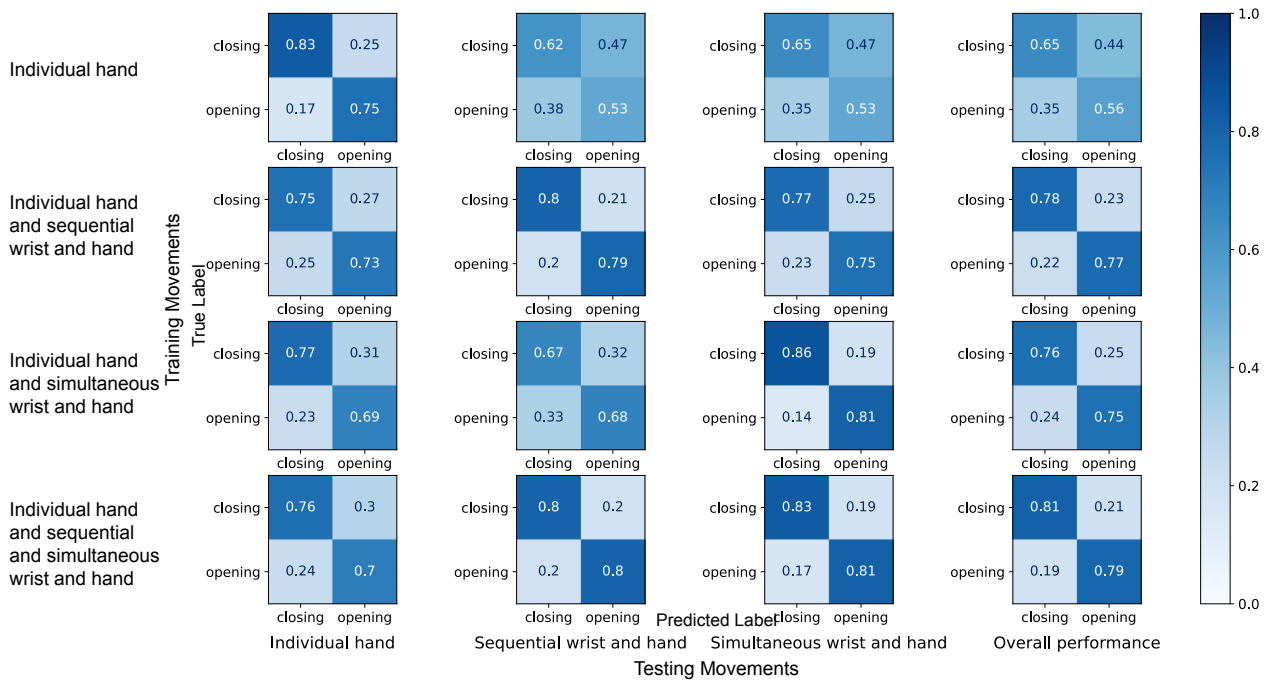
(overall performance ‘closing’ recalls: 0.84, 0.83, 0.85; ‘opening’ recalls: 0.7, 0.65, 0.74, respectively; ‘closing’ precision: 0.78, 0.76, 0.81; ‘opening’ precision: 0.77, 0.75, 0.79). The overall classification accuracy for each training and testing set are listed in Table 2.3.

Table 2.3: Classification of closing and opening accuracy of the 3-DoF system.

		Testing Sets			
	Accuracy (%)	Individual hand	Sequential wrist and hand	Simultaneous wrist and hand	Overall performance
Training Sets	Individual hand	79.74	58.02	59.83	61.18
	Individual hand and sequential wrist and hand	74.33	79.60	76.38	77.64
	Individual hand and simultaneous wrist and hand	73.58	67.61	83.58	75.15
	Individual hand and sequential and simultaneous wrist and hand	73.34	79.99	82.27	80.25



(a) Recall



(b) Precision

Figure 2.13: Normalised confusion matrices for the 3-DoF system. a) recall b) precision for classes predicted when different movement sets are used to train an SVM classifier regressor, tested on individual wrist movements, sequential wrist and hand movements, and simultaneous wrist and hand movements, for 8 subjects.

2.5 Discussion

With a complete training set of both individual and combined movements, we robustly inferred the motions of the wrist as a combination of flexion-extension and supination-pronation for a 2-DoF wrist controller, regardless of confounding factors such as the hand movements. At the same time, we robustly distinguished between two classes of hand movements - hand closing and opening - regardless of the wrist position and movement. The 3-DoF system is thus able to control a total of three DoFs independently and simultaneously using a combination of two models, regression and classification. The results show that the system performance decreases when using only individual movements to train the regression and classification systems with respect to training with all the individual, sequential, and simultaneous movements. The difference in performance between training with sequential and with simultaneous movements was relatively small (0.08 for regression r^2 ; 0.01 and 0.04 for classification closing/opening recall), such that removing the simultaneous movements from the complete training set may be justified in terms of the benefits it offers. Replacing simultaneous movements with sequential movements offered a similar performance, accounting for a wider range of movement combinations, without losing performance in control over the hand closing and opening. These results confirm that it is not needed to train simultaneous movements in order to predict them, due to the linearity of the system in muscle space [32]. This has a practical impact - shorter training time is essential, as is including movements that people with limb loss are comfortable performing repetitively, for the uptake and intuitive use of prostheses. The regression and classification features and techniques selected are also of low computational complexity, such that a model can be trained and used in a real-time setting.

With the 4-DoF system we are also able to infer the motions of the wrist as a combination of flexion-extension, supination-pronation, and radial-ulnar deviation, regardless of the hand movements. We are also able to distinguish between hand closing and opening regardless of the wrist position and movement. The results do show potential for extending the system to a greater range of motion; however, the performance is not as successful or robust as with the 3-DoF regression models, even though the performance is similar for classifying hand closing/opening. As such the 3-DoF Kinematic

Prediction Models will feed into the development of an online, real-time system, with only individual and sequential movement tasks used to train the regression and classification models (see Section 4.1.2).

While the results are comparable to other research, the models can always be improved. The way in which the data is segmented into trials and the separate trials are rearranged could have led to reduced performance as the movements are no longer continuous. We could also use regression to control the hand opening/closing instead of classification. In Figure 2.9 we can see that we are also able to predict the MCPJ2 angles. In that case we could apply the thresholding post regression to use as a switch to activate the CV system.

While there are some other studies mapping EMG to kinematics, rather than to force, very few record both HDsEMG and Motion capture concurrently for dynamic movements of the hand and wrist, [33, 106] and this was not very common prior to the start of this PhD. The data set we have collated is extensive and can be used to explore and investigate alternative EMG control mechanisms. The number of channels used can be explored, as well as the features, and the models. It would be interesting to extend the research in this thesis to the full dataset collected, including the different grasp types, and exploring even more training set reductions. The setup and protocol have also formed a basis for collaborations and further experimentation in our research group. The dataset and the setup and protocol can also be expanded to explore other sensors that can be used for human intention detection, and have led to further research with Ultrasound [118].

The study in this chapter was performed on able-bodied subjects, and an extension to individuals with upper limb loss would require mirror movements for training [33], due to lack of kinematics of the target limb. Though the offline system was not tested with people with upper limb loss, it is useful to determine a baseline with precise kinematics, that can then be tested in a real-time system with both able-bodied individuals as well as those with limb-loss using estimated kinematics, without motion capture.

The methods developed in this chapter underpin the control for the complete integrated shared prosthetic controller. They allow the user to control the wrist while also initiating the application of the grasp type chosen by the Computer Vision Assistance system through closing the hand. The

hardware involved is not yet clinically transferable, but research and development of more wearable HDsEMG sensors is in process, and future work into our research should incorporate these. For now, the HDsEMG, along with our Kinematic Prediction Models, should allow for a robust control with individuals with limb loss in a real-time scenario.

In order to further validate the Human Intention Detection system for use by individuals with upper limb loss, the 3-DoF Kinematic Prediction Models were incorporated into a real-time system, as described in Section 4.1.2.

Chapter 3

Computer Vision Assistance

In this chapter we propose a novel CV approach to determine the grasp type of a given object, using contour-based recognition [119] as an intermediate step between the object and the grasp prediction. Our method uses a CNN to extract global shape features from an RGB image of the object to be then used for grasp classification. The algorithm is trained and validated offline before being adapted for use in a continuous, real-time, system that would translate well into a shared control strategy with a myoelectric prosthesis. The framework is validated in a study with human participants to establish its parameters and validate its use, during which a dataset of human reaching and grasping videos is also collated for future use to further improve this algorithm and similar ones in this area.

3.1 Introduction

As lower arm prosthetic design has developed, the dexterity of the prostheses has improved, allowing users to choose between a range of grasp types. For example, the TASKA prosthetic hand has a total of 20 grasp type configurations, as well as the option to define further grasp types (Taska Prosthetics, New Zealand) [60]. The increased complexity in prosthetic design is reflected by a complexity in control. One main challenge is making the user's control of their prosthetic arm as intuitive as possible. However, switching between grasping configurations of prosthetic hands is most commonly performed either by pressing a button or by co-contracting the muscles [60, 120, 121]. As described in Section 1.2.4, recently, to address this problem, CV techniques used in automated robotics are

being translated into usage for object recognition and grasp type identification in prosthetics [54, 70–72, 75, 77, 79, 81], since the physiological process of grasping an object relies heavily on sensory information from a number of sources, including vision.

Vision sensors are mostly applied in robotics to aid in reaching and grasping tasks. In robotics, vision sensors can provide global information, such as the location of objects in the environment, to aid in visually guided reaching control and visual serving [122]. Vision sensors can also be used to perform object detection and recognition [122]. This is not necessarily applied for use in grasp recognition, but instead to determine successful grasping locations. Most of the work within automated robotics is targeted at robotic end-effectors, usually of 2-3 ‘fingers’; thus they focus on grasping locations as opposed to targeting human ‘hand’-like end-effectors. The problem addressed in robotic grasping is thus where to grasp the object as well as how to perform the reaching task, whereas our question is how the object is going to be picked up, focusing solely on the grasping task. The user can choose the best location to place the hand prosthesis when grasping an item. Thus, any grasp recognition algorithm used should be specific to our use case within the prosthetic hand.

Some studies have proposed CNNs for object recognition within prosthetic use and a one-to-one correspondence with grasp types [75, 76, 79]. A more recent approach consists of classifying objects according to grasp patterns without explicitly identifying them or measuring their dimensions [77]. The main difficulty faced by the current methods is grasping unseen objects. When the CNN encounters a novel object, the grasp success rate falls as it was not present in the training phases [77]. In addition, although some of the previously proposed CV prosthetics use EMG control, they mostly use them for taking a picture [77] or just closing/opening the hand [70, 71, 79, 81], with minimal wrist control [54]. We want to free up the EMG to be used for more general wrist and hand control. This also means the CV systems are not run continuously, which has both positive and negative implications that should be considered.

3.1.1 Camera selection and placement

Multiple sensors can make use of spatial distribution and coordination to provide complex global information about both the robot end-effector and its workspace, such as creating 3D laser range point clouds of objects and scaling and transforming objects [122]. However, fixed external sensors are not a suitable possibility for this research due to the flexible requirements of moving around with the user. As demonstrated in Figure 3.1, the configuration of having the camera located or mounted on a fixed place external to the end-effector is known as eye-to-hand configuration [122]. On the other hand, visual feedback from a camera is nominated as eye-in-hand if the camera is positioned on the end-effector of the robot's arm [122]. An eye-in-hand configuration provides a constant relation between the position of both camera and the end-effector.

One option for use in our system could be placing the camera on an accessory such as glasses or a hat, such as in [54] and [72], but we chose instead to place the camera on the hand itself, for an eye-in-hand point of view. This is for two main reasons, the first being that we aim to develop a fully integrated system without additional accessories that might be cumbersome or hinder the user, the second being that placing the camera on the end-effector allows for a more direct point of view during the reaching movement [62], providing a more intuitive positioning than those proposed by [66, 72]. This also avoids the occlusion that may occur with static sensors, where objects may be out of the field of view [65], potentially hidden by the prosthesis itself. More specifically, we chose to place the camera on the palmar side of the wrist, to provide a more direct view of the object upon approach during reaching, as opposed to the dorsal side of the hand (as in [77]), which would require an unintuitive approach to the object. However, our chosen positioning provides its own challenges that must be overcome, including noise, instability due to movement, and the proximity to the object which could cause the camera to be out of focus. Thus, certain parameters for any approach we chose would require establishment in an experimental setting, in order to be integrated with a prosthetic hand. Based on the eye-in-hand placement and the integration with a prosthetic hand, size is another constraint for the sensor.

Although our system could benefit from the integration of multiple sensors such as an IMU or a

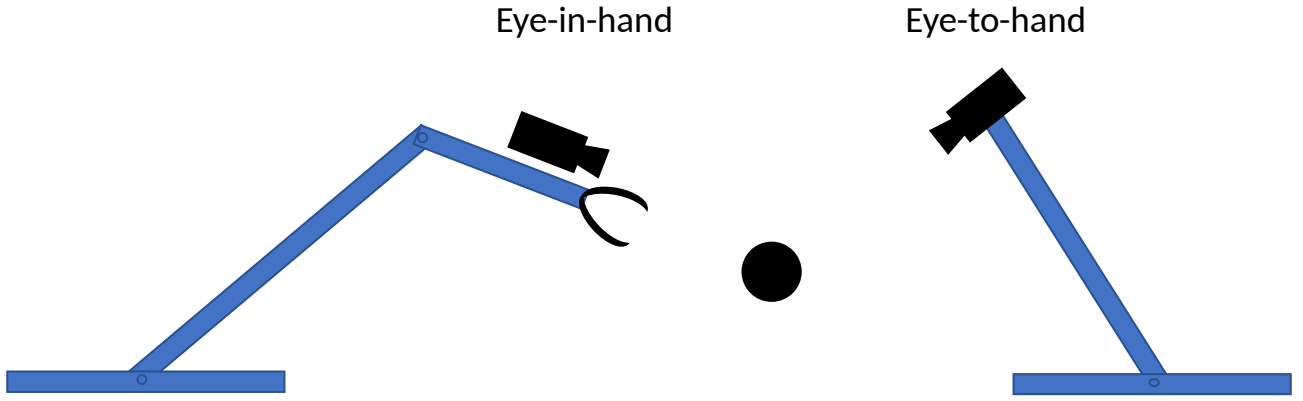


Figure 3.1: Eye-in-hand vs eye-to-hand robotic vision configurations.

depth sensor alongside the camera as in other research [54, 70–74, 83], we choose to limit our research to one sensor type. This simplifies the system, reducing the processing complexity and simplifying the integration. Our research provides a starting point, from whereon out we can investigate whether additional sensors are required or how they might contribute to the shared control. Considering all the other more complex research systems out there, we aim to investigate how successful we could be with a smaller sensor suite, while also focusing on the shared control between the automated system and the user control. Although we lose out on the advantages associated with insight into the movement of the arm as a whole as well as distance to the object to be grasped, we investigate whether we can compensate for these via calibration of a more advanced computer vision algorithm alongside the more intuitive EMG control.

3.1.2 Grasp type classification

Several attempts have been made to classify the range of human grasp types into a taxonomy. While Feix [57] focuses on grasp types in human movements and manipulations, Cutkosky [56] attempts to put this in a frame of reference of transforming into robotic grasping and end-effector manipulators. The grasp recognition model used in our Computer Vision Assistance system is based on an adaptation of the grasp taxonomy developed by [56], as shown in Figure 3.2, where the grasp type is determined according to the precision of the task and the size of the object. We chose to focus on 4 different grasp types based on some of the most commonly used grasp types [123–126] available with the TASKA prosthetic hand. Thus, the system should classify images into one of the following 4 grasp types:

General (power grasp), Tripod (3 finger pinch), Pinch (2 finger pinch), Spherical.

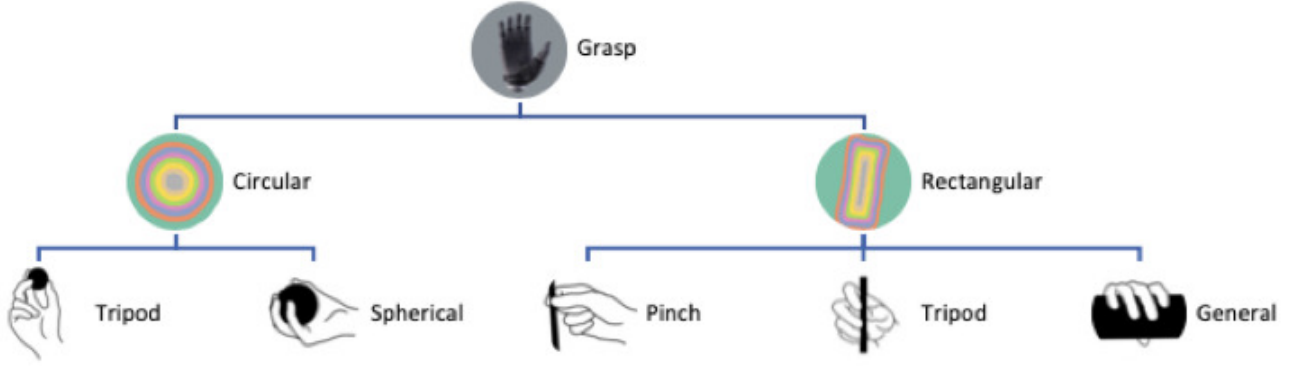


Figure 3.2: Adaptation of the Cutkosky grasp taxonomy for the Taska hand.

3.1.3 Aims

Our aim is to develop and validate a proof-of-concept Computer Vision Assistance system, which can be further refined at a later stage. The current literature and the context in which we are designing the Computer Vision Assistance system lead to the following system specifications and aims:

- Develop an algorithm and system to determine grasp type of object,
- The grasp recognition algorithm should be adaptable to novel objects, and thus should not depend on the object recognition itself,
- The algorithm should classify images into the following 4 grasp types: ‘General’, ‘Tripod’, ‘Pinch’, and ‘Spherical’,
- The algorithm should be trained and validated in an offline setting,
- The algorithm should be suitable for use with an eye-in-hand camera,
- The camera should be suitably integrated with a prosthetic hand,
- The algorithm should be suitable for use in a continuous, real-time, system, further referred to as the Computer Vision - Real-time System (CV-RS) that would adapt well into a shared control strategy with a myoelectric prosthetic hand, for continuous myoelectric wrist control as well as opening and closing the hand,

- The CV-RS should be validated with suitable use cases,
- Parameters for the integration of the CV-RS with a prosthetic hand should be established in a real-time study,
- A suitable dataset for human reaching and grasping tasks should be collated for future use to further improve this system and similar ones in the field.

3.2 Offline algorithm development and training

Grasping¹ unseen objects is the main challenge faced by current CV-based algorithms used in conjunction with prosthetic devices. When these algorithms, mainly CNNs, encounter a novel object, the grasp success rate falls since the novel object was not present in the training phase [77]. To address this, and ensure our algorithm generalises to unseen objects, we focus on features surrounding the shape of the visible part of the object rather than linking the object itself with a specific grasp type. We propose a novel CV algorithm approach using contour-based [119] recognition which creates an intermediate step between the object and the grasping classification, as shown in Figure 3.3. Our algorithm uses a CNN to extract global shape related features from an RGB image of the object. These features are used to classify the initial images into one of 4 grasp types via a K-means clustering model for grasp classification.

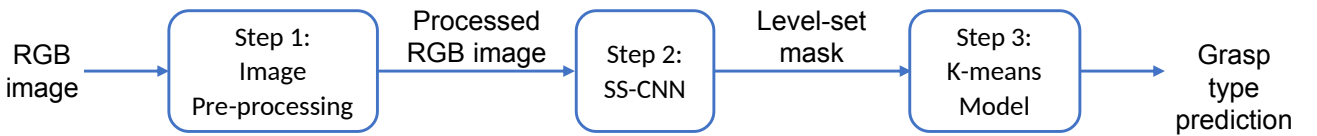


Figure 3.3: Overview of CV offline algorithm.

3.2.1 Image dataset

Training a neural network requires a large amount of annotated data. Several ‘Object Grasping’ Datasets exist, which either provide 2D or 3D grasping points or instances, 2D grasping rectangles,

¹This section is based on work devised by Rachel Azulay for completion of her Masters Thesis, in collaboration with and as supervised by the author of this PhD thesis. Here we present a summary of the offline algorithm development and training described in [127].

or directly a grasp type associated with each image. Some examples include the Stanford synthetic object grasping point data [63, 128], Dex-Net 2.0 [129], The Columbia Grasp Database [130], the Cornell Database [131], the Jacquard Dataset [132], and the Newcastle Grasp Library [77]. None of the datasets listed here are designed specifically to meet the requirements of our task. Although the problem addressed here is different from robotic grasping, as in the context of the first 5 datasets listed above, we can still extract relevant information/features from the grasp locations. Although the Newcastle Grasp Library is based on human grasping, relying only on the annotated grasp type to train a CNN may not allow for the extrapolation to novel objects through intermediary features. Thus, the Jacquard dataset [132] was deemed most suitable for use in our algorithm development, as a total of 54,485 images are provided for multiple angles for each of the 11,619 objects. Millions of grasp locations are also provided, defined by grasping rectangles, which were exploited to train our algorithm. Each object image comes with an RGB image, a binary mask, as well as other images not used in this project, along with grasp annotations as rectangles in a text file. To ensure the dataset is valid in the context of grasping with a prosthetic hand, the dataset was refined in to exclude all irrelevant objects which cannot be picked up with one hand. Objects with fewer than four rotation images were also removed.

3.2.2 Image pre-processing

To improve consistency in data and avoid losing information about the object size, the RGB object images were first cropped to a square centred on the object and then padded to consistent sizes. Each cropped image was overlaid with the grasping rectangles, such that each pixel value was assigned a number equal to the number of grasping rectangles overlayed over that pixel. The grasping rectangle overlay densities create a rough outline of the grasping locations for each object, defining a Region of Interest (ROI) for grasping the object. To homogenise the metrics obtained by the grasp rectangle overlay, a fast-marching method [133–135] was applied, where the shortest distance from each pixel to the boundary of the ROI is calculated to create a level-set mask. The distances are globally scaled across all images to values between 255 for the innermost pixels, with the largest distance to the

boundary, to 0 for the background. To simplify the semantic segmentation mask, the level-set image is segmented into contour bands, such that each band had a width of 25 pixels. The pixel labels are now the band to which each pixel belongs. An example of an image being processed is shown in Figure 3.4

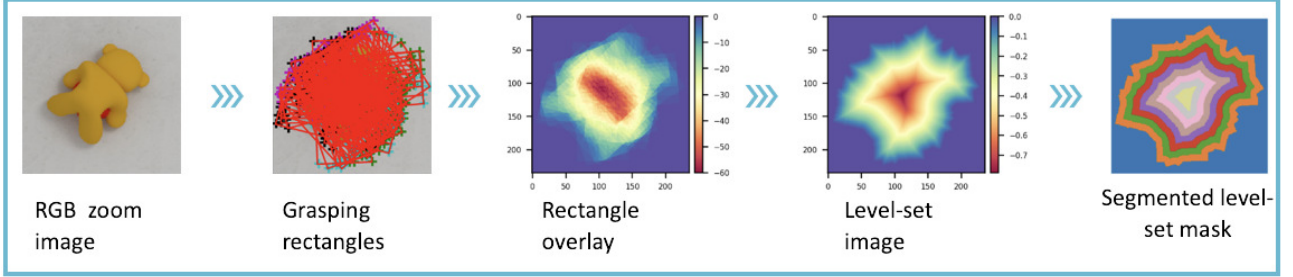


Figure 3.4: Step 1: Image pre-processing steps.

3.2.3 Algorithm development

Semantic Segmentation-Convolutional Neural Network

From here on the problem becomes one of image semantic segmentation, where we are classifying each pixel in the image into one of 12 classes (one for each band and the background). The neural network follows a U-Net [136] architecture, initially designed for biomedical semantic segmentation [137, 138]. This architecture has been effective in creating segmentation masks for a wide variety of images [139]. It is represented in a U-like manner, with a contracting path which downsamples the image, and an expansive path to upsample the image back to the desired resolution. We used the Fastai [140] implementation of the U-Net[136] architecture set up with a Resnet34 [141] encoder, pre-trained on the ImageNet dataset. We then trained it with our own dataset as described above. From an input 224x224 RGB image, the Semantic Segmentation-Convolutional Neural Network (SS-CNN) computes a class for every pixel to create a segmentation mask.

We choose to use a neural network for the contour-based recognition in the prediction pipeline, rather than a more simple contour recognition algorithm similar to that in the pre-processing stage. This is because once we move to an online setting with new images, we will not have access to the grasping rectangles associated with each object or to the real size of the object, both of which are

used to generate the level-set masks that train the SS-CNN. Without these parameters, the SS-CNN allows us to predict segmented level-set masks that are scaled across all images, accommodating for new images as well.

The hyper parameters for the SS-CNN model were chosen following the methods suggested by [142–144], and the training was broken down into two stages. GPUs were required to train the SS-CNN. To measure the efficiency of the SS-CNN, the relative error of each classified pixel is calculated as

$$\delta = \left| \frac{v_A - v_E}{v_E} \right| \quad (3.1)$$

where δ is the relative error, v_A is the observed value v_E is the ground truth. Since v_E increases from the contour to the centre, for the same absolute error, the relative error is higher for the external bands and lower for the internal bands. This prioritises the correct classification of the pixels in the outer bands over that of the pixels in the inner bands. The outer bands provide more relevant information about the shape and relative proportions of the object in the image, and as such the outer bands will be used in the feature extraction for the K-means clustering. This relative error also penalises more severely a greater gap in the classes.

The dataset was split into training, validation, and testing sets, ensuring a balanced distribution of two metrics across each set - the number of nonzero pixels in the level-set image, reflective of the ‘area’ defined by the object, and the maximum value of the level-set, reflective of the ‘width’ of the object as seen from this angle. Each of these metrics was divided into 3 bins to split the dataset into 9 categories of images. This results in a distribution of shape information – for example, an image with a high number of pixels and low maximum pixel value has a long and narrow shape, whereas an image with a high number of pixels and high maximum pixel value is large and bulky. 100 objects from each group were used in the training and validation set, for a total of 3,600 images with an 80-20 training-validation split. The remaining images were allocated to the testing set. The distribution is demonstrated in Figure 3.5.

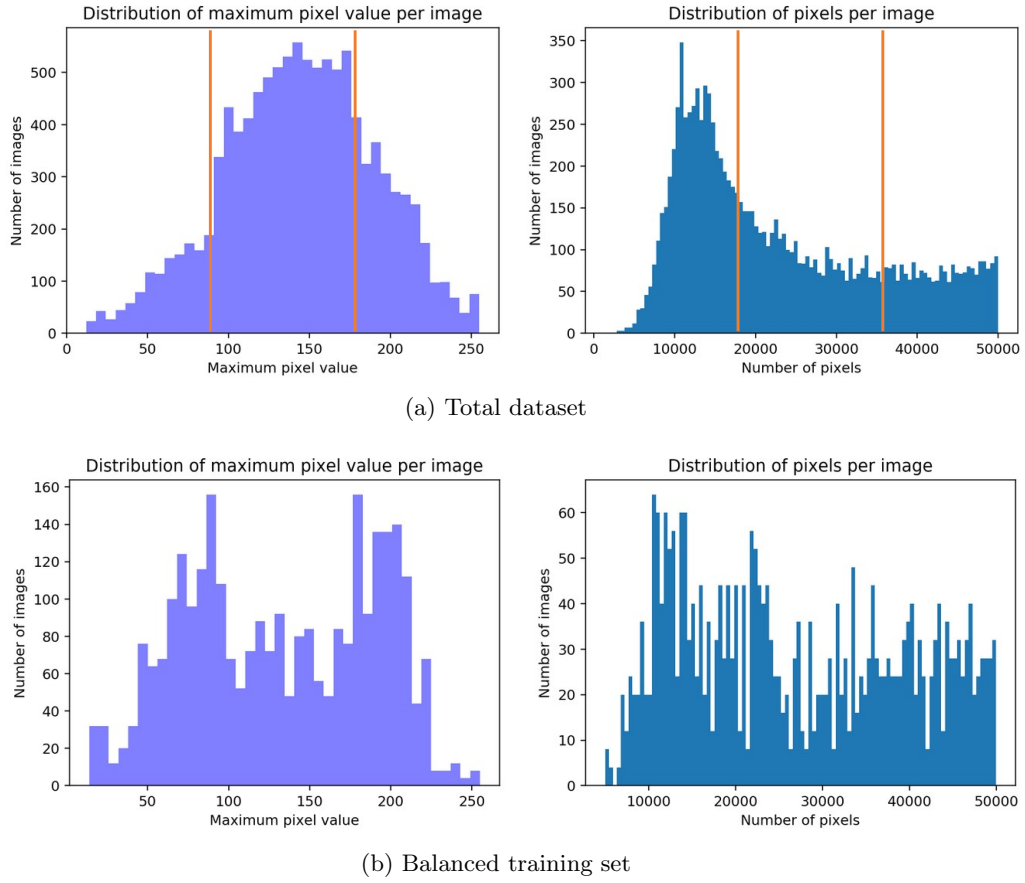


Figure 3.5: Distribution of images based on 2 metrics: maximum pixel value in the mapped range 1-255 and the number of nonzero pixels, in the original dataset as well as the balanced dataset.

Unsupervised K-means clustering for grasp classification

Feature extraction: Our grasp taxonomy is based on precision of the task and the shape and size of the object, so we extract relevant features from the predicted level-set mask. Contour recognition of the outermost band is used to obtain relevant shape information about the object. The Hu-moment invariant values of the level-set image, values that are invariant to translation, scale, rotation, and reflection, and give a score indicating the difference between a regular shape and the level-set image,) are compared to those of a regular circle, and this similarity score is used as a feature. The proportion of pixels in each of the first 3 bands out of the total number of pixels in the object were also used as features. The features are then standardised by removing the mean and scaling to unit variance.

The dataset we used is not labelled by grasp type, thus we fed the features into an unsupervised K-means classifier to cluster the dataset.

Using a K-means model allows us to focus on features around the properties of the object or part

of object in the image rather than the object label itself. An unsupervised method means we do not need to manually label a large number of images in a subjective manner, and the clustering can take place based on the features we have chosen.

We aim to classify the images into 4 grasp types, but, as per our taxonomy of grasps in Figure 3.2, we would expect 5 clusters, such that the tripod grasp could be assigned to 2 clusters. However, upon plotting the inertia for clustering analysis, 4 is determined as the optimum number of clusters, as indicated by the elbow in Figure 3.6. The k-means algorithm separates samples into groups of equal variance by minimising a criterion known as inertia or within-cluster sum-of-squares [116]. This is defined by

$$inertia = \sum_{i=0}^n \min_{\mu_j \in C} (||x_i - \mu_j||^2) \quad (3.2)$$

where x are the samples with n samples and μ_j is the mean of the samples in the cluster C , also known as the cluster centroid.

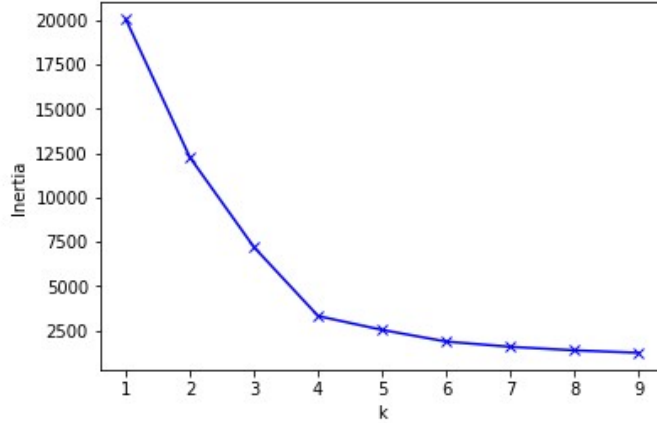


Figure 3.6: Elbow plot for the Kmeans clustering model indicating 4 as the most appropriate number of clusters. k is the number of clusters. Inertia is the within-cluster sum-of-square as defined in Equation 3.2.

The reduction of the clusters from the expected 5 to the optimal 4 could be due to the image dataset not containing a significant number of small round objects, one of the branches in Figure 3.2 that could be classified as a tripod grasp. As such, we decide to cluster the images into 4 clusters. This also corresponds to the 4 grasps we want to use for classification. Examples of the images and levelset images in each cluster of the K-means model are shown in Figure 3.7. A grasp type was then assigned to each of these 4 clusters, based on our taxonomy.

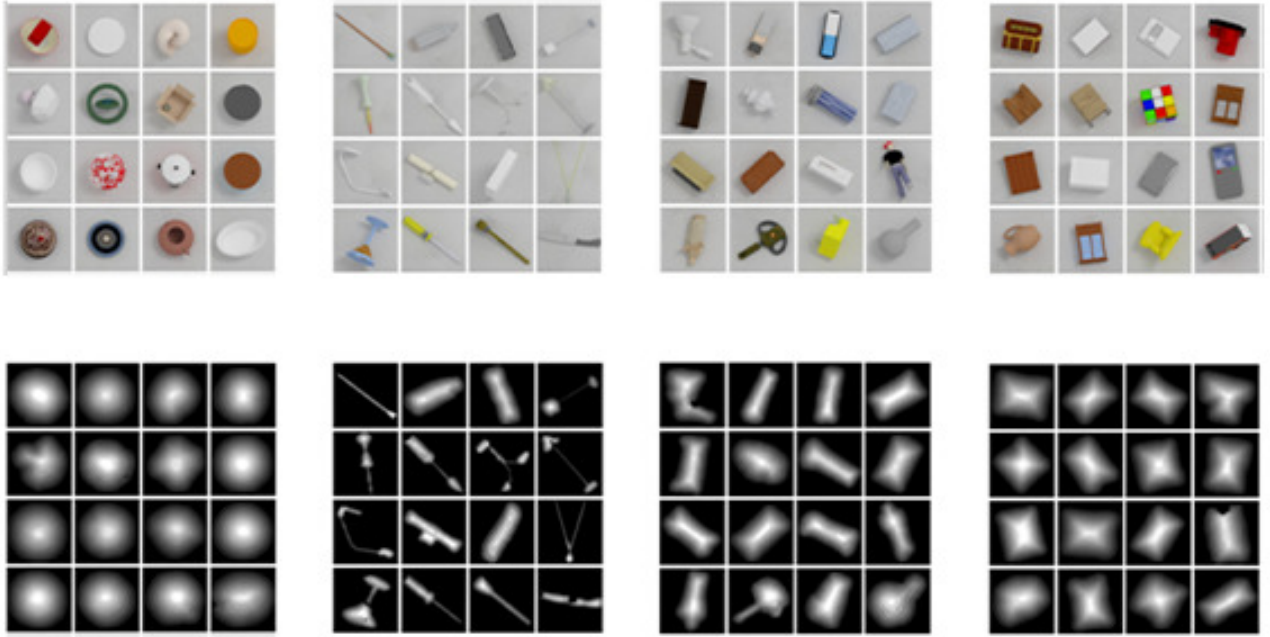


Figure 3.7: Examples of K-means clustering (Step 3).

3.2.4 Results

We obtained 21% relative error on the predictions of the trained SS-CNN on the testing set. Some examples of the predicted level-sets are shown in Figure 3.8. To validate the proposed K-means clustering method, a balanced sample of the dataset was manually annotated with appropriate grasps. We obtained 88% of correctly classified grasps. A normalised confusion matrix for the K-means clustering is shown in Figure 3.9.

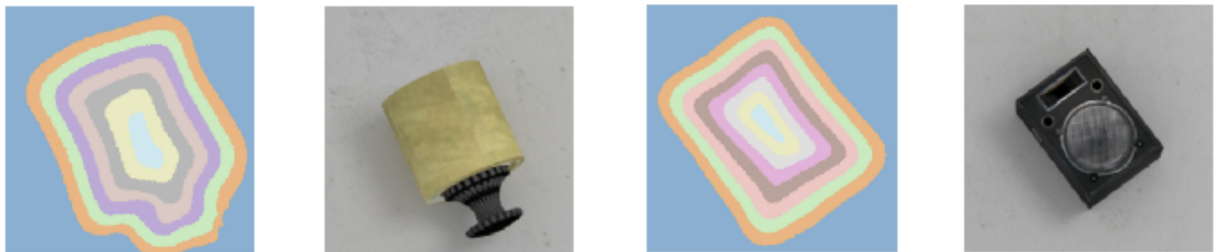


Figure 3.8: Step 2: Predicted level-sets from SS-CNN with associated RGB image.

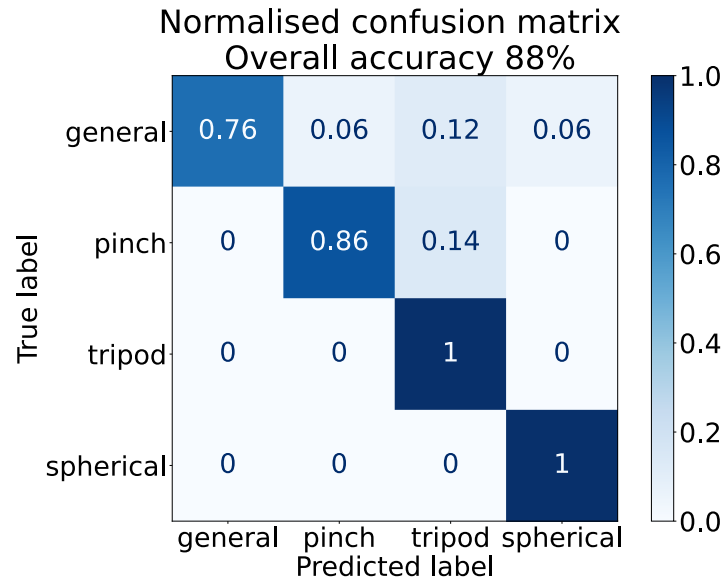


Figure 3.9: Normalised confusion matrix for grasp classification after having labelled a sample of the dataset.

3.3 Computer Vision Real-time System development

Once the offline algorithm was developed, and its parameters established, I adapted and incorporated the trained models of the algorithm into a real-time system – referred to as the CV-RS. The CV-RS includes a webcam to capture continuous recordings. The CV-RS takes an RGB image frame as input and outputs a grasp state.

3.3.1 Hardware

In most robotic systems, the camera configuration is eye-to hand, such that the camera is mounted on a fixed place. In this configuration, the camera can observe the manipulation area as well as the end-effector and is not affected by the movement of the end-effector [122]. When it comes to assessing human grasping, or the use of cameras within prosthesis control, the camera is often used with a ‘first-person’ view, which would also provide a view of the working space, by, for example, being mounted on a headband or on glasses [54, 63, 145]. However, in order to avoid the need for additional accessories, which may discourage the user, and, to account for the approach of the user in the classification of the grasp type, we decide to mount the camera on the end-effector in an ‘eye-in-hand’ configuration. More specifically, the 5MP 1080p webcam (ELP-USB500W05G-FD100, Ailipu Technology Co, Ltd., China) is attached to an adjustable wristband via a custom 3D-printed setting. The camera is chosen to be

on the palmar side to provide a direct view of the object to be grasped as the hand approaches. The size of the board for the webcam (3.8x3.8cm) is small enough to be incorporated within the proof-of-concept system. Using an adjustable wristband allows the camera to be tested by able-bodied subjects as well as prosthesis users, without interfering with grasping objects. The webcam is connected to a computer via cable (though this could be modified in future iterations to be wireless).

3.3.2 System overview

The CV-RS can be summarised in the diagram in Figure 3.11. The CV-RS system is run on Python 3.7. The process begins with system initialisation, which involves loading all the previously trained models onto the GPU. The system then runs as a continuous cycle, beginning with the capture of an image. The image is passed through an object detection and recognition algorithm. Though we do not currently use the resulting object label to inform the rest of the grasp recognition algorithm, this object label provides insight into the observations of the camera and the ROIs for all the objects seen in the image. If no object is detected by the algorithm, we use the full image from the camera as there still might be an object, or part thereof, to be grasped within the frame. If at least one object is detected, the centremost object is selected and cropped, before being passed as input to the SS-CNN to predict a level-set mask.

Features extracted from this level-set mask are scaled and used to classify the grasp type from a previously trained K-means clustering model. Majority probability voting is used to smooth the classification output. If a new grasp type is selected for two consecutive instances, then the grasp state machine is updated with this grasp type as the state. As soon as the image is processed, another frame is captured and processed. It takes approximately 0.08s to process each image from time of capture to updating the state machine. The process is continuously run once initiated and is run on a laptop utilising a GPU (NVIDIA GeForce RTX 3050 Ti). The process is broken down into steps in Figure 3.11 and in Section 3.3.3. The images are given labels to aid with the description. An example of the image transformation throughout a cycle is shown in Figure 3.10.

For research analysis, a frame is saved for each of the following: original image, SS-CNN input,

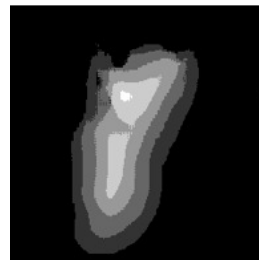
SS-CNN output. Videos showing the live image feed, the level-set output, and the state machine output are displayed in real-time to the assessor for observation of system performance. The tasks of saving files and displaying videos are offloaded onto a background thread. This frees up the main thread for interacting with the GPU therefore achieving higher overall throughput.



(a) Image captured by webcam.



(b) Cropping of region of interest for the centre-most object. This is the image that is passed into the SS-CNN.



(c) Levelset mask output by the SS-CNN.

Figure 3.10: Example of an image processing cycle.

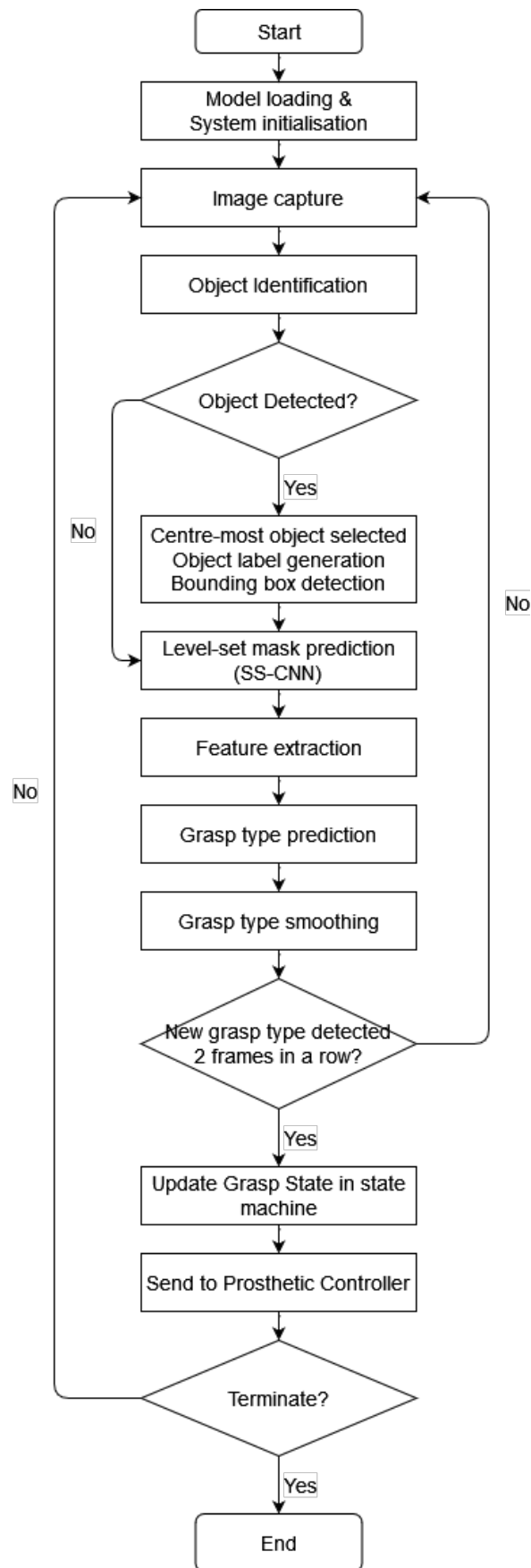


Figure 3.11: Computer Vision Real-time System Flowchart.

3.3.3 Image processing cycle

- **Image capture** The webcam captures an RGB image of 640x480 pixels.
- **Object detection and recognition** Since the SS-CNN and K-means models were trained on images containing only one object, an off-the-shelf object detection and recognition model, YOLO (YOLOv5s [146]), is used to localise and label any objects in the image. YOLO is chosen because of its speed and ease to implement. It is a pre-trained CNN, trained to detect and recognise objects in the COCO-80 set [147]. The COCO-80 set is a list of 80 objects, including several objects used in activities of daily living that users may interact with using their hands. Although more comprehensive object sets exist, they were deemed unnecessary as they could add complexity and worsen the performance of the system, especially since object recognition is not our aim in this system. Thus, for speed, we do not choose a more comprehensive set, as we do not require an extensive list of objects. The YOLO model predicts a bounding box for each object along with an object label. The object label is also stored, but not used in the grasp selection. If no objects are detected, the original image is used as input to the SS-CNN. We still use the original image because there could be an object in the image, but the object may not have been detected by the YOLO model. Thus, we still proceed with predicting a grasp type based on such an image.
- **Image pre-processing** If one or more objects are detected by YOLO, the object with the centre of the bounding box closest to the centre of the entire image is selected as the object of interest. This is determined using the Euclidean distance between the centres of the ROIs and that of the image. The original image is then cropped to an image of size 224x224. The dimensions and localisation of the cropped image are initially determined by the bounding box provided by the YOLO, but they increased to 224 pixels where necessary.
- **SS-CNN** Either the original image or the cropped image, as appropriate, is then passed as input to the previously trained SS-CNN. We use the SS-CNN model trained on a band width value of 25. The output is a level-set mask of size 224x224 pixels.

- **K-means clustering** Features for use in the K-means classification are extracted from the level-set image, as described in the offline training. The features are scaled and transformed as above. The pairwise Euclidean distance of the features to each of the 4 cluster centres is calculated, normalised, and used in majority voting to determine the smoothed probability output class over 3 frames. Each cluster centroid is mapped to a grasp type label. The output here is the ‘grasp type’ predicted for the original image by the overall system.

3.3.4 Updating the state machine

The system is continuously updating the predicted grasp type, and this grasp type may change as the hand moves around. To avoid changing the Grasp State when not focusing on a specific object, we apply a smoothing to the grasp type output. At the same time, we need to balance this with the response time of the system in changing grasp type when necessary. Thus, if a change in grasp type is detected, consecutively for 2 frames, then this new grasp type is taken as the Grasp State stored in the state machine. This can then be passed on to a prosthetic hand controller, where decisions on when to use the grasp in the state machine are taken separately. Additional algorithm parameters have been tuned based on empirical results from online experiments as described in Section 4.3. The Grasp State is the final output of the CV-RS.

Supplementary Video 3.1 shows a demonstration of the Computer Vision system where the Grasp State automatically updates the connected prosthetic hand.

3.4 Validation of the real-time system

To validate the real-time implementation of the CV system, we designed a study for able-bodied subjects. By attaching a camera to the subject’s wrist during a reaching and grasping task, we can compare the CV-RS output to the grasp type actually used by the subject. The study also allows us to establish certain parameters of the system that could later be used for general or personal calibration, when implemented in a prosthetic hand.

Investigating reaching and grasping mannerisms of able-bodied users when interacting with differ-



Supplementary Video 3.1: Demonstration of the Computer Vision system being used to automatically control the grasp type of a prosthetic hand. The grasp type changes from target to target: ‘pinch’ for the pen; ‘general’ for the bottle; ‘spherical’ for the ball. Video can be viewed at: <https://youtu.be/sMm7I66R-jQ>

ent objects is an interesting research area, as understanding such tasks is essential for feeding back into prosthetic and robotic designs. This study enabled us to create a dataset of grasping task recordings, from the eye-in-hand perspective as well as an external view, which provides a good basis for a number of different analyses. As such, this study allowed us to collate a mapping of natural grasp types associated with different objects, as a baseline definition for ‘correct’ grasp types. This dataset can be used in future work to create, train, test, and validate algorithms that can identify the grasp type from the image of an object for a range of reach and grasp applications.

Thus, we hypothesise that different users may use alternative, though similar, grasp types for the same object and context. We further hypothesise that we can successfully discern a valid grasp intention from the CV algorithm in real-time for each grasping task.

3.4.1 Experimental protocol

2 able-bodied subjects participated in the experiment. All subjects consented to participate in the experiment, the protocol for which was approved by the Imperial College Research Ethics Committee (JRCO number: 18IC4685). The subject information can be found in Table 3.1. In the experiment, the subjects were asked to perform a series of reaching and grasping tasks involving their dominant

arm. A wristband with a webcam attached was strapped to their wrist, as described in Figure 3.12. Each subject was asked to stand on a fixed spot at 10cm distance in front of the table. At the start of each task the subject assumed the neutral starting position, with their elbow bent and their palm facing down. The objects were placed on a fixed point on the table, directly in front of the subject's arm, such that the distance from the end of the table was the same as the forearm length (EL to US), with standardised orientations for each object. The table had a custom backdrop setup to ensure consistency between images. The table height was adjusted such that there was an 11cm height gap between the top of the table and the bottom of the wristband on the subject's arm when the subject was stood in the neutral starting position. 2 cameras recorded the subject's arm movements from the front and the side. The setup is shown in Figure 3.13.

Table 3.1: Subject information for CV-RS experiment.

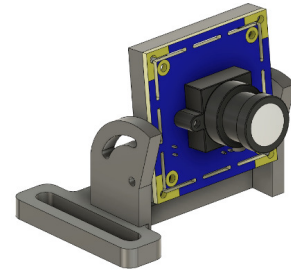
Subject	1	2
Gender	Female	Male
Forearm length (EL to US)	28cm	25cm
Hand length (wrist to tip of middle finger)	18.5cm	17.5cm
Hand width (M1 to M5)	11cm	13cm



(a) Webcam mounted on wristband, shown worn by able-bodied user.



(b) Webcam is attached to an adjustable Velcro strap



(c) CAD render of 3-D printed adjustable attachment with camera placement.

Figure 3.12: Webcam with attachment and as worn by a participant

Task set

The subject was asked to naturally grasp and lift the object just off the table (under 10cm, hold for around 2s) before returning it to its original position. Then, the subject was instructed to return their hand to the starting position and rest for 2s, before returning to grasp the object again. Each task per condition was repeated three times. Usually, no instruction for grasping was given, with care taken not

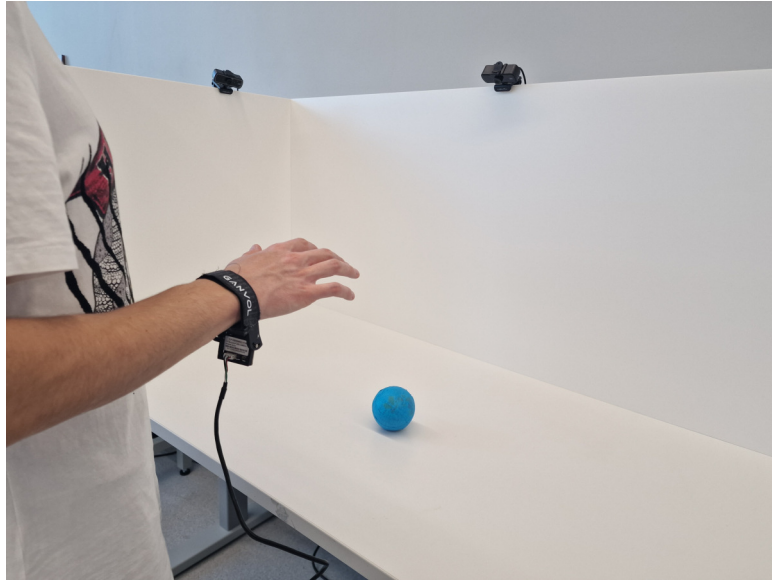


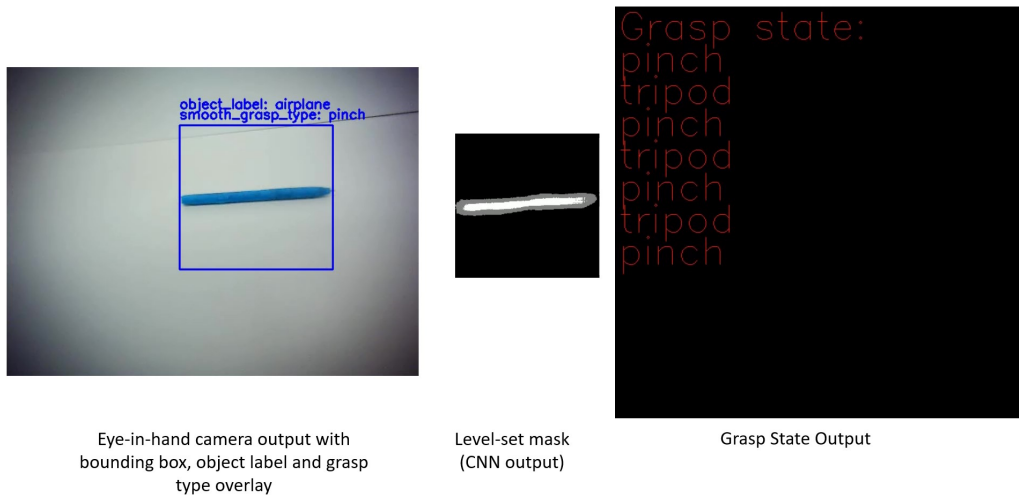
Figure 3.13: Setup of the experiment, showing subject’s starting position, webcam wristband, object placement, custom backdrop, and external cameras.

to influence the choice of grip type. However, for some objects, different approach contexts are given. A list of objects and their grasping contexts can be found in Table 3.2. The order of the objects will be the same for all subjects. The camera on the wristband was continuously sending images to a laptop running the CV-RS, to predict 1 of 4 grasp types as the grasp state. This procedure was repeated three times. During the second round, participants was introduced to the 4 grasp types recognised by the CV-RS. The participants were restricted to using only these grasp types during this round. For each subject round there was a total of 11 objects, 14 task contexts, each task per condition repeated 3 times (39 reach and grasp tasks in total, per round, 117 tasks in total per participant).

An example of a subject reaching and grasping the pen during the second round of recordings is shown in Supplementary Video 3.2 and Supplementary Video 3.3.

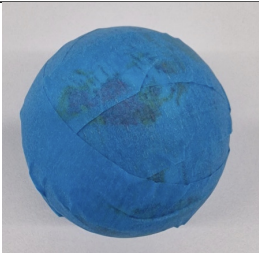


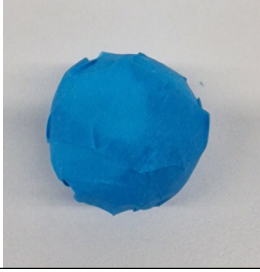
Supplementary Video 3.2: Video Output from the fixed external cameras during the second round of reaching and grasping tasks. The participant reaches for the pen and grasps it 3 times. Video available at: <https://youtu.be/U7OCb6oHigo>



Supplementary Video 3.3: Video Output from the eye-in-hand camera during the second round of reaching and grasping tasks. The participant reaches for the pen and grasps it 3 times. The corresponding level-set mask which is output from the CNN is also shown along with the continuously output Grasp States. Video available at: <https://youtu.be/Xw4qdvd9K1I>

Table 3.2: List of objects, their dimensions, and grasping contexts.

Image of object in fixed orientation	Object name	Dimensions (cm)	Grasp Context 1	Grasp Context 2
	Pen	L: 14 $\varnothing$: 0.8	Grasp as if to use to write	N/A
	Ball 1 (large)	$\varnothing$: 7.5	N/A	N/A
	Clothespin	L: 95 W: 4.9 H: 3.1	Grasp as if to use	N/A
	Knife	L: 22.5 W: 1.1 at narrowest point (2 at widest point) H: 0.7	Grasp as if to lift the handle	N/A
	Bottle	H: 31cm $\varnothing_{bottom}$: 7.5 $\varnothing_{top}$: 4	Grasp from main body	Grasp the lid

	Ball 2 (small)	$\varnothing$: 2	N/A	N/A
	Scissors	L: 21 $W_{handles}$: 8.6 W_{joint} : 2cm H: 2.5	Grasp as if to use (from handles)	Grasp as if to pickup (from joint)
	Jar	H: 6.8 $\varnothing_{top}$: 8.5 $\varnothing_{bottom}$: 6.3	Approach from top	N/A
	Tweezers	L: 12.5 W: 1 H: 1	Grasp as if to use	N/A
	Cup	H: 8.4 $\varnothing_{top}$: 7.5 $\varnothing_{bottom}$: 5.5	Approach from top	Approach from side
	Ball 3 (medium)	$\varnothing$: 3	N/A	N/A

Data analysis

The study has two purposes: establishing parameters and validating the classification. In order to validate the classification, we need to identify the point at which the camera has the object or object part to be grasped in reasonable view such that it functions as designed and outputs the relevant grasp type. As such, we first perform a frequency analysis of the grasp class output by the CV-RS.

Establishing parameters

We define the point at which the person grasps the object for each trial as $t_f = 0$, and then all prior frames f that have been processed are assigned t values in receding order, i.e. $t_f = -1, -2, -3, \dots, -n$ where n is the number of frames in the movement sequence. As such, there is a Grasp State (GS) output by the CV-RS at each frame, GS_f . We calculate the frequency of each class of the GS for frames over a specific window of frames, W , starting from frame f such that $GS_f \leq GS \leq GS_{f+W}$.

We only used the second, or “main”, round for now, in which the grasp options were limited to those produced by the TASKA hand. We vary the window size and plot the frequencies for each trial for $-n \leq f \leq 0$. In Figure 3.14, we can see this plot for an average classification prediction frequency over all trials for both subjects. This plot allows us to determine the W and f that we should use to determine the overall predicted class of the CV-RS for each trial. The overall predicted class is the class with the highest prediction frequency within the chosen window. This predicted class is to be used here to examine the accuracy of the CV-RS, and could also be used in the future for calibration of parameters for each participant.

Accuracy calculation

As expected, the subjects do not use consistent grasp types for the same object. As such, there may be multiple “true labels” associated with grasping each object grasp context. The overall predicted class of the CV-RS for each trial for each grasp context is compared with all the true labels associated with that grasp context. It is considered a correct classification if it matches either of the true labels. The accuracy is thus calculated as:

$$\text{Accuracy} = \frac{\text{Number of correct classifications}}{\text{Total number of trials}} \quad (3.3)$$

3.4.2 Results

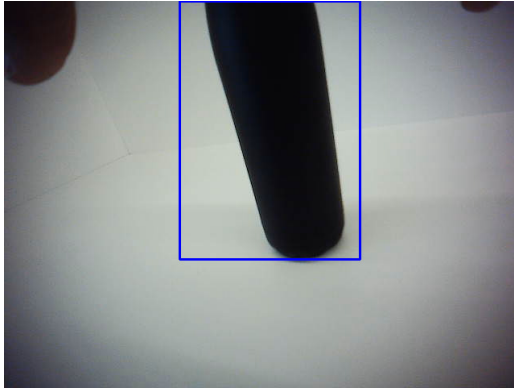
Figure 3.14 shows some example outputs of the CV-RS in different scenarios. The list of “true” labels for each object were generated from the grasp types chosen by the subjects, as labelled manually from the external video feed. These are shown in Table 3.3, along with the grasps used by the subjects in the unrestricted first round.

Table 3.3: Grasp type used by each subject for each grasp context in the both rounds of recording where they were first unrestricted to any grasp type and then limited to the 4 grasp types used in the CV-RS.

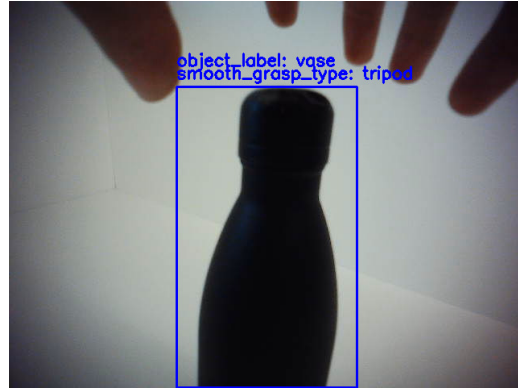
Task context	Subject 1 grasp - unrestricted	Subject 2 grasp - unrestricted	Subject 1 grasp - restricted	Subject 2 grasp - restricted
Pen	4-finger pinch	tripod	pinch	tripod
Ball_large	spherical	spherical	general	general
Clothespin	general	4-finger pinch	tripod	tripod
Knife	5-finger pinch	4-finger pinch	tripod	tripod
Bottle_body	general	general	general	general
Bottle_top	general	spherical	tripod	general
Ball_small	general	tripod	pinch	pinch
Scissors_handles	general	5-finger pinch	tripod	tripod
Scissors_pickup	5-finger pinch	5-finger pinch	general	tripod
Jar	general	general	general	spherical
Tweezers	4-finger pinch	4-finger pinch	pinch	tripod
Cup_top	general	general	general	tripod
Cup_side	general	general	tripod	general
Ball_medium	5-finger pinch	tripod	pinch	tripod

Frequency analysis

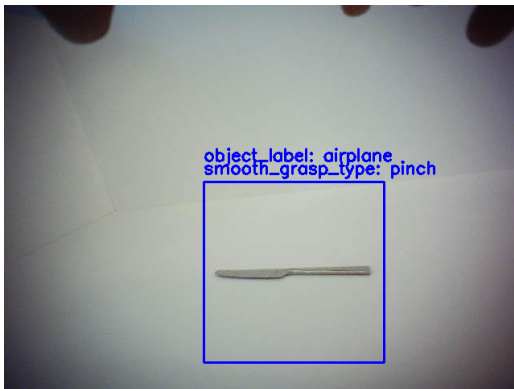
Figure 3.15 shows the frequency of class output for the different predicted classes in windows of size 10 frames beginning at the frame number indicated. The ‘true’ classes associated with the object are indicated in each plot. Data for 3 trials averaged for both subjects is shown. It appears that the data windowed at start samples between -20 and -15 are most stable and consistent with the true labels. A similar exercise is repeated for different window sizes, and the results are shown in Table 3.4. Based on these results, we choose an f of -15 and a W of 10. These parameters produce an overall accuracy



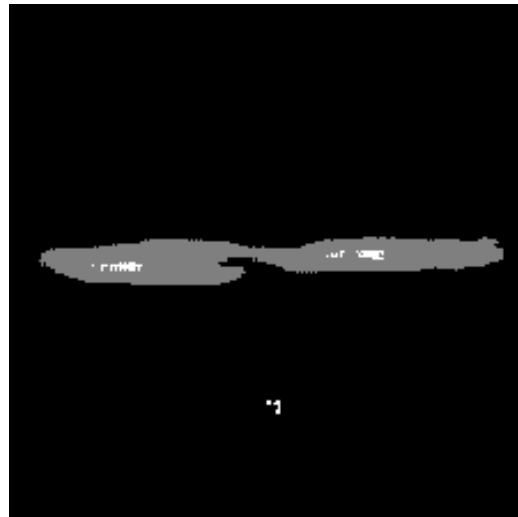
(a) When the camera sees the body of the bottle it generates a GS = general



(b) When the camera sees the lid of the bottle it generates a GS = tripod



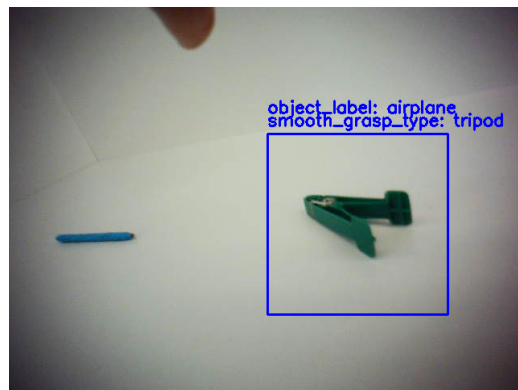
(c) Knife labelled with the incorrect GS = pinch



(d) Levelset mask of knife output from CNN. The shape is consistent with the requirements for a GS = pinch



(e) Jar mislabeled as cup, generates. GS = spherical.



(f) Object closes to the centre is selected by the CV-RS. Clothespin is mislabeled as an airplane by the object detection model. GS = tripod

Figure 3.14: Example outputs from the CV-RS showing the overlay of the object bounding box, object label, and grasp state overlayed over the original input image.

of 79%. These parameters will be set to be used in the complete EMG-CV system.

3.5 Discussion

As hypothesised, people are not limited to only one grasp type when grasping the same object. This depends on the context for use, and the part of the object being grasped. This reaffirms our choice to not select the grasp type based on the object itself, but rather to do so based on features relating to the shape of the part we want to grasp. On top of that, even within the same grasp context, there are multiple successful grasping type options. The grasp type used not only varied between participants, but also sometimes even for a specific participant. Even when limited by the choice of grasp type, though the person may have used a different grasp type initially, they were still able to successfully complete a grasp. The grasp recognition algorithm is also developed on an individual's perception of associating grasp types with shape, following on from Cutkosky's taxonomy [56] so it does not necessarily reflect every individual's choice of grasp context. However, the flexibility of grasp type used displayed by the subjects suggests that, even with limited grasping options, as with our complete EMG-CV prosthetic system, users should be able to grasp objects successfully. A range of items were selected to provide a range of grasp types based on expected CV-RS output. However, it seems our participants did not use the spherical grasp with objects we may have expected them to do so with, e.g. with the ball and jar. The difference between the spherical grasp and the general grasp is quite subtle when used by a person, thus it is quite difficult to manually determine which of the two grasp types has been used for annotating the data. The difference is based on whether the thumb is abducted or in the palmar position.

In general, the CV-RS is successful in identifying a correct grasp type to be used with a grasp

Table 3.4: Accuracies attained when varying f and W .

Frame f	Window size W	Subject 1 accuracy	Subject 2 accuracy
-15	5	0.69	0.81
-15	10	0.76	0.81
-20	5	0.71	0.83
-20	10	0.69	0.83

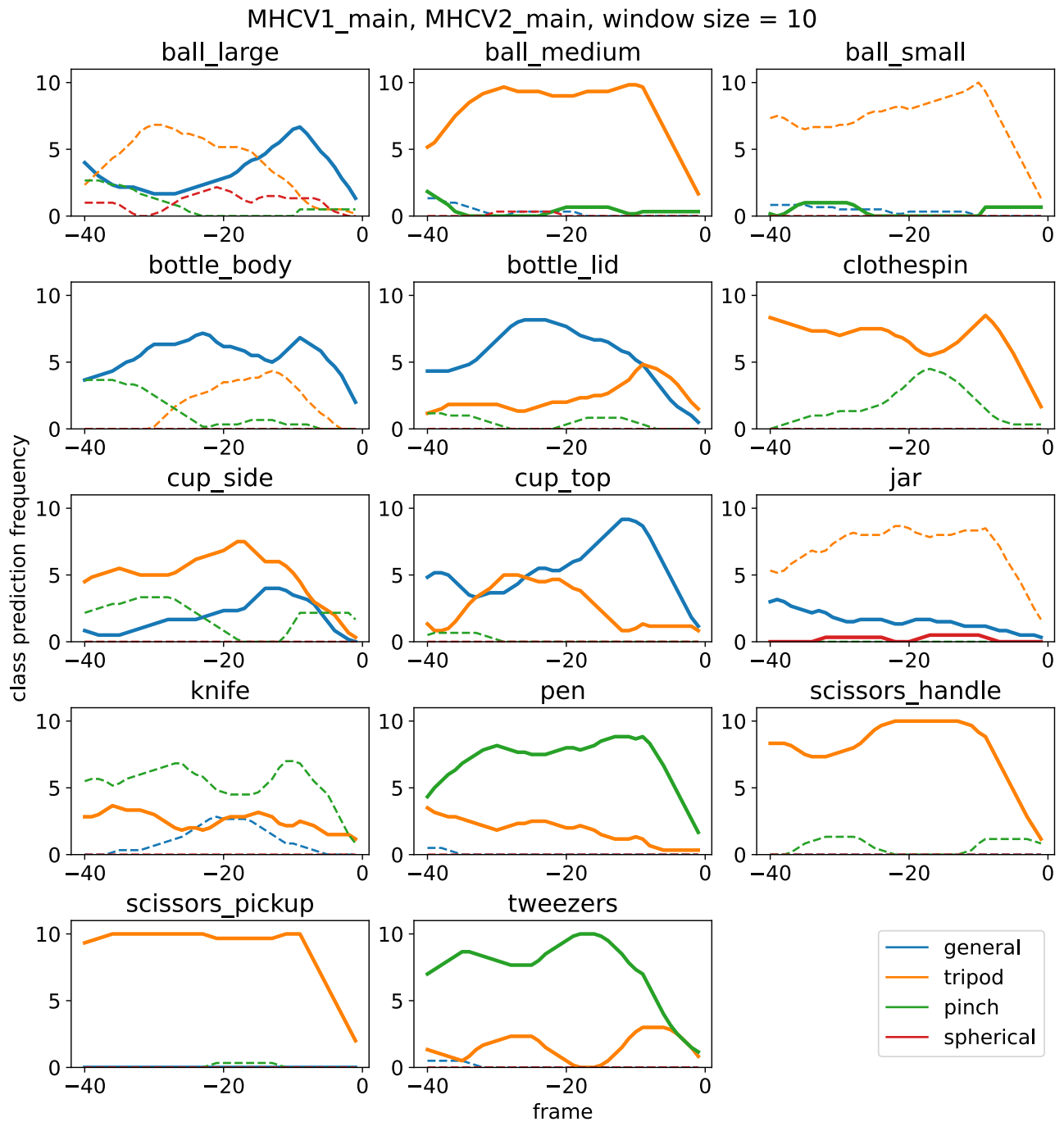


Figure 3.15: Frequency of class output for the different predicted classes in windows of size 10 frames beginning at the frame number indicated. The ‘true’ classes associated with each object are indicated as solid lines in each plot, whereas the dashed lines represent incorrect predictions. Data is averaged over 3 trials and both subjects.

context, following the criteria set out in Section 3.1.3. The grasp state successfully changes, not only based on the object, but also based on the grasp context, with 79% accuracy. This is a slight decrease in performance from the offline validation at 88% accuracy, but it is comparable to other systems. Ghazaei et al. [77] report 84% classification accuracy in an offline setting, but 88% successful trials with a user in the loop. Bu et al. [79] report 90% accuracy in an offline setting. In our proposed system, even if the object label is incorrect, or no object in the COCO list is detected, we are still able to continue with the algorithm, and identify the grasp type. This shows we can extrapolate to unknown objects and removes the reliance on object recognition.

Some trials do result in incorrect classifications. The most common errors in classifications in both the offline and online validations occur between pinch and tripod and between general and spherical. This is not surprising as there is potential overlap in the parameters that are used to distinguish between these classes. For example, the knife is always classified with the pinch grasp type. This grasp type is actually expected based on how the algorithm is trained – thin, long objects are assigned a ‘pinch’ class. However, the algorithm does not take into account other parameters such as the weight of the object. Thus, a human would assess that the knife is expected to be heavy, and so would use at least three fingers to grasp it. A similar concept applies to the small ball – the GS output by the CV-RS is tripod, which is what we would actually expect based solely on our taxonomy as a small round object. The ‘Spherical’ grasp is applied to round objects, but these object are also often larger and with similar width:length ratios, as associated with the ‘General’ grasp.

The placement of the camera allows for limited obstruction of the field of vision, and for a grasp state to be continuously generated up until the object is grasped. The CV-RS is also robust to other objects present in the field of view and is not affected by fingers partially obstructing the field of view, thanks to the selection of the object closest to the centre of the field of vision. However, further improvements could be made to ensure the system is robust to changes in the background, and further obstructions.

We also process the entire frame captured by the camera for images where no object is detected by the object recognition model, instead of skipping it. This is to ensure that any objects present that

are not detected by the object recognition model are still classified into a grasp type. It also allows for sections of objects to still be considered for classification which is a unique advantage in our system.

There are a few cases where this could happen:

- There is no object in the image. The user may not be reaching towards an object for grasping, or the object has already been grasped. In this case any grasp state output by the system will not be used anyway.
- Part of an object is in the image:
 - The user could be reaching towards grasping part of an object, e.g. lid of bottle. In this case, processing the frame is important and allows for associating different grasp types with different parts of an object.
 - The prosthetic hand is too close to the object, and the user has almost grasped the object. In this case the grasp states classified from these images are discounted anyway because the overall grasp chosen for each object is determined based on a previous window of Grasp States.
- There is a full object in the image, but it was not detected by the object detection and recognition model:
 - The object is typically detected, but was not detected in this instance. This results in spot areas not correlated with the object position appearing in the level-set mask. This could result in incorrect classification. However, the grasp types are smoothed out throughout the reaching task.
 - The object is rarely detected by the object detection algorithm. Even though there may still be spot areas resulting from using the complete frame, the user still wants to grasp the object. So classifying a grasp type is required, which would not be possible if the frames were skipped.

As such, at this stage, we decide to use the entire image for processing when no object is detected, to ensure all objects and parts of objects are captured, even though this runs the risk of mis-classification of a few frames.

The CV-RS is continuously producing a grasp type that is being used to update the Grasp State. Based on the frequency analysis performed, we have identified a suitable time to select the GS to be used. Obviously, capturing the right moment for GS decision depends on multiple factors such as velocity of the user’s arm, as well as the distance from the object, etc. Without sensors that can provide information on velocity and distance, and to validate our current CV-RS, we have created an empirical method for making this decision.

From the experiment we have performed, we have collated 3 simultaneous video streams of the reaching and grasping tasks within these object and grasp contexts. This provides an interesting and unique dataset lending itself for further examination. The dataset could be further grown and the video feed from the eye-in-hand camera could be used to re-train the SS-CNN based on different approaches, with data captured for our specific use case from real human grasping tasks.

Several recent studies consider reaching and grasping tasks in different contexts to create datasets of human reaching and grasping. Motion capture is commonly used for such datasets [148] for object and hand pose estimation, but images are also often captured [149]. Reaching and grasping task datasets are also collated and studied as part of varied experiments for different purposes (e.g. measuring EEG activity [150]) or simply reaching and grasping for CV classification [72, 77, 151]. Some datasets contain just images captured separately to experimentation, such as the Newcastle Grasp Library [77] which was developed for human grasping tasks. However, the images are static and are taken from a fixed distance to the objects. They do not include the reaching aspect or even the grasping, just the images labelled by a person. Experimental tasks can contain static reach and hold tasks [150, 152], pick and place tasks, or more dynamic manipulations of objects [152], while some research considers more general activities of daily living in a living environment [145, 153]. In some experiments the grasp type to use is indicated to the participant [152], while in others, the participant is not restricted in any way [145, 153]. Several reaching and grasping datasets are recorded from a first-person point-of-view,

mainly with a camera on the head of the participant or at eye-level [145, 152, 153]. Han et al. [151] record continuously from both the person’s eye-level as well as from their palm (along with EMG and IMU sensors). They record dynamic reaching, but the hand initially starts on the table and moves towards the object, which does not generalise to reaching tasks where the hand does not start on the same surface as the object, and prediction of grasp type was assessed at the initial position.

As far as I am aware our research presents a unique set that considers dynamic movements recorded continuously throughout reaching and grasping tasks from an eye-in-hand point of view with human participants alongside fixed external cameras. We also consider grasp classification at multiple stages of the reaching task. He et al. [81] also consider classification at different frames during a task, but this is for object classification rather than grasp classification. Of course, however, we should extend our study to greater number of participants, objects, and grasp contexts in order to create a more complete dataset that encompasses a greater variation across different scenarios. The dataset would also be more complete with the incorporation of kinematics, IMUs, or even EMG. In this case, we could gather more information about reaching and grasping tasks simultaneously, that could provide insight for a number of analyses. They could also provide a basis for more complex movement prediction and control systems.

In conclusion, we have developed an algorithm that is successful in classifying the grasp type of an object based on an image in an offline setting. This has been integrated with a real-time system, which has also been validated as successful for the premise of the overall EMG-CV system. This should now be integrated with the rest of the main system.

Limitations and future work

One main limitation of the system is the reliance on a mainly uncluttered and single-coloured background, with consistent lighting. Reflections affect the level-set image created, and although the system can localise and detect the central object in an image, a very cluttered background would be unsuitable. As such, further validation needs to take place in a non-laboratory setting, under real case conditions such as in a person’s home as in [145, 153]. The background limitation is also a common

CV problem, not unique to our task, and is an area which needs further development and refining for our use case.

Determining the point at which the CV-RS has output the correct GS via the software allows for continuous GS detection, and avoids the user having to give additional input – in [77] for example, the EMG is used to capture an image, which is then used as the basis for grasp classification. As we have indicated, our aim is to provide intuitive control for the user. Thus, we need to keep the EMG free to be used for wrist control and closing and opening the hand. For semi-autonomous control, the camera and any additional sensors should operate autonomously without requiring direct input from the user themselves. Instead of relying on time for selecting the ideal grasp state selection, it might be prudent to integrate an IMU as in [154] to help with the intention to grasp from the user. There are downsides to continuously running the CV-RS in terms of power and memory consumption. Additional sensors could potentially help determine when the CV-RS should be running. To more accurately determine the point at which the CV-RS has output the correct GS – we could use additional sensors to determine the speed and the distance from the object. IMUs could help determine the speed of the movement, and a depth sensor would provide useful information to help determine the size of the object, which could further improve the grasp type classification.

As a result, an additional area of research entails investigating the optimal positioning and arrangement of sensors on the prosthetic as well as the number and type of sensors needed to maximise the field of view. Thus, several different sensor arrangements could be studied to test for a number of parameters including, but not limited to: number of sensors, type of sensors, placement, distance from surface of prosthetic device, etc. To allow for a systematic investigation, we propose that these parameters could be tested in a simulated environment. Using a simulated environment could also help provide additional datasets for the training of a CV-RS, as multiple images could be produced of a range of objects with various backgrounds and with various reaching angles under various lighting conditions. This would help the algorithm become more robust to changes in background, light, and the presence of clutter, which are currently limitations of the CV-RS. The current speed of the CV-RS (around 0.08s per frame processing) is acceptable, but could be further improved. This leads to the

question of whether we need to use the YOLO model for object detection and recognition at all. The object label is not used in the rest of the processing, and the bounding box could be determined with other, less time-consuming methods.

When it comes to processing complete frames when no object is detected, skipping certain frames may help reduce the processing complexity and speed. There is room here for further investigation into what percentage of images contain no object detected and what part of the reaching/grasping cycle they belong to. A balance needs to be found between ensuring all parts of objects and objects intended to be grasped are detected and processed when necessary, while also reducing the processing time and complexity when classification is not needed. As such, using an alternate method to capture the bounding box for each image could also be a way to contribute to detecting objects and parts of objects more appropriately.

Looking forwards, we could also investigate incorporating a depth sensor or an IMU for a more complete sensor suite. Additional sensors could help determine at which point the grasp decision should be made, without relying on a fixed number of frames. The reaching movement could be detected with an IMU, which could also be a good indicator for when to initiate the Computer Vision Assistance. This could potentially address power concerns for a continuously running CV analysis. A depth sensor could also aid with determining the correct grasp type before the object is too close. It could also provide additional information about the size of the object, which could inform the grasp classification algorithm. Within our current research we have still been able to accommodate for the absence of additional sensors using calibration parameters, leading to a simpler but still effective system.

Finally, the system currently runs with wires and through a laptop. Ideally, the webcam, or any other sensors, would communicate directly with a microcontroller integrated within the prosthetic hand or socket.

In terms of this research project, the CV-RS should be first incorporated with a prosthetic hand and further with the EMG system. This is described in Chapter 4.

Chapter 4

Integration and validation of the shared EMG-Computer Vision system

In the previous chapters, I described an offline EMG Kinematic Prediction Models and a Computer Vision Real-time System. In this Chapter, I explain how the two systems are brought together into one continuous real-time system to control a prosthetic hand and wrist. The HDsEMG Kinematic Prediction Models developed in Chapter 2 are first integrated with a real-time framework. This framework allows for labelling of the EMG data and training of the models while the user follows the movement of a cursor on a screen. The Human Intention Detection system then utilises the trained models to allow the user to control a cursor on a screen via their real-time HDsEMG recordings. The position of the cursor is then used to control a prosthetic wrist and hand in 2-DoFs. Then, the CV-RS developed in Chapter 3 is integrated with this real-time framework to provide complementary information on the grasp type to be used. The information from both the Human Intention Detection system and CV-RS is sent to the prosthetic hand and wrist via a control framework which is described in Section 4.1.1.

The aims and realisation method of the complete integrated system are previously set out in Section 1.3.1. Integrating two real-time control systems into one shared semi-autonomous system is a challenging task. There are multiple systems running in parallel, at different rates, which need to communicate together and function as one in real-time. The decisions need to be made in a centralised

manner, with instructions sent to a prosthetic hand. This hand is controlled via 3 separate motor groups that are each controlled individually. 2 of these 3 motor groups do not make use of positional encoders, providing an additional challenge for our positional myoelectric control framework.

In Section 4.2, we then proceed to validate the myoelectric human intention detection system in real-time. In the study, 3 participants with trans-radial limb loss trained our Kinematic Prediction Models to control a cursor on a screen in both 2- and 3-DoF systems. The training movements consisted of individual and sequential movements in up to 2 DoFs. The system was then validated with target-reaching tasks on the screen in individual and combined DoFs for both systems. The control was also tested in the context of controlling a prosthetic hand. In Section 4.3 we test the complete shared control with both HDsEMG and CV to control a prosthetic hand.

4.1 EMG-CV controller integrated system design

The EMG-CV complete system diagram is shown in Figure 4.1. The Electromyography - Real-time System (EMG-RS) and CV-RS run in parallel. Acquisition of both EMG signals and image frames take place continuously and concurrently. Both inputs are processed to produce predictions - positions for each DoF are predicted with the EMG-RS while Grasp States are predicted with the CV-RS. Both predictions are converted such that they can be used to control the prosthetic hand and wrist via an Arduino micro-controller. Figure 4.2 shows an overview of all the equipment involved in the complete system.

The continuously running framework for the EMG control is based on a MATLAB (Matlab2019b, The MathWorks, Inc) platform previously developed in the research group [155] that we adapt for the purposes of this research. The EMG-RS is described in Section 4.1.2. The CV-RS runs on Python as described previously in Chapter 3. Both subsystems run on the same laptop (Windows 10, Intel Core i9 2.5GHz, 64 GB RAM). The grasp states are sent from the CV-RS to the MATLAB framework, for integration into the shared control, which in turn, sends a packet to the Arduino with information on both wrist and hand movement, and grasping. The Arduino controller then communicates with the TASKA hand and wrist.

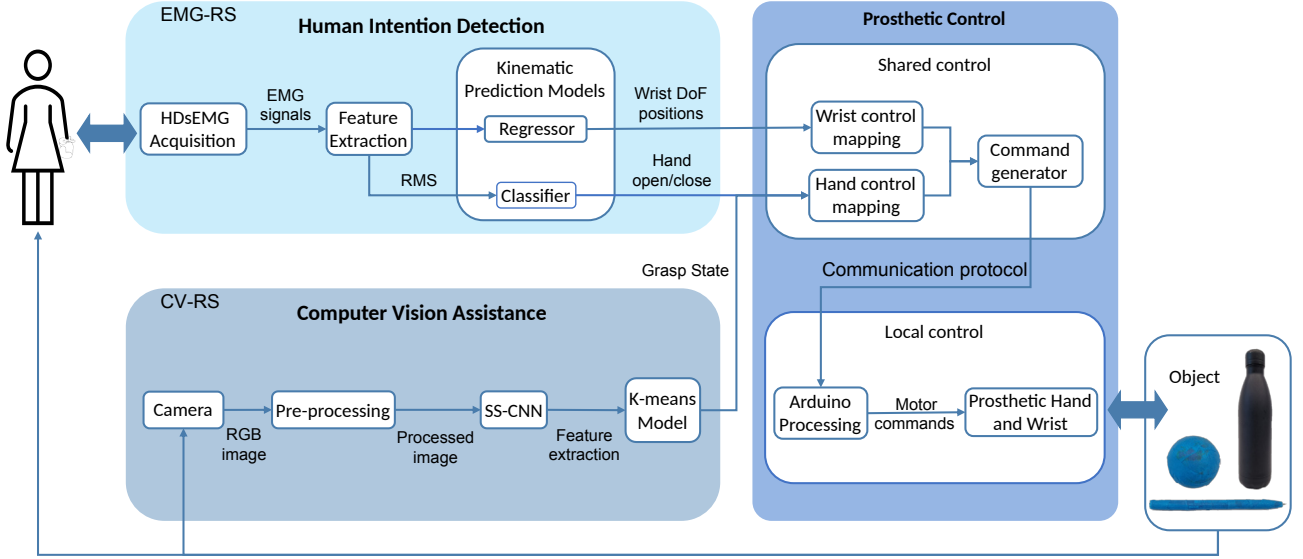


Figure 4.1: EMG-CV shared controller complete system diagram. The user and the camera both observe an object during a reaching task. HDsEMG is recorded from the user and processed to produce instructions for wrist positions and closing/opening the hand per the user’s intention. Visual input from the camera is processed to produce Grasp State outputs that are sent to the shared controller, where wrist and hand control mapping takes place. The commands for the prosthetic movement are sent to the local controller to generate movement in the prosthetic hand and wrist. The user then grasps the object via the prosthesis upon their intention to close the hand.

4.1.1 TASKA prosthetic hand

The TASKA HandGen1 (TASKA Prosthetics, New Zealand) is a commercially available myoelectric multi-articulating prosthetic hand.¹ One unique aspect of the TASKA hand is that the finger positions can be finely controlled and used to create a large range of grip types. The grip patterns available on the TASKA hand include: Relaxed, Open Palm, Keyboard, Don & Doff, Pointer, Key Grip, General Grasp, Apple, Tripod, Grab & Lock, Pincer, Hook, Flexitool, Precision Grasp, Grab & Go, Mouse, 1 Finger Trigger, 2 Finger Trigger, OK, and Rock On. In this research we utilise the four following grasp types to be predicted from the CV-RS, as shown in Figure 1.3: Pincer/Pinch, Tripod, General Grasp, Apple/Spherical. These are among the most commonly used grasp types according to [123–126].

Commercially, buttons are used to cycle through these grip types. Two sets of preset grips can be determined using an accompanying mobile phone application (MyTASKA, TASKA Prosthetics LTD) [59], which connects to the hand via Bluetooth. One set has 3 built-in grips loaded (General Grasp, Flexitool, Pincer) while the other has 5 grip types which can be customised by the user. These preset

¹We were generously loaned 2 research based TASKA HandGen1’s from TASKA Prosthetics, www.taskaprosthetics.com. One hand is in the left hand configuration and the other is in the right hand configuration.

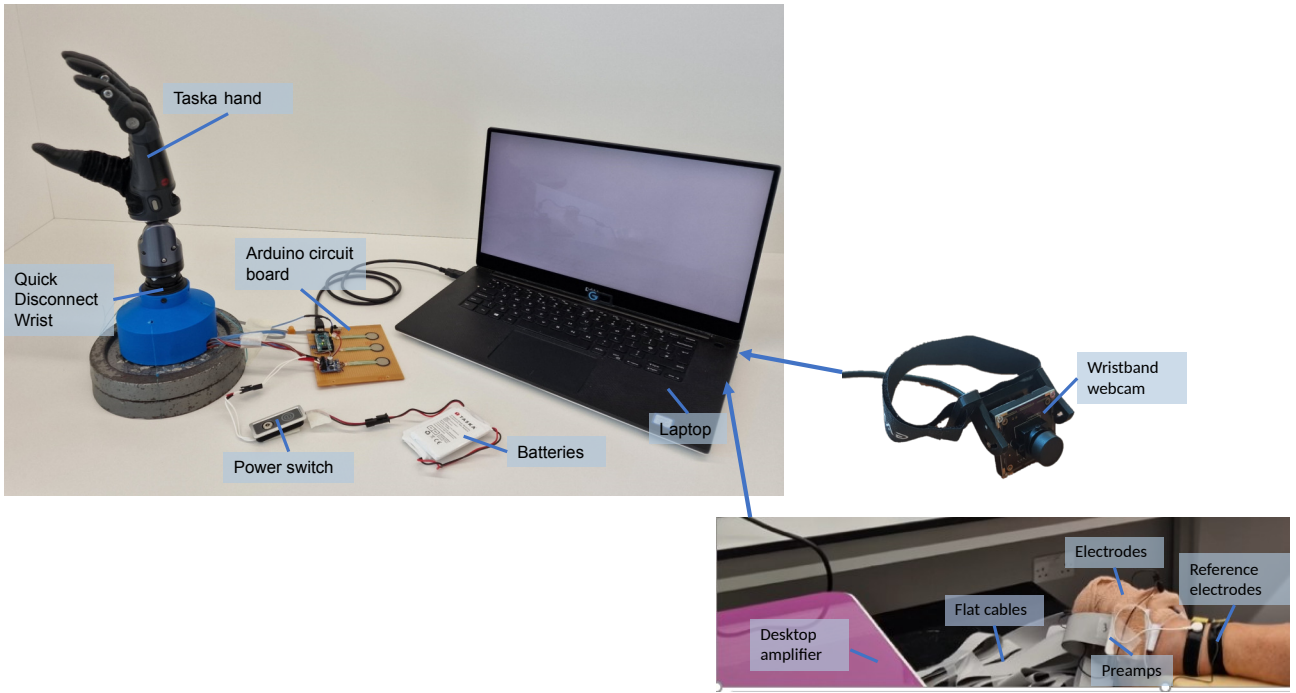


Figure 4.2: Overview of the equipment used in the EMG-CV shared control system.

grips can then be cycled through by pressing respective buttons on the back of the hand, one button for each set of grips. However, in our system, to change the grasp type, a command can be sent via the Arduino to the hand via serial communication, determining the grasp type to be used. A command can also be sent to determine the position of the hand on a scale of fully closed to fully open, where the fully closed and fully open positions are determined by the grip type selected.

In addition, the TASKA hands we are using are connected to two wrist units:

- a Quick Disconnect Wrist (QDW) that allows for wrist rotation. This is the unit that normally fits in a socket and then connects and disconnects from the TASKA hand. All the connections to the TASKA pass through this unit.
- Flexion-Extension wrist unit, the Motion Control Powered Flexion Wrist (Motion Control division of Fillauer, Utah). This unit is a fixed attachment to the TASKA hand, and the combination of the two is a research-based combination, not yet in production.

The prosthetic hand can be placed and attached to a 3-D printed support for the purpose of experiments, when it is not worn and attached to a socket. The whole device is powered by a 7.4V battery unit ². The hand and wrist system are shown in Figure 4.3.

²Battery provided by TASKA Prosthetics



Figure 4.3: TASKA hand and wrist units on the custom stand.

Wrist control

The two wrist units on the TASKA prosthesis are motors that are controlled separately and independently from each other and from the hand itself. The rotation at the QDW is controlled via the Arduino and a bridge driver. The direction of the motor can be switched, and the motor can be set to run at a certain Pulse-Width Modulation (PWM) for a certain length of time. The flexion-extension wrist motor can be switched on in a specific direction separately, and the motor can be set to run for a certain length of time before being switched off, via direct control from the Arduino. The motor can only function in a small range of voltages, but the speed can be changed manually using the Motor Control User Interface (MCUI) iOS application (Motion Control Inc.). As such, in the EMG-CV system a fixed speed of this motor is used.

The outputs from the Human Intention Detection system are positions for each DoF. However, neither of the wrist motors has positional encoders. This proved challenging for control, and additional

time would have been needed to design a more advanced and smooth control of the wrist motors. As such we created linear approximations of the systems such that change in position is estimated and converted into a length of time in which to turn each motor ‘on’, and in which direction. The speed of both motors is kept unchanged. The control of the TASKA wrists for rotation and flexion-extension were approximated into linear systems. The transfer functions were calculated experimentally. Three markers were placed on the TASKA hand and the relevant motor was set to run for specific amounts of time while video was recorded. The positions of the markers were found and tracked based on proximity and used to calculate rotation and flexion-extension angles. From here a line was fitted to the angle data vs time.

Local prosthetic control³

The local control for the TASKA hand and wrists passes through an Arduino circuit board. The Arduino receives a series of values over serial communication from the MATLAB framework running the EMG-CV shared controller, at a baud rate of 9600bps. The availability of a received input is regularly checked (every 30ms). The input packet received from MATLAB is a series of 2 byte integers representing flags for each of the possible DoF movements in the desired direction (grip change, closing, opening, flexion, extension, clockwise rotation, anticlockwise rotation), along with values representing the grip ID, and the value in ms for the duration of the desired movement.

If any flags are raised, then this triggers the start of a change or movement in the desired DoF and direction, before resetting the flag. For flexion, extension, and rotation, the task is ended according to the corresponding duration value for the raised flag(s), and the elapsed time since the start of the movement.

If a flag indicating a change of grasp type is raised, then the grip ID that is received, is set as the new current grip, which is applied directly via serial communication to the TASKA hand. If an instruction to close the hand is received, then the serial command to set the hand to the fully closed position in the indicated grip ID is sent via serial communication to the TASKA hand. This triggers

³The main circuitry and design was provided by TASKA Prosthetics, but was then modified to the system described in this section.

the hand to close fully, until an instruction to open the hand is received. In this case, the TASKA hand is sent, via serial communications, an instruction to fully open. A baud rate of 4800bps is used for the serial communications from the Arduino to the TASKA hand.

For the flexion-extension wrist control, flexion and extension are controlled separately through sending a fixed ‘high’ voltage (5V) to the relevant pins on the connector on the QDW via the respective Arduino Pin. A ‘low’ voltage (0V) is sent to stop the motor. For rotation, the direction of the motor is set through a bridge driver and the motor is activated at a single PWM level for the desired duration.

4.1.2 EMG-RS framework

Our research group has developed a real-time control framework on MATLAB for EMG recording and processing, training prediction systems, and controlling a cursor on a screen [155]. I adapted this framework for my research purposes, including extending the control to 4-DoFs, integrating my own prediction models, and communicating with the TASKA hand.

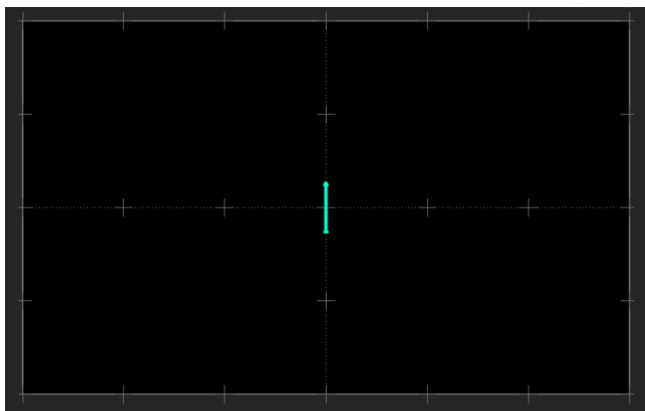
The framework consists of a GUI for the researcher that can be used to set up experimental parameters and run the framework. Alongside this is a GUI for the user of the control system. This user-GUI displays a cursor on a screen. The cursor is an arrow which moves along the x-axis to represent flexion-extension movements, and along the y-axis to represent radial-ulnar deviation (although we do not utilise this feature in our real-time system, but it is available if we do decide to use it). The arrow rotates around its centre to represent pronation-supination, and the arrow shrinks in size to represent closing the hand. Combinations of movements are also possible. The cursor is first used for training the prediction models - the cursor position provides a label that can be assigned to the relevant EMG envelope samples. Then, once the prediction models are trained, the the cursor position and size represent the predicted output.

While for validation with able-bodied participants, recording kinematics and EMG from the arm of interest is possible, this is not the case for real applications of the technology to prosthetic users. Even though we use kinematics for a more accurate algorithm development in Chapter 3, this is not possible in this scenario without using mirrored movements and recording from the intact anatomical

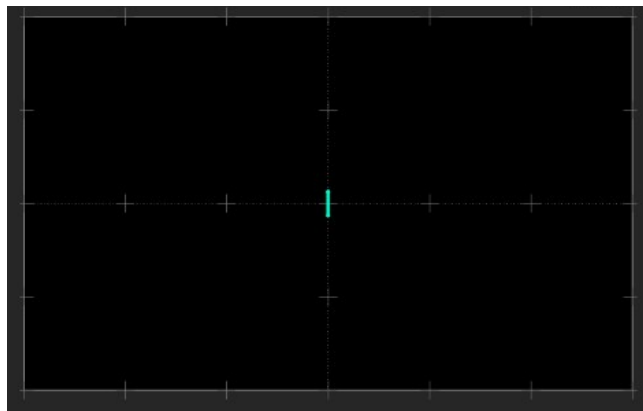
limb [33]. It is also impractical in an every-day scenario for a user to have access to motion capture or perform mirrored movements. As such, this cursor space is the space used for training and testing. The cursor space GUI is shown in Figure 4.4.

For the user to find the myocontrol predictable and intuitive, the interface should map the EMG inputs to the output control commands (cursor movements or prosthetic motions) in a way that is consistent and clear to understand [102]. To achieve this, the mapping should be close to linear and direct, and the system robustness should be ensured through the signal conditioning and training procedures [102]. This is also why the cursor mapping for each DoF is assigned to reflect the actual physiological DoFs as closely as possible. For example, we mentioned how the y-axis is unused in our cursor representation, but remains for representation of radial-ulnar deviation. In some systems this axis could be mapped to another function, e.g. closing/opening. However, this would provide a less intuitive interface for the user and the cursor space would be less complete.

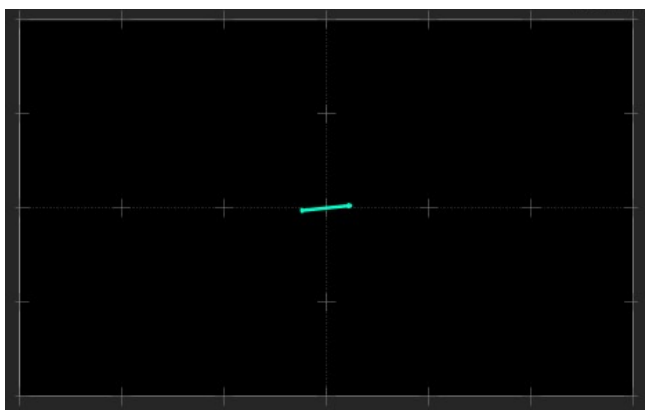
The real-time EMG capture and processing is performed very similarly to what is described in the offline system testing in Chapter 2, but is adapted for a real-time setting. The EMG-RS system flowchart is shown in Figure 4.5. In order to provide robust inputs to the motion estimation algorithm, the raw acquired EMG signal is filtered by a 5th order Butterworth bandpass filter in the frequency range of 20Hz-500Hz. Furthermore, in order to compensate for power-line interference, a 2nd order Butterworth band stop filter with cut-off frequencies 45Hz and 55Hz is applied to the signal. Then, the envelope of the filtered EMG signals is used as input to the regression and classification Kinematic Prediction Models. EMG amplitude envelope was selected as it is a robust, commonly used signal feature in myocontrol, which is linearly related to force [18]. As investigated in Appendix A, the RMS feature was found to be the most robust out of those compared. The envelope is extracted from the raw EMG by estimates of the RMS values in 126ms intervals with an overlap of 63ms. The interval is double that used in the offline analysis in Chapter 2; however the overlap allows us to retain a 63ms acquisition and processing time for use in real-time, while accessing a larger amount of data. No channel-wise normalisation is performed in order to avoid amplification of channels with low signal-to-noise ratio, which could negatively influence the controller. EMG channels that were considered



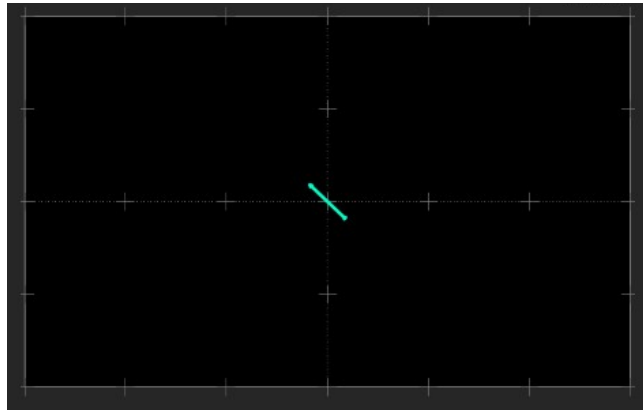
(a) Neutral cursor position



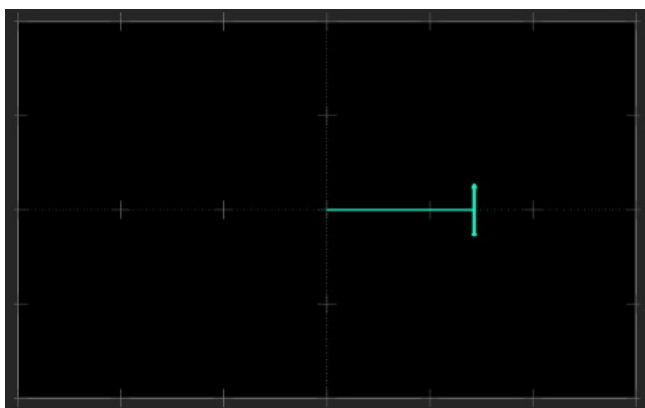
(b) Fully closed cursor indicates a fully closed hand



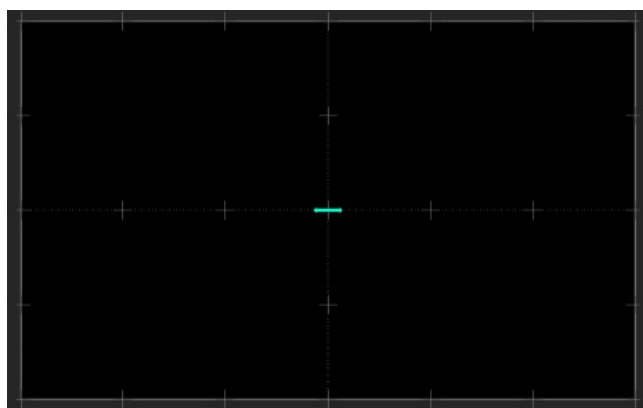
(c) Clockwise rotation of the cursor indicates pronation for the left arm or supination for the right arm



(d) Anticlockwise rotation of the cursor indicates supination for the left arm or pronation for the right arm



(e) Cursor translation to the right indicates flexion for the left arm or extension for the right arm



(f) Fully rotated and closed cursor

Figure 4.4: Graphic User Interface with cursor that can be either followed by the user for training or controlled by the user. The cursor is a blue arrow that moves in the x-axis space for flexion-extension, rotates to indicate pronation-supination, and becomes smaller to indicate hand closing.

noisy are set to 0.

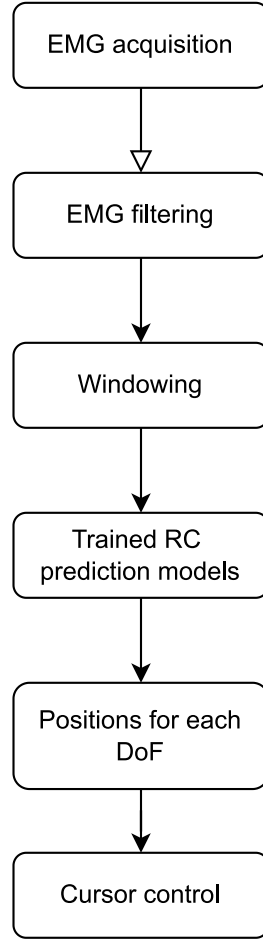


Figure 4.5: EMG Real-time System Flowchart. Acquired EMG is processed in real-time to produce positions for flexion-extension, pronation-supination, and hand closing/opening, which are used to control a cursor on the screen.

As in the developed Kinematic Prediction Modelling system in the offline setting, for the purpose of providing control over multiple DoFs, each DoF for the wrist control is assigned its own Linear Regressor, which is trained on both extends of the DoF, e.g. the DoF wrist rotation is trained on both wrist pronation and wrist supination. The DoF ‘grasping’ is assigned its classifier which predicts whether the hand is fully open/neutral/opening or fully closed/closing (as in Chapter 2). The movements are mapped in the cursor space to values between -100 to 100 with neutral being set to 0. Flexion-extension values extend to only 86 as the TASKA flexion-extension wrist unit is limited to 86 degrees in flexion. These values are also mapped to the TASKA control, as will be described in Section 4.1.3. These values are then normalised to a range of -1 to 1 for the target reaching tasks analysis. As described above, the cursor moves along the x-axis for flexion/extension, rotates

for pronation/supination and becomes smaller to indicate closing, and returns to size to indicate opening. In order to further improve user experience, a 2Hz low pass filter is applied to the predicted outputs from the regression, effectively reducing the control jitter. The classification outputs were also smoothed, through majority class selection over 3 windows. With the windowing overlap of 63ms, the perceived delay remains well under the real-time threshold of 300ms [156]. The filtering and smoothing parameters were chosen based on pilot trials. However, an improvement would be to adjust the filter based on individual subject preference and the specific training on the day. The predicted output updates the cursor on the screen as well as the prosthesis when it is connected.

4.1.3 CV-RS integration with EMG-RS framework

The EMG-RS system runs on the custom MATLAB framework as described in Section 4.1.2. The EMG-RS runs at intervals of 63ms, where new EMG data is read every 63ms and new position outputs for the DoFs of hand closing/opening, pronation-supination, and flexion-extension are generated every 63ms. The CV-RS system runs on the custom Python framework as described in Section 3.3. A grasp type is predicted and a new frame captured by the camera approximately every 0.08s. A new Grasp State is only generated when a new grasp type has been predicted for 2 frames in a row. This Grasp State needs to be passed on to the EMG-RS system as part of the combined EMG-CV framework.

Since both systems are run concurrently on the same laptop, User Datagram Protocol (UDP) is used to communicate between the two systems. Thus when a new grasp state is generated by the CV-RS system, the grasp state is sent to the EMG-CV conversion section. The EMG-CV thus continues to run at an acceptable processing rate of 63ms.

Shared control and command mapping

As we can see in Figure 4.1, the EMG-CV shared control framework integrates positional information received from the Human Intention Detection system for rotation-supination, flexion-extension, and closing/opening with information on the Grasp State from the EMG-CV. These are both combined and mapped into the format to be sent to the Arduino for TASKA control (as described in Section 4.1.1).

Hand control mapping If a new Grasp State is received from the CV-RS, then this new Grasp State is stored by the framework. The closing flag is raised when an EMG-RS output of fully closed is received (after the smoothing stage). The opening flag is raised when a classification of fully open is received. Once the flag for closing is raised, then the grip change flag is raised and the grasp ID is changed according to the grasp state stored. (In the Arduino, if the same grasp state is sent then nothing changes).

As an outcome of Chapter 3, we should consider the desired Grasp State to be the most frequent class predicted during the 10-frame window starting 15 frames before the intention to grasp is received. This has not yet been fully integrated. Currently, a grasp state is only sent to the MATLAB framework once a change is indicated. Instead, the system should be configured to send all the predictions to MATLAB, and incorporate the windowing selection with the shared control section. We can also choose to set the grip change flag to raise as soon as a new grasp state is received, without necessarily requiring the hand to close. For example, this was used for validation in Supplementary Video 3.1 and Supplementary Video 4.2.

Wrist control mapping The difference in flexion-extension angle δ_{fe} is calculated based on the input from the EMG-RS, comparing the current input to the input from the previous cycle. If $|\delta_{fe}| > 0$, then the flag for either flexion or extension is raised. The sign of the difference δ_{fe} indicates the direction of the movement. The activation time, t_{fe} of the movement, is then calculated as:

$$t_{fe} = \frac{|\delta_{fe}| - b_{fe}}{m_{fe}} \quad (4.1)$$

where m_{fe} is the slope of the flexion-extension motor transfer function and b_{fe} is a constant. The flexion-extension motor has a minimum activation time, because it is not in fact a linear system. In addition, to reduce jitters, we apply a threshold to remove very small movements. The threshold for each direction is variable and is applied based on the user control that session. Due to the lack of encoders and the approximation to linear control for a nonlinear system, the flexion-extension control of the TASKA is not ideal, and we reduce how much we rely on in our research. While not ideal,

removing the reliance on this DoF in our proof-of-concept system is acceptable because of the reduced performance in this DoF by the participants in Section 4.2.

Similarly, for rotation, the difference in angle δ_{rot} along the rotation DoF in the EMG-RS is calculated based on the cursor position. If $|\delta_{rot}| > 0$, then the flag for either clockwise or anticlockwise rotation is raised. The sign of δ_{rot} indicates the direction of the movement. The activation time, t_{rot} of the movement, is then calculated as follows

$$t_{rot} = \frac{|\delta_{rot}|}{m_{rot}} \quad (4.2)$$

where m_{rot} is the slope of the linear approximation. Again, a threshold is applied to remove small movements and reduce jittering. The threshold for each direction is variable and is applied based on the user control that session. All flags are reset within one processing loop.

To note, the EMG-RS can directly control the TASKA even when there is no input from the EMG-CV (not connected), as in Section 4.2. Likewise, the TASKA can be connected directly to the EMG-CV system separately, as in Chapter 3.

4.1.4 Socket design

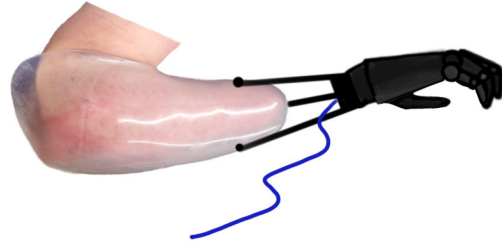
For a user to be able to use the prosthetic hand as a wearbale system, we need to attach it to a socket. As this thesis presents a proof-of-concept system, our aim was to create a simple connecting mechanism and socket. With support from Dorset Orthopaedic⁴, we devised a simplistic test socket to be used in further experimentation. Casts were taken of each of the participants' arms and used to create clear plastic (polypropylene) sockets, as in Figure 4.6a. The initial idea was to add supporting rods to connect the socket to the QDW, while leaving space for any wiring to connect to any external equipment. This is shown in Figure 4.6b.

However, this configuration reduced the number of EMG electrode arrays that we could use. To maintain consistency between the experiment in Section 4.2 and validation of the complete system in Section 4.3, we instead decided to use the Mitt sleeve from Koalaa (Koalaa Ltd., UK) [58] that utilises

⁴Dorset Orthopaedic is a Prosthetics and Orthotics company, and we were supported by their clinic in Surrey, UK



(a) Plastic test socket



(b) Initial prototype design

Figure 4.6: Initial Socket Design with a clear plastic test socket connected to the QWD for TASKA connection.

the BOA[®] system which is easily adjustable. The socket should thus be suitable for use with different participants, while still allowing for a high number of electrode grids. A connector was designed and 3-D printed to connect the sleeve to the TASKA at the QDW and flexion-extension wrist level⁵. This is shown in Figure 4.7. In this scenario, the flexion-extension capability is reduced. This is an early stage prototype for validating the system.

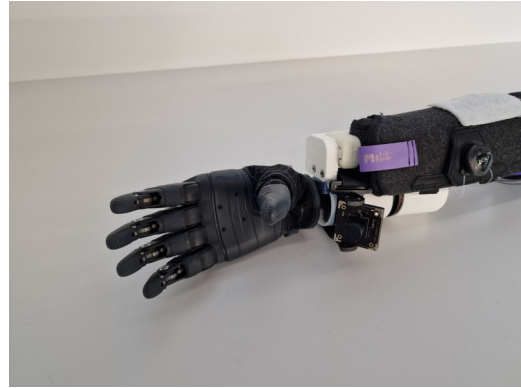
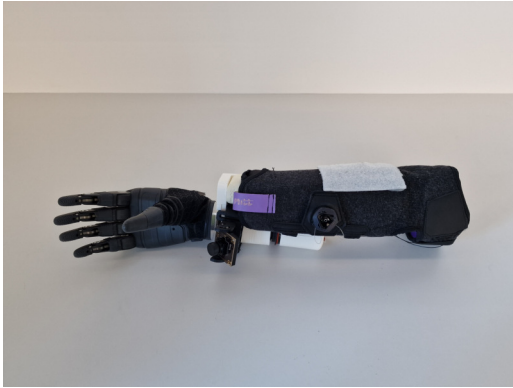


Figure 4.7: Proof-of-concept prototype design. The prosthesis is connected to the Mitt Sleeve via a custom 3-D printed connector.

4.2 Validation of the EMG-RS with participants with upper limb loss

In Chapter 2, we developed and tested an algorithm offline to map HDsEMG signals to kinematics of the wrist and hand, thus developing the Human Intention Detection system. In Section 4.1.2, we

⁵The connector was designed by my colleague Patrick Sagastegui Alva

described how this controller was integrated and adapted to be used in a real-time setting to control a cursor on a screen or a prosthetic hand. In this section, we will describe a study used to validate the performance of this Human Intention Detection system in real-time for people with limb loss, who are the intended target users of the controller.

In the offline experiment we were able to use motion capture to accurately define the kinematics of each movement for training of the controller. However, to adapt this to test the framework on participants with limb loss, real-time myoelectric control algorithms can be trained by having the user imagine performing a set of movements by following a cursor on a screen while the HDsEMG is recorded. The position of the cursor on the screen is used to label the EMG data for training. The trained algorithm can then be used to predict, in real-time, the user's intended movement of the cursor from the continuous HDsEMG signal. As such, we use the EMG-RS system as described above in Section 4.1.2.

Aims and hypothesis

As part of the study we performed a general HDsEMG assessment for each individual, to identify the activity areas on their residual limbs. This will influence the ideal electrode placement for further research and integration with a socket. We also aim to validate the use of the Kinematic Prediction Models for hand and wrist control in a real-time setting, both with cursor-based tasks on a screen, as well as in translation to the control of a prosthetic hand. We hypothesise that the users will be able to successfully control the cursor in at least 2 DoFs, wrist rotation and hand opening/closing.

Participant information

The experimental procedure was in accordance with the Declaration of Helsinki and approved by the Imperial College Research Ethics Committee (JRICO number: 18IC4685). Participants were recruited via advertising on relevant research networks. Individuals who reacted to the advert received general information about the experiment via email. All participants were guided through the experimental information and consent documentation, and consented to participate, before the experiment. We

recruited 3 participants (female, mean age 44.67 ± 17.67): 2 trans-radial unilateral amputees and 1 participant with congenital limb loss below the elbow. Participants visited over at least 2 sessions. The participant information is listed in Table 4.1, as measured in the first experimental session.

4.2.1 Experiment design and evaluation

The proposed Human Intention Detection system was tested in a real-time virtual environment. In the experiment, each subject was asked to perform a series of tasks involving their arm with the limb deficiency, while HDsEMG of the forearm muscles of the residual limb was recorded. The HDsEMG was used to control a cursor on a screen, and it was also used to assess the electrical activity at different location points of the arm. After training, the recruited study participants were asked to complete a number of target-reaching tasks requiring proportional articulation of up to two DoFs of the wrist and classification-based articulation of a power grasp, either individually or in combinations of DoFs.

Experimental setup

The main setup components consisted of a high-density acquisition device, a laptop used for computation and execution of the control algorithm, and an external screen used to show visual cues and the online prediction gestures to the participant. A TASKA prosthetic hand supported by a custom stand on the table was also used in some sessions. The experimental setup is shown in Figure 4.8.

High Density electrode grids were applied to the forearms of each participant, covering areas of perceived activity based on skin palpitations. Due to differences in limb size and shape, the electrode grid size, number, and placement was subject specific and varied between participants. The types of electrodes used are shown in Table 4.2. The electrode arrangements for each participant are shown in Table 4.3. There were only minor adjustments to the electrode arrangements between sessions. Conductive paste was used to improve the electrode-skin contact. Prior to electrode placement, the skin of the residual limb was cleansed with abrasive cream and ethanol. The electrodes were secured in place with medical tape. The monopolar HDsEMG signals were recorded, amplified, sampled at 2048Hz, A/D converted to 16 bits, and digitally band-pass filtered (10-500 Hz).

Table 4.1: Participant information for real-time HDsEMG control study.

	LL 1	LL 2	LL 3
Gender	Female	Female	Female
Age (years)	65	33	36
Arm with limb loss	Left	Right	Right
Dominant arm	Right	Ambidexterous	Left
Level of limb loss	Trans-radial	Trans-radial	Below Elbow
Reason for limb loss	Neurofibromatosis	Trauma	Congenital
Time since amputation (years)	6	8	N/A
Current Prosthetic Use	N/A	Open Bionics Hero Arm – 4 months	beBionic (broken) Cosmetic passive prosthetic
Previous prosthetic Use	Briefly used a body powered prosthetic hand.	N/A	Myoelectric / split hook / cosmetic
Length of residual limb (LE to end of limb) (cm)	23	21	10
Picture			

Table 4.2: Properties of the HDsEMG electrodes used in the experiments. Three different types of electrodes were used: GR10MM0808, GR08MM1305, GR04MM1305 (OTBioelettronica, Italy). i.e.d refers to interelectrode distance. The sensor material is Cu with chemical gilding. Electrode images from [157].

		
GR10MM0808	GR08MM1305	GR04MM1305
64 electrodes	64 electrodes	64 electrodes
10mm i.e.d	10mm i.e.d	4mm i.e.d
8x8	5x13	5x13



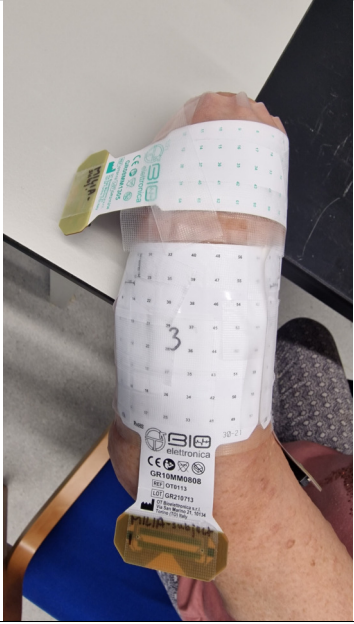
Figure 4.8: Experimental Setup for offline validation of EMG-RS. High Density electrodes on the participant’s arm are connected to the Quattrocento amplifier. The EMG signals are sent to the laptop where they are processed and used to control a cursor on the screen, as well as the TASKA hand, which is shown here supported on the table.

The acquisition of the monopolar HDsEMG channels was realised through the custom-made MATLAB framework (as described in Section 4.1.2) using the Quattrocento amplifier (OTBioelettronica, Italy). Reference electrodes were placed on the elbows of the participants on the same arm as the grid electrodes.

4.2.2 Experimental protocol

Based on user surveys [6], and the fact that the prosthesis setup we are using does not have the capability for ulnar-radial deviation on top of the current functionality, six gestures along 3 DoFs of the hand and wrist were selected for investigation. These gestures are: neutral state, power grasp, wrist flexion, wrist extension, pronation and supination, as well as combinations of different gestures. As the participants are missing a wrist or hand, all movements are imagined, while moving the necessary

Table 4.3: High Density surface EMG electrode grid configurations for each participant.

Participant	LL 1	LL 2	LL 3
Electrodes	3x GR10MM0808 1x GR08MM1305	2x GR10MM0808 2x GR08MM1305	2x GR08MM1305 1x GR04MM1305
Picture			
Number of channels	192	192	192

muscles, as they would have to naturally perform the movement. The cursor on the screen is used to help guide the subject. The experiment was divided into two main protocols – one for 3-DoF control and one for 2-DoF control, as these required separate training of the Kinematic Prediction Models:

Protocol 1: Assessing training and control with 3 DoFs (wrist flexion-extension, pronation-supination, closing/opening):

- Train on individual DoFs (Training Task Set 1) and sequential DoFs (wrist followed by hand movements) (Training Task Set 2).
- Test 1.1: Virtual target reaching task with EMG Control, 3-DoF, individually tested.
- Test 1.2: Virtual target reaching task with EMG Control, 3-DoF, combinations of DoFs tested.

Protocol 2: Assessing training and control with 2-DoFs (pronation-supination, closing/opening):

- Train on individual DoFs (Training Task Set 1) and sequential DoFs (wrist followed by hand movements) (Training Task Set 2).

- Test 2.1: Virtual target reaching task with EMG Control of a prosthesis, 2-DoF individual and combined movements.
- Test 2.2: EMG Control of a prosthesis, 2-DoF individual and combined movements, while following video prompts.
- The testing was then followed by a questionnaire to gain insight into the participants' use of the system.

The training and testing protocols are described in more detail below. The survey questions can be found in Figure 4.13. Due to scheduling difficulties, tiredness levels of participants, and the continual development and evolution of the system, not all participants performed all tasks. A breakdown of the testing protocols performed by each participant is shown in Table 4.4

Training

In order to train the Kinematic Prediction Models, the subjects were first instructed, using visual cues, to perform 4 repetitions of the desired gestures at a strong, but comfortable level. The recordings were then synthetically labelled denoting the cued gesture (no kinematics or kinetics were acquired), as for use with persons with limb loss. During these contractions, all the EMG data was used for training, including the dynamic movements, the held contractions, and the rest periods.

Training Task Set 1 - tasks in 1-DoF of the wrist or hand The first task set involved the subject imagining moving their wrist or hand from neutral to the required position in one DoF, holding

Table 4.4: Summary of the different experimental protocols performed with each participant. LL represents participant with Limb Loss.

	Test 1.1	Test 1.2	Test 2.1	Test 2.2	2 DoF control survey
LL1	Done		Done x2 (No ± 0.6 level pro-sup), once with TASKA and once without.		Done
LL2	Done		Done x 3	Done x 2	Done
LL3	Done	Done	Done x 3	Done x 2	Done

it there, and moving it back to neutral, resting there before moving to the other extreme of the DoF. This was done following a cursor on a screen with a movement transition time of 2s, a contraction period of 3s, and 2s of rest in between directions. The movement was repeated 4 times, with repetitions recorded separately, with rest between each repetition, as required by the participant. An example of one repetition is shown in Figure 4.9. The DoFs included in Protocol 1 were: flexion and extension of the wrist, pronation and supination of the forearm, and closing and opening the hand. The DoFs included in Protocol 2 are: pronation and supination of the forearm, and closing and opening the hand.

Training Task Set 2 – closing the hand in different positions For the second training task set, the subject was asked to imagine closing their fingers (general grasp) while the hand is in different positions. They first imagined moving the wrist to the extreme position of one DoF, for example flexed wrist, and then closed the hand while the wrist was flexed. They held the closed position, before opening the hand again, all while maintaining the flexed (or other starting) position. An example of one repetition is shown in Figure 2.5. This is done following a cursor on a screen with a movement transition time of 2s to move the wrist, followed by 2s transition for closing the hand and a contraction period of 3s, followed by 2s to open the wrist and returning the wrist to the neutral position over a period of 2s. There were 4 repetitions of the movement, with rest between each repetition as required by the participant. The wrist movements for the initial position were the following for protocol 1: flexion and extension of the wrist, and pronation and supination of the forearm. The initial DoFs included in protocol 2 were: pronation and supination of the forearm.

Testing To test the control algorithm, the user is tasked with controlling a cursor on a screen in target reaching tasks in an abstract fashion as in previous studies [102, 158]. The subject is instructed to complete each test as quickly and as naturally as possible. A test requires the user to move the cursor from a neutral position to a randomly ordered target location. The targets are ellipses drawn 8 dpi around the target cursor position. The test is considered complete when the user successfully places and keeps their cursor within the target range for 0.3s. Leaving the target area led to a reset

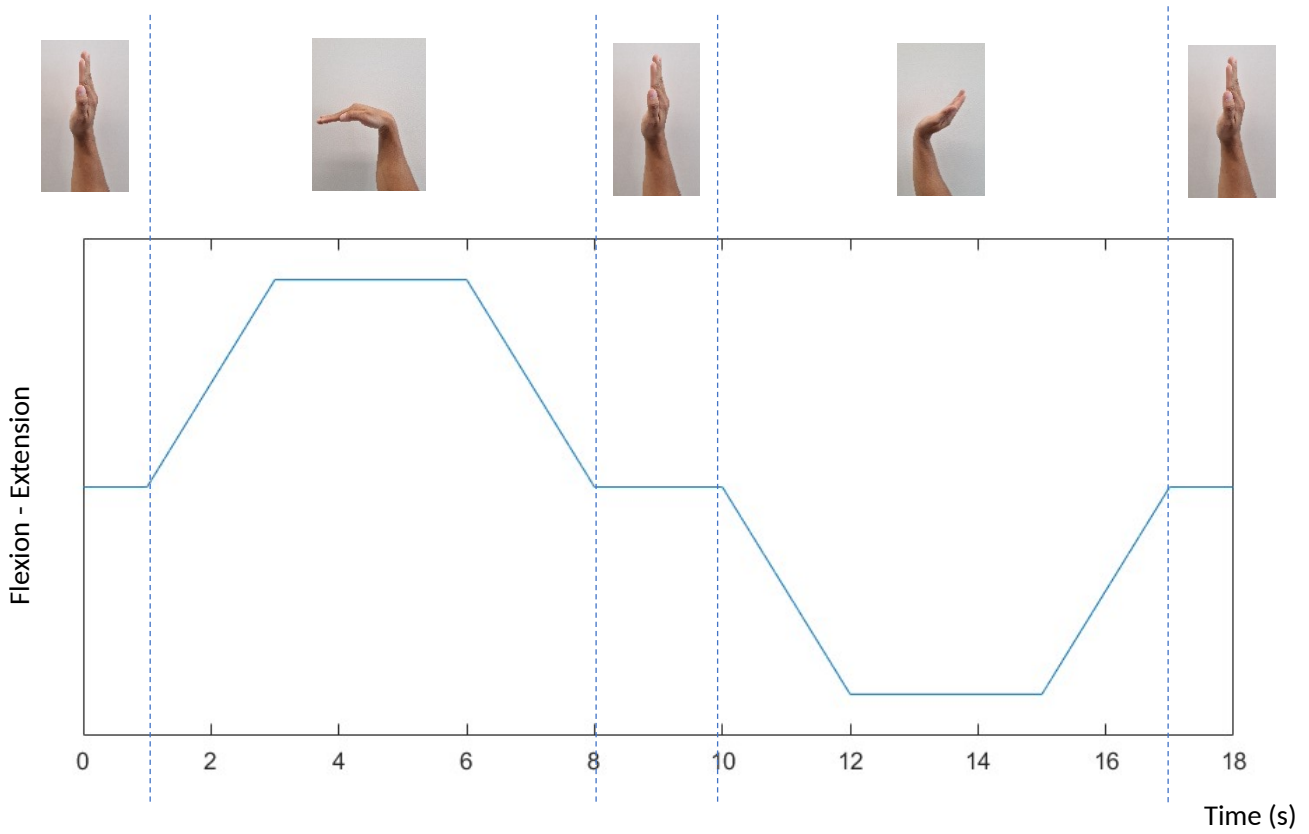


Figure 4.9: Example of 1 repetition of a training movement, here flexion and extension of the wrist.

of this dwell time. If unsuccessful after 20s, the test times out and is considered incomplete, and the next target is displayed. Reaching the target area is visually indicated to the participant by a change of colour of the target from magenta to yellow. The colour changed back to magenta if the user left the target area or changed to green if the user was successful in holding the arrow within the target area for the required 0.3s. In case the task was not successfully accomplished and the timeout of 20s was reached, the colour of the target changed to red and the user was instructed to fully relax before resetting for the next task.

In test set 1.1, the targets consisted of the following DoFs, individually tested, at different distances and sizes for a total of **18** targets: flexion and extension of the wrist, pronation and supination of the forearm, closing and opening the hand. In test set 1.2, the same DoFs were tested, but in combinations of 2 or 3 DoFs simultaneously, for a total of **22** targets. Only subject LL3 performed test set 1.2.

In test set 2.1, the targets consisted of the following DoFs, tested individually and in combination, at different distances and sizes for a total of **12** targets: pronation and supination of the forearm, closing and opening the hand. The TASKA hand was also connected and controlled at the same time,

but the participants were asked to focus on the cursor on the screen as the main feedback for control. Here the TASKA hand was placed on a custom-made support on the table. Subject LL2 did not perform 4 of the targets in the task set (± 0.60 pronation-supination with and without closing) and only repeated the test twice, on separate occasions, the first of which did not include the TASKA hand connected. Subjects LL2 and LL3 repeated the test set 2.1 for 3 separate times, with a different random order of targets each repetition. In test set 2.2, the participant was asked to follow a video on a screen along with prompts from the researcher while the only feedback for their EMG control was the TASKA hand. Participant LL1 did not perform test set 2.2.

An added difficulty to the testing was included by introducing intermediate levels of the actions. This was done for both individual gestures, e.g. 40% wrist pronation, as well as for combinations, e.g. 40% wrist pronation combined with 70% flexion. The targets with the highest difficulty were those which involved a combination of gestures along all three DoFs. A total of 196 target reaching tasks were recorded across participants and sessions. A full breakdown of training actions and testing goals with further categorical information is listed in Table 4.5. Here the values for each DoF are normalised in the range -1 to 1, where 0 is the neutral position. *TargetNoDoFs* is a factor with three levels which denotes the number of DoFs involved in a task (1 stands for individual gesture goals, while 2 and 3 stand for combinations of 2 or 3 gestures in a goal). The Activation level (*ActLvl*) is a summation of the levels of activation for each gesture involved in a goal. This parameter indicates the extent to which each gesture involved in the task should be activated. *NoSamples* represents the number of samples recorded for each target based on subject participation. Figure 4.10 also shows all the different targets in the cursor space.

Performance measures

The primary performance measure was *Success*, which determines whether the participant is able to successfully complete the task of reaching a goal within the required timeframe. Thus, the *Completion Rate* is the percentage of successful tasks. Successful tasks were further analysed with five secondary performance measures, which build upon measures from the Target Achievement Control test and

Table 4.5: Complete set of training and testing tasks, with further categorical information for the target reaching tasks.

		Legend		Used in both testing paradigms			
				Used only in 3DoF training/testing			
				Used only in 2DoF training/testing			
		Flex-Ext	Sup-Pro	Grasp	TargetNoDoFs	ActLvl	NoSamples
Training data	Set 1 – 3DoF	1.00	0.00	0.00	-	-	-
		-1.00	0.00	0.00	-	-	-
		0.00	1.00	0.00	-	-	-
		0.00	-1.00	0.00	-	-	-
		0.00	0.00	0.00	-	-	-
		0.00	0.00	-1.00	-	-	-
		1.00	0.00	-1.00	-	-	-
		-1.00	0.00	-1.00	-	-	-
	Set 2 – 2DoF	0.00	1.00	0.00	-	-	-
		0.00	-1.00	0.00	-	-	-
		0.00	0.00	0.00	-	-	-
		0.00	0.00	-1.00	-	-	-
		0.00	1.00	-1.00	-	-	-
		0.00	-1.00	-1.00	-	-	-
Testing data		0.47	0.00	0.00	1	0.47	3
		0.70	0.00	0.00	1	0.70	3
		0.93	0.00	0.00	1	0.93	3
		-0.47	0.00	0.00	1	0.47	3
		-0.70	0.00	0.00	1	0.70	3
		-0.93	0.00	0.00	1	0.93	3
		0.00	0.40	0.00	1	0.40	11
		0.00	0.60	0.00	1	0.60	9
		0.00	0.80	0.00	1	0.80	11
		0.00	-0.40	0.00	1	0.40	11
		0.00	-0.60	0.00	1	0.60	9
		0.00	-0.80	0.00	1	0.80	11
		0.00	0.00	-1.00	1	1.00	11
		0.00	0.00	-1.00	1	1.00	11
		0.00	0.00	-1.00	1	1.00	11
		0.00	0.00	-1.00	1	1.00	11
		0.00	0.00	-1.00	1	1.00	3
		0.00	0.00	-1.00	1	1.00	3
		0.00	0.40	-1.00	2	1.40	8
		0.00	0.60	-1.00	2	1.60	7
		0.00	0.80	-1.00	2	1.80	8
		0.00	-0.40	-1.00	2	1.40	8
		0.00	-0.60	-1.00	2	1.60	7
		0.00	-0.80	-1.00	2	1.80	8
	only performed by 1 subject	0.47	0.60	0.00	2	1.07	1
		0.47	-0.60	0.00	2	1.07	1
		-0.47	0.60	0.00	2	1.07	1
		-0.47	-0.60	0.00	2	1.07	1
		0.93	0.60	0.00	2	1.53	1
		0.93	-0.60	0.00	2	1.53	1
		-0.93	0.60	0.00	2	1.53	1
		-0.93	-0.60	0.00	2	1.53	1
		0.47	0.00	-1.00	2	1.47	1
		-0.47	0.00	-1.00	2	1.47	1
		0.93	0.00	-1.00	2	1.93	1
		-0.93	0.00	-1.00	2	1.93	1
		0.47	0.60	-1.00	3	2.07	1
		0.47	-0.60	-1.00	3	2.07	1
		-0.47	0.60	-1.00	3	2.07	1
		-0.47	-0.60	-1.00	3	2.07	1
		0.93	0.60	-1.00	3	2.53	1
		0.93	-0.60	-1.00	3	2.53	1
		-0.93	0.60	-1.00	3	2.53	1
		-0.93	-0.60	-1.00	3	2.53	1

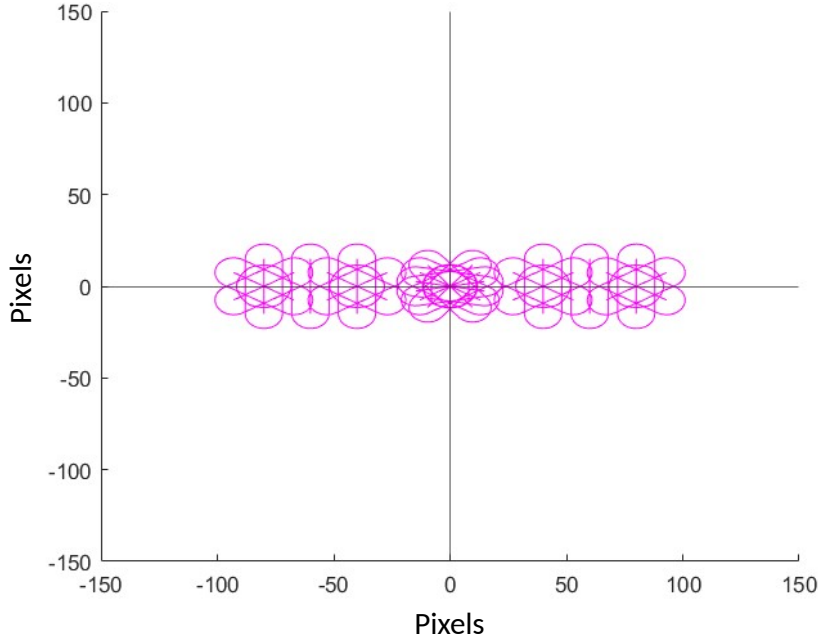


Figure 4.10: All targets shown in the cursor space.

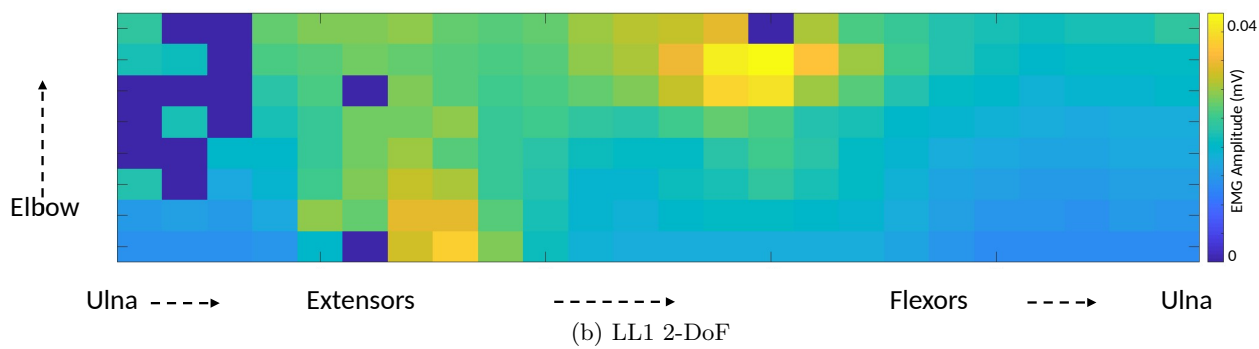
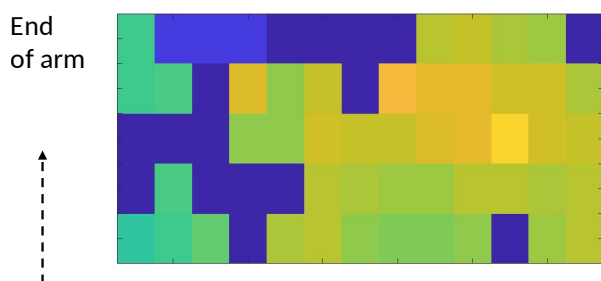
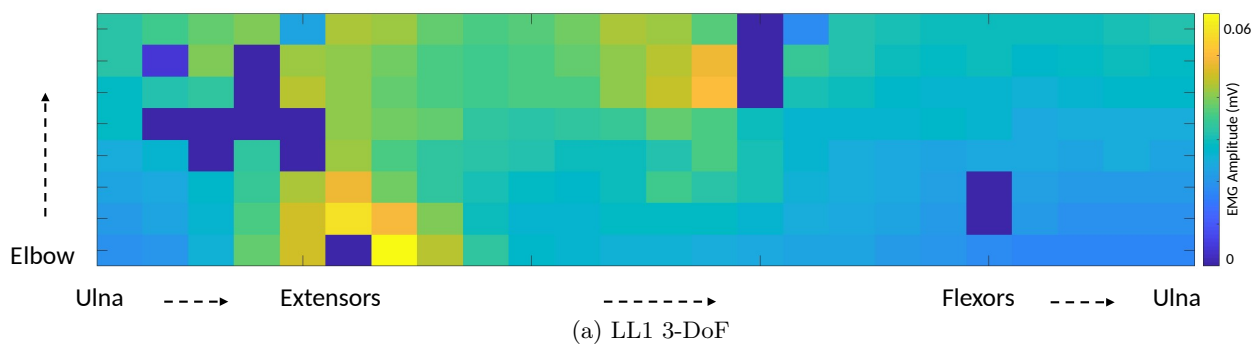
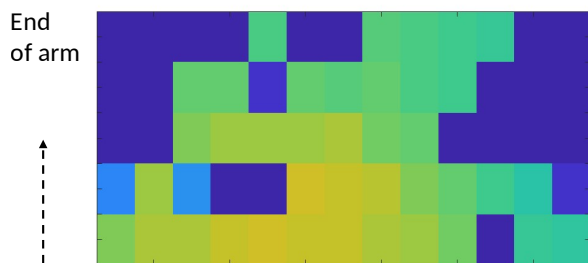
Fitt's Law testing [158, 159]: *Task completion time*, *Speed*, *Overshoot*, *Time to first hit*, and *Path Efficiency*. *Task completion time* indicates the time it takes the user to successfully accomplish a task, *Speed* is the ratio between the length of the path travelled and *Task completion time*, and *Path Efficiency* is the ratio of the length of the shortest path from start (origin of the GUI shown in Figure 4.4) to the endpoint (target location) and the length of the path actually taken by the user. The distances are Euclidean distances calculated based on the normalised values of the DoF positions for each target. *Overshoot* is the number of passes through a target prior to a successful hit, while *Time to first hit* is the time taken to reach the target for the first time.

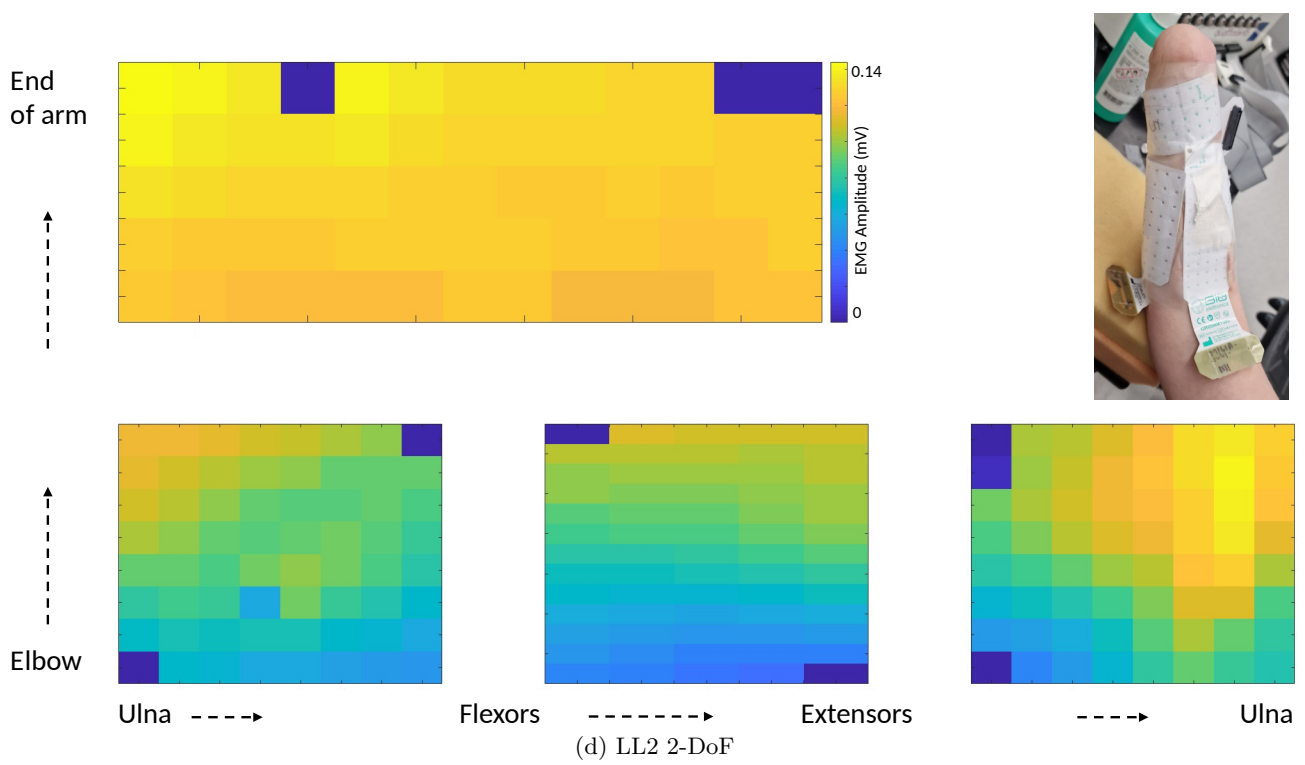
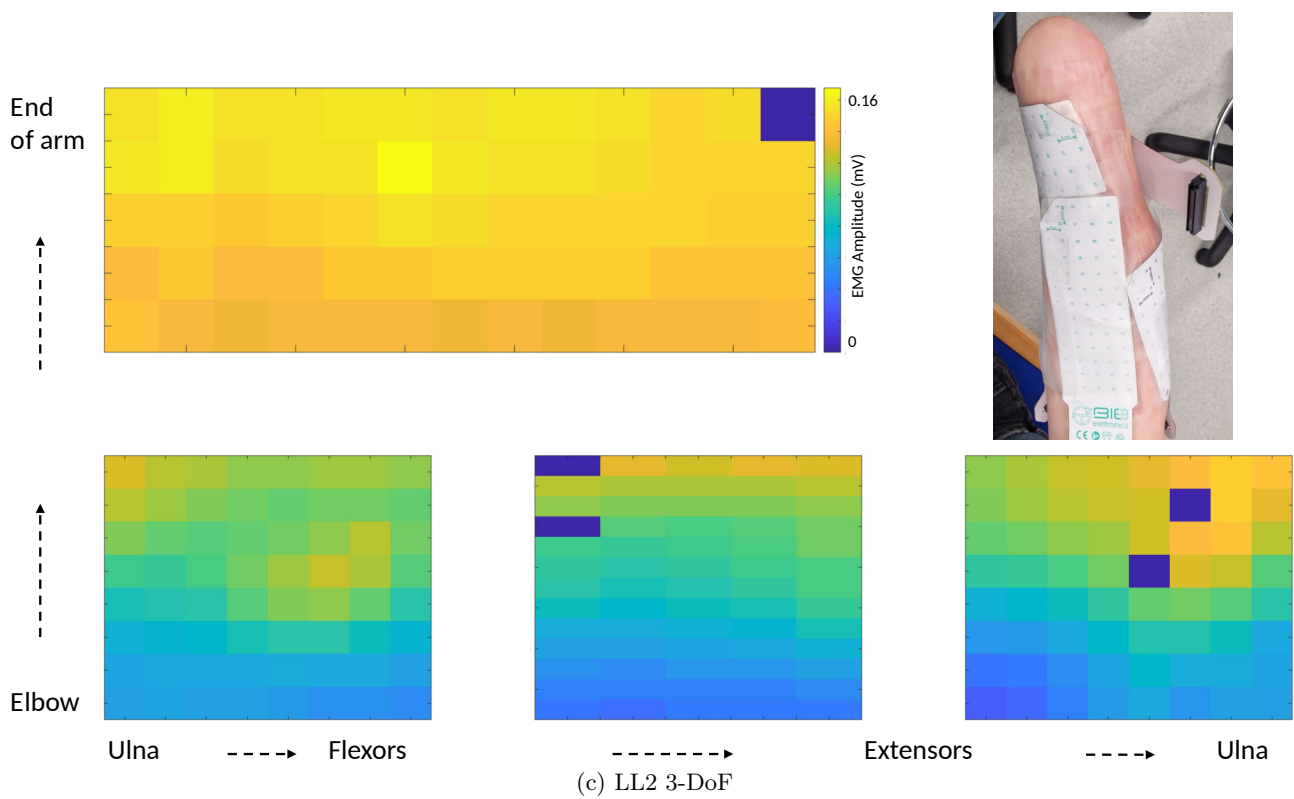
4.2.3 Results

EMG data

Mapping the EMG activity across the arm for each individual subject is important to ensure that we are covering areas of activity, especially when we eventually reduce the matrices in order to accommodate a socket. Figure 4.9 shows the average activity over the held contraction activities during training over each channel.

Clearly each participant has a unique anatomy, which produces different activity levels across the





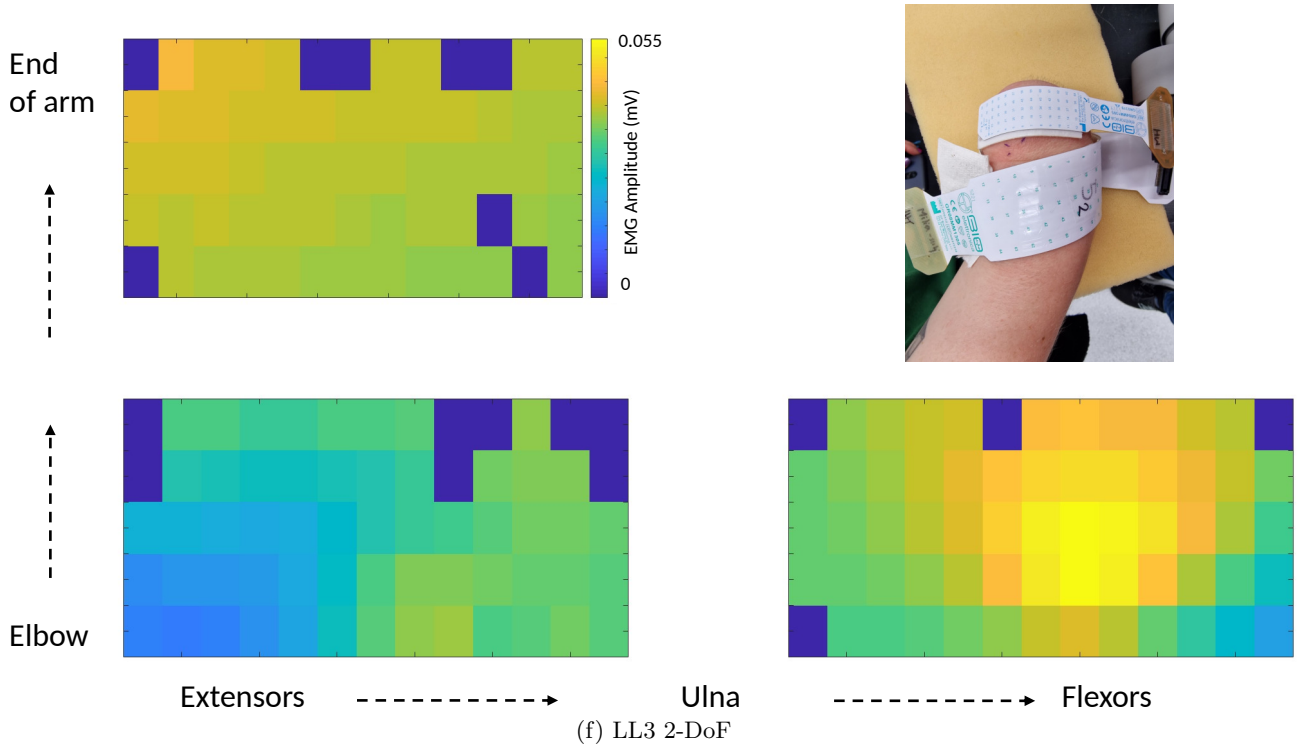
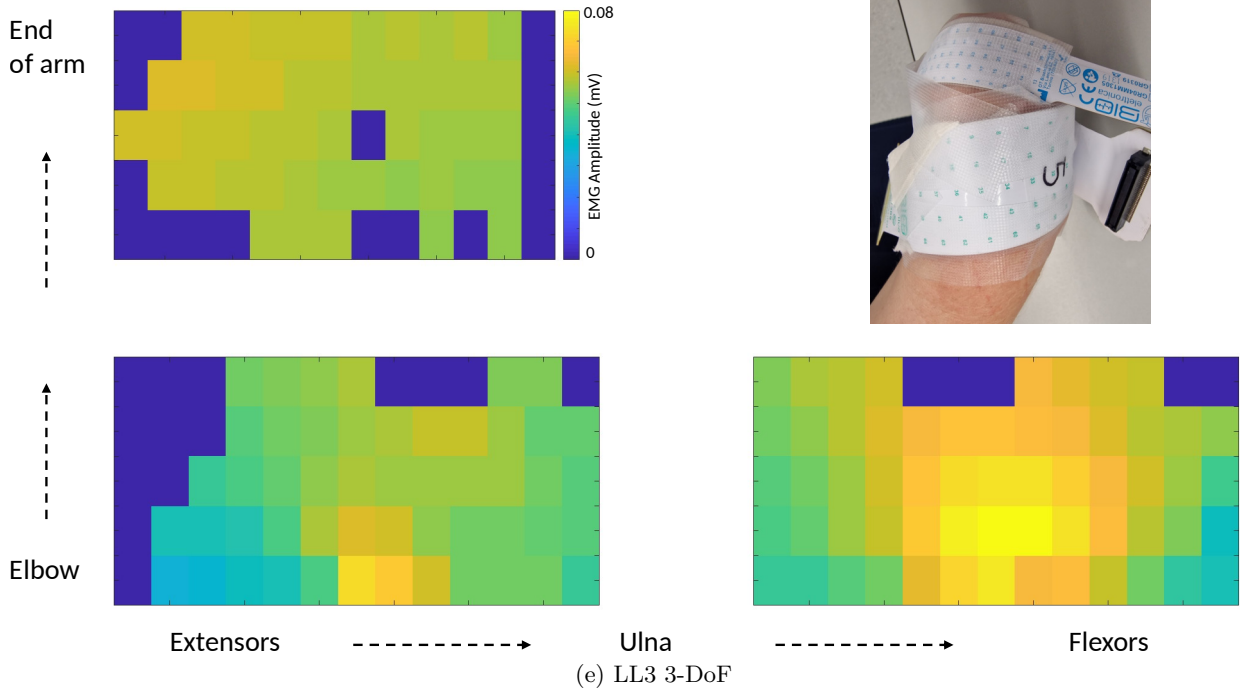


Figure 4.9: Heatmaps of the mean EMG activity during held contractions during training tasks for each participant, for both the 3- and 2-DoF control systems. The EMG channels are arranged according to the grid arrangements. Channels with 0mV activity are those that have been manually removed to due noisy signals.

forearm. The shape of the arm also affects how well the electrode grids attach to the skin, and there are clear areas of electrode detachment. LL1 has more scar tissue and more skin that does not directly overlay over any muscles. She has the lowest overall activity level, while LL2 has the highest (0.06mV highest average EMG amplitude for LL1, 0.08mV for LL3 and 1.6mV for LL2). LL2 has increased activity more distally on the arm. More proximally on the anterior side there is an area of skin with limited electrical muscular activity. LL3 has clear distinct areas of activity. These extend proximally, however the proximal region is difficult to capture activity from due to the nature of the shape of the arm. As the nature of the limb loss is congenital, the muscular activity may not directly map to movements as would be expected in an able-bodied person without limb loss.

Target-reaching tasks

The performance metrics described above were calculated for each of the target-reaching tasks. The overview of observed performances across all measures grouped by the number of target DoFs is shown in Figure 4.10. The overall completion rate with respect to the number of DoFs included in the task is displayed as a bar plot. These are further separated based on whether the control system is trained and tested on 2 or 3 DoFs – i.e. protocol 2 or protocol 1. At the same time, individual measures for *Task completion time*, *Speed*, *Path Efficiency*, *Overshoot*, and *Time to first hit* for successful tasks are shown as dots on top of box plots indicating the distribution of the data.

The *Completion Rate* for the 3-DoF control system was 48% across all targets, with a lower performance for 1-DoF tasks. However, the *Completion Rate* for the 2-DoF control system is 100% across all tasks. The performance over all metrics is generally worse for the 3-DoF control system than the 2-DoF system, although the speed is faster for the 3-DoF system.

In order to gain insight into different aspects of the observed real-time performance of the system, further factors were compared for protocol 2 (2-DoF trained system). The focus on the 2-DoF system stems from the larger number of data points available for comparison. The *Completion Time*, *Speed*, and *Path Efficiencies* were compared over different activation levels in Figure 4.11. *ActLvls* less than 1.0 represent individual movements in pronation-supination. *ActLvls* equal to 1.0 represent grasping targets. *ActLvls* greater than 1.0 represent movements in combination of grasping and wrist rotation at different levels of activation. Generally, the task performance worsens with increased *ActLvl*, with the exception of the grasping only tasks (*ActLvl* = 1.0).

Figure 4.12 shows the individual measures for *Task completion time*, *Speed*, *Path Efficiency*, *Overshoot*, and *Time to first hit* for successful tasks based on the performance of the individual participants with the 2-DoF system, separated over the number of target DoFs. It is difficult to identify meaningful differences between the performance of all 3 participants in these tasks.

Performance of one participant with the direct control of the TASKA hand in 2-DoF (Test 2.2) is provided in Supplementary Video 4.1.

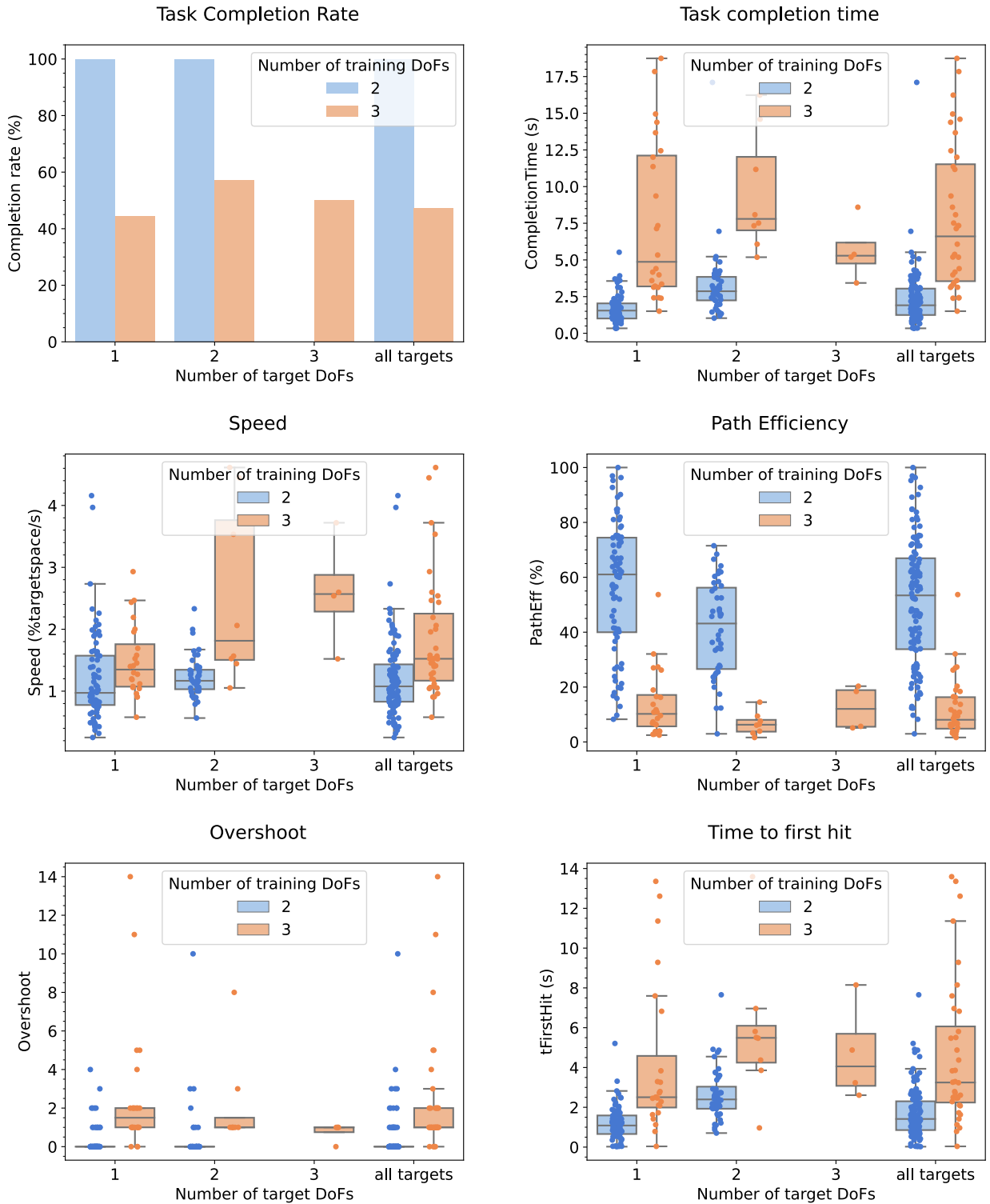
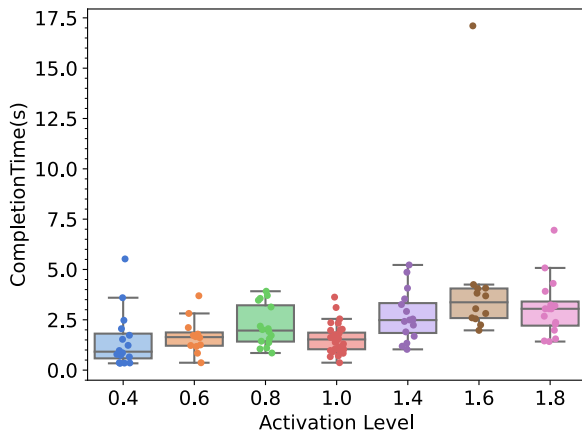
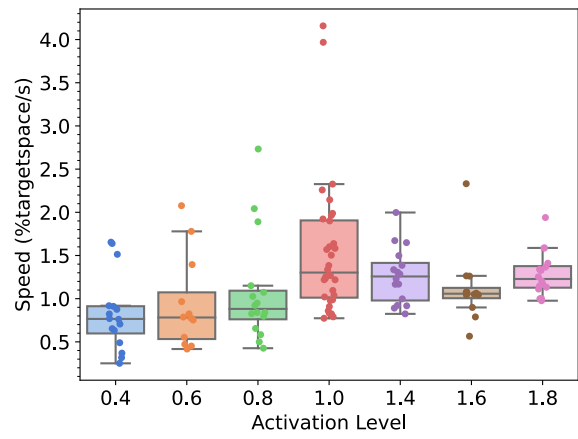


Figure 4.10: Overview of the observed real-time target-reaching task performance across 6 performance metrics separated by number of target DoFs and the number of DoFs involved in training the system. All metrics besides completion rate are based on successful attempts.

Task Completion time vs ActLvl for 2DoF training tasks



Speed vs ActLvl for 2DoF training tasks



Path Efficiency vs ActLvl for 2DoF training tasks

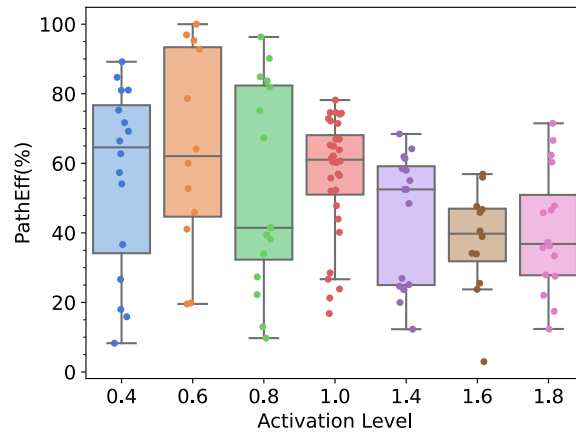


Figure 4.11: Observed real-time target-reaching task performance across 3 metrics separated by Activation Level (*ActLvl*) for tasks in the 2-DoF trained system.

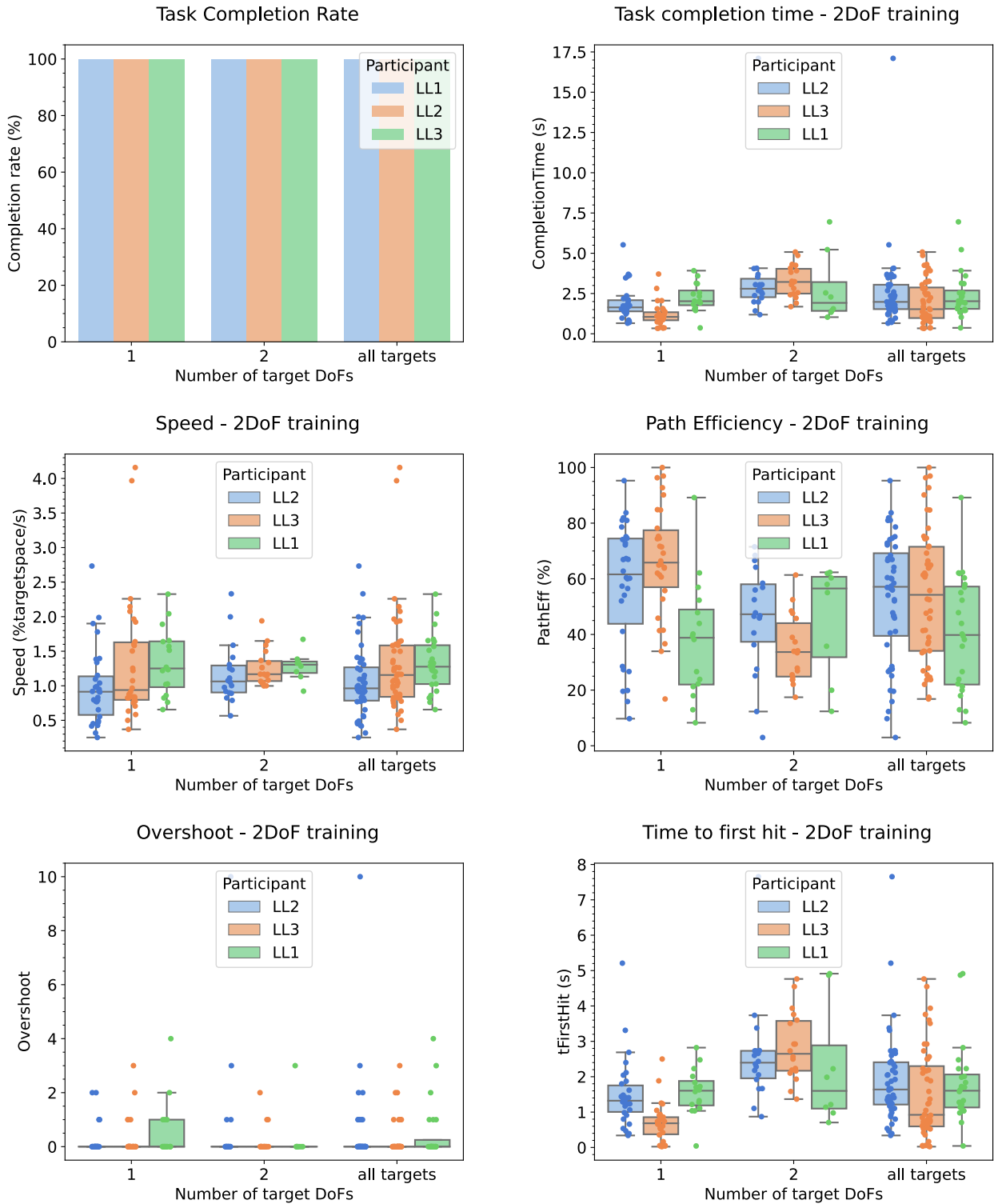


Figure 4.12: Overview of the individual participant performance observed real-time target-reaching tasks in the 2-DoF system across 6 performance metrics separated by number of target DoFs. All metrics besides completion rate are based on successful attempts.

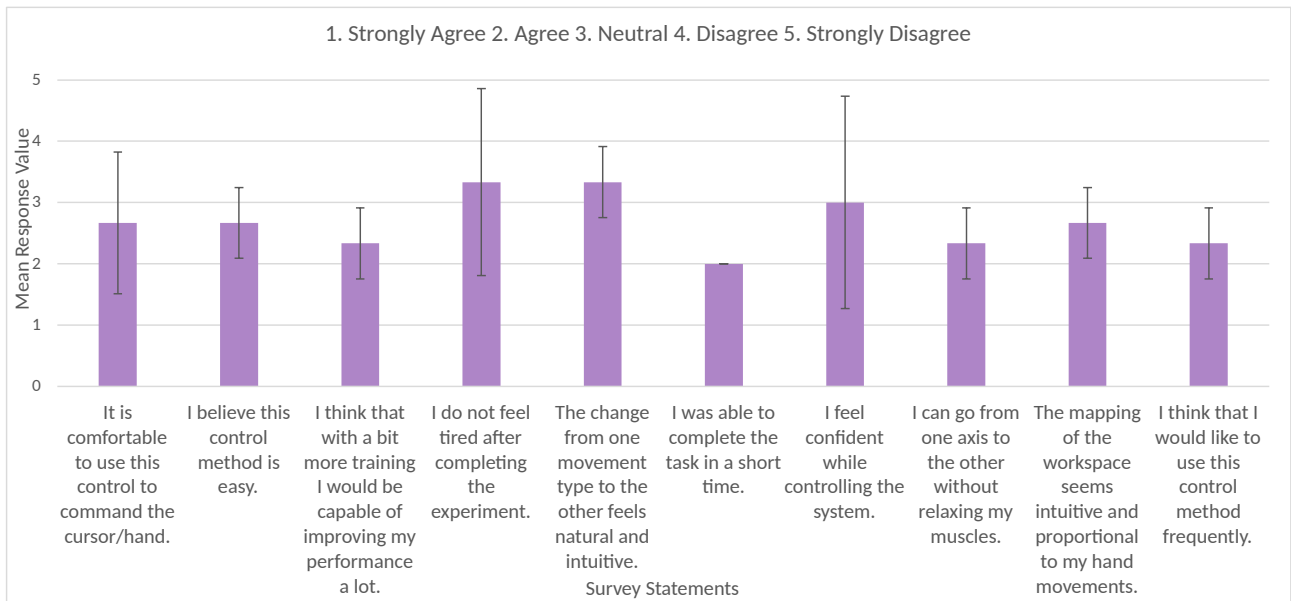


Supplementary Video 4.1: A participant with limb loss at the transradial level controls the TASKA hand via HDsEMG and the Kinematic Prediction Models of the Human Intention Detection System. The intended movements are displayed on the screen, but the participant is following instructions from the researcher resulting in a perceived delay. Video available at: <https://youtu.be/xa-9GJdbfBI>

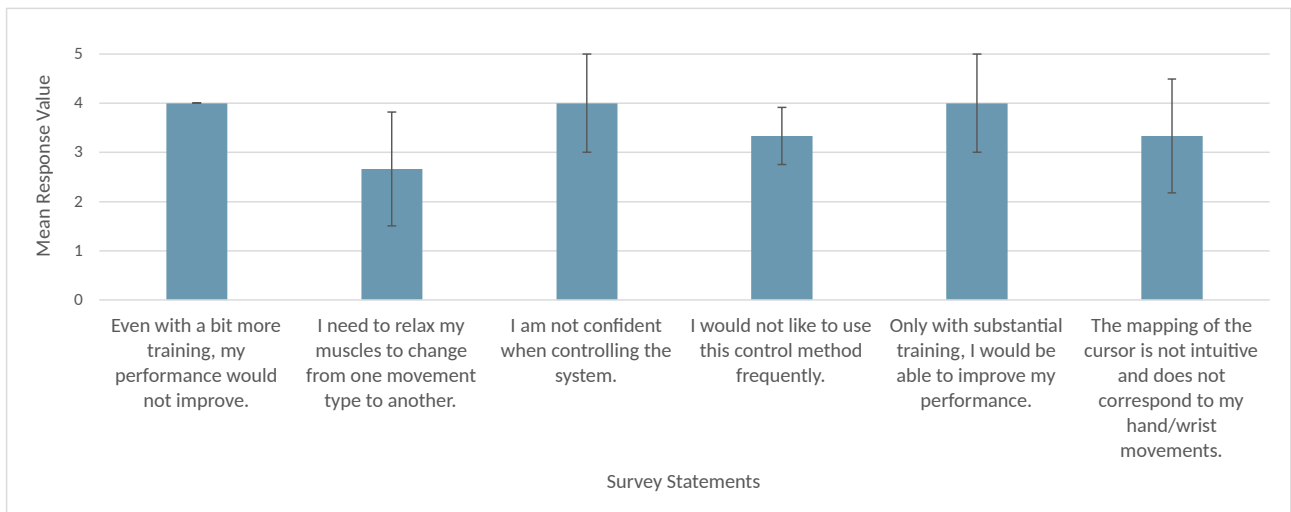
Survey results

The questionnaire was given to the participants after performing the test sets with the trained 2-DoF control system in real-time. Its 23 statements aim to capture the comfort and confidence levels, perceived performance levels, and the intuitiveness of the control of using the control system. The results of the survey are shown in Figure 4.13, where Figure 4.13a shows the mean responses for statements with positive implications, such that lower value responses that tend towards agreeing with the statement indicate easier perceived use by the participant, while Figure 4.13b and Figure 4.13c show the means of the statements with negative implications such that higher value responses that tend towards disagreeing with the statement indicate easier perceived use by the participant. Similarly, Figure 4.14 shows histograms of responses to the statements with positive and negative implications.

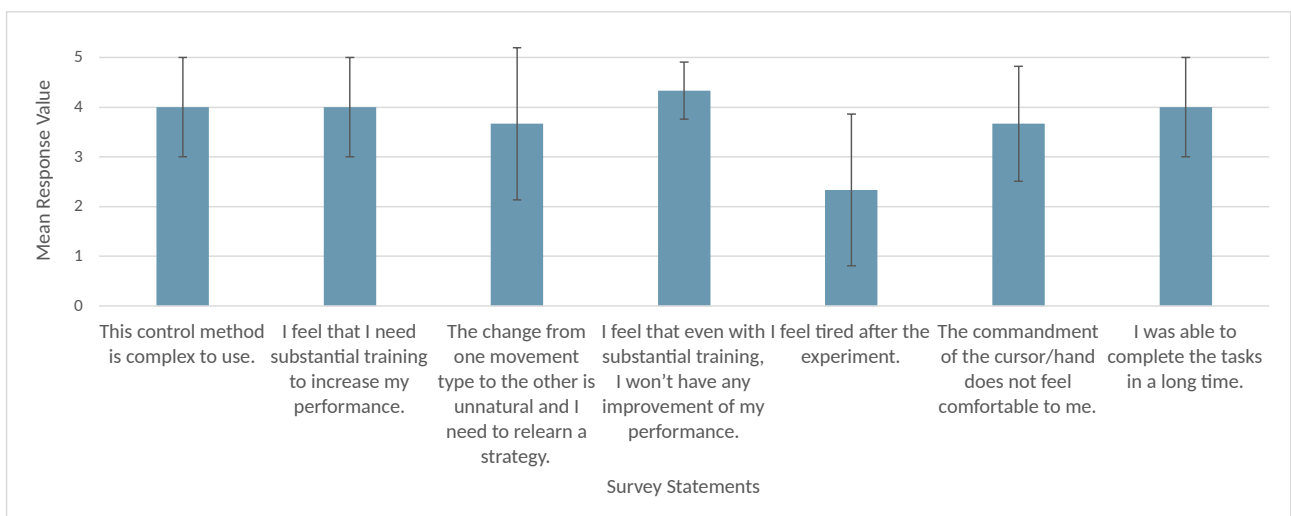
We can see that the responses for the statements with positive implications are generally skewed towards the lower values, i.e. participants tend to agree with statements about the system being easy to use. Likewise, the responses to the statements with negative implications are generally skewed towards the higher values, i.e. participants tend to disagree with the statements about the system being difficult to use. This shows generally a perceived confidence and intuitiveness in use of the real-time human intention detection system.



(a) Statements with positive implications



(b) Statements with negative implications



(c) Statements with negative implications

Figure 4.13: Bar plots of mean values for responses to statements in the user survey for the 2-DoF real-time control system. Responses from Strongly Agree to Strongly Disagree are mapped with values from 1 to 5. Error bars show standard deviation.

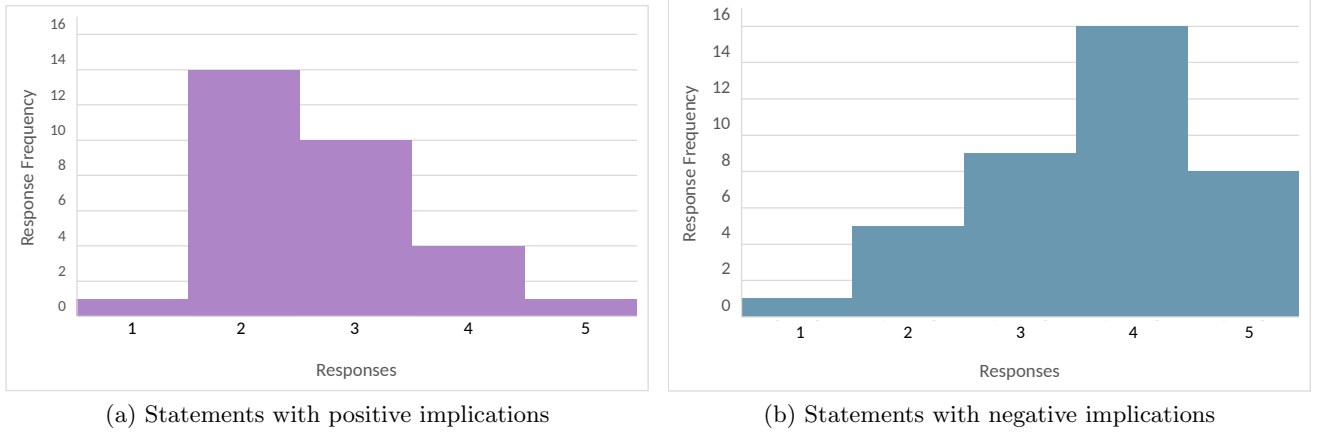


Figure 4.14: Histogram of survey responses. Response map: 1. Strongly Agree 2. Agree 3. Neutral 4. Disagree 5. Strongly Disagree.

4.2.4 Discussion

In this study we demonstrated real-time simultaneous and proportional myocontrol over 2 DoFs and 3 DoFs of the wrist as well as hand grasping using EMG signals from the forearm muscles of individuals with limb loss. The proposed paradigm enables intuitive control by relying on linear mapping between input commands and the target output gestures. This is achieved using a linear regression motion estimation algorithm for wrist control and a classifier for estimating hand closing trained on only four repetitions of individual and sequential movements.

The system allows for very accurate control of rotation and grasping, with a 100% success rate for the 2-DoF system. The completion rate drops to around 50% for the 3-DoF system. While our system will have to be limited to 2-DoFs, rotation and grasping are the most important DoFs for prosthesis users. The flexion-extension movement produced some discomfort for some participants, as it is not a movement they practice regularly or can imagine as easily. Participant LL2 for example found it very difficult to even train the regressor on flexion and extension due to discomfort, especially when in combination with grasping, which is one reason she did not test the system in combinations of 3 DoFs. In addition, the system was not trained on combined movements of rotation with flexion-extension which could explain a drop in the performance for tasks combining these 2 DoFs, similarly to the offline control algorithm. This reduction in performance between the 2-DoF and 3-DoF systems is clear across all metrics, although the speed in 1 DoF tasks is similar in both systems. Despite

the lower performance over 3 DoF compared to 2 DoFs, our 3-DoF system still performs better at tasks involving all 3 DoFs than in [102], the system in which produced a *Completion Rate* of 37% for combined movements across 3 DoF. Any comparisons to the research in [102] are limited as their system also involved proportional control over the grasping, which is not the case in this study. However, their participants involved mainly able-bodied participants, with only 1 amputee, who was able to successfully complete only 25% of the combined tasks in 3 DoF. We also only have one amputee who attempted the 3-DoF combined control. The performance in combinations of 2 DoFs over the 3-DoF trained system are comparable between our participants (58% Completion Rate) and the amputee in [102] (62% Completion Rate). Another thing to note is that, for both systems, despite the training over sequential movements, without explicitly training over simultaneous movements of the wrist and hand, the users were able to reach the combined targets in simultaneous control over movements in multiple DoFs. This further validates the results in Chapter 2.

There are more data samples for the 2-DoF system which better enables us to draw conclusions, even though sample sizes do not currently allow for statistical evaluation. The results for the 3-DoF system may yet vary with additional data points. The 2-DoF system performance is robust and accurate – most tasks were completed in under 5s, with a median overshoot of 0 – meaning that participants directly reached the target most of the time without too much jitter around the target. The overall *Path Efficiency* values of the 2-DoF control system are over 54%, in comparison to less than 20% path efficiency for the 3-DoF system, suggesting that users are taking relatively direct paths towards reaching the targets. These values are comparable to those in [102] (25-50%). Generally, the task performance worsens with increased *ActLvl*, with the exception of the grasping only tasks (*ActLvl* = 1.0). This is expected as the increased activation may reflect an increase in difficulty, and is also seen in [102]. I have been limiting comparisons to [102] so far, as they explicitly evaluate simultaneous combinations of 3 DoFs over wrist and hand with proportional capabilities using positional control. Most other similar studies rely on velocity control, for example in [36, 160]. Both studies have reached success rates of more than 90% for proportional and simultaneous control. However, velocity control may allow participants to complete movements in combined DoFs without

necessarily using simultaneous activation of these DoFs - the DoFs may be activated sequentially [102, 161]. Following on to the comparisons with other research, we specifically decided not to test on able bodied participants because there are so many studies out there considering only able-bodied participants, including [36, 160]. Results in such cases do not always translate the same to individuals with limb loss and are difficult to compare. The Kinematic Prediction Models used in this study were previously validated in an offline setting with able-bodied participants Chapter 2, so the study in this chapter is a continuation of that work, applied specifically to the target users.

Another interesting observation is that results for the 2-DoF system are fairly consistent across participants, despite differences in their anatomy. Therefore, the system is robust against variation across muscle tissue, scar tissue, residuum length and even across a variation in the reason for limb loss, from amputation to congenital limb absence. We can see from Figure 4.9 that the EMG activity varies in amplitude and location. Yet all participants are able to control the 2-DoF system with similar performance levels and 100% successful completion of tasks. The participants also have varying experience with traditional myoelectric control, but were still able to adapt to the simultaneous proportional control fairly quickly. Due to time constraints participants were only given around 1-2 minutes to familiarise with the control system after training and before beginning the target-reaching tasks, compared to around 10 minutes in [102].

Participants also noted various comments about the training and testing experience. One participant enjoys such activities as it gives her a chance to stretch, train, and rehabilitate her muscles, even though she experiences slight discomfort performing some of the tasks as they are unfamiliar to her and she does not use them in her every day activities. For all participants it was sometimes difficult to imagine the movements. As such they relied on the visual feedback from the cursor or prosthesis. I found it interesting that for example, while LL3 preferred the feedback from the cursor on the screen, for LL2 for example, controlling a prosthesis seemed more intuitive. Perhaps the prosthesis control was more intuitive as a replacement for the lost limb with strategies already having been previously used for mapping muscle activity to the required movements. Whereas, perhaps mapping to the cursor space allows for flexibility in which muscle activities map to which cursor movements. Either way, the

survey results indicate a general comfort and intuitiveness with use of the HDsEMG control strategy in our human intention detection system in a real-time setting for individuals with limb loss.

4.3 EMG-CV Shared Control System validation

Previously in this chapter we described the functionality of the shared EMG-CV controller, in Figure 4.1. This shared EMG-CV controller was also tested in a real-time setting, with the equipment and set up shown in Figure 4.2 and the socket in Figure 4.7. Supplementary Video 4.2 shows a demonstration with a participant with limb loss utilising the EMG-CV shared control to control the prosthetic hand while wearing the combined proof-of-concept socket. The grasp type is automatically selected by the Computer Vision Assistance, but is only applied once the user decides to close the hand, as inferred from the Human Intention Detection system. The user trained the Human Intention Detection system, and we also calibrated the Computer Vision Assistance system to select the appropriate frame number for the grasp state selection window. In the first parts of the demonstration, the Computer Vision system is validated using a simpler EMG setup. 2 sets of bipolar electrodes are used on the flexors and extensors of the arm, and the signals are passed through the classifier from our Kinematic Prediction Models. Thus, the user is only able to activate closing and opening of the hand, and we can demonstrate the functionality of the Computer Vision Assistance. The user is asked to direct their hand towards the object intended for grasping. As the camera’s visual perspective changes during the reaching task, a Grasp State is continuously output and stored as the current state of the system. Once the user activates their muscles to closes the hand, the grasp type is changed accordingly, and the user grasps the object with the prosthetic hand. When the user activates their muscles to open the hand, the prosthetic hand opens and reverts back to the pose for the General Grasp. For each object, and appropriate grasp type is used. Another aspect demonstrated is that our Computer Vision Assistance is not specific to each object. Different grasp types are selected based on the part of the object to be grasped for different uses. The system also runs continuously without having to be reset or restarted for each new object. We also briefly demonstrate the complete shared control with HDsEMG, where the user is also able to rotate their wrist in order to bring the prosthetic

hand into the appropriate orientation for grasping, while the CV system selects and applies the grasp type.



Supplementary Video 4.2: Demonstration of the shared control during a pilot experiment with a user with limb loss. The user controls the rotation of the prosthesis as well as closing and opening the prosthetic hand via the Human Intention Detection system. The prosthesis changes grasp type automatically based on the Computer Vision Assistance. The image shows the user wearing the complete proof-of-concept system and reaching towards a bottle. Video available at: <https://youtu.be/CGhTeKkKpFw>



Figure 4.15: User wearing the complete proof-of-concept system with shared control.

4.4 Conclusion

In this chapter, we have successfully integrated a human intention detection system for the control of a prosthetic hand in real-time, with a Computer Vision Assistance to generate the intended grasp types for different objects. The shared controller maps positions of the wrist and hand closing/opening from HDsEMG along with grasp states from the image processing and classification into commands for the control of a prosthetic hand and wrist.

In terms of the real-time Human Intention Detection system, we have developed a successful 2-DoF controller for proportional rotation of the wrist and simultaneous grasping, and even our 3-DoF control system shows promise with further improvements. This developed EMG-RS aligns well with the concept of the shared control framework, and although it would be interesting to investigate with more individuals to conduct a full statistical analysis, the performance is robust enough to use in our proof of concept semi-autonomous prosthetic control system integrated with CV.

The shared EMG-CV controller is currently very promising in terms of combining the outputs from the individual systems into the functionality of a user within the context of a reaching and grasping task. The shared controller will continue to be further refined and updated based on the experimental parameters in Chapter 3.

In order to further integrate the EMG control within the prototype, we must continue to refine the electrode placement for individual participants based on electrical activity as well as socket restrictions. To do so, we can refer to the EMG activity mapping. In further work, we would perform some offline or online testing of different arrangements of potentially reduced electrodes (to work with socket limitations). Offline, this could be selecting a limited number of channels and using the training tasks in cross validation for validation of the system. Online testing would require a repetition of the target tasks, but in any case would be validated along with a complete system validation.

The HDsEMG sensors may seem complex with respect to the CV assistance which consists of a single sensor. We chose to focus on a robust myocontrol as this underpins the entire control of the prosthesis. One of our aims was to focus on the user control, such that the user directly controls most of the functionality of the prosthesis through the Human Intention Detection system, as opposed

to conceding more control to the automated system. Achieving this level of myocontrol with users with limb loss is at the moment only achievable through HDsEMG, which is the current state of the art within research environments. This hardware is not yet clinically transferable, but research and development of more wearable HDsEMG sensors is in process, and these could be used in our research in the future. The automated control is complementary to the human control, so we pursued our research with a simpler sensor array, that forms a basis for an automated control that can be further refined in the future with additional sensors.

The proof-of-concept system could be further refined with the use of a more portable EMG acquisition to test how well the system would work in a more realistic environment. However, reducing the number of electrodes and increasing the separation between them would reduce the regions of electrical activity available. For example for participant LL1, as we can see from Figure 4.11a and Figure 4.11b, the amount of electrical activity is quite reduced. This could be due to scarring and bulky tissue on the skin. An armband around the circumference of the arm at any level is likely to only capture a couple isolated sites of activity. Thus, further improvements to wearable electrode systems in general is necessary for the translation of this research into more clinical applications.

There is also space within the framework development and the cursor space, regardless of the prosthetic attachment, to expand the research into control of 4 DoFs, where radial-ulnar deviation would be the represented on the y-axis. We have not mapped this axis to opening and closing (as in other research) in order to preserve an intuitive user interface with as close a linear mapping as possible from movements to the cursor space. However, this would need to be complemented with the necessary mechanical technology incorporated within a wearable hand and wrist prosthesis.

Chapter 5

Conclusion

This chapter is a review of the studies presented in this thesis and a discussion of the research contributions, limitations and suggestions for future work.

5.1 Thesis contributions and achievements

1. A 3-DoF human intention detection control system has been developed. The system can be extended to 4-DoF, for individual or simultaneous control of hand and wrist movements using HDsEMG.

Main outcomes:

- A new experimental setup has been established and a novel study protocol developed for recording HDsEMG and motion capture simultaneously. This protocol has been later used by many of my colleagues, with the establishment of collaborations on several research topics.
- A unique dataset of concurrent HDsEMG and arm kinematics has been captured.
- A HDsEMG control system has been designed to successfully control a total of 3 or 4 DoFs independently and simultaneously using a combination of two model types, regression for wrist movements and classification hand movements, in an offline setting. Wrist control is proportional and can be combined with closing or opening of the hand.

- A reduced training set of sequential wrist followed by hand movements was found to be sufficient for controlling more complicated simultaneous wrist and hand movements, successfully reducing the training time and complexity required.
- The control system was extended to a real-time setting for the control of a prosthesis. The control was validated with participants with limb loss with successful target reaching tasks in the cursor space in 2 DoFs, providing robust control across different limb properties.

Limitations:

- The Kinematic Prediction Models were not validated in the offline setting with individuals with limb loss.
- The Kinematic Prediction Models were not validated in the offline setting for simultaneous wrist control
- The performance of the real-time human intention detection controller is not as robust for a 3-DoF system as it is with a 2-DoF system, i.e when flexion-extension is included.
- The hand closing and opening could be expanded for use with proportional control. Our results do show successful regression to the finger angles (Figure 2.9), and angle thresholding could still be used to determine a closing trigger for the shared control. However, as it stands the control is robust and sufficient for a proof-of concept system.
- The real-time control system does not include radial-ulnar deviation. Following on from the offline study, the control method has potential to extend to radial-ulnar deviation, and the online GUI framework would support this. However, this functionality would need to be reflected in the mechanical design of a prosthetic with an active wrist in 3 DoFs.

2. A novel real-time CV algorithm for predicting the grasp types of objects for use with a myoelectrically controlled prosthetic device was developed.

Main outcomes:

- A novel CV algorithm was developed to determine the grasp type from a given object image,

using contour-based recognition from which features for grasp prediction are extracted.

- The algorithm was trained and validated offline before being adapted for use in a continuous, real-time, system for incorporation with a myoelectric prosthesis.
- The system was validated in a study with real-time reaching and grasping tasks with a 79% accuracy, while its operational parameters were established. The system generalises to novel objects and does not depend on object recognition
- A unique dataset of human reaching and grasping videos was collected from three camera point of views, which can be used to further improve this algorithm and similar ones in this area.

Limitations:

- The CV algorithm is currently limited to 4 classes, compared to the full range of grasp types used by humans and even available in more advanced prosthetic hands.
- The most common mis-classification occurs between the pinch and tripod classes. Potentially the features between these two classes are not distinct enough. However, in some cases both grasp types can be used to successfully grasp the object.
- The CV system only has access to RGB information. As such it is limited in that it misses out on other information humans would have such as weight of an object, depth perception, or other experiences of using it, along with some of the manipulation intentions.
- The system has not been validated in a realistic setting with a coloured or cluttered background.
- The continuous running of the system allows for ease of use without the need for the user to operate the CV aspect. However, this affects power and memory consumption and may not be valid in a more practical scenario.
- The current speed of the system seems to be sufficient, but it can always be improved.

3. An integrated shared control framework was developed for the control of a prosthetic hand which meets the specifications outlined in our aims: Dexterous, Real-time, User-control-centric,

Intuitive, Classifies 4 grasp types, Detects the user's intentions for 2-DoF proportional wrist control. This was realised through a Human Intention Detection System utilising HDsEMG for the simultaneous control of 2-DoFs of the wrist as well as activation of the hand commands; which shares its control with Computer Vision Assistance for visual input and analysis for grasp type selection, applied upon hand closing.

Main outcomes:

- This system presents shared control where the human intention detection is not solely used to operate the automated aspect.
- This system presents semi-autonomous control, such that the user is in control of the wrist even during grasping tasks, rather than allowing the automated system to fully take over.
- Grasp classification generalises to novel objects and is not restricted to object class.

Limitations:

- Although the system functions in 2-DoF of the wrist, it is more successful in 1-DoF proportional wrist control alongside closing and opening of the hand.
- The developed proof-of-concept system utilises a laptop and all the components are wired together. This is sufficient for a laboratory setting but would need to be scaled down in order to be used in practice.
- As part of the shared control, the current grasp state of the system is chosen based on established parameters based on frame number. This is general and could be different for different speeds of grasping and inconsistent reaching activities.

5.2 Future outlook

Further validation of the proof-of-concept shared control system is required with persons with limb loss. A validation experiment is in the process of being prepared, with the aim of having participants with limb loss use the system in a series of reaching and grasping tasks as well as other functional tests.

Ideally measures of reduction of cognitive complexity, time, and compensatory movements would be recorded and compared with other methods of controlling a prosthetic hand.

As it stands we have developed a proof-of-concept integrated system which would benefit from integration into an isolated system for more practical use. Ideally the electronics and processing should be incorporated into the socket itself and/or communicate wirelessly. In terms of the myoelectric signals, the acquisition technology would need to be improved to accommodate full integration of high density surface electrodes with a practical system.

The data collected from the experimental studies provides a great resource for further development of both the kinematic prediction models as well as the CV algorithm. More advanced signal processing techniques could be used to extract features from the surface EMG, and more complex models could be applied for the kinematic predictions. The kinematic prediction models could also be extended to proportional control of the closing/opening functionality. Further, the CV system could be enhanced by increasing the number of predicted classes.

To further improve the shared control strategy, additional sensors may be helpful in selecting the correct point at which to update the grasp state in the state machine. This could include investigating the use of depth sensors to better gauge the size of objects as well as IMUs to better gauge the speed of the arm or to decide if a reaching and grasping motion is intended by the user. Thus, several different sensor arrangements could be studied to test for a number of parameters including, but not limited to: number of sensors, type of sensors, placement, distance from surface of prosthetic device, etc. To allow for a systematic investigation, we propose that these parameters could be tested in a simulated environment. Even with just two sensory inputs, the system could benefit from more advanced sensor fusion techniques and drawing on Artificial Intelligence (AI) techniques such as adaptive learning and Kalman filtering.

More broadly, in this thesis we choose to focus on the control aspect of prosthesis use, but there are other concerns users have with prostheses which we do not address. It would be great to extend this proof-of-concept system, adapt it and improve it for more practical and clinical use to benefit people with limb loss.

The advancement that our research has brought in utilising the advanced myocontrol within the shared control presents a promising development in the research setting; but the clinical application is still restricted by the available hardware. Electrode shifts, the need to retrain, using wet electrodes, and portability of the acquisition systems are all challenges that are yet to be addressed for clinical use of current HDsEMG technology. However, our research has provided a control method which can be applied once the sensor technology improves and more individualised integration of EMG sensors within prostheses can be implemented.

The Computer Vision system can be implemented into an integrated wearable system, which is required for the further testing required in more clinical settings. The main remaining considerations are the limitations of CV in terms of object segmentation in cluttered backgrounds, which is an area being advanced in the fields of robotics and CV. The power implications of a continuously running system also need to be addressed. The system also currently requires calibration and training for approaching objects in a consistent manner.

The core next step in the research would be to design a more translational proof-of-concept prototype to investigate the clinical feasibility of the shared controller. This involves:

- Fully integrating the CV system with a prosthesis, from the sensor to the processing and power.
- Using a prosthetic hand which is more lightweight but still able to perform the range of grasps.

A prosthetic device more suited to the individual's anatomy in terms of overall length of the device and socket would also help with proprioception during reaching and grasping tasks.

- Designing individualised sockets for each participant.
- Incorporating a more wearable EMG sensor suite, despite its limitations in myocontrol. The CV Assistance will still be beneficial and add a new dimension to improve the overall control. We could even test the CV Assistance alongside currently commercially available prostheses with myocontrol.
- Continuously improving the algorithms involved.

- Designing an experimental study for validation of the shared controller with multiple people with limb loss in a more clinical setting.

Finally, even though my target user in this thesis is transradial prosthesis users, different aspects of the research have the potential to be applied in several other areas. Human intention detection has been used in several human-machine interface systems for controlling any number of systems, even for able-bodied users. This includes controlling a cursors on screens, human augmentation to control additional limbs or surgical tools, [162], and computer vision could complement several of these use-cases. The CV algorithm can also be expanded to other areas such as rehabilitation engineering, for use with exoskeletons, or in more automated environments or humanoid robotics.

Appendix A

EMG feature extraction and evaluation

The investigated features are time-domain and spatial domain features: root mean square for the reduction of data, 2-dimensional cross-correlation, and variogram. In addition to their robust performance in myoelectric control [28, 42] spatial filters are being evaluated because of the notion that antagonistic co-contraction occurs during the closing of the hand, such as during a power grip [163]. Hence, spatial features should highlight this co-contraction. All the features were calculated after the RMS calculation.

A.1 Variogram

After the RMS, the second feature investigated is the variogram, which is a function that describes spatial correlation between observations [42]. The experimental variogram is defined as follows [42]:

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2n(\mathbf{h})} \sum_{i=1}^{n(\mathbf{h})} [z(\mathbf{x}_i + \mathbf{h})]^2 \quad (\text{A.1})$$

Where $z(\mathbf{x}_i)$ is a measurement taken at location $\mathbf{x}_i$ and $n(\mathbf{h})$ is the number of pairs $\mathbf{h}$ units apart in the direction of the vector $\mathbf{h}$ [164] as cited by [42]. Thus, a relation is found between every electrode upon shifting positions. Utilising such a spatial feature space was shown in [42] to improve the robustness to electrode shift or change in myoelectric control, when compared to classical time-domain and autoregressive features. The variogram was implemented as in [42] and calculated in both

the x and y directions along the electrode grids.

A.2 2D cross-correlation

2D cross-correlation is a measure of correlation, or similarity, between two matrices. The 2-D cross-correlation of an M -by- N matrix, X , and a P -by- Q matrix, H , is a matrix, C , of size $M+P-1$ by $N+Q-1$ [165]. Its elements are given by

$$C(k, l) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} X(m, n) \bar{H}(m-k, n-l)$$

where $-(P-1) \leq k \leq M-1$
 $-(Q-1) \leq l \leq N-1$

where the bar over H denotes complex conjugation [165].

For each RMS window, the RMS values of all the channels were rearranged into their electrode grid format (3 8x8 matrices). 2D cross-correlation was performed between these matrices, and the peaks of these cross-correlograms were used as a feature (3 values for each time window). Using the RMS values for one window at a time will be more practical for real-time analysis. This is demonstrated in Figure A.1.

A.3 Feature selection

The kinematic prediction models were trained on the different features extracted from the HDsEMG data, as well as combinations of features, in order to evaluate the performance of the prediction models. For the evaluation, the prediction models were trained on individual wrist movements. The r^2 and RMSE scores were calculated for regression across different angles, features, and testing movements in the 4-DoF (Figure A.2 and Figure A.3) and 3-DoF (Figure A.4 and Figure A.5) systems.

The RMS feature generally outperforms the other features, except for combination of features.

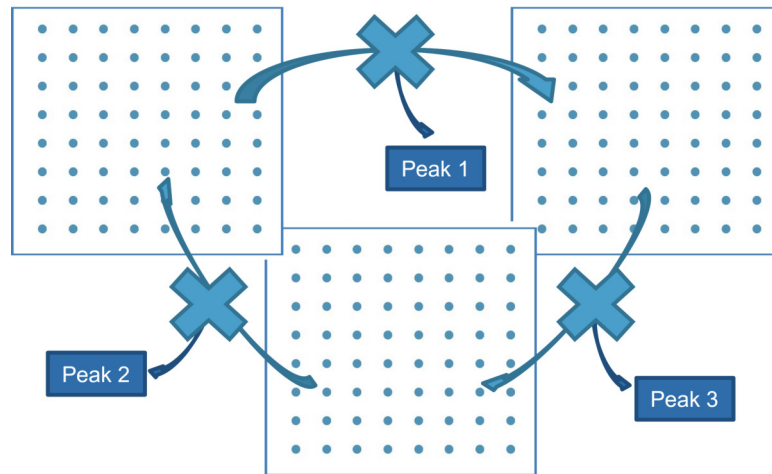


Figure A.1: 2D cross-correlation, as used as a feature for kinematic prediction. Each dot represents one RMS value over a specific window. These values are arranged in grids based on the original EMG channel alignment. X represents cross-correlation between the RMS values of the grids. The three peak values of the resulting cross-correlograms were used as features.

The scores are not much improved with the combinations of features though. As such, the RMS was chosen as the preferred feature for use in further investigations.

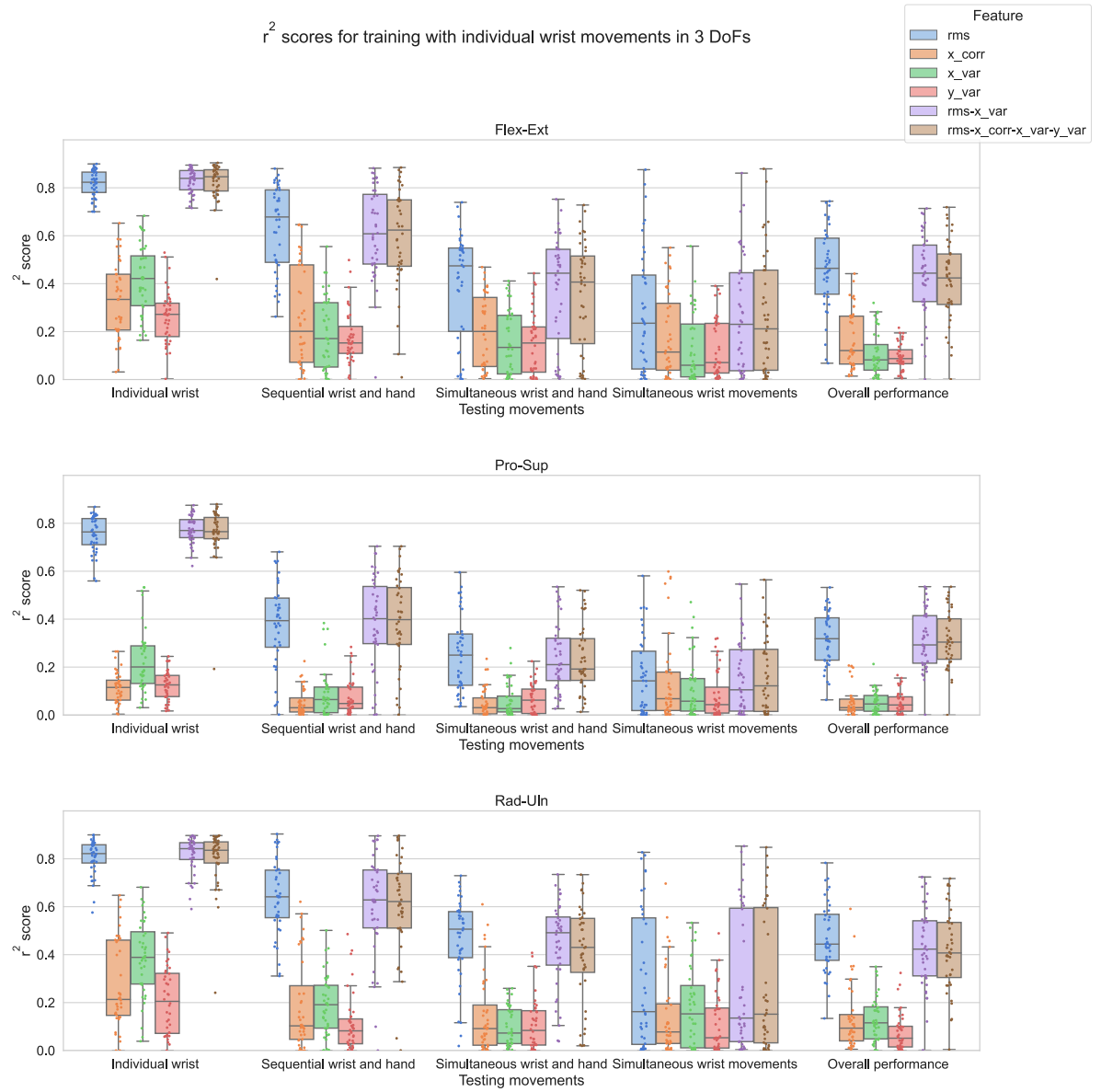


Figure A.2: 4-DoF system showing r^2 scores across wrist angles when using different features to train the linear regressor with the individual wrist movements

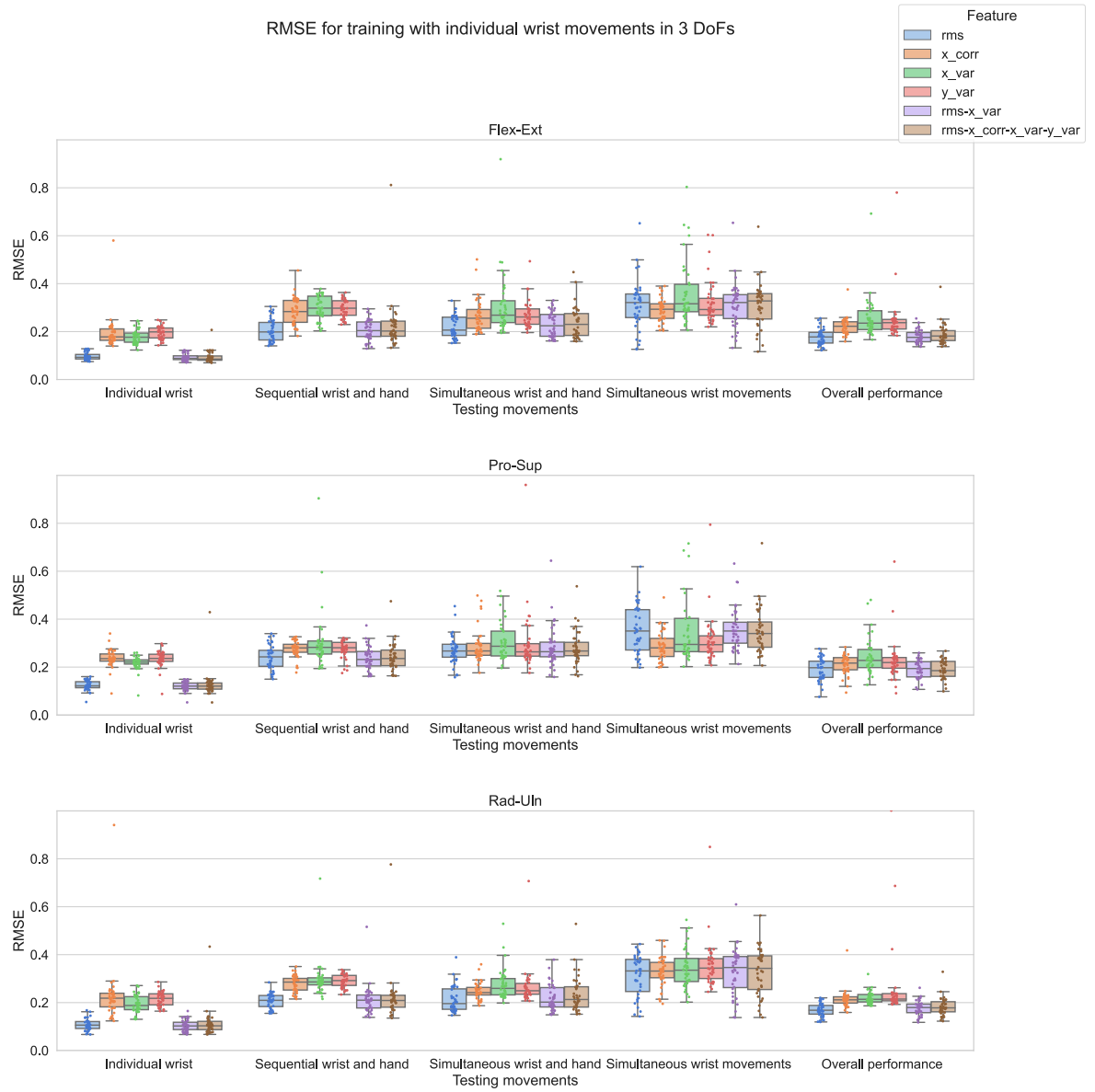


Figure A.3: 4-DoF system showing normalised RMSE scores across wrist angles when using different features to train the linear regressor with the individual wrist movements

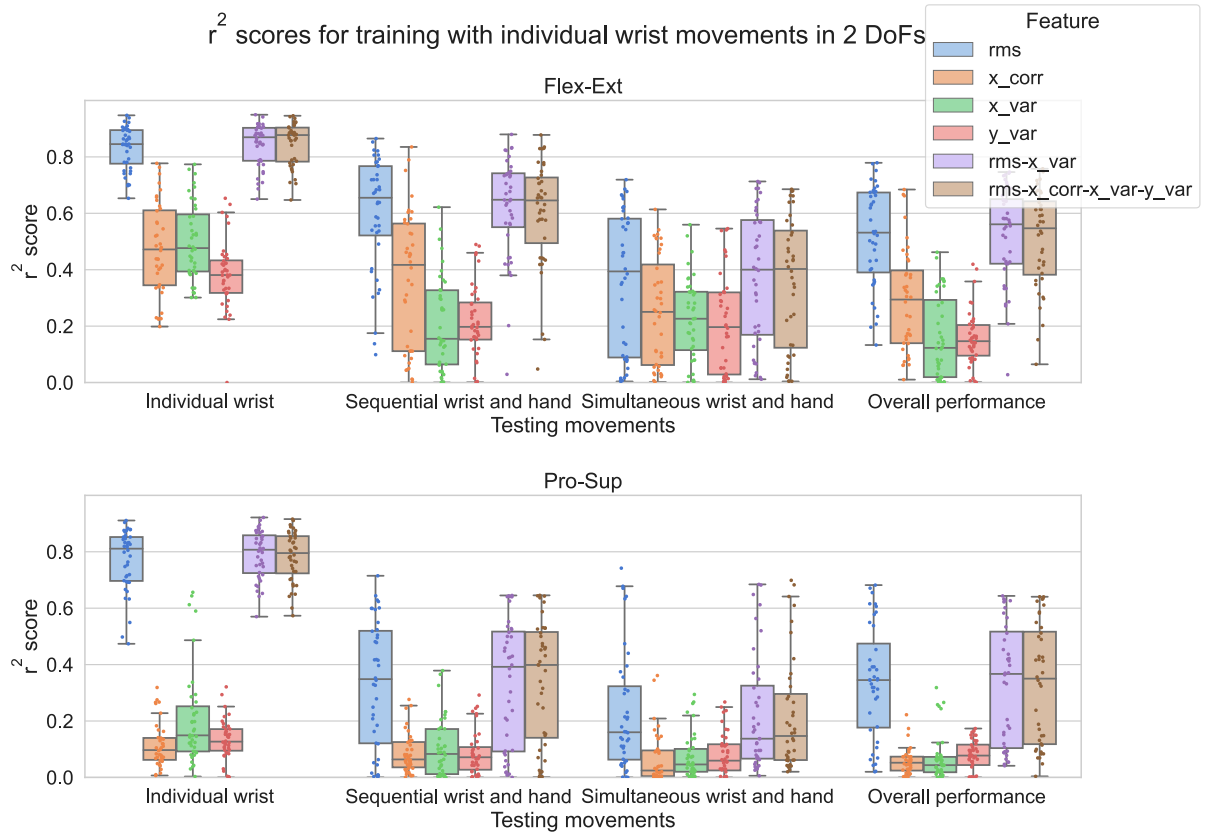


Figure A.4: 3-DoF system showing r^2 scores across wrist angles when using different features to train the linear regressor with the individual wrist movements

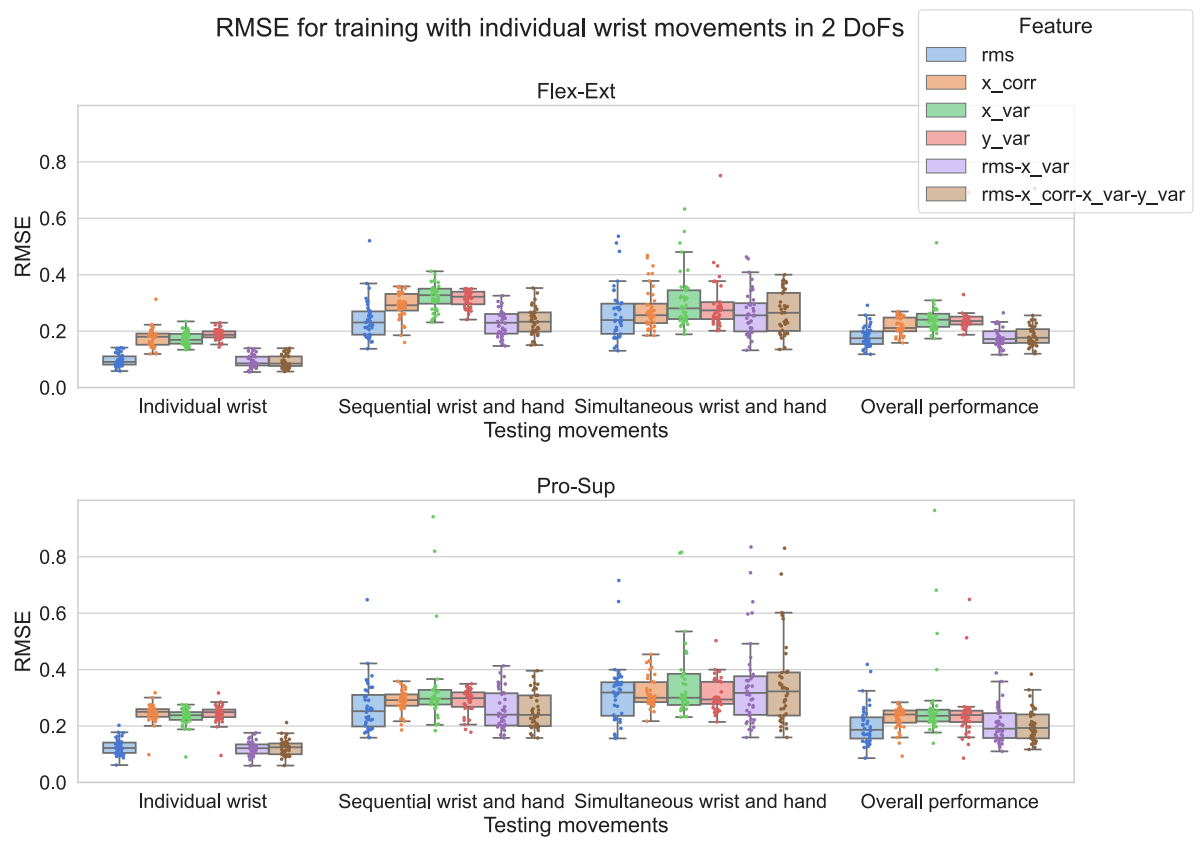


Figure A.5: 3-DoF system showing normalised RMSE scores across wrist angles when using different features to train the linear regressor with the individual wrist movements

Appendix B

Permissions for reuse

Figure 1.1: *The LEGO Group does not object to scanning of limited extracts of materials in unaltered form for non-commercial purposes of exchange of information or good faith commentary. This thesis is independent research not sponsored by the LEGO Group.*

Figure 1.2: *The IEEE does not require individuals working on a thesis to obtain a formal reuse license.*

Figure 2.2: *The IEEE does not require individuals working on a thesis to obtain a formal reuse license. Springer Nature does not require individuals to obtain permission to reproduce open access articles distributed under the terms of the Creative Commons CC by license.*

E-theses letter template: request to reproduce an extract from a third party's published work

06/06/2023

Dear Matthew,

I have completed my PhD thesis at Imperial College London entitled 'Integrating Computer Vision with NeuroMuscular Interfacing for the Semi-Autonomous Control of Robotic Limbs'.

I seek your permission to reprint, in my thesis an extract from: MyTASKA (v22.11.1) The extract to be reproduced is attached (Figure 1.3 in my thesis) below:

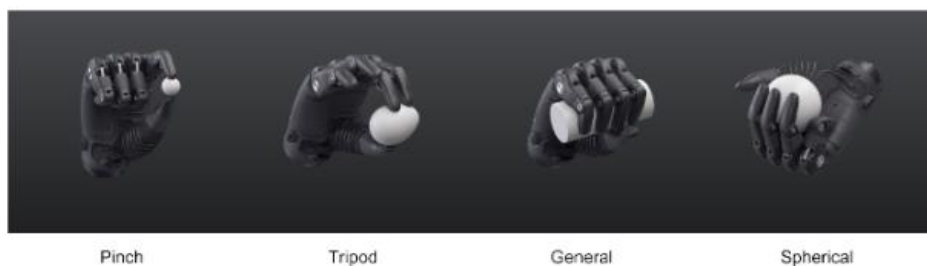


Figure 1.3: Representation of the selected grasp type options as performed by the TASKA prosthetic hand [59]. ©TASKA Prosthetics 2023.

I would like to include the extract in my thesis which will be added to Spiral, Imperial's institutional repository <http://spiral.imperial.ac.uk/> and made available to the public under a Creative Commons Attribution-Non Commercial-No Derivatives 4.0 International Licence (CC BY-NC-ND).

If you are happy to grant me all the permissions requested, please return a signed copy of this letter. If you wish to grant only some of the permissions requested, please list these and then sign.

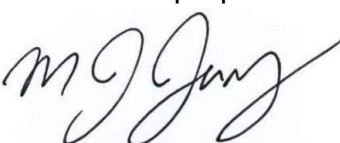
Yours sincerely,

Milia Helena Hasbani

Permission granted for the use requested above:

I confirm that I am the copyright holder of the extract above and hereby give permission to include it in your thesis which will be made available, via the internet, for non-commercial purposes under the terms of the user licence.

Signed:



Name: Mathew Jury

Organisation: Taska Prosthetics Ltd

Job title: Product Director

Bibliography

- [1] The LEGO Group, “The lego group reveals a new generation of lego® friends - about us - lego.com,” 10 2022. [Online]. Available: <https://www.lego.com/en-us/aboutus/news/2022/october/the-lego-group-reveals-a-new-generation-of-lego-friends>
- [2] —, “Autumn – lego® friends character — official lego® shop gb,” 2023. [Online]. Available: <https://www.lego.com/en-gb/themes/friends/characters/autumn>
- [3] NHS England Specialised Commissioning Team, “Hand and upper limb reconstruction using vascularised composite allotransplantation (haul-vca),” 2015. [Online]. Available: <https://www.england.nhs.uk/commissioning/wp-content/uploads/sites/12/2015/07/d00pa-hand-transplnt-forearm-loss.pdf>
- [4] NHS, “Nhs commissioning » prosthetics service review,” 2020. [Online]. Available: <https://www.england.nhs.uk/commissioning/spec-services/npc-crg/group-d/rehabilitation-and-disability/prosthetics-review/>
- [5] K. Ziegler-Graham, E. J. MacKenzie, P. L. Ephraim, T. G. Trivison, and R. Brookmeyer, “Estimating the prevalence of limb loss in the united states: 2005 to 2050,” *Archives of Physical Medicine and Rehabilitation*, vol. 89, pp. 422–429, 3 2008.
- [6] F. Cordella, A. L. Ciano, R. Sacchetti, A. Davalli, A. G. Cutti, E. Guglielmelli, and L. Zollo, “Literature review on needs of upper limb prosthesis users,” *Frontiers in Neuroscience*, vol. 10, p. 209, 5 2016. [Online]. Available: <http://journal.frontiersin.org/Article/10.3389/fnins.2016.00209/abstract>

- [7] C. H. Jang, H. S. Yang, H. E. Yang, S. Y. Lee, J. W. Kwon, B. D. Yun, J. Y. Choi, S. N. Kim, and H. W. Jeong, "A survey on activities of daily living and occupations of upper extremity amputees," *Annals of Rehabilitation Medicine*, vol. 35, p. 907, 12 2011. [Online]. Available: <http://e-arm.org/journal/view.php?doi=10.5535/arm.2011.35.6.907>
- [8] D. Farina, I. Vujaklija, R. Brånemark, A. M. J. Bull, H. Dietl, B. Graimann, L. J. Hargrove, K.-P. Hoffmann, H. Huang, T. Ingvarsson, H. B. Janusson, K. Kristjánsson, T. Kuiken, S. Micera, T. Stieglitz, A. Sturma, D. Tyler, R. F. ff. Weir, and O. C. Aszmann, "Toward higher-performance bionic limbs for wider clinical use," *Nature Biomedical Engineering*, vol. 10, p. 20, 5 2021. [Online]. Available: <https://www.nature.com/articles/s41551-021-00732-x>
- [9] L. V. McFarland, S. L. H. Winkler, A. W. Heinemann, M. Jones, and A. Esquenazi, "Unilateral upper-limb loss: satisfaction and prosthetic-device use in veterans and servicemembers from vietnam and oif/oef conflicts." *Journal of rehabilitation research and development*, vol. 47, pp. 299–316, 2010. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/20803400>
- [10] E. Biddiss and T. Chau, "Upper limb prosthesis use and abandonment: A survey of the last 25 years," pp. 236–257, 9 2007. [Online]. Available: <http://journals.sagepub.com/doi/10.1080/03093640600994581>
- [11] T. R. Makin, F. de Vignemont, and A. A. Faisal, "Neurocognitive barriers to the embodiment of technology," *Nature Biomedical Engineering*, vol. 1, p. 0014, 1 2017. [Online]. Available: <http://www.nature.com/articles/s41551-016-0014>
- [12] P. Beckerle, R. Kõiva, E. A. Kirchner, R. Bekrater-Bodmann, S. Dosen, O. Christ, D. A. Abbink, C. Castellini, and B. Lenggenhager, "Feel-good robotics: Requirements on touch for embodiment in assistive robotics." *Frontiers in neurorobotics*, vol. 12, p. 84, 2018. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/30618706http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC6297195>
- [13] L. Resnik, M. Borgia, and M. Clark, "Function and quality of life of unilateral major upper limb

- amputees: Effect of prosthesis use and type,” *Archives of Physical Medicine and Rehabilitation*, vol. 101, pp. 1396–1406, 8 2020.
- [14] A. Sturma, L. A. Hruby, A. Boesendorfer, A. Pittermann, S. Salminger, C. Gstoettner, O. Politikou, I. Vujaklija, D. Farina, and O. C. Aszmann, “Prosthetic embodiment and body image changes in patients undergoing bionic reconstruction following brachial plexus injury,” *Frontiers in Neurorobotics*, vol. 15, p. 45, 4 2021. [Online]. Available: <https://www.frontiersin.org/articles/10.3389/fnbot.2021.645261/full>
- [15] NHS England, “Clinical commissioning policy multi-grip prosthetic hand (all ages) (2009 [220801p] commissioning position,” 2022. [Online]. Available: <https://www.england.nhs.uk/publication/clinical-commissioning-policy-multi-grip-prosthetic-hand-all-ages/>
- [16] D. Farina, N. Jiang, H. Rehbaum, A. Holobar, B. Graimann, H. Dietl, and O. C. Aszmann, “The extraction of neural information from the surface emg for the control of upper-limb prostheses: Emerging avenues and challenges,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, pp. 797–809, 7 2014. [Online]. Available: <http://ieeexplore.ieee.org/document/6737308/>
- [17] M. A. Oskoei and H. Hu, “Myoelectric control systems—a survey,” *Biomedical Signal Processing and Control*, vol. 2, pp. 275–294, 10 2007. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1746809407000547>
- [18] D. Farina, R. Merletti, and R. M. Enoka, “The extraction of neural strategies from the surface emg,” *Journal of Applied Physiology*, vol. 96, pp. 1486–1495, 4 2004. [Online]. Available: <http://www.physiology.org/doi/10.1152/japplphysiol.01070.2003>
- [19] N. Jiang, S. Dosen, K. R. Muller, and D. Farina, “Myoelectric control of artificial limbs: there a need to change focus? [in the spotlight],” pp. 148–152, 2012.
- [20] V. S. Ramachandran and W. Hirstein, “The perception of phantom limbs. the

- d. o. hebb lecture.” *Brain*, vol. 121, pp. 1603–1630, 9 1998. [Online]. Available: <https://academic.oup.com/brain/article/121/9/1603/283360>
- [21] J. M. Hahne, M. A. Schweisfurth, M. Koppe, and D. Farina, “Simultaneous control of multiple functions of bionic hand prostheses: Performance and robustness in end users,” *Science Robotics*, vol. 3, 6 2018. [Online]. Available: <https://www.science.org>
- [22] M. Ison, I. Vujaklija, B. Whitsell, D. Farina, and P. Artemiadis, “High-density electromyography and motor skill learning for robust long-term control of a 7-dof robot arm,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, pp. 424–433, 4 2016. [Online]. Available: <http://ieeexplore.ieee.org/document/7073629/>
- [23] E. D. Sherman, “A russian bioelectric-controlled prosthesis: Report of a research team from the rehabilitation institute of montreal,” *Canadian Medical Association Journal*, vol. 91, p. 1268, 12 1964. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1927453/>
- [24] P. Parker, K. Englehart, and B. Hudgins, “Myoelectric signal processing for control of powered limb prostheses,” *Journal of Electromyography and Kinesiology*, vol. 16, pp. 541–548, 12 2006.
- [25] B. Peerdeman, D. Boere, H. Witteveen, R. H. I. ’t Veld, H. Hermens, S. Stramigioli, H. Rietman, P. Veltink, and S. Misra, “Myoelectric forearm prostheses: State of the art from a user-centered perspective,” vol. 48, 2011.
- [26] K. Englehart and B. Hudgins, “A robust, real-time control scheme for multifunction myoelectric control,” *IEEE Transactions on Biomedical Engineering*, vol. 50, pp. 848–854, 7 2003. [Online]. Available: <http://ieeexplore.ieee.org/document/1206493/>
- [27] C. Castellini and P. van der Smagt, “Surface emg in advanced hand prosthetics,” *Biological Cybernetics*, vol. 100, pp. 35–47, 1 2009. [Online]. Available: <http://link.springer.com/10.1007/s00422-008-0278-1>
- [28] J. M. Hahne, B. Graimann, and K. R. Muller, “Spatial filtering for robust myoelectric control,”

- IEEE Transactions on Biomedical Engineering*, vol. 59, pp. 1436–1443, 5 2012. [Online]. Available: <http://ieeexplore.ieee.org/document/6156755/>
- [29] S. Amsuess, P. Goebel, B. Graimann, and D. Farina, “A multi-class proportional myocontrol algorithm for upper limb prosthesis control: Validation in real-life scenarios on amputees,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 23, pp. 827–836, 9 2015. [Online]. Available: <http://ieeexplore.ieee.org/document/6915873/>
- [30] A. J. Young, L. H. Smith, E. J. Rouse, and L. J. Hargrove, “Classification of simultaneous movements using surface emg pattern recognition,” *IEEE Transactions on Biomedical Engineering*, vol. 60, pp. 1250–1258, 2013.
- [31] N. Jiang, K. B. Englehart, and P. A. Parker, “Extracting simultaneous and proportional neural control information for multiple-dof prostheses from the surface electromyographic signal,” *IEEE Transactions on Biomedical Engineering*, vol. 56, pp. 1070–1080, 4 2009. [Online]. Available: <https://pubmed.ncbi.nlm.nih.gov/19272889/>
- [32] J. M. Hahne, F. Bießmann, N. Jiang, H. Rehbaum, D. Farina, F. C. Meinecke, K. R. Muller, and L. C. Parra, *IEEE Transactions on Neural Systems and Rehabilitation Engineering*.
- [33] S. Muceli and D. Farina, “Simultaneous and proportional estimation of hand kinematics from emg during mirrored movements at multiple degrees-of-freedom,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 20, pp. 371–378, 5 2012. [Online]. Available: <https://ieeexplore.ieee.org/document/6104157/>
- [34] A. Ameri, E. N. Kamavuako, E. J. Scheme, K. B. Englehart, and P. A. Parker, “Support vector regression for improved real-time, simultaneous myoelectric control,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, pp. 1198–1209, 11 2014.
- [35] —, “Real-time, simultaneous myoelectric control using visual target-based training paradigm,” *Biomedical Signal Processing and Control*, vol. 13, pp. 8–14, 9 2014.

- [36] M. Ortiz-Catalan, B. Håkansson, and R. Brånemark, “Real-time and simultaneous control of artificial limbs based on pattern recognition algorithms,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, pp. 756–764, 2014.
- [37] M. Ison and P. Artemiadis, “The role of muscle synergies in myoelectric control: trends and challenges for simultaneous multifunction control,” *Journal of Neural Engineering*, vol. 11, p. 051001, 10 2014. [Online]. Available: <http://stacks.iop.org/1741-2552/11/i=5/a=051001?key=crossref.8802780ae2923251ec7d10f6b568ca80>
- [38] M. Ariyanto, W. Caesarendra, K. A. Mustaqim, M. Irfan, J. A. Pakpahan, J. D. Setiawan, and A. R. Winoto, “Finger movement pattern recognition method using artificial neural network based on electromyography (emg) sensor,” *Proceedings of the 2015 International Conference on Automation, Cognitive Science, Optics, Micro Electro-Mechanical System, and Information Technology, ICACOMIT 2015*, pp. 12–17, 3 2016.
- [39] J. L. Nielsen, S. Holmgård, N. Jiang, K. B. Englehart, D. Farina, and P. A. Parker, “Simultaneous and proportional force estimation for multifunction myoelectric prostheses using mirrored bilateral training,” *IEEE Transactions on Biomedical Engineering*, vol. 58, pp. 681–688, 3 2011.
- [40] J. Ekstedt and E. Stålberg, “Single fibre electromyography for the study of the microphysiology of the human muscle,” *New Concepts of the Motor Unit, Neuromuscular Disorders, Electromyographic Kinesiology*, vol. 1, pp. 89–112, 4 1973. [Online]. Available: <https://www.karger.com/Article/FullText/394031https://www.karger.com/Article/Abstract/394031>
- [41] “Welcome to myo support,” 2018. [Online]. Available: <https://support.getmyo.com/hc/en-us>
- [42] A. Stango, F. Negro, and D. Farina, “Spatial correlation of high density emg signals provides features robust to electrode number and shift in pattern recognition for myocontrol,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 23, pp. 189–198, 3 2015. [Online]. Available: <http://ieeexplore.ieee.org/document/6949119/>
- [43] G. Drost, D. F. Stegeman, B. G. van Engelen, and M. J. Zwarts, “Clinical applications of

- high-density surface emg: A systematic review,” *Journal of Electromyography and Kinesiology*, vol. 16, pp. 586–602, 12 2006. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1050641106001209?via%3Dihub>
- [44] A. Holobar and D. Farina, “Blind source identification from the multichannel surface electromyogram,” *Physiological Measurement*, vol. 35, p. R143, 6 2014. [Online]. Available: <https://iopscience.iop.org/article/10.1088/0967-3334/35/7/R143https://iopscience.iop.org/article/10.1088/0967-3334/35/7/R143/meta>
- [45] COAPT LLC.
- [46] Ottobock, “Myo plus — upper limb prosthetics — prosthetics — ottobock us shop.” [Online]. Available: <https://shop.ottobock.us/Prosthetics/Upper-Limb-Prosthetics/Myo-Plus/c/2901>
- [47] A. J. Young, L. J. Hargrove, and T. A. Kuiken, “The effects of electrode size and orientation on the sensitivity of myoelectric pattern recognition systems to electrode shift,” *IEEE Transactions on Biomedical Engineering*, vol. 58, pp. 2537–2544, 9 2011.
- [48] H. J. Hwang, J. M. Hahne, and K. R. Müller, “Real-time robustness evaluation of regression based myoelectric control against arm position change and donning/doffing,” *PLOS ONE*, vol. 12, p. e0186318, 11 2017. [Online]. Available: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0186318>
- [49] A. Fougner, E. Scheme, A. D. Chan, K. Englehart, and Øyvind Stavdahl, “Resolving the limb position effect in myoelectric pattern recognition,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 19, pp. 644–651, 12 2011.
- [50] C. Cipriani, R. Sassu, M. Controzzi, and M. C. Carrozza, “Influence of the weight actions of the hand prosthesis on the performance of pattern recognition based myoelectric control: Preliminary study,” *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, pp. 1620–1623, 2011.

- [51] S. Amsuss, L. P. Paredes, N. Rudigkeit, B. Graimann, M. J. Herrmann, and D. Farina, “Long term stability of surface emg pattern classification for prosthetic control,” *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, pp. 3622–3625, 2013.
- [52] J. R. Napier, “The prehensile movements of the human hand,” *The Journal of Bone and Joint Surgery*, vol. 38 B, pp. 902–913, 11 1956. [Online]. Available: <https://online.boneandjoint.org.uk/doi/abs/10.1302/0301-620X.38B4.902>
- [53] R. J. Bootsma and P. C. van Wieringen, “Spatio-temporal organisation of natural prehension,” *Human Movement Science*, vol. 11, pp. 205–215, 2 1992.
- [54] M. Markovic, S. Dosen, C. Cipriani, D. Popovic, and D. Farina, “Stereovision and augmented reality for closed-loop control of grasping in hand prostheses,” *Journal of Neural Engineering*, vol. 11, p. 046001, 8 2014. [Online]. Available: <http://stacks.iop.org/1741-2552/11/i=4/a=046001?key=crossref.920cc30034714dfe1da6df3291ab5200>
- [55] A. E. Flatt, “Grasp,” *Proceedings (Baylor University. Medical Center)*, vol. 13, p. 343, 10 2000. [Online]. Available: </pmc/articles/PMC1312228/https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1312228/>
- [56] M. R. Cutkosky, “On grasp choice, grasp models, and the design of hands for manufacturing tasks,” *IEEE Transactions on Robotics and Automation*, vol. 5, pp. 269–279, 1989.
- [57] T. Feix, J. Romero, H.-B. Schmiedmayer, A. M. Dollar, and D. Kragic, “The grasp taxonomy of human grasp types,” *IEEE Transactions on Human-Machine Systems*, vol. 46, pp. 66–77, 2 2016. [Online]. Available: <http://ieeexplore.ieee.org/document/7243327/>
- [58] Koalaa Ltd., “The alx — koalaa.” [Online]. Available: <https://www.yourkoalaa.com/the-alx>
- [59] TASKA Prosthetics LTD, “MyTASKA (v22.11.1).”
- [60] TASKA Prosthetics, “Advanced prosthetic hand — taska prosthetics.” [Online]. Avail-

able: <https://www.taskaprosthetics.com/the-taska/features/#http://www.taskaprosthetics.com/the-taska/features/>

- [61] J. Redmon and A. Angelova, “Real-time grasp detection using convolutional neural networks.” IEEE, 5 2015, pp. 1316–1322. [Online]. Available: <http://ieeexplore.ieee.org/document/7139361/>
- [62] M. R. Maymó, A. Shafti, and A. A. Faisal, “Fast orient: Lightweight computer vision for wrist control in assistive robotic grasping,” pp. 207–212, 2018. [Online]. Available: <http://www.enhance-motion.eu>
- [63] A. Saxena, J. Driemeyer, and A. Y. Ng, “Robotic grasping of novel objects using vision,” *The International Journal of Robotics Research*, vol. 27, pp. 157–173, 2008. [Online]. Available: <http://ijr.sagepub.com>
- [64] Q. V. Le, D. Kamm, A. F. Kara, and A. Y. Ng, “Learning to grasp objects with multiple contact points,” pp. 5062–5069, 2010. [Online]. Available: <http://www.robotics.stanford.edu/~ang/papers/icra10-LearningGraspMultipleContactPoints.pdf>
- [65] P. Sanz, A. D. Pobil, J. Inesta, and G. Recatala, “Vision-guided grasping of unknown objects for service robots,” vol. 4. IEEE, 1998, pp. 3018–3025. [Online]. Available: <http://ieeexplore.ieee.org/document/680889/>
- [66] D. P. McMullen, G. Hotson, K. D. Katyal, B. A. Wester, M. S. Fifer, T. G. McGee, A. Harris, M. S. Johannes, R. J. Vogelstein, A. D. Ravitz, W. S. Anderson, N. V. Thakor, and N. E. Crone, “Demonstration of a semi-autonomous hybrid brain–machine interface using human intracranial eeg, eye tracking, and computer vision to control a robotic upper limb prosthetic,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, pp. 784–796, 7 2014. [Online]. Available: <http://ieeexplore.ieee.org/document/6683036/>
- [67] D. G. K. Madusanka, L. N. S. Wijayasingha, R. A. R. C. Gopura, Y. W. R. Amarasinghe, and

- G. K. I. Mann, "A review on hybrid myoelectric control systems for upper limb prosthesis." IEEE, 4 2015, pp. 136–141. [Online]. Available: <http://ieeexplore.ieee.org/document/7112334/>
- [68] D. Novak and R. Riener, "A survey of sensor fusion methods in wearable robotics," *Robotics and Autonomous Systems*, vol. 73, pp. 155–170, 11 2015. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0921889014001705?via%3Dihub>
- [69] O. Fukuda, D. Sakaguchi, Y. He, N. Yamaguchi, and H. Okumura, "Bimodal control of a vision-based myoelectric hand," *IEEE Access*, vol. 9, pp. 98 369–98 380, 2021.
- [70] S. Dosen, C. Cipriani, M. Kostic, M. Controzzi, M. C. Carrozza, and D. B. Popovic, "Cognitive vision system for control of dexterous prosthetic hands: Experimental evaluation," *Journal of NeuroEngineering and Rehabilitation*, vol. 7, p. 42, 8 2010. [Online]. Available: <http://jneuroengrehab.biomedcentral.com/articles/10.1186/1743-0003-7-42>
- [71] S. Došen and D. B. Popović, "Transradial prosthesis: Artificial vision for control of prehension," *Artificial Organs*, vol. 35, pp. 37–48, 1 2010. [Online]. Available: <http://doi.wiley.com/10.1111/j.1525-1594.2010.01040.x>
- [72] M. Markovic, S. Dosen, D. Popovic, B. Graimann, and D. Farina, "Sensor fusion and computer vision for context-aware control of a multi degree-of-freedom prosthesis," *Journal of Neural Engineering*, vol. 12, p. 066022, 12 2015. [Online]. Available: <http://stacks.iop.org/1741-2552/12/i=6/a=066022?key=crossref.e839c819ffabd7b7ef86d80f40749f1f>
- [73] M. Gardner, R. Woodward, R. Vaidyanathan, E. Burdet, and B. C. Khoo, "An unobtrusive vision system to reduce the cognitive burden of hand prosthesis control." IEEE, 2014, pp. 1279–1284.
- [74] M. J. L. Gardner, R. Vaidyanathan, and P. E. Burdet, "A heterogeneous sensor suite for reducing the cognitive burden of upper limb robotic control," 2017. [Online]. Available: <http://spiral.imperial.ac.uk/handle/10044/1/45047>

- [75] J. Degol, A. Akhtar, B. Manja, and T. Bretl, “Automatic grasp selection using a camera in a hand prosthesis,” vol. 2016-October. IEEE, 8 2016, pp. 431–434. [Online]. Available: <http://ieeexplore.ieee.org/document/7590732/>
- [76] D. T. G. Andrade, A. Ishikawa, A. D. Munoz, and E. Rohmer, “A hybrid approach for the actuation of upper limb prostheses based on computer vision,” vol. 2017-December. IEEE, 11 2017, pp. 1–6. [Online]. Available: <http://ieeexplore.ieee.org/document/8215308/>
- [77] G. Ghazaei, A. Alameer, P. Degenaar, G. Morgan, and K. Nazarpour, *Journal of Neural Engineering*, 6.
- [78] G. note, “Deep vision for prosthetic grasp,” 9 2019. [Online]. Available: <http://theses.ncl.ac.uk/jspui/handle/10443/4793>
- [79] N. Bu, Y. Bandou, O. Fukuda, H. Okumura, and K. Arai, “A semi-automatic control method for myoelectric prosthetic hand based on image information of objects.” IEEE, 11 2017, pp. 23–28. [Online]. Available: <http://ieeexplore.ieee.org/document/8279702/>
- [80] R. Shima, Y. He, O. Fukuda, N. Bu, H. Okumura, and N. Yamaguchi, “Object shape classification using spatial information in myoelectric prosthetic control,” *International Journal of Computer Software Engineering*, vol. 3, 2018.
- [81] Y. He, R. Shima, O. Fukuda, N. Bu, N. Yamaguchi, and H. Okumura, “Development of distributed control system for vision-based myoelectric prosthetic hand,” *IEEE Access*, vol. 7, pp. 54 542–54 549, 2019.
- [82] Y. He, R. Kubozono, O. Fukuda, N. Yamaguchi, and H. Okumura, “Vision-based assistance for myoelectric hand control,” *IEEE Access*, vol. 8, pp. 201 956–201 965, 2020.
- [83] J. Mouchoux, S. Carisi, S. Dosen, D. Farina, A. F. Schilling, and M. Markovic, “Artificial perception and semiautonomous control in myoelectric hand prostheses increases performance and decreases effort,” *IEEE Transactions on Robotics*, vol. 37, pp. 1298–1312, 8 2021.

- [84] D. Kim, B. B. Kang, K. B. Kim, H. Choi, J. Ha, K. J. Cho, and S. Jo, “Eyes are faster than hands: A soft wearable robot learns user intention from the egocentric view,” *Science Robotics*, vol. 4, pp. 101–121, 1 2019. [Online]. Available: <https://www.science.org>
- [85] M. Strbac, S. Kočović, M. Marković, and D. B. Popović, “Microsoft kinect-based artificial perception system for control of functional electrical stimulation assisted grasping.” *BioMed research international*, vol. 2014, p. 740469, 8 2014. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/25202707><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4151575>
- [86] S. Salminger, H. Stino, L. H. Pichler, C. Gstoettner, A. Sturma, J. A. Mayer, M. Szivak, and O. C. Aszmann, “Current rates of prosthetic usage in upper-limb amputees – have innovations had an impact on device acceptance?” <https://doi.org/10.1080/09638288.2020.1866684>, vol. 44, pp. 3708–3713, 2020. [Online]. Available: <https://www.tandfonline.com/doi/abs/10.1080/09638288.2020.1866684>
- [87] D. K. Kumar, B. Jelfs, X. Sui, and S. P. Arjunan, “Prosthetic hand control: A multidisciplinary review to identify strengths, shortcomings, and the future,” *Biomedical Signal Processing and Control*, vol. 53, p. 101588, 8 2019.
- [88] I. Vujaklija, M. K. Jung, T. Hasenoehrl, A. D. Roche, A. Sturma, S. Muceli, R. Crevenna, O. C. Aszmann, and D. Farina, “Biomechanical analysis of body movements of myoelectric prosthesis users during standardized clinical tests,” *IEEE Transactions on Biomedical Engineering*, 2022.
- [89] N. R. Olsen, J. A. George, M. R. Brinton, M. D. Paskett, D. T. Kluger, T. N. Tully, C. C. Duncan, and G. A. Clark, “An adaptable prosthetic wrist reduces subjective workload,” *bioRxiv*, p. 808634, 10 2019. [Online]. Available: <https://www.biorxiv.org/content/10.1101/808634v1><https://www.biorxiv.org/content/10.1101/808634v1.abstract>
- [90] M. Atzori, A. Gijssberts, S. Heynen, A.-G. M. Hager, O. Deriaz, P. V. D. Smagt, C. Castellini, B. Caputo, and H. Müller, “Building the ninapro database: A resource for the biorobotics community,” 2012. [Online]. Available: <http://www.cyberglovesystems>.

- [91] M. H. Hasbani, D. Y. Barsakcioglu, M. K. Jung, and D. Farina, "Simultaneous and proportional control of wrist and hand degrees of freedom with kinematic prediction models from high-density emg," *2022 44th Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*, pp. 764–767, 7 2022. [Online]. Available: <https://ieeexplore.ieee.org/document/9871346/>
- [92] A. D. Roche, B. Lakey, I. Mendez, I. Vujaklija, D. Farina, Oskar, and C. Aszmann, "Clinical perspectives in upper limb prostheses: An update," 2019. [Online]. Available: <https://doi.org/10.1007/s40137-019-0227-z>
- [93] D. Datta, K. Selvarajah, and N. Davey, "Functional outcome of patients with proximal upper limb deficiency–acquired and congenital," *Clinical Rehabilitation*, vol. 18, pp. 172–177, 3 2004. [Online]. Available: <http://journals.sagepub.com/doi/10.1191/0269215504cr716oa>
- [94] P. J. Kyberd and W. Hill, "Survey of upper limb prosthesis users in sweden, the united kingdom and canada," pp. 234–241, 6 2011.
- [95] N. Chinn, L. Kenney, A. Chadwell, A. Williams, and J. Head, "Use of prostheses in sport by adolescents with upper limb absence," 2018. [Online]. Available: https://issuu.com/nchinn/docs/ispo-conference-poster_a0_with_refe
- [96] E. Scheme and K. Englehart, "Electromyogram pattern recognition for control of powered upper-limb prostheses: State of the art and challenges for clinical use," *Journal of Rehabilitation Research and Development*, vol. 48, pp. 643–660, 2011.
- [97] N. Jiang, K. Englehart, and P. Parker, "Estimating forces at multiple degrees of freedom from surface emg using non-negative matrix factorization for myoelectric control." IEEE, 10 2008, pp. 1–5. [Online]. Available: <http://ieeexplore.ieee.org/document/4712582/>
- [98] N. Jiang, J. L. Vest-Nielsen, S. Muceli, and D. Farina, "Emg-based simultaneous and proportional estimation of wrist/hand kinematics in uni-lateral trans-radial amputees,"

Journal of NeuroEngineering and Rehabilitation, vol. 9, p. 42, 12 2012. [Online]. Available: <https://jneuroengrehab.biomedcentral.com/articles/10.1186/1743-0003-9-42>

- [99] I. Vujaklija, V. Shalchyan, E. N. Kamavuako, N. Jiang, H. R. Marateb, and D. Farina, "Online mapping of emg signals into kinematics by autoencoding," *Journal of NeuroEngineering and Rehabilitation*, vol. 15, pp. 1–9, 3 2018. [Online]. Available: <https://jneuroengrehab.biomedcentral.com/articles/10.1186/s12984-018-0363-1>
- [100] L. H. Smith, T. A. Kuiken, and L. J. Hargrove, "Real-time simultaneous myoelectric control by transradial amputees using linear and probability-weighted regression," *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, vol. 2015–November, pp. 1119–1123, 11 2015.
- [101] N. Jiang, H. Rehbaum, I. Vujaklija, B. Graimann, and D. Farina, "Intuitive, online, simultaneous, and proportional myoelectric control over two degrees-of-freedom in upper limb amputees," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, pp. 501–510, 2014.
- [102] M. Nowak, I. Vujaklija, A. Sturma, C. Castellini, and D. Farina, "Simultaneous and proportional real-time myocontrol of up to three degrees of freedom of the wrist and hand," *IEEE Transactions on Biomedical Engineering*, 2022.
- [103] H. Dantas, T. C. Hansen, D. J. Warren, and V. J. Mathews, "Shared prosthetic control based on multiple movement intent decoders," *IEEE Transactions on Biomedical Engineering*, vol. 68, pp. 1547–1556, 5 2021.
- [104] S. Amsuess, I. Vujaklija, P. Goebel, A. D. Roche, B. Graimann, O. C. Aszmann, and D. Farina, "Context-dependent upper limb prosthesis control for natural and robust use," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, pp. 744–753, 7 2016. [Online]. Available: <https://ieeexplore.ieee.org/document/7153553/>

- [105] A. . P. Pictures, “Primal’s 3d atlas of human anatomy of the forearm and hand.” [Online]. Available: <https://www.anatomy.tv/gotoview.aspx?embeddedcode=759292>
- [106] S. Tanzarella and D. Farina, “A non-invasive human-machine interfacing framework for investigating dexterous control of hand muscles,” 2021. [Online]. Available: <http://spiral.imperial.ac.uk/handle/10044/1/100263>
- [107] A. Gustus, G. Stillfried, J. Visser, H. Jörntell, and P. van der Smagt, “Human hand modelling: kinematics, dynamics, applications,” *Biological Cybernetics*, vol. 106, pp. 741–755, 2012.
- [108] L. Y. Chang, N. S. Pollard, T. M. Mitchell, and E. P. Xing, “Feature selection for grasp recognition from optical markers.” IEEE, 10 2007, pp. 2944–2950. [Online]. Available: <http://ieeexplore.ieee.org/document/4399115/>
- [109] P. Cerveri, E. D. Momi, N. Lopomo, G. Baud-Bovy, R. M. L. Barros, and G. Ferrigno, “Finger kinematic modeling and real-time hand motion estimation,” *Annals of Biomedical Engineering*, vol. 35, pp. 1989–2002, 11 2007. [Online]. Available: <http://link.springer.com/10.1007/s10439-007-9364-0>
- [110] I. Carpinella, J. Jonsdottir, and M. Ferrarin, “Multi-finger coordination in healthy subjects and stroke patients: a mathematical modelling approach,” *Journal Of Neuroengineering*, vol. 19, 2011. [Online]. Available: <http://www.jneuroengrehab.com/content/pdf/1743-0003-8-19.pdf>
- [111] G. Wu, F. C. T. V. D. Helm, M. Makhsous, P. V. Roy, C. Anglin, J. Nagels, A. R. Karduna, K. Mcquade, X. Wang, F. W. Werner, and B. Buchholz, “Isb recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion-part ii: shoulder, elbow, wrist and hand,” *Journal of Biomechanics*, vol. 38, pp. 981–992, 2005. [Online]. Available: www.elsevier.com/locate/jbiomech
- [112] “Full body modeling with plug-in gait - nexus 2.6 documentation - vicon documentation.” [Online]. Available: <https://docs.vicon.com/display/Nexus26/Full+body+modeling+with+Plug-in+Gait>

- [113] M. Duarte and R. N. Watanabe, “Notes on scientific computing for biomechanics and motor control,” 3 2021. [Online]. Available: <https://zenodo.org/record/4599319>
- [114] D. A. Winter, *Biomechanics and Motor Control of Human Movement: Fourth Edition*. John Wiley Sons, Inc., 2009, vol. 7. [Online]. Available: www.wiley.com.
- [115] M. Duarte, “detecta: A python module to detect events in data,” 3 2021. [Online]. Available: <https://zenodo.org/record/4598962>
- [116] F. Pedregosa, V. Michel, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, J. Vanderplas, D. Cournapeau, F. Pedregosa, G. Varoquaux, A. Gramfort, B. Thirion, O. Grisel, V. Dubourg, A. Passos, M. Brucher, M. P. and Édouardand, and Édouard Duchesnay, “Scikit-learn: Machine learning in python gaël varoquaux bertrand thirion vincent dubourg alexandre passos pedregosa, varoquaux, gramfort et al. matthieu perrot,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011. [Online]. Available: <http://scikit-learn.sourceforge.net>.
- [117] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İlhan Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, A. Vijaykumar, A. P. Bardelli, A. Rothberg, A. Hilboll, A. Kloeckner, A. Scopatz, A. Lee, A. Rokem, C. N. Woods, C. Fulton, C. Masson, C. Häggström, C. Fitzgerald, D. A. Nicholson, D. R. Hagen, D. V. Pasechnik, E. Olivetti, E. Martin, E. Wieser, F. Silva, F. Lenders, F. Wilhelm, G. Young, G. A. Price, G. L. Ingold, G. E. Allen, G. R. Lee, H. Audren, I. Probst, J. P. Dietrich, J. Silterra, J. T. Webber, J. Slavič, J. Nothman, J. Buchner, J. Kulick, J. L. Schönberger, J. V. de Miranda Cardoso, J. Reimer, J. Harrington, J. L. C. Rodríguez, J. Nunez-Iglesias, J. Kuczynski, K. Tritz, M. Thoma, M. Newville, M. Kümmerer, M. Bolingbroke, M. Tartre, M. Pak, N. J. Smith, N. Nowaczyk, N. Shebanov, O. Pavlyk, P. A. Brodtkorb, P. Lee, R. T. McGibbon, R. Feldbauer, S. Lewis,

- S. Tygier, S. Sievert, S. Vigna, S. Peterson, S. More, T. Pudlik, T. Oshima, T. J. Pingel, T. P. Robitaille, T. Spura, T. R. Jones, T. Cera, T. Leslie, T. Zito, T. Krauss, U. Upadhyay, Y. O. Halchenko, and Y. Vázquez-Baeza, “Scipy 1.0: fundamental algorithms for scientific computing in python,” *Nature Methods*, vol. 17, pp. 261–272, 3 2020. [Online]. Available: <https://doi.org/10.1038/s41592-019-0686-2>
- [118] B. G. Sgambato, M. H. Hasbani, D. Y. Barsakcioglu, J. Ibanez, A. Jakob, M. Fournelle, M.-X. Tang, and D. Farina, “Ultrasound as a wearable high performance human machine interface for wrist and hand kinematic tracking,” submitted, 2023.
- [119] L. Sun, J. chun Xing, Z. yu Wang, X. Zhang, and L. Liu, “Virtual reality of recognition technologies of the improved contour coding image based on level set and neural network models,” *Neural Computing and Applications*, vol. 29, pp. 1311–1330, 3 2018. [Online]. Available: <https://link.springer.com/article/10.1007/s00521-017-2856-4>
- [120] I. Vujaklija, D. Farina, and O. C. Aszmann, “New developments in prosthetic arm systems,” *Orthopedic Research and Reviews*, vol. 8, pp. 31–39, 7 2016. [Online]. Available: <https://www.dovepress.com/new-developments-in-prosthetic-arm-systems-peer-reviewed-fulltext-article-ORR>
- [121] R. N. Scott and P. A. Parker, “Myoelectric prostheses: State of the art,” *Journal of Medical Engineering and Technology*, vol. 12, pp. 143–151, 1988. [Online]. Available: <https://www.tandfonline.com/action/journalInformation?journalCode=ijmt20>
- [122] Q. M. Marwan, S. C. Chua, and L. C. Kwek, “Comprehensive review on reaching and grasping of objects in robotics,” *Robotica*, vol. 39, pp. 1849–1882, 2021. [Online]. Available: <https://doi.org/10.1017/S0263574721000023>
- [123] D. A. Bennett, S. A. Dalley, D. Truex, and M. Goldfarb, “A multigrasp hand prosthesis for providing precision and conformal grasps,” *IEEE/ASME Transactions on Mechatronics*, vol. 20, pp. 1697–1704, 8 2015.

- [124] T. E. Wiste, S. A. Dalley, H. A. Varol, and M. Goldfarb, "Design of a multigrasp transradial prosthesis," *Journal of Medical Devices, Transactions of the ASME*, vol. 5, 8 2011. [Online]. Available: <https://asmedigitalcollection.asme.org/medicaldevices/article/5/3/031009/475534/Design-of-a-Multigrasp-Transradial-Prosthesis>
- [125] C. L. Taylor and R. J. Schwarz, "The anatomy and mechanics of the human hand." *Artificial limbs*, vol. 2, pp. 22–35, 5 1955.
- [126] C. Sollerman and A. Ejeskär, "Sollerman hand function test: A standardised method and its use in tetraplegic patients," <https://doi.org/10.3109/02844319509034334>, vol. 29, pp. 167–176, 2009. [Online]. Available: <https://www.tandfonline.com/doi/abs/10.3109/02844319509034334>
- [127] R. Azulay and M. Hasbani, "Computer vision system for semi-autonomous control of prehension in myoelectric upper limb prosthesis," 2020.
- [128] A. Saxena, J. Driemeyer, J. Kearns, C. Osondu, and A. Y. Ng, "Learning to grasp novel objects using vision," *Experimental Robotics*, 2006.
- [129] J. Mahler, J. Liang, S. Niyaz, M. Laskey, R. Doan, X. Liu, J. A. Ojea, and K. Goldberg, "Dex-net 2.0: Deep learning to plan robust grasps with synthetic point clouds and analytic grasp metrics," vol. 13, 2017. [Online]. Available: <http://berkeleyautomation.github.io/dex-net>.
- [130] C. Goldfeder, M. Ciocarlie, H. Dang, and P. K. Allen, "The columbia grasp database," 2009, pp. 1710–1716.
- [131] I. Lenz, H. Lee, and A. Saxena, "Deep learning for detecting robotic grasps," *The International Journal of Robotics Research*, vol. 34, pp. 705–724, 4 2015. [Online]. Available: <http://journals.sagepub.com/doi/10.1177/0278364914549607>
- [132] A. Depierre, E. Dellandrea, and L. Chen, "Jacquard: A large scale dataset for robotic grasp detection." Institute of Electrical and Electronics Engineers Inc., 3 2018, pp. 3511–3516. [Online]. Available: <http://arxiv.org/abs/1803.11469>

- [133] J. A. Sethian, “A fast marching level set method for monotonically advancing fronts.” *Proceedings of the National Academy of Sciences*, vol. 93, pp. 1591–1595, 2 1996. [Online]. Available: <https://www.pnas.org/doi/abs/10.1073/pnas.93.4.1591>
- [134] J. A. Sethian and A. M. Popovici, “3-d traveltime computation using the fast marching method,” <https://doi.org/10.1190/1.1444558>, vol. 64, pp. 516–523, 2 2012. [Online]. Available: <https://library.seg.org/doi/10.1190/1.1444558>
- [135] D. L. Chopp, S. J. S. Comput, J. Sethian, and A. Popovici, “Some improvements of the fast marching method,” <https://doi.org/10.1137/S106482750037617X>, vol. 23, pp. 230–244, 7 2006. [Online]. Available: <https://epubs.siam.org/doi/10.1137/S106482750037617X>
- [136] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” vol. 9351. Springer Verlag, 5 2015, pp. 234–241.
- [137] S. Moradi, M. G. Oghli, A. Alizadehasl, I. Shiri, N. Oveisi, M. Oveisi, M. Maleki, and J. Dhooge, “Mfp-unet: A novel deep learning based approach for left ventricle segmentation in echocardiography,” *Physica Medica*, vol. 67, pp. 58–69, 11 2019.
- [138] X. Li, H. Chen, X. Qi, Q. Dou, C. W. Fu, and P. A. Heng, “H-denseunet: Hybrid densely connected unet for liver and tumor segmentation from ct volumes,” *IEEE Transactions on Medical Imaging*, vol. 37, pp. 2663–2674, 12 2018.
- [139] X. Liu, Y. Zhang, H. Jing, L. Wang, and S. Zhao, “Ore image segmentation method using u-net and res_unet convolutional networks,” *RSC Advances*, vol. 10, pp. 9396–9406, 3 2020. [Online]. Available: <https://pubs.rsc.org/en/content/articlehtml/2020/ra/c9ra05877j>
- [140] J. Howard and S. Gugger, “Fastai: A layered api for deep learning,” *Information 2020*, Vol. 11, Page 108, vol. 11, p. 108, 2 2020. [Online]. Available: <https://www.mdpi.com/2078-2489/11/2/108/htmlhttps://www.mdpi.com/2078-2489/11/2/108>
- [141] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” 2015. [Online]. Available: <http://image-net.org/challenges/LSVRC/2015/>

- [142] L. N. Smith, “Cyclical learning rates for training neural networks,” 2017. [Online]. Available: www.cs.toronto.edu/
- [143] —, “A disciplined approach to neural network hyper-parameters: Part 1 - learning rate, batch size, momentum, and weight decay,” 3 2018. [Online]. Available: <https://arxiv.org/abs/1803.09820v2>
- [144] E. Barshan and P. Fieguth, “Stage-wise training: An improved feature learning strategy for deep models,” vol. 44. Proceedings of Machine Learning Research 44, 12 2015, pp. 49–59. [Online]. Available: <https://proceedings.mlr.press/v44/Barshan2015.html>
- [145] G. Garcia-Hernando, S. Yuan, S. Baek, and T. K. Kim, “First-person hand action benchmark with rgb-d videos and 3d hand pose annotations,” 2018, pp. 409–419. [Online]. Available: <https://youtu.be/U5gleNWjz44>
- [146] G. Jocher, A. Stoken, J. Borovec, NanoCode012, A. Chaurasia, TaoXie, L. Changyu, A. V, Laughing, tkianai, yxNONG, A. Hogan, lorenzomammana, AlexWang1900, J. Hajek, L. Diaconu, Marc, Y. Kwon, oleg, wanghaoyang0106, Y. Defretin, A. Lohia, ml5ah, B. Milanko, B. Fineran, D. Khromov, D. Yiwei, Doug, Durgesh, and F. Ingham, “ultralytics/yolov5: v5.0 - yolov5-p6 1280 models, aws, supervise.ly and youtube integrations,” 4 2021. [Online]. Available: <https://zenodo.org/record/4679653>
- [147] T.-Y. Lin, M. Maire, S. Belongie, L. Bourdev, R. Girshick, J. Hays, P. Perona, D. Ramanan, C. L. Zitnick, and P. Dollár, “Microsoft coco: Common objects in context,” 5 2014. [Online]. Available: <http://arxiv.org/abs/1405.0312>
- [148] O. Taheri, N. Ghorbani, M. J. Black, and D. Tzionas, “Grab: A dataset of whole-body human grasping of objects,” 8 2020. [Online]. Available: <http://arxiv.org/abs/2008.11200>http://dx.doi.org/10.1007/978-3-030-58548-8_34
- [149] S. Hampali, M. Rad, M. Oberweger, and V. Lepetit, “Honnotate: A method for 3d annotation of hand and object poses,” 7 2019. [Online]. Available: <http://arxiv.org/abs/1907.01481>

- [150] A. Schwarz, C. Escolano, L. Montesano, and G. R. Müller-Putz, “Analyzing and decoding natural reach-and-grasp actions using gel, water and dry eeg systems,” *Frontiers in Neuroscience*, vol. 14, 8 2020. [Online]. Available: [/pmc/articles/PMC7438923/](https://pmc/articles/PMC7438923/)[https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7438923/](https://pmc/articles/PMC7438923/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC7438923/)
- [151] M. Han, S. Y. Günay, G. Schirner, T. Padır, and D. Erdoğan, “Hands: a multimodal dataset for modeling toward human grasp intent inference in prosthetic hands,” *Intelligent Service Robotics*, vol. 13, pp. 179–185, 1 2020. [Online]. Available: <https://doi.org/10.1007/s11370-019-00293-8>
- [152] M. Cognolato, A. Gijsberts, V. Gregori, G. Saetta, K. Giacomino, A. G. M. Hager, A. Gigli, D. Faccio, C. Tiengo, F. Bassetto, B. Caputo, P. Brugger, M. Atzori, and H. Müller, “Gaze, visual, myoelectric, and inertial data of grasps for intelligent prosthetics,” *Scientific Data*, vol. 7, pp. 1–15, 12 2020. [Online]. Available: www.nature.com/scientificdata
- [153] A. Saudabayev, Z. Rysbek, R. Khassenova, and H. A. Varol, “Data descriptor: Human grasping database for activities of daily living with depth, color and kinematic data streams,” *Scientific Data*, vol. 5, pp. 1–13, 5 2018. [Online]. Available: www.nature.com/sdata/
- [154] M. Gardner, C. S. M. Castillo, S. Wilson, D. Farina, E. Burdet, B. C. Khoo, S. F. Atashzar, and R. Vaidyanathan, “A multimodal intention detection sensor suite for shared autonomy of upper-limb robotic prostheses,” *Sensors*, vol. 20, p. 6097, 10 2020. [Online]. Available: <https://www.mdpi.com/1424-8220/20/21/6097>
- [155] I. M. Guerra, D. Y. Barsakcioglu, I. Vujaklija, D. Z. Wetmore, and D. Farina, “Far-field electric potentials provide access to the output from the spinal cord from wrist-mounted sensors,” 2022. [Online]. Available: <https://doi.org/10.1088/1741-2552/ac5f1a>
- [156] T. R. Farrell and R. F. Weir, “The optimal controller delay for myoelectric prostheses,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 15, pp. 111–118, 3 2007.
- [157] OTBioelettronica.

- [158] A. M. Simon, L. J. Hargrove, B. A. Lock, and T. A. Kuiken, “Target achievement control test: evaluating real-time myoelectric pattern-recognition control of multifunctional upper-limb prostheses,” *Journal of rehabilitation research and development*, vol. 48, pp. 619–628, 2011. [Online]. Available: <https://pubmed.ncbi.nlm.nih.gov/21938650/>
- [159] E. N. Kamavuako, E. J. Scheme, and K. B. Englehart, “On the usability of intramuscular emg for prosthetic control: A fitts’ law approach,” *Journal of Electromyography and Kinesiology*, vol. 24, pp. 770–777, 10 2014.
- [160] L. H. Smith, T. A. Kuiken, and L. J. Hargrove, “Evaluation of linear regression simultaneous myoelectric control using intramuscular emg,” *IEEE Transactions on Biomedical Engineering*, vol. 63, pp. 737–746, 4 2016.
- [161] M. D. Paskett, M. R. Brinton, T. C. Hansen, J. A. George, T. S. Davis, C. C. Duncan, and G. A. Clark, “Activities of daily living with bionic arm improved by combination training and latching filter in prosthesis control comparison,” *Journal of NeuroEngineering and Rehabilitation*, vol. 18, pp. 1–18, 12 2021. [Online]. Available: <https://jneuroengrehab.biomedcentral.com/articles/10.1186/s12984-021-00839-x>
- [162] J. Eden, M. Bräcklein, J. Ibáñez, D. Y. Barsakcioglu, G. D. Pino, D. Farina, E. Burdet, and C. Mehring, “Principles of human movement augmentation and the challenges in making it a reality,” *Nature Communications*, vol. 13, p. 1345, 3 2022. [Online]. Available: <https://www.nature.com/articles/s41467-022-28725-7>
- [163] C. Charissou, D. Amarantini, R. Baurès, E. Berton, and L. Vigouroux, “Effects of hand configuration on muscle force coordination, co-contraction and concomitant intermuscular coupling during maximal isometric flexion of the fingers,” *European Journal of Applied Physiology*, vol. 117, pp. 2309–2320, 11 2017. [Online]. Available: <http://link.springer.com/10.1007/s00421-017-3718-6>

- [164] R. Webster and M. A. Oliver, *Geostatistics for Environmental Scientists*. John Wiley Sons, Ltd, 1 2007. [Online]. Available: <http://doi.wiley.com/10.1002/9780470517277>
- [165] MathWorks Inc., “Mathworks - makers of matlab and simulink.” [Online]. Available: <https://uk.mathworks.com/>