THE UNIVERSITY *of* EDINBURGH

Mitigating Confounding Factors in Myoelectric Control Through Adaptive Modelling and Learning

Katarzyna Szymaniak

School of Informatics

University of Edinburgh

This dissertation is submitted for the degree of

Doctor of Philosophy

January 2026

To my loving parents . . .

Dla moich kochanych rodziców . . .

Lay Summary

Hands are essential for how we interact with and navigate the world. They enable us to perform everyday tasks like writing, eating, and grasping objects, profoundly impacting our independence and quality of life. Losing an upper limb, therefore, disrupts these abilities, as the human hand serves as our primary tool for physical interaction and intricate movements necessary for daily living.

Current advancements in technology have led to myoelectric prostheses, which use signals from remaining muscles to control artificial limbs. However, these devices often struggle to perform consistently in real-world scenarios due to various factors that affect the accuracy of signal decoding. These factors, both internal and external, significantly challenge the reliability and precision of prosthetic control systems.

This thesis addresses these critical limitations by proposing innovative approaches to enhance the adaptability and accuracy of myoelectric control systems. One major focus is on long-term adaptation, where the thesis introduces a novel active learning framework that improves the prosthetic's ability to adapt to changes in muscle signals over time. By employing selective sampling techniques and human-in-the-loop simulations, significant improvements in control accuracy have been achieved with minimal training data.

Another key challenge addressed is the position effect, which refers to how limb positions affect the muscle signals used for prosthetic control. To tackle this, the thesis introduces a comprehensive dataset that captures the variability of muscle signals across different arm positions. In addition, we propose a statistical model,

Domain-Regularised class-conditional Gaussian Mixture Model (DR-cGMM), that demonstrates superior performance in maintaining accuracy across various limb positions compared to traditional deep learning methods.

By overcoming these fundamental challenges, this thesis lays the groundwork for more reliable and adaptable myoelectric prosthetic interfaces. These advancements promise to enhance the quality of life for individuals relying on upper limb prostheses by making them more intuitive and capable of supporting daily activities with greater confidence and precision.

Acknowledgements

No words can truly capture the depth of my gratitude to those who stood by me throughout this long and often challenging journey. This thesis is not just a culmination of academic effort, it is a testament to the unwavering support, love, mentorship, and kindness of so many people who shaped my path along the way. Without them, this would have been impossible.

To my parents, thank you for always being there for me. Your unconditional love and steadfast belief in me, even when I doubted myself, have been my anchor throughout life. Your support has never wavered, and for that, I am endlessly grateful.

To Fady, I do not know how to put into words what your presence has meant to me. Thank you for being there for me like no one else ever was. Your strength carried me in the moments I felt weakest. You pushed me to break out of my shell, to believe in myself, to fight for the life I want. I am so thankful for your unwavering belief in me and the quiet power with which you stood by me.

To Arushi, thank you for being a constant source of light. Your long conversations, honest reflections, and the warmth of your heart have meant more to me than I can say. You've been the one I could always turn to for advice, for laughter, or simply to be heard. Thank you for seeing me through so many ups and downs.

To my collaborators, Kia, Agamemnon, Iris, Henry, and Ondrej, thank you for enriching this journey with your insights, energy, and companionship. Kia, I am immensely grateful for your mentorship and for introducing me to not only the world of academia but also entrepreneurship, creativity, and big ideas. Agamemnon, thank

you for the spark you bring into everything you do. Working with you was both inspiring and energising. Iris, our conversations and collaboration were truly beautiful, thoughtful, creative, and always honest. Henry, thank you for opening my eyes to new ways of thinking, for challenging me to view problems differently. And Ondrej, thank you for the dark humour, the laughs, and the ease you brought into our work.

To my incredible friends, Nikita, Ruchika, and Raman, thank you for your genuine support, the chats that went on for hours, and the amazing food that comforted both the body and soul. Ola, thank you for expanding my bubble tea tastes and for the sweet, meaningful conversations that always left me smiling. Matus, your energy, kind smile, and sharp banter brightened many of my days. Domas, thank you for being the everyday face I looked forward to seeing in the office. You made the mundane feel like something I could anticipate. Arek thank you, especially for the countless airport rides; Klaudia and Lia, thank you for your bright energy and gentle kindness. You all made life richer. Tiffany, your beautiful soul and heartfelt conversations reminded me of parts of myself I had forgotten. Dominik, thank you for the fun, deep chats on AI and entrepreneurship. Connor, thank you for sticking around all these years. Your consistency and friendship meant more than I ever said out loud.

To my lab, thank you for letting me be part of something greater than myself. The shared mission, the collaborations, and the small moments of camaraderie are memories I'll always hold dear. Milad and Aidan, thank you for always being there when the servers went down (because, let's face it, they always did!). Eric, your fresh perspective and good humour made MEC discussions far more enjoyable.

To the people who did me wrong, thank you, too. You taught me lessons in resilience, boundaries, and growth. Without the shadows, the light would not shine so brightly.

To my internship friends, Sasha, Arna, and many more, thank you for the shared moments, the laughter, and the learning. You made those months unforgettable.

To Fabio Stefanini, my manager at Meta Reality Labs, thank you for being such an exceptional mentor. Your guidance, support, and belief in me had a profound impact.

And to Dan Wetmore, thank you for giving me the chance at Meta Reality Labs. Your trust opened the door to one of the most meaningful professional experiences I've had.

This thesis carries my name on the cover, but it is built on the shoulders of a village. Thank you, all of you, for walking this path with me.

Abstract

Despite significant advancements in myoelectric control for upper limb prostheses, which utilise electromyography (EMG) signals to control prosthetic movements, these devices continue to face performance limitations in real-world settings. The decoding of EMG signals is fundamentally challenged by confounding factors, which introduce unwanted variability that alters signal characteristics and degrade the decoding capabilities of the gesture recognition model, leading to performance deterioration. Consequently, myoelectric control systems for upper limb prosthetics struggle to maintain robustness and dexterity in practical applications, resulting in a gap between controlled laboratory settings and everyday use. This thesis addresses these critical limitations through adaptive learning and modelling approaches that can effectively mitigate these confounding factors, with the ultimate goal of developing more reliable prosthetic interfaces.

The thesis focuses on two critical confounding factors: long-term adaptation and the position effect. First, we focus on long-term adaptation to address the challenge of gradual data drift over time. We propose an active learning framework with simulated human-in-the-loop (HITL) methodology to enhance the adaptability of a myoelectric decoder over time. We implement sampling strategies, including least confidence and smallest margin techniques, demonstrating significant performance improvements. With just 3.2 minutes of strategically selected training data, this framework improved accuracy by 4-5% for able-bodied participants and a notable 2% for individuals with

limb differences compared to random sampling. Based on the significance of our findings, we propose a structured pipeline for future real-time deployment.

Second, the position effect refers to variations in limb positions that alter EMG signals by influencing muscle co-activation patterns through gravitational forces and electrode-muscle fiber interactions. To address the position effect, we first introduce GREAT, a comprehensive surface electromyographic dataset that addresses a critical gap in understanding signal variability across different arm positions. The protocol's design extends beyond single-plane movements to encompass the full Cartesian space, thereby significantly enhancing the representativeness of the collected EMG data. The dataset comprises EMG and hand kinematics data collected from 8 able-bodied participants performing 6 distinct hand gestures (power, lateral, pointer, open, tripod, and rest) across 9 unique arm positions arranged in a 3×3 spatial grid. This data provides a robust resource for developing position-invariant myoelectric control algorithms. Our analysis demonstrates significant performance degradation (approximately 10%) during inter-position gesture decoding, underscoring arm position variability critical influence on sEMG signal characteristics.

Building on this foundation, we investigate position domain generalisation by implementing a rigorous benchmark design with nested cross-validation for fair and reliable model comparison. We conduct a comprehensive evaluation across statistical models, including Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA), and deep learning models, such as Empirical Risk Minimisation (ERM), Maximum Mean Discrepancy (MMD), and CORrelation ALignment (CORAL), under a domain generalisation framework. We introduce our Domain-Regularised class-conditional Gaussian Mixture Model (DR-cGMM) that incorporates limb position knowledge via domain covariance regularisation to generalise across varying arm positions. Our analysis concludes that DR-cGMM achieves the highest overall average accuracy 94.29% across all domain generalisation tasks and statistically surpasses deep learning domain generalisation methods in the within-subject, low-data regime.

This thesis advances the field of myoelectric control by directly addressing two critical confounding factors through complementary approaches: active learning for long-term adaptation and domain generalisation through domain-regularised statistical modelling for position effect mitigation. Together, these contributions offer critical insights for enhancing the robustness of myoelectric interfaces. By tackling these fundamental challenges, this thesis lays the groundwork for more reliable, intuitive, and adaptable EMG-based myoelectric control systems that remain robust against the confounding factors encountered in daily prosthetic use.

Contents

Chapter 1

Introduction

The human experience is fundamentally shaped by our physical capabilities, which enable us to interact with and adapt to the world around us. When these physical attributes are compromised, individuals face significant challenges in maintaining independence and quality of life. The loss of an upper limb, in particular, represents a profound disruption, since the human hand serves as the primary interface for physical interaction, enabling motor skills, sensory feedback, and complex manipulations essential for everyday living (Shahsavari et al., 2020). An artificial device like a prosthesis is designed to replace a missing body part to restore its function or appearance, and can be external or implanted depending on the medical need (Guo et al., 2024). The field of prosthetics has evolved significantly over the centuries, from simple wooden replacements to sophisticated robotic appendages (Paré, 1728; Sherman, 1964; Trent et al., 2020). Today, cutting-edge robotics and artificial intelligence enable the creation of assistive technologies that can restore functionality and enhance the quality of life for people with upper limb impairments.

A truly robust and dexterous prosthetic should provide intuitive, seamless control that closely mimics natural hand function. As a non-invasive technique, surface electromyography (sEMG) is the most widely used technique to capture muscle activity through subcutaneous motor action potentials during muscular contraction (Farina et al., 2014). This modality is often incorporated with pattern recognition-based myoelectric control systems, which translate electrical signals generated by

muscle contractions into commands for prosthetic device movement (Englehart and Hudgins, 2003). These systems decode the signals into discrete hand gestures or continuous movements through advanced feature extraction techniques using machine learning algorithms. While advances in machine learning have enabled increasingly sophisticated classifiers, translating benchmarks achieved in controlled laboratory settings into real-world prosthetic control remains challenging (Vujaklija et al., 2017).

One of the most pressing issues in upper limb prosthetics is the high rate of device abandonment, with studies reporting rejection rates as high as 44% (Biddiss and Chau, 2007; Salminger et al., 2022). Research investigating user needs and priorities consistently identifies three primary areas requiring improvement: wearing comfort, device weight, and functional capabilities (Biddiss and Chau, 2007; Østlie et al., 2012; Schultz et al., 2007; Smail et al., 2021). While advances in materials science and mechanical design can address many comfort and weight concerns, functional capabilities, particularly the intuitiveness and reliability of control systems, remain a significant limitation in prosthetic adoption.

Myoelectric control systems often attribute these challenges to the stochastic nature of sEMG signals and the presence of confounding factors that degrade control performance (Li et al., 2016; Roche et al., 2014; Samuel et al., 2019). Confounding factors introduce unwanted variability in the recorded signals, disrupting the feature representation of the muscle activity and decoding capabilities of the model (Eddy et al., 2024). These factors can be intrinsic, such as muscle fatigue and physiological changes, or extrinsic, including electrode displacement, limb position changes, and environmental noise (Cipriani et al., 2011; Farina et al., 2014; Kyranou and Erden, 2018; Pereira et al., 2024; Scheme and Englehart, 2013; Yatsenko et al., 2007). Their combined effect complicates the development of robust control systems and limits the practical usability of myoelectric prostheses.

The impact of these confounding factors extends beyond technical performance; it directly influences user satisfaction and prosthesis acceptance. Variability in signal quality can lead to inconsistent control, increased cognitive load, and frustration, ultimately contributing to device rejection (Amsüss et al., 2013b; Burrough and Brook,

1985; Glynn et al., 1986; Østlie et al., 2012). Among the various confounding factors, two have emerged as particularly critical: limb position variation and long-term adaptability. Limb position effect refers to the alteration of sEMG signal patterns caused by changes in arm posture. Distribution shifts in the signal due to the gravitational forces and electrode-muscle fiber interactions result in a performance decrease for everyday movements (Chen et al., 2011; Fougner et al., 2011; Khushaba et al., 2016; Liu et al., 2012; Scheme et al., 2010). Long-term adaptability addresses the system's ability to maintain performance over time despite physiological changes, electrode shifts, and user learning or fatigue (He et al., 2015a; Hu et al., 2023; Kyranou and Erden, 2018). Without effective adaptation mechanisms, control accuracy deteriorates, undermining user confidence and prosthesis reliability.

Addressing these challenges is essential for advancing myoelectric control toward practical, robust and dexterous prosthetic solutions. This research focuses on mitigating the confounding effects of limb position variation and enhancing long-term adaptability in myoelectric control systems. By improving robustness against these factors, the goal is to bridge the gap between high offline classification accuracies reported in research and the real-world functional performance required for daily prosthesis use. Ultimately, this work aims to contribute to the development of more intuitive, reliable, and acceptable upper limb prostheses that better meet the needs of users in their everyday environments.

1.1 Contributions

We highlight the importance of adaptability in myoelectric control systems as an indispensable component for the dexterity of human-machine interactions. To that end, we investigate confounding factors that contribute to signal modifications, and in consequence, to the quality of EMG-based systems for signal decoding in upper limb prosthetics and wearables. Specifically, we make the following hypotheses:

Hypothesis 1: We hypothesise that active learning using selective sample querying achieves higher classification accuracy with reduced training data compared to

random sampling in myoelectric control systems, demonstrating a data-efficient recalibration approach with implications for long-term adaptation scenarios.

We list our first contribution as follows:

Long-term Adaptation Through Active Learning. We present a novel framework for long-term adaptation of myoelectric control systems for upper limb prosthetics by implementing active learning with human-in-the-loop methodology. Our approach systematically optimises the data acquisition process through uncertainty-based sampling strategies, addressing a critical challenge of reliable and personalised prosthetic control.

Through comprehensive evaluation of least confidence, smallest margin, and entropy reduction sampling techniques in both single and batch-mode configurations, we demonstrate substantial improvements in model adaptation efficiency. Our simulated real-time validation confirms that strategically selected data samples significantly reduce the required training data volume while enhancing decoder performance. With merely 3.2 minutes of actively selected training data, our approach outperforms random sampling by 4-5% for able-bodied participants and approximately 2% for individuals with limb differences.

The experimental results establish that uncertainty-based active learning techniques systematically enhance decoder adaptation while optimising training data requirements on a class-specific basis. Particularly, the smallest margin and the least confidence sampling strategies emerged as superior methods for uncertainty quantification in this domain.

This research introduces the first implementation of an active learning framework specifically designed for long-term adaptation in myoelectric control systems. Our work provides a rigorously tested offline simulation of a closed-loop environment with a structured pipeline for future real-time deployment.

Hypothesis 2: We hypothesise that systematically capturing surface electromyographic signals across a spatial 3 $\times$ 3 grid will demonstrate significant inter-position performance differences, indicating position-dependent signal

variability, and enable the development of standardised data acquisition protocols for position-aware myoelectric control research.

This brings us to our next contribution:

Novel Arm Translation Dataset. We present a novel dataset, GREAT, and propose a protocol for data acquisition across multiple arm positions. The GREAT dataset represents a comprehensive collection of surface electromyographic signals recorded from eight participants performing six distinct hand gestures across various spatial positions in both vertical and horizontal planes. This dataset serves as an essential resource for researchers developing and evaluating position-invariant and position-induced myoelectric control algorithms.

In addition to the dataset, we have established a data acquisition protocol that systematically captures EMG signal variability across diverse arm positions. The protocol's design extends beyond single-plane movements to encompass the full Cartesian space, thereby significantly enhancing the representativeness of the collected EMG data.

Our analysis demonstrates significant performance degradation (approximately 10%) during inter-position classification, underscoring arm positioning's critical influence on sEMG signal characteristics. We confirmed the presence of position-specific information by successfully decoding arm position from sEMG data with 80% accuracy for individual grasps. While our hierarchical approach leveraged position knowledge to enhance gesture classification, the position encoder achieved 58% accuracy. Model performance surpassed 92% only when employing a 'soft' evaluation metric that disregarded position correctness. Under this metric, accuracy depended solely on correct final grasp classification, regardless of initial position identification. Importantly, this performance level was attained when all positions were included in the training data, representing an i.i.d. scenario that is not feasible for real-world deployment. These findings emphasised the discrepancy in gesture recognition performance when the position effect is present.

Finally, that leads us to the last contribution:

Hypothesis 3: We hypothesise that formulating position-induced covariate shift as a domain generalisation problem enables the learning of position-invariant features for accurate myoelectric control across unseen arm positions, and that domain-regularised statistical approaches will demonstrate superior generalisation performance compared to deep learning methods under rigorous benchmarking in low-data regimes.

Position Domain Generalisation. We frame the problem of invariant feature learning as a domain generalisation problem since we assume no access to data from new, unseen positions at deployment time. Subsequently, we propose a novel approach, a Domain-Regularised class-conditional Gaussian Mixture Model (DR-cGMM) that incorporates limb position knowledge via domain covariance regularisation, to generalise across varying arm translations in a low-data, within-subject setting for static limb position variations.

To support fair and unbiased model selection and evaluation, we design a rigorous benchmark framework with nested cross-validation, ensuring that hyperparameter tuning remains isolated from performance estimation. This structure enables principled, leakage-free comparison across both statistical and deep learning approaches, offering a reliable basis for generalisation analysis.

Using this framework, we conduct a comprehensive investigation across statistical and deep learning models under a domain generalisation framework. We provide descriptive analysis and statistical testing to evaluate the significance of the investigated algorithms. Our method, DR-cGMM achieved the highest overall average accuracy of $94.29\% \pm 0.008$ across all DG tasks. Contrary to expectations, Empirical Risk Minimization (ERM) and modern DG techniques fail to outperform simpler baselines (LDA, QDA), particularly in low-data settings. Moreover, deep learning-based DG approaches, Maximum Mean Discrepancy (MMD), and CORrelation ALignment (CORAL) fail to outperform ERM with statistical significance, suggesting that in *static-limb positioning, and low-data* regimes, the complexity and assumptions underlying deep DG techniques may not yield tangible gains over simpler ERM.

Through our statistical analysis, we observed the following: (a) Friedman test indicated significant differences in performance across methods ($p = 0.0009$); (b) LDA and QDA did not exhibit statistical superiority to deep learning methods, reinforcing the observation that statistical methods retain competitiveness in a low-data regime; (c) subsequent Nemenyi post-hoc test showed that our approach, DR-cGMM was uniquely and significantly superior to all evaluated deep learning approaches.

Our study underscores the critical role of statistical modelling in low-data regimes, particularly when enhanced with domain-aware regularisation, for within-subject myoelectric control in static-limb positioning scenarios.

1.2 Thesis Outline

This chapter introduces the readers to the objectives of this thesis. We list the contents of the remaining chapters of this thesis.

- Chapter 2 provides essential background on upper limb prostheses, discussing their challenges and the role of electromyography in recording muscle activity. It covers principles of myoelectric control, machine learning techniques, and explores confounding factors like long-term adaptation and position effects, emphasising responsible research practices.

- Chapter 3 investigates active learning as a scalable approach to improve long-term myoelectric control without frequent recalibration. Using iterative methods and sampling strategies, the study demonstrates that minimal, carefully selected data can significantly enhance decoder performance, paving the way for real-time long-term adaptation.

- Chapter 4 introduces the GREAT dataset. This chapter offers a comprehensive collection of EMG signals from multiple arm positions, aimed at developing position-invariant decoding algorithms. It also proposes a novel data acquisition protocol to facilitate future research on EMG variability across different arm configurations.

- Chapter 5 explores the potential of domain generalisation algorithms to extract invariant features across different arm positions, especially under a low-data regime. It represents a pioneering effort to address position-induced variability in myoelectric control through this framework.

- Chapter 6, the final chapter summarises key findings and outlines promising avenues for future research to advance the field further.

The publications included in this thesis are ($*$ denotes equal contribution):

- Szymaniak K, Krasoulis A and Nazarpour K (2022) Recalibration of myoelectric control with active learning. Front. Neurorobot. 16:1061201.
  doi: 10.3389/fnbot.2022.1061201 [Chapter 3]

  I led this work in all aspects: designing the experiment, data analysis, writing the manuscript, and proposing a pipeline for future real-time deployment of a human-in-the-loop setting with active learning. Agamemnon Krasoulis collected the data, and Kianoush Nazarpour supervised the project and contributed to writing the document.

- * Szymaniak, K. * Kyranou, I., and Nazarpour, K. EMG Dataset for Gesture Recognition with Arm Translation. Sci Data 12, 100 (2025).
  doi: 10.1038/s41597-024-04296-8 [Chapter 4]

  All authors collaboratively designed the data acquisition protocol, incorporating the novel grid component. The implementation of the data collection code was carried out by Iris Kyranou, who also collected data from all participants following my initial data collection phase. Subsequently, I took responsibility for implementing data quality checks, preprocessing procedures, and developing the machine learning pipeline for signal analysis within this novel context of investigating position effects. Throughout the project, Kianoush Nazarpour provided supervision and contributed to the preparation of the manuscript.

Chapter 2

Background

This chapter examines upper limb prosthetic fundamentals, highlighting current solutions and challenges. We explore electromyography for muscle activity recording, data preprocessing protocols, and feature extraction methodologies. We detail myoelectric control systems and machine learning applications before presenting a comprehensive taxonomy of confounding factors, with particular emphasis on long-term adaptation and position effects. The chapter concludes by highlighting the importance of responsible research practices in this field.

2.1 Upper Limb Prostheses

2.1.1 Upper Limb Loss

Upper limb amputation statistics revealed 1.6 million cases in 2005, with projections suggesting this number could reach twice that by 2050. The majority of amputations in the US (82%) stem from vascular conditions (Behrend et al., 2011; Shahsavari et al., 2020). In particular, for upper limb loss, trauma is listed as the leading cause representing about 68.8% of all cases. Workplace-related accidents account for the vast majority (more than 90%) of these trauma-induced amputations, predominantly for males (Fitzgibbons and Medvedev, 2015). Subsequent causes include mostly neoplasia and dysvascularity. Upper limb anomalies affect approximately 15 in every 100,000 newborns. Among those affected by upper extremity deformities, significant

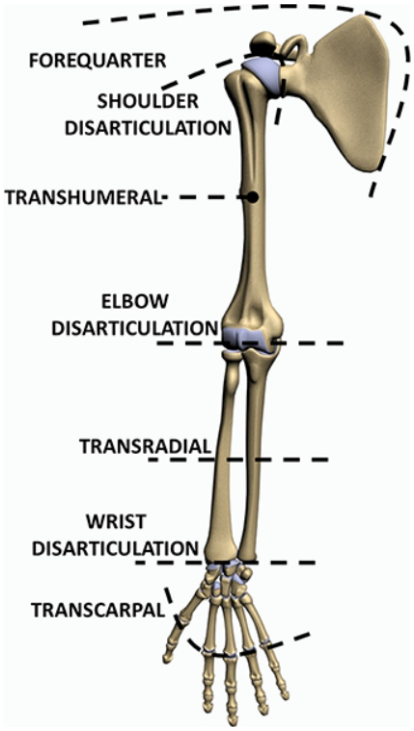


Figure 2.1 Levels of upper limb amputation/absence. Figure has been adapted from (Cordella et al., 2016) and is distributed under a CC BY 4.0 International License (https://creativecommons.org/licenses/by/4.0/).

cases comprise 8% of the population, representing around 41,000 individuals (Ziegler-Graham et al., 2008).

The severity and location of upper limb loss varies significantly, with Figure 2.1 depicting different levels of amputation: from trans-carpal at the hand level, through wrist disarticulation, trans-radial, elbow disarticulation, and trans-humeral, up to the more proximal shoulder disarticulation and forequarter amputation (Cordella et al., 2016; Maduri and Akhondi, 2019). When an amputation occurs, it disrupts the closed-loop brain's natural feedback system, breaking the connection between motor control signals and sensory feedback pathways. Studies indicate that the loss of an upper limb has a more profound impact on quality of life than lower limb amputation (Behrend et al., 2011).

(a) (b) (c) (d)

Figure 2.2 Commercially available prosthetic hands. (a) Vincent Hand (Vincent Systems, 2022); (b) i-Limb (Van Der Niet and van der Sluis, 2013); (c) Taska Hand (Khan and Desmulliez, 2019); (d) Bebionic Hand (Cordella et al., 2016; Mastinu et al., 2019). The figure has been adapted from (Guo et al., 2024) and is distributed under a CC BY 4.0 International License (https://creativecommons.org/licenses/by/4.0/).

2.1.2 Upper Limb Prosthetic Devices

The first known prosthetic device was found in ancient Egypt in the form of an artificial toe (Finch, 2011). Later, an iron hand was made for Marcus Sergius after he lost his hand in the Second Punic War (218-201 B.C.). In the 16th century, Götz von Berlichingen owned the first advanced example of a prosthetic arm equipped with jointed fingers enabling grasping a sword (Zuo and Olson, 2014). Around the same period, Ambroise Paré designed a prosthetic hand with catches and springs (Paré, 1728). During World War II time (1948) Munich University physics student, Reinhold Reiter, developed the first myoelectric prostheses that amplified sEMG to power motorised parts using a vacuum tube amplifier (Reiter, 1948). However, this invention did not gain popularity and only in 1960 did Russian scientist, Alexander Kobrinski, make a significant advancement in myoelectric prosthesis. He incorporated transistors, allowing a more compact, portable design with electronics and batteries worn on a belt, connected by wires to the prosthesis, which also included a skin-toned cosmetic glove (Sherman, 1964). Despite issues such as heaviness, slow movement, weak pinch force, vulnerability of wire connections, and reliability challenges due to electrical interference, these scientific advancements became the foundation for modern solutions in upper limb prosthetics (Zuo and Olson, 2014).

Currently, prostheses can be categorised depending on the type of power source as passive, body-powered, electrically powered, hybrid and activity/task-specific (Trent et al., 2020). Passive prostheses restore limb length for carrying or stabilising objects and provide positional support, yet they lack active movement. In contrast, body-powered prostheses use a harness-driven cable system to control movement, offering advantages such as durability, lower cost, and proprioceptive feedback; however, they can restrict movement and cause discomfort. Electrically powered prostheses, on the other hand, use motors controlled by electromyography (EMG) signals to enable movement, powered by rechargeable batteries. Additionally, hybrid prostheses combine body-powered and electrically powered elements to balance strength, functionality, and control. Finally, activity-specific prostheses are designed for particular tasks or activities, optimised to perform effectively in specific functions or environments.

Commercially available prosthetic hands exhibit distinct features tailored to user needs. The Vincent Hand (Figure 2.2a) is cost-effective and durable, providing multiple grip settings but lacks sensory feedback, which may limit dexterity. The i-Limb (Figure 2.2b) offers individually powered fingers for complex movements and customizable grip patterns via a mobile app, though it comes with a higher price and maintenance requirements. The Taska Hand (Figure 2.2c), the first waterproof prosthetic, is robust and designed for wet conditions, but its weight can cause fatigue, and its cost may restrict accessibility. Lastly, the Bebionic Hand (Figure 2.2d) features extensive grip options and customisation, but its complexity may pose challenges for new users, alongside significant power demands requiring frequent recharging. Each model is engineered to meet specific functional requirements and enhance user experience in various applications (Guo et al., 2024).

2.1.3 Prosthesis Rejection

The acceptance of upper limb prostheses remains a significant challenge in prosthetic development. Multiple factors influence user adoption, including device design, functional capabilities, usability, comfort level, longevity, and aesthetic qualities. Despite technological advances since Biddiss' early work (Biddiss and Chau, 2007),

current research shows rejection rates as high as 44% (Salminger et al., 2022). Studies consistently identify three primary areas requiring improvement: wearing comfort, device weight, and functional capabilities (Biddiss and Chau, 2007; Østlie et al., 2012; Schultz et al., 2007; Smail et al., 2021).

In Section 2.4, various factors that influence functional capabilities are described, with particular emphasis placed on adaptation over time (Section 2.5) and variations in limb positioning (Section 2.6). The section provides a foundation for understanding how these confounding factors impact overall functionality assessment.

2.2 Electromyography

The predominant way myoelectric control for upper limb prosthetic devices operates is through muscle activity patterns gathered from the skin surface using EMG. Its history can be traced back to 1771 when Galvani demonstrated that animal muscles would contract in response to electrical stimuli (Kazamel and Warren, 2017). By 1950, commercially available systems were introduced, and since 1993, EMG data acquisition, processing, and documentation became ubiquitous (Ladegaard, 2002).

This section provides background on the underlying muscle physiology of EMG signal and discusses the signal analysis approaches used in myoelectric control systems.

2.2.1 Neurophysiology of Movement

Motor control is the process by which the nervous system regulates and coordinates movements. It involves complex interactions between various neural subsystems from motor regions in the cortex to the spinal cord, including the brainstem, basal ganglia, and cerebellum (Purves, 2012).

In 1925, Charles Sherrington introduced the concept of the motor unit. He recognised that a motor neuron and the skeletal muscle fibres it innervates using the neuron's axon terminals, function as an inseparable anatomical and functional unit (Buchthal and Schmalbruch, 1980). The term *unit* reflects how all muscle fibers within a motor

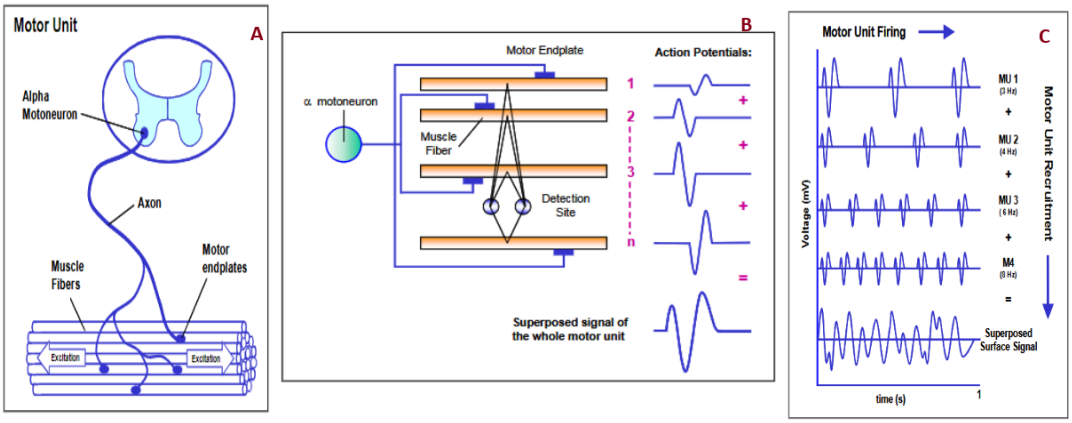


Figure 2.3 Figure presents: A. Motor unit (Basmajian, 1985); B. MUAPs superposition (Basmajian, 1985); C. Surface signal (Kumar, 2017). The figure has been adapted from the cited sources and is distributed under a CC BY 4.0 International License (https://creativecommons.org/licenses/by/4.0/).

unit activate together during nerve stimulation. Specifically, when motor neurons fire action potentials through axons presynaptic terminal to a motor endplate, also known as a neuromuscular junction, they cause all their associated muscle fibers to contract, as shown in Figure 2.3 (Basmajian, 1985; Enoka, 1994).

These contractions create force against the skeleton, leading to body movement in isotonic contractions. The magnitude of the force created by the muscle is enforced by the number of active motor units and their firing rates. Motor units contain varying numbers of muscle fibers; this variation exists within individual muscles and between different muscles, as shown by Buchthal and Schmalbruch (1980). In addition, muscles typically contain hundreds of motor neurons and tens to hundreds of thousands of fibers (Feinstein et al., 1955). These physiological structures generate electrical signals that can be captured and analysed through electromyography.

2.2.2 Composition of Electromyographic Signal

The electromyographic signal represents the electrical activity produced by skeletal muscles during contraction. Recording can be done either noninvasively using surface electrodes on the skin or invasively by placing electrodes directly in muscle tissue.

Since surface EMG is noninvasive and carries minimal risk to the subject, it is a widely used technique for measuring myoelectric activity.

The EMG signal is the sum of the electrical activity of the muscle fibers, created by the superposition of motor unit action potentials (MUAPs) (Farina et al., 2014). Based on the recruitment order described by Henneman's size principle (Uchida and Delp, 2021), and the firing frequency, the magnitude and density of the signal can be modulated to produce the force output of a muscle, Figure 2.3C. EMG signal can be represented by the following equation:

$$x(n) = \sum_{r=0}^{N-1} h(r)e(n-r) + w(n). \tag{2.1}$$

In this context, $x(n)$ denotes the modelled EMG signal, $e(n)$ is the point-processed firing impulse, $h(r)$ represents the MUAP, $w(n)$ is zero-mean additive white Gaussian noise, and N indicates the total number of motor unit firings (Reaz et al., 2006). The signal's amplitude, time, and frequency domain characteristics are influenced by various intrinsic and extrinsic factors. Some of these include muscle contraction timing and intensity, electrode placement relative to the active muscle area, properties of overlying tissues (such as skin thickness), electrode and amplifier specifications, and the quality of electrode-skin contact (more details in Section 2.4).

Usually, the signal gathered from the electrode is amplified to improve the signal-to-noise ratio (SNR) by making the signal amplitude much larger than the noise floor of subsequent processing stages (Gerdle et al., 1999). To enhance signal quality, various filtering techniques can be employed to eliminate unwanted noise. Specifically, power line interference can be mitigated using a Notch filter, while low-frequency and high-frequency noise can be effectively removed using filters such as the Butterworth filter. These filtering methods contribute to the overall improvement of the signal-to-noise ratio, facilitating more accurate analysis of the EMG data (Reaz et al., 2006).

2.2.3 Signal Processing for Prosthetic Control

Signal preprocessing and feature extraction are imperative for the dexterity of machine learning-based myoelectric control systems. Given the data-driven nature of machine learning algorithms, the representation of EMG features plays a pivotal role in creating myoelectric control interfaces.

Due to the stochastic nature of the EMG signal, feature extraction commonly employs sliding windows to segment the signal into sequential batches. This approach can be implemented using either disjoint or overlapping windows. Disjoint windows place each segment's starting point immediately after its predecessor's endpoint while overlapping windows include signal samples in multiple consecutive segments (Xiong et al., 2021).

Numerous studies have conducted exhaustive analyses to evaluate and compare various EMG features, contributing significantly to our understanding of their efficacy in different contexts (Boostani and Moradi, 2003; Englehart et al., 1999; Phinyomark et al., 2018, 2012, 2013; Phinyomark and Scheme, 2018b; Xiong et al., 2021; Zardoshti-Kermani et al., 1995). These investigations have generally categorised EMG features into three main domains: time-domain, frequency-domain, and time-frequency domain.

Let N denote the length of the window with raw EMG signal $x = x_1, x_2, \ldots, x_N$, and $\bar{x} = \frac{1}{N}\sum_{n=1}^{N} x_n$ be the empirical mean of the vector x_N. Some examples of the most frequently employed time-domain features include:

- **Mean absolute value:**

$$MAV = \frac{1}{N}\sum_{n=1}^{N} |x_n|. \tag{2.2}$$

- **Variance:**

$$VAR = \frac{1}{N-1}\sum_{n=1}^{N} (x_n - \bar{x})^2. \tag{2.3}$$

- **Log-variance:**

$$LogVar = \log\left(\frac{1}{N-1}\sum_{n=1}^{N} (x_n - \bar{x})^2\right). \tag{2.4}$$

- **Root mean square:**

$$RMS = \sqrt{\frac{1}{N}\sum_{n=1}^{N} x_n^2}. \tag{2.5}$$

- **Waveform length:**

$$WL = \sum_{n=1}^{N-1} |x_{n+1} - x_n|. \tag{2.6}$$

- **Wilson amplitude:**

$$WAMP = \sum_{n=1}^{N-1} f(|x_{n+1} - x_n|), \tag{2.7}$$

where

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

- **Slope sign change:**

$$SSC = \sum_{n=2}^{N-1} f\left((x_n - x_{n-1})(x_n - x_{n+1})\right), \tag{2.8}$$

where

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

- **Zero-crossing:**

$$ZC = \sum_{n=1}^{N-1} \left[\text{sgn}(-x_n x_{n+1}) \cap |x_n - x_{n+1}| \geq \text{threshold}\right], \tag{2.9}$$

where

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

Frequency-domain features provide valuable insights that can complement time-domain characteristics (Azhiri et al., 2021a). Let M denote the number of frequency components, f_k the frequency at component k, and $P_k = |X(f_k)|^2$ the power spectral

density at f_k, where $X(f_k)$ is the Fourier transform of the EMG signal x. Some frequency-domain features include:

- **Skewness:**
$$SKW = \frac{\sum_{k=1}^{M}(f_k - MNF)^3 P_k}{\sum_{k=1}^{M} P_k}, \tag{2.10}$$
where MNF is the mean frequency defined below.

- **Mean frequency:**
$$MNF = \frac{\sum_{k=1}^{M} f_k P_k}{\sum_{k=1}^{M} P_k}. \tag{2.11}$$

- **Peak frequency:**
$$PKF = \arg\max_{f_k} P_k. \tag{2.12}$$

- **Variance of the central frequency:**
$$VC = \frac{\sum_{k=1}^{M}(f_k - MNF)^2 P_k}{\sum_{k=1}^{M} P_k}. \tag{2.13}$$

Interestingly, some time-domain features like zero-crossing rate and waveform length also capture frequency-related information. Time-frequency domain analysis, employing techniques such as short-time Fourier transform and wavelet transforms, offers a dynamic view of the signal's spectral content over time.

Consequently, careful consideration and optimisation of EMG feature extraction methods are essential for learning pattern recognition, as the quality and relevance of these extracted features impact how effectively myoelectric control systems can translate signals into intended prosthetic device commands.

2.3 Myoelectric Control

Myoelectric control, which uses EMG signals to operate devices (Oskoei and Hu, 2007), has evolved primarily through prosthetic applications for people with limb differences. Early systems employed a simple one-muscle one-function approach with antagonistic muscles acting as threshold-activated switches (Scheme and Englehart,

2011). This restricted control scheme employed non-intuitive and sequential mode switching across the list of gestures, necessitating expanding functionality (Farina et al., 2014). To overcome these limitations, pattern recognition was introduced (Englehart and Hudgins, 2003; Parker et al., 2006), utilising machine learning to detect EMG signal patterns across multiple muscle sites.

Pattern recognition-based myoelectric control emerged in the mid-to-late 1900s to address these shortcomings (Englehart and Hudgins, 2003) by incorporating advanced feature extraction and algorithms to capture how muscles work synergistically and interpret their activity patterns. Conventionally, the EMG-based pattern recognition pipeline includes data preprocessing to filter noise, segmenting data into overlapping windows, extracting features to enhance information density, and a machine learning model to discriminate between patterns of muscle co-activations.

Depending on the event generation patterns, EMG-based control schemes can be categorised as either continuous or discrete (Eddy et al., 2023). Continuous control generates a sequence of events (based on EMG window length and stride/increment) to create a one-to-one mapping of each event/window to the decision event, enabling fluid control of prosthetics or cursors but requiring sustained user focus. In contrast, discrete control generates a single event once a gesture sequence is completed (i.e. many-to-one mapping), offering increased robustness against unintended activations while reducing cognitive demands (Scheme and Englehart, 2011).

Since continuous control has a one-to-one mapping and adjacent windows are usually extracted in very short temporal increments (typically in milliseconds), giving an assumption of the static contractions, classification using stationary models is utilised (Eddy et al., 2023). Due to its simplicity and low-data requirements, traditional machine learning approaches have been heavily investigated in the research community. Algorithms such as Linear Discriminant Analysis (LDA) (Campbell et al., 2019), Quadratic Discriminant Analysis (QDA) (Ghojogh and Crowley, 2019), Support Vector Machines (SVM) (Chang and Lin, 2011), Random Forest (RF) (Biau and Scornet, 2016) and K-Nearest Neighbour (KNN) (Jiang et al., 2007) have shown promising

results. However, as the complexity of the tasks increases, deep learning solutions such as Artificial Neural Networks (ANN) (Azhiri et al., 2021a) and Convolutional Neural Networks (CNN) (Zhai et al., 2017) are utilised. Additionally, temporal models such as Recurrent Neural Networks (RNN) (Azhiri et al., 2021b), Long Short-Term Memory (LSTM) (Jabbari et al., 2020), Temporal Convolutional Networks (TCN) (Bai et al., 2018) and Dynamic Time Warping (DTW) (Huang et al., 2010) have been proposed in myoelectric control.

2.3.1 Traditional Machine Learning

We briefly mention popular statistical approaches that have been used in this work. The challenge of EMG-based gesture recognition differs fundamentally from conventional pattern recognition tasks due to the non-stationary, stochastic nature of EMG signals and their susceptibility to confounding factors, including electrode shift, muscle fatigue, limb position variability, and inter-subject anatomical differences.

Linear Discriminant Analysis (LDA) is a well-known and established decoder for myoelectric control due to its computational efficiency and interpretability, critical factors for embedded prosthetic systems with limited processing power. Unlike general classification tasks, EMG control requires sub-100ms response times to maintain natural interaction Farrell and Weir (2007). LDA's closed-form solution and linear decision boundaries enable this rapid inference without iterative optimisation.

One of the key assumptions of LDA is that all classes share a common covariance matrix. In the context of myoelectric control, these classes correspond to different grasp types. This shared covariance matrix is often referred to as the pooled covariance or within-class scatter matrix, which leads to linear decision boundaries (Fisher, 1936; Hastie et al., 2009). Based on the training data D_s, LDA models the data using the prior probability π_c, and the class mean μ_c for each class $c \in \{1,\ldots,C\}$, and the pooled

covariance matrix Σ. These are given by:

$$\hat{\pi}_c = \frac{|D_c|}{|D_s|}, \tag{2.14}$$

$$\hat{\mu}_c = \frac{1}{|D_c|} \sum_{y_n=c} \mathbf{x}_n, \tag{2.15}$$

$$\hat{\Sigma} = \frac{1}{|D_s| - C} \sum_{c=1}^{C} \sum_{y_n=c} (\mathbf{x}_n - \hat{\mu}_c)(\mathbf{x}_n - \hat{\mu}_c)^\top \tag{2.16}$$

where $|D_c|$ denotes the number of samples in D_s where y_i equals the specified class is given by $|D_c| = \sum_{i=1}^{|D_s|} \mathbb{I}(y_i = c)$, where c is a class and $\mathbb{I}(\cdot)$ is the indicator function.

Based on these estimate, a new test sample $\mathbf{x}^*$ is assigned to the class that maximizes the discriminant function $\underset{c}{\operatorname{argmax}}\ \delta_c(\mathbf{x}*)$, i.e,

$$\delta_c(\mathbf{x}*) = \mathbf{x}*^\top \Sigma^{-1} \mu_c - \frac{1}{2} \mu_c^\top \Sigma^{-1} \mu_c + \log \pi_c, \tag{2.17}$$

EMG-Specific Limitations: While LDA performs well under controlled conditions, its linear decision boundaries struggle with the heterogeneous (non-linear) variability of real-world sEMG data (due to electrode shift or muscle fatigue), which violates the shared covariance assumption. Additionally, muscle fatigue introduces time-varying drift in the feature space that LDA cannot adapt to without retraining.

Another closely related approach, *Quadratic Discriminant Analysis* (QDA), is a class-conditional Gaussian model that does not assume a shared covariance matrix Hastie et al. (2009). Instead, it computes a separate covariance matrix Σ_c for every class c. This leads to quadratic decision boundaries, which, in theory, can model more complex class distributions than LDA.

Both LDA and QDA belong to the family of supervised generative models, which assume class-conditional multivariate Gaussian densities (Hastie et al., 2009). These two methods represent opposite ends of the spectrum in how they define the covariance structure. Under this assumption, the probability density function (PDF) of a point x given class c is:

$$p(x \mid y = c) = \mathcal{N}(x; \mu_c, \Sigma_c) \tag{2.18}$$

Practical Trade-off: In EMG-based classification, the choice between LDA and QDA represents a fundamental bias-variance trade-off. Studies have shown LDA often outperforms QDA in within-session evaluation but degrades more severely across sessions due to its inability to capture class-specific non-stationarity (Campbell et al., 2020b).

These algorithms are thoroughly studied in Chapter 5.4. Therefore, more details can be found therein.

2.3.2 Deep Learning

Artificial Neural Networks (ANNs) implement computational models inspired by biological neural systems to approximate complex functions. A fundamental ANN architecture, the Multi-Layer Perceptron (MLP), consists of an input layer, multiple hidden layers, and an output layer, as depicted in Figure 2.4. Each neuron applies a non-linear activation function f to a weighted sum of its inputs:

$$y = f\left(\sum_{i=1}^{n} w_i x_i + b\right), \tag{2.19}$$

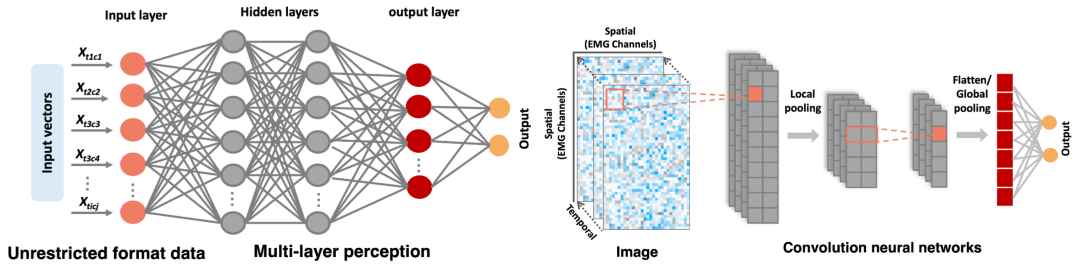
where w_i are the weights, x_i are the inputs, and b is the bias.

Using gradient-based optimisation techniques, training MLPs involves minimising a non-convex loss function $L(\theta)$, such as cross-entropy or mean squared error. This process updates the network's parameters θ, typically weights w_i, to minimise the loss function. The update rule follows from the gradient descent method:

$$\theta^{(t+1)} = \theta^{(t)} - \eta \nabla L(\theta), \tag{2.20}$$

where η is the learning rate, and $\nabla L(\theta)$ denotes the gradient of L with respect to θ. This gradient is computed using the chain rule, propagating errors backwards through the network. Optimisation aims to adjust the model's parameters to improve its performance on a given task. It iteratively refines the network's predictions by reducing the discrepancy between predicted and actual outputs, thereby enhancing

the model's ability to generalise to unseen data. Hyperparameters such as learning rate, network depth, width (number of neurons per layer), and batch size significantly influence convergence and generalisation capabilities. In low-data regimes, particularly, careful tuning of these hyperparameters is crucial to prevent issues such as overfitting or vanishing/exploding gradients. Techniques like gradient clipping or using smaller batch sizes can mitigate these problems.



(a) (b)

Figure 2.4 Figure (a) presents an artificial neural network, while Figure (b) depicts an architecture example of a convolutional neural network.

For EMG signals, MLPs traditionally process data in an unrestricted format. This means they treat input features (often hand-crafted features or raw samples) as an unordered collection, effectively disregarding the inherent spatial and temporal arrangements of the EMG signals. While simple, this approach may overlook valuable spatio-temporal dependencies present in the muscle contractions.

While Multi-Layer Perceptrons (MLPs) lack spatial awareness, *Convolutional Neural Networks (CNNs)* overcome these challenges by introducing hierarchical feature extraction through convolutional layers, enabling the model to capture spatial dependencies while maintaining computational efficiency.

Unlike fully connected layers, which process all input features simultaneously, CNNs utilize kernel-based convolutions to extract localized features, allowing the network to focus on spatial hierarchies in the data. A convolutional layer applies a learnable filter K over an input X, producing a feature map Y:

$$Y_{ij} = \sum_m \sum_n K_{mn} X_{(i+m)(j+n)} + b, \tag{2.21}$$

where Y_{ij} represents the output feature map at position (i, j), computed as the weighted sum of input values $X_{(i+m)(j+n)}$ using the learnable filter K_{mn}, where m and n define the kernel's receptive field, plus a bias term b. During training, the formulation of the kernel space is learned, effectively acting as the weight set to be optimized using back-propagation. To capture complex hierarchical features, multiple kernels are typically stacked within each convolutional layer, enabling deeper feature representations. Each convolutional layer is typically followed by a non-linear activation function, such as ReLU, to introduce non-linearity, and often by a pooling layer, which reduces spatial dimensions while retaining essential features.

CNNs are particularly effective for structured data representations such as images and time-series signals. In 1D CNNs, convolutional filters slide along a single dimension, making them well-suited for sequential data such as low-density electromyography (EMG) signals recorded from linearly arranged electrodes. 2D CNNs process grid-arranged data such as images or high-density EMG signals with spatially mapped electrodes, where shallow layers detect local features while deeper layers form abstract spatial representations. Variations of CNNs have been explored in myoelectric control to enhance spatial feature representations (Hu et al., 2018; Wang et al., 2024; Wei et al., 2019; Zou and Cheng, 2021).

Depending on data representation, different architectures extract meaningful features from EMG signals. For temporal sequencing, recurrent networks (LSTMs/GRUs) (Khushaba et al., 2020; Zhang et al., 2023a) and Temporal Convolutional Networks (TCN) (Tsinganos et al., 2019; Zanghieri et al., 2019) effectively model time dependencies in signal waveforms. Spectral transformation approaches convert signals to visual representations through techniques like Short-Time Fourier Transform (STFT) (Tripathi et al., 2023; Yamanoi et al., 2020), wavelet transforms (Buelvas et al., 2023; Oh and Jo, 2021), or the Mel spectrograms (Qureshi et al., 2022), enabling powerful 2D convolutional architectures to identify spatial patterns (Dere and Lee, 2023). Hybrid frameworks combine paradigms, such as CNN-MLP-Mixer systems (Shen et al.,

2024) that process signals through convolutional feature extraction before applying mixing operations for robust classification.

2.4 Confounding Factors in Myoelectric Control

Having explored various architectures for extracting meaningful signal representations, it is crucial to acknowledge that pattern recognition-based myoelectric control relies on data quality to ensure system reliability and robustness, directly influencing prosthesis abandonment rates and user satisfaction (Burrough and Brook, 1985; Eddy et al., 2024; Glynn et al., 1986; Østlie et al., 2012; Tabor et al., 2018). In this context, confounding factors represent variables that introduce unwanted signal variations, degrading classification accuracy and interfering with the intended relationship between muscle activity and prosthetic control commands. Performance degradation due to these confounding factors presents a significant challenge in myoelectric systems.

These factors can be broadly categorised based on their origin as either *intrinsic* or *extrinsic*, each introducing distinct signal variations that compromise functional performance. Additionally, *compounded factors* are distinguished as a hybrid of the former two confounding types.

2.4.1 Intrinsic Factors

Intrinsic factors emerge in the process of signal generation, when MUAPs are created through the temporal and spatial superposition of individual muscle fiber action potentials. The amplitude and shape of these MUAPs serve as distinctive indicators of motor unit properties, including functionality, fiber arrangement, and fiber diameter. These intrinsic variations contribute significantly to the diverse characteristics observed across muscles, directly influencing critical performance factors such as strength, endurance, muscle topology and recruitment patterns. Several elements, including individual variations, contraction intensity, muscle fatigue, and limb position, can be categorised as intrinsic factors.

While *contraction intensity* in a controlled environment provides consistent EMG signals (repeated measurements achieving correlation coefficients of approximately 0.98), it remains challenging once deployed in clinical applications (Buskirk and Komi, 1970; Moritani and DeVries, 1978, 1979; Moritani and Muro, 1987). In real-world scenarios, contraction intensity varies naturally based on the anticipated effort required for a given task. This variability directly affects EMG amplitude through both linear and non-linear relationships, while also altering frequency characteristics, which, although used to proportional control to influence the speed to force (Scheme and Englehart, 2013; Yatsenko et al., 2007), remains challenging in practical settings.

Prolonged prosthetic usage inevitably leads to *muscle fatigue*, manifesting as a temporary decrease in the ability to perform physical movements. Fatigue development can be measured by tracking the reduction in maximum muscle force or power output, which causes EMG electrode readings to change over time (Kyranou and Erden, 2018). While during data acquisition protocol fatigue is often minimised by keeping trials short and scheduling adequate rest periods between them, this approach is not practical for real-world prosthetic device users who need continuous functionality.

Changes in *arm posture* throughout daily activities dramatically alter underlying muscle dynamics. These postural variations modify muscle topology and recruitment patterns under different conditions (Kyranou and Erden, 2018). This concept is well observed in small wrist deviations (supination or pronation), but also investigated for large position shifts which occur in activities of daily living (ADL), such as grabbing a glass from a table or a book from a top shelf. In Section 2.6, limb position effect is described in more detail.

Lastly, signal contamination through *muscle crosstalk* occurs when electrodes capture unintended activity from adjacent muscles (Farina et al., 2014). This interference is particularly problematic in residual limbs post-amputation, where altered anatomy increases signal overlap between neighbouring muscle groups (Scheme and Englehart, 2011). Crosstalk severity depends on electrode configuration, inter-electrode distance, and tissue composition, significantly impacting classification accuracy (Germer et al., 2021; Mesin, 2020; Winter et al., 1994). Modern mitigation approaches include high-

density electrode arrays, spatial filtering techniques, and machine learning algorithms specifically designed to function despite signal contamination (Muceli et al., 2015).

2.4.2 Extrinsic Factors

Extrinsic factors encompass all external variables that interfere with EMG signal acquisition during the data collection process, such as electrode shifts, skin conditions, variation in collection devices, variation in downstream task requirements, external noises and mechanical load.

Among these factors, *electrode placement* is particularly crucial, as it is present in the everyday use of prosthetic devices. Whether donning/doffing or readjusting the prosthetic socket, the electrodes can shift from their initial position. This can occur as a gradual change in a span of a day or different positions across two days, where shifts in sensor position relative to the underlying muscle fibers, can significantly alter the signal characteristics (Kyranou and Erden, 2018; Pereira et al., 2024). Based on the studies, the common displacement scenario in everyday use of a prosthetic hand is approximately 1 cm (Boschmann and Platzner, 2012, 2014; Hargrove et al., 2008; Muceli et al., 2013; Pan et al., 2015; Stango et al., 2014).

The *mechanical load* enforced by a prosthetic weight represents a separate challenge, creating additional muscle strain and compensatory muscle co-activation patterns caused by gravitational forces imposed on the limb (Cipriani et al., 2011). Furthermore, *skin conditions* such as moisture levels, tissue thickness, impedance fluctuations, and inconsistent electrode contact pressure can dramatically affect signal conductivity and amplitude (Kato et al., 2006).

Unlike intrinsic factors, which are physiologically inevitable, many extrinsic factors could theoretically be controlled with improved engineering solutions.

2.4.3 Compounded Factors

Commonly studied settings emerging from a combination of confounding factors focus on performance degradation over time or contextual factors. These factors manifest

across several dimensions: electrode variations, long-term adaptation, inter-subject and inter-session differences.

Creating off-the-shelf, generic and functional models for every user could potentially eliminate costly and complex, high-quality signal acquisition processes. However, *inter-subject* variability in EMG signals presents significant challenges, arising from anthropometric differences such as subcutaneous fat distribution, muscle fiber volume, and subjective perception of generating force and motion execution (De Luca, 1997; Zhao et al., 2021). While calibration techniques, transfer learning, and meta-learning frameworks (Campbell et al., 2021; Côté-Allard et al., 2019a; Prorokovic et al., 2020) have been explored, this remains a challenging problem primarily due to the data scarcity and the inherent heterogeneity of EMG signals.

Inter-session variability further exacerbates these challenges due to subtle differences in gesture execution, electrode positioning variations, and accumulated muscle fatigue, leading to distribution shifts in the data (Jaber et al., 2021). Common *electrode variations* in EMG-based systems happen due to placement shifts and channel variations (Wu et al., 2023). The latter one occurs when channels are missing at inference time as opposed to the data during model training. While electrode positioning is critical for EMG applications, displacement is unavoidable during donning and doffing acquisition devices, whether armband-style or socket-based systems.

Crucially, the temporal deterioration of system performance necessitates robust *inter-day adaptation* strategies to maintain performance and minimise abandonment rates (Amsüss et al., 2013b). This compounded factor, extensively analysed in Section 2.5, represents a fundamental challenge requiring systematic recalibration protocols to preserve functional performance across extended usage periods.

2.5 Long-Term Adaptation

The first concept investigated in this thesis is the long-term adaptation of EMG-based myoelectric systems. This section provides an overview of the topic, outlines the moti-

vation behind this research, and highlights shortcomings and areas for improvement in existing research.

2.5.1 What and Why?

The cornerstone of effective EMG-based prosthetic control is the ability to maintain a reliable and intuitive control of upper limb prostheses over time. However, various physiological and non-physiological factors, including electrode shift, impedance variations, muscle fatigue, and alterations in muscle activation patterns as users adapt to prosthetic devices, cause *concept drift*, characterised by temporal shifts in EMG signal distributions (Gama et al., 2014; Kyranou and Erden, 2018). Consequently, classification algorithms trained on initial EMG signals gradually experience performance degradation when confronted with changing data properties during inference time. This deterioration in prosthetic control accuracy substantially impacts the user, often leading to device abandonment (Amsüss et al., 2013b; Burrough and Brook, 1985; Glynn et al., 1986; Østlie et al., 2012).

Achieving long-term adaptation and mitigating performance degradation in myoelectric control systems requires a synergistic interplay of user and machine adaptation (He et al., 2015a). As signal properties change over time, the model must adapt to the incoming data (Côté-Allard et al., 2020b). Concurrently, users undergo their adaptation process, progressively refining their muscle co-activation patterns as they become accustomed to the prosthesis (He et al., 2015a). Thus, long-term adaptation in EMG-based myoelectric control can be comprehensively studied from two complementary perspectives: algorithm-based and human-centred (Hu et al., 2023).

2.5.2 Algorithm-Based Adaptation

A prevalent strategy for addressing performance degradation in real-time applications involves implementing *post-hoc error detection* mechanisms that evaluate prediction confidence levels. Specifically, confidence-based rejection operates by selectively outputting gesture predictions only when a pre-established confidence criterion is satisfied (Al-Timemy et al., 2018; Amsüss et al., 2013a; Scheme et al., 2013). While

computationally lightweight, this approach fails to facilitate classifier adaptation over time as the model remains fixed.

Supervised recalibration represents a potential solution where the myoelectric control system is periodically retrained using new labelled EMG data collected through a predefined set of tasks performed by the user. Various adaptation strategies have been explored to improve gesture decoder robustness (Amsüss et al., 2013a; Chen et al., 2013; Scheme and Englehart, 2011; Sensinger et al., 2009; Vidovic et al., 2015). Amsüss et al. (2014) developed a neural network post-processing approach for misclassification detection, achieving accuracy improvements ranging from 4.8% to 31.6%, however, without modifying the underlying classifier structure. Chen et al. (2013) proposed direct adaptation using prediction results from previous sessions, while Liu et al. (2014a) introduced reduced recalibration by utilising prior models from previous sessions (days) to recalibrate the decoder. In contrast, Vidovic et al. (2015) focused on optimising the decoder using collected daily new data subsets (single trials) to fine-tune the initially trained model. The experiment demonstrated classification accuracy above 92% over five days compared to 75% without adaptation in the offline data analysis and 25% performance improvements in online studies. However, as daily data collection and recalibration are inefficient and inconvenient for the user, periodic recalibration aims to reduce this burden (Côté-Allard et al., 2019b; Prahm et al., 2017).

Unsupervised calibration using pseudo-labels takes this a step further by eliminating annotation requirements. While majority voting is based on the assumption that adjacent samples share the same label, it introduces significant latency in real-time applications during gesture transitions. Huang et al. (2017) proposed the Particle Adaptive Classifier to combat performance degradation in unsupervised adaptive learning, while mitigating computational inefficiencies in SVM-based methods through an incremental least-square support vector classifier. Zhai et al. (2017) developed a self-recalibrating CNN classifier with label correction mechanisms, achieving $\sim$10.18% and $\sim$2.99% accuracy improvements for intact subjects and amputees, respectively. Côté-Allard et al. (2020b) introduced a self-calibrating domain adversarial network

(SCADANN) demonstrating 14-day adaptation capabilities, though supervised recalibration still outperformed unsupervised methods at extended time intervals in both offline and dynamic evaluation contexts.

2.5.3 Human-Centred Adaptation

Having explored algorithm-based adaptation strategies, user adaptation in EMG-based myoelectric control emphasises the importance of user experience and feedback in refining control performance over time (He et al., 2015a). Open-loop calibration (OLC) in myoelectric control follows a straightforward training-testing protocol where users generate EMG signals according to predefined gesture instructions during a training phase, without receiving feedback to alter their muscle contractions to adjust to the myoelectric feature drift (He et al., 2015b; Hu et al., 2023). Bridging this gap with closed-loop adaptive calibration (CLAC) allows the user to adapt thanks to instantaneous feedback.

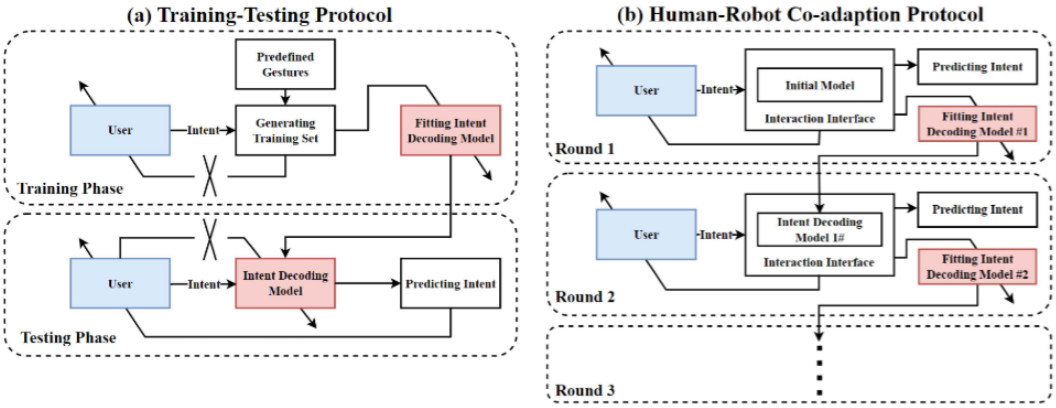


Figure 2.5 Figure presents two training protocols: (a) the training-testing, and (b) the human-robot co-adaptation protocol Hu et al. (2023). The figure has been adapted from the cited sources and is distributed under a CC BY 4.0 International License (https://creativecommons.org/licenses/by/4.0/).

Various feedback mechanisms and training strategies enhance user proficiency in controlling myoelectric prostheses. Somatosensory feedback, frequently requested by users Peerdeman et al. (2011), employs vibrations to close the control loop and enable user adaptation (Tchimino et al., 2021). Visual feedback approaches include EMG space similarity feedback for improving muscle synergy (Barradas et al., 2023)

and clustering feedback strategies that visualise online sEMG signals against training sample centroids (Fang et al., 2016). Advanced training environments employ gamification and virtual reality to significantly accelerate skill acquisition by increasing user engagement while progressively improving movement completion and classifier classification Powell et al. (2013).

This *human-centred* approach improves performance beyond what machine adaptation alone can achieve, as demonstrated by simultaneous 2-D proportional control in comparison to the open-loop paradigm (Hahne et al., 2015a). Together, these approaches form a 'bidirectional spiral of machine intelligence and human adaptability' (Hu et al., 2023) that enhances prosthesis control while accommodating inevitable physiological changes over time. *Co-adaptation* enables the human to actively interact with the prosthesis during data collection, incorporating their adaptive behaviours into the training process while the classifier simultaneously adapts through online learning (Vujaklija et al., 2017).

2.6 Limb Position Effect

The second aspect investigated in this work is the limb position effect, a critical factor impacting the reliability of EMG-based myoelectric systems. This section provides a comprehensive overview and motivation behind investigating this phenomenon, discussing typical scenario setups used in research, reviewing current approaches, and exploring potential solutions involving experimental protocols and sensor fusion techniques.

Variations in limb positions affect EMG signals by altering muscle activation patterns due to gravitational forces and electrode-muscle fiber interactions. These directly impact the performance of myoelectric control systems, affecting their robustness and reliability for real-world use (Chen et al., 2011; Fougner et al., 2011; Khushaba et al., 2016; Liu et al., 2012; Scheme et al., 2010).

2.6.1 What and Why?

Myoelectric control systems dexterity depends on precise muscle activity measurement, which is predominantly reliable only under controlled experimental settings. Standardised data acquisition protocols are employed to ensure data reliability and robust signal decoding. These typically restrict subjects to a specific neutral limb position, characterised by the forearm positioned adjacent to the body at a 90-degree angle to the upper arm and aligned parallel to the ground.

During muscle contraction, the position of muscle fibers can shift in relation to the mounted sensors and distort the tissue filter effect, impacting the recorded signal (Campbell et al., 2020b). This displacement occurs as the muscle generates the necessary tension for movement or limb stabilisation. As a result, the spatial relationship between the muscle fibers and the sensors may not remain constant throughout the contraction process. This issue amplifies when position shifts change muscle co-activation, subcutaneous tissue and muscle topology (Jiang et al., 2013a; Scheme et al., 2010). In particular, the brachioradialis muscle supports both forearm and elbow flexion, which heavily influences muscle co-activation as the gravitational forces are imposed on a limb. These position-dependent variations create artifacts that distort the interpretation of muscle activity. Hence, inter-position classification degrades in both precision and accuracy (Fougner et al., 2011; Scheme et al., 2010) compared to intra-position setting, where the model is trained and tested within the same position.

Research has shown that variations in limb positioning degrade performance for able-bodied and amputees. However, this effect appears to be less significant in amputee populations, which may be attributed to their anatomical differences, specifically their shorter muscle lengths and the occasional direct attachment of muscles to bone. In contrast to people with limb loss, the gravitational load during arm position shifts is directly applied to the tendons of the muscles for able-bodied (Jiang et al., 2013a). While the challenge of position effect impacts able-bodied more prominently, when the added weight of the prosthetic device compresses the underlying muscles of an amputee, it changes their contraction patterns as well. These can cause relief com-

pression in certain positions, redistribute the pressure points in others, or increase the overall compression on the muscular tissues. Addressing the challenge posed by limb positioning is crucial for establishing consistently reliable myoelectric control systems that work effectively for both amputee and able-bodied users.

2.6.2 Experiment Setup

Position effect can be categorised into several taxonomic groups based on underlying mechanisms, objectives and observable outcomes. These include scenario setup, experimental protocols, data modelling, and sensor fusion.

Scenario Setup

Scenario setup can be categorised into the following groups: *x* vs. *x*: (intra-position comparison), where both training and testing are conducted within the same position; *x* vs. *y* (inter-position comparison), where training and testing take place in different positions; *x* vs. *all* (single-position generalisation), which involves training in a single position and testing across all positions; *N* vs. *all* (multi-position generalisation), where training is performed in multiple positions, and testing is done across all positions; and *N-1* vs. *x* (leave-one-position-out generalisation) where the model is trained on all positions except one that is used for testing (Campbell et al., 2020b). Depending on the chosen frameworks, leave-one-trial-out cross-validation might be used when *x* is included in both training and testing, such as *x* vs. *x*, *x* vs. *all* or *N* vs. *all* to ensure consistent evaluation.

Experimental Protocols

Research exploring limb position effect has typically utilised four primary experimental protocols. First, *the static forearm orientation* protocol focuses on maintaining a consistent shoulder and elbow joint angles during data collection while introducing positional variability through wrist joint supination and pronation (Adewuyi et al., 2017; Ishii et al., 2017; Khushaba et al., 2015, 2016; Yang et al., 2017, 2015; You et al., 2010). This forearm rotation is often referred to in the literature as a wrist rotation.

The positioning of the forearm can significantly impact the detected muscle activation patterns, as pronator and supinator muscles are located in areas commonly used for muscle activity measurement (Yang et al., 2015). Notably, hand and finger movements engage both the intrinsic muscles within the hand and the extrinsic muscles positioned in the forearm (placed closely to the pronator and supinator muscles), introducing variability in muscular engagement and movement patterns. The impact of forearm orientation varies among transradial amputees, depending on the remaining limb length. Individuals with longer residual limbs can typically maintain forearm rotation, which introduces potential complications in muscle activity interpretation.

Second, the *static limb position* protocol involves holding the limb in predetermined, fixed configurations that prompt a subject to maintain specific angles between wrist, elbow and shoulder joints. These positions are supposed to imitate activities of daily living, making the problem significant for people with limb loss. The main objective of static limb position experiments is to examine changes in muscle co-activation patterns and their reliability for a myoelectric control system. Based on a comprehensive review of existing research, static limb positioning emerges as the predominant experimental approach in scientific literature (Al-Angari et al., 2016; Atzori et al., 2012; Betthauser et al., 2016; Boschmann and Platzner, 2014; Chen et al., 2011; Fougner et al., 2011; Geng et al., 2012a, 2017, 2012b,c; Ishii et al., 2017; Jiang et al., 2013a; Kanitz et al., 2018; Khushaba, 2014; Khushaba et al., 2012; Liu et al., 2012, 2014b; Masters et al., 2014; Radmand et al., 2014a,c; Scheme et al., 2010; Yang et al., 2017, 2015; You et al., 2010; Yu et al., 2017).

Third, *the dynamic two-dimensional (2D) space* protocol, which unlike experimental methods that maintain fixed forearm orientation or stationary limb positions, introduces movement by incorporating systematic changes in elbow and wrist joint angles throughout the data collection period (Beaulieu et al., 2017; Gu et al., 2018; Krasoulis et al., 2017; Liu et al., 2012, 2014b; Radmand et al., 2014a,c; Urra et al., 2014; Yang et al., 2017, 2015). These protocols capture the complexity and the reliability challenges inherent in myoelectric control systems during real-world applications. By incorporating transition periods in limb positioning, they capture both the body's

natural stabilisation mechanisms and muscle activation patterns associated with wrist and elbow joint movements. The comprehensive approach of tracking these dynamic elements enables more precise differentiation between various motion categories and the underlying postural adjustments (Liu et al., 2012).

Fourth, *the dynamic three-dimensional (3D) space* protocol examines transitions across shoulder, elbow, and wrist joints during specific motion tasks, falling primarily into reach-to-grasp sequences and activities of daily living (ADLs) categories (Batzianoulis et al., 2018; González et al., 2010; Horiuchi et al., 2009; Liarokapis et al., 2012; Radmand et al., 2014a; Rivela et al., 2015). While the former approach involves a structured progression through phases (rest, arm-extension to object, contact with object, grasp of object, unhand object, return to initial position and rest), the latter simulates everyday functional movements like drinking or reaching into cabinets (Radmand et al., 2014a). The objective of this protocol is to investigate muscle co-activation patterns across the entire limb during realistic tasks, with research showing consistent patterns under identical conditions but significant variations when object position or grip type changes (Liarokapis et al., 2012).

2.6.3 Sensor Fusion

Sensor fusion approaches have emerged as a promising alternative to address the challenges posed by limb position effects in myoelectric control systems. Fougner et al. (2011) established two approaches for sensor fusion: a cascaded classifier approach that first identifies limb position using accelerometer data before applying position-specific grasp classifiers, and a single-stage classifier that integrates features from both EMG and accelerometer signals simultaneously. While the cascaded approach requires substantial training data across discrete positions, the single-stage approach demands well-thought-through feature extraction techniques to effectively differentiate between position and motion information. Additionally, with more features, the amount of training data increases to avoid the curse of dimensionality (Radmand et al., 2014c). Notably, Shahzad et al. (2019) demonstrated that single-stage classi-

fiers trained with continuous position data outperformed discrete position classifiers, achieving accuracies of 98.7% and 97.6%, respectively.

Mechanomyography (MMG) signals have been investigated as a complementary data source to EMG, with Geng et al. (2012c) demonstrating improved position specificity when MMG data augmented traditional EMG-based classifiers. Interestingly, Khushaba et al. (2018) investigated graph Laplacian-based feature extraction techniques for both accelerometer and MMG modalities, achieving remarkable classification accuracies of 93.8% and 94.1%, respectively, substantially outperforming traditional EMG-based Hudgins' feature set, which only reached 66% accuracy (across 40 motion classes).

Inertial measurement units (accelerometers, gyroscopes, and magnetometers) provide motion data that navigation systems convert to orientation, velocity, and position through integration and filters, like Kalman and particle methods (Mourikis and Roumeliotis, 2007). However, these sensors, especially accelerometers and gyroscopes, suffer from inherent bias errors causing drift errors that accumulate over time, potentially limiting their long-term reliability without appropriate drift correction mechanisms (Tommasi et al., 2012).

2.7 Responsible Research and Innovation

AI integration in prosthetic control systems introduces several ethical concerns related to data privacy, security, and accountability (Keskinbora, 2019). These systems collect sensitive biometric data, including surface EMG signals and movement patterns, to customise device control (Gordon, 2021; Nayak and Das, 2020). While this data collection enables personalised functionality, it simultaneously creates privacy vulnerabilities that must be addressed through robust security protocols and transparent data governance frameworks (Weiner et al., 2022).

The autonomous decision-making capabilities of AI-driven prosthetics raise questions about responsibility when system failures occur. Unexpected hand movements or malfunctions could potentially harm users, creating ambiguity regarding accountability

among manufacturers, clinicians, and software developers (Stahl, 2021). Establishing clear liability frameworks is essential for maintaining user trust while encouraging continued innovation (Dignum, 2018).

Current laboratory-based assessment methods for myoelectric control systems often fail to translate effectively to real-world clinical outcomes (Vujaklija et al., 2017). This disconnection between research metrics and functional utility creates an ethical responsibility to develop more representative evaluation protocols that prioritise user experience in daily living conditions rather than controlled environments. Offline validation protocols typically overlook crucial performance indicators such as real-time responsiveness (Englehart and Hudgins, 2003; Farrell and Weir, 2007; Graupe et al., 1983), individual physiological variations (Zheng et al., 2022), and environmental adaptability (Tam et al., 2024). While conventional machine learning metrics like accuracy and precision are widely used in hand gesture recognition, these assessments fail to capture performance in online scenarios as they neglect users' adaptive responses to non-stationary EMG signals (Gusman et al., 2017; Jiang et al., 2013b; Lock et al., 2005; Ortiz-Catalan et al., 2015; Vujaklija et al., 2017).

To address these limitations, research should incorporate standardised online validation methods, including the box and blocks test (Farrell and Weir, 2007; Mathiowetz et al., 1985), target achievement control test (Simon et al., 2011), and Fitts' law test (Fitts, 1954), which provide more comprehensive insights into practical usability. Additionally, human-in-the-loop approaches are essential, integrating user feedback throughout development and deployment phases to continuously refine control algorithms based on real-world performance and subjective user experience. This collaborative approach acknowledges users as active participants rather than passive recipients of technology, enabling more responsive and adaptable prosthetic systems that better address individual needs.

2.7.1 Responsible Solutions and Future Directions

Addressing these challenges requires multilevel approaches:

1. Implementing privacy-by-design principles during development to minimise data vulnerability while maintaining functionality (Weiner et al., 2022).

2. Establishing collaborative governance frameworks involving users, clinicians, engineers, and ethicists to ensure diverse perspectives inform technology development (Dignum, 2018).

3. Developing transparent algorithms with explainable decision-making processes to build trust and enable informed user consent (Stahl, 2021).

4. Creating standardised clinical assessment protocols that prioritise functional outcomes over technical metrics, incorporating comprehensive online validation strategies that assess real-time performance, diverse subject populations, continuous adaptation mechanisms, and varied environmental conditions (Jiang et al., 2013b; Vujaklija et al., 2017).

The advancement of EMG-based myoelectric control systems for upper limb prosthetics requires balancing technological innovation with ethical responsibility. By implementing robust privacy protections, clear accountability structures, and meaningful assessment protocols, researchers can ensure these technologies serve user needs while maintaining ethical standards.

Chapter 3

Recalibration of Myoelectric Control with Active Learning

In the previous Chapter 2, we explored the fundamental limitations of upper limb prosthetics, the characteristics of sEMG signals, and various confounding factors causing significant performance disparities between laboratory settings and real-world applications. In particular, Section 2.5 examined the critical issue of performance deterioration over time, which contributes to high abandonment rates among prosthesis users with limb differences.

Improving the robustness of myoelectric control to work over many months without the need for recalibration could reduce prosthesis abandonment. However, current approaches rely on post-hoc error detection to verify the certainty of a decoder's prediction using a predefined threshold value. Since the decoder is fixed, the performance decline over time is inevitable. Other approaches, such as supervised recalibration and unsupervised self-recalibration, entail limitations in scaling up and computational resources. The objective of this chapter is to study active learning as a scalable, human-in-the-loop framework to improve the robustness of myoelectric control.

We use active learning with linear discriminant analysis to create an iterative learning process and modify decision boundaries based on changes in the data. We simulate a real-time scenario and exploit least confidence, smallest margin, and entropy reduction sampling strategies in single and batch-mode sample selection. We consider

optimal batch-mode sampling using ranked batch-mode active learning. With only 3.2 minutes of data carefully selected by the active learner, the decoder outperforms random sampling by 4-5% for able-bodied individuals and approximately 2% for people with limb differences. We observe active learning strategies systematically and significantly enhance the decoder's adaptation while optimising the amount of training data on a class-specific basis. Smallest margin and least confidence uncertainty prove to be the most effective. We introduce, for the first time, an active learning framework for long-term adaptation in myoelectric control. This study simulates a closed-loop environment offline and proposes a pipeline for future real-time deployment.

This chapter begins by reviewing recalibration approaches in Section 3.1, followed by Section 3.2, which outlines the methodology, including data acquisition protocols, an active learning framework, and experimental design. Section 3.3 presents a comparative analysis of various query strategies, while Section 3.4 and 3.5 propose a framework for real-time implementation, address current limitations, and offer concluding insights into the potential of active learning for myoelectric control adaptation.

3.1 Introduction

Benefiting from an exponential increase in computational power and the availability of data, machine learning stands as one of the main pillars of the digital revolution. Not surprisingly, the literature on myoelectric control has also seen the reemergence of academic interest in the use of machine learning for the classification of the myoelectric signals (Buongiorno et al., 2019). Likewise, the myoelectric industry has recently developed or adopted machine learning-based solutions, although large-scale deployment of such methods is challenging because the performance of current machine learning algorithms degrades over time (Amsüss et al., 2013b; He et al., 2013; Kaufmann et al., 2010). It is caused by a range of intrinsic factors, e.g. changes in muscle physiology (atrophy or hypertrophy) (Kyranou and Erden, 2018) and motor behaviour (Hahne et al., 2020) as well as extrinsic perturbations such as electrode displacement and arm position (Radmand et al., 2014c; Scheme and Englehart, 2011).

A conventional way to enhance the robustness of machine learning-based decoding is to use a post-hoc error detection system that determines whether a decoder's prediction is certain enough. Based on predefined criteria of certainty, the model outputs a movement class label, and the prosthesis executes the movement if the predicted probability attains a preset level of confidence (Al-Timemy et al., 2018; Amsüss et al., 2014; Krasoulis et al., 2020; Scheme et al., 2013). In this approach, the classifier is fixed; that is, it is not updated upon the detection and rejection of errors. Hence, these approaches do not address, but partially circumvent, the issue of performance decline over time. To address this concern, ideally, both human motor function and the machine learning system should modify their behaviour over time, a process known as co-adaptation. However, in all earlier literature that advocates for motor learning/adaptation for prosthesis control (Antuvan, 2019; Dyson et al., 2018, 2020; Radhakrishnan et al., 2008; Segil et al., 2020; Segil and Weir, 2015), a fixed decoder was used and learning/adaptation was quantified by probing the human's motor behaviour. On the other hand, despite the use of terms such as co-adaptive, machine adaptation has been shown to work for short-term data and in a fixed environment, without any direct involvement from the user beyond the operation of the interface (Hahne et al., 2015b; Igual et al., 2019; Vidovic et al., 2016).

To the best of our knowledge, there is no platform that enables and explicitly verifies a human user and the machine co-adapt. Nevertheless, research on the development of closed-loop frameworks that allow machine adaptation is ongoing. Broadly, these can be grouped into supervised recalibration and unsupervised self-recalibration.

In *supervised re-calibration*, a set of predefined tasks is performed at every iteration of decoder update (He et al., 2015a). This naive approach puts the onerous of regular system updates on a user. An alternative approach uses a new set of labelled samples (Liu et al., 2014a; Vidovic et al., 2015). Specifically, (Liu et al., 2014a) modified the classifier using all prior models from previous days to recalibrate the decoder. In contrast, Vidovic *et al.* (Vidovic et al., 2015) focused on optimising the decoder using small runs of daily data. Examples of periodic recalibration using transfer learning were also used in an attempt to decrease the amount of needed data (Côté-Allard et al.,

2019b; Prahm et al., 2017). Moreover, (Gu et al., 2018) used an incremental learning scheme to redefine a representative sample set for model training. These scenarios however, can account for small shifts in the distribution of the data only, and hence are more suited for short-term laboratory research.

System recalibration shows performance improvement, but retraining is not a sustainable solution for long-term prosthesis usage. It disrupts the prosthesis use, requires memory, and is costly computationally. To address these shortcomings, the method of pseudo-labelling for *unsupervised self-calibrating* has been proposed. By using models' predictions to annotate/label the data, the system is able to adapt using only estimation of the user intents (Chen et al., 2013; Côté-Allard et al., 2020b; Sensinger et al., 2009; Zhai et al., 2017). Given its unsupervised nature, it removes the burden of successive retraining processes and hence becomes easier for the user. Initial attempts towards system recalibration involved post-hoc comparison analysis, by enlarging the training pool with respect to controller confidence using unsupervised entropy-based confidence threshold (Sensinger et al., 2009). Using the predictions of testing data from previous sessions, Chen et al. (2013) updated the model to maintain the performance. Such a self-enhancing classifier is a self-recalibrating system because it continuously updates its parameters without enlarging the dataset; that is, the testing data is discarded after each model update. Preliminary results were promising, but they were based on single-day data acquisition (with 2$\sim$3 and 6$\sim$7 hours time span between training and testing data collection). Recently, Zhai et al. (2017) suggested refining the decoder by using the predictions from the previous training sessions and those based on adjacent window segments. This concept was based on the assumption that neighbouring segments share the same class movement. To enhance relabelling using context-based predictions, domain adaptation was applied in (Côté-Allard et al., 2020b). They report performance improvement over the no calibration setting. Yet, when compared with recalibration, the performance degradation was significant.

The aim of this paper is to demonstrate, in an offline setting, this bottleneck can be addressed fundamentally using active learning (Settles, 2009). We offer a structured perspective and methodology, via which a user can interact with the machine. The

intuition behind this approach is based on selecting new (unlabelled) data samples that are representative of the underlying real-world data distribution or that provide new information for the model during the training process. Specifically, the objective of active learning algorithms is to define samples for optimal model training by minimising labelling costs while maximising a model's performance.

We will argue that this approach can provide a sustainable solution for the shortcomings of all previously stated approaches. Active learning is capable of detecting new regions of interest and modifying decision boundaries based on changes in data (possible distribution shifts and others). Rejection-based calibration dismisses those samples. This is a feasible solution in a short-term scenario, however over time, it will lead to discrepancies and system limitations. This problem can be solved by simply retraining the decoder using more data (supervised recalibration). Yet, active learning increases robustness of the model by choosing informative samples and simultaneously using fewer samples for training (a.k.a. decreasing computational costs). Unsupervised calibration tries to tackle the problem from the annotation point of view. However, as it often relies on initial calibration being well optimised, if impaired, it can lead to a drastic decrease in the performance of the model over time. Active learning can optimise this performance with as few samples as possible, leaving this approach supreme in all aspects.

3.2 Method

3.2.1 Ethics

All experimental procedures were in accordance with the Declaration of Helsinki and approved by the local Ethics Committees of the School of Informatics, University of Edinburgh (#201507160854) and School of Engineering, Newcastle University (#14-NAZ- 056). All participants read an information sheet and gave written informed consent prior to the experimental sessions.

3.2.2 Data Collection

Data was collected as part of Krasoulis et al. (2020). For completeness, we briefly described the experimental process. The myoelectric signals were collected from the forearm of 12 able-bodied and the stump of two subjects with trans-radial limb loss. See Table 3.1 for a description of participants with limb loss. Prior to the placement of myoelectric sensors, we cleansed participants' skin using 70% isopropyl alcohol. In the first group, we placed 16 Tringo^{TM} sensors (Delsys, USA) on the participants' in two rows of eight equidistant electrodes. For subjects with limb loss, we used 13 and 12 sensors based on their physiology of the stump after amputation. We used an adhesive elastic bandage to secure the locations of the electrodes throughout the sessions.

Sampling rates in data recording were originally 1111 Hz and 128 Hz for the myoelectric channels and the inertial measurement data, respectively. This sampling rate was the standard and fixed hardware sampling rate of Delsys Trigno system when inertial data was recorded simultaneously. We did not use the recorded inertial data in this study. In the process of data cleaning, the Hampel filter was applied to up-sampled (2kHz) data to eliminate the power line interference. This pre-processing followed the process introduced in Atzori et al. (2014b). The data was then band-pass filtered in the range 30-400 Hz using a 4th-order Butterworth filter. In this paper, we used only the myoelectric data. In a sliding window of 128 ms (256 raw signal samples) with an increment of 50 ms (100 raw signal samples), the waveform length for each channel was calculated, resulting in 16 values per window (one per sensor), and 13 and 12 participants, respectively, with limb loss. In the discussion, we described why we chose only one feature.

Data was collected in two identical sessions, namely $T1$ and $T2$, which were conducted sequentially on the same day. During the recording session, each subject was instructed to perform the grips that were shown on a computer screen. The experiment involved five grips, namely the power grip, lateral grasp, tripod grasp, index pointer, and hand open. During two blocks of data recording, each grip was repeated 10 times, with 5s of muscle activation followed by 3s of rest.

Gender	Age	Type of amputation	Cause of amputation	Years	Missing limb	Hand dominance (prior to amputation)	Prosthesis use
Male	28	Transradial	Car accident	6	Right	Right	Split hook
Male	54	Transradial	Cancer (epitheliod sarcoma)	18	Right	Right	Split hook

Table 3.1 Medical records of amputee subjects

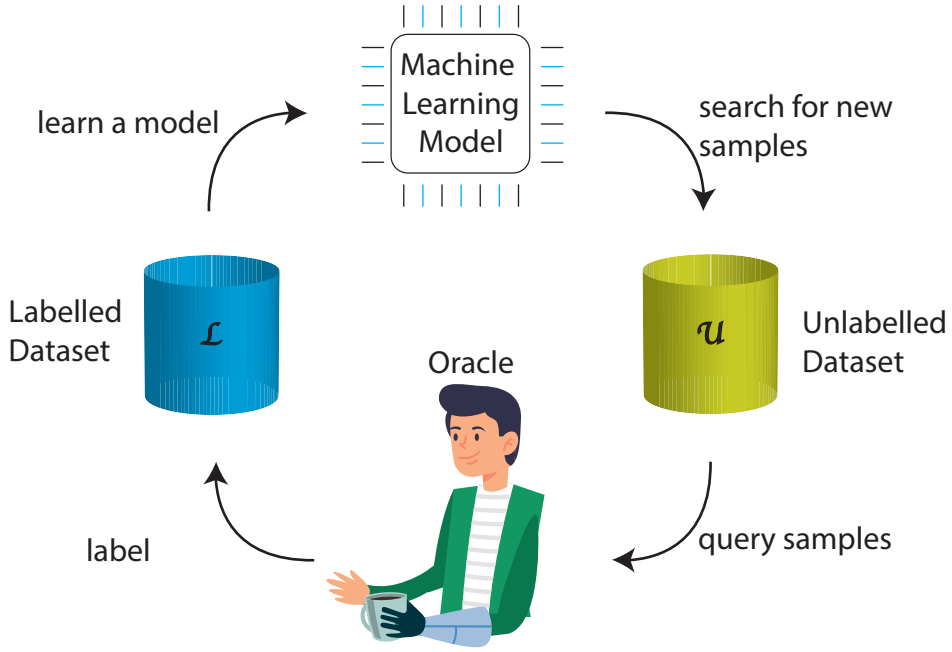


Figure 3.1 The generic framework of Active learning. The feedback loop includes a small set of labelled data L that is used to train a machine learning model. Based on that, the active learner explores a large pool of unlabelled data U and selects new samples. These samples are then passed to an oracle, which would be a prosthesis user in this paper, for labelling. These samples join the training set for model retraining.

The first dataset, $T1$, was used for training, and the second dataset, $T2$, was used for testing the decoder.

3.2.3 Active Learning

A general framework for an active learner includes two sets of data, namely an initial small set of labelled data L and a large pool of unlabelled samples U. Based on initial subset of data, $L^* \subset L$, a model $f(x|L^*)$ is trained. The aim of an active learner is to choose an L^* such that $f(x|L^*) \approx f(x|L)$. Note that these two sets are not the same sets as $T1$ and $T2$.

To obtain L^*, the active learner adopts a query strategy and a current model $f(x|L')$, where L' is a transitional annotated set, and selects new samples for annotation by an oracle, e.g. the prosthesis user. The oracle has the task of labelling the unlabelled data. Figure 3.1 presents a generic active learning framework. Starting from the initial dataset, the model is trained. Adopting a query strategy, the active learner selects new instances x_U from the pool of unlabelled data for labelling by the oracle and moving them from U to L. The model is then updated.

There are three main strategies of sampling from the pool of unlabelled data, namely, stream-based selective sampling (Les E. Atlas, 1990), membership query synthesis (Angluin, 1988), and pool-based active learning (Lewis and Gale, 1994). We adopted the pool-based sampling method in this proof-of-concept, offline study. We also shared our views as to how active learning can be implemented in a real-time myoelectric control setting.

3.2.4 Query Strategy

Active learning is based on defining optimal criteria for deciding whether the oracle needs to label an unlabelled sample, that is, a query. Using measures that determine the *informativeness* and/or *representativeness* of a sample, there are many algorithms for querying the label. Informativeness is typically measured with uncertainty sampling. In this method, the active learner finds samples in the pool about which the decoder is least certain. We used the following metrics:

1. *Least Confidence Uncertainty* (Lewis and Gale, 1994) considers querying a sample that it is the least certain about the predicted class:

$$x^*_{LC} = \arg\max_x \; 1 - P_\theta(\hat{y}|x) \tag{3.1}$$

where $\hat{y} = \arg\max_y P_\theta(y|x)$. This approach is intuitive; however, it focuses on the most probable label without considering the remaining label distribution.

2. *Smallest Margin Uncertainty* (Scheffer et al., 2001) considers the uncertainty between two most likely labels:

$$x_M^* = \arg\min_x \left[P_\theta(\hat{y_1}|x) - P_\theta(\hat{y_2}|x)\right] \tag{3.2}$$

where $\hat{y_1}$ and $\hat{y_2}$ are the two labels with the highest probabilities predicted by the model. Intuitively, large margins are easy to identify as they are expected to be far from the decision boundary. Instances with small margins dictate potential ambiguity in discriminating between the two classes. Acquiring such instances improves classification within the decision boundary area.

3. *Entropy Reduction* (Shannon, 1948) sampling strategy uses entropy to measure the amount of information necessary to depict a distribution. As such often described in machine learning as a measure of impurity.

$$x_H^* = \arg\max_x -\sum_i P_\theta(y_i|x)\log P_\theta(y_i|x) \tag{3.3}$$

where y_i indicates all plausible annotations.

Alternative to the single sample selection, a batch-mode active learning can be applied, which allows querying larger sets of samples. We used ranked batch-mode active learning (RBMAL) (Cardoso et al., 2017), which generates an optimised ranked list (Q) of unlabelled samples. Specifically, the uncertainty estimation step calculates two sets. First, uncertainty score U_{score} for all unlabelled samples in U, and $D_{estimate}$ set with L samples and already included (if any) samples within the ranking Q. The set $D_{estimate}$ presents the expected training set given all samples within the ranking Q will be annotated. Then, for instance, ranking score value for each sample within U (Equation 3.4) is calculated using similarity score Φ and U_{score}. To define Φ, the highest similarity between samples from the expected training set $D_{estimate}$ and each U instance was calculated. Combining this information enables the calculation of *score* to determine an instance with the highest value to be added to Q (and removed from

U). This process for ranking construction is repeated until $U = \varnothing$ or $|Q| = batchSize$.

$$score(x) = \alpha \times \left[1.0 - \Phi(x, D_{estimate})\right] + (1.0 - \alpha) \times U_{score}(x), \tag{3.4}$$

where $\alpha = \frac{|U|}{|U|+|L'|}$ balances the trade-off between exploration and exploitation.

3.2.5 Experiment Design

We ensured that all classes, including the rest and grasp classes, had the same number of samples, i.e. balanced classification. This was achieved by random undersampling of the rest class, which appeared between each grip repetition. The initial training set $L \subset U$ included 300 samples in total, that is, 50 samples per class. We sought to simulate a real-life scenario and hence picked the first 50 samples of data in each class in $T1$. This is equivalent to approximately 2.6s of myoelectric data per class, in the time domain. We treated the rest of the data in U as unlabelled samples. All training data was from the first recording session, $T1$.

For single instance sample selection, we compared the four approaches, namely random sampling (i.e. passive learning), least confidence uncertainty, smallest margin uncertainty, and entropy reduction. We used the results of the random sampling approach as a benchmark. We made 1500 queries iteratively, and at each iteration the training set was normalised to its z-scores. Then a linear discriminant analysis model was fitted on the normalised training set, L'. All 1500 resultant models were tested with the data from the second session, $T2$.

The second analysis compared the performance of single instance sampling with that of two batch-mode sampling methods, namely naive and ranked. The former approach focuses on finding n best samples to query within a single iteration. The latter learns to obtain an optimal set of samples to query from. The ranked method removes potential sample redundancy within a queried batch (if queried samples are too similar) by emphasising sample diversification using the euclidean distance measure.

To avoid any bias in training the decoder using the random batch sampling method, we imposed the constraint of sampling one sample per class within the same batch,

keeping the batch size to 6. This ensured an independent and identically distributed random sampling condition, imposed on the random sampling scenario, throughout all the iterations. All previously mentioned query strategies were used in this analysis, too.

3.2.6 Analysis

Metrics used for the evaluation of the proposed approaches included accuracy and F1-score (i.e., harmonic mean of precision and recall). Accuracy is the percentage of samples that have their labels correctly recognised. F1-score was calculated per class using a one-vs-all scheme and then averaged across the classes, that is, macro-averaging. These measures, together, reflect the overall success of the proposed approach.

3.3 Results

3.3.1 Single Query Sampling

Four query strategies were compared across 1500 single queries (equivalent to 3.2 min of data). Figure 3.2 presents the accuracy scores of all query strategies plotted against a number of queried samples and averaged across all participants. The classifier was initially trained with 50 samples from each class (giving 15.9 sec of training data per subject). The mean accuracy and standard deviation for able-bodied participants and people with limb difference, before the first query was 63.66 $\pm$16.13% and 54.24 $\pm$ 2.65%, respectively.

Analysis of data from able-bodied participants showed that active learning, irrespective of the adopted sampling method, outperforms random sampling by 4-5 %. This improvement for people with limb differences was $\sim$ 2 %. Although entropy sampling outperformed the baseline, it required more samples to level off with the remaining active learners. These improvements might seem small, but we would like to remind the reader that they were achieved with only 1500 samples (only 3.2 min of

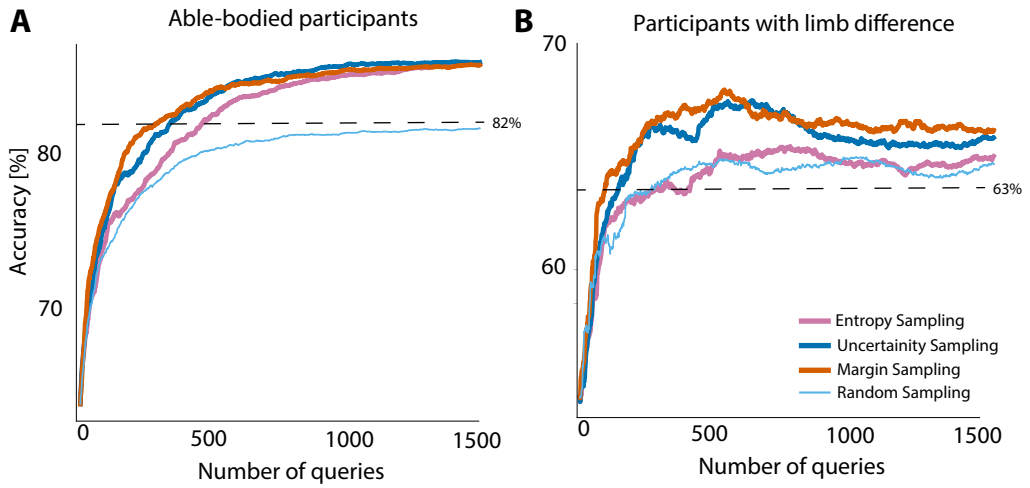


Figure 3.2 Averaged test accuracy across all able-bodied participants (A) and people with limb difference (B). Initial training sets for subjects in both groups contained 50 samples from each class. Initial accuracy without active learning is presented at query step zero. Offline baseline: 82% (A) and 63% (B) for able-bodied and people with limb loss, respectively.

	Able-bodied				Amputee 1				Amputee 2			
	Q200		Q1500		Q200		Q1500		Q200		Q1500	
	Acc (std)	F1	Acc (std)	F1	Acc	F1	Acc	F1	Acc	F1	Acc	F1
Entropy	77.2 (12.7)	0.76	85.9 (7.3)	**0.86**	65.1	0.65	67.8	0.66	61.1	**0.58**	62.3	0.59
Uncertainty	79.1 (11.8)	0.78	**86.1 (7.3)**	**0.86**	68.4	**0.67**	**68.1**	**0.66**	**62.1**	**0.58**	63.7	0.61
Margin	**80.9 (10.9)**	**0.8**	85.9 (7.6)	**0.86**	**69.6**	**0.68**	67.8	**0.66**	61.6	**0.58**	**64.7**	**0.62**
Random	76.7 (11.9)	0.76	81.7 (9.1)	0.82	66.8	0.66	67	0.65	60.1	**0.58**	62.4	0.6

Table 3.2 Method Comparison between query 200 and 1500. All accuracies are reported in %.

data) and very little additional computation. For completeness, we also considered the case of 200 queries only. Details on obtained results for both query numbers and all sampling strategies are presented in Table 3.2.

3.3.2 Comparison of queries

Figure 3.3 depicts the accuracy scores obtained by all the subjects after querying the first 200 and after all 1500 samples. For a comprehensive analysis, all combinations are reported. In the first row, three adjacent plots map out random sampling against active learning query strategies. In all cases, active learning outperformed random sampling in almost all subjects, that is, markers representing accuracy results from individual participants lie above the unity line in Figure 3.3A. When active learning

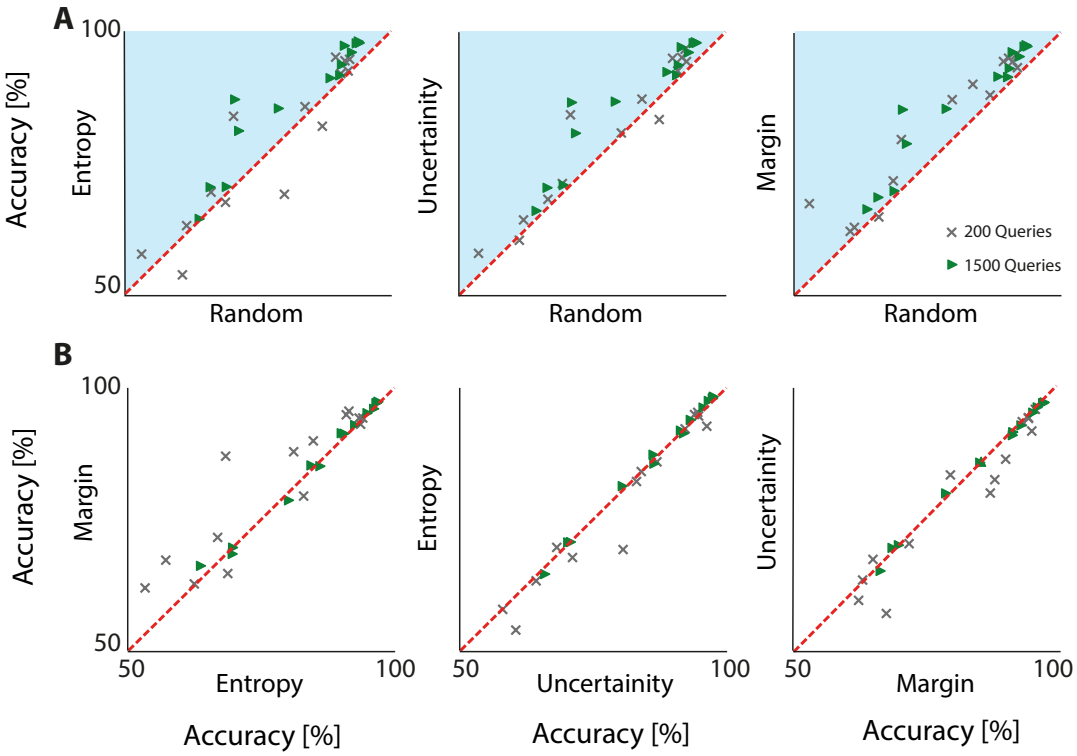


Figure 3.3 Comparative analysis of classification accuracy for different active learning query strategies. Each point on the scatter plots represents the accuracy for a single participant under two different query strategies (one for each axis). The analysis is shown at two key stages of the recalibration process: an early stage after 200 queries (x markers) and a final stage after all 1500 queries (▲ markers).
(A) Performance of active learning versus the random sampling baseline. The results for Entropy, Uncertainty, and Margin sampling are plotted against the random sampling results. For the final 1500-query stage, nearly all participants fall within the blue shaded area, indicating that all active learning strategies consistently and significantly outperformed the conventional random sampling approach.
(B) Inter-comparison of the active learning strategies. These plots directly compare the performance of Margin vs. Entropy, Entropy vs. Uncertainty, and Uncertainty vs. Margin sampling. While the results are closely clustered around the unity line (dashed red), a slight trend can be observed where Margin sampling offers a modest performance advantage for most participants.

approaches were compared against each other, results were relatively similar (*c.f.* Table 3.1). Interestingly, for most subjects, margin sampling proved slightly more successful.

Figure 3.4 represents the number of queried instances from each grasp per query strategy (A), averaged across all able-bodied participants. In random sampling, the data is picked with a discrete and uniform probability distribution across all the classes. When 1500 samples are being queried, $\sim$ 250 samples are acquired from each grip

class. These are then averaged across all participants. However, with active learning, we can observe a class-specific sampling. When inspecting grip trends for active learning query strategies, we observed the rest class being acquired the most by entropy sampling for both participant groups. Characteristic for all strategies, most samples were queried for the lateral grasp class. This is explainable from an anatomical point of view: the abductor pollicis brevis longus muscle, which is responsible for thumb movement and stabilisation, is a deep muscle in the forearm. Hence, the myoelectric data from this muscle is affected significantly by volume conduction and interference from other muscles. As such, active learning prioritised querying this class.

Figure 3.4 (B) shows the same analysis averaged across the two participants with limb difference. A notable trend emerges when comparing these results to the able-bodied cohort: all active learning strategies queried substantially fewer samples for the rest class. This suggests that, for these participants, the neural signature for the rest state was more distinct and required less data for the model to learn, allowing the active learner to focus its queries on more complex or ambiguous active grasps like tripod and lateral. However, it is critical to note that with a small sample size of two participants, the observed inter-subject variability was high. Therefore, while this trend is apparent and interesting, it cannot be considered a statistically significant finding and requires further investigation with a larger cohort of individuals with limb differences.

3.3.3 Batch Mode

For batch mode analysis, two scenarios were considered: naive batch sampling and ranked batch sampling. Both approaches followed a similar protocol for experimental design as in single instance sampling. An active learning framework was constructed with a linear discriminator analysis decoder, 50 initial samples (acquired in the same manner as in the previous experiment) and 250 queries (with 6 samples per batch, corresponding to 3ms), which is the equivalent of the number of samples obtained after 1500 queries for single instance sampling (250 queries $\times$ 6 samples).

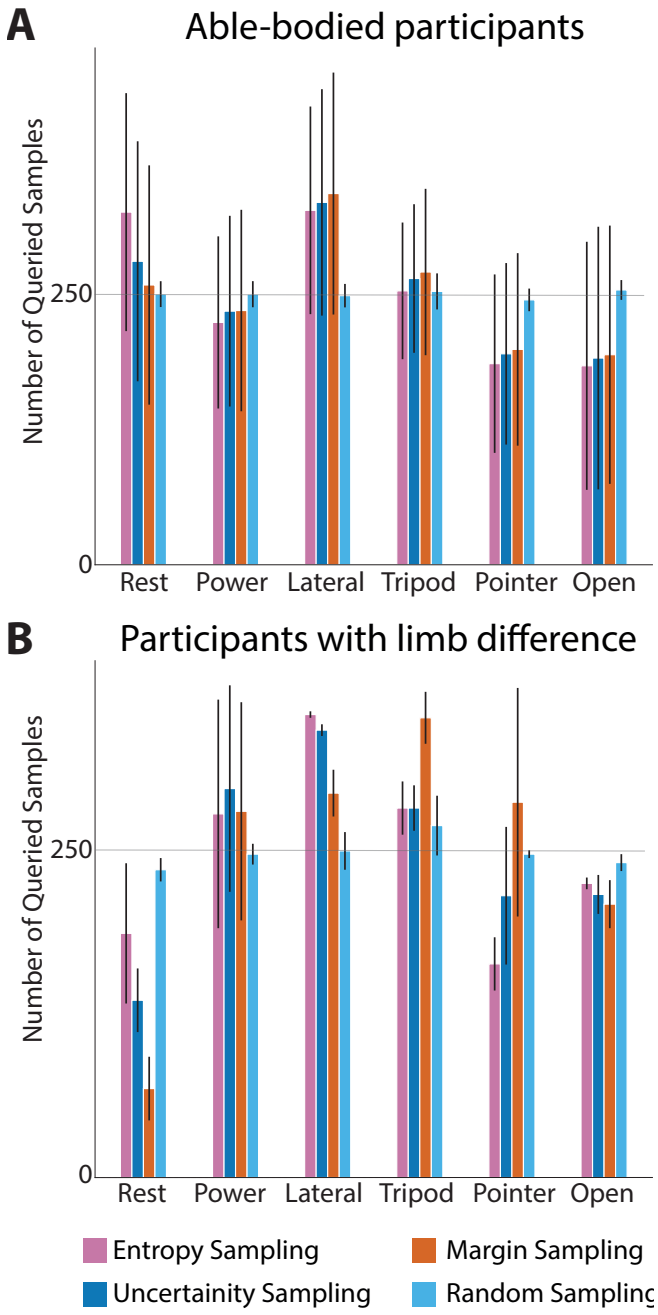


Figure 3.4 Grip instances acquisition during 1500 querying process across all the subjects: able-bodied participants (A) and people with limb difference. Query strategies are compared against class. In (A) the highest demand for new instances within 1500 queries was for the lateral grip. The pointer grip and hand open classes needed the fewest samples. In contrast to able-bodied participants, active learning requested fewest samples for the rest class. In all cases, random sampling drew 250 samples on average.

Observed results in Figure 3.5 indicate single instance sampling and batch mode sampling behave in a similar manner. Margin and uncertainty sampling perform

the best. However, ranked batch mode sampling presents divergent results. In the initial stages of querying, random batch sampling outperformed active learning query strategies. Entropy and uncertainty sampling slowly converged and outperformed random batch sampling (as the number of iterations increased). Important to note, random batch sampling was constrained to querying a sample from each class (within a single batch) at each iteration to ensure no bias across classes occurs in further iterations. Based on this constraint, the batch size for sampling new instances was set to six (number of classes) for all query approaches. In contrast with the previous approaches, margin ranked batch sampling had the lowest performance across all sampling methods.

3.4 Discussion

We introduced active learning for myoelectric control as a potential approach to ease re-calibration and enhance long-term decoding stability. We reported the results of an offline, feasibility study in a simulated human-in-the-loop setting. We will draw conclusions as to how the proposed approach can be implemented in real-time. We envisage that the machine learning decoder, supplied with only a small subset of labelled data, using an active learning boost, can efficiently exploit an abundance of unlabelled data in closed-loop, real-life settings.

We utilised a linear decoder, as an example. Nevertheless, we could have used active learning with any machine learning decoder. Results reported during the experiments prove the effectiveness of this approach with such a simple decoder.

While a direct comparison with existing supervised and unsupervised recalibration methods would be valuable for future work, this study was designed to establish the fundamental viability of active learning as a novel paradigm for myoelectric control. The comparison against random sampling was deliberately chosen to isolate and quantify the benefit of intelligent sample selection, the core principle of active learning, versus passive data collection. Existing recalibration methods represent two ends of a spectrum with well-documented trade-offs: supervised recalibration achieves high

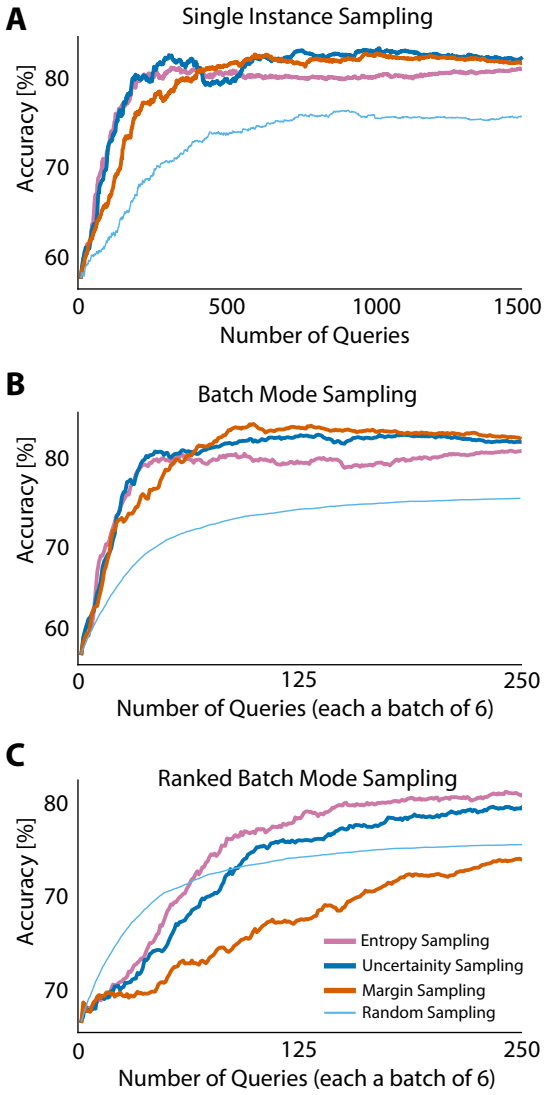


Figure 3.5 Comparison of batch mode sampling and single instance sampling in one representative participant.

accuracy but imposes significant user burden through repetitive data collection, while unsupervised self-calibration eliminates this burden but risks performance degradation through error accumulation when models reinforce their own mistakes. Our active learning framework occupies a strategic middle ground, incorporating human confirmation specifically for uncertain samples to ensure reliability while maintaining minimal user disruption (3.2 minutes of carefully selected data). Having demonstrated that this approach is significantly more data-efficient than random sampling, future studies can conduct direct comparative analyses to quantify the precise trade-offs in

accuracy, user burden, and long-term stability against state-of-the-art techniques across multi-day deployments.

The input feature was the waveform length of the myoelectric signals. We chose it because it has been shown as one of the most effective features for myoelectric signal classification. This approach also enabled us to avoid the curse of dimensionality, especially in the initial phase of learning when the decoder was trained on the 300 samples only.

Within a pool-based setting, three sampling methods were considered: single instance sample selection, batch-mode and ranked batch-mode. Although active learning can consider a cold start with no initial dataset, using a small number of labelled samples can initially boost active learning performance. We included the first 50 samples from each class, equivalent to 2.65 sec of data, mimicking real-life scenarios.

We demonstrated that active learning outperforms random sampling. The system demonstrated rapid adaptation in single query sampling from a baseline of only 300 samples (15.9 sec of training data), and within the first 200–500 queried samples (bringing the total accumulated data to 25.9 to 40.9 sec), demonstrated a high learning curve. This efficiency resulted in a $4-5\%$ and ~ 2 % performance improvement for able-bodied and people with limb difference, respectively, over random sampling after incorporating just 1500 strategically selected queries, equivalent to only 3.2 minutes (192 sec) of accumulated data. Within proposed query strategies, the smallest margin sampling and the least confidence uncertainty achieved higher accuracy in most participants. Intuitively, we expected large margins to be further away from the decision boundary, which consequently means instances with a small margin could potentially provide information to improve the decision boundary. Least confidence uncertainty focuses on sampling cases which the model is most unsure about. In the initial stages of training, entropy sampling presents a decrease in performance for able-bodied participants and is comparable to baseline results for amputees. This type of behaviour can be caused by emphasising feature space exploration over exploitation by query strategy.

As stated earlier, the objective of an active learner is to optimise the amount of data needed for training the decoder while maximising its performance. Using passive learning, i.e. random sampling, gave us a performance comparison against the baseline. Since our experiment is static, that is, the pool of data does not change with time, the active learners after querying all the samples will converge to the same performance as the random sampling solution. This would provide the results equivalent to offline training on the entire block $T1$ (initial and pool data set altogether). Figure 3.2 presents the results after training the decoder on the entire block $T1$. As noticed, the performance of the active learner at query no. 1500 is greater than when the entire data set is being used. We presume the potential noise of the signal and transient periods created noisy input data, leading to decreased performance.

When considering query sampling size, we applied batch-mode active learning (BMAL) and ranked batch-mode active learning (RBMAL). Performances of single instance sampling and BMAL were comparable. In RBMAL, the design requires a trade-off where exploration (the uncertainty measure) is balanced against exploitation (the distance measure). This design caused active learning to show a lower performance than the benchmark random batch sampling during the initial querying stages.

Among the RBMAL strategies, margin sampling performed the worst. Smallest margin sampling looks for samples that potentially lie on the decision boundaries, meaning the difference between the two most probable classes is the smallest. We presume that RBMAL's similarity score (distance measure) can compromise this strategy, as the samples with the smallest margin may be quite similar/close to each other, leading to the rejection of potentially informative, boundary-proximate samples in favour of less relevant, distant ones to avoid redundancy. This dilutes the quality of the batch, diminishing overall efficiency.

The performance disparity in batch mode is likely associated with the intrinsic redundancy profile of each uncertainty metric. Entropy sampling tends to select clustered points within highly confused regions of the feature space, which, if sampled naively (as in simple BMAL), would lead to high redundancy. Consequently, Entropy's strong performance with RBMAL suggests it is heavily dependent on the explicit diversity

constraint within the RBMAL scoring function. This constraint plausibly forces the batch to select samples spanning geometrically distinct areas of ambiguity, maximising the net information gain. Conversely, margin and least confidence sampling, by focusing on individual, low-ambiguity decision boundaries, appear to select samples that are already naturally dispersed. This suggests that the optimal batch selection method must be carefully matched to the sampling strategy's native redundancy.

3.4.1 From Offline to Real-Time Deployment

Creating a human-in-the-loop environment has the potential to motivate the users to engage more with the decoder and ultimately lead to enhanced control of the prosthesis. We presented a closed-loop setting in which an oracle and a prostheses user are the same. Figure 3.6 offers one perspective by which an active learning paradigm could be implemented in real-time.

One might assume that asking the user for specific labels in a real-time setting, e.g. in the stream-based active learning paradigm, can prove difficult and indeed, no hardware setting is readily available to capture such user input. To minimise user input when the active learner is requesting incoming data to be labelled, the user has to specify only whether the decoder prediction was correct or not (a binary choice). Real-time prosthesis movement is performed only after unlabelled samples from movement intention pass the safety check (certainty/confidence criteria) or the user has confirmed the decoder prediction. The additional *R&I* criteria on incorrect predictions (false positives) would work as a filter for noise/outlier detection and sample re-selection for weak supervision.

This way, samples will augment additional information and can be re-integrated into the pool. Furthermore, this approach has the potential to minimise the annotation/labelling noise. It is possible that additional steps may be required to resolve issues around the user bias.

Active learning offers a modular structure that should scale effectively with growing data volume, additional sensing modalities, expanding movement classes and new underlying models. This scalability, however, is constrained by the computational

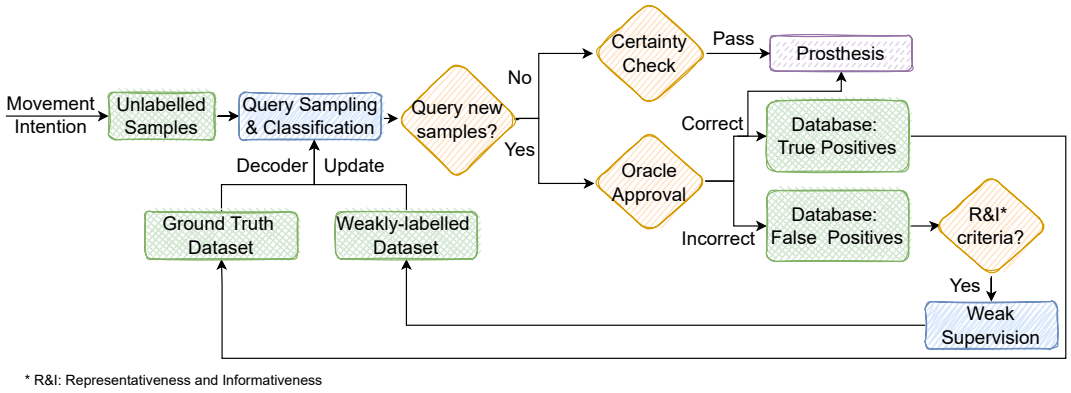


Figure 3.6 The envisaged real-time operation of a user-in-the-loop setting with active learning. Movement intention generates new signal data each second (w.r.t. the sampling rate). Unlabelled samples after being preprocessed are classified by a pre-trained decoder. An active learner using query sampling criteria decides whether to request the label or not. Confident samples pass the certainty check criteria to be sent to the prosthesis actuator. If a sample is queried by the active learner, the oracle dis-/approves the decoder prediction. Incorrect predictions are filtered with R&I (representativeness and informativeness criteria) for outlier detection and *interesting* samples to be weakly-labelled. These, combined with ground truth samples, are used to update the decoder.

cost tightly coupled to the complexity of the underlying model and the real-time requirements of the prosthetic control environment.

3.4.2 Limitations

A limitation of this work is in using a simulated human annotator as an oracle for the active learning framework, meaning the entire dataset contains pairs of $\{x, y\}$ data. Nonetheless, labels were introduced to the learner only when requested by the query strategy.

In this study, we focused on the informativeness of the samples. Further experiments with alternative query strategies may be needed for a more detailed understanding of the potential benefits of active learning for prosthetic control. Considering long-term adaptation of the prosthesis in the real-world, the exploration versus exploitation paradigm could be further investigated. The trade-off between data informativeness and representativeness could potentially enhance the decoder functionality. While initially greedy feature space exploration enables the model to rapidly boost its per-

formance, with time, defining specific grip characteristics could possibly establish its singularity aspects.

While our framework is inherently data-agnostic and could accommodate multi-modal inputs, we deliberately limited this study to EMG-only data. Although sensor fusion with inertial measurement units (IMUs) has demonstrated performance improvements in gesture recognition (Krasoulis et al., 2017), IMUs could introduce additional challenges, potentially increasing recalibration frequency and user burden (Tommasi et al., 2012), which should be further investigated.

Data used during this study presented an example of a high-quality dataset with multiple subjects, including people with limb differences. From each participant, data was recorded in two distinct, but consecutive, sessions in one day. Between the two sessions, the electrodes remained stuck to the skin with the help of adhesive tape. This arrangement minimised the between-session variability in the data, as one would expect to see with dry electrodes, especially when considering the donning and doffing of a prosthesis socket. This is neither a limitation of the proposed active learning method nor a shortcoming of the paper. Nevertheless, this repeatable recording method minimised the additional benefit of using active learning.

Future work should address ambiguous data through two key extensions: implementing an explicit reject class to handle co-contractions and unintended movements safely, and mitigating anomalies through more sophisticated query sampling strategies combining uncertainty sampling with density-weighting (Settles and Craven, 2008) to down-weight isolated outliers while prioritising informative samples. Additionally, tracking and clustering outlier patterns could enable incremental discovery of emergent gesture variations during long-term adaptation with oracle confirmation.

Commercial myoelectric controllers such as Sense (IBT, USA), Myo Plus (Ottobock, Germany) and CoAPT Gen2 (CoAPT, USA) provide supervised recalibration options with varying levels of user-friendliness. CoAPT also offers Adaptive Advance™as an automatic, continuous-learning algorithm that combines new calibration data into the existing state of users' control. However, no technical information about

this adaptive calibration method is available, beyond the generic text in paragraph [0011] of CoAPT pattern recognition patent (WO2020172261).

3.5 Conclusions

We explored the feasibility of using active learning for pattern recognition-based myoelectric control. We observed that all adopted active learning strategies improve decoder adaptation and significantly reduce or optimise the amount of training data on a class-specific basis. Query strategies, such as smallest margin sampling and least confidence uncertainty, were found to improve decoders' performance when compared to conventional random sampling. Future work will include the development of a setup for testing the active learning in real-time and with a prosthesis user in the loop.

Chapter 4

It's GREAT: Gesture REcognition for Arm Translation

This chapter addresses a critical challenge in the field of myoelectric control systems: maintaining consistent performance across various arm positions. In our previous Chapter 3, we explored the problem of performance deterioration over time and proposed an active learning framework to address long-term adaptation challenges. However, real-world deployment of myoelectric control systems faces additional complexities that require further investigation.

Activities of daily living demand a wide range of movements, and myoelectric control accuracy is significantly affected by confounding factors such as muscle fatigue, perspiration, electrode position shifts and changes in arm position. The latter factor has received comparatively less attention in the literature despite its substantial impact on signal quality and decoding accuracy.

To address this gap, we present GREAT, a novel dataset of surface electromyographic (EMG) signals captured from multiple arm positions. This dataset, comprising EMG and hand kinematics data from 8 participants performing 6 different hand gestures, provides a comprehensive resource for investigating position-invariant myoelectric control decoding algorithms. We envision this dataset to serve as a valuable resource for both training and benchmarking arm position-invariant myoelectric control algorithms.

Furthermore, we introduce an innovative data acquisition protocol that can be utilised for future data collection efforts, expanding the availability of EMG signal data that captures variability across diverse arm positions. This protocol represents a significant methodological contribution to the field.

The chapter is structured as follows: in Section 4.1, we explore the various use cases for myoelectric control and review current approaches that attempt to resolve issues arising during position variation. Section 4.2 details our methodology, including the acquisition apparatus and data collection protocol we developed for this study. In Section 4.3, we present a thorough analysis of position effects on model performance, examining position classifiability of each grasp and proposing a hierarchical approach that leverages position knowledge to improve gesture classification accuracy.

Finally, Section 4.4 emphasises the significance of our experimental findings and discusses how our data collection protocol extends beyond a single plane to a full Cartesian space, thereby increasing the variability and representativeness of the EMG data. We conclude by highlighting the broader applications of the GREAT dataset, which extends beyond assistive technologies for individuals with limb loss to include extended reality (XR) applications and other domains requiring robust gesture recognition across varying spatial contexts.

4.1 Introduction

Surface electromyographic (EMG) signals are widely employed in academic research on hand gesture recognition for upper-limb prosthetics control (Atzori et al., 2014a; Castellini and Van Der Smagt, 2009; Englehart et al., 1999; Fougner et al., 2012; Phinyomark and Scheme, 2018a; Scheme and Englehart, 2011), rehabilitation exoskeletons (Burns et al., 2019; Cisnal et al., 2021; Leonardis et al., 2015; Topini et al., 2022), robotics teleoperation (Antuvan et al., 2014; Côté Allard et al., 2016; Fukuda et al., 2003; Hassan et al., 2020; Luo et al., 2020) and more recently incorporated in the extended reality (XR) applications (Côté-Allard et al., 2021; Dwivedi et al., 2020; Gaballa et al., 2022; Huang et al., 2021).

Research in upper-limb prosthetics has traditionally utilised the EMG signals of the participant's forearm muscles, while they simulate performing different grasps. These signals are then used to train a grasp classifier, which in turn identifies the grasp intentions and drives the prosthesis. Similarly, the EMG signals are used in diverse teleoperation applications. Even though academic research has resulted in successful classification of up to 52 grasp classes (Atzori et al., 2012) in offline settings, clinical solutions give only limited capabilities to the user, and that entails frequent recalibrations, as can be seen in the solutions offered by Coapt, LLC. This mismatch is induced by various factors that in controlled settings are minimised.

More recently, gesture recognition for human-computer interactions (HCI) has served as a source of inspiration for the extended reality applications and the metaverse (Lyu, 2023). The use of EMG-based grasp classification in immersive interfaces that are using Virtual/Augmented Reality (VR/AR) has seen an increase in popularity over the last few years (Blana et al., 2016; Dash and Lahiri, 2020; Dwivedi et al., 2020; Kwon et al., 2021; Pereira et al., 2020; Toledo-Peral et al., 2022). As research interest in this area grows, Meta in 2021 revealed their vision of the metaverse and introduced their EMG-powered wrist-based platform (Ctrl-labs at Reality Labs et al., 2024). Designing multimodal HCI systems (vision and sensor data combined) is seen as synergistic, leading to investigating sensor data robustness, such as glove, inertial measurement units (IMU), EMG, and other haptic technologies.

Sensor-based solutions can provide a rich source of information. However, signal quality and pattern variability can be influenced by many intrinsic and extrinsic factors. Specifically in the case of EMG signals, these factors include muscle fatigue, perspiration, shifts in electrode positions, changes in the participant's arm positioning, or a combination thereof (Cifrek et al., 2009; Hargrove et al., 2006; Kyranou and Erden, 2018; Young et al., 2011). Several studies focused on how to mitigate the effect of such changes in inter-subject settings when the signal is recorded over a long period of time, between different sessions on the same day, between different days or between different arm positions in space (Geng et al., 2012c; Hargrove et al., 2008; Jiang et al., 2013a). Specifically, the effect of posture change and arm position has been

repeatedly shown to reduce the robustness of the myoelectric control system, which in effect hinders the adaptation of such systems for real-life applications (Campbell et al., 2020b; Fougner et al., 2011; Kaufmann et al., 2010; Khushaba et al., 2016, 2014; Liu et al., 2012; Park et al., 2015; Radmand et al., 2014b,c; Rajapriya et al., 2021; Stuttaford et al., 2024; Teh and Hargrove, 2020; Williams et al., 2022).

To limit the effect of arm translation, previous research has focused on two main factors. First, participant training combined with feedback has been shown to minimise the signal variability caused by limb position (Stuttaford et al., 2024). Second, from a computational perspective, the most popular approach is to record data that captures the aforementioned variability, by recording data from various positions and utilising machine learning techniques that compensate for the effect of arm position (Fougner et al., 2011; Geng et al., 2012c; Hahne et al., 2012; Khushaba et al., 2014; Liu et al., 2012). Those two approaches are not mutually exclusive, and both benefit from a large database of EMG signals during which the arm position factor has been considered. Such a dataset will lead to the development of more advanced and more robust machine learning paradigms, which in turn would lead to biofeedback paradigms faster and possibly more intuitive.

The first step in creating robust algorithms in the presence of signal drifts is creating a dataset that captures such shifts. A number of datasets from a range of recording conditions have been compiled and published. Table 4.1 summarises the available online datasets containing EMG recordings of multiple participants, performing a set of predetermined grasps under the previously mentioned causes of signal variation. For the scope of this paper, we are focusing only on the datasets that are recorded in any of the three following cases: A) two or more sessions in a day, B) a period of more than one day, and C) various spatial arm orientations.

As of the time of writing, there are only three datasets containing EMG data recorded from various limb positions created with the goal of investigating the effect of limb position, also known as arm translation, on the recorded EMG signal. Khushaba et al. (2014) and Hahne et al. (2012) record EMG signals from five distinct limb positions in a vertical plane. The biggest difference between these two datasets is the

amount of sensors used. Khushaba et al. (2014) utilised seven sensors located around the forearm, while Hahne et al. (2012) employed a forearm band consisting of an array of 96 sensors. Finally, the SeNic (Zhu et al., 2022) dataset includes recordings of 7 grasps from 3 different arm positions: with the elbow bent and supported on a table, an unsupported bent elbow and adjacent to the participant's body. In terms of datasets involving recordings from multiple sessions or spanning across various days, there is a broader selection available online. The duration of these recordings ranges from two to a maximum of ten days span (Cote-Allard et al., 2021; Du et al., 2017; Fang et al., 2018; Kaczmarek et al., 2019; Palermo et al., 2017; Pradhan et al., 2022). The majority of those datasets are recorded in a static setting, with the arm adjacent to the body over the whole period of the experiment. Notably, in Cote-Allard et al. (2021), the training recording happens in one position, but the test data include recordings from a range of angles in space. Since the target of the arm position is determined with respect to the Virtual Reality reference point, it gives the participant freedom of movement to reach the target without many constraints. Consequently, it lacks structured information about arm positioning, which makes it difficult to use in future studies for arm translation.

A more dynamic approach is followed for the recording of the MeganePro DB1 dataset (Cognolato et al., 2020), where the recordings capture dynamic movements in both vertical and horizontal planes. Due to the object's location, the experiment does not explore the spatial resolution in a structured way. The primary focus of the MeganePro DB1 is not the arm translation analysis, but rather combining eye tracking with EMG for the purposes of object detection and grasp activation.

Our contribution to the recorded dataset encompasses several key aspects in capturing the shifts in the EMG signals due to arm translation. Firstly, we designed a 3 $\times$3 grid that includes a horizontal plane, complementing the data sets and experiments that only include a vertical plane movement (Hahne et al., 2012; Khushaba et al., 2014). We provide a structured and systematic format to facilitate extensive data analysis with precise annotations for arm translation and grasp activation. Furthermore, to minimise bias and optimise our dataset's utility, we ensure the data is balanced across all the

Dataset	Limb Positions	# Days	# Sessions per day	Trials per session	Gestures	Subjects	EMG Channels
Khushaba et al. Khushaba et al. (2014)	5	1	1	6	8	11	7
Hahne et al. Hahne et al. (2012)	5	1	1	25-35	6	10	96
Palermo et al. Palermo et al. (2017)	1	5	2	12	7	10	14
GRABMyo Pradhan et al. (2022)	1	3	1	7	17	43	28
PutEMG Kaczmarek et al. (2019)	1	2	1	20	8	44	24
Côté-Allard et al. Cote-Allard et al. (2021)	1	3-4	1	4	11	20	10
ISRMYO-I Fang et al. (2018)	1	10	2†	1	13	6	16
CapgMyo DB-b Du et al. (2017)	1	2	1	10	8	10	128
MeganePro DB1 Cognolato et al. (2020)	2‡	1	1	8	10	45	12
MyoBitChen et al. (2023)	2	3	1	1	7	24	16
SeNicZhu et al. (2022)	3	3/10	1	3	8	36§	8
GREATSzymaniak et al. (2024)	9	2†	2†	5	6	8	16

† not removed the sensors between 2 sessions of same day.
† only one position is the same between different days
‡ this refers to them recording in seated and standing positions. They also allow the grabbing objests from different locations on the table.
§ only a subset of the participants are used to record different aspects of the dataset.

Table 4.1 Table summarising the existing online datasets that contain recordings from different positions, from different sessions in one day or over the period of multiple days.

positions. Lastly, we conducted our experiments over a span of two days, allowing for more variability of the data and its implications.

4.2 Method

4.2.1 Subjects and Ethical Requirements

The data was recorded from the right hand of 8 intact subjects, all male, over the period of two consecutive days. All the participants were given a written and oral explanation of the experiment protocol and signed an informed consent form (example provided at Appendix A). The experiment was conducted according to the Declaration of Helsinki principles for medical research involving human subjects (Association, 2013), and it was approved by the Ethics Commission of the University of Edinburgh, ethics approval number: 2019/89177.

4.2.2 Acquisition Apparatus

The dataset consists of recordings of hand kinematics and electrical muscular activity, while the subjects are performing specific hand gestures. The apparatus used in the study comprised a motion capture dataglove and a surface EMG recording system connected to a single laptop responsible for the data acquisition.

For the **hand kinematics** recordings a wireless motion capture data glove was utilised. More specifically, the 18-sensor CyberGlove II data glove (CyberGlove Systems LLC), which can be seen in Figure 4.1a, was used to record data from two bend sensors on each finger, four abduction sensors, plus sensors measuring thumb crossover, palm arch, wrist flexion, and wrist abduction. The data glove is calibrated with respect to the anatomic differences of every participant to ensure accurate recordings. We performed the calibration once at the beginning of the experiment for each participant, following a protocol provided by the CyberGlove Systems LLC.

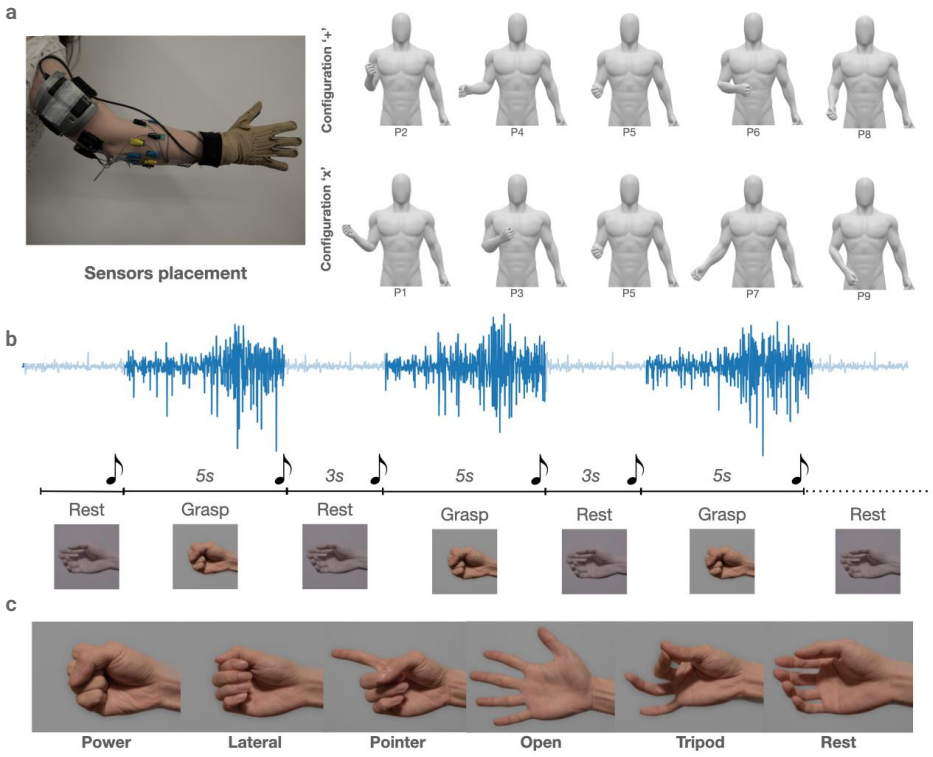


Figure 4.1 Figure representing the recording protocol. The first row (Part a) shows the sensors positioning on the hand and arm of the participant (we have acquired informed consent to publish the participant's likeness) and the nine different positions that the participant is asked to move their arm while recording the dataset. Configuration Con_+ is always recorded on day 1 and Configuration Con_x is always recorded on day 2. Part b shows the signal recording process; after a period of rest, during which the signal is not recorded (indicated with light blue colour) the participant is prompted to perform a grasp indicated by a visual cue and a beep sound. The participant is required to hold the grasp for 5 sec, followed by 3 sec of rest. This process corresponds to one trial and is repeated 5 times for a predefined grasp and position. The grasps that are recorded during this experiment are shown in part c.

The **muscular activity** is recorded by 4 Trigno Quattro Sensors (Delsys, Inc), each one consisting of a reference electrode and 4 sEMG recording channels, which totals 16 acquisition channels. The sEMG signals are sampled at a rate of 2kHz and were fixed on the forearm using the Delsys adhesive interface. Prior to electrode placement, the participants' skin is cleansed using 70% isopropyl alcohol. We then placed the acquisition electrodes in two rows of 8. The first electrode was positioned on the extensor carpi ulnaris muscle, identified by palpitation, and the rest of the sensors were placed equidistantly around the forearm without targeting specific muscles. For

the second band, the electrodes were positioned between two electrodes from the first row. Two of the reference electrodes of the 4 Quattros we used were placed close to the elbow, and the other two next to the wrist (as shown in Figure 4.1a. The reference electrodes on the wrist are hidden by the data glove in this figure). After visually inspecting the signal quality of all EMG channels in the acquisition software, an elastic band was placed around the electrodes to keep them securely fixed during the acquisition period.

4.2.3 Acquisition Protocol

Following the sensors positioning and data glove calibration, participants were presented with a computer interface that instructed them to move their arm at a specific position and perform one of the following grasps: power, lateral, pointer, tripod, open and rest (Figure 4.1c shows a picture of those grasps). The software responsible for the computer interface and data acquisition was implemented in Python using the Axopy package (Lyons and Margolis, 2019).

The participant is visually instructed to position their forearm in one out of nine different orientations, a 3x3 grid appearing on the monitor. Let P_n be the position, where $n \in \{1,\ldots,9\}$ starting from the top left corner with P_1. Consequently, P_5 refers to the central part of the grid. It's a *neutral* position that corresponds to the forearm being at a 90° angle with respect to the upper arm and parallel to the ground. The rest of the positions are reached by moving the forearm by a 45° angle from this *neutral* position towards the direction indicated by the computer prompt. For example, P_2 requires the participant to move their forearm 45° up, and P_1 requires 45° up and 45° left, always with the *neutral* P_5 position as reference. The participant is instructed to keep their upper arm in a stable but relaxed position. The nine different positions are presented in the right half of Figure 4.1a.

The participant gets familiar with the positions and the six grasps that we expect them to perform for each experiment (Figure 4.1c) before the recording starts. Then, the graphical user interface is launched, where the image of the desired grasp appears only on the cell in the 3x3 grid that corresponds to the prompted position. In order to

eliminate collecting data corresponding to transitioning from the rest phase to grasp, we ensure that prior to each recording, the participant's arm is located in the desired position and the hand forms the prompted grasp correctly. Once the participant is ready, we press the enter key on the keyboard to initiate the recording of the data.

One ***trial*** in our experiment consists of 5*sec* of holding the prompted grasp, followed by a resting period lasting 3*sec*. This sequence is repeated five times for every grasp in each specific position. Each ***block*** of five trials is then followed by a longer period of rest, during which the participant can relax their arm at any position they like, for their preferred duration of time. The positions sequence is predetermined for each day: Con_+: {2,4,5,6,8} and Con_x: {1,3,5,7,9} for day 1 and day 2, respectively.

In addition to the visual cue, where the participant is presented with the grasp image in one cell of the grid and the remaining time/repetitions for completing the experiment, an audio cue is used to indicate the start of each 5*sec* recording (sine wave; frequency 400Hz; duration 100ms) and a different audio cue (sine wave; frequency 300Hz; duration 100ms) to indicate the end of the recording and transition to rest pose. The audio cue to start the recording starts 250*ms* before the time our program starts to record, in order to allow enough reaction time for the participant and avoid recording transitional data from rest to grasp.

The grasp sequence between subjects is pseudo-randomised to minimise potential biases. Our data recording protocol shuffles grasps and positions to reduce participants' fatigue as a result of numerous repetitions of the same movement. Moreover, this makes the experiment less monotonous and boring based on participants' feedback.

The first ***session*** of recordings is completed after 30 trials (5 repetitions of 6 grasps) across all positions, giving us a total of 150 grasp trials. After a longer rest period, the experiment is repeated one more time (second session). The break between the two blocks of recordings was about half an hour long (including the doffing/donning of the dataglove), allowing the subject to leave the room. When the subject returned to the room, they were more rested, and after some activity (often involving stretching), we reported the two blocks recorded in one day as two separate sessions to account for any potential drift that could be present in the signal after the break. The final dataset

consists of 2 sessions of 5 grasp repetitions (trials), on 5 different positions per day (Con_+ or Con_x), lasting 2 days.

The between-days recording was not performed with the purpose of capturing the effect of a two-day span experiment, but with the goal of adding variability on the arm positions. However, position 5 (neutral) is one that repeats over the two different days.

4.3 Technical Validation

4.3.1 Task Overview

Our objective to demonstrate the validity and significance of the collected dataset. Specifically, we propose a set of experiments which in the future should be referred to as a baseline for arm translation tasks. Overall, our experiments are as follows: 1) Standard classification with data only from a single position, 2) Naive transfer learning classification between positions, 3) Classifying positions for each grasp, 4) Hierarchical multi-label classifier (HMC) with position and grasp dependencies.

4.3.2 Signal Processing and Feature Extraction

The protocol for signal pre-processing was applied to the acquired data, ensuring high-quality data modelling. During the recording, the data was filtered using a 4th-order Butterworth filter (with lowcut and highcut frequencies of 20 Hz and 450 Hz, respectively). To remove powerline interference, a Notch filter (Orfanidis, 1995) was applied. After offset correction, we apply a sliding window of 128 ms with a stride of 50 ms. From each window, a set of features and a class label were calculated. To emphasise the universality and feasibility of the data, we choose to extract a well-established set of features commonly used in myoelectric control. Hudgins et al. (1991) features constitute mean absolute value, zero crossing, slope sign changes and waveform length. The grasp activity of a window is assigned a class label based on the majority vote, similar to Wei et al. (2019). Each participant's data has been processed separately for every session based on the settings we presented above. We prepare

data by separating it by trials, shuffling and then splitting it into training and testing data at an 80:20 ratio. Lastly, we normalise data for each sensor using the z-score.

4.3.3 Classification

In all the scenarios, we apply a Linear Discriminant Analysis (LDA) classifier, which has been shown over the years as one of the most reliable algorithms used in myoelectric control (Atzori et al., 2014a; Englehart and Hudgins, 2003; Fougner et al., 2011; Hargrove et al., 2010; Scheme and Englehart, 2011; Young et al., 2012).

Standard classification: We start with a naive approach of grasp classification within the same position. This scenario is comparable to other offline pattern recognition-based encoders (Campbell et al., 2020b; Fougner et al., 2011; Khushaba et al., 2014; Liu et al., 2012). For each position, an independent classifier encodes six gestures using 5-fold cross-validation. The results from all subjects are then averaged (the same protocol applies to the remaining scenarios). We achieve on average, an accuracy of around 96% in each position. While these results illustrate an impressive level of classifiability in the conventional approach, we implicitly include certain assumptions, such as data coming from the same distribution.

Con_x	Accuracy	Con_+	Accuracy
Position 1	0.96 σ: 0.03	Position 2	0.95 σ: 0.04
Position 3	0.96 σ: 0.02	Position 4	0.96 σ: 0.03
Position 5	0.96 σ: 0.02	Position 5	0.96 σ: 0.02
Position 7	0.96 σ: 0.02	Position 6	0.96 σ: 0.02
Position 9	0.96 σ: 0.02	Position 8	0.95 σ: 0.02

Table 4.2 This table presents a classic scenario of grasp classification with data coming from the same distribution (in our case, the arm position).

Naive transfer learning: The assumption of i.i.d. (independent and identically distributed) in standard machine learning refers to data being independently and randomly sampled from the same underlying distribution. This becomes problematic

when the model is presented with new, out-of-distribution data, struggling to encode the unfamiliar patterns in the data. We investigate the effect of arm positioning variance and its significance in translating the knowledge between positions. Hence, we conduct an experiment where we train the model on data from one position and test on another position from the same configuration (Con_+ or Con_x).

Arm Translation in Con_+

	Position 2	Position 4	Position 5	Position 6	Position 8	OVR $P \setminus P_s$
Position 2		0.88; σ: 0.08	**0.91; σ: 0.06**	0.87; σ: 0.08	0.86; σ: 0.08	0.88
Position 4	0.90; σ: 0.07		**0.92; σ: 0.03**	0.87; σ: 0.08	0.88; σ: 0.03	0.89
Position 5	0.90; σ: 0.06	0.90; σ: 0.05		**0.92; σ: 0.03**	0.91; σ: 0.03	0.91
Position 6	0.89; σ: 0.05	0.87; σ: 0.04	**0.92; σ: 0.03**		0.89; σ: 0.03	0.89
Position 8	0.85; σ: 0.08	0.87; σ: 0.05	**0.90; σ: 0.04**	0.88; σ: 0.07		0.87

Arm Translation in Con_x

	Position 1	Position 3	Position 5	Position 7	Position 9	OVR $P \setminus P_s$
Position 1		0.88; σ: 0.08	**0.91; σ: 0.05**	0.85; σ: 0.07	0.85; σ: 0.10	0.87
Position 3	0.90; σ: 0.06		**0.92; σ: 0.03**	0.84; σ: 0.07	0.91; σ: 0.03	0.89
Position 5	**0.91; σ: 0.05**	**0.91; σ: 0.05**		0.89; σ: 0.05	**0.91; σ: 0.05**	0.90
Position 7	0.88; σ: 0.06	0.85; σ: 0.08	**0.90; σ: 0.07**		0.88; σ: 0.05	0.88
Position 9	0.88; σ: 0.05	**0.91; σ: 0.05**	**0.91; σ: 0.04**	0.88; σ: 0.04		0.90

Table 4.3 Arm translation results for independent models trained on the source position P_s and tested on target position P_t given in each row and column, respectively. One versus the Rest (OVR) presents a scenario where a model trained on a P_s is tested on remaining positions within this configuration.

The presented results in Table 4.3 show the accuracy of independent models trained on source position P_s (row) and tested on target position P_t (column). We demonstrate a performance drop of up to 10% in comparison with a standard protocol (training and testing dataset coming from the same position). As expected, in most cases testing on the central position (P_5) results in the highest classification score, between 89-91%. In the One versus the Rest (OVR) setting, the model is trained on a single position (P_S) and tested on the remaining positions within this configuration ($P \setminus P_S$). In the OVR setting, we also observe a performance drop with central position (P_5), giving the best results.

Classifying positions for each grasp: We observed a significant performance drop when a model was tested on the data from another domain, namely different position

P_t. In transfer learning two domains are defined as different when we observe either discrepancies in feature space or in marginal distributions. To investigate the cause of the observed behaviour, we take a closer look at positioning and its classifiability. By classifying positions (within the given configuration) for every grasp activation, we observe that, irrespective of the gesture, the performance accuracy is on average around 80%.

Position classification						
	Power	Lateral	Tripod	Pointer	Open	Rest
Configuration x						
Acc.	0.82	0.83	0.81	0.86	0.86	0.79
Configuration +						
Acc.	0.78	0.80	0.73	0.85	0.82	0.73

Table 4.4 Position classification in each grasp for both configurations, Con_+ and Con_x.

Hierarchical Multi-label Classification (HMC): Finally, we look at how critical it is to account for differences in arm positioning when trying to identify what type of grasp is being used in everyday situations (such as reaching for a glass of water on a table versus a high shelf). A hierarchical multi-label classifier (HMC) (Giunchiglia and Lukasiewicz, 2020) is applied to leverage position knowledge to classify the gesture. We define the problem as a hierarchically organised tree where each prediction must be coherent with respect to a hierarchy constraint. The hierarchy constraint imposes that a data point assigned to a specific class must also be assigned to all of its predecessors within the hierarchy (Giunchiglia and Lukasiewicz, 2020). Our HMC, shown in Figure 4.2, consists of a position encoder (parent node) and $|z|$ leaf nodes, corresponding to the nine positions. During the training, data from all positions across both configurations (with position 5 being used only from one configuration) are merged and split in the same manner as in the prior experiments (a.k.a 80:20 ratio on trials with 5-fold cross-validation). Firstly, we train the position encoder E_p with $\{x_i, z_i\}$ data pairs from $\{x_i, y_i, z_i\}_{i=1}^{|D|}$ where z represents the position and y the grasp of each sample x for all dataset samples $|D|$. The augmented features (through concatenation: $\{x_i \oplus norm(z_i')\}$) are grouped with respect to the ground truth z. Based

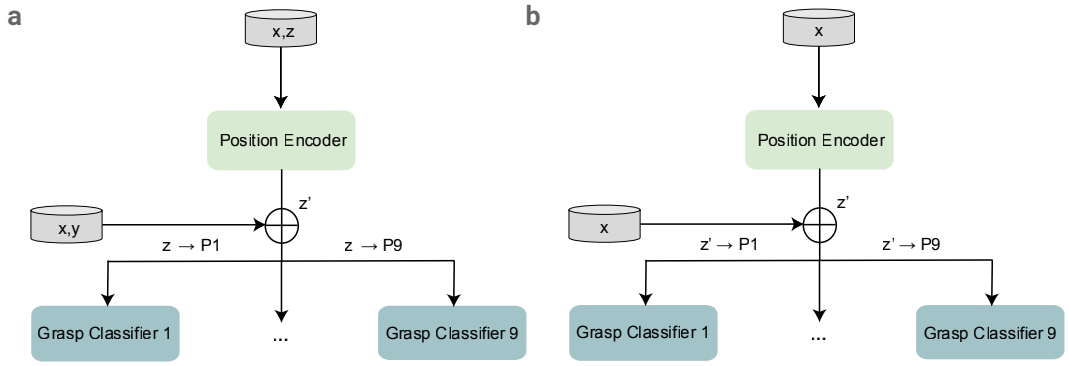


Figure 4.2 The architecture of the Hierarchical Multi-Label Classification model. In Part a we show the first level in which we train the position encoder E_p with $\{x,z\}$ data pairs to predict position z'. Here z represents the position and y the grasp of each sample x. We then modify the input vector by concatenating x with normalised z' and pass it to train the grasp classifier E_g, where z value is equivalent to the ground truth in each sample. Part b shows the process in the testing phase. Concretely, once E_p predicts z', the sample $\{x,z'\}$ is then input to the grasp encoder E_g, where $z' = g$ (g being a grasp). Soft HMC, freezes the constraint $z' \equiv z$ when calculating the accuracy of the model.

on the z corresponding to the nine grasp encoders $E_{g:n}$, nine independent models where trained for grasp $y_i \in \{1,...,6\}$ classification.

During the inference time, based on the prediction z' of E_p the sample is assigned to the encoder $E_{g:z'}$ to infer the grasp activity. We verify the overall performance by calculating accuracy for every $\{x_i \oplus norm(z'_i)\}$, using multi-label pairs and their predictions $\{y'_i, z'_i\}$. In the Soft Hierarchical Classification, the evaluation metric is modified by freezing the condition on position accuracy, namely $z'_i \equiv z_i$.

Interestingly, we observe a drastic performance drop for HMC compared to standard classification (Table 4.2), which does not hold position knowledge. Instead of performing grasp classification using train and test data coming from a single position, in the case of HMC, the model is conditioned on the positioning information. Despite data augmentation and incorporating position information into the data, the multi-label evaluation metric causes a significant drop in performance (Table 4.5). This is primarily because the initial position encoder (E_p) struggles to separate the data reliably based solely on position (z), achieving only 58% accuracy. The HMC's design imposes a strict hierarchical constraint: an error in the initial position prediction stage

Hierarchical Multi-Label Classification

P_1: 0.63; σ: 0.10	P_2: 0.51; σ: 0.07	P_3: 0.63; σ: 0.11
P_4: 0.53; σ: 0.11	P_5: 0.39; σ: 0.07	P_6: 0.53; σ: 0.14
P_7: 0.68; σ: 0.11	P_8: 0.55; σ: 0.12	P_9: 0.61; σ: 0.09

Soft Hierarchical Classification

P_1: 0.94; σ: 0.3	P_2: 0.92; σ: 0.06	P_3: 0.93; σ: 0.04
P_4: 0.94; σ: 0.02	P_5: 0.94; σ: 0.03	P_6: 0.93; σ: 0.03
P_7: 0.95; σ: 0.02	P_8: 0.94; σ: 0.03	P_9: 0.94; σ: 0.04

Table 4.5 Results from Hierarchical Multi-Label Classifier (A). In the first level, Position encoder E_p, classifies data with an accuracy of 58% and std. 0.05. Below, there is a results table for Soft Hierarchical Classifier (B). The tables present classification scores for individual grasp encoders E_g at every position w.r.t. y and z (A) or only y (B).

inevitably leads to an error in the final multi-label output ($\{y', z'\}$). This difficulty arises because position-related variations in sEMG though sufficient to cause covariate shift affecting grasp classification, do not produce sufficiently distinct patterns for the linear position encoder to reliably separate.

When relaxing the constraint on position accuracy through a 'soft' evaluation metric, the model achieved accuracy exceeding 92% (Table 4.5), with the model assessed exclusively on grasp classification outcomes. This performance demonstrates that the grasp-specific classifiers (E_g) learned the gesture features effectively, but the system's performance under the strict metric was bottlenecked by the difficulty of the initial position classification stage. It is crucial to note that this result was obtained under an i.i.d. assumption where all positions were present in the training set, a condition that cannot be guaranteed in practical deployment scenarios. Yet, we still do not grasp the characteristics and information within the model of arm translation.

4.4 Conclusion and Discussion

In this paper, we present the scenario of arm translation in a static setting with a 3x3 grid point of reference. To the best of our knowledge, our dataset is the first of its kind, where such extensive position resolution has been presented. This claim refers specifically to datasets integrating multi-channel EMG with detailed, wearable hand kinematic sensors for myoelectric control research. While image-based hand tracking systems may offer alternative forms of high-resolution spatial or kinematic tracking, they operate under different principles and face distinct challenges, such as line-of-sight occlusion, not present in our wearable sensor approach (Ahmad et al., 2019; Lim et al., 2020).

We hope the GREAT dataset will provide new inspiration and use cases for bridging the gap between data collected in a controlled environment and real-life scenarios. We demonstrated the importance of data variability in arm translation and its impact on myoelectric control signal decoding. We envision the following possible applications.

Most of the online available datasets in the myoelectric control are collected in a single position. While these established datasets provide reliable information, they lack the representation of variability as a result of arm translation. The effects of this are observed in Table 4.3. Our aim with the GREAT dataset is to provide a new source of training to address this issue and assist in bridging the gap between myoelectric-based grasp recognition in an experimental setting versus a real-life scenario.

We proposed a novel design protocol to capture the intricacies of EMG signal in arm translation. Using a 3x3 grid as a reference, we prompt the participant to explore movement in both vertical and horizontal planes.

As obtaining data from the target population (individuals with limb loss) is challenging, it often necessitates initial research using able-bodied participants. This approach is justified and widely adopted in myoelectric control research for characterising fundamental confounding factors (Atzori et al., 2014b; Campbell et al., 2020a; Fougner et al., 2011; Khushaba et al., 2014). Both populations experience position-induced changes in EMG signals through shared physiological mechanisms, such as muscle topology alterations, fiber-electrode geometry changes, and gravitational

effects on the muscle co-activation (Chen et al., 2011; Fougner et al., 2011; Scheme et al., 2010). Although the magnitude may differ due to shorter residual muscle lengths and prosthesis-specific factors such as socket compression and device weight (Cipriani et al., 2011; Jiang et al., 2013a), prior studies report similar directions of performance degradation across both populations during inter-position tasks (Fougner et al., 2011; Geng et al., 2012b; Jiang et al., 2013a), supporting the fundamental translatability of our findings. Nevertheless, future validation with individuals with limb loss remains essential to quantify population-specific differences and adapt our protocol to account for unique biomechanical constraints of prosthetic use.

The current benchmark datasets that include data from different arm positions, such as Khushaba et al. (2014) and Hahne et al. (2012) have played a critical role in myoelectric control. We believe that extending the data collection protocol from a single plane to a Cartesian increases the variability of the EMG data. It can serve as a new and challenging benchmark dataset for future research in the mitigation of the adverse effects of arm translation.

Moreover, XR applications are adopting EMG to create a more immersive experience for the user. Such integration of EMG with XR inherits the problems that arise from the EMG instability due to arm translation and variability over different days. Until now, the only publicly available datasets created for XR control that include EMG data have recorded gestures from one specific posture. GREAT dataset can be utilised as a benchmark dataset by the related XR research that attempts to mitigate the disturbances resulting from the previously mentioned variability.

The total recording duration per day of our dataset is constrained by hardware and human limitations. The myoelectric sensors need to be removed and charged after a maximum of three hours of consecutive recording in order to avoid malfunction. This limit was coupled to the participants' limits of focus and performance before mental and physical fatigue were present. The existing datasets that are recording data from a large number of gestures or more repetitions are most commonly recording from one supported arm position (Atzori et al., 2014a; Palermo et al., 2017). This posture has a different impact on the effort that is required by the subject to perform one trial

compared to our experiment. Our subjects are not resting their hand on the desk, and they are asked to keep their arm for a period of approximately 40 seconds at a position that they might not be used to holding their hand in (for example, position 3) while performing the grasp sequence. Our solution to the participant's fatigue was to allow for enough self-determined breaks and resting phases and break the recording into two sessions, which allowed the participant to move around and relax. This increased the total timing of the experiment, which affects the total mental fatigue of the subject and the draining of the sensors' battery. We perceive this to be a potential limitation to all datasets that attempt to record a larger amount of data from a wide range of positions.

One important aspect in the creation of datasets that are used in the field of prosthesis control is the inclusion of data from people with limb differences. The importance of inclusion of people with limb differences in myoelectric control experiments has been repeatedly emphasised, and the recordings from able-bodied people and people with limb differences typically come from very distinct distributions (Campbell et al., 2020a; Waris et al., 2018). The importance of such inclusivity is not limited to the prosthesis research and can benefit the XR research as well. Moreover, our current subject pool consists solely of male participants, a practice that may introduce specific biases in the data collection and analysis.

In the future, we hope to extend this dataset by including information from participants with limb differences and female subjects.

4.5 Code Availability

Examples with data read/write, preprocessing, and a simple classification pipeline are provided in a detailed manner in the following repository: https://github.com/MoveR-Digital-Health-and-Care-Hub/MoveR_AT_GREAT. The data collection experiment code can be found at GitHub (https://github.com/MoveR-Digital-Health-and-Care-Hub/posture_dataset_collection) and was implemented in Python using the Axopy package Lyons and Margolis (2019). Data is publicly available at Szymaniak et al. (2024).

Chapter 5

Position Domain Generalisation in Myoelectric Control

This chapter builds upon our previous work in Chapter 4, where we identified limb position effect as a significant yet frequently overlooked confounding factor in myoelectric control systems. Our initial analysis of the GREAT dataset revealed a substantial $\sim 10\%$ performance degradation when evaluating models across different limb positions, highlighting the challenge of developing position-invariant systems.

The ability to accurately decode gestures regardless of limb position directly affects the quality of life for prosthetic users during activities of daily living. Current approaches to address this challenge present significant practical limitations. Transfer learning techniques require additional data collection for each new position. At the same time, domain adaptation methods necessitate access to target domain data, which may not always be available in real-world scenarios. These constraints create substantial barriers to the practical deployment of myoelectric control systems.

To address these challenges, this chapter investigates how to improve within-subject generalisation for limb position variations by leveraging domain generalisation methods to enhance model robustness across different limb configurations.

We state our contributions in Section 5.1, then we review related work in Section 5.2 and define the domain generalisation problem in Section 5.3. Sections 5.4–5.5 present

our framework and the proposed DR-cGMM approach, followed by methodology and results in Sections 5.6 and 5.7.

5.1 Introduction

Recent years have witnessed a surge of interest in human-computer interfaces, with wearable technologies gaining significant traction. Among these, myoelectric control-based gesture recognition systems have emerged as a focal point of research beyond just the limb prosthetics community. While these systems demonstrate nearly flawless classification accuracy in highly constrained laboratory settings, their real-world implementation faces numerous challenges that compromise performance (Vujaklija et al., 2017).

In controlled settings, standardised protocols establish a specific set of parameters for data collection. Although these controlled variables facilitate task optimisation and ensure high signal quality, many constraints become impractical in real-world applications and need further investigation (Kyranou and Erden, 2018). Typically, experimental data is collected in static positions where subjects maintain their arm adjacent to the body with the elbow bent at a 90-degree angle relative to the upper arm and forearm parallel to the ground throughout the experiment.

Meanwhile, confounding factors such as arm translation, often referred to as position effect, are unavoidable when these models are deployed in practical applications. Activities of daily living produce variations in muscle activation patterns due to changes in limb kinematics, sensor contact, and muscle recruitment (Batzianoulis et al., 2018; Radmand et al., 2014a). This position effect causes a distribution shift that leads to substantial degradation in classification accuracy of approximately 10% as shown in the previous Chapter 4. While supplementary modalities such as Inertial Measurement Unit (IMU) data are often incorporated to address this issue, they introduce their own complications, including the need for periodic recalibration and signal precision limitations (Tommasi et al., 2012).

Therefore, in this work, we focus on the real-world motivated and underexplored scenario: within-subject cross-position generalisation during gesture recognition. The position effect acts as a major confounding factor hindering the dexterity and robustness of the myoelectric control gesture classification system. The key assumption is that during deployment, we do not have access to data from the new positions. This rules out common adaptation-based strategies, such as domain adaptation or transfer learning, for recalibration and fine-tuning the model (Wu et al., 2023). Instead, we require models to generalise solely from training data collected in a limited number of positions for a given subject. To elaborate, such a setting is unrealistic in user-oriented wearable systems, where collecting labelled data for every possible posture is infeasible. As the primary goal is to deliver a calibration-free user experience, repeatedly prompting users to calibrate the device can lead to its abandonment (Amsüss et al., 2013b; Burrough and Brook, 1985; Eddy et al., 2024; Tabor et al., 2018).

To tackle this challenge, we formulate the problem as a domain generalisation (DG) question. We show that different deep learning DG methods and statistical models can all be understood under the domain generalisation framing, given different assumptions taken by different models. Using this understanding, we motivate and propose our novel domain-regularised class-conditioned Gaussian Mixture Model (DR-cGMM).

Our main contributions are as follows:

1. **Rigorous benchmark design with nested cross-validation for fair and reliable model comparison.** To ensure unbiased model selection and evaluation, we implement a nested cross-validation protocol that isolates hyperparameter optimisation from final performance assessment. This design supports a robust comparative analysis across all methods, effectively mitigating data leakage while facilitating principled, generalisable performance estimation across heterogeneous model families.

2. **Comprehensive investigation and evaluation across statistical and deep learning models under a domain generalisation framework.** We reinterpret both classical statistical methods (e.g., LDA, QDA) and recent deep learning

domain generalisation (DG) algorithms under a unified framework and conduct rigorous, pairwise statistical comparisons. Our benchmark reveals that empirical risk minimisation (ERM) fails to outperform simpler baselines and that DG methods do not yield clear benefits in this low-data regime.

3. **A novel domain-regularised class-conditional Gaussian Mixture Model (DR-cGMM).** We introduce a new statistical model that integrates domain information into a GMM via a domain regularisation term. DR-cGMM is the only method in our study that significantly outperforms all deep learning methods, highlighting the continued relevance and superiority of principled statistical approaches in low-resource, real-world myoelectric control settings.

In Section 5.2, we introduce related work on limb position effect and domain generalisation methods. Subsequently, we introduce the DG problem formulation in Section 5.3. Using this conceptual understanding of DG, we show that statistical and deep learning methods are instantiations of the DG framework under different assumptions in Section 5.4. Using this framework and understanding, we introduce our novel baseline DR-cGMM in Section 5.5. Finally, investigation design is explained in Section 5.6 with results discussed in Section 5.7.

5.2 Related work

5.2.1 Limb Position Effect

Variations in arm posture, often referred to as the position effect, introduce distribution shifts in EMG signals due to changes in muscle recruitment, sensor alignment, and gravitational influence (Chen et al., 2011; Fougner et al., 2011; Khushaba et al., 2016; Liu et al., 2012; Scheme et al., 2010). These shifts substantially degrade classification performance and are a key obstacle for reliable myoelectric control in daily life scenarios.

Several strategies have been explored to address this issue:

Comprehensive training protocols have emerged as effective strategies by incorporating multiple arm positions during model development. Studies implementing multi-positional training across eight distinct limb positions have demonstrated remarkable accuracy improvements, achieving 86-95% accuracy compared to merely 65% for single-position training paradigms (Scheme et al., 2010). Including diverse positional data in the training protocol effectively helps capture the inherent variability in EMG signals, enabling classifiers to learn more robust representations that generalise across the position space (Beaulieu et al., 2017). Nevertheless, these methods typically assume that the model will only be tested on the same positions it was trained on. In other words, there is no generalisation to unseen positions at deployment. This is a key limitation for real-world use cases where collecting labelled data for every possible posture is impractical.

Knowledge transfer methodologies leverage transfer learning and domain adaptation to bridge the gap between positional domains. Advanced techniques such as Canonical Correlation Analysis (CCA) optimised linear projections between position-specific EMG feature spaces, with enhanced performance when combined with Uncorrelated Linear Discriminant Analysis (ULDA) (Cheng et al., 2018). Other approaches include two-stage cascaded bilinear transforms specifically engineered to mitigate both subject-specific and position-dependent effects from EMG signals (Ishii et al., 2017). Mechanomyogram-based domain adaptation across ten arm postures achieved 64.29% inter-posture classification accuracy in comparison with 61.34% and 58.81% for CNN and SVM, respectively, with performance improving to 71.75% when extending segment windows from 200ms to 600ms (Wattanasiri et al., 2024).

Robust algorithmic frameworks have demonstrated considerable promise in creating position-resilient systems through advanced feature extraction and classification methods. Techniques such as Sparse Representation Classification (SRC) (Betthauser et al., 2016) and naive conditional Gaussian Mixture Models (cGMM) have significantly improved inter-position classification performance (Liu et al., 2014b). Notably, mixed-LDA classifiers have achieved 93.6% accuracy across five static positions compared to 82.5% with standard LDA classifiers (Yu et al., 2017). These approaches

focus on extracting position-invariant signal characteristics while maintaining gesture discriminability.

Multimodal sensing integration enhances EMG by adding sensors like accelerometers or IMUs to capture limb position. These systems typically use cascaded classifiers (limb position detection followed by gesture recognition) or unified classifiers with integrated positional context (Fougner et al., 2011). Other modalities, such as mechanomyography (MMG), have shown promise, with methods like graph Laplacian-based feature extraction achieving up to 94.1% accuracy across 40 motion classes (Geng et al., 2012c; Khushaba et al., 2018). However, integrating these sensors increases dataset size requirements and introduces challenges such as sensor drift and bias (Radmand et al., 2014c; Tommasi et al., 2012).

These complementary approaches reflect the field's multifaceted effort to create reliable myoelectric control systems that maintain accuracy despite arm positional variations. Each approach offers unique advantages while presenting specific implementation challenges for practical myoelectric control applications. In this work, we exploit the advantages of comprehensive training protocols and investigate adaptation-free transferability of learned features through the lens of domain generalisation.

5.2.2 Domain Generalization

Domain generalization aims to create models inherently capable of generalizing to unseen domains without adaptation, eliminating the need for target domain data during training or frequent recalibration during deployment (Wang et al., 2022).

Recent work has applied DG in a cross-subject/user-generic framework. Li et al. (2023) demonstrated an integration of domain generalization with unsupervised domain adaptation (UDA), by leveraging mix-up training and adversarial strategies (Xu et al., 2020). Their proposed MAT-DGA method achieved 95.71±4.17% classification accuracy, significantly outperforming conventional approaches with 69.64%, 93.54%, and 91.75% performance accuracy for ConvNet, MAT-DA, and ADANN (Côté-Allard et al., 2020a), respectively. In a related study, a multi-source framework treating individual users as separate domains showed improved cross-user performance by aligning

each source user with the target user in individual feature spaces, achieving 87.93% accuracy with domain adaptation and 73.04% with domain generalization (Zhang et al., 2023b). Notably, Yang et al. (2024) established the EMGBench, a comprehensive benchmark for out-of-distribution generalization in myoelectric control, demonstrating intersubject classification and adaptation capabilities across nine diverse datasets.

In the context of the position effect in myoelectric control systems, domain generalization could offer distinct advantages by enabling models to maintain stable performance across different arm positions without requiring extensive recalibration. This capability would not only reduce the calibration burden on users but also facilitate more natural and seamless human-computer interaction in dynamic environments where arm posture frequently varies during daily activities. To the best of our knowledge, this work represents the first explicit identification of domain generalization as a promising framework to effectively mitigate the challenges associated with position-induced variability in myoelectric control.

5.3 Problem Formulation

In this section, we define domain generalization both informally and formally, and distinguish it from related paradigms such as domain adaptation. We then present a motivating example that illustrates distribution shifts in our data.

Informally, assume we have access to a fixed number of datasets, each sampled from a different distribution. In this context, each dataset corresponds to a separate domain. Domain generalization refers to a class of methods that aim to learn feature representations invariant across these diverse source domains. By extracting features that remain stable despite domain-specific variations, these methods attempt to capture underlying factors that are both predictive of the target label and transferable to unseen domains. More formally, we define the notions of domain and domain generalization as follows:

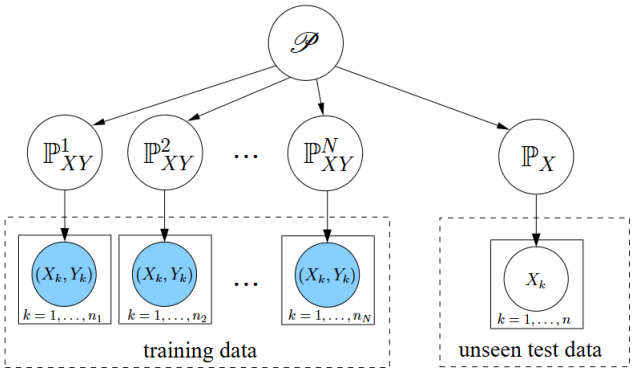


Figure 5.1 A simplified schematic diagram of the domain generalisation framework. A major difference between our framework and most previous work in domain adaptation is that we do not observe the test domains during training time. See text for a detailed description of how the data is generated. (Muandet et al., 2013)

Definition 1 (Domain)**.** A domain comprises a data collection derived from a distribution, defined as $S = \{(x_i, y_i)\}_{i=1}^{n} \sim P_{XY}$, where $x \in X \subset \mathbb{R}^d$ representing the input space, $y \in Y \subset \mathbb{R}$ representing the output space, and P_{XY} signifies the joint probability distribution between input samples and their corresponding output labels. The random variables associated with these spaces are denoted by X and Y.

Definition 2 (Domain Generalization)**.** In the domain generalization framework, as depicted in Fig. 2, we work with N source domains for training, represented as $S_{train} = \{S^i | i = 1, \ldots, N\}$ where $S^i = \{(x^i_j, y^i_j)\}_{j=1}^{n_i}$ denotes the i-th domain. The joint distributions across different domain pairs exhibit dissimilarity: $P^i_{XY} \neq P^j_{XY}$ for $1 \leq i \neq j \leq N$. The fundamental objective of domain generalization is to develop a robust and generalizable predictive function $h : X \to Y$ using these N training domains that minimises prediction error on an unseen test domain.

In Table 5.2 we demonstrate the key difference between related learning paradigms. The key distinction for domain generalization is that, unlike most other paradigms, it does not have access to target/test domain data during training, making it particularly challenging. The only other method without access to test data access is zero-shot learning. In our case, zero-shot learning assumes never seeing data on a given subject, which is outside the scope of this study.

Comparison between domain generalization and some related learning paradigms.

Learning paradigm	Training data	Test data	Condition	Test access
Multi-task learning	$\mathcal{S}^1, \cdots, \mathcal{S}^n$	$\mathcal{S}^1, \cdots, \mathcal{S}^n$	$\mathcal{Y}^i \neq \mathcal{Y}^j, 1 \leq i \neq j \leq n$	✓
Transfer learning	$\mathcal{S}^{src}, \mathcal{S}^{tar}$	$\mathcal{S}^{tar}$	$\mathcal{Y}^{src} \neq \mathcal{Y}^{tar}$	✓
Domain adaptation	$\mathcal{S}^{src}, \mathcal{S}^{tar}$	$\mathcal{S}^{tar}$	$P(\mathcal{X}^{src}) \neq P(\mathcal{X}^{tar})$	✓
Meta-learning	$\mathcal{S}^1, \cdots, \mathcal{S}^n$	$\mathcal{S}^{n+1}$	$\mathcal{Y}^i \neq \mathcal{Y}^j, 1 \leq i \neq j \leq n+1$	✓
Lifelong learning	$\mathcal{S}^1, \cdots, \mathcal{S}^n$	$\mathcal{S}^1, \cdots, \mathcal{S}^n$	$\mathcal{S}^i$ arrives sequentially	✓
Zero-shot learning	$\mathcal{S}^1, \cdots, \mathcal{S}^n$	$\mathcal{S}^{n+1}$	$\mathcal{Y}^{n+1} \neq \mathcal{Y}^i, 1 \leq i \leq n$	×
Domain generalization	$\mathcal{S}^1, \cdots, \mathcal{S}^n$	$\mathcal{S}^{n+1}$	$P(\mathcal{S}^i) \neq P(\mathcal{S}^j), 1 \leq i \neq j \leq n+1$	×

Figure 5.2 Comparison between domain generalisation and some related learning paradigms (Wang et al., 2022). Figure has been adapted from the cited sources and is distributed under a CC BY 4.0 International License (https://creativecommons.org/licenses/by/4.0/).

5.3.1 Motivating Example

The phenomenon of distribution shift across different limb positions manifests as changes in the marginal probability P(X), commonly referred to as covariate shift or domain shift. We operate under the assumption that P(Y|X) remains stable or exhibits smooth variations corresponding to changes in P(X). Based on the GREAT dataset (Szymaniak et al., 2024) (see Chapter 4), the Figure 5.3 demonstrates the t-SNE visualisation of distinct clustering patterns for different gesture classes across two limb positions, P_1 and P_7.

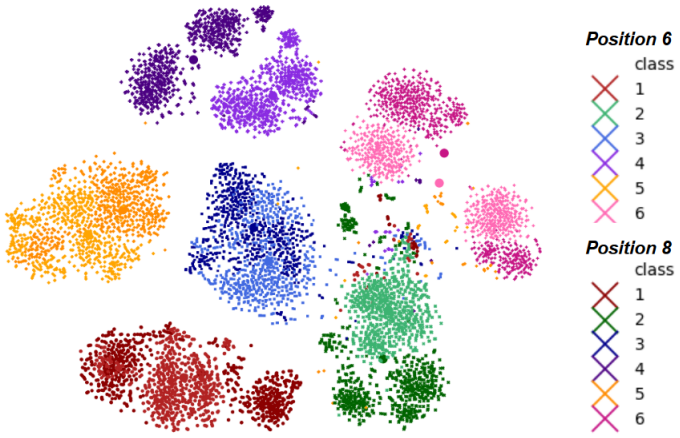


Figure 5.3 Example t-SNE plot depicting inter-position distribution shifts between P_6 and P_8 for 6 classes (gestures). Different shades of the same colour represent the same grasp with an emphasised centroid (bold dot).

Our analysis reveals that variations in limb positioning induce significant distribution shifts while preserving the underlying functional relationship between features and labels. The t-SNE visualisation demonstrates well-defined clusters and their centroids, confirming that these represent separate but related distributions. Each grasp category maintains its distinctive colour identity across both positions, with varying shades differentiating between positions.

5.4 Algorithms

In this section, we present the theoretical foundations of the algorithms selected for addressing the limb position effect as a domain shift problem in myoelectric control. We evaluate both traditional statistical methods and modern neural network-based approaches that have demonstrated success in various domain generalisation tasks. We show that all these algorithms make different assumptions about how invariant features are learnable from source domains.

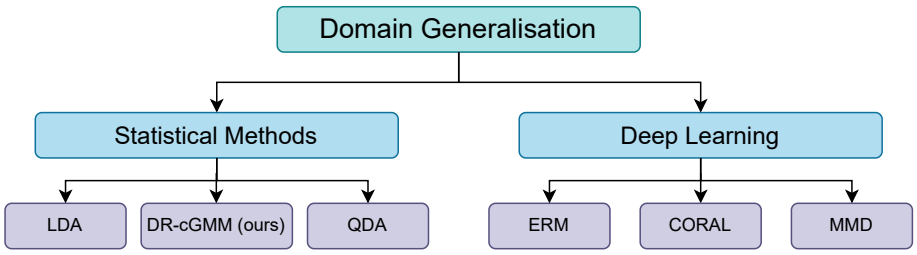


Figure 5.4 The taxonomy of selected methods in our comparative analysis for domain generalisation.

Linear Discriminant Analysis (LDA)

Linear Discriminant Analysis remains one of the most widely used algorithms in myoelectric pattern recognition due to its computational efficiency and robust performance (Fisher, 1936). LDA operates by projecting data onto a lower-dimensional space while maximising between-class separability and minimising within-class variability. In the context of domain generalisation, LDA's strength lies in its simplified covariance structure assumption. By using a shared covariance matrix across all classes, LDA

inherently regularises the solution space, potentially making it more robust to domain shifts. The decision boundary in LDA is given by:

$$\delta_k(x) = x^T \Sigma^{-1} \mu_k - \frac{1}{2} \mu_k^T \Sigma^{-1} \mu_k + \log(\pi_k) \tag{5.1}$$

where Σ is the pooled covariance matrix, μ_k is the mean of class k, and π_k is the prior probability of class k.

Domain Generalisation Perspective: LDA assumes that the invariant structure across all training domains lies in a low-dimensional linear subspace. Specifically, domain shifts affect only irrelevant or non-discriminative directions (i.e., orthogonal to the discriminative subspace).

Quadratic Discriminant Analysis (QDA)

Unlike LDA, QDA allows for different covariance matrices for each class, resulting in quadratic decision boundaries (Hastie et al., 2009). While this provides greater flexibility in modelling class distributions, it may be more susceptible to overfitting when faced with domain shifts. The decision function for QDA is:

$$\delta_k(x) = -\frac{1}{2} \log |\Sigma_k| - \frac{1}{2} (x - \mu_k)^T \Sigma_k^{-1} (x - \mu_k) + \log(\pi_k) \tag{5.2}$$

where Σ_k is the covariance matrix specific to class k.

Domain Generalisation Perspective: QDA assumes that class-specific covariances are invariant across domains. That is:

1. For all domains, the class-conditional distributions $P(Y|X)$ follow Gaussians with domain-invariant means μ_k and class-specific covariances Σ_k.

2. Domain shift is assumed to occur only in the directions that do not alter the intra-class covariance structure or class-mean relationships.

Empirical Risk Minimisation (ERM)

ERM serves as our baseline neural approach, where the model is trained to minimise the empirical risk across all source domains without any explicit domain adaptation mechanism (Gulrajani and Lopez-Paz, 2021). The objective function is:

$$\min_{\theta} \frac{1}{|\mathscr{D}|} \sum_{(x,y)\in\mathscr{D}} \mathscr{L}(f_\theta(x), y) \tag{5.3}$$

where $\mathscr{D}$ represents the combined data from all source domains, f_θ is the neural network with parameters θ, and $\mathscr{L}$ is the classification loss function. For our implementation, ERM is paired with a Convolutional Neural Network (CNN) architecture suitable for processing myoelectric signals.

Domain Generalisation Perspective: ERM assumes that minimising the average empirical risk over multiple source domains will result in a model that captures features useful across all domains. In doing so, ERM implicitly relies on the presence of domain-invariant predictive features and assumes that these dominate the learning process. While it lacks explicit mechanisms to separate invariant from domain-specific signals, the aggregation of diverse domains encourages the model to ignore domain-specific noise and focus on shared structure, assuming such structure exists.

Maximum Mean Discrepancy (MMD)

MMD is a DG method by design rather than having a corresponding DG perspective (Gulrajani and Lopez-Paz, 2021; Li et al., 2018). MMD minimises the distance between feature distributions from different domains in a Reproducing Kernel Hilbert Space (RKHS). Unlike adversarial approaches, MMD provides a closed-form solution for distribution matching. The MMD loss is defined as:

$$\mathrm{MMD}^2(\mathscr{D}_s, \mathscr{D}_t) = \left\| \frac{1}{n_s} \sum_{i=1}^{n_s} \phi(x_i^s) - \frac{1}{n_t} \sum_{j=1}^{n_t} \phi(x_j^t) \right\|_{\mathscr{H}}^2 \tag{5.4}$$

where ϕ maps inputs to a RKHS $\mathscr{H}$, and $\mathscr{D}_s$ and $\mathscr{D}_t$ represent source and target domains respectively. In practice, we use the kernel trick to compute MMD without explicitly defining ϕ.

Domain Generalisation Perspective: MMD assumes that domain shifts can be mitigated by aligning the marginal distributions of features across domains in a high-dimensional RKHS. The underlying belief is that if the feature representations of different domains become statistically similar, a classifier trained on one domain will generalise to others. This assumes the existence of a shared feature space in which class-discriminative information is preserved and domain-specific variations can be minimised through distribution matching.

CORrelation ALignment (CORAL)

CORAL aims to align the second-order statistics (covariance) of source and target feature distributions (Gulrajani and Lopez-Paz, 2020; Sun and Saenko, 2016). The loss function minimises the distance between covariance matrices:

$$\mathscr{L}_{\text{CORAL}} = \frac{1}{4d^2}\|C_s - C_t\|_F^2 \tag{5.5}$$

where C_s and C_t are the covariance matrices of source and target feature distributions, d is the feature dimension, and $\|\cdot\|_F$ denotes the Frobenius norm. CORAL offers a computationally efficient approach to aligning domain distributions without requiring adversarial training.

Domain Generalisation Perspective: CORAL assumes that domain shift arises primarily from differences in second-order statistics (i.e., feature covariances) between domains. By aligning covariance matrices in the feature space, CORAL encourages the model to learn representations where class structure is preserved and domain-specific variations are minimised. This method implicitly assumes that aligning covariances is sufficient to achieve domain invariance, and that class-discriminative information lies in the aligned subspace.

5.5 Domain-Regularised Class-Conditional Gaussian Mixture Model (DR-cGMM)

Our motivation for exploring class-conditioned Gaussian Mixture Models (cGMM) stems from their ability to model each domain (arm position) as a distinct Gaussian distribution. To clarify, we note that both Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) are special cases of cGMM. Specifically, LDA assumes each class follows a single Gaussian distribution with a shared covariance matrix across classes, while QDA allows for class-specific covariance matrices. In contrast, cGMM models each class as a mixture of Gaussian components, offering greater flexibility. This flexibility is especially beneficial for electromyographic (EMG) gesture recognition, where a single gesture class may be performed in different arm positions. By modelling each arm position (domain) as a separate Gaussian component within the same gesture (class), cGMM can better capture intra-class variability due to position changes, resulting in more robust classification, as shown in Liu et al. (2014b).

In the rest of this section, we refer to each mixture component as a 'domain' (arm position) to emphasise that each component represents a distinct position. We continue to use the term 'class' as in standard cGMM terminology to refer to the gesture.

5.5.1 Classical cGMM

Formally, the class-conditioned Gaussian Mixture Model (cGMM) models the likelihood of a data point x belonging to class k as a mixture of J_k Gaussian components:

$$p(x \mid y = k) = \sum_{j=1}^{J_k} \pi_{jk} \mathcal{N}(x \mid \mu_{jk}, \Sigma_{jk}) \tag{5.6}$$

where π_{jk} is the mixing coefficient of the j-th component in class k, satisfying $\sum_{j=1}^{J_k} \pi_{jk} = 1$, and $\mathcal{N}(x \mid \mu_{jk}, \Sigma_{jk})$ denotes a multivariate normal distribution with mean μ_{jk} and covariance matrix Σ_{jk} and J_k is the total number of mixture components at training time.

By applying regularisation on μ_{jk} and Σ_{jk}, cGMM provides a continuum of models between simple parametric classifiers like LDA/QDA and more expressive, multi-modal representations suitable for domain generalisation. Inspired by this approach's flexibility and adaptability, we believe that the regularisation concept can boost domain generalisation for the position effect in myoelectric control.

5.5.2 DR-cGMM

Our method is motivated by the empirical observation that different limb positions usually form smooth transitions in feature space rather than occupying distinct, isolated regions (see Figure 5.3). This reflects the underlying continuity of muscle activation patterns, suggesting that new, unseen positions for a given gesture are likely interpolations of previously observed ones rather than entirely new domains.

To capture this structure, we adopt a mixture model perspective in which each class (e.g., a movement intent) is represented by a set of components corresponding to different positions. However, treating each position as fully independent risks overfitting in low-data regimes, while enforcing full parameter sharing can underfit by ignoring position-specific nuances.

To address this, we introduce the Domain-regularised class-conditional Gaussian Mixture Model (DR-cGMM). This model includes three tunable regularisation parameters that control how statistical information is shared across global, class, and domain levels. These parameters act as interpolation weights, allowing the model to:

- Share information across similar positions to improve generalisation,
- Preserve distinct patterns where needed to avoid underfitting,
- Interpolate smoothly to handle novel or intermediate positions.

Our method, DR-cGMM, formalizes this intuition through three levels of regularization. These are global covariance regularization (α), class mean regularization (β), and domain covariance regularization (γ). Next we introduce each regularizer and it's motivation.

(1) Global covariance regularisation (α):

This parameter controls interpolation between global and class-specific covariance matrices:

$$\Sigma_k = \alpha\hat{\Sigma}_k + (1-\alpha)\Sigma_{\text{global}} \tag{5.7}$$

where $\hat{\Sigma}_k$ is the empirical class covariance for class k and Σ_{global} is the weighted average of all class covariances. Setting $\alpha = 0$ recovers a shared structure (as in LDA), while $\alpha = 1$ allows for full class specificity (as in QDA).

This mirrors shrinkage estimators in statistics, particularly Regularised Discriminant Analysis (RDA) (Friedman, 1989), which balances flexibility and generalisation by interpolating between LDA and QDA.

(2) Class mean regularisation (β):

This parameter interpolates between class-level and domain-specific component means:

$$\mu_{jk} = \beta\hat{\mu}_{jk} + (1-\beta)\mu_k \tag{5.8}$$

where $\hat{\mu}_{jk}$ is the mean for domain j within class k, and μ_k is the overall class mean. When $\beta = 0$, all domain means within a class collapse to the class mean; when $\beta = 1$, domain means are estimated independently from the class mean, i.e, one mean per domain.

This strategy reflects techniques in domain adaptation that align class-conditional distributions across domains. In particular, Maximum Mean Discrepancy (MMD) methods align domain means to a shared structure in domain adaptation (Long et al., 2015; Tzeng et al., 2014) and domain generalisation (Li et al., 2018). Our parametric approach achieves similar alignment by softly regularising domain means toward a global class prototype. The idea also aligns with multi-task learning (Evgeniou and Pontil, 2004) and hierarchical Bayesian modelling (Gelman et al., 2013), where task-specific estimates are regularised toward shared priors.

(3) Domain covariance regularisation (γ):

This is our **novel contribution**. It controls the degree to which domain-specific covariances deviate from their corresponding class-level covariances:

$$\Sigma_{jk} = \gamma\hat{\Sigma}_{jk} + (1-\gamma)\Sigma_k \tag{5.9}$$

where $\hat{\Sigma}_{jk}$ is the domain-specific covariance estimated from the data, and Σ_k is the class covariance (which itself is regularized by α). When $\gamma = 0$, all domains within a class share the same covariance; when $\gamma = 1$, each domain has a fully independent estimate.

This parameter is particularly important for generalisation across positions. Lower values of γ promote shared structure across positions, capturing invariant feature relationships within a class. Higher values allow for more position-specific flexibility. By interpolating between these extremes, the model learns stable patterns that generalise smoothly, even to unseen, intermediate positions.

Unlike α, which promotes global sharing of covariance structure across all classes, γ operates at a finer level, preserving distinctions within each class while allowing structured variation across positions. This makes DR-cGMM especially well-suited for tasks like gesture recognition in myoelectric control, where limb positions lie on a continuous manifold rather than forming discrete clusters.

Intuitively, this can be seen as a form of hierarchical regularisation over covariance structures: α controls sharing across classes, while γ operates one level below, regulating covariance sharing across domains within each class.

Together, the α, β, and γ parameters enable DR-cGMM to balance shared structure with necessary flexibility, supporting robust generalisation across the continuous landscape of limb configurations.

5.6 Study Design

Common inconsistencies in experimental conditions in the Domain Generalisation field lead to biased and unreliable outcomes. To troubleshoot those, we define and clarify datasets, architectures, hyperparameter optimisation (HPO) protocol, and other model selection criteria in detail.

5.6.1 Dataset and Preprocessing

The data collected and presented as part of the thesis in the Chapter 4 were used for analysis. The GREAT dataset consists of 8 subjects, and the analysis is conducted in an inter-subject manner. For a subject's dataset, the respective positions in cross configuration Con_+ are {2,4,5,6,8} and diagonal configuration Con_x are {1,3,5,7,9}. For each position, participants performed 10 trials of six distinct hand gestures: power grasp, lateral grasp, pointer gesture, tripod grasp, open hand, and rest position.

To comprehensively characterise the EMG signals, we extracted both time-domain and frequency-domain features. The extraction process employed a sliding window technique with a window length of 128 ms and a stride of 5 ms. Our feature set incorporates both temporal and frequency-domain characteristics, providing comprehensive multi-domain signal representation. We deliberately employed this standardised feature set uniformly across all methods to ensure fair comparison of generalisation capabilities independent of data preprocessing choices. In the process, we extracted the following features: mean absolute value, waveform length, log variance, zero crossing, slope sign changes, skewness, mean frequency, peak frequency, and variance of the central frequency. This preprocessing step yielded (16 channels x 9 features) features extracted from a single window.

Please note that the data was normalised dynamically during each training process. In each experimental run, normalisation was applied independently on the training data to prevent any leakage from the validation or test sets.

5.6.2 Domain Split

Within each configuration, we have 5 positions. We split the domains using leave-one-position-out cross-validation, giving us 4 training and 1 testing domain. We denote each training/testing domain split as a scenario. This structure arises from the fact that we treat each of the five domains as a distinct test domain. This setup naturally leads to a combinatorial arrangement of training and testing splits. By cycling through all possible held-out domains, this approach accounts for the variability in model performance that may arise due to the relative difficulty of each domain. For instance, if a held-out domain happens to be less (or more) challenging to generalise to, its influence is averaged out across all the scenarios. This leads to a more robust and unbiased estimate of generalisation performance. More details in Appendix B.2.2.

5.6.3 Hyperparameter Optimisation (HPO)

We perform hyperparameter optimisation (HPO) for every scenario using random search, as advocated in Bergstra and Bengio (2012), to sample a diverse set of hyperparameter configurations. Each algorithm is tuned over a predefined hyperparameter space (see Table 5.1). For each scenario, we run 20 trials sampled from the associated distribution.

Regarding Table 5.1, each algorithm features a distinct set of hyperparameters requiring optimisation. For traditional statistical models such as LDA and QDA, parameters are fixed. In contrast, methods such as DR-cGMM and neural network demand more extensive hyperparameter search, especially when parameters must be optimised per subject (denoted as *DS). Neural network methods, in particular, involve parameters that span several orders of magnitude (e.g., learning rates ranging from 10^{-6} to 10^{-2}).

5.6.4 Model Selection

Rigorous model selection is essential for DG methods, particularly during HPO (Gulrajani and Lopez-Paz, 2020). Given that each scenario includes only 4 training domains,

Condition	Parameter	Default value	Random distribution
DR-cGMM	α: global covariance	*DS	Uniform(0,1)
	β: class mean	*DS	Uniform(0,1)
	γ: domain covariance	*DS	Uniform(0,1)
LDA	α: global covariance	0	NaN
	β: class mean	0	NaN
	γ: domain covariance	0	NaN
QDA	α: global covariance	1	NaN
	β: class mean	0	NaN
	γ: domain covariance	0	NaN
ERM	weight decay	*DS	$10^{\text{Uniform}(-1,-4)}$
	learning rate	*DS	$10^{\text{Uniform}(-6,-2)}$
MMD	gamma	*DS	$10^{\text{Uniform}(-1,1)}$
	weight decay	*DS	$10^{\text{Uniform}(-1,-4)}$
	learning rate	*DS	$10^{\text{Uniform}(-6,-2)}$
CORAL	weight decay	*DS	$10^{\text{Uniform}(-1,-4)}$
	learning rate	*DS	$10^{\text{Uniform}(-6,-2)}$

Table 5.1 Hyperparameters, their default values and distributions for random search. *DS: defined per subject

we further split them using a 2-fold cross-validation to estimate hyperparameter performance. When combined with the held-out domain in the outer loop, this results in a nested cross-validation scheme.

Nested Cross-Validation addresses a critical challenge in machine learning model development: preventing data leakage during hyperparameter optimisation, model performance assessment and estimating its generalisation error. This two-level validation structure ensures that the final evaluation of model generalisation ability is completely independent from the hyperparameter tuning process.

With reference to Figure 5.5, note our nested cross-validation framework setup. The outer loop defines the K folds as one of the five positions is held out as the test set once, while the remaining $K-1$ positions are used for the inner cross-validation loop. Within this inner loop, we conduct 2-fold cross-validation, splitting the $K-1$ positions into training and validation sets to perform HPO.

Subsequently, during the hyperparameter optimisation phase inside each inner loop, we select the optimal hyperparameter across 20 HPO trials using the best validation metric (more details in Appendix B.2.2). The final model assessment is derived by combining the inner-loop domains to train a final model (with the selected hyperparameters) and is evaluated on the held-out domain.

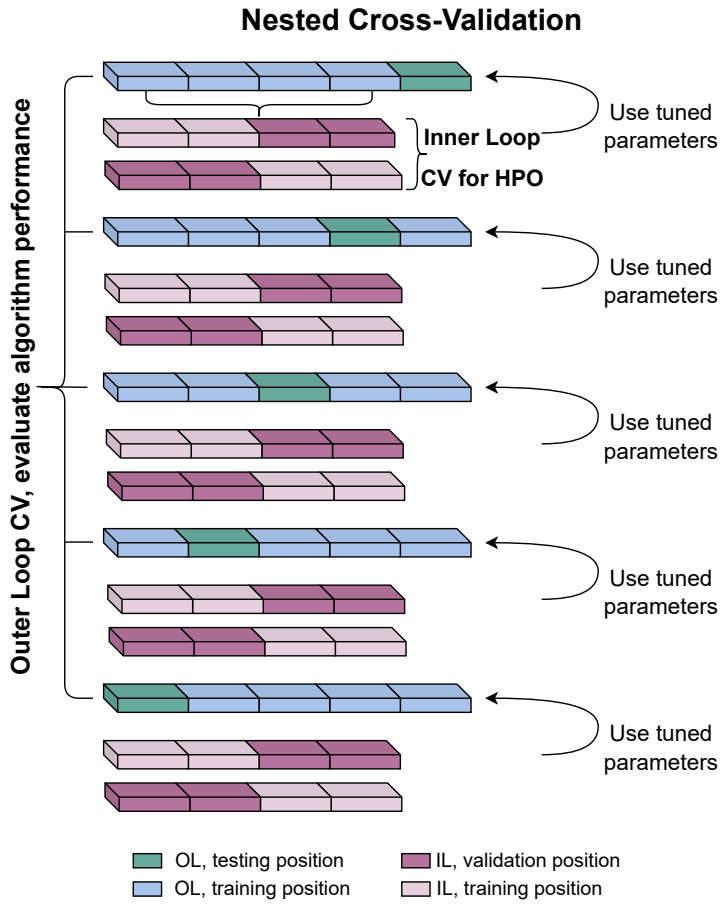


Figure 5.5 The figure illustrates our nested cross-validation framework, which implements a hierarchical validation structure with two distinct loops. The outer loop evaluates algorithm performance across different positions, while the inner loop partitions are utilised for hyperparameter optimisation (HPO) on each combination of training and validation partitions.
The colour-coded partitions correspond to the following: outer loop testing positions (OL testing position, green), outer loop training positions (OL training position, blue), inner loop validation positions (IL validation position, purple), and inner loop training positions (IL training position, pink). This visualisation demonstrates how data is systematically partitioned to ensure rigorous validation while preventing information leakage between optimisation and evaluation phases. The figure showcases our implementation of nested cross-validation across single configuration (Con_+ and Con_x).

Please note that by enumerating the positions as held-out sets, we end up with $K \times F \times C$ distinct scenarios, where K represents the number of positions (5), F denotes the number of folds in the inner loop (2), and C indicates the number of configurations (Con_+ and Con_x). This yields 20 comprehensive experiments for hyperparameter optimisation with a single set of hyperparameters (one trial).

For a single algorithm that required HPO, we conclude 3,280 experiments (details in Appendix B.2.2). To account for variability from random seeds, initialisations, data splits, and hyperparameter search, we repeat the entire study three times with independently sampled randomness. All reported results represent the average across the three independent runs, accompanied by their corresponding standard error. This experimental protocol involved training a total of 39,520 models.

5.6.5 Evaluation Metrics and Hypothesis Testing

As explained in the previous section, each algorithm trained on a subject's dataset produces 10 generalisation estimates, stemming from using each of 10 positions as a held-out set once. To get a final generalisation measurement for a method on a subject's dataset, we average the 10 held-out set measurements.

We note that this experimental design yields performances of 6 algorithms over 8 datasets (subjects). Comparing the final metrics on their own can be misleading since we do not conclude with statistical significance without hypothesis testing. Nevertheless, running many hypothesis tests using the same experimental data increases the Type-I Error on any given test, and we must correct for the family-wise error rate (Benjamini and Braun, 2002).

We first run the Friedman omnibus test (Demšar, 2006; Friedman, 1937) to assess whether there are any statistically significant differences among the algorithms across the 8 datasets. The Friedman test is a non-parametric alternative to repeated-measures ANOVA and is suitable in our context because we cannot assume normality of performance differences.

If the Friedman test detects a significant difference (i.e., the null hypothesis that all algorithms perform equally is rejected), we proceed with post-hoc pairwise compar-

isons to identify which algorithms differ significantly from each other. For this, we use the Nemenyi test (Demšar, 2006), which controls the family-wise error rate and compares average ranks of algorithms. Two algorithms are significantly different if the difference in their average ranks exceeds the critical difference (CD), which depends on the number of algorithms, the number of datasets, and the desired confidence level.

All statistical tests are conducted at a significance level of $\alpha = 0.05$, and the CD is computed accordingly. We provide formal definitions in Appendix B.2.3.

5.6.6 Other Implementation Choices

As the objective of this work is to investigate domain generalisation for position effect in gesture recognition, we control for architecture design and optimiser choice as factors of variation. These two factors were not considered for HPO.

The architecture blocks are depicted in Section B.2.1 with the featurizer and classifier blocks being described therein. All models were trained for 10 epochs, with AdamW optimiser (Loshchilov and Hutter, 2017) and cross-entropy loss (Cox, 1958; Mao et al., 2023). LDA and QDA have closed-form solutions.

5.7 Results

We evaluated the performance of six different approaches for domain generalization tasks: LDA, DR-cGMM, QDA, ERM, MMD, and CORAL. Before presenting our formal hypothesis tests, we first offer a descriptive overview of the raw performance and variability of all six methods across the eight domain generalization scenarios. We explore family-level behaviour, comparing the performance of deep learning approaches (ERM, MMD, CORAL) with that of classical statistical methods (LDA, QDA, DR-cGMM). We also analyse per-subject variability and cross-scenario stability. Finally, having established these empirical patterns, we proceed in Section 5.7.2 to the Friedman omnibus test with Nemenyi post-hoc comparisons to identify statistically significant pairwise performance differences.

5.7.1 Descriptive Analysis

Table 5.2 presents the performance results across 10 domain generalisation scenarios for these six algorithms. The reported results represent the average performance and standard deviation for each scenario. For each subject, we calculate the average performance across 3 repeated nested cross-validation runs, and then average the results across all subjects. The reported standard deviation reflects the variability across subjects for each method in a given scenario. Both the raw performance metrics and the average ranks for each method are provided in the supplementary material (Appendix B).

Scenario	LDA	DR-cGMM	QDA	ERM	MMD	CORAL	AVG
$P_{3,5,7,9} \rightarrow P_1$	92.93 ± 0.03	93.29 ± 0.042	92.35 ± 0.06	93.37 ± 0.036	93.38 ± 0.038	93.28 ± 0.035	93.10 ± 0.004
$P_{1,5,7,9} \rightarrow P_3$	94.34 ± 0.03	94.52 ± 0.027	92.67 ± 0.04	92.36 ± 0.047	92.39 ± 0.049	92.39 ± 0.052	93.11 ± 0.009
$P_{1,3,7,9} \rightarrow P_5$	94.89 ± 0.02	95.52 ± 0.020	95.52 ± 0.02	96.01 ± 0.018	95.72 ± 0.019	95.75 ± 0.023	95.57 ± 0.003
$P_{1,3,5,9} \rightarrow P_7$	93.14 ± 0.03	93.48 ± 0.037	91.38 ± 0.06	92.68 ± 0.045	93.29 ± 0.041	92.97 ± 0.038	92.82 ± 0.007
$P_{1,3,5,7} \rightarrow P_9$	94.47 ± 0.03	94.51 ± 0.032	93.28 ± 0.04	94.07 ± 0.039	94.03 ± 0.039	94.08 ± 0.042	94.07 ± 0.004
$P_{4,5,6,8} \rightarrow P_2$	93.19 ± 0.05	93.43 ± 0.053	92.37 ± 0.06	90.82 ± 0.054	90.60 ± 0.055	91.53 ± 0.049	91.99 ± 0.011
$P_{2,5,6,8} \rightarrow P_4$	93.01 ± 0.05	93.26 ± 0.049	93.00 ± 0.04	92.47 ± 0.040	92.39 ± 0.044	92.59 ± 0.037	92.79 ± 0.003
$P_{2,4,6,8} \rightarrow P_5$	95.38 ± 0.02	95.37 ± 0.023	94.76 ± 0.03	95.23 ± 0.024	95.63 ± 0.021	95.58 ± 0.021	95.32 ± 0.003
$P_{2,4,5,8} \rightarrow P_6$	93.81 ± 0.03	94.54 ± 0.031	92.31 ± 0.07	93.59 ± 0.033	92.84 ± 0.030	93.47 ± 0.033	93.43 ± 0.007
$P_{2,4,5,6} \rightarrow P_8$	94.59 ± 0.02	94.95 ± 0.019	92.81 ± 0.05	93.74 ± 0.023	93.81 ± 0.023	93.49 ± 0.024	93.90 ± 0.007
AVG	93.98 ± 0.03	94.29 ± 0.008	93.05 ± 0.05	93.43 ± 0.014	93.41 ± 0.015	93.51 ± 0.013	

Table 5.2 Domain generalisation across 10 scenarios. The aggregated results correspond to classification accuracy across all subjects.

Deep Learning Methods

Our deep learning baseline method, ERM with 1DCNN architecture, achieved competitive performance with an average accuracy of $93.43\% \pm 0.014$ across all domain generalisation scenarios. ERM directly minimises the training error on the source domains without any explicit domain adaptation mechanisms, yet it demonstrates robust generalisation capabilities.

When examining specialised domain generalisation methods, we found MMD and CORAL exhibited comparable effectiveness to the ERM baseline. MMD achieved an average accuracy of $93.41\% \pm 0.015$ while CORAL performed slightly better at

93.51% ± 0.013. Interestingly, while these specialised domain generalisation approaches, MMD and CORAL, are designed to better handle distribution shifts, they only marginally outperformed the simple ERM baseline in our experiments.

Statistical Methods

DR-cGMM achieved the highest average accuracy of 94.29% ± 0.008 across all scenarios, demonstrating superior performance compared to LDA and QDA with 93.98% ± 0.03 and 93.05% ± 0.05 accuracy, respectively. Our proposed method outperformed LDA in 9 out of 10 scenarios and consistently surpassed QDA by a significant margin in most cases. Notably, DR-cGMM showed particular strength in challenging generalisation scenarios, achieving 94.54% accuracy compared to 93.81% for LDA and 92.31% for QDA in the $P_{2,4,5,8} \rightarrow P_6$ transition, and 93.48% accuracy compared to 93.14% for LDA and 91.38% for QDA in the $P_{1,3,5,9} \rightarrow P_7$ transition.

The lowest standard deviation across all methods (±0.008) further indicates DR-cGMM consistent performance across different position transitions. These results validate our hypothesis that the three-tiered regularisation approach effectively captures the interpolative nature of arm positions within the same gesture class, allowing for better generalisation to unseen positions compared to traditional discriminant analysis methods that constrain (LDA) or often overfit (QDA) the covariance structure.

Subject Variability

Our experimental results demonstrate considerable inter-subject variability across all DG scenarios, as illustrated in Figure 5.6. This figure presents a quantitative analysis of classification accuracy across eight subjects for the investigated algorithms on 10 domain generalisation tasks. We depict the distribution of accuracy scores for each subject. Details are given in the Appendix B.1.

Subject-specific analysis reveals variability with some individuals, such as S_1, S_2, and S_6 showing overall higher accuracy and tighter distributions across all algorithms. In contrast, subjects S_4, S_7, and S_8 generally presented more challenging cases for classification with greater spread and lower median performance.

DR-cGMM algorithm consistently achieves high and stable classification accuracies across all subjects, with median values generally close to or above 95% and narrow interquartile ranges, indicating robust performance and low variability. In contrast, QDA combined with CORAL and MMD shows greater variability across subjects, particularly for S_4, S_7, S_8, and S_3 (for QDA), where the spread is wider and more outliers are present. This pattern was consistent across all the evaluated approaches, suggesting intrinsic differences in the discriminability of the data from different subjects.

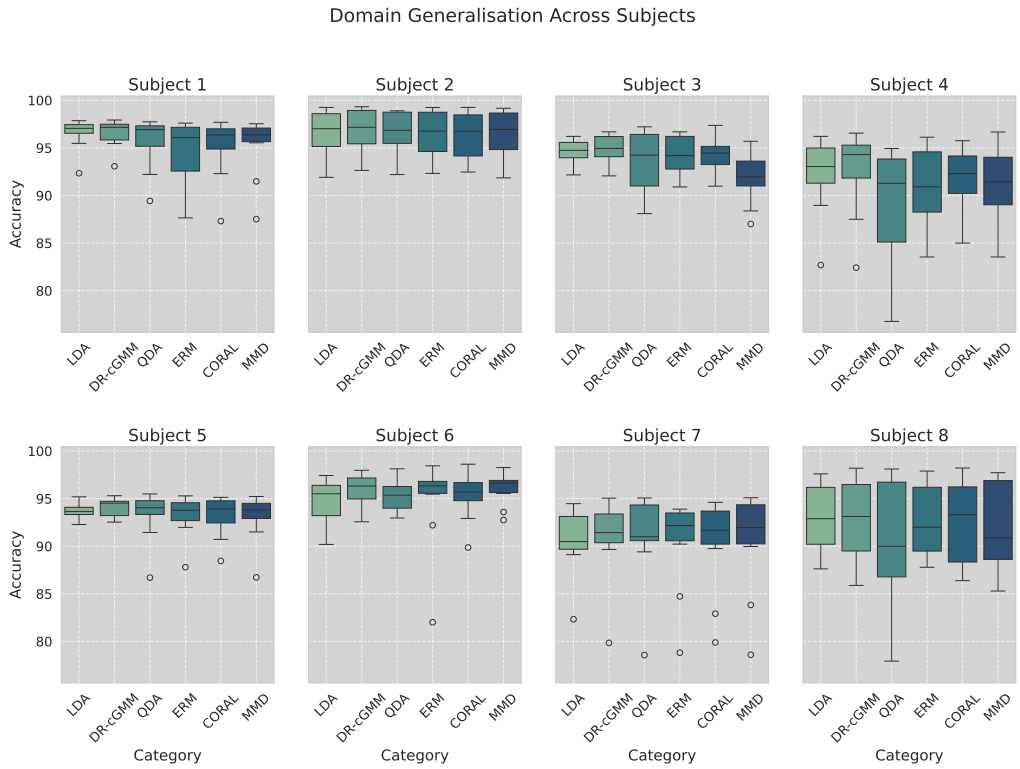


Figure 5.6 Subject-wise domain generalisation variability. Subplots depict classification accuracy distributions for investigated algorithms across 10 DG tasks.

Scenario Generalisation

In our investigation of scenario generalisation in the context of position effect, we observe interesting variability in performance across different held-out scenarios, as detailed in Table 5.2 and the Appendix B.1. While all algorithms exhibit some degree of performance fluctuation, a consistent pattern emerges across the DG scenarios. Specif-

ically, generalisation to the central position P_5 yields the highest average accuracy across all methods, with average performances of $95.57\% \pm 0.003$ and $95.32\% \pm 0.003$ in the Con_x and Con_+, respectively.

Conversely, the P_2 scenario emerged as the most challenging, with MMD exhibiting the weakest performance at $90.60\% \pm 0.055$, while LDA and DR-cGMM maintained more robust performance, achieving $93.19\% \pm 0.05$ and 93.43 ± 0.053, respectively. Within the Con_x setting, generalization to P_7 proved to be the most difficult, yielding an average performance of $92.82\% \pm 0.004$.

5.7.2 Statistical Testing

For statistical testing, we need to look at the performance of each method and every subject. As previously explained, each subject's final performance for a given method is calculated as the average over their 10 different scenarios, i.e, using each position as a held-out set. For each subject, we rank the methods from best (1) to worst (6). Then, we calculate the average rank of each method across all subjects. These results are reported in Table 5.3.

Subject	LDA	DR-cGMM	QDA	ERM	CORAL	MMD
S_1	96.52 (1)	96.47 (2)	95.60 (3)	94.61 (6)	94.88 (5)	95.17 (4)
S_2	96.59 (2)	96.83 (1)	96.52 (3)	96.49 (4)	96.37 (6)	96.44 (5)
S_3	94.53 (3)	94.84 (1)	93.62 (5)	94.49 (4)	94.66 (2)	93.60 (6)
S_4	92.24 (2)	92.50 (1)	88.70 (6)	91.17 (5)	91.33 (3)	91.18 (4)
S_5	93.69 (2)	94.01 (1)	93.29 (3)	93.17 (4)	92.99 (5)	92.98 (6)
S_6	94.62 (6)	95.88 (1)	95.34 (3)	95.03 (5)	95.24 (4)	95.50 (2)
S_7	90.65 (3)	90.99 (1)	90.95 (2)	90.12 (6)	90.32 (4)	90.18 (5)
S_8	92.97 (1)	92.77 (2)	90.36 (6)	92.40 (3)	92.29 (4)	92.21 (5)
Avg Ranks	2.5	1.25	3.875	4.625	4.125	4.625

Table 5.3 Statistical ranking of the investigated methods across the datasets (subjects).

Using the average ranks reported in Table 5.3, we conducted the omnibus Friedman test at a significance level of 0.05. The Friedman test yielded a test statistic of 20.85. Under the null hypothesis, this corresponds to a p-value of 0.0009. Therefore, we

reject the null hypothesis that all considered algorithms have equivalent performances and proceed with post-hoc pairwise comparisons.

To identify which methods differ, we proceeded to post-hoc pairwise comparisons using the Nemenyi test. The results are visualised in the critical difference diagram shown in Figure 5.7, where methods not significantly different from each other at the 0.05 level are connected by a horizontal line. This provides a clear summary of which methods perform similarly and which show statistically significant differences in average rank. In other words, for connected methods, the differences in their average rank are not statistically significant.

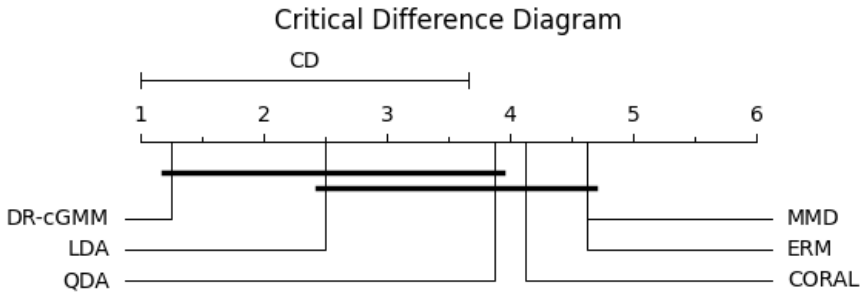


Figure 5.7 Critical Difference (CD) diagram depicting average ranks across all methods. Bolded horizontal bars highlight groups of algorithms that are not significantly different from each other.

We first note that no deep learning domain generalization method significantly outperforms ERM. While the average ranks vary slightly across these methods, the differences are not statistically significant. Similarly, although LDA and QDA achieve higher average ranks than the deep learning methods, these differences are also not statistically significant. Therefore, we conclude that neither LDA nor QDA significantly outperforms deep learning methods.

Additionally, our post-hoc analysis shows that all statistical methods are statistically indistinguishable from each other at the 0.05 significance level. It is important to emphasise that (1) all statistical methods being mutually equivalent, and (2) only LDA and QDA are equivalent to deep learning methods, which are not logically contradictory, since statistical significance is not a transitive property. Based on this

analysis, the only method that significantly outperforms deep learning methods is our proposed DR-cGMM.

5.8 Discussion

This work presents a comprehensive study of *within-subject cross-position generalisation* for myoelectric gesture recognition, addressing a critical challenge of mitigating confounding factors in myoelectric control. The position effect, a well-documented confounding factor in EMG-based gesture classification, often impedes generalisation due to the high variability introduced by changes in limb posture. By assuming no access to data from unseen positions during deployment, we frame this task as a *domain generalisation (DG)* problem, offering a more realistic and user-centric approach than traditional adaptation or fine-tuning strategies.

Our empirical findings underscore several important conclusions. First, the Friedman test confirmed significant performance differences across methods ($p = 0.0009$), and the Nemenyi post-hoc test further identified DR-cGMM as the only method significantly better than deep learning approaches. DR-cGMM achieved the highest overall average accuracy of $94.29\% \pm 0.008$ across all DG tasks.

Interestingly, LDA and QDA did not exhibit statistical superiority to deep learning methods, reinforcing the observation that statistical methods retain competitiveness in low-data regime but may not sufficiently capture position-related variability without additional modelling, as introduced in DR-cGMM.

Moreover, deep learning-based DG approaches (MMD, CORAL), while widely adopted for domain shift scenarios, offer only marginal improvements over the baseline ERM. All three methods achieved, on average across all the DG tasks, closely clustered accuracy values of 93.41%–93.51% and showed no statistically significant differences in performance. This suggests that in *static-limb positioning, low-data, subject-specific* regimes, the complexity and assumptions underlying deep DG techniques may not yield tangible gains over simpler empirical risk minimisation.

Our analysis of subject-level variability further revealed key insights. DR-cGMM maintained stable, high accuracy across all users, with low interquartile ranges and few outliers, highlighting its robustness to individual subjects' differences in EMG signal quality or muscle activation patterns. Conversely, deep learning methods, especially in conjunction with CORAL and MMD, show greater variance and susceptibility to outlier performance in difficult subjects such as S_4, S_7, and S_8. This robustness across individuals makes DR-cGMM well-suited for real-world use, where ensuring a stable and reliable user experience is important.

From a scenario generalisation perspective, our findings confirm that generalisation difficulty varies substantially with the held-out position, with central positions (P_5) proving easier to predict than extreme postures (like P_2, P_7). This pattern reinforces the importance of *position-aware modelling*, especially under limited training data, and further illustrates the strengths of DR-cGMM in adapting to positional interpolation.

5.8.1 Limitations and Future Work

While our study demonstrates strong performance and highlights the effectiveness of domain-regularised statistical modelling for low-data, static limb position scenarios, there remain several promising directions for future work.

Our study focuses exclusively on able-bodied participants, a common initial approach in myoelectric control research that enables systematic characterisation of confounding factors with larger sample sizes (in comparison to people with limb loss) (Atzori et al., 2014b; Khushaba et al., 2014). While shared physiological mechanisms underlying the position effect suggest our domain generalisation methods provide a valid foundation for amputee applications (Fougner et al., 2011; Jiang et al., 2013a), important differences exist: individuals with limb loss experience attenuated position effects due to shorter residual muscles and altered biomechanics, alongside unique challenges from prosthesis weight and socket compression (Cipriani et al., 2011; Jiang et al., 2013a; Teh and Hargrove, 2020). Future work should validate our DR-cGMM approach with amputee participants and investigate whether the domain regularisation

framework requires population-specific adaptations to account for these biomechanical differences in real-world prosthetic control scenarios.

A key missing analysis in this investigation is an ablation study of our proposed DR-cGMM. We introduced three different regularizers and demonstrated that the method outperforms deep learning approaches. To better understand the source of these performance gains, we would need to evaluate all combinations of the regularizers (both individual and pairwise) by running the full training and evaluation protocol. Afterwards, we could conduct Friedman and Nemenyi post-hoc tests to determine whether enabling all regularizers outperforms using only a subset of them.

Regarding the external validity of our hypothesis testing, we must clarify that the current study was conducted under a low-data regime, which is the de facto standard for these applications, as large datasets are often not available. This low-data setting may inherently favour simpler statistical models like DR-cGMM over more data-intensive deep learning methods. Expanding the dataset or incorporating data augmentation techniques could allow for a more thorough evaluation of deep models' ability to learn robust, position-invariant representations. However, as the dataset grows, our conclusions may not hold when scaling up these methods. With more data, deep learning models might (or might not) show improvements. Therefore, we emphasise that our conclusions are specific to the architecture and dataset used in this study.

Moreover, our experiments focused on controlled, static-limb positions in a laboratory setting. While this provided a well-defined environment to isolate the position effect, future work will aim to extend this framework to dynamic, real-world movements and continuous transitions, bringing the evaluation closer to practical prosthetic use.

Finally, due to the computational cost of nested cross-validation, we explored only a single variant per model with limited hyperparameter tuning. As a next step, we plan to explore knowledge transfer techniques, such as pretraining on related EMG datasets or multi-task learning frameworks, to enhance generalisation performance while maintaining low-data applicability.

By addressing these areas, we aim to further strengthen the generalisability and practicality of myoelectric control systems by mitigating the confounding factor, position effect, in real-world assistive technologies.

5.8.2 Summary

All things considered, our study highlights a key message: *in low-data regime, user-specific myoelectric control scenarios for static-limb positioning*, principled statistical modelling, when augmented with domain-aware regularisation, can outperform complex deep learning methods in both accuracy and consistency. DR-cGMM provides a promising path forward for building *generalisation-capable, calibration-free EMG systems*, paving the way for practical deployment in daily-use assistive and prosthetic technologies.

The results of our study provide valuable insights into the challenges and strategies for addressing within-subject, cross-position generalisation in myoelectric gesture recognition. Our findings highlight several key points that warrant further discussion.

Our novel contribution, the Domain Regularised class-conditional Gaussian Mixture Model (DR-cGMM), emerged as the top-performing method in our study. It consistently outperformed deep learning approaches across most domain generalisation scenarios. The superior performance of DR-cGMM can be attributed to its ability to effectively integrate domain knowledge into the model via regularisation, thereby enhancing its capacity to generalise across different arm positions without explicit recalibration.

5.9 Conclusion

This work provides a comprehensive and rigorous investigation into the challenge of *within-subject cross-position generalisation* in myoelectric gesture recognition, reframing it through the lens of domain generalisation. Our contributions span from a carefully designed benchmarking protocol with nested cross-validation to a unified analysis of statistical and deep learning models under the DG framework. The

proposed DR-cGMM model, which introduces domain-aware regularisation into a class-conditional Gaussian Mixture Model, emerged as the most effective approach, significantly outperforming both classical statistical and deep learning-based DG methods.

Chapter 6

Discussion and Future work

In this chapter, we summarise our contributions and provide recommendations for further exploration. We list some interesting directions for future work.

6.1 Summary of contributions

In Chapter 1 and Chapter 2, we highlighted the core challenges for deployment of upper limb prosthetics and the importance of mitigating confounding factors. Here, we review the contributions in the light of the challenges discussed.

1. **Active learning for long-term adaptation in myoelectric control.** In Chapter 3 (Szymaniak et al., 2022), we developed a framework for long-term adaptation of myoelectric control systems for upper limb prosthetics through active learning with simulated human-in-the-loop. The novel contributions of this research can be summarised as follows:

    - We introduced for the first time an active learning framework for myoelectric control adaptation, addressing performance deterioration over time. This approach enables the development of personalized prosthetic control system through *intelligent, targeted data acquisition* that significantly reduces the burden on users during the adaptation process.
    - Our systematic evaluation of uncertainty-based sampling strategies demonstrated that carefully selected training data can substantially improve de-

coder performance while minimising data collection requirements. Specifically, we established that with only 3.2 minutes of strategically selected training data, our approach achieved performance improvements of 4-5% for able-bodied participants and approximately 2% for individuals with limb differences compared to random sampling methods.

- The research identified the smallest margin and least confidence sampling as optimal uncertainty quantification methods for myoelectric signal classification, providing valuable guidance for future implementations in this domain.
- Finally, we proposed a modular framework for real-time HITL implementation of our adaptive approach that accommodates different decoder models, user profiles, and query strategies, creating a flexible foundation for future real-time implementations. This design enables balanced exploration and exploitation for optimal prosthetic control dexterity while maintaining adaptability to various clinical and research contexts.

This work represents a significant advancement in adaptive myoelectric control systems, with direct implications for improving the usability and functionality of upper limb prosthetics in real-world applications.

2. **Arm Translation Dataset, GREAT.** In Chapter 4 (Szymaniak et al., 2025), we addressed a critical challenge of arm translation in myoelectric control systems by investigating arm position effects on signal characteristics and gesture classification accuracy, with significant implications for standardised recording protocols. Our research yielded several important contributions to the field:

   - We introduced the GREAT dataset, a comprehensive collection of surface electromyographic signals recorded across multiple spatial positions in a 3x3 grid manner.
   - Our research proposed an innovative data acquisition protocol that systematically captures EMG signal variability across diverse arm positions.

This methodological advancement extends beyond traditional single-plane movement protocols to encompass the full Cartesian space (across horizontal and vertical planes), significantly enhancing the viability of collected EMG data and simultaneously questioning the standardised approaches with single-position recordings for future research efforts.

- Through rigorous analysis, we demonstrated a significant performance degradation of approximately 10% during inter-position classification tasks, quantifying the critical influence of arm positioning on sEMG signal characteristics. We confirmed the presence of position-specific information by successfully decoding arm position from sEMG data with approximately 80% accuracy for individual grasps, providing empirical evidence of this confounding factor.
- While our hierarchical approach leveraged position knowledge to enhance gesture classification, we identified persistent challenges in developing truly position-invariant systems. Despite implementing data augmentation strategies and position information integration, our position encoder achieved only 58% accuracy when conditioned on position rather than grasp type. Notably, when position correctness was disregarded, accuracy increased substantially to over 92%.

This dataset represents a valuable resource for researchers developing position-invariant and position-induced myoelectric control algorithms and establishes a benchmark for evaluating algorithmic performance under position variability conditions. The significance of this work extends beyond assistive technologies for individuals with limb differences to include extended reality applications and other domains requiring robust gesture recognition across varying spatial contexts.

3. **Domain Generalisation for Position Effect in Myoelectric Control:** We addressed distribution shifts caused by arm translation in EMG myoelectric

control by domain generalisation in gesture recognition. Our study concluded with the following contributions:

- Benchmark design with nested cross-validation: We implement nested cross-validation to ensure fair and unbiased model comparison, isolating hyperparameter optimisation from performance evaluation. This establishes a clear benchmark for comparing different algorithms in domain generalisation using our GREAT dataset.
- Evaluation of statistical and deep learning models under domain generalisation: We examine classical statistical methods (LDA, QDA) and deep learning DG algorithms within a unified framework, revealing that ERM fails to outperform simpler baselines and DG methods do not offer clear advantages in low-data settings over ERM.
- Domain-Regularised class-conditional Gaussian Mixture Model (DR-cGMM): We introduce DR-cGMM, integrating hierarchical regularisation over classes and domains into classical cGMM. DR-cGMM is the only method that significantly outperforms all deep learning models, demonstrating the continued relevance of statistical approaches in low-resource myoelectric control.

This work presents important insights into addressing challenges in myoelectric gesture recognition across different arm positions. Our findings emphasise that in a low-data regime, within-subject myoelectric control can benefit more from statistical approaches, surpassing deep learning techniques. By incorporating domain-aware regularisation, DR-cGMM delivers both reliable and accurate performance, making it a strong candidate for real-world EMG systems that require minimal calibration, an important step toward practical use in assistive and prosthetic devices.

6.2 Future work

Our research has revealed several promising avenues for further investigation that build upon our current findings. We have identified specific areas that warrant deeper exploration to provide robust and reliable myoelectric control under realistic environmental conditions.

1. **Real-Time Closed Loop Active Learning framework with HITL**. Building upon our demonstrated active learning framework for myoelectric control, we propose several directions for future research to enable the real-time closed-loop system with human-in-the-loop (HITL) capabilities. These advancements aim to create a truly co-adaptive framework that balances system performance with user experience over extended periods.

    - **Human as an Oracle and Optimised User Interaction Paradigms.** Our work has simulated the human-in-the-loop setting to investigate the feasibility of our approach. Next steps would include incorporating the user as an oracle to enable the model to truly adapt to the user. While HITL offers personalised adaptation, the interaction design must balance information quality with potential user burden. We will explore different labelling strategies, from simple binary responses (as presented in Section 3.4.1) to implicit feedback through task performance metrics (Madduri et al., 2022; Osborn et al., 2021). Additionally, we will aim to develop context-aware query timing strategies that minimise interruption during critical control tasks. Drawing inspiration from HCI and context-informed learning (Abowd et al., 1999; Campbell et al., 2024; Hong et al., 2009), *context is 'any information that can be used to characterise a situation or entity'*. Information such as past history, user location can be utilised to optimise the query sampling process (Kulkarni and Rodd, 2020).

    - **Long-Term Co-adaptation Studies.** Our preliminary work has demonstrated the efficacy of active learning for myoelectric control in an offline setting (as shown in Chapter 3, (Szymaniak et al., 2022)). However, we

must extend the deployment of the studies over an extended period to validate real-world applicability. The proposed active learning framework for real-time study would be implemented and deployed. This would evaluate how the system adapts to gradual physiological changes, muscle fatigue and alterations in muscle activation patterns as the user adapts over time. Following Markovic et al. (2014), we would develop metrics that quantify both algorithmic adaptation and user learning to provide comprehensive measures of co-adaptation success in everyday use scenarios.

- **Enhanced Biofeedback Integration.** True co-adaptation requires bidirectional information flow between user and system (Hu et al., 2023). While our current work has focused primarily on algorithm adaptation, we recognise the critical importance of user feedback pathways in myoelectric control systems (He et al., 2015a). We aim to move beyond traditional open-loop calibration approaches by implementing closed-loop adaptive calibration mechanisms that allow users to modify their muscle contractions in response to system performance. Our future research will investigate multiple feedback modalities, including vibrotactile somatosensory feedback that users consistently express a preference for in prosthetic applications (Peerdeman et al., 2011). We will explore visual feedback approaches that communicate EMG space similarity and clustering information to help users develop improved muscle synergy patterns based on work from (Barradas et al., 2023; Fang et al., 2016). Additionally, we plan to incorporate gamification and virtual reality elements that have shown promising results in accelerating skill acquisition while maintaining user engagement (Powell et al., 2013).

- **Algorithm Flexibility and Computational Efficiency.** To enable practical deployment while allowing for research flexibility, we will develop a modular, plug-and-play frameworks that facilitate comparison of active learning algorithms, thereby balancing computational efficiency with classification

performance. This modular approach will facilitate systematic evaluation of different algorithmic components in real-world settings.

- **Risk Mitigation and Limitations.** Several challenges must be addressed for successful clinical translation. The system may develop biases toward specific user movement patterns, limiting adaptability to new contexts or tasks not encountered during training, hence why we will investigate the exploration vs. exploitation dilemma between query strategies. Real-time processing requirements for closed-loop systems could exceed the computational capabilities of wearable prosthetics, creating latency that irritates users. We will develop approaches to manage user frustration during adaptation periods and address privacy concerns regarding continual data collection, creating responsible deployment protocols that prioritise user autonomy while maximising system performance.

2. **Re-evaluate Research Pipeline to Realistic Environmental Conditions.** Consecutively, in Chapter 4 and Chapter 5, we delved into the issue of position effect and its effect on myoelectric control and struggles in daily activities. Our experience with the GREAT dataset has highlighted the following: a) the need for re-evaluating the typical setup for data acquisition and emphasised the need to redefine standardised validation protocols for sEMG datasets, b) the critical importance of position variability on system feasibility for real-world applications. Based on this work, we raise two main goals:

   - **Realistic and Standardised Protocols.** Current myoelectric research frequently employs simplified protocols that vary in their structure and fail to capture real-world complexity. We propose developing comprehensive standardised frameworks that incorporate multiple confounding factors encountered in daily living scenarios. Specifically, these protocols should include position variations during data acquisition to better reflect everyday use cases. Standardisation would significantly enhance cross-study com-

parability and research reproducibility, allowing the field to make more coordinated progress through effective validation and system evaluation.

- **Position Effect: from Static to Dynamic Data.** While existing research has primarily focused on static positioning, real-world applications require understanding muscle activation patterns during continuous movement and reach-to-grasp actions (Batzianoulis et al., 2018). Future studies should implement dynamic protocols using two- and three-dimensional protocols in a 3×3 and $3 \times 3 \times 3$ grid, respectively, that better represent natural joint kinematics across the shoulder, elbow, and wrist. This approach will enable characterisation of muscle activation patterns during continuous movement transitions, facilitating robust development of myoelectric control systems under varied conditions.

3. **Position Domain Generalisation for Myoelectric Control.** Finally, in Chapter 5 we investigate domain generalisation approaches for limb position variations and propose DR-cGMM approach for position-invariant myoelectric gesture recognition in low-data regimes. Building on this work, we propose several promising directions for future work that remain to be explored.

    - **Comprehensive Ablation Study.** A key missing analysis in this investigation is an ablation study of our proposed DR-cGMM. We introduced three different regularizers and demonstrated that the method outperforms deep learning approaches. To better understand the source of these performance gains, we would need to evaluate all combinations of the regularizers (both individual and pairwise) by running the full training and evaluation protocol. Afterwards, we could conduct Friedman and Nemenyi post-hoc tests to determine whether enabling all regularizers outperforms using only a subset of them. This systematic evaluation would provide valuable insights into which components of our approach contribute most significantly to its generalisation capabilities and potentially identify opportunities for computational optimisation without sacrificing performance.

- **Extending Low-Data Regime Assessment.** Our study was conducted under a low-data regime. Future work should focus on expanding the dataset or incorporating data augmentation techniques that could allow for a more thorough evaluation of deep models' ability to learn robust, position-invariant representations. This investigation would establish clearer guidelines on method selection based on available training data and determine whether a crossover point exists where deep learning approaches begin to outperform statistical methods.
- **Advanced Hyperparameter Optimization.** Due to computational constraints, we explored only a single variant per model with limited hyperparameter tuning. Future work should implement more sophisticated optimization strategies by: (1) expanding the search space to include different network architectures, optimization algorithms, and regularization techniques; (2) employing alternative evaluation metrics beyond accuracy, such as loss or AUC; and (3) investigating knowledge transfer techniques, including pretraining DL methods on related EMG datasets to enhance generalization while maintaining low-data applicability.
- **Navigation Tasks and Dynamic Environments.** Our current position-invariant feature learning supports daily activities effectively, but future work should incorporate continuous limb position tracking for navigation-focused applications. Moreover, extending this framework to dynamic, real-world movements and transitions would better approximate practical prosthetic use.

By addressing these areas, we aim to further strengthen the generalisability and dexterity of myoelectric control systems by mitigating the confounding factor of position effect in real-world assistive technologies, ultimately improving the reliability and usability of these systems for end users.

Appendix A

Ethics Approval

Below we present an example of an ethics approval document that has been filled in and used during our data acquisition protocol for project Szymaniak et al. (2024).

The experiment was conducted according to the Declaration of Helsinki principles for medical research involving human subjects Association (2013) and it was approved by the Ethics Commission of the University of Edinburgh, ethics approval number: 2019/89177.

THE UNIVERSITY of EDINBURGH
informatics

Participant Consent Form

Project title:	Limb Position Dataset Recording
Principal investigator (PI):	Kia Nazarpour
Researcher(s):	Kyranou Iris
PI contact details:	kianoush.nazarpour@ed.ac.uk

By participating in the study you agree that:

- I have read and understood below participant information sheets (PIS) and I have had the opportunity to ask questions, and that any questions I had were answered to my satisfaction.

PIS1	PIS2	PIS3	PIS4	PIS5
✓				

- My participation is voluntary, and that I can withdraw at any time without giving a reason. Withdrawing will not affect any of my rights. In addition, should I not wish to answer any particular question or questions, I am free to decline.
- I consent to my anonymised data being used in academic publications and presentations.
- I understand that my anonymised data will be stored for the duration outlined in the Participant Information Sheet.

Please tick yes or no for each of these statements.

		Yes	No
1.	I agree to being video recorded.		
2.	I allow my data to be used in future ethically approved research.		
3.	I understand that my responses will be kept strictly confidential. I give permission for members of the research team to have access to my anonymised data and responses. I understand that my name will not be linked with the research materials, and I will not be identified or identifiable in the report or reports that result from the research.		
4.	I agree to take part in this study.		

[redacted] 23/05/2023 [redacted]
Name of Participant — Date — Signature

[redacted] [redacted]
Name of person taking consent — Date — Signature
To be signed and dated in presence of the participant

[redacted] [redacted]
Name of Chief Investigator — Date — Signature

Once this has been signed by all parties the participant can receive a copy of the signed and dated consent form, participant information sheets and any other written information provided to the participants. A copy of the signed and dated consent form will be kept with the project's main documents in a secure location.

Participant Information Sheet 1

Project title:	Limb Position Dataset Recording
Principal investigator:	Kia Nazarpour
Researcher collecting data:	Iris Kyranou, Jiechong Zhang
Funder (if applicable):	Engineering and Physical Sciences Research Council

This study was certified according to the Informatics Research Ethics Process, RT number 2019/89177. Please take time to read the following information carefully. You should keep this page for your records.

Who are the researchers?
The team are members of Edinburgh Neuroprosthetics Laboratory. Kia Nazarpour is Principal Investigator. Kyranou Iris is the researcher of this project within the team. They may be accompanied by PhD and Master's students while conducting this project.

What is the purpose of the study?
We work to develop prosthetic hands that are fit for purpose. You are being asked to take part in a research study on how your brain can learn to activate your muscles in novel ways and how muscle activity could be used to control robotic or prosthetic hands.

Do I have to take part?
No – participation in this study is entirely up to you. You may decide to stop being a part of the research study at any time without explanation. You have the right to ask that any data you have supplied to that point be withdrawn or destroyed. You have the right to refuse to answer or respond to any question that is asked of you. You have the right to have your questions about the procedures answered (unless answering these questions would interfere with the study's outcome). If you have any questions as a result of reading this information sheet, you should ask the researcher before the study begins. You will have the option of taking part in the longer or shorter experiments.

What will happen if I decide to take part?
The technique that we use is non-invasive electromyography. In this method we measure electrical activity of muscles from the surface of the skin. Electromyography is a well-known technique and is widely used in both academia and the NHS. As such

the risks associated with our research is minimal. The experimental protocol will be as the following:

After skin preparation, e.g. wiping off any dirt from the skin with clinical level (NHS-approved) wipes, electronic sensors will be placed on your skin to measure the activity of muscles in your hand and forearm. Sensors will usually be placed by an experimental operator; however, you will also have the option to place them yourself.

Once sensors are placed, a brief calibration routine will be performed that will require you to perform specific grasps with your hand.

You will be prompted by visual and auditory cues to perform a specific sequence of grasps while holding your arm in specific positions. At any point the experimental operator will be present to make sure that everything works as expected and to answer your questions. You will have various opportunities to rest between different recordings and they will be made obvious to you by the operator in the beginning and during the recording.

Time Commitment
Between different experimental sessions, there will be rest period of up to 30 minutes during which you can relax. The sensors may remain attached to you during the break. Therefore, we may ask that your body movement will have to be under the supervision of the research team. If the experiment takes more than 90 minutes, we will provide light refreshment, e.g. fruits, biscuits and tea/coffee. In the case of long recording sessions, buffet lunch will be provided. Please see the Participants' Rights section for more details.

Depending on the study question, you may be asked to return to the laboratory, up to five time, to repeat the same experiment. We are interested in quantifying the differences of the recordings over the period of many days. Before you sign up in study, you will know whether it is a one-day or a multi-day study.

Are there any risks associated with taking part?
There are no risks associated with participation. All equipment and systems have been tested for safety. The experimental protocol is safe.

If recommended by the UK and/or Scottish Government or The University of Edinburgh Health and Safety guidelines, all experimental operators will be wearing personal protection equipment (PPE). You will also be given the option of wearing such equipment. Social distancing measures will be maintained when contact is not necessary. The laboratory space will be clean. Staff are required to wear gloves and masks when entering the space for three days prior to experiments. Any surfaces you may come into contact with will have been cleaned prior to the experiment.

What will happen to the results of this study?
The results of this study may be summarised in published articles, reports and presentations. Quotes or key findings will be anonymised. With your consent, information can also be used for future research. Your experimental data will be anonymised and archived on a public data repository as per the requirement of the funding agency.

Data protection and confidentiality.
The data we collect do not contain any personal information about you except your initials for data management and your age for statistical analysis purposes. If you would like your data to be completely anonymised you must inform us as soon as possible, preferably before signing this document. No one will link the data you provided to the identifying information you supply.

Your data will be processed in accordance with Data Protection Law. All information collected about you will be kept strictly confidential. Your data will be referred to by a unique code, e.g. participant number. Your data will only be viewed by the researcher/research team, including Kia Nazarpour, Iris Kyranou, and Jiechong Zhang.

All electronic data will be stored on a password-protected encrypted computer, on the School of Informatics' secure file servers and all paper records will be stored in a locked filing cabinet in the PI's office. Your consent information will be kept separately from your responses in order to minimise risk.

What are my data protection rights?
The University of Edinburgh is a Data Controller for the information you provide. You have the right to access information held about you. Your right of access can be exercised in accordance Data Protection Law. You also have other rights including

rights of correction, erasure and objection. For more details, including the right to lodge a complaint with the Information Commissioner's Office, please visit www.ico.org.uk. Questions, comments and requests about your personal data can also be sent to the University Data Protection Officer at dpo@ed.ac.uk.

Who can I contact?
Kia Nazarpour will be glad to answer your questions about this study at any time. You may contact him at kianoush.nazarpour@ed.ac.uk. If you want to find out about the final results of this study, you can contact Dr Nazarpour directly. If you wish to make a complaint about the study, please contact inf-ethics@inf.ed.ac.uk. When you contact, please provide the study title and detail the nature of your complaint.

Signature
By signing below, you are agreeing that:

- you have read and understood the Participant Information Sheet 1,
- questions about your participation in this study have been answered satisfactorily,
- you are aware of the potential risks (if any), and
- you are taking part in this research study voluntarily (without coercion).

Participant's Name (Printed)*: [redacted]
Participant's signature*:
Date:

**Participants wishing to preserve some degree of anonymity may use their initials.*

Appendix B

Additional Experiments on Arm Translation Dataset

B.1 Aggregated Results

The reported tables in this section present algorithm-specific performance on 10 DG tasks across all subjects. Each table includes the outer-loop averaged test accuracy and standard deviation over 3 seeds. The results are presented for both Con_+ and Con_x.

LDA

Task	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	Avg ± std
$P_{3,5,7,9} \rightarrow P_1$	97.47	98.13	94.47	88.96	92.28	90.19	89.85	92.07	92.93 ± 0.03
$P_{1,5,7,9} \rightarrow P_3$	95.48	99.05	94.74	92.47	94.20	92.67	89.63	96.46	94.34 ± 0.03
$P_{1,3,7,9} \rightarrow P_5$	97.36	97.98	95.99	95.00	93.42	94.85	90.85	93.71	94.89 ± 0.02
$P_{1,3,5,9} \rightarrow P_7$	92.34	99.25	92.16	90.90	93.30	90.85	90.12	96.24	93.14 ± 0.03
$P_{1,3,5,7} \rightarrow P_9$	96.53	98.74	93.79	95.14	93.13	96.29	93.13	89.03	94.47 ± 0.03
$P_{4,5,6,8} \rightarrow P_2$	97.06	96.06	96.21	93.23	94.00	95.71	82.33	90.94	93.19 ± 0.05
$P_{2,5,6,8} \rightarrow P_4$	96.55	94.29	94.74	82.69	93.81	95.31	89.12	97.60	93.01 ± 0.05
$P_{2,4,6,8} \rightarrow P_5$	97.86	95.07	92.47	96.21	94.12	97.43	93.88	96.02	95.38 ± 0.02
$P_{2,4,5,8} \rightarrow P_6$	97.55	91.90	95.80	92.84	95.19	96.43	93.11	87.62	93.81 ± 0.03
$P_{2,4,5,6} \rightarrow P_8$	97.04	95.39	94.90	94.95	93.47	96.51	94.47	89.97	94.59 ± 0.02
Avg ± Std	96.52 ± 0.02	96.59 ± 0.02	94.53 ± 0.01	92.24 ± 0.04	93.69 ± 0.01	94.62 ± 0.02	90.65 ± 0.03	92.97 ± 0.03	93.98 ± 0.03

Table B.1 LDA performance on the outer loop with leave-one-position-out generalisation.

QDA

Task	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	Avg ± Std
$P_{3,5,7,9} \rightarrow P_1$	97.28	98.52	95.92	90.03	93.40	94.51	91.23	77.93	92.35 ± 0.06
$P_{1,5,7,9} \rightarrow P_3$	92.21	98.89	94.86	87.21	95.49	92.98	89.41	90.31	92.67 ± 0.04
$P_{1,3,7,9} \rightarrow P_5$	97.33	97.57	96.58	92.53	94.30	94.86	94.34	96.62	95.52 ± 0.02
$P_{1,3,5,9} \rightarrow P_7$	89.43	98.88	88.09	79.23	94.63	92.96	90.78	97.04	91.38 ± 0.06
$P_{1,3,5,7} \rightarrow P_9$	95.43	98.83	93.13	94.93	86.70	96.16	94.54	86.53	93.28 ± 0.04
$P_{4,5,6,8} \rightarrow P_2$	97.04	96.12	96.68	93.28	93.78	93.83	78.57	89.68	92.37 ± 0.06
$P_{2,5,6,8} \rightarrow P_4$	95.09	92.96	93.62	84.40	93.32	95.85	90.66	98.11	93.00 ± 0.04
$P_{2,4,6,8} \rightarrow P_5$	97.65	95.31	90.29	94.56	94.83	98.13	90.56	96.77	94.76 ± 0.03
$P_{2,4,5,8} \rightarrow P_6$	97.74	92.19	97.21	76.77	94.97	97.79	94.32	87.52	92.31 ± 0.07
$P_{2,4,5,6} \rightarrow P_8$	96.80	95.97	89.76	94.01	91.45	96.31	95.07	83.11	92.81 ± 0.05
Avg ± Std	95.60 ± 0.03	96.52 ± 0.02	93.62 ± 0.03	88.70 ± 0.06	93.29 ± 0.02	95.34 ± 0.02	90.95 ± 0.05	90.36 ± 0.06	93.05 ± 0.05

Table B.2 QDA performance on the outer loop with leave-one-position-out generalisation.

DR-cGMM

Task	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	Avg ± Std
$P_{3,10,7,9} \rightarrow P_1$	97.56 ± 0.000	98.62 ± 0.000	95.58 ± 0.004	90.54 ± 0.004	92.88 ± 0.000	94.60 ± 0.001	90.87 ± 0.003	85.72 ± 0.008	93.29 ± 0.042
$P_{1,10,7,9} \rightarrow P_3$	95.00 ± 0.006	99.11 ± 0.000	94.86 ± 0.001	94.11 ± 0.002	94.85 ± 0.001	93.00 ± 0.001	89.52 ± 0.002	95.67 ± 0.005	94.52 ± 0.027
$P_{1,3,7,9} \rightarrow P_{10}$	97.44 ± 0.001	98.05 ± 0.000	96.70 ± 0.000	94.40 ± 0.002	94.41 ± 0.003	96.84 ± 0.001	92.05 ± 0.001	94.29 ± 0.004	95.52 ± 0.020
$P_{1,3,10,9} \rightarrow P_7$	92.99 ± 0.001	99.31 ± 0.000	92.60 ± 0.002	87.89 ± 0.010	94.47 ± 0.001	93.27 ± 0.007	90.04 ± 0.002	97.28 ± 0.004	93.48 ± 0.037
$P_{1,3,10,7} \rightarrow P_9$	96.42 ± 0.001	99.01 ± 0.000	93.96 ± 0.001	95.22 ± 0.002	92.00 ± 0.008	97.10 ± 0.001	93.57 ± 0.002	88.83 ± 0.002	94.51 ± 0.032
$P_{4,5,6,8} \rightarrow P_2$	97.36 ± 0.000	96.08 ± 0.002	96.56 ± 0.000	94.80 ± 0.002	94.14 ± 0.000	95.83 ± 0.002	81.22 ± 0.014	91.45 ± 0.003	93.43 ± 0.053
$P_{2,5,6,8} \rightarrow P_4$	95.27 ± 0.002	94.49 ± 0.003	94.85 ± 0.002	82.23 ± 0.003	94.66 ± 0.001	95.77 ± 0.000	90.66 ± 0.005	98.15 ± 0.002	93.26 ± 0.049
$P_{2,4,6,8} \rightarrow P_5$	97.94 ± 0.000	95.47 ± 0.001	92.09 ± 0.001	96.37 ± 0.002	94.43 ± 0.003	97.93 ± 0.000	92.23 ± 0.010	96.50 ± 0.005	95.37 ± 0.023
$P_{2,4,5,8} \rightarrow P_6$	97.68 ± 0.000	92.62 ± 0.000	96.58 ± 0.000	94.09 ± 0.003	95.30 ± 0.001	97.24 ± 0.001	94.62 ± 0.002	88.22 ± 0.001	94.54 ± 0.031
$P_{2,4,5,6} \rightarrow P_8$	97.05 ± 0.001	95.53 ± 0.001	94.65 ± 0.002	95.41 ± 0.003	92.98 ± 0.001	97.23 ± 0.002	95.16 ± 0.001	91.56 ± 0.002	94.95 ± 0.019
Avg ± Std	96.47 ± 0.015	96.83 ± 0.022	94.84 ± 0.015	92.5 ± 0.042	94.01 ± 0.010	95.88 ± 0.016	90.99 ± 0.037	92.77 ± 0.040	94.29 ± 0.008

Table B.3 DR-cGMM performance on the outer loop with leave-one-position-out generalisation.

ERM

Task	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	Avg ± Std
$P_{3,10,7,9} \rightarrow P_1$	96.35 ± 0.008	98.38 ± 0.002	95.80 ± 0.003	88.76 ± 0.030	92.54 ± 0.007	95.32 ± 0.005	91.16 ± 0.005	88.68 ± 0.016	93.37 ± 0.036
$P_{1,10,7,9} \rightarrow P_3$	89.70 ± 0.024	99.17 ± 0.001	96.07 ± 0.007	89.76 ± 0.009	94.47 ± 0.004	95.17 ± 0.005	84.50 ± 0.011	90.06 ± 0.038	92.36 ± 0.047
$P_{1,3,7,9} \rightarrow P_{10}$	97.00 ± 0.003	98.51 ± 0.001	96.90 ± 0.003	94.54 ± 0.002	93.64 ± 0.004	96.84 ± 0.002	93.61 ± 0.011	97.03 ± 0.004	96.01 ± 0.018
$P_{1,3,10,9} \rightarrow P_7$	93.63 ± 0.023	98.51 ± 0.003	92.81 ± 0.022	87.94 ± 0.022	94.12 ± 0.004	85.35 ± 0.044	91.28 ± 0.007	97.81 ± 0.001	92.68 ± 0.045
$P_{1,3,10,7} \rightarrow P_9$	96.60 ± 0.002	99.01 ± 0.001	95.73 ± 0.013	94.37 ± 0.006	88.28 ± 0.007	96.79 ± 0.002	93.33 ± 0.007	88.46 ± 0.007	94.07 ± 0.039
$P_{4,5,6,8} \rightarrow P_2$	90.55 ± 0.026	94.00 ± 0.002	94.30 ± 0.004	93.02 ± 0.004	94.01 ± 0.002	92.11 ± 0.014	77.97 ± 0.014	90.62 ± 0.001	90.82 ± 0.054
$P_{2,5,6,8} \rightarrow P_4$	92.08 ± 0.022	94.05 ± 0.006	92.02 ± 0.028	84.50 ± 0.008	93.49 ± 0.004	96.48 ± 0.001	90.08 ± 0.011	97.02 ± 0.012	92.47 ± 0.040
$P_{2,4,6,8} \rightarrow P_5$	97.46 ± 0.003	95.55 ± 0.002	91.71 ± 0.007	95.86 ± 0.003	94.45 ± 0.007	98.60 ± 0.001	92.11 ± 0.024	96.09 ± 0.007	95.23 ± 0.024
$P_{2,4,5,8} \rightarrow P_6$	97.50 ± 0.001	92.32 ± 0.000	96.19 ± 0.004	89.31 ± 0.011	94.81 ± 0.000	96.89 ± 0.007	92.91 ± 0.010	88.80 ± 0.025	93.59 ± 0.033
$P_{2,4,5,6} \rightarrow P_8$	95.25 ± 0.005	95.40 ± 0.007	93.34 ± 0.008	93.66 ± 0.008	91.89 ± 0.005	96.71 ± 0.003	94.28 ± 0.005	89.42 ± 0.016	93.74 ± 0.023
Avg ± Std	94.61 ± 0.028	96.49 ± 0.024	94.49 ± 0.018	91.17 ± 0.035	93.17 ± 0.018	95.03 ± 0.036	90.12 ± 0.048	92.40 ± 0.038	93.43 ± 0.014

Table B.4 ERM performance on the outer loop with leave-one-position-out generalisation.

CORAL

Task	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	Avg ± Std
$P_{3,10,7,9} \rightarrow P_1$	96.88 ± 0.001	98.42 ± 0.003	95.03 ± 0.004	90.56 ± 0.004	91.02 ± 0.011	94.43 ± 0.010	91.61 ± 0.002	88.31 ± 0.020	93.28 ± 0.035
$P_{1,10,7,9} \rightarrow P_3$	87.66 ± 0.012	99.18 ± 0.000	97.01 ± 0.003	89.66 ± 0.015	94.97 ± 0.003	95.75 ± 0.006	83.97 ± 0.011	90.91 ± 0.042	92.39 ± 0.052
$P_{1,3,7,9} \rightarrow P_{10}$	97.32 ± 0.003	98.66 ± 0.003	97.06 ± 0.003	93.06 ± 0.011	93.52 ± 0.004	96.81 ± 0.005	92.71 ± 0.018	96.84 ± 0.009	95.75 ± 0.023
$P_{1,3,10,9} \rightarrow P_7$	93.61 ± 0.054	98.39 ± 0.002	91.75 ± 0.010	88.89 ± 0.011	94.16 ± 0.006	88.51 ± 0.039	90.45 ± 0.011	97.99 ± 0.002	92.97 ± 0.038
$P_{1,3,10,7} \rightarrow P_9$	96.32 ± 0.004	99.09 ± 0.002	95.13 ± 0.008	94.75 ± 0.002	87.38 ± 0.016	97.21 ± 0.003	94.66 ± 0.002	88.10 ± 0.007	94.08 ± 0.042
$P_{4,5,6,8} \rightarrow P_2$	93.92 ± 0.031	93.86 ± 0.002	94.44 ± 0.002	93.14 ± 0.006	94.00 ± 0.003	92.46 ± 0.008	79.81 ± 0.001	90.62 ± 0.019	91.53 ± 0.049
$P_{2,5,6,8} \rightarrow P_4$	92.88 ± 0.006	94.14 ± 0.007	93.16 ± 0.009	84.92 ± 0.016	93.77 ± 0.011	96.46 ± 0.004	89.77 ± 0.006	95.59 ± 0.011	92.59 ± 0.037
$P_{2,4,6,8} \rightarrow P_5$	97.70 ± 0.000	95.25 ± 0.003	92.51 ± 0.018	95.89 ± 0.005	94.93 ± 0.002	98.35 ± 0.002	93.08 ± 0.013	96.91 ± 0.007	95.58 ± 0.021
$P_{2,4,5,8} \rightarrow P_6$	97.40 ± 0.004	92.39 ± 0.001	96.25 ± 0.008	88.48 ± 0.009	94.71 ± 0.010	96.19 ± 0.011	93.16 ± 0.005	89.17 ± 0.023	93.47 ± 0.033
$P_{2,4,5,6} \rightarrow P_8$	95.12 ± 0.006	94.30 ± 0.006	94.29 ± 0.011	93.98 ± 0.005	91.45 ± 0.007	96.28 ± 0.011	94.03 ± 0.004	88.49 ± 0.013	93.49 ± 0.024
Avg ± Std	94.88 ± 0.029	96.37 ± 0.025	94.66 ± 0.017	91.33 ± 0.032	92.99 ± 0.023	95.24 ± 0.027	90.32 ± 0.045	92.29 ± 0.038	93.51 ± 0.013

Table B.5 CORAL performance on the outer loop with leave-one-position-out generalisation.

MMD

Task	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	Avg ± Std
$P_{3,10,7,9} \rightarrow P_1$	96.79 ± 0.004	98.67 ± 0.001	94.52 ± 0.013	90.02 ± 0.009	92.46 ± 0.004	96.02 ± 0.006	91.28 ± 0.012	87.27 ± 0.027	93.38 ± 0.038
$P_{1,10,7,9} \rightarrow P_3$	89.14 ± 0.014	98.74 ± 0.008	96.50 ± 0.007	90.08 ± 0.001	94.41 ± 0.010	95.70 ± 0.011	83.48 ± 0.006	91.03 ± 0.008	92.39 ± 0.049
$P_{1,3,7,9} \rightarrow P_{10}$	96.84 ± 0.003	98.37 ± 0.000	95.58 ± 0.026	94.33 ± 0.004	93.41 ± 0.002	97.11 ± 0.003	93.17 ± 0.020	96.99 ± 0.003	95.72 ± 0.019
$P_{1,3,10,9} \rightarrow P_7$	96.88 ± 0.003	98.34 ± 0.004	89.20 ± 0.019	88.36 ± 0.006	94.66 ± 0.004	90.21 ± 0.041	90.90 ± 0.002	97.81 ± 0.001	93.29 ± 0.041
$P_{1,3,10,7} \rightarrow P_9$	96.09 ± 0.001	98.79 ± 0.005	95.71 ± 0.009	93.75 ± 0.009	87.95 ± 0.011	97.08 ± 0.006	94.08 ± 0.003	88.75 ± 0.004	94.03 ± 0.039
$P_{4,5,6,8} \rightarrow P_2$	94.33 ± 0.034	94.84 ± 0.004	91.61 ± 0.035	93.27 ± 0.007	94.18 ± 0.002	91.02 ± 0.019	78.45 ± 0.002	87.10 ± 0.025	90.60 ± 0.055
$P_{2,5,6,8} \rightarrow P_4$	90.95 ± 0.013	94.71 ± 0.005	94.19 ± 0.004	83.17 ± 0.006	93.02 ± 0.007	96.53 ± 0.002	90.03 ± 0.002	96.52 ± 0.008	92.39 ± 0.044
$P_{2,4,6,8} \rightarrow P_5$	97.47 ± 0.004	95.40 ± 0.003	91.70 ± 0.010	96.15 ± 0.007	94.91 ± 0.004	98.30 ± 0.001	93.97 ± 0.006	97.10 ± 0.006	95.63 ± 0.021
$P_{2,4,5,8} \rightarrow P_6$	97.36 ± 0.002	92.13 ± 0.003	94.08 ± 0.050	88.57 ± 0.005	93.18 ± 0.014	95.93 ± 0.005	91.69 ± 0.024	89.75 ± 0.013	92.84 ± 0.030
$P_{2,4,5,6} \rightarrow P_8$	95.87 ± 0.005	94.40 ± 0.018	92.88 ± 0.014	94.09 ± 0.006	91.65 ± 0.003	97.12 ± 0.002	94.71 ± 0.004	89.76 ± 0.012	93.81 ± 0.023
Avg ± Std	95.17 ± 0.027	96.44 ± 0.023	93.6 ± 0.021	91.18 ± 0.037	92.98 ± 0.019	95.5 ± 0.026	90.18 ± 0.05	92.21 ± 0.042	93.41 ± 0.015

Table B.6 MMD performance on the outer loop with leave-one-position-out generalisation.

B.2 Implementation Details

B.2.1 Model Architecture

Layer	**Type**	**Parameters**	**Output Shape (L=9)**
Input	-	16 channels, length = 9	(B, 16, 9)
Conv1	1D Convolution	in=16, out=64, kernel=3, padding=1	(B, 64, 9)
BatchNorm1	BatchNorm1d	64 channels	(B, 64, 9)
ReLU	Activation	-	(B, 64, 9)
MaxPool1	Pooling	kernel=2, stride=2	(B, 64, 4)
Conv2	1D Convolution	in=64, out=128, kernel=3, padding=1	(B, 128, 4)
BatchNorm2	BatchNorm1d	128 channels	(B, 128, 4)
ReLU	Activation	-	(B, 128, 4)
MaxPool2	Pooling	kernel=2, stride=2	(B, 128, 2)
Conv3	1D Convolution	in=128, out=256, kernel=3, padding=1	(B, 256, 2)
BatchNorm3	BatchNorm1d	256 channels	(B, 256, 2)
ReLU	Activation	-	(B, 256, 2)
MaxPool3	Pooling	kernel=2, stride=2	(B, 256, 1)
Flatten	-	-	(B, 256)

Table B.7 CNN1D architecture overview

We use a 1D convolutional neural network (CNN1D) to extract temporal features from time-series data. The network takes 16-channel input sequences (e.g., from EMG sensors) of length 9 and consists of three convolutional blocks. Each block includes a convolutional layer, batch normalization, ReLU activation, and max pooling. The first block outputs 64 feature maps. Each subsequent block multiplies the feature maps by a factor of $2\times$. Therefore, the final output is a 256-dimensional feature vector per input sequence. The full architecture is described in Table B.7.

Note: B denotes the batch size.

B.2.2 HPO

We list all hyperparameters, their default values, and the search distribution for each hyperparameter in our random hyperparameter trials (sweeps), in Table 5.1.

During HPO, each inner loop calculates average validation accuracy for a single trial. Based on these, we select the optimal set of hyperparameters across 20 HPO trials. These values are then used in the corresponding outer loop to first train and then test the algorithm on the held-out position.

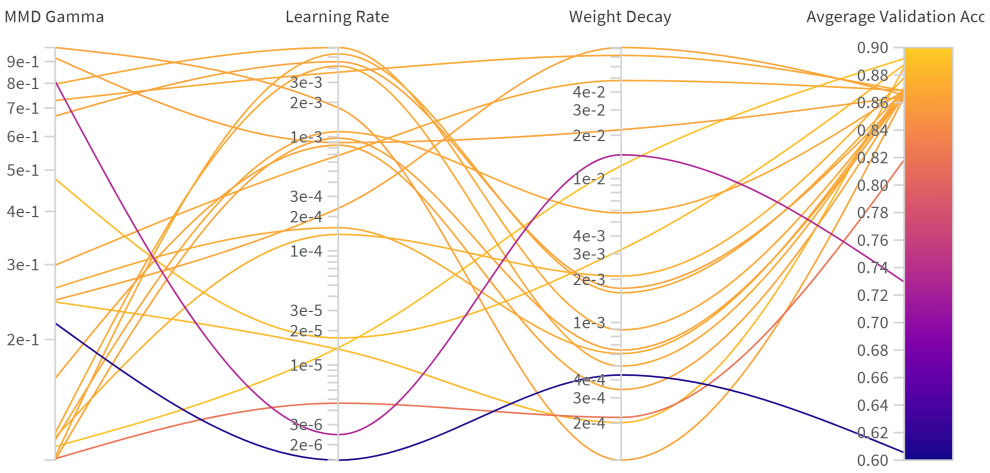


Figure B.1 Example of a HPO sweep, Subject 7: MMD.

Scenario ID	$P_{train} \rightarrow P_{val}$	P_{test}
0	$P_{3,7} \rightarrow P_{5,9}$	P_1
1	$P_{5,9} \rightarrow P_{3,7}$	P_1
2	$P_{1,7} \rightarrow P_{5,9}$	P_3
3	$P_{5,9} \rightarrow P_{1,7}$	P_3
4	$P_{1,7} \rightarrow P_{3,9}$	P_5
5	$P_{3,9} \rightarrow P_{1,7}$	P_5
6	$P_{1,5} \rightarrow P_{3,9}$	P_7
7	$P_{3,9} \rightarrow P_{1,5}$	P_7
8	$P_{1,5} \rightarrow P_{3,7}$	P_9
9	$P_{3,7} \rightarrow P_{1,5}$	P_9
10	$P_{4,6} \rightarrow P_{5,8}$	P_2
11	$P_{5,8} \rightarrow P_{4,6}$	P_2
12	$P_{2,6} \rightarrow P_{5,8}$	P_4
13	$P_{5,8} \rightarrow P_{2,6}$	P_4
14	$P_{2,6} \rightarrow P_{4,8}$	P_5
15	$P_{4,8} \rightarrow P_{2,6}$	P_5
16	$P_{2,5} \rightarrow P_{4,8}$	P_6
17	$P_{4,8} \rightarrow P_{2,5}$	P_6
18	$P_{2,5} \rightarrow P_{4,6}$	P_8
19	$P_{4,6} \rightarrow P_{2,5}$	P_8

Table B.8 Nested cross validation: inner loop. During hyperparameter optimisation with N trials, for each trial the proposed parameters were trained and validated on the given sets. The average performance was calculated across all 20 scenarios. Based on predefined criteria (i.e. argmax validation accuracy or argmin validation loss, the best set of parameters were chosen and used at the inference time in the outer loop. Test positions were not used in the inner loop.

Model Training Specifications

Our comprehensive DG framework necessitated extensive computational resources for model optimisation across diverse scenarios. The total number of trained models M can be formalised as:

$$M = \sum_{s=1}^{S} \sum_{m=1}^{Q} \sum_{l=1}^{L} \sum_{i=1}^{I} (n_{s,m,l,i} + \alpha_{s,m,l,i})$$

where:

- $S = 3$ represents the random seeds employed for statistical robustness,
- $Q = 4$ denotes the hyperparameter optimisation methodologies evaluated, namely DR-cGMM, ERM, CORAL and MMD,
- $L = 8$ corresponds to the subject cohorts examined,
- $I = 20$ indicates the total of inner loop scenarios,
- $n_{s,m,l,i} = 20$ reflects the number of trials per scenario,
- $\alpha_{s,m,l,i} = 10$ accounts for the outer loop evaluations.

This yielded 3,280 models per optimisation algorithm, culminating in 39,360 models across all HPO approaches. With the inclusion of LDA and QDA as baseline comparators, the final experimental corpus comprised 39,520 models, providing substantial statistical power for our comparative analysis.

B.2.3 Hypothesis Testing

Let k be the number of algorithms (in our case, $k = 6$) and N the number of datasets (subjects, $N = 8$). For each dataset $i \in \{1,\ldots,N\}$, we evaluate all algorithms and assign a rank r_{ij} to algorithm $j \in \{1,\ldots,k\}$, where the best-performing algorithm receives rank 1, the second-best rank 2, and so on. In case of ties, average ranks are used.

The average rank of each algorithm across all datasets is computed as:

$$\bar{r}_j = \frac{1}{N}\sum_{i=1}^{N} r_{ij}.$$

Friedman Test. To assess whether there are statistically significant differences among the algorithms, we apply the Friedman test. The test statistic is given by:

$$\chi_F^2 = \frac{12N}{k(k+1)}\left[\sum_j R_j^2 - \frac{k(k+1)^2}{4}\right],$$

where $R_j = \sum_{i=1}^{N} r_{ij}$ is the total rank of algorithm j. This statistic follows a chi-square distribution with $k-1$ degrees of freedom under the null hypothesis that all algorithms perform equally.

Post-hoc Nemenyi Test. If the Friedman test rejects the null hypothesis, we proceed with the Nemenyi post-hoc test to compare algorithms pairwise. The critical difference (CD) that determines statistical significance between average ranks is computed as:

$$CD = q_\alpha \sqrt{\frac{k(k+1)}{6N}},$$

where q_α is the critical value based on the Studentized range distribution for a given significance level α. Two algorithms are considered significantly different if the difference in their average ranks exceeds this critical difference.

Bibliography

Abowd, G. D., Dey, A. K., Brown, P. J., Davies, N., Smith, M., and Steggles, P. (1999). Towards a better understanding of context and context-awareness. In *Handheld and Ubiquitous Computing: First International Symposium, HUC'99 Karlsruhe, Germany, September 27–29, 1999 Proceedings 1*, pages 304–307. Springer.

Adewuyi, A. A., Hargrove, L. J., and Kuiken, T. A. (2017). Resolving the effect of wrist position on myoelectric pattern recognition control. *Journal of neuroengineering and rehabilitation*, 14:1–11.

Ahmad, A., Migniot, C., and Dipanda, A. (2019). Hand pose estimation and tracking in real and virtual interaction: A review. *Image and Vision Computing*, 89:35–49.

Al-Angari, H. M., Kanitz, G., Tarantino, S., and Cipriani, C. (2016). Distance and mutual information methods for emg feature and channel subset selection for classification of hand movements. *Biomedical Signal Processing and Control*, 27:24–31.

Al-Timemy, A. H., Bugmann, G., and Escudero, J. (2018). Adaptive windowing framework for surface electromyogram-based pattern recognition system for transradial amputees. *Sensors*, 18(8):2402.

Amsüss, S., Goebel, P. M., Jiang, N., Graimann, B., Paredes, L., and Farina, D. (2013a). Self-correcting pattern recognition system of surface emg signals for upper limb prosthesis control. *IEEE Transactions on Biomedical Engineering*, 61(4):1167–1176.

Amsüss, S., Paredes, L. P., Rudigkeit, N., Graimann, B., Herrmann, M. J., and Farina, D. (2013b). Long term stability of surface emg pattern classification for prosthetic control. In *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 3622–3625. IEEE.

Amsüss, S., Goebel, P. M., Jiang, N., Graimann, B., Paredes, L., and Farina, D. (2014). Self-correcting pattern recognition system of surface emg signals for upper limb prosthesis control. *IEEE Transactions on Biomedical Engineering*, 61(4):1167–1176.

Angluin, D. (1988). Queries and concept learning. *Machine learning*, 2(4):319–342.

Antuvan, C. W. (2019). *Decoding human motion intention using myoelectric signals for assistive technologies*. PhD thesis, Nanyang Technological University, Singapore.

Antuvan, C. W., Ison, M., and Artemiadis, P. (2014). Embedded human control of robots using myoelectric interfaces. *IEEE Transactions on neural systems and rehabilitation engineering*, 22(4):820–827.

Association, W. M. (2013). World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects. *JAMA*, 310(20):2191–2194.

Atzori, M., Gijsberts, A., Castellini, C., Caputo, B., Hager, A.-G. M., Elsig, S., Giatsidis, G., Bassetto, F., and Müller, H. (2014a). Electromyography data for non-invasive naturally-controlled robotic hand prostheses. *Scientific data*, 1(1):1–13.

Atzori, M., Gijsberts, A., Castellini, C., Caputo, B., Mittaz Hager, A., Elsig, S., Giatsidis, G., Bassetto, F., and Müller, H. (2014b). Electromyography data for non-invasive naturally-controlled robotic hand prostheses. *Sci Data*, 1:140053.

Atzori, M., Gijsberts, A., Heynen, S., Hager, A.-G. M., Deriaz, O., van der Smagt, P., Castellini, C., Caputo, B., and Müller, H. (2012). Building the Ninapro database: A resource for the biorobotics community. In *2012 4th IEEE RAS 'I&' EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob)*, pages 1258–1265.

Azhiri, R. B., Esmaeili, M., and Nourani, M. (2021a). Emg-based feature extraction and classification for prosthetic hand control. *arXiv preprint arXiv:2107.00733*.

Azhiri, R. B., Esmaeili, M., and Nourani, M. (2021b). Real-time emg signal classification via recurrent neural networks. In *2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 2628–2635. IEEE.

Bai, S., Kolter, J. Z., and Koltun, V. (2018). An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv preprint arXiv:1803.01271*.

Barradas, V. R., Cho, W., and Koike, Y. (2023). Emg space similarity feedback promotes learning of expert-like muscle activation patterns in a complex motor skill. *Frontiers in Human Neuroscience*, 16:805867.

Basmajian, J. V. (1985). Muscles alive. *Their functions revealed by electromyography*.

Batzianoulis, I., Krausz, N. E., Simon, A. M., Hargrove, L., and Billard, A. (2018). Decoding the grasping intention from electromyography during reaching motions. *Journal of neuroengineering and rehabilitation*, 15(1):1–13.

Beaulieu, R. J., Masters, M. R., Betthauser, J., Smith, R. J., Kaliki, R., Thakor, N. V., and Soares, A. B. (2017). Multi-position training improves robustness of pattern recognition and reduces limb-position effect in prosthetic control. *JPO: Journal of Prosthetics and Orthotics*, 29(2):54–62.

Behrend, C., Reizner, W., Marchessault, J. A., and Hammert, W. C. (2011). Update on advances in upper extremity prosthetics. *The Journal of hand surgery*, 36(10):1711–1717.

Benjamini, Y. and Braun, H. I. (2002). John w. tukey's contributions to multiple comparisons. *Annals of Statistics*, 30.

Bergstra, J. and Bengio, Y. (2012). Random search for hyper-parameter optimization. *The journal of machine learning research*, 13(1):281–305.

Betthauser, J. L., Hunt, C. L., Osborn, L. E., Kaliki, R. R., and Thakor, N. V. (2016). Limb-position robust classification of myoelectric signals for prosthesis control using sparse representations. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 6373–6376. IEEE.

Biau, G. and Scornet, E. (2016). A random forest guided tour. *Test*, 25(2):197–227.

Biddiss, E. A. and Chau, T. T. (2007). Upper limb prosthesis use and abandonment: a survey of the last 25 years. *Prosthetics and orthotics international*, 31(3):236–257.

Blana, D., Kyriacou, T., Lambrecht, J. M., and Chadwick, E. K. (2016). Feasibility of using combined EMG and kinematic signals for prosthesis control: A simulation study using a virtual reality environment. *Journal of Electromyography and Kinesiology*, 29:21–27.

Boostani, R. and Moradi, M. H. (2003). Evaluation of the forearm emg signal features for the control of a prosthetic hand. *Physiological measurement*, 24(2):309.

Boschmann, A. and Platzner, M. (2012). Reducing classification accuracy degradation of pattern recognition based myoelectric control caused by electrode shift using a high density electrode array. In *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4324–4327. IEEE.

Boschmann, A. and Platzner, M. (2014). Towards robust hd emg pattern recognition: Reducing electrode displacement effect using structural similarity. In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4547–4550. IEEE.

Buchthal, F. and Schmalbruch, H. (1980). Motor unit of mammalian muscle. *Physiological reviews*, 60(1):90–142.

Buelvas, H. E. P., Montaña, J. D. T., and Serrezuela, R. R. (2023). Hand gesture classification using deep learning and cwt images based on multi-channel surface emg signals. In *2023 3rd International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME)*, pages 1–7. IEEE.

Buongiorno, D., Cascarano, G. D., and Brunetti, A. B. D. F. B. (2019). A survey on deep learning in electromyographic signal analysis. In Huang, D.-S., Huang, Z.-K., and Hussain, A., editors, *International Conference on Intelligent Computing*, pages 751–761. Springer International Publishing.

Burns, M. K., Pei, D., and Vinjamuri, R. (2019). Myoelectric control of a soft hand exoskeleton using kinematic synergies. *IEEE Transactions on Biomedical Circuits and Systems*, 13(6):1351–1361.

Burrough, S. F. and Brook, J. A. (1985). Patterns of acceptance and rejection of upper limb prostheses. *Orthot Prosthet*, 39(2):40–47.

Buskirk, E. and Komi, P. (1970). Reproducibility of electromyographic measurements with inserted wire electrodes and surface electrodes. *Acta Physiologica Scandinavica*, 79(2):29A.

Campbell, E., Chang, J., Phinyomark, A., and Scheme, E. (2020a). A comparison of amputee and able-bodied inter-subject variability in myoelectric control. In *MEC Symposium Conference*.

Campbell, E., Eddy, E., Bateman, S., Côté-Allard, U., and Scheme, E. (2024). Context-informed incremental learning improves both the performance and resilience of myoelectric control. *Journal of NeuroEngineering and Rehabilitation*, 21(1):70.

Campbell, E., Phinyomark, A., and Scheme, E. (2019). Linear discriminant analysis with bayesian risk parameters for myoelectric control. In *2019 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pages 1–5. IEEE.

Campbell, E., Phinyomark, A., and Scheme, E. (2020b). Current trends and confounding factors in myoelectric control: Limb position and contraction intensity. *Sensors*, 20(6):1613.

Campbell, E., Phinyomark, A., and Scheme, E. (2021). Deep cross-user models reduce the training burden in myoelectric control. *Frontiers in Neuroscience*, 15:657958.

Cardoso, T. N., Silva, R. M., Canuto, S., Moro, M. M., and Gonçalves, M. A. (2017). Ranked batch-mode active learning. *Information Sciences*, 379:313–337.

Castellini, C. and Van Der Smagt, P. (2009). Surface EMG in advanced hand prosthetics. *Biological cybernetics*, 100:35–47.

Chang, C.-C. and Lin, C.-J. (2011). Libsvm: a library for support vector machines. *ACM transactions on intelligent systems and technology (TIST)*, 2(3):1–27.

Chen, L., Geng, Y., and Li, G. (2011). Effect of upper-limb positions on motion pattern recognition using electromyography. In *2011 4th International Congress on Image and Signal Processing*, volume 1, pages 139–142. IEEE.

Chen, W., Feng, L., Lu, J., Wu, B., and Liu, D. (2023). MyoBit: A public dataset based on an armband with 16 sEMG channels for gesture recognition under non-ideal conditions. In *2023 27th International Conference on Methods and Models in Automation and Robotics (MMAR)*, pages 69–74.

Chen, X., Zhang, D., and Zhu, X. (2013). Application of a self-enhancing classification method to electromyography pattern recognition for multifunctional prosthesis control. *Journal of neuroengineering and rehabilitation*, 10(1):1–13.

Cheng, J., Wei, F., Li, C., Liu, Y., Liu, A., and Chen, X. (2018). Position-independent gesture recognition using semg signals via canonical correlation analysis. *Computers in biology and medicine*, 103:44–54.

Cifrek, M., Medved, V., Tonković, S., and Ostojić, S. (2009). Surface EMG based muscle fatigue evaluation in biomechanics. *Clinical biomechanics*, 24(4):327–340.

Cipriani, C., Sassu, R., Controzzi, M., and Carrozza, M. C. (2011). Influence of the weight actions of the hand prosthesis on the performance of pattern recognition based myoelectric control: preliminary study. In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 1620–1623. IEEE.

Cisnal, A., Pérez-Turiel, J., Fraile, J.-C., Sierra, D., and de la Fuente, E. (2021). RobHand: A hand exoskeleton with real-time EMG-driven embedded control. quantifying hand gesture recognition delays for bilateral rehabilitation. *IEEE Access*, 9:137809–137823.

Cognolato, M., Gijsberts, A., Gregori, V., Saetta, G., Giacomino, K., Hager, A.-G. M., Gigli, A., Faccio, D., Tiengo, C., Bassetto, F., et al. (2020). Gaze, visual, myoelectric, and inertial data of grasps for intelligent prosthetics. *Scientific data*, 7(1):43.

Cordella, F., Ciancio, A. L., Sacchetti, R., Davalli, A., Cutti, A. G., Guglielmelli, E., and Zollo, L. (2016). Literature review on needs of upper limb prosthesis users. *Frontiers in neuroscience*, 10:209.

Côté-Allard, U., Campbell, E., Phinyomark, A., Laviolette, F., Gosselin, B., and Scheme, E. (2020a). Interpreting deep learning features for myoelectric control: A comparison with handcrafted features. *Frontiers in bioengineering and biotechnology*, 8:158.

Côté-Allard, U., Fall, C. L., Drouin, A., Campeau-Lecours, A., Gosselin, C., Glette, K., Laviolette, F., and Gosselin, B. (2019a). Deep learning for electromyographic hand gesture signal classification using transfer learning. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 27(4):760–771.

Côté-Allard, U., Gagnon-Turcotte, G., Phinyomark, A., Glette, K., Scheme, E., Laviolette, F., and Gosselin, B. (2019b). Virtual reality to study the gap between offline and real-time emg-based gesture recognition. *arXiv preprint*.

Cote-Allard, U., Gagnon-Turcotte, G., Phinyomark, A., Glette, K., Scheme, E., Laviolette, F., and Gosselin, B. (2021). A transferable adaptive domain adversarial neural network for virtual reality augmented EMG-based gesture recognition. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 29:546–555.

Côté-Allard, U., Gagnon-Turcotte, G., Phinyomark, A., Glette, K., Scheme, E. J., Laviolette, F., and Gosselin, B. (2020b). Unsupervised domain adversarial self-calibration for electromyography-based gesture recognition. *IEEE Access*, 8:177941–177955.

Cox, D. R. (1958). The regression analysis of binary sequences. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 20(2):215–232.

Ctrl-labs at Reality Labs, Sussillo, D., Kaifosh, P., and Reardon, T. (2024). A generic noninvasive neuromotor interface for human-computer interaction. *bioRxiv*, pages 2024–02.

Côté-Allard, U., Gagnon-Turcotte, G., Phinyomark, A., Glette, K., Scheme, E., Laviolette, F., and Gosselin, B. (2021). A transferable adaptive domain adversarial neural network for virtual reality augmented EMG-based gesture recognition. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 29:546–555.

Côté Allard, U., Nougarou, F., Fall, C. L., Giguère, P., Gosselin, C., Laviolette, F., and Gosselin, B. (2016). A convolutional neural network for robotic arm guidance using sEMG based frequency-features. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2464–2470.

Dash, A. and Lahiri, U. (2020). Design of virtual reality-enabled surface electromyogram-triggered grip exercise platform. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 28(2):444–452.

De Luca, C. J. (1997). The use of surface electromyography in biomechanics. *Journal of applied biomechanics*, 13(2):135–163.

Demšar, J. (2006). Statistical comparisons of classifiers over multiple data sets. *Journal of Machine Learning Research*, 7(1):1–30.

Dere, M. D. and Lee, B. (2023). A novel approach to surface emg-based gesture classification using a vision transformer integrated with convolutive blind source separation. *IEEE Journal of Biomedical and Health Informatics*, 28(1):181–192.

Dignum, V. (2018). Ethics in artificial intelligence: introduction to the special issue. *Ethics and Information Technology*, 20(1):1–3.

Du, Y., Jin, W., Wei, W., Hu, Y., and Geng, W. (2017). Surface EMG-based inter-session gesture recognition enhanced by deep domain adaptation. *Sensors*, 17(3):458.

Dwivedi, A., Kwon, Y., and Liarokapis, M. (2020). EMG-based decoding of manipulation motions in virtual reality: Towards immersive interfaces. In *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 3296–3303.

Dyson, M., Barnes, J., and Nazarpour, K. (2018). Myoelectric control with abstract decoders. *Journal of Neural Engineering*, 15(5):056003.

Dyson, M., Dupan, S., Jones, H., and Nazarpour, K. (2020). Learning, generalization, and scalability of abstract myoelectric control. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 28(7):1539–1547.

Eddy, E., Campbell, E., Bateman, S., and Scheme, E. (2024). Understanding the influence of confounding factors in myoelectric control for discrete gesture recognition. *Journal of neural engineering*, 21(3):036015.

Eddy, E., Scheme, E. J., and Bateman, S. (2023). A framework and call to action for the future development of emg-based input in hci. In *Proceedings of the 2023 CHI conference on human factors in computing systems*, pages 1–23.

Englehart, K. and Hudgins, B. (2003). A robust, real-time control scheme for multifunction myoelectric control. *IEEE transactions on biomedical engineering*, 50(7):848–854.

Englehart, K., Hudgins, B., Parker, P. A., and Stevenson, M. (1999). Classification of the myoelectric signal using time-frequency based representations. *Medical engineering & physics*, 21(6-7):431–438.

Enoka, R. M. (1994). *Neuromechanical basis of kinesiology*. Human kinetics Champaign, IL.

Evgeniou, T. and Pontil, M. (2004). Regularized multi–task learning. *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 109–117.

Fang, Y., Zhou, D., Li, K., and Liu, H. (2016). Interface prostheses with classifier-feedback-based user training. *IEEE transactions on biomedical engineering*, 64(11):2575–2583.

Fang, Y., Zhou, D., Li, K., and Liu, H. (2018). ISRMyo-I: A database for sEMG-based hand gesture recognition. *IEEE Dataport*.

Farina, D., Jiang, N., Rehbaum, H., Holobar, A., Graimann, B., Dietl, H., and Aszmann, O. C. (2014). 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*, 22(4):797–809.

Farrell, T. R. and Weir, R. F. (2007). The optimal controller delay for myoelectric prostheses. *IEEE Transactions on neural systems and rehabilitation engineering*, 15(1):111–118.

Feinstein, B., Lindegård, B., Nyman, E., and Wohlfart, G. (1955). Morphologic studies of motor units in normal human muscles. *Cells Tissues Organs*, 23(2):127–142.

Finch, J. (2011). The ancient origins of prosthetic medicine. *The Lancet*, 377(9765):548–549.

Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems. *Ann. Eugen.*, 7(2):179–188.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology*, 47(6):381.

Fitzgibbons, P. and Medvedev, G. (2015). Functional and clinical outcomes of upper extremity amputation. *JAAOS-Journal of the American Academy of Orthopaedic Surgeons*, 23(12):751–760.

Fougner, A., Scheme, E., Chan, A. D., Englehart, K., and Stavdahl, Ø. (2011). Resolving the limb position effect in myoelectric pattern recognition. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 19(6):644–651.

Fougner, A., Stavdahl, O., Kyberd, P. J., Losier, Y. G., and Parker, P. A. (2012). Control of upper limb prostheses: Terminology and proportional myoelectric control—a review. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 20(5):663–677.

Friedman, J. H. (1989). Regularized discriminant analysis. *J. Am. Stat. Assoc.*, 84(405):165.

Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association*, 32(200):675–701.

Fukuda, O., Tsuji, T., Kaneko, M., and Otsuka, A. (2003). A human-assisting manipulator teleoperated by EMG signals and arm motions. *IEEE Transactions on Robotics and Automation*, 19(2):210–222.

Gaballa, A., Cavalcante, R. S., Lamounier, E., Soares, A., and Cabibihan, J.-J. (2022). Extended reality "X-Reality" for prosthesis training of upper-limb amputees: A review on current and future clinical potential. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 30:1652–1663.

Gama, J., Žliobaitė, I., Bifet, A., Pechenizkiy, M., and Bouchachia, A. (2014). A survey on concept drift adaptation. *ACM computing surveys (CSUR)*, 46(4):1–37.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2013). *Bayesian Data Analysis*. CRC press.

Geng, Y., Chen, L., Tian, L., and Li, G. (2012a). Comparison of electromyography and mechanomyogram in control of prosthetic system in multiple limb positions. In *Proceedings of 2012 IEEE-EMBS International Conference on Biomedical and Health Informatics*, pages 788–791. IEEE.

Geng, Y., Samuel, O. W., Wei, Y., and Li, G. (2017). Improving the robustness of real-time myoelectric pattern recognition against arm position changes in transradial amputees. *BioMed research international*, 2017(1):5090454.

Geng, Y., Zhang, F., Yang, L., Zhang, Y., and Li, G. (2012b). Reduction of the effect of arm position variation on real-time performance of motion classification. In *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 2772–2775. IEEE.

Geng, Y., Zhou, P., and Li, G. (2012c). Toward attenuating the impact of arm positions on electromyography pattern-recognition based motion classification in transradial amputees. *Journal of neuroengineering and rehabilitation*, 9:1–11.

Gerdle, B., Karlsson, S., Day, S., and Djupsjöbacka, M. (1999). Acquisition, processing and analysis of the surface electromyogram. In *Modern techniques in neuroscience research*, pages 705–755. Springer.

Germer, C. M., Farina, D., Elias, L. A., Nuccio, S., Hug, F., and Del Vecchio, A. (2021). Surface emg cross talk quantified at the motor unit population level for muscles of the hand, thigh, and calf. *Journal of applied physiology*, 131(2):808–820.

Ghojogh, B. and Crowley, M. (2019). Linear and quadratic discriminant analysis: Tutorial. *arXiv preprint arXiv:1906.02590*.

Giunchiglia, E. and Lukasiewicz, T. (2020). Coherent hierarchical multi-label classification networks. *Advances in neural information processing systems*, 33:9662–9673.

Glynn, M., Galway, H., Hunter, G., and Sauter, W. (1986). Management of the upper-limb-deficient child with a powered prosthetic device. *Clinical Orthopaedics and Related Research®*, 209:202–205.

González, J., Horiuchi, Y., and Yu, W. (2010). Classification of upper limb motions from around-shoulder muscle activities: hand biofeedback. *The open medical informatics journal*, 4:74.

Gordon, J.-S. (2021). Ai and law: ethical, legal, and socio-political implications.

Graupe, D., Kohn, K. H., Kralj, A., and Basseas, S. (1983). Patient controlled electrical stimulation via emg signature discrimination for providing certain paraplegics with primitive walking functions. *Journal of biomedical engineering*, 5(3):220–226.

Gu, Y., Yang, D., Huang, Q., Yang, W., and Liu, H. (2018). Robust emg pattern recognition in the presence of confounding factors: features, classifiers and adaptive learning. *Expert Systems with Applications*, 96:208–217.

Gulrajani, I. and Lopez-Paz, D. (2020). In search of lost domain generalization. *arXiv preprint arXiv:2007.01434*.

Gulrajani, I. and Lopez-Paz, D. (2021). In search of lost domain generalization. In *International Conference on Learning Representations*.

Guo, K., Lu, J., Wu, Y., Hu, X., and Yang, H. (2024). The latest research progress on bionic artificial hands: A systematic review. *Micromachines*, 15(7):891.

Gusman, J., Mastinu, E., and Ortiz-Catalan, M. (2017). Evaluation of computer-based target achievement tests for myoelectric control. *IEEE journal of translational engineering in health and medicine*, 5:1–10.

Hahne, J. M., Dähne, S., Hwang, H.-J., Müller, K.-R., and Parra, L. C. (2015a). Concurrent adaptation of human and machine improves simultaneous and proportional myoelectric control. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 23(4):618–627.

Hahne, J. M., Dähne, S., Hwang, H.-J., Müller, K.-R., and Parra, L. C. (2015b). Concurrent adaptation of human and machine improves simultaneous and proportional myoelectric control. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 23(4):618–627.

Hahne, J. M., Graimann, B., and Muller, K.-R. (2012). Spatial filtering for robust myoelectric control. *IEEE Transactions on Biomedical Engineering*, 59(5):1436–1443.

Hahne, J. M., Wilke, M. A., Koppe, M., Farina, D., and Schilling, A. F. (2020). Longitudinal case study of regression-based hand prosthesis control in daily life. *Frontiers in Neuroscience*, 14:600.

Hargrove, L., Englehart, K., and Hudgins, B. (2006). The effect of electrode displacements on pattern recognition based myoelectric control. In *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 2203–2206. IEEE.

Hargrove, L., Englehart, K., and Hudgins, B. (2008). A training strategy to reduce classification degradation due to electrode displacements in pattern recognition based myoelectric control. *Biomedical signal processing and control*, 3(2):175–180.

Hargrove, L. J., Scheme, E. J., Englehart, K. B., and Hudgins, B. S. (2010). Multiple binary classifications via linear discriminant analysis for improved controllability of a powered prosthesis. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 18(1):49–57.

Hassan, H. F., Abou-Loukh, S. J., and Ibraheem, I. K. (2020). Teleoperated robotic arm movement using electromyography signal with wearable Myo armband. *Journal of King Saud University - Engineering Sciences*, 32(6):378–387.

Hastie, T., Tibshirani, R., Friedman, J., et al. (2009). The elements of statistical learning.

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

He, J., Zhang, D., Sheng, X., and Zhu, X. (2013). Effects of long-term myoelectric signals on pattern recognition. In Lee, J., Lee, M. C., Liu, H., and Ryu, J.-H., editors, *International Conference on Intelligent Robotics and Applications*, pages 396–404. Springer Berlin Heidelberg.

He, J., Zhang, D., Sheng, X., and Zhu, X. (2015b). A comparison of open-loop and closed-loop adaptive calibration for pattern recognition based myoelectric control. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 1144–1147. IEEE.

Hong, J.-y., Suh, E.-h., and Kim, S.-J. (2009). Context-aware systems: A literature review and classification. *Expert Systems with applications*, 36(4):8509–8522.

Horiuchi, Y., Kishi, T., Gonzalez, J., and Yu, W. (2009). A study on classification of upper limb motions from around-shoulder muscle activities. In *2009 IEEE International Conference on Rehabilitation Robotics*, pages 311–315. IEEE.

Hu, X., Song, A., Zeng, H., Wei, Z., Deng, H., and Chen, D. (2023). Bridging human-robot co-adaptation via biofeedback for continuous myoelectric control. *IEEE Robotics and Automation Letters*, 8(12):8573–8580.

Hu, Y., Wong, Y., Wei, W., Du, Y., Kankanhalli, M., and Geng, W. (2018). A novel attention-based hybrid cnn-rnn architecture for semg-based gesture recognition. *PloS one*, 13(10):e0206049.

Huang, G., Zhang, D., Zheng, X., and Zhu, X. (2010). An emg-based handwriting recognition through dynamic time warping. In *2010 annual international conference of the IEEE engineering in medicine and biology*, pages 4902–4905. IEEE.

Huang, J., Lin, M., Fu, J., Sun, Y., and Fang, Q. (2021). An immersive motor imagery training system for post-stroke rehabilitation combining vr and EMG-based real-time feedback. In *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pages 7590–7593.

Huang, Q., Yang, D., Jiang, L., Zhang, H., Liu, H., and Kotani, K. (2017). A novel unsupervised adaptive learning method for long-term electromyography (emg) pattern recognition. *Sensors*, 17(6):1370.

Hudgins, B., Parker, P., and Scott, R. (1991). The recognition of myoelectric patterns for prosthetic limb control. In *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society Volume 13: 1991*, pages 2040–2041.

Igual, C., Igual, J., Hahne, J. M., and Parra, L. C. (2019). Adaptive auto-regressive proportional myoelectric control. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 27(2):314–322.

Ishii, A., Kondo, T., and Yano, S. (2017). Improvement of emg pattern recognition by eliminating posture-dependent components. In *Intelligent Autonomous Systems 14: Proceedings of the 14th International Conference IAS-14 14*, pages 19–30. Springer.

Jabbari, M., Khushaba, R. N., and Nazarpour, K. (2020). Emg-based hand gesture classification with long short-term memory deep recurrent neural networks. In *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pages 3302–3305. IEEE.

Jaber, H. A., Rashid, M. T., and Fortuna, L. (2021). Online myoelectric pattern recognition based on hybrid spatial features. *Biomedical Signal Processing and Control*, 66:102482.

Jiang, L., Cai, Z., Wang, D., and Jiang, S. (2007). Survey of improving k-nearest-neighbor for classification. In *Fourth international conference on fuzzy systems and knowledge discovery (FSKD 2007)*, volume 1, pages 679–683. IEEE.

Jiang, N., Muceli, S., Graimann, B., and Farina, D. (2013a). Effect of arm position on the prediction of kinematics from emg in amputees. *Medical & biological engineering & computing*, 51:143–151.

Jiang, N., Rehbaum, H., Vujaklija, I., Graimann, B., and Farina, D. (2013b). Intuitive, online, simultaneous, and proportional myoelectric control over two degrees-of-freedom in upper limb amputees. *IEEE transactions on neural systems and rehabilitation engineering*, 22(3):501–510.

Kaczmarek, P., Mańkowski, T., and Tomczyński, J. (2019). putEMG - a surface electromyography hand gesture recognition dataset. *Sensors*, 19(16):3548.

Kanitz, G., Cipriani, C., and Edin, B. B. (2018). Classification of transient myoelectric signals for the control of multi-grasp hand prostheses. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 26(9):1756–1764.

Kato, R., Fujita, T., Yokoi, H., and Arai, T. (2006). Adaptable emg prosthetic hand using on-line learning method-investigation of mutual adaptation between human and adaptable machine. In *ROMAN 2006-The 15th IEEE International Symposium on Robot and Human Interactive Communication*, pages 599–604. IEEE.

Kaufmann, P., Englehart, K., and Platzner, M. (2010). Fluctuating emg signals: Investigating long-term effects of pattern matching algorithms. In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, pages 6357–6360. IEEE.

Kazamel, M. and Warren, P. P. (2017). History of electromyography and nerve conduction studies: A tribute to the founding fathers. *Journal of Clinical Neuroscience*, 43:54–60.

Keskinbora, K. H. (2019). Medical ethics considerations on artificial intelligence. *Journal of clinical neuroscience*, 64:277–282.

Khan, S. R. and Desmulliez, M. P. (2019). Implementation of a dual wireless power transfer and rotation monitoring system for prosthetic hands. *IEEE Access*, 7:107616–107625.

Khushaba, R. N. (2014). Correlation analysis of electromyogram signals for multiuser myoelectric interfaces. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 22(4):745–755.

Khushaba, R. N., Al-Timemy, A., and Kodagoda, S. (2015). Influence of multiple dynamic factors on the performance of myoelectric pattern recognition. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 1679–1682. IEEE.

Khushaba, R. N., Al-Timemy, A., Kodagoda, S., and Nazarpour, K. (2016). Combined influence of forearm orientation and muscular contraction on emg pattern recognition. *Expert Systems with Applications*, 61:154–161.

Khushaba, R. N., Krasoulis, A., Al-Jumaily, A., and Nazarpour, K. (2018). Spatio-temporal inertial measurements feature extraction improves hand movement pattern recognition without electromyography. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 2108–2111. IEEE.

Khushaba, R. N., Phinyomark, A., Al-Timemy, A. H., and Scheme, E. (2020). Recursive multi-signal temporal fusions with attention mechanism improves emg feature extraction. *IEEE Transactions on Artificial Intelligence*, 1(2):139–150.

Khushaba, R. N., Shi, L., and Kodagoda, S. (2012). Time-dependent spectral features for limb position invariant myoelectric pattern recognition. In *2012 International Symposium on Communications and Information Technologies (ISCIT)*, pages 1015–1020. IEEE.

Khushaba, R. N., Takruri, M., Miro, J. V., and Kodagoda, S. (2014). Towards limb position invariant myoelectric pattern recognition using time-dependent spectral features. *Neural Networks*, 55:42–58.

Krasoulis, A., Kyranou, I., Erden, M. S., Nazarpour, K., and Vijayakumar, S. (2017). Improved prosthetic hand control with concurrent use of myoelectric and inertial measurements. *Journal of neuroengineering and rehabilitation*, 14:1–14.

Krasoulis, A., Vijayakumar, S., and Nazarpour, K. (2020). Multi-grip classification-based prosthesis control with two EMG-IMU sensors. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 28(2):508–518.

Kulkarni, S. and Rodd, S. F. (2020). Context aware recommendation systems: A review of the state of the art techniques. *Computer Science Review*, 37:100255.

Kumar, S. (2017). Electromyography in ergonomics. In *Electromyography in ergonomics*, pages 1–50. Routledge.

Kwon, Y., Dwivedi, A., McDaid, A. J., and Liarokapis, M. (2021). Electromyography-based decoding of dexterous, in-hand manipulation of objects: Comparing task execution in real world and virtual reality. *IEEE Access*, 9:37297–37310.

Kyranou, I. and Erden, S. V. M. S. (2018). Causes of performance degradation in non-invasive electromyographic pattern recognition in upper limb prostheses. *Frontiers in neurorobotics*, 12:58.

Ladegaard, J. (2002). Story of electromyography equipment. *Muscle & Nerve: Official Journal of the American Association of Electrodiagnostic Medicine*, 25(S11):S128–S133.

Leonardis, D., Barsotti, M., Loconsole, C., Solazzi, M., Troncossi, M., Mazzotti, C., Castelli, V. P., Procopio, C., Lamola, G., Chisari, C., Bergamasco, M., and Frisoli, A. (2015). An EMG-controlled robotic hand exoskeleton for bilateral rehabilitation. *IEEE Transactions on Haptics*, 8(2):140–151.

Les E. Atlas, David A. Cohn, R. E. L. (1990). Training connectionist networks with queries and selective sampling. In *Advances in neural information processing systems*, pages 566–573. Citeseer.

Lewis, D. D. and Gale, W. A. (1994). A sequential algorithm for training text classifiers. In *SIGIR'94*, pages 3–12. Springer.

Li, H., Pan, S. J., Wang, S., and Kot, A. C. (2018). Domain generalization with adversarial feature learning. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE.

Li, X., Chen, S., Zhang, H., Samuel, O. W., Wang, H., Fang, P., Zhang, X., and Li, G. (2016). Towards reducing the impacts of unwanted movements on identification of motion intentions. *Journal of Electromyography and Kinesiology*, 28:90–98.

Li, X., Zhang, X., Chen, X., Chen, X., and Zhang, L. (2023). A unified user-generic framework for myoelectric pattern recognition: Mix-up and adversarial training for domain generalization and adaptation. *IEEE Transactions on Biomedical Engineering*, 70(8):2248–2257.

Liarokapis, M. V., Artemiadis, P. K., Katsiaris, P. T., and Kyriakopoulos, K. J. (2012). Learning task-specific models for reach to grasp movements: Towards emg-based teleoperation of robotic arm-hand systems. In *2012 4th IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob)*, pages 1287–1292. IEEE.

Lim, G. M., Jatesiktat, P., Kuah, C. W. K., and Ang, W. T. (2020). Camera-based hand tracking using a mirror-based multi-view setup. In *2020 42nd annual international conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pages 5789–5793. IEEE.

Liu, J., Sheng, X., Zhang, D., He, J., and Zhu, X. (2014a). Reduced daily recalibration of myoelectric prosthesis classifiers based on domain adaptation. *IEEE journal of biomedical and health informatics*, 20(1):166–176.

Liu, J., Zhang, D., He, J., and Zhu, X. (2012). Effect of dynamic change of arm position on myoelectric pattern recognition. In *2012 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pages 1470–1475. IEEE.

Liu, J., Zhang, D., Sheng, X., and Zhu, X. (2014b). Quantification and solutions of arm movements effect on semg pattern recognition. *Biomedical Signal Processing and Control*, 13:189–197.

Lock, B., Englehart, K., and Hudgins, B. (2005). Real-time myoelectric control in a virtual environment to relate usability vs. accuracy. In *Myoelectric Symposium*, pages 122–127. Citeseer.

Long, M., Cao, Y., Wang, J., and Jordan, M. I. (2015). Learning transferable features with deep adaptation networks. In *International conference on machine learning*, pages 97–105. PMLR.

Loshchilov, I. and Hutter, F. (2017). Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*.

Luo, J., Lin, Z., Li, Y., and Yang, C. (2020). A teleoperation framework for mobile robots based on shared control. *IEEE Robotics and Automation Letters*, 5(2):377–384.

Lyons, K. R. and Margolis, B. W. L. (2019). AxoPy: A Python library for implementing human-computer interface experiments. *Journal of Open Source Software*, 4(34):1191.

Lyu, Z. (2023). State-of-the-art human-computer-interaction in metaverse. *International Journal of Human–Computer Interaction*, pages 1–19.

Madduri, M. M., Yamagami, M., Millevolte, A. X., Li, S. J., Burckhardt, S. N., Burden, S. A., and Orsborn, A. L. (2022). Co-adaptive myoelectric interface for continuous control. *IFAC-PapersOnLine*, 55(41):95–100.

Maduri, P. and Akhondi, H. (2019). Upper limb amputation.

Mao, A., Mohri, M., and Zhong, Y. (2023). Cross-entropy loss functions: Theoretical analysis and applications. In *International conference on Machine learning*, pages 23803–23828. PMLR.

Markovic, M., Dosen, S., Cipriani, C., Popovic, D., and Farina, D. (2014). Stereovision and augmented reality for closed-loop control of grasping in hand prostheses. *Journal of neural engineering*, 11(4):046001.

Masters, M. R., Smith, R. J., Soares, A. B., and Thakor, N. V. (2014). Towards better understanding and reducing the effect of limb position on myoelectric upper-limb prostheses. In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 2577–2580. IEEE.

Mastinu, E., Clemente, F., Sassu, P., Aszmann, O., Brånemark, R., Håkansson, B., Controzzi, M., Cipriani, C., and Ortiz-Catalan, M. (2019). Grip control and motor coordination with implanted and surface electrodes while grasping with an osseointegrated prosthetic hand. *Journal of neuroengineering and rehabilitation*, 16:1–10.

Mathiowetz, V., Volland, G., Kashman, N., and Weber, K. (1985). Adult norms for the box and block test of manual dexterity. *The American journal of occupational therapy*, 39(6):386–391.

Mesin, L. (2020). Crosstalk in surface electromyogram: literature review and some insights. *Physical and Engineering Sciences in Medicine*, 43:481–492.

Moritani, T. and DeVries, H. A. (1978). Reexamination of the relationship between the surface integrated electromyogram (iemg) and force of isometric contraction. *American Journal of Physical Medicine & Rehabilitation*, 57(6):263–277.

Moritani, T. and DeVries, H. A. (1979). Neural factors versus hypertrophy in the time course of muscle strength gain. *American journal of physical medicine & rehabilitation*, 58(3):115–130.

Moritani, T. and Muro, M. (1987). Motor unit activity and surface electromyogram power spectrum during increasing force of contraction. *European journal of applied physiology and occupational physiology*, 56:260–265.

Mourikis, A. I. and Roumeliotis, S. I. (2007). A multi-state constraint kalman filter for vision-aided inertial navigation. In *Proceedings 2007 IEEE international conference on robotics and automation*, pages 3565–3572. IEEE.

Muandet, K., Balduzzi, D., and Schölkopf, B. (2013). Domain generalization via invariant feature representation. In *International conference on machine learning*, pages 10–18. PMLR.

Muceli, S., Jiang, N., and Farina, D. (2013). Extracting signals robust to electrode number and shift for online simultaneous and proportional myoelectric control by factorization algorithms. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 22(3):623–633.

Muceli, S., Poppendieck, W., Negro, F., Yoshida, K., Hoffmann, K. P., Butler, J. E., Gandevia, S. C., and Farina, D. (2015). Accurate and representative decoding of the neural drive to muscles in humans with multi-channel intramuscular thin-film electrodes. *The Journal of physiology*, 593(17):3789–3804.

Nayak, S. and Das, R. K. (2020). Application of artificial intelligence (ai) in prosthetic and orthotic rehabilitation. In *Service Robotics*. intechopen.

Oh, D.-C. and Jo, Y.-U. (2021). Classification of hand gestures based on multi-channel emg by scale average wavelet transform and convolutional neural network. *International Journal of Control, Automation and Systems*, 19(3):1443–1450.

Orfanidis, S. J. (1995). Introduction to signal processing.

Ortiz-Catalan, M., Rouhani, F., Brånemark, R., and Håkansson, B. (2015). Offline accuracy: A potentially misleading metric in myoelectric pattern recognition for prosthetic control. In *2015 37th annual international conference of the IEEE engineering in medicine and biology society (EMBC)*, pages 1140–1143. IEEE.

Osborn, L. E., Moran, C. W., Johannes, M. S., Sutton, E. E., Wormley, J. M., Dohopolski, C., Nordstrom, M. J., Butkus, J. A., Chi, A., Pasquina, P. F., et al. (2021). Extended home use of an advanced osseointegrated prosthetic arm improves function, performance, and control efficiency. *Journal of neural engineering*, 18(2):026020.

Oskoei, M. A. and Hu, H. (2007). Myoelectric control systems—a survey. *Biomedical signal processing and control*, 2(4):275–294.

Østlie, K., Lesjø, I. M., Franklin, R. J., Garfelt, B., Skjeldal, O. H., and Magnus, P. (2012). Prosthesis rejection in acquired major upper-limb amputees: a population-based survey. *Disability and Rehabilitation: Assistive Technology*, 7(4):294–303.

Palermo, F., Cognolato, M., Gijsberts, A., Müller, H., Caputo, B., and Atzori, M. (2017). Repeatability of grasp recognition for robotic hand prosthesis control based on sEMG data. In *2017 International Conference on Rehabilitation Robotics (ICORR)*, pages 1154–1159. IEEE.

Pan, L., Zhang, D., Jiang, N., Sheng, X., and Zhu, X. (2015). Improving robustness against electrode shift of high density emg for myoelectric control through common spatial patterns. *Journal of neuroengineering and rehabilitation*, 12:1–16.

Paré, A. (1728). *Les oeuvres d'Ambroise Paré*. Buon.

Park, K.-H., Suk, H.-I., and Lee, S.-W. (2015). Position-independent decoding of movement intention for proportional myoelectric interfaces. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 24(9):928–939.

Parker, P., Englehart, K., and Hudgins, B. (2006). Myoelectric signal processing for control of powered limb prostheses. *Journal of electromyography and kinesiology*, 16(6):541–548.

Peerdeman, B., Boere, D., Witteveen, H., Hermens, H., Stramigioli, S., Rietman, H., Veltink, P., Misra, S., et al. (2011). Myoelectric forearm prostheses: State of the art from a user-centered perspective. *Journal of rehabilitation research and development*, 48(6):719–738.

Pereira, J., Halatsis, D., Hodossy, B., and Farina, D. (2024). Tackling electrode shift in gesture recognition with hd-emg electrode subsets. In *ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1786–1790. IEEE.

Pereira, M. F., Prahm, C., Kolbenschlag, J., Oliveira, E., and Rodrigues, N. F. (2020). Application of AR and VR in hand rehabilitation: A systematic review. *Journal of Biomedical Informatics*, 111:103584.

Phinyomark, A., N. Khushaba, R., and Scheme, E. (2018). Feature extraction and selection for myoelectric control based on wearable emg sensors. *Sensors*, 18(5):1615.

Phinyomark, A., Phukpattaranont, P., and Limsakul, C. (2012). Feature reduction and selection for emg signal classification. *Expert systems with applications*, 39(8):7420–7431.

Phinyomark, A., Quaine, F., Charbonnier, S., Serviere, C., Tarpin-Bernard, F., and Laurillau, Y. (2013). Emg feature evaluation for improving myoelectric pattern recognition robustness. *Expert Systems with applications*, 40(12):4832–4840.

Phinyomark, A. and Scheme, E. (2018a). Emg pattern recognition in the era of big data and deep learning. *Big Data and Cognitive Computing*, 2(3):21.

Phinyomark, A. and Scheme, E. (2018b). A feature extraction issue for myoelectric control based on wearable emg sensors. In *2018 IEEE Sensors Applications Symposium (SAS)*, pages 1–6. IEEE.

Powell, M. A., Kaliki, R. R., and Thakor, N. V. (2013). User training for pattern recognition-based myoelectric prostheses: Improving phantom limb movement consistency and distinguishability. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 22(3):522–532.

Pradhan, A., He, J., and Jiang, N. (2022). Gesture recognition and biometrics electromyography (GRABMyo) dataset.

Prahm, C., Paassen, B., Schulz, A., Hammer, B., and Aszmann, O. (2017). Transfer learning for rapid re-calibration of a myoelectric prosthesis after electrode shift. In *Converging clinical and engineering research on neurorehabilitation II*, pages 153–157.

Proroković, K., Wand, M., and Schmidhuber, J. (2020). Meta-learning for recalibration of emg-based upper limb prostheses. In *4th Lifelong Machine Learning Workshop at ICML 2020*.

Purves, D. (2012). *Neuroscience*. Oxford University Press.

Qureshi, M. F., Mushtaq, Z., ur Rehman, M. Z., and Kamavuako, E. N. (2022). Spectral image-based multiday surface electromyography classification of hand motions using cnn for human–computer interaction. *IEEE Sensors Journal*, 22(21):20676–20683.

Radhakrishnan, S. M., Baker, S. N., and Jackson, A. (2008). Learning a novel myoelectric-controlled interface task. *Journal of Neurophysiology*, 100(4):2397–2408.

Radmand, A., Scheme, E., and Englehart, K. (2014a). A characterization of the effect of limb position on emg features to guide the development of effective prosthetic control schemes. In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 662–667. IEEE.

Radmand, A., Scheme, E., and Englehart, K. (2014b). A characterization of the effect of limb position on EMG features to guide the development of effective prosthetic control schemes. In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 662–667.

Radmand, A., Scheme, E., and Englehart, K. (2014c). On the suitability of integrating accelerometry data with electromyography signals for resolving the effect of changes in limb position during dynamic limb movement. *JPO: Journal of Prosthetics and Orthotics*, 26(4):185–193.

Rajapriya, R., Rajeswari, K., and Thiruvengadam, S. (2021). Deep learning and machine learning techniques to improve hand movement classification in myoelectric control system. *Biocybernetics and Biomedical Engineering*, 41(2):554–571.

Reaz, M., Hussain, M., and Mohd-Yasin, F. (2006). Techniques of emg signal analysis: detection, processing, classification and applications (correction). *Biological procedures online*, 8:163–163.

Reiter, R. (1948). Eine neue eltronkunsthand. *Grenzbebiete der Medizin*, 4.

Rivela, D., Scannella, A., Pavan, E. E., Frigo, C. A., Belluco, P., and Gini, G. (2015). Processing of surface emg through pattern recognition techniques aimed at classifying shoulder joint movements. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 2107–2110. IEEE.

Roche, A. D., Rehbaum, H., Farina, D., and Aszmann, O. C. (2014). Prosthetic myoelectric control strategies: a clinical perspective. *Current Surgery Reports*, 2:1–11.

Salminger, S., Stino, H., Pichler, L. H., Gstoettner, C., Sturma, A., Mayer, J. A., Szivak, M., and Aszmann, O. C. (2022). Current rates of prosthetic usage in upper-limb amputees–have innovations had an impact on device acceptance? *Disability and Rehabilitation*, 44(14):3708–3713.

Samuel, O. W., Asogbon, M. G., Geng, Y., Al-Timemy, A. H., Pirbhulal, S., Ji, N., Chen, S., Fang, P., and Li, G. (2019). Intelligent emg pattern recognition control method for upper-limb multifunctional prostheses: advances, current challenges, and future prospects. *Ieee Access*, 7:10150–10165.

Scheffer, T., Decomain, C., and Wrobel, S. (2001). Active hidden markov models for information extraction. In *International Symposium on Intelligent Data Analysis*, pages 309–318. Springer Berlin Heidelberg.

Scheme, E. and Englehart, K. (2011). Electromyogram pattern recognition for control of powered upper-limb prostheses: state of the art and challenges for clinical use. *Journal of Rehabilitation Research & Development*, 48(6).

Scheme, E. and Englehart, K. (2013). Training strategies for mitigating the effect of proportional control on classification in pattern recognition–based myoelectric control. *JPO: Journal of Prosthetics and Orthotics*, 25(2):76–83.

Scheme, E., Fougner, A., Stavdahl, Ø., Chan, A. D., and Englehart, K. (2010). Examining the adverse effects of limb position on pattern recognition based myoelectric control. In *2010 annual international conference of the IEEE engineering in medicine and biology*, pages 6337–6340. IEEE.

Scheme, E. J., Hudgins, B. S., and Englehart, K. B. (2013). Confidence-based rejection for improved pattern recognition myoelectric control. *IEEE Transactions on Biomedical Engineering*, 60(6):1563–1570.

Schultz, A. E., Baade, S. P., and Kuiken, T. A. (2007). Expert opinions on success factors for upper-limb prostheses. *Journal of Rehabilitation Research and Development*, 44(4):483.

Segil, J., Kaliki, R., Uellendahl, J., and Weir, R. F. (2020). A myoelectric postural control algorithm for persons with transradial amputation: A consideration of clinical readiness. *IEEE Robotics and Automation Magazine*.

Segil, J. L. and Weir, R. F. (2015). Novel postural control algorithm for control of multifunctional myoelectric prosthetic hands. *Journal of Rehabilitation Research and Development*, 52(4):449–466.

Sensinger, J. W., Lock, B. A., and Kuiken, T. A. (2009). Adaptive pattern recognition of myoelectric signals: exploration of conceptual framework and practical algorithms. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 17(3):270–278.

Settles, B. (2009). Active learning literature survey. Technical report, University of Wisconsin-Madison Department of Computer Sciences.

Settles, B. and Craven, M. (2008). An analysis of active learning strategies for sequence labeling tasks. In *proceedings of the 2008 conference on empirical methods in natural language processing*, pages 1070–1079.

Shahsavari, H., Matourypour, P., Ghiyasvandian, S., Ghorbani, A., Bakhshi, F., Mahmoudi, M., and Golestannejad, M. (2020). Upper limb amputation; care needs for reintegration to life: An integrative review. *International journal of orthopaedic and trauma nursing*, 38:100773.

Shahzad, W., Ayaz, Y., Khan, M. J., Naseer, N., and Khan, M. (2019). Enhanced performance for multi-forearm movement decoding using hybrid imu–semg interface. *Frontiers in neurorobotics*, 13:43.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, 27(3):379–423.

Shen, S., Li, M., Mao, F., Chen, X., and Ran, R. (2024). Gesture recognition using mlp-mixer with cnn and stacking ensemble for semg signals. *IEEE Sensors Journal*, 24(4):4960–4968.

Sherman, E. D. (1964). A russian bioelectric-controlled prosthesis: Report of a research team from the rehabilitation institute of montreal. *Canadian Medical Association Journal*, 91(24):1268.

Simon, A. M., Hargrove, L. J., Lock, B. A., and Kuiken, T. A. (2011). The target achievement control test: Evaluating real-time myoelectric pattern recognition control of a multifunctional upper-limb prosthesis. *Journal of rehabilitation research and development*, 48(6):619.

Smail, L. C., Neal, C., Wilkins, C., and Packham, T. L. (2021). Comfort and function remain key factors in upper limb prosthetic abandonment: findings of a scoping review. *Disability and rehabilitation: Assistive technology*, 16(8):821–830.

Stahl, B. C. (2021). *Artificial intelligence for a better future: an ecosystem perspective on the ethics of AI and emerging digital technologies*. Springer Nature.

Stango, A., Negro, F., and Farina, D. (2014). 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*, 23(2):189–198.

Stuttaford, S. A., Dyson, M., Nazarpour, K., and Dupan, S. S. G. (2024). Reducing motor variability enhances myoelectric control robustness across untrained limb positions. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 32:23–32.

Sun, B. and Saenko, K. (2016). Deep CORAL: Correlation alignment for deep domain adaptation. In *Lecture Notes in Computer Science*, Lecture notes in computer science, pages 443–450. Springer International Publishing, Cham.

Szymaniak, K., Krasoulis, A., and Nazarpour, K. (2022). Recalibration of myoelectric control with active learning. *Frontiers in Neurorobotics*, 16:1061201.

Szymaniak, K., Kyranou, I., and Nazarpour, K. (2024). It's great: Gesture recognition for arm translation.

Szymaniak, K., Kyranou, I., and Nazarpour, K. (2025). Emg dataset for gesture recognition with arm translation. *Scientific Data*, 12(1):100.

Tabor, A., Bateman, S., and Scheme, E. (2018). Evaluation of myoelectric control learning using multi-session game-based training. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 26(9):1680–1689.

Tam, S., Raghu, S. T. P., Buteau, É., Scheme, E., Boukadoum, M., Campeau-Lecours, A., and Gosselin, B. (2024). Towards robust and interpretable emg-based hand gesture recognition using deep metric meta learning. *arXiv preprint arXiv:2404.15360*.

Tchimino, J., Markovic, M., Dideriksen, J. L., and Dosen, S. (2021). The effect of calibration parameters on the control of a myoelectric hand prosthesis using emg feedback. *Journal of Neural Engineering*, 18(4):046091.

Teh, Y. and Hargrove, L. J. (2020). Understanding limb position and external load effects on real-time pattern recognition control in amputees. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 28(7):1605–1613.

Toledo-Peral, C. L., Vega-Martínez, G., Mercado-Gutiérrez, J. A., Rodríguez-Reyes, G., Vera-Hernández, A., Leija-Salas, L., and Gutiérrez-Martínez, J. (2022). Virtual/augmented reality for rehabilitation applications using electromyography as control/biofeedback: systematic literature review. *Electronics*, 11(14):2271.

Tommasi, T., Orabona, F., Castellini, C., and Caputo, B. (2012). Improving control of dexterous hand prostheses using adaptive learning. *IEEE Transactions on Robotics*, 29(1):207–219.

Topini, A., Sansom, W., Secciani, N., Bartalucci, L., Ridolfi, A., and Allotta, B. (2022). Variable admittance control of a hand exoskeleton for virtual reality-based rehabilitation tasks. *Frontiers in neurorobotics*, 15:789743.

Trent, L., Intintoli, M., Prigge, P., Bollinger, C., Walters, L. S., Conyers, D., Miguelez, J., and Ryan, T. (2020). A narrative review: current upper limb prosthetic options and design. *Disability and Rehabilitation: Assistive Technology*.

Tripathi, A., Prathosh, A., Muthukrishnan, S. P., and Kumar, L. (2023). Surfmyoair: A surface electromyography-based framework for airwriting recognition. *IEEE Transactions on Instrumentation and Measurement*, 72:1–12.

Tsinganos, P., Cornelis, B., Cornelis, J., Jansen, B., and Skodras, A. (2019). Improved gesture recognition based on semg signals and tcn. In *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1169–1173. IEEE.

Tzeng, E., Hoffman, J., Zhang, N., Saenko, K., and Darrell, T. (2014). Deep domain confusion: Maximizing for domain invariance. *arXiv preprint arXiv:1412.3474*.

Uchida, T. K. and Delp, S. L. (2021). *Biomechanics of movement: the science of sports, robotics, and rehabilitation*. Mit Press.

Urra, O., Casals, A., and Jané, R. (2014). Evaluating spatial characteristics of upper-limb movements from emg signals. In *XIII Mediterranean Conference on Medical and Biological Engineering and Computing 2013: MEDICON 2013, 25-28 September 2013, Seville, Spain*, pages 1795–1798. Springer.

Van Der Niet, O. and van der Sluis, C. K. (2013). Functionality of i-limb and i-limb pulse hands: Case report. *Journal of rehabilitation research and development*, 50(8):1123.

Vidovic, M. M.-C., Hwang, H.-J., Amsüss, S., Hahne, J. M., Farina, D., and Müller, K.-R. (2015). Improving the robustness of myoelectric pattern recognition for upper limb prostheses by covariate shift adaptation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 24(9):961–970.

Vidovic, M. M.-C., Hwang, H.-J., Amsüss, S., Hahne, J. M., Farina, D., and Müller, K.-R. (2016). Improving the robustness of myoelectric pattern recognition for upper limb prostheses by covariate shift adaptation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 24(9):961–970.

Vincent Systems (2022). VINCENTevolution 4. [Online; accessed 21-02-2022].

Vujaklija, I., Roche, A. D., Hasenoehrl, T., Sturma, A., Amsuess, S., Farina, D., and Aszmann, O. C. (2017). Translating research on myoelectric control into clinics—are the performance assessment methods adequate? *Frontiers in neurorobotics*, 11:7.

Wang, F., Ao, X., Wu, M., Kawata, S., and She, J. (2024). Explainable deep learning for semg-based similar gesture recognition: A shapley-value-based solution. *Information Sciences*, 672:120667.

Wang, J., Lan, C., Liu, C., Ouyang, Y., Qin, T., Lu, W., Chen, Y., Zeng, W., and Philip, S. Y. (2022). Generalizing to unseen domains: A survey on domain generalization. *IEEE transactions on knowledge and data engineering*, 35(8):8052–8072.

Waris, A., Niazi, I. K., Jamil, M., Gilani, O., Englehart, K., Jensen, W., Shafique, M., and Kamavuako, E. N. (2018). The effect of time on EMG classification of hand motions in able-bodied and transradial amputees. *Journal of Electromyography and Kinesiology*, 40:72–80.

Wattanasiri, P., Wilson, S., Huo, W., and Vaidyanathan, R. (2024). Gesture recognition through mechanomyogram signals: An adaptive framework for arm posture variability. *IEEE Journal of Biomedical and Health Informatics*.

Wei, W., Wong, Y., Du, Y., Hu, Y., Kankanhalli, M., and Geng, W. (2019). A multi-stream convolutional neural network for semg-based gesture recognition in muscle-computer interface. *Pattern Recognition Letters*, 119:131–138.

Weiner, P., Starke, J., Rader, S., Hundhausen, F., and Asfour, T. (2022). Designing prosthetic hands with embodied intelligence: The kit prosthetic hands. *Frontiers in neurorobotics*, 16:815716.

Williams, H. E., Shehata, A. W., Dawson, M. R., Scheme, E., Hebert, J. S., and Pilarski, P. M. (2022). Recurrent convolutional neural networks as an approach to position-aware myoelectric prosthesis control. *IEEE Transactions on Biomedical Engineering*, 69(7):2243–2255.

Winter, D., Fuglevand, A. J., and Archer, S. (1994). Crosstalk in surface electromyography: theoretical and practical estimates. *Journal of Electromyography and Kinesiology*, 4(1):15–26.

Wu, D., Yang, J., and Sawan, M. (2023). Transfer learning on electromyography (emg) tasks: Approaches and beyond. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 31:3015–3034.

Xiong, D., Zhang, D., Zhao, X., and Zhao, Y. (2021). Deep learning for emg-based human-machine interaction: A review. *IEEE/CAA Journal of Automatica Sinica*, 8(3):512–533.

Xu, M., Zhang, J., Ni, B., Li, T., Wang, C., Tian, Q., and Zhang, W. (2020). Adversarial domain adaptation with domain mixup. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, pages 6502–6509.

Yamanoi, Y., Ogiri, Y., and Kato, R. (2020). Emg-based posture classification using a convolutional neural network for a myoelectric hand. *Biomedical Signal Processing and Control*, 55:101574.

Yang, D., Gu, Y., Jiang, L., Osborn, L., and Liu, H. (2017). Dynamic training protocol improves the robustness of pr-based myoelectric control. *Biomedical Signal Processing and Control*, 31:249–256.

Yang, D., Yang, W., Huang, Q., and Liu, H. (2015). Classification of multiple finger motions during dynamic upper limb movements. *IEEE journal of biomedical and health informatics*, 21(1):134–141.

Yang, J., Soh, M., Lieu, V., Weber, D., and Erickson, Z. (2024). Emgbench: Benchmarking out-of-distribution generalization and adaptation for electromyography. *Advances in Neural Information Processing Systems*, 37:50313–50342.

Yatsenko, D., McDonnall, D., and Guillory, K. S. (2007). Simultaneous, proportional, multi-axis prosthesis control using multichannel surface emg. In *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 6133–6136. IEEE.

You, K.-J., Rhee, K.-W., and Shin, H.-C. (2010). Finger motion decoding using emg signals corresponding various arm postures. *Experimental neurobiology*, 19(1):54.

Young, A. J., Hargrove, L. J., and Kuiken, T. A. (2011). The effects of electrode size and orientation on the sensitivity of myoelectric pattern recognition systems to electrode shift. *IEEE transactions on biomedical engineering*, 58(9):2537–2544.

Young, A. J., Smith, L. H., Rouse, E. J., and Hargrove, L. J. (2012). Classification of simultaneous movements using surface EMG pattern recognition. *IEEE Transactions on Biomedical Engineering*, 60(5):1250–1258.

Yu, Y., Sheng, X., Guo, W., and Zhu, X. (2017). Attenuating the impact of limb position on surface emg pattern recognition using a mixed-lda classifier. In *2017 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pages 1497–1502. IEEE.

Zanghieri, M., Benatti, S., Burrello, A., Kartsch, V., Conti, F., and Benini, L. (2019). Robust real-time embedded emg recognition framework using temporal convolutional networks on a multicore iot processor. *IEEE transactions on biomedical circuits and systems*, 14(2):244–256.

Zardoshti-Kermani, M., Wheeler, B. C., Badie, K., and Hashemi, R. M. (1995). Emg feature evaluation for movement control of upper extremity prostheses. *IEEE Transactions on Rehabilitation Engineering*, 3(4):324–333.

Zhai, X., Jelfs, B., Chan, R. H., and Tin, C. (2017). Self-recalibrating surface EMG pattern recognition for neuroprosthesis control based on convolutional neural network. *Frontiers in Neuroscience*, 11:379.

Zhang, H., Qu, H., Teng, L., and Tang, C.-Y. (2023a). Lstm-msa: A novel deep learning model with dual-stage attention mechanisms forearm emg-based hand gesture recognition. *IEEE transactions on neural systems and rehabilitation engineering*, 31:4749–4759.

Zhang, X., Wu, L., Zhang, X., Chen, X., Li, C., and Chen, X. (2023b). Multi-source domain generalization and adaptation toward cross-subject myoelectric pattern recognition. *Journal of Neural Engineering*, 20(1):016050.

Zhao, K., Zhang, Z., Wen, H., and Scano, A. (2021). Intra-subject and inter-subject movement variability quantified with muscle synergies in upper-limb reaching movements. *Biomimetics*, 6(4):63.

Zheng, N., Li, Y., Zhang, W., and Du, M. (2022). User-independent emg gesture recognition method based on adaptive learning. *Frontiers in Neuroscience*, 16:847180.

Zhu, B., Zhang, D., Chu, Y., Gu, Y., and Zhao, X. (2022). SeNic: An open source dataset for sEMG-based gesture recognition in non-ideal conditions. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 30:1252–1260.

Ziegler-Graham, K., MacKenzie, E. J., Ephraim, P. L., Travison, T. G., and Brookmeyer, R. (2008). Estimating the prevalence of limb loss in the united states: 2005 to 2050. *Archives of physical medicine and rehabilitation*, 89(3):422–429.

Zou, Y. and Cheng, L. (2021). A transfer learning model for gesture recognition based on the deep features extracted by cnn. *IEEE Transactions on Artificial Intelligence*, 2(5):447–458.

Zuo, K. J. and Olson, J. L. (2014). The evolution of functional hand replacement: From iron prostheses to hand transplantation. *Plastic Surgery*, 22(1):44–51.