Prosthetic Hand Control: Phase-based Grasping Pattern Recognition using sEMG and Computer Vision

by

© **Shuo Wang**

A thesis submitted to the School of Graduate Studies in partial fulfillment of the requirements for the degree of Master of Science.

Department of Computer Science

Memorial University

December 2022

St. John's, Newfoundland and Labrador, Canada

Abstract

Pattern recognition using surface Electromyography (sEMG) applied on prosthesis control has attracted much attention. The strong relationship between visual perception and hand manipulation makes vision play an essential role in prosthetic hand control. Utilizing both sEMG and visual information to improve prosthetic hand control became a promising research direction. In most existing hand grasping classification research using sEMG, the signals collected during the firmly grasped period were used for classification because stable signals facilitated classification performance. However, using signals collected from the firm grasp period may cause a delay in controlling the prosthetic hand. Targeting this issue, we explored a new way for grasp classification using signals collected before the firm grasp. We examined accuracy changes during the reaching and grasping process and identified an sEMG sweet period, starts at 1100 ms and ends at 1400 ms in the early grasping phase, that can leverage the grasp classification accuracy for the earlier grasp detection. Although Surface Electromyography (sEMG) achieved a feasible solution in a laboratory environment, the classification accuracy is not high enough for real-time application. Researchers proposed integrating sEMG signals with another feature not affected by amputation. The muscular coordination between vision and hand manipulation makes us consider including the visual information in prosthetic hand control. In this study, we identify another sweet period, starts at 0 ms and ends at 320 ms during the early reaching phase, in which the

vision data could better classify the grasp patterns. Moreover, the visual classification results from the sweet period could be naturally integrated with sEMG data collected during the grasp phase. After the integration, the accuracy of grasp classification increased from 85.5% (only sEMG) to 90.06% (integrated). Knowledge gained from this study encourages us to further explore the methods for incorporating computer vision into myoelectric data to enhance the movement control of prosthetic hands.

Acknowledgements

I want to express my sincere thanks to my supervisors, Dr. Xianta Jiang and Dr. Bin Zheng, for providing me with the opportunity to be involved in machine learning for human activity detection. My detailed discussions with them and their encouragement of innovative ideas and critical thinking were essential for guiding me to be an independent researcher.

I also want to acknowledge the financial support from the Ubiquitous Computing and Machine Learning Research Lab of Memorial University, Department of Computer Science of Memorial University, School of Graduate Study of Memorial University and Surgical Simulation Research Lab of Alberta University.

Additionally, I would like to give my special thanks to my colleagues Mr. Owen Combden, Mr. Mohammad Asfour, Dr. Jingjing Zheng and Dr. Wanglong Lu, for providing me with professional suggestions about programming, data processing and math derivations. Moreover, I would like to thank my girlfriend, Ellie, for providing mental support during my entire master's program. Finally, I also appreciate all the reviews and suggestions from the examiners.

Statement of contributions

Contribution 1 and 3 relate to publication 1, contribution 2 relates to publication 2 and 3.

1. Found the sEMG sweet period, which can significantly reduce the delay of prosthetic hand control in real-life conditions with a slightly increased classification accuracy.
2. Found the vision sweet period, suitable for integration with sEMG. Found the best integration strategies for vision and sEMG. The integration results further increased the classification accuracy, thus improving the prosthesis performance.
3. Found the best training and testing strategies for prosthetic grasp classification using the sEMG sweet period.

Related Publications:

1. Wang, Shuo, Jingjing Zheng, Bin Zheng, and Xianta Jiang. 2022. "Phase-Based Grasp Classification for Prosthetic Hand Control Using SEMG." Biosensors 12 (2): 57. https://doi.org/10.3390/bios12020057.

Author Contributions: Conceptualization, S.W., X.J.; methodology, S.W., J.Z., X.J.; software, S.W.; validation, S.W., J.Z., B.Z., X.J.; formal analysis, S.W., J.Z., X.J.; investigation, S.W., J.Z.; resources, S.W., X.J.; data curation, S.W.; writing—original

draft preparation, S.W., X.J.; writing—review and editing, J.Z., X.J., B.Z., S.W.; visualization, S.W.; supervision, X.J., B.Z.; project administration, X.J., B.Z,; funding acquisition, X.J., B.Z.

2. Wang, Shuo, Jingjing Zheng, Ziwei Huang, Xiaoqin Zhang, Vinicius Prado Da Fonseca, Bin Zheng, and Xianta Jiang. "Integrating Computer Vision to Prosthetic Hand Control with sEMG: Preliminary Results in Grasp Classification." Frontiers in Robotics and AI: 271.

Author Contributions: Conceptualization, S.W., X.J.; methodology, S.W., J.Z., Z.H., X.J.; software, S.W.; validation, S.W., J.Z., B.Z., X.J., V.P., X.Z.; formal analysis, S.W., X.J.; investigation, S.W., J.Z.; resources, S.W., X.J.; data curation, S.W.; writing—original draft preparation, S.W., J.Z.; writing—review and editing, J.Z., V.P., X.J., B.Z., X.Z., S.W.; visualization, S.W.; supervision, X.J., B.Z., V.P.; project administration, X.J., B.Z,; funding acquisition, X.J., B.Z.

3. Zhang, Xiaoqin, Ziwei Huang, Jingjing Zheng, Shuo Wang, and Xianta Jiang. 2022. "DcnnGrasp: Towards Accurate Grasp Pattern Recognition with Adaptive Regularizer Learning." ArXiv:2205.05218 [Cs], May. http://arxiv.org/abs/2205.05218.

Author Contributions: Conceptualization, X.Z., Z.H., J.Z., X.J.; methodology, X.Z., Z.H., J.Z., S.W.; software, X.Z., Z.H.; validation, X.Z., Z.H., J.Z., X.J.; formal analysis, X.Z., Z.H., J.Z.; investigation, X.Z., Z.H., J.Z.; resources, X.Z., Z.H., J.Z., X.J., S.W.; data curation, X.Z., Z.H., S.W.; writing—original draft preparation, Z.H., J.Z., S.W.; writing—review and editing, X.Z., X.J. ; visualization, X.Z., Z.H., J.Z.; supervision, X.Z., X.J.; project administration, X.Z., X.J.; funding acquisition, X.Z., X.J.

Table of contents

List of tables

List of figures

Chapter 1

Introduction

Hands are one of the essential tools for humans to achieve a wide variety of manipulations. The loss of hands can devastate a person, depriving them of their ability to study, work or even live a daily life [64]. After amputation, people may have to change their careers or stay unemployed, leading to more severe problems such as social isolation [43, 12].

Amputees often choose to wear non-invasive prosthetic hands to restore their fundamental abilities and increase their independence in their daily life. There are three types of non-invasive prosthetic hands: cosmetic hand, body-powered hand, and myoelectric controlled hand [53, 15], as shown in Figure 1.1. The first type is the passive prosthetic hand, which is only for decoration; the rest are active prosthetic hands that can move and implement specific hand gestures [13]. The body-powered device comprises a mechanical structure that allows the wearer to open and close the hand by tightening or releasing the wire. The myoelectric prosthesis can achieve a more realistic simulation of the grasping process than body-powered hands, providing a better

user experience [53, 15]. Nowadays, an advanced myoelectric prosthesis is actuated by a classifier, which converts muscular signals into corresponding grasp gestures. The surface electromyography (sEMG) sensors on the upper limb collect the muscular signals.

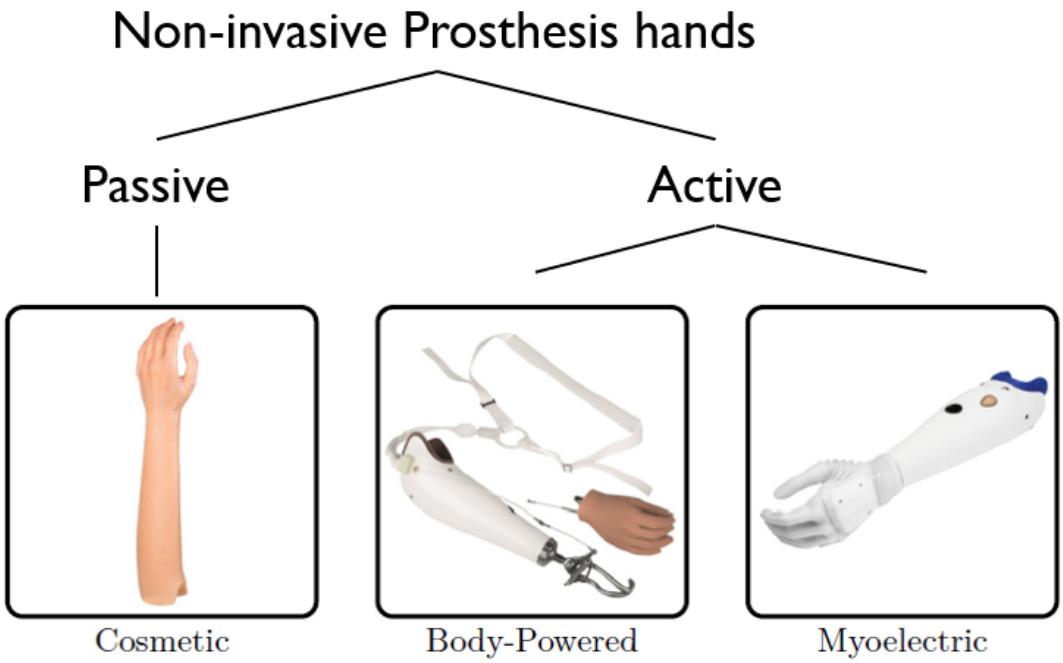

Figure 1.1: Examples of three types of non-invasive prosthesis hands. The cosmetic hand is passive type. Body-powered and myoelectric hand are active type [32].

In recent research, pattern recognition was widely used to recognize hand gestures by analyzing and classifying the muscle sEMG signals, which makes myoelectric prosthetic hands promising for precise hand movement control. However, research by Carey et al. showed that body-powered prosthetic hands are more practical [13] and

have less production cost and learning cost in real-life conditions [69, 39]. Most research on myoelectric prosthesis did not provide enough technical support for effective application improvement in the clinical and real-life environment [25, 71, 68, 74]. The major issue with myoelectric prosthetic hands is that sEMG signals are hard to be decoded to an appropriate level, and the control process is time-consuming [4, 14].

One of the control problems is the delay between the onset of the movement of an amputee and the start of prosthetic hand manipulation. The root cause for the delay is the use of sEMG signals collected during the firm grasp period for classification; sEMG signals in this period are relatively stable and can yield a better classification outcome. However, the firm grasp occurs late in the grasping process [83]. In this work, we consider using the sEMG signal captured in the early grasping period to reduce the delay.

Another control problem is how to address the variability in the sEMG signals. The variation of sEMG signals not only comes from the muscle condition of amputees [24], such as muscle fatigue, but is also affected by other factors, such as displacement of the electrodes and sweat [85, 14, 78]. To solve this problem and further improve pattern recognition performance, researchers have proposed several strategies [14, 25, 54, 37]. These strategies integrated sEMG signals with another feature, such as visual information, which is not affected by amputation. A better classification result can be obtained after the incorporation of other features.

1.1 Improve Prosthetic Hand Control by Phase-based sEMG Analysis

Interpretation of muscle signals is essential for controlling electric-powered prosthetic hands, which requires machine learning algorithms to classify muscular electric signals into corresponding hand movement patterns. When signals during the whole grasp period (including Reaching, Early Grasping and Firm Grasping) were used, the accuracy was not high enough to control the prosthetic hands. In Cognolato et al.'s report [19], the classification accuracy for ten grasp gestures was approximately 63% to 82% using the sEMG signals during the whole grasp period.

Therefore, in most of the published papers, myoelectric signals recorded during firmly grasped periods for grasp classification were used, which yielded satisfactory classification outcomes [45, 44, 3, 18, 17, 19]. For instance, the research done by Jiang et al. [45] using 3 s firm grasp sEMG signals achieved approximately 85% accuracy for classifying 16 grasp gestures. For more examples, please see Table 1.1. However, the firmly grasped periods occur at the end of reaching and grasping, giving no time to control arm movement in a real-life environment [81].

Table 1.1: Examples of previous studies for myoelectric signal classification.

Researcher and Year	**Signal Type**	**Gesture Amount**	**Accuracy**
Chen et al., 2007	sEMG	5 wrist gestures	93.5%
Chen et al., 2007	sEMG	6 wrist motions	88%
Jiang et al., 2017	sEMG	48 hand gestures	84.6%
Jiang et al., 2018	FMG	16 hand gestures	82%
Cognolato et al., 2020	sEMG	10 hand gestures	82.46%
Asfour et al., 2021	FMG	16 hand gestures	86.4%

To solve this problem, developing a method to classify grasp patterns using sEMG data recorded in the earlier grasp period with high accuracy is necessary. This study investigates how the grasp classification accuracy changes over the entire reaching and grasping process and identifies a period in the early grasp phase that can achieve the best classification outcome. We call this period as *sEMG sweet period*. Once the sweet period is identified, we can develop a better classification strategy used in the real-time environment.

Specifically, we first apply and compare several processing methods for the feature extraction of the sEMG signals. Then, we design an experiment to find the sEMG sweet period suitable for early grasp classification with the best classification outcome. Finally, we will conduct another experiment to compare several typical training and

testing strategies to identify an effective strategy for better real-time grasp classification.

1.2 Improve Prosthetic Hand Control by Adding Vision Information

In humans, vision is critical in performing hand gestures before the activity and guiding the activity itself. Humans use visual information to understand and predict future actions [46]. Moreover, in the study of Hebert et al. [34], it has been found that the visual interaction of amputees is more active than intact subjects. The strong relationship between vision and action makes integrating vision and muscle signals a promising prospect.

Some researchers have integrated visual information with a myoelectric prosthesis to improve their performance [33, 55, 56, 30]. In these experiments, the subjects often wore an eye-tracking device, which could also record the first-person video using the integrated camera. The main idea behind these studies is to identify objects to be grasped in the video and then select the corresponding grasp gestures. However, the subjects were asked to stare at the object [10, 77, 34] or manually take a photo [30] until it was recognized and then grasped it. In these cases, the visual information is obtained by established rules that the subject must follow, such as staring at the object for at least 3 seconds, which is not a natural way to perform the grasp action.

We explore how the grasp classification accuracy changes over the entire grasping

process while identifying a period that can achieve the best grasp classification outcome using visual data. We call this interval *vision sweet period*. The sweet period should also be short and located in the early phases to speed up the control process. Once the sweet period is identified, grasp classification by the camera can be automatically conducted during this interval without purposed confirmation. As mentioned previously, in Section 1.1, a similar sweet period (for sEMG) right before the hand grasps the target object was identified for hand grasp classification using sEMG. It will be interesting to explore the vision sweet period again during the reach-and-grasp process and utilize both the vision and sEMG sweet periods for better prosthetic hand control.

In order to achieve the above analysis, we design an experiment to analyze the vision performance and find the vision sweet period with the best grasp type classification outcome. We first extract object photos from the original dataset to build a new dataset. Then we fed a sequence of object images during the reach-and-grasp process to a deep learning model and output classified grasp types. The grasp classification accuracy and the ratio of the number of images containing objects to the total images are analyzed along the whole reach-and-grasp process to identify the vision sweet period. Finally, we integrated sEMG and vision classification outcomes to identify a better classification strategy.

1.3 Hypotheses

We hypothesize that the muscle activities recorded in the early period of the hand grasping process can provide sufficient information to achieve the same or higher accuracy of grasp classification with a reduced delay for prosthetic hand control. We also

hypothesize that the vision sweet period is at the beginning of the reaching phase when the target object has a higher probability of being visible to the participant. We also hypothesize that the integration can provide higher accuracy of grasp classification.

1.4 Thesis Organization

In Chapter 1, we have introduced the two main research topics, including the research questions, hypotheses, and the summary of our work.

In the second chapter, we will make a general introduction on related topics on which the work is based. The first content will be the introduction of surface electromyography and its application for a myoelectric prosthetic hand, followed by the current research status of gesture recognition by computer vision technologies. Subsequently, we will introduce the general mechanism of machine learning and its utilization in the related fields.

In Chapter 3, we will introduce the methodology of this research, including the grasp phases segmentation, eleven sEMG filters, sEMG signal feature engineering, sEMG and vision classification model. In addition, we will introduce the object detection model.

In Chapter 4, we will first introduce the data collection and pre-processing. Then we will provide the details of the experiments, analysis and results, which will start with the sEMG part and be followed by the vision part.

In Chapter 5, we will discuss the findings and possible reasons behind the experiment results and explain the contribution of this research. Finally, the conclusion and future work will also be presented.

Chapter 2

Background

The research of this thesis is multi-disciplinary, including the topics of machine learning, kinesiology of forearm muscle and computer vision. This chapter aims to provide the concepts and mechanisms behind the related research. In the first section, we introduced the EMG signals and the signal acquisition technologies. Then we explained the mechanism for controlling the myoelectric prosthetic hands by EMG signals. Moreover, we introduced object and gesture classification by computer vision technologies and the research status of this field. In recent years, pattern recognition for prosthesis control has been highly developed by many machine learning methods. Therefore, in this section, we also provide the general concepts of machine learning and its application to prosthesis control.

2.1 Surface Electromyography for Myoelectric prosthesis

The human motion system is a complex system composed of control and implementation units, which work together to produce human motions [70, 5]. Hand grasp begins with producing electrical signals in the brain. These signals travel to skeletal muscles through the nervous system and cause the contraction of muscles, resulting in hand and finger movement. The Motor unit action potentials (MUAPs) are the fundamental components of muscle electrical signals, which can be obtained by sensors invasively or non-invasively [62], as shown in Figure 2.1.

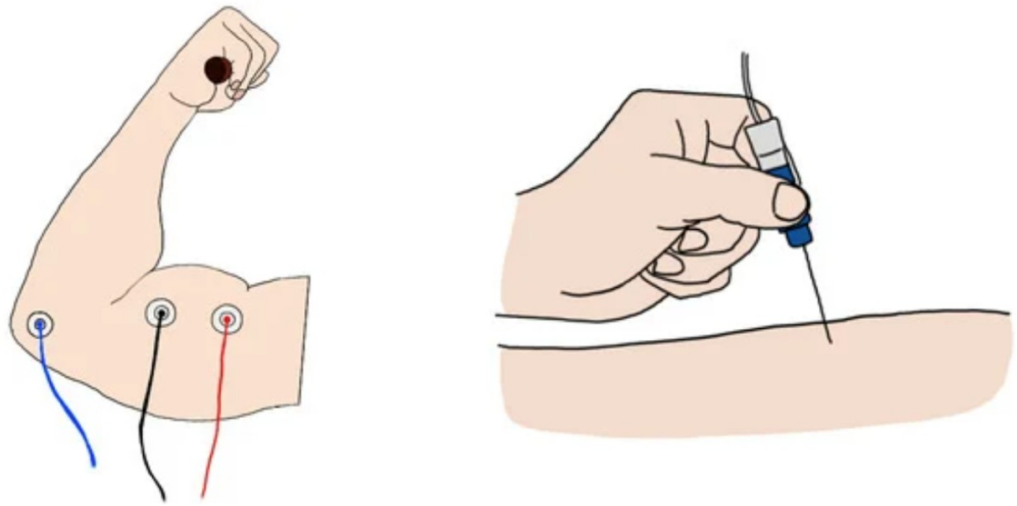

Figure 2.1: Illustrations of invasive and non-invasive EMG acquirement methods [62]. In (a) the sensor is attached on the surface of the skin. In (b) the sensor is placed into the muscle by a needle.

Invasive EMG sensors can acquire more localized and accurate EMG signals because the sensor is placed in the deep muscles by the needle, significantly reducing the distance between the muscle and the sensor. For example, EMG can provide accurate information in the research of investigating muscle pain [16, 75]. However, sEMG cannot be used in that study because it cannot be accurate enough as an invasive EMG. However, sEMG is used more widely due to its stability and convenience because most researchers do not require such high precision that invasive EMG can provide. Although precision limits its utilization in the specific field, sEMG is promising and widely used in wearable devices. Several studies have proposed strategies for where to place the sEMG sensors to detect the specific muscle activity [31, 21].

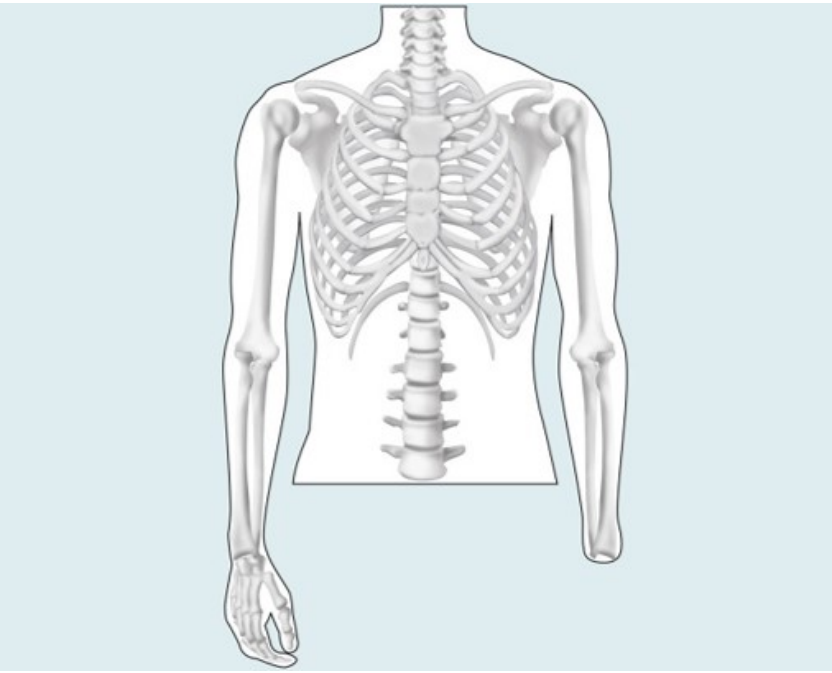

Figure 2.2: Example of a upper-limb amputee. The upper-limb amputee lost their hand and the area close to the hand, most muscles on the forearm still remain [1].

As shown in Figure 2.2, most upper-limb amputees still have muscles on their forearms. When these residual muscles contract, we can obtain sEMG signals from muscles on their forearms [29]. Most amputees retain muscle memory for different grasping actions, making it possible for us to detect their movement intentions from sEMG signals and translate them into grasping gesture commands [65, 76].

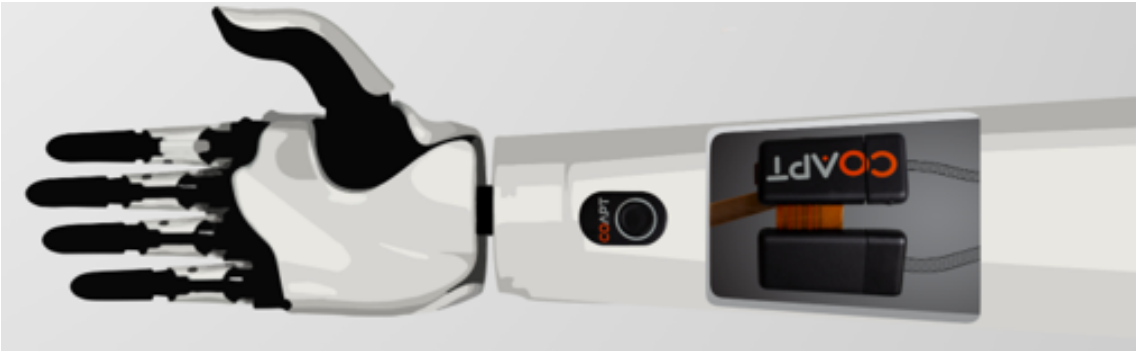

Figure 2.3: Example of a myoelectric prosthetic hand with two sensors from the company COAPT [2].

In the early studies, only two sensors were placed on the surface of the forearm muscle, as shown in Figure 2.3. The number of signals obtained was minimal; therefore, the muscle activity was divided into contraction and relaxation. Through the analysis of signal amplitudes, threshold values of muscle contraction and relaxation were found and used by Zecca et al. to control the switch of myoelectric prosthetic hands [87]. When the muscles contract, the signal amplitude increases and myoelectric prosthetic hands close after exceeding the threshold value. When the muscles relax, the signal

amplitude returns to a lower level below the threshold value, and the myoelectric prosthetic hands open. In a subsequent study, different grasping gestures were introduced into the function of myoelectric prosthetic hands. In the systems designed by Belter et al., Van et al. and Mastinu et al., the user is allowed to switch the grasping gesture manually and perform hand open and close by using the threshold value method mentioned above [7, 82, 58].

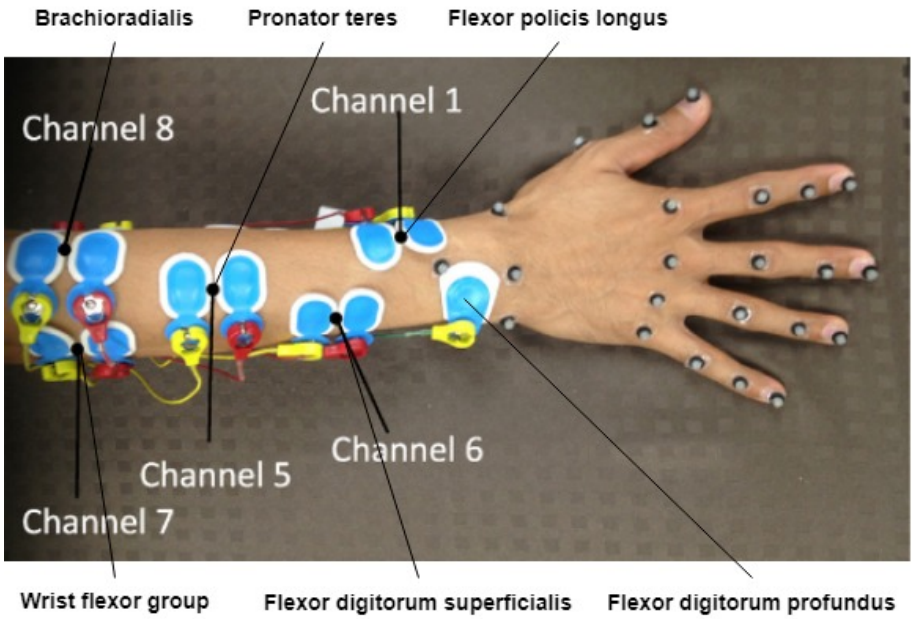

Figure 2.4: Example of multiple sEMG sensors placement in the laboratory environment [63].

The control method mentioned above is cumbersome, and the learning cost for users is high [35]. In recent years, due to the development of machine learning, pattern

recognition has been applied to a myoelectric prosthetic hand. Since all the applications mentioned above use only two sensors, the signal richness is far from the level required for pattern recognition applications. Therefore, in recent studies, more sensors have been placed on users' forearms, as shown in Figure 2.4. In the subsequent development, the array-based sensor placement method was used instead of specific placement for each individual sensor [59, 36]. This sEMG sensor placement method makes the classification of signals more widely applicable, not only to experimental subjects but also to other amputees. At the same time, it also promotes the development of research on the universality of prosthetic grasp classification.

By obtaining the multi-dimensional signals on the forearm and using pattern recognition technology, users no longer need to manually or rely on particular methods to select grasping gestures. Therefore, the myoelectric prosthetic hand is expected to be more natural when performing grasping actions. However, so far, most of the pattern recognition applications are still in the research stage and have not fundamentally improved the user experience of amputees. As mentioned in the Section 1.1, the quality of pattern recognition application depends on how we record and decode the information in sEMG signals and how we address noises caused by other factors, such as body condition, muscle flexibility, muscle fatigue, and sweat [85, 14, 78].

2.2 Gesture Recognition by Computer Vision

There are some limitations in using only sEMG for gesture recognition. First, the sEMG signal fluctuates greatly in the early period of the grasping process, which means it is hard to decode in this period. Therefore, most researchers choose to implement classification during the later period when the subject firmly holds the object

because the signal is relatively steady in this period. Although using the later period significantly increases the classification accuracy, it causes the second limitation that the prosthetic grasping process has a long delay. The reason is that, in real-life applications, the prosthesis controller has to wait until reaching the later period to produce the correct grasp command. Moreover, the sEMG signal easily interferes with other electronic devices, which produce signals in a similar frequency range. Due to decoding difficulty, these problems are hard to be solved only by using sEMG signals because the sEMG classification accuracy might reach its ceiling. Instead of focusing solely on the sEMG, researchers have proposed several strategies that integrate sEMG signals with other information channels, such as visual information, which is not affected by amputation. Therefore, grasping gesture recognition by computer vision has become a new research direction.

Grasping gesture recognition based on computer vision is a technology that generates the corresponding grasping gesture according to the visual information of the target object. The technology can be applied to a myoelectric prosthetic hand to guide the execution of the grasp movements. Typical visual gesture recognition and grasping systems consist of the visual gesture recognition module and the manipulator control module. The former obtains the visual information of the captured object through the camera, analyzes and classifies it, determines the corresponding grasping action, and transmits the information to the manipulator control module. The latter monitors the sEMG signals in real time and triggers the grasping action by the signal changes.

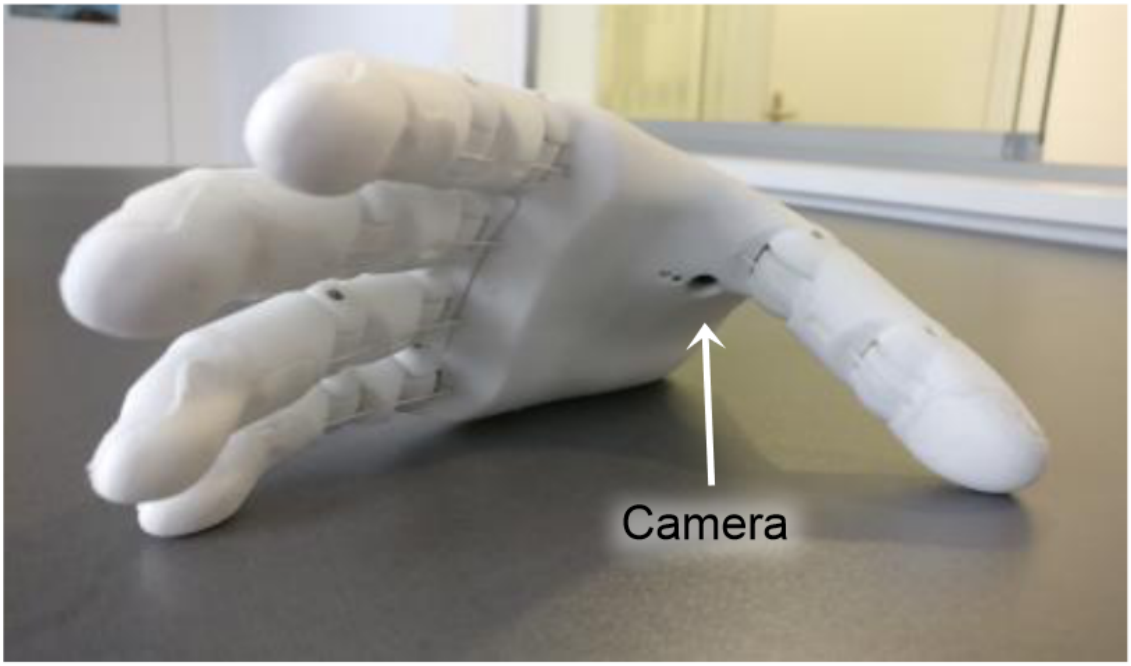

Figure 2.5: Camera placement. The camera was placed in the palm of the prosthetic hand [41].

Compared with the previous manual selection of grasping gestures, Hundhausen et al. [41] tried to reduce user input instructions as much as possible by using computer vision technology for gesture recognition. The study simplified the operation process and switched the pure manual gesture selection into the semi-autonomous grasping process, which can effectively accelerate the grasping process. In their experiment, a camera was placed in the palm, as shown in Figure 2.5. When the hand approached the object, the visual information obtained by the camera was processed, and the corresponding grasping gesture was determined for final prosthesis execution. Although this method significantly reduced human manual intervention, it cannot entirely rely on computer vision to determine the final grasping gesture.

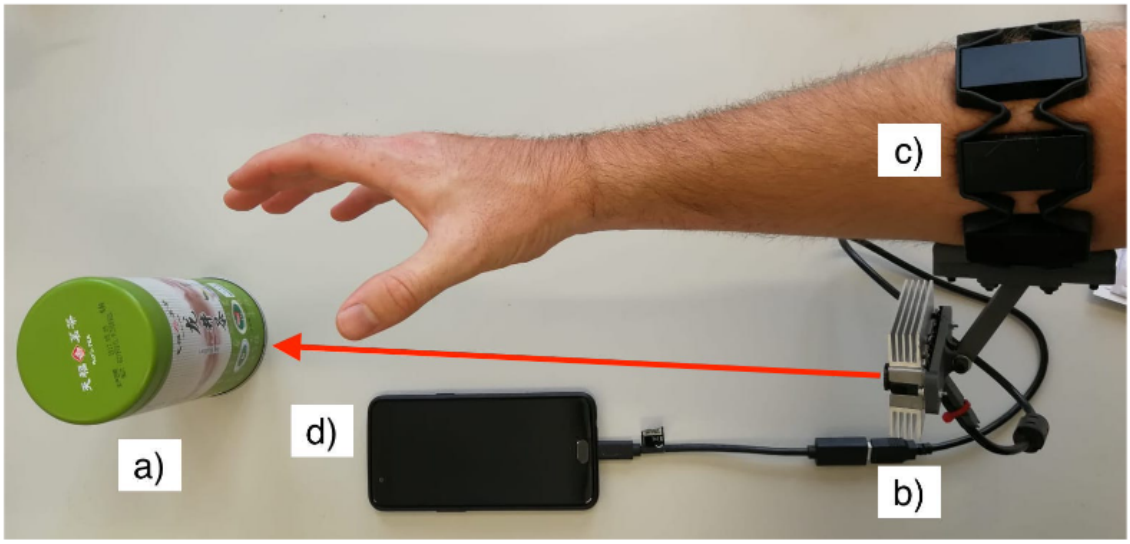

Figure 2.6: Camera placement. The camera was placed on the forearm at the same location of sEMG sensors. This system work process is a) the video camera obtained the object photo, b) the connector transferred the data into a smartphone phone application, c) the sEMG sensor armband read sEMG signals, d) the system generated the corresponding grasp gesture [80].

In contrast to the above study, Taverne et al. [80] developed a gesture recognition system purely based on computer vision, which did not require any manual intervention by users, significantly reduced the delay and improved the robustness. In their experiment, a camera was placed on the forearm in the same position as the sEMG sensors, as shown in Figure 2.6. The system can generate predictive results for each video frame and automatically select the corresponding grasping gesture when the hand touches the object. The gesture recognition rate on objects that appeared in the dataset reached 95.90%. In addition to recognizing objects in the data set, the

system also included a gesture recognition function for entirely new objects, and the accuracy reached 88.65%.

From the above two studies, we can find that the visual pattern recognition accuracy achieves a better performance than sEMG with stable object photos as the input. Nevertheless, in a natural grasp process, the input of object photos is not as stable as in the laboratory environment. However, the sEMG signal is a stable source in a laboratory and real-life environment. Therefore, a new research direction arises, which is using sEMG as a basis and integrating vision as a support to improve the overall prosthetic control performance.

2.3 Machine Learning

As a branch of artificial intelligence, machine learning provides a variety of methods to deal with complex problems, such as classification and regression. Machine learning studies computational algorithms that can make predictions or decisions using training data to find the solutions to problems [60]. The essence of machine learning algorithms is mathematical modelling [60, 73, 61]. Based on the existing form of labels in the data set, machine learning can be divided into three categories: supervised learning, unsupervised learning and reinforcement learning. In supervised learning, the machine is constantly training itself using data and corresponding labels to get a mapping from input to output. In unsupervised learning, there are no labels in the data set, and the model only relies on the data for pattern recognition. The main idea in reinforcement learning is to interact with the environment and determine the decision to maximize the benefit in the current environment. The diversity of machine learning makes it a powerful tool for solving complex problems.

Machine learning is applied to prosthesis control to predict the grasping gestures of users. Myoelectric prosthesis controlled by various machine learning methods has more intuitive control ability. Each machine learning method is a model that can predict the grasping action from the captured sEMG signal. Therefore, these models can be regarded as a mapping function that receives input signals from the sEMG sensor and outputs the motion instructions to the manipulator.

The classification model occupies the most significant proportion of machine learning models in the myoelectric prosthesis. This model can gradually form the prediction ability of unlabelled data by learning existing data, so this model is also called a data-driven model. The model's input is a part of the existing data, also known as training data, including observations and corresponding classes, also known as labels. For the sEMG signal data, the observation is the signal recorded by sEMG sensors, and the label is the actual grasp gesture corresponding to the signal. A classifier can convert an input signal into an output instruction. However, before that, the user must implement data collection to form training data and then train the model. The prediction can only be made after the model has been trained. Through continuous learning of the training data, the parameters in the model are constantly improved until labels can be generated for the unknown sample data. For example, we need to collect the corresponding ten muscle signal patterns for a myoelectric prosthetic hand that can perform ten different grasping gestures. Each signal pattern corresponds to each grasping gesture. Another portion of the existing data is used to measure the model's performance, known as testing data. The model's performance can be measured by feeding observations from testing data into the classifier and comparing the output labels with actual labels. Once the classifier has completed training and

testing, it can be used in the controller of the myoelectric prosthetic hand.

Machine learning approaches have shown an excellent potential for myoelectric prosthesis study. However, original data cannot be applied directly to machine learning models because of the highly fluctuated signals and noise problem in the raw data. It is a challenge for machine learning algorithms to classify noisy biological signals. Researchers have put plenty of effort into sEMG signal processing and proposed feasible methods to increase the signal to noise ratio of the original data. Signal denoising is one of the data processing methods, including modular maximal reconstruction filtering, spatial correlation filtering and threshold filtering. Using these technologies, the wavelet components generated by noise at each scale can be removed, and an accurate estimate of the original signal can be obtained [84]. Feature selection is another efficient machine learning data processing method to extract discriminatory features, which can significantly increase the classification accuracy and reduce the training time and over-fitting [42].

Chapter 3

Methods

This chapter discusses the terminology and methodology used in the thesis studies. We will start by introducing the challenges we met, followed by grasp phases that an entire grasp process includes. Then we will introduce the sEMG signal processing methods and the sEMG classification model. At last, we will discuss some methods related to computer vision, such as object detection and visual classification models.

3.1 Research Challenges

The first challenge is to figure out the suitable overlapped window size and step. As the increase of window size, the classification accuracy increased, but the delay also increased. This is because when we increase the window size, more data was used to derive better features, and betters features led to higher classification accuracy. It is challenging to find the balance that can obtain an acceptable classification accuracy and keep a low latency time at the same time. To solve this problem, we implement an experiment to find the correlation between windows length, step and classification

accuracy. A balance with 200 ms window length and 50 ms step was chosen in our research. The second challenge is building the object detection dataset from the recorded video. Considering the high volume of videos, it is impossible to manually extract each object to build the dataset. In this condition, we trained a RetinaNet object detection model and build a pipeline to detect and extract each object frame by frame from the video. The details of how we solve these challenges will be discussed in Chapter 3.

3.2 Grasp Phases

Normally, a typical reaching and grasping process can be divided into three phases [57, 79]:

1. The Reaching Phase: starts from the hand lifting off and ends by touching the object. During this phase, the hand is accelerated to a peak velocity and then is decelerated and brought to touch the target object. The hand is usually configured to the target grasp gesture (pre-shape) [6].

2. The Early Grasping Phase begins when the hand initially contacts the object and gradually closes the fingers until the hand starts to grasp the object firmly.

3. The Firm Grasping Phase: the target object is firmly grasped, and the hand shape is maintained relatively steady.

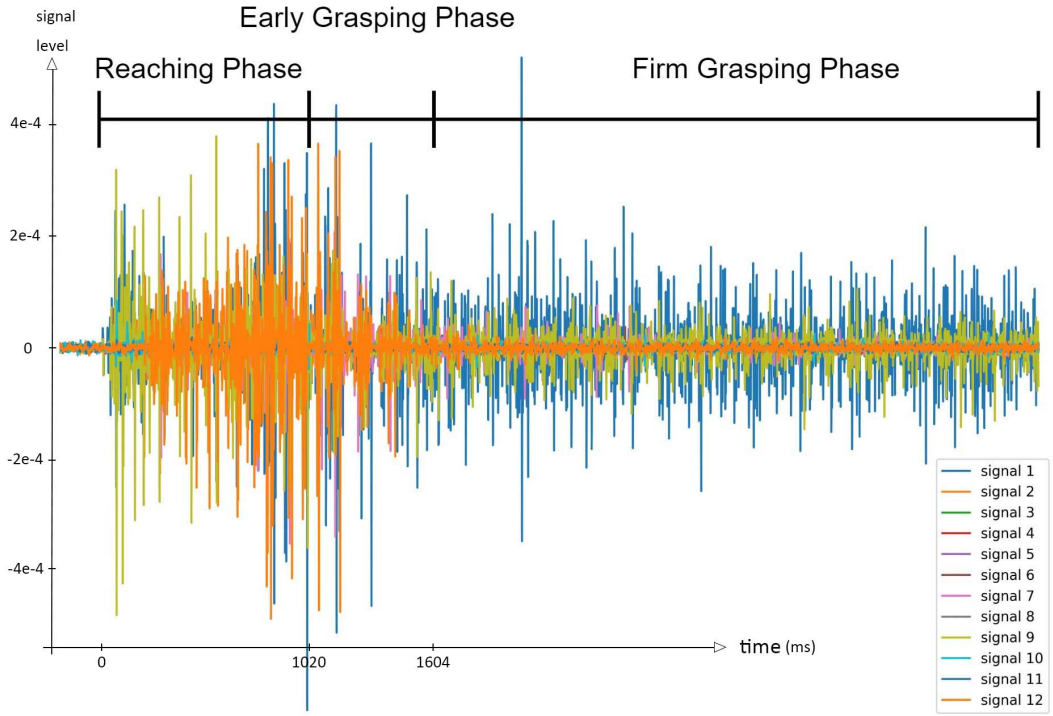

Figure 3.1: An example of grasp phases overlaid with sEMG signals during a full grasp trial. The start and end positions of these three phases were determined by observing corresponding videos frame by frame.

We segmented the Reaching, Early grasping, and Firm Grasping phases of each grasp gesture from each subject by observing the corresponding video frame by frame and calculated the average duration of each phase from all the observations. The judgment criteria for entering an Early Grasping Phase was the moment that the hand started to touch the target object, and the judgement criteria for entering a Firm Grasping Phase was the moment that the target grasp gesture was completely formed and the hand started to keep relatively steady. According to the segmentation, Early Grasping Phase and Firm Grasping Phase started at 1020 ms and 1604 ms from the beginning of the Reaching phase, respectively. The start and end positions of these three phases

were determined by inter-rater reliability testing between two people. An example of grasp phases overlaid with sEMG signals during an entire grasp trail is shown in Figure 3.1.

3.3 sEMG Signal Processing and Classification Model

3.3.1 sEMG Filters

In this section, we listed the details of the eleven different sEMG filters from literature [66, 47] because these filters are proved to have potential to be suitable for sEMG signals.

- **Standard Deviation (STD)** calculates the amount of variation or dispersion of sEMG signals. The mathematical equation can be defined as

$$STD = \sqrt{\sum_{i=1}^{N}(x_i - \bar{x})^2} \tag{3.1}$$

where N represents the length of the sEMG signal, x_i is the current sEMG signal reading in a segment i, $\bar{x}$ is the signals' mean value.

- **Root Mean Square (RMS)** is similar to standard deviation and widely used for sEMG signal analysis [8, 49]. It calculates the summation of the square of the signals inside the sliding window. The mathematical equation can be defined as

$$RMS = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2} \tag{3.2}$$

where N represents the length of the sEMG signal, x_i is the current sEMG signal reading in a segment i.

- **Integrated EMG (IEMG)** is usually used to detect early symptoms of diseases in clinical environment [38]. It calculates summation of the absolute values in the window. The mathematical equation can be defined as

$$IEMG = \sum_{i=1}^{N} |x_i| \tag{3.3}$$

where N represents the length of the sEMG signal, x_i is the current sEMG signal reading in a segment i.

- **Mean Absolute Value (MAV)** is another popular method used for the sEMG signal analysis. [40, 86]. Besides IEMG, it is also used for detecting early symptoms of diseases in clinical environment. It computes the mean of the absolute values inside the sliding window. The mathematical equation can be defined as

$$MAV = \frac{1}{N} \sum_{i=1}^{N} |x_i| \tag{3.4}$$

where N represents the length of the sEMG signal, x_i is the current sEMG signal reading in a segment i.

- **Waveform Length (WL)** is normally used for measuring the complexity of the sEMG signal [40]. It calculates the cumulative length of the signals in the

window as its representation. The mathematical equation can be defined as

$$WL = \sum_{i=1}^{N-1} |x_{i+1} - x_i| \tag{3.5}$$

where N represents the length of the sEMG signal, x_i is the current sEMG signal reading in a segment i and x_{i+1} is the next sEMG signal reading in a segment *i+1*.

- **Log Detector (LOG)** is a non-linear method for providing the measurement of the muscle contraction force. The mathematical equation can be defined as

$$LOG = \exp\left(\frac{1}{N}\sum_{i=1}^{N} \log(|x_i|)\right) \tag{3.6}$$

where N represents the length of the sEMG signal, x_i is the current sEMG signal reading in a segment i.

- **Simple Square Integral (SSI)** is a measurement of the energy of the sEMG signals. It calculates the summation of the square values of the sEMG sginals. The mathematical equation can be defined as

$$SSI = \sum_{i=1}^{N} x_i^2 \tag{3.7}$$

where N represents the length of the sEMG signal, x_i is the current sEMG signal reading in a segment i.

- **Skewness (SKW)** is one of the High Order Statistics (HOS) parameters and

has been widely used for detecting the shape features in sEMG analysis [9]. It computes the lack of symmetry. The mathematical equation can be defined as

$$SKW = \frac{\sum_{i=1}^{N}(x_i - \bar{x})^3/N}{\delta^3} \tag{3.8}$$

where N represents the length of the sEMG signal, x_i is the current sEMG signal reading in a segment i, $\bar{x}$ and and δ are the signals' mean value and standard deviation, respectively.

- **Kurtosis (KURT)** is another HOS parameters for detecting the shape features in sEMG analysis [88]. It calculates the degree to which the data is heavy-tailed or light-tailed relative to a Gaussian. The mathematical equation can be defined as

$$KURT = \frac{\sum_{i=1}^{N}(x_i - \bar{x})^4/N}{\delta^4} \tag{3.9}$$

where N represents the length of the sEMG signal, x_i is the current sEMG signal reading in a segment i, $\bar{x}$ and and δ are the signals' mean value and standard deviation, respectively.

- **Average Amplitude Change (AAC)** is a similar measurement method to Waveform Length, the only difference is that AAC is averaged. [27]. The mathematical equation can be defined as

$$AAC = \frac{1}{N}\sum_{i=1}^{N-1}|x_{i+1} - x_i| \tag{3.10}$$

where N represents the length of the sEMG signal, x_i is the current sEMG signal reading in a segment i and x_{i+1} is the next sEMG signal reading in a segment

i+1.

- **Difference Absolute Standard Deviation Value (DASDV)** is similar to Root Mean Square, it is a measurement of the standard deviation value of the wavelength.

$$DASDV = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N-1}(x_{i+1} - x_i)^2} \tag{3.11}$$

where N represents the length of the sEMG signal, x_i is the current sEMG signal reading in a segment i.

3.3.2 Electromyography Feature Extraction and Selection

In the feature extraction process, we first determined the suitable window size for deriving features [23]. As shown in Table 3.1, several sizes of the overlapped window were tested, which are 50 ms, 100 ms, 200 ms, 500 ms, and 1000 ms. As the increase of the window size, the accuracy keeps increasing, which means that the more data we used to derive features, the better performance we could get. However, considering the capability of Myoelectric prosthesis in real-life conditions, a large window would delay the grasp action from the prosthetic hand. On the other hand, it can be seen that, when increasing the window size over 200 ms, the accuracy increase is less than 1%, which is a tiny increase. Therefore, to keep the balance between accuracy and implementation speed, we chose 200 ms as the window length with the step of 50 ms, which is a 75% overlap between successive windows.

Table 3.1: Window Length Analysis. Both training and test data used the whole grasp period. The classifier used was lightGBM. The features used were STD, RMS, IEMG, MAV, WL, SSI, AAC, and DASDV mentioned in Figure 3.2. The cross-validation method used was leave-one-repetition-out cross-validation which used one repetition data for testing and the rest three repetitions for training the model, and repeated this process four times to cover all repetitions for testing.

Window Length	**Accuracy**
50 ms	77.02%
100 ms	78.79%
200 ms	79.98%
500 ms	80.04%
1000 ms	80.33%

To assure the recognition accuracy by using proper features, we tested eleven commonly used features, which are Standard Deviation (STD), Root Mean Square (RMS), Integrated EMG (IEMG), Mean Absolute Value (MAV), Waveform Length (WL), Log Detector (LOG), Simple Square Integral (SSI), Skewness (SKW), Kurtosis (KURT), Average Amplitude Change (AAC) and Difference Absolute Standard Deviation Value (DASDV) [67]. We dropped the three lowest performance features, which are LOG, SKW and KURT and chose the rest eight with the highest accuracy as the final features for the following research. The performance of these features is shown in Figure 3.2. After applying the eight features to the sEMG signals, the data set was converted

from 12 to 96 columns. Due to the sensor hardware issue mentioned in the first subsection, the sEMG data of subject S024 was changed from 11 to 88 columns.

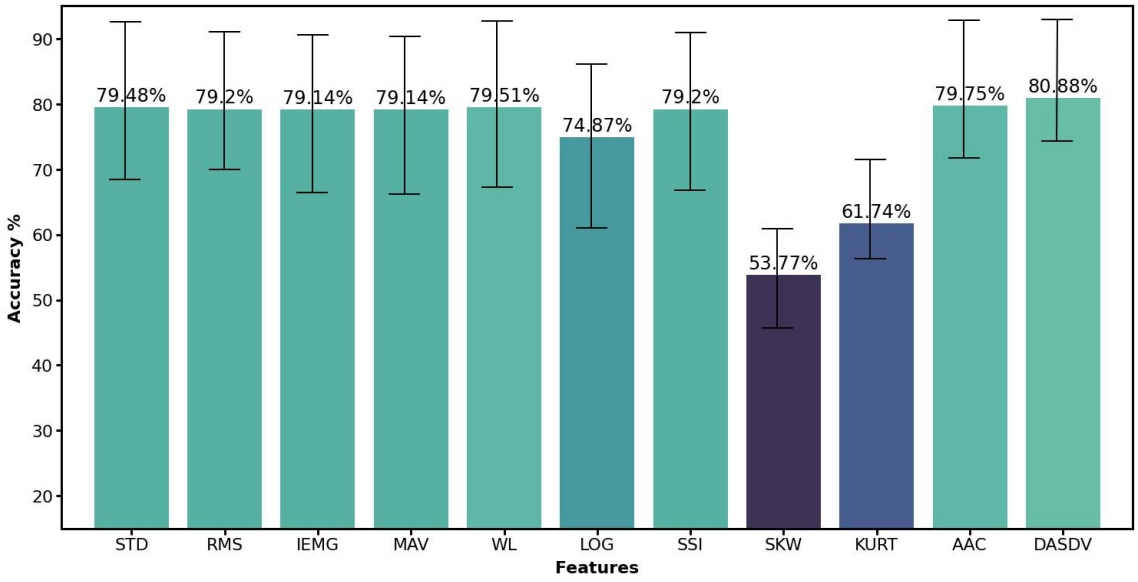

Figure 3.2: Single feature performance with window size 200 ms. The eleven features are Standard Deviation (STD), Root Mean Square (RMS), Integrated EMG (IEMG), Mean Absolute Value (MAV), Waveform Length (WL), Log Detector (LOG), Simple Square Integral (SSI), Skewness (SKW), Kurtosis (KURT), Average Amplitude Change (AAC) and Difference Absolute Standard Deviation Value (DASDV). The classifier used was lightGBM. The cross-validation method used was leave-one-repetition-out cross-validation which used one repetition data for testing and the rest three repetitions for training the model, and repeated this process four times to cover all repetitions for testing.

3.3.3 sEMG Classification Model

Gradient boosting decision tree, such as XGBoost [28] and Light Gradient Boosting Machine (LightGBM) [48], is a popular machine learning algorithm used by a large amount of data scientists recently, which can achieve high performance by using decision trees as weak learners and assembling them to come up with one strong learner. Considering the high feature dimensions and large data size, we chose LightGBM as the classifier which runs faster while maintaining a high level of accuracy by utilizing two novel techniques called Gradient-Based One-Side Sampling (GOSS) and Exclusive Feature Bunding (EFB) [48]. In the experiment of Ke et al. (2017), LightGBM can accelerate the training process up to twenty times more than XGBoost. The tool we used to implement the LightGBM classifier is an open-source python package LightGBM developed by Microsoft Corporation.

We tuned the hyperparameters by using the training set of all the subjects and obtained the best results as follows: the learning rate is 0.1; no limit was set for the maximum depth; the number of estimators is 100; the number of leaves is 31; the remaining parameters were set to the default values. In the parameter turning process, the training set was split into sub-training and validation set with the default ratio of 80% for training and 20% for validation.

3.4 Object Detection and Classification Model

3.4.1 RetinaNet Detection Model

We used RetinaNet to detect objects from the frames of the video. RetinaNet is a one-stage convolutional neural network model widely used for object detection, which utilizes a focal loss function to address class imbalance during training [51]. Considering the high volume of the dataset and the heavy labelling work, in this study, we chose the RetinaNet model pre-trained by the MS COCO (Microsoft Common Objects in Context) dataset to reduce the required size of the training dataset. COCO dataset included photos of 91 object types that would be easily recognizable by a four-year-old, and it contained a total of 2.5 million labelled instances in 328k images [52].

Since the COCO dataset did not fully cover the object types in the dataset we used, fine-tuning is required to make it fit our object types. We build a dataset by collecting and labelling 1186 photos from the videos of the 30 subjects. There are 3-6 objects in each photo, and each object showed approximately 200 times in these 1186 photos. 80% of this dataset was used for training, and 20% was used for validation. Then we created a new output layer to replace the previous output layer in the pre-trained model and trained it using the training and testing data mentioned above (this training and testing data has no overlap with the data for classification training and testing in the following work). Therefore, the final model we obtained can be used to detect the object in each frame.

There could be multiple objects that appeared in a frame of the first-person video.

However, we only needed to detect the target object that the participant was trying to grasp using the fine-tuned RetinaNet model (as described in the Section 2.1), where the target object was cropped using a bounding box. If the target object was shown in the frame, we regarded this frame as a valid one. We regarded it invalid if the target object was not shown in the frame or entirely blocked by the hands. Setting up the valid frame where the target object is shown within the frame was necessary for further analysis. The object and gesture recognition model we applied only worked under the condition where one object is presented in the frame. Objects could be detected by the RetinaNet model from approximately 90% frames. We reviewed and manually detected the object on the rest of the frames.

3.4.2 Dual-Channel CNN Classification Model

Zhang et al. proposed a dual-channel convolutional neural network (DcnnGrasp), in which object category information was adopted to improve the accuracy of the grasp pattern recognition [89]. To maximize the collaborative learning of object category classification and grasp pattern recognition, they proposed a loss function called Joint Cross-Entropy with Adaptive Regularizer (JCEAR) derived from maximizing a posterior. A developed training strategy updated the regularization coefficient and trainable parameters in the loss function JCEAR and DcnnGrasp. From the experiments given in their paper, it can be seen that, compared with SOTA methods, DcnnGrasp achieved the best accuracy in most cases [89].

In this study, we trained the DcnnGrasp model using the object photos from the first three repetitions and tested using the left one repetition. When applying this model to testing data, the input photos in the same grasp trial were fed to the model

chronologically so that we could obtain the outcome in a time sequence.

Chapter 4

Experiments and Results

4.1 Data Collection and Pre-processing

4.1.1 Data Collection

The data used in this study were from an open-source dataset collected by Cognolato et al. [19], where the sEMG and vision data were recorded from 30 healthy subjects (27 male and 3 female), with an average age of 46.63 $\pm$ 15.11 years.

Twelve sEMG sensors were placed on the forearm of each subject, producing twelve columns of sEMG data, respectively. The head-mounted camera recorded the first-person videos simultaneously with the sEMG data. Due to a hardware problem, no myoelectric data were received from electrode number eight during the acquisition of subject S024. Therefore, the sEMG data for this subject was recorded from eleven electrodes instead of twelve [19].

Ten grasp gestures were performed in this data collection which was selected based on the hand taxonomies [22, 72, 20, 26] and grasp frequency in Activities of Daily Living [11]. As shown in Figure 4.1, the participant performed each gesture for four repetitions, and in each repetition, the same gesture was performed three times using three different objects, respectively. A designated experimenter vocally guided the participant to perform which gestures and grasp which objects. The data were labelled according to the vocal instruction. Table 4.1 shows the list of gestures and objects.

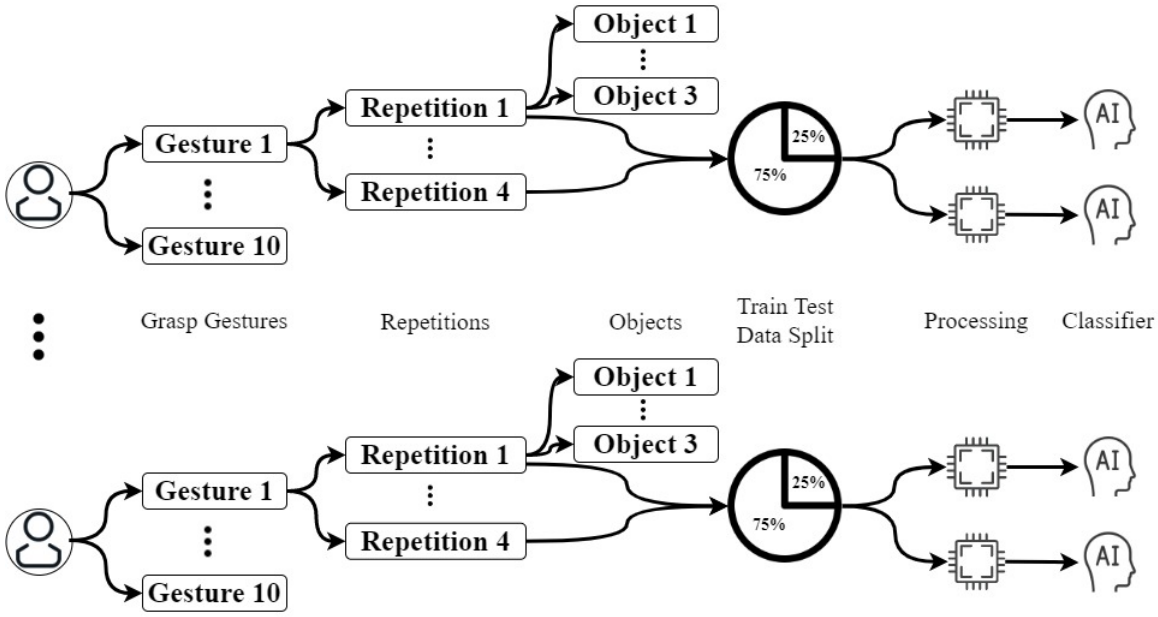

Figure 4.1: The data structure and processing steps.

Table 4.1: The columns indicate the ID and name of the grasp gestures, the name of the object, and the name of the part of the object involved in the grasping [19].

ID	Grasp Gesture	Object	Grasp Location
1	medium wrap	bottle can door handle	bottle body can body door handle stick
2	lateral	mug key pencil case	mug handle key body case zip
3	parallel extension	plate book drawer	plate edge book body drawer edge
4	tripod grasp	bottle mug drawer	bottle cap mug body drawer knob
5	power sphere	ball bulb key	ball body bulb body key chain
6	precision disk	jar bulb ball	jar lid bulb body ball body
7	prismatic pinch	clothespin key can	clothespin body key ring can pull tab
8	index finger extension	remote knife fork	remote button knife body fork body
9	adducted thumb	screwdriver remote wrench	screwdriver body remote body wrench body
10	prismatic four finger	knife fork wrench	knife handle fork handle wrench handle

4.1.2 Data Pre-processing and Splitting

As part of data pre-processing, the abnormal samples were replaced with the precedent valid samples when filtering outliers [19]. As there might be a delay between the participant's response to the vocal instructions [19], the sEMG activation time might not be matched perfectly with the stimulus time. Therefore, relabeling was performed to calibrate this difference using the method described by Kuzborskij et al. [50].

In this study, each participant performed one grasp gesture four times (repetitions) which allowed us to split the sEMG data by repetitions to validate testing results. For all the cases in this study, we used three repetitions (75%) for training and one repetition (25%) for testing with leave-one-repetition-out cross-validation, which used one repetition data for testing and the rest three repetitions for training the model. We repeated this process four times to cover all repetitions for testing. The data organization and processing steps can be easily understood from Figure 4.1.

4.2 Grasp Phase Analysis for sEMG and Experiments

We conducted two experiments. The first aimed to analyze the grasp classification accuracy during the three grasping phases and find the best position and length of the sweet period; another was to find the best training strategy.

4.2.1 Data Processing

The grasp trials performed by the participants lasted approximately 4.5–5 s [19]. We removed the data after 4.5 s to align all the trials to the same length. Because the overlapped window step is 50 ms and the grasp period length is 4.5 s, 90 samples of data reminded for each trial.

As mentioned in the Section 4.1.1, three objects were grasped in each repetition with the same gesture to increase the reliability of the sEMG data set. In other words, there were 324,000 data samples (90 samples/grasp $\times$ 10 grasp gestures $\times$ 4 repetitions $\times$ 3 objects $\times$ 30 subjects) in the data set.

4.2.2 Phases and Sweet Period Analysis

Figure 4.2 shows the mean changes in testing accuracy of grasp classification during all three phases. Each data point is averaged across all 900 trials from 30 participants.

From Figure 4.2 we can see that the accuracy increases from 42% to 84% during the Reaching phase and then becomes stable at the start of the Early Grasping phase at around the time of 1000 ms, fluctuating between 84% and 87% during the rest of the grasp period. The mean accuracy further increases to relatively stable at around 1250 ms, where we then define the location of the sweet period.

To find the optimal length of the sweet period, we designed different sliding windows with sizes of 300 ms, 400 ms, 500 ms, 600 ms, 700 ms, 800 ms, 900 ms and 1000 ms. The sliding window moved along with the time with step 50 ms; in each move, it calculated and recorded the mean accuracy. We analyzed the records from the sliding

window, and the results are given in Figure 4.3.

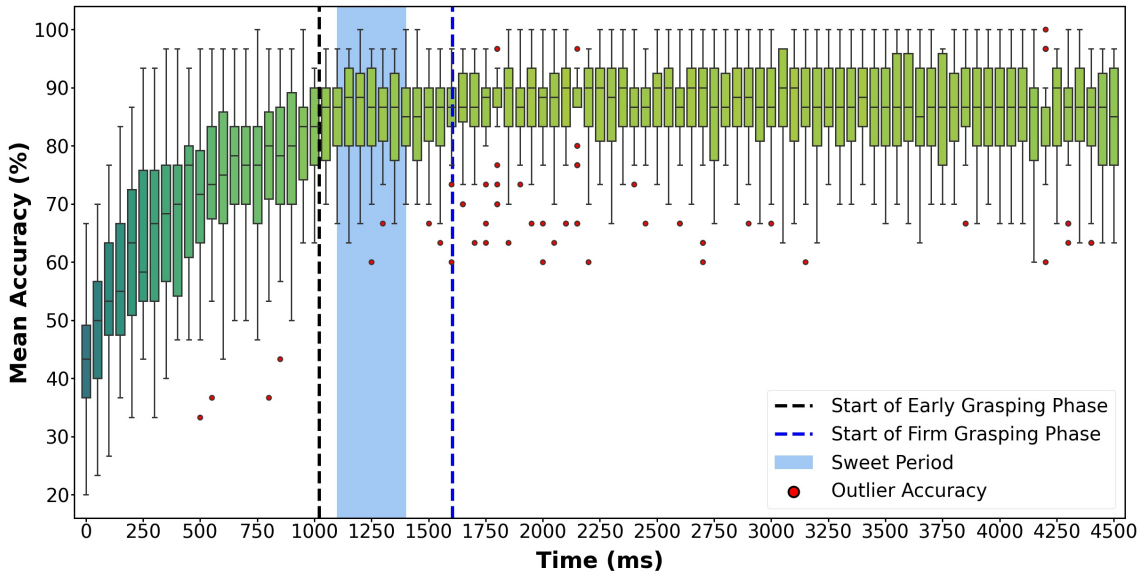

Figure 4.2: Mean accuracy at each time point during the entire grasp period. This result is from the model which was trained using all three phases data using leave-one-repetition-out cross-validation, and the mean accuracy represents the average accuracy of 30 subjects. The blue region, starts from 1100 ms and ends from 1400 ms, is the sweet period which was confirmed from the first experiment. The vertical dashed lines are averaged starting times of Early Grasping and Firm Grasping phases, which locates at 1020 ms and 1604 ms, respectively. The red dots are outliers.

From Figure 4.3, we can see that the mean accuracy increases with the increase of window length significantly during the Reaching phase and beginning of the Early Grasping phase (at about 1100 ms) but not significantly afterward. For instance, although the window length of 1000 ms can reach the highest accuracy of 86.3%, it

takes much longer than the length of 300 ms with an accuracy of 85.5%. Therefore, the sweet period length is set to 300 ms, and the position is set from 1100 ms to 1400 ms, which makes it entirely located in the Early Grasping phase as the blue region shown in Figure 4.2.

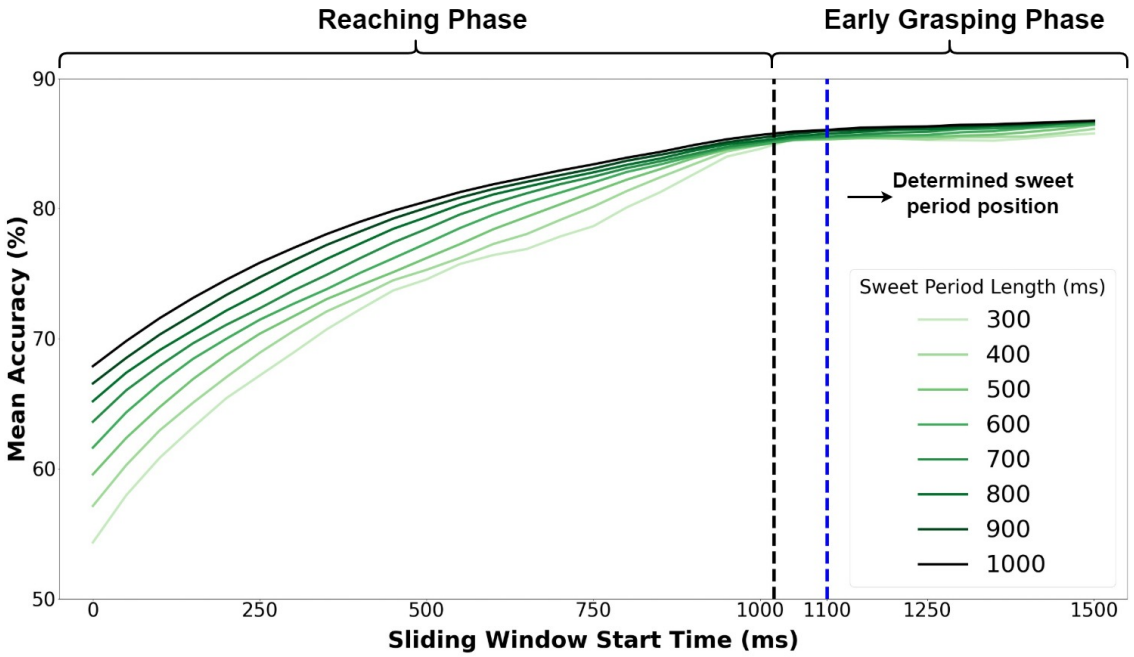

Figure 4.3: Mean accuracy with different sweet period lengths at different start time.

4.2.3 Comparison Experiment

In the comparison experiment, we tested six strategies using different training and testing data, as shown in Table 4.2.

Table 4.2: Analysis Results for Six Cases. All Three Phases include signal from the time of 0 ms to 4500 ms, Firm Grasping Phase is from the time of 2000 ms to 4500 ms, sweet period is from the time of 1100 ms to 1400 ms. Leave-one-repetition-out cross-validation was employed for all cases, such that all testing data was excluded from training the model.

Case Number	**Training Data**	**Testing Data**	**Accuracy**
1	All Three Phases	All Three Phases	79.98%
2	All Three Phases	Firm Grasping Phase	81.68%
3	All Three Phases	Sweet Period	85.50%
4	Firm Grasping Phase	Firm Grasping Phase	80.39%
5	Firm Grasping Phase	Sweet Period	60.80%
6	Reaching Phase and Early Grasping Phase	Sweet Period	81.01%
7	Early and Firm Grasping Phase	Sweet Period	82.51%
8	Sweet Period	Sweet Period	74.99%

In cases 1–3, we used all three grasp phases as training data and reduced the testing data size from all three phases to only the firm grasping phase, then to the sweet period. The purpose of performing these three comparisons was to study which phase/period was better for testing data when using all grasp phases as training data. Besides, we studied another five cases to figure out which phase played a better role in model training. For cases 4–5, we used Firm Grasping Phase for training and reduced

the testing data size. In cases 6–7, we used combined phases for training and the sweet period for testing. In case 8, we used the data in the sweet period both for training and testing. It is worth mentioning again that the cross-validation method used for all the cases was leave-one-repetition-out cross-validation which used one repetition of data for testing and the rest three repetitions for training the model, and repeated this process four times to cover all repetitions for testing, such that all testing data was excluded from training the model. For example, in one testing repetition of case 8, the data from the sweet period of three repetitions were used for training the model and the rest for testing. The results are presented in Table 4.2.

As shown in Table 4.2, we get the highest accuracy of 85.50% when we train with all grasp phases and test with the only sweet period. Besides, from cases 1 to 3, we find that if we keep the training data unchanged, the accuracy increases as the decrease of testing data size.

4.3 Grasp Phase Analysis for Visual Information and Integration with sEMG

The main workflow of this part is shown in Figure 4.4. Besides the RetinaNet model, DcnnGrasp model and frame extraction discussed in the last section, another four steps will be introduced.

4.3.1 Visual Dataset Building

Each subject performed 10 grasp gestures in the original video, and each grasp was acting on 3 different objects. On each object, the same grasp gesture was required

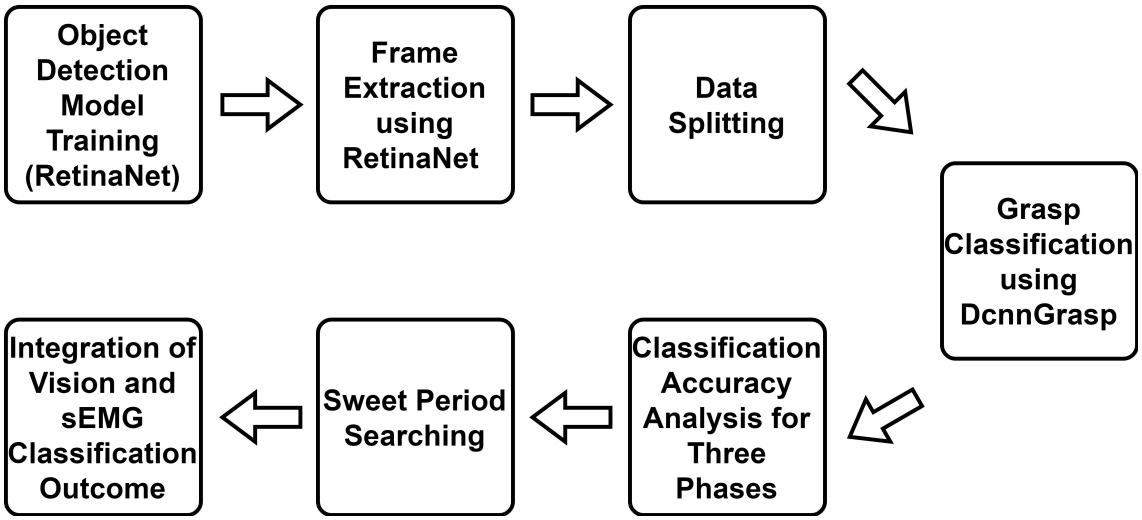

Figure 4.4: Work flow of analysis for visual information. Object Detection Model Training, Frame Extraction, DcnnGrasp Model are discussed in Methods section. The rest four parts are discussed in Experiments and Results section.

to perform 4 times. There were 30 subjects in the database. Therefore, there were 3600 grasp trials in total (120 trials x 30 subjects) in this dataset. Therefore, there were 120 ($10 \times 4 \times 3$) grasp trials in total for each subject. Each grasp trial lasted approximately 4.5–5 s [19]. To keep all the trials to the same length, we removed the frames after 4.48 s. Please be noted that the video was recorded with a frame rate of 25 Hz (one frame per 0.04 s), and there were 112 frames in each trial of video.

4.3.2 Sweet Period Analysis

After extracting the frames from the first-person videos, we calculated the proportion of valid frames and drew Figure 4.5 to illustrate the changes during the entire grasp process. There are 900 trials in testing data (30 subjects x 10 grasp gestures x 1 test repetition x 3 objects), which means that there are 900 frames at each time point. The percentage in Figure 4.5 represents the proportion of valid frames among these 900 frames at each time point. The result is calculated with leave-one-repetition-out

cross-validation.

We can see that the valid frame proportion increases at the beginning of the Reaching phase (Figure 4.5), reaching the peak at 160 ms. The valid frame starts to decrease until the late Early Grasping phase, keeping a stable low level during the entire Firm Grasping phase. The high percentage of the valid frame at the early Reaching phase allowed us to define the location of the sweet period.

In searching for the sweet period, we defined several windows with different lengths and calculated the average percentage of valid frames in these windows. Since we wanted the sweet period to locate at as early period as possible and the percentage is high enough at the start of the Reaching phase (0 ms in Figure 4.5), we made all the windows start from 0 ms and end at different times. After calculation, the window with the second highest average accuracy was chosen as the sweet period shown in the blue zone, which was from 0 ms to 320 ms (Figure 4.5). The window with the highest average accuracy (from 0ms to 160 ms) was dropped because it only contained four frames which were not enough to make it reliable.

In our previous research [83], we achieved the best grasp classification outcome using the sEMG sweet period between 1100 ms and 1400 ms in the Early Grasping phase (pink zone in Figure 3). The sEMG sweet period was 800 ms behind the vision sweet period. The time gap between these two sweet periods makes it possible for us to integrate the classification outcome using vision and sEMG in a real-life situation. Although the definition of the sweet period for sEMG and vision was the same, the methods to determine the sweet period were different. The sEMG sweet period was identified by analyzing classification accuracy, while the vision sweet period was found

by analyzing the percentage of valid frames.

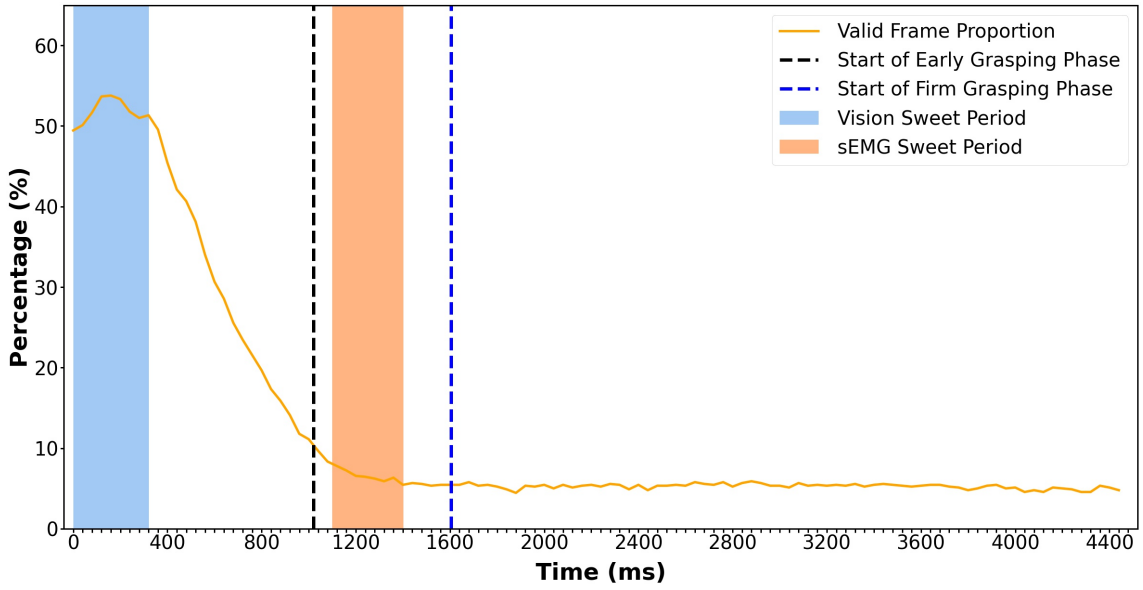

Figure 4.5: Valid frame proportion at each time point during the entire grasping process. The percentage represents the average proportion of valid frames in 900 trials from 30 subjects. The x-axis contains 112 points representing 112 frames in a grasp trial (40 ms for each frame, 4480 ms in total). The vision sweet period starts from 0 ms and ends at 320 ms. The sEMG sweet period starts at 1100 ms and ends at 1400 ms. The vertical dashed lines are averaged starting times of the Early Grasping and Firm Grasping phases, which locates at 1020 ms and 1604 ms, respectively. The result is calculated with leave-one-repetition-out cross-validation.

4.3.3 Comparison of Grasp Classification Performance

Figure 4.6 shows mean accuracy rates for grasp classification by three different methods (object recognition, gesture recognition, and sEMG) over the entire grasping phases. The value at each time point is averaged across all 900 trials from 30 participants. From Figure 4.6 we can see that object and gesture recognition accuracy

keep relatively stable, fluctuating slightly between 97% and 91% and are both higher than sEMG recognition accuracy in most circumstances. The object classification yields higher accuracy than the gesture classification during the entire grasp process, only reverses once at 3200 ms. The accuracy achieved by the gesture classification only goes below the sEMG classification on three occasions during the Firm Grasping phase. Overall, the gesture classification accuracy is much higher than sEMG recognition.

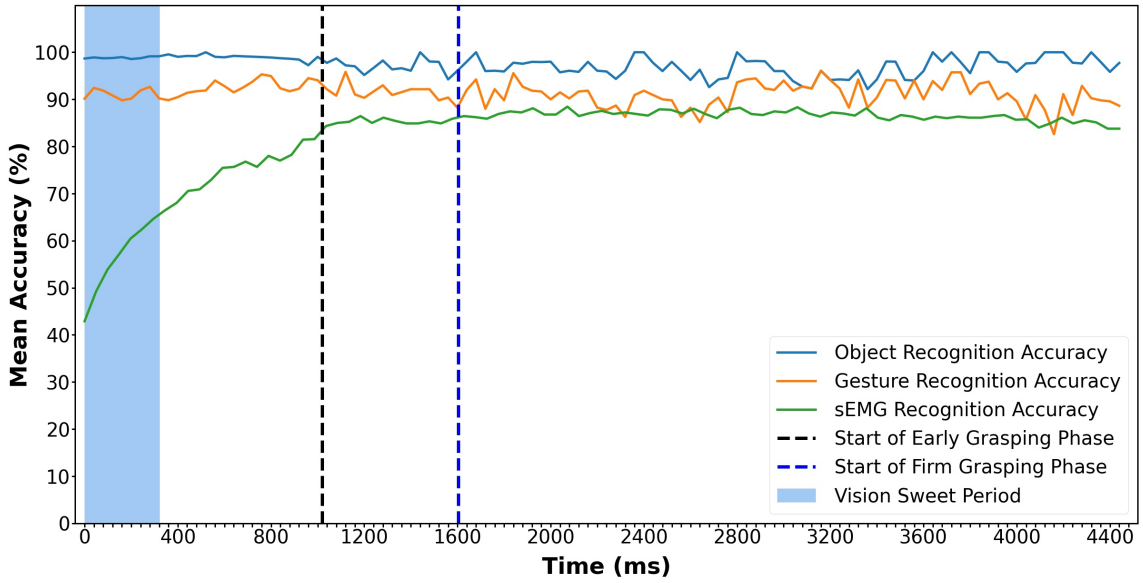

Figure 4.6: Mean accuracy for sEMG, object and gesture classification at each time point during the entire grasping process. The object and gesture recognition results are from the trained Dual-channel CNN model with leave-one-repetition-out cross-validation among valid frames. The mean accuracy represents the average accuracy of 900 trials from 30 subjects. The result is also calculated with leave-one-repetition-out cross-validation.

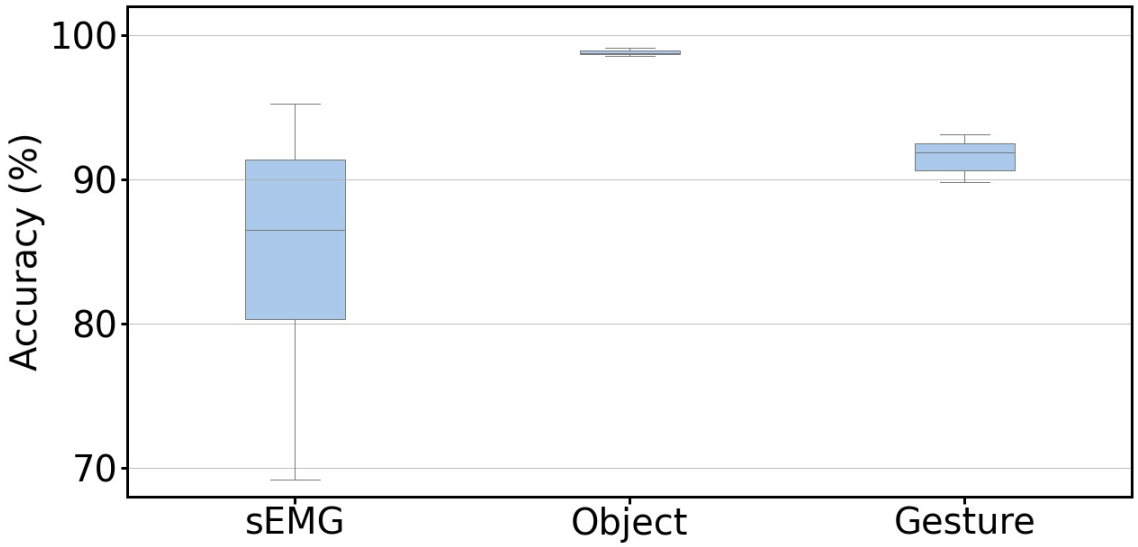

Figure 4.7: Mean classification accuracy of object and gesture during sweet period among 30 subjects. The sEMG result is from the best strategy we achieved in our last research which also utilized the sweet period for sEMG signals. The values for sEMG, object and gesture classification accuracy are 85.50%, 98.81%, 91.59%, respectively. The result is calculated with leave-one-repetition-out cross-validation.

Our last comparison of grasp classification over three methods focused on the sweet periods. Specifically, we calculated the mean accuracy of object and gesture recognition using data collected in their sweet periods. Compared to sEMG classification from the best strategy we have achieved in the previous report [83]. The results are shown in Figure 4.7, in which the object and gesture classification accuracy was calculated from valid frames in the sweet period. The mean accuracy reaches 98.81% and 91.59% for the object and gesture classification, respectively. Both are much higher than the sEMG classification accuracy of around 85.50%.

4.3.4 Integration of Vision and sEMG Classification Outcome

Satisfied grasp classification outcome from visual data encouraged us to integrate the gesture data from the vision sweet period with the sEMG data from the sEMG sweet period [83]. The simplest and most effective method to integrate these two outcomes was comparing the plurality vote probabilities during their respective sweet period. For the visual part, the probability was calculated from the plurality vote results of gesture recognition only for the valid frames during the sweet periods. For the sEMG part, the probability was also calculated from the plurality vote results during the sweet period, but each time point in the sweet period was valid. Both probabilities illustrated the confidence of the grasp recognition results classified from sweet periods. We calculated and compared the probabilities for each grasp trial in the testing dataset and chose the outcome with the higher probability as the final classification result. After calculating the mean accuracy of all the 900 grasp trials in the testing dataset with leave-one-repetition-out cross-validation, we obtained the results in Table 4.3. After the integration, the grasp classification accuracy was increased from 85.50% to 90.06% as shown in Table 4.3; however, the result was 1.5% lower than the visual gesture classification.

The vision sweet period lasts for 320 ms, which means there are 8 frames in this period. Due to the valid frame proportion being approximately 53%, around half of the frames are invalid. For the circumstance that the number of valid frames is less than 5 and at least one frame failed the recognition, the probability would be equal to or less than 75%, in which the sEMG would dominate this result because the probabilities of sEMG are stable and higher than 75% in most circumstances according to

our previous research [83]. If no frame failed the recognition, the vision with 91.59% probability would dominate the result.

Table 4.3: Gesture Classification Comparison. The result is calculated during sweet period among 900 grasp trials from 30 subjects with leave-one-repetition-out cross-validation.

Gesture Classification Basis	**Mean Accuracy**
sEMG	85.50%
Visual information	91.59%
Integration of sEMG and visual information	90.06%

Chapter 5

Discussion and Conclusion

Our hypothesis is supported by the results that there is an sEMG sweet period located in the Early Grasping Phase where sEMG signals can be used to achieve a similar or higher accuracy and lower delay of grasp classification than other windows, which would help to improve the performance of robotic hand manipulation in the real-life applications. This is important as the classifier can get the data much earlier instead of waiting for the muscle to get into the Firm Grasping Phase. We also identified a vision sweet period ahead of grasping the object where visual information can be used to classify gestures and integrated with sEMG. This is also important as the integration can break the accuracy limit by only using sEMG signals. Furthermore, this vision sweet period can be used to extract images automatically for grasp classification without a pause that is required when images are acquired manually.. This would make the prosthetic hand control more natural.

5.1 sEMG Phase-based Analysis for Delay Reducing

From the the Section 4.2.2, we obtained the result that the sEMG sweet period may be used to accelerate the speed of prosthetic hand control. In this part, we will discuss the details behind this result.

We found that the mean accuracy of sEMG during the Reaching phase is only about 63%. This is because, in this period, the subjects moved their hands to reach the object and started to perform the grasping gesture, keeping the hand moving. Therefore, the sEMG signal patterns in this period are not related to grasps, making it difficult to decode the sEMG signals, see Figure 3.1.

When getting into the Early Grasping phase, the accuracy of sEMG reaches approximately 85%, which is as high as that in the Firm Grasping phase. The possible reason is that the hand has already fully formed into the target gesture during the Early Grasping phase. Although this formed gesture may be slightly different from the final target gesture, it can provide sufficient information for the classification. After the subject firmly grasps the object (getting into the Firm Grasping phase), the accuracy keeps stable at around 85% because the sEMG signals started to be stable, which also makes the classification performance stable.

Notice that the sEMG signal is more active in the Reaching and Early Grasping phases with high amplitude of the sEMG waveform as shown in Figure 3.1. This is because the hand starts to perform the corresponding grasping gestures, such as hand

aperture, where the sEMG signals from the forearm are usually active with higher amplitude than other phases [6]. However, the hand has not grasped the object during the Reaching Phase. In contrast, starting from the mid-Early Grasping phase to the Firm Grasping Phase, the muscle status keeps relatively unchanged, which makes the amplitude sEMG signal slightly lower than that in the reaching and grasping phase; this is also why better grasp classification performance was achieved during the Early Grasping phase and the Firm Grasping Phase where the sEMG signal patterns are relatively similar. Therefore, using the sEMG sweet period can obtain similar or even higher accuracy than using the Firm Grasping phase. Since the sweet period is located much earlier than the Firm Grasping phase, using the sweet period will reduce the delay significantly. Overall, using a sweet period can simultaneously reduce the delay and maintain a high classification performance.

5.2 Vision Phase-based Analysis

We found that the valid frame proportion peaks at the start of the Reaching phase, as shown in Figure 4.5. This is because, before the grasp manipulation, the subject would look at the object before executing the grasp action, making it possible to identify hand movement using the visual signal during this period. After reaching the peak, the proportion quickly decreases to approximately 10% at the end of the Reaching phase. This is because, once the hand touches the object, some part of the object starts to be fully covered by the hand, blocking the object from showing in the subjects' vision. Therefore, the RetinaNet model cannot detect the objects, and the classifier cannot process the recognition. Since the hand starts to cover the object in the camera's view during the Reaching phase, the highest decline of valid frame proportion occurs in this period. During the Firm Grasping phase, hand occlusion

happened frequently; only a few objects with considerable volume can be recognized, thus making the valid frame proportion keep a low level of less than 10%. Although the proportion is the highest during the sweet period, it is only 53%. This is because the value for each time point is calculated across the 900 grasp trials (from 30 subjects), in which some objects are not shown in the subjects' vision at the current time point, or the objects are blocked by hand, makes this frame an invalid one at this time point. For this proportion level, we can find that it is impossible to implement recognition at a particular time point for different subjects and objects. However, it is feasible to expand the time point to a period to implement the recognition. In this research, we call this period the sweet period and find it between 0 ms and 320 ms. The probability of obtaining valid frames is higher in this period than in other periods. This period can provide a stable input to the classifier when performing a grasp action naturally.

From Figure 4.6 we can find that the visual recognition accuracy is much more stable than sEMG recognition accuracy in the Reaching phase. This is because sEMG signals change very much with muscle contraction, but visual information changes are rare, only some minor changes of visual angle. Therefore, as long as the visual information input is enough, the classification outcome would be stable.

5.3 Accuracy Improvement

Using all three grasp phases for training the model and only using the sEMG sweet period for controlling is found to achieve the highest classification accuracy, which implies that it can be the best strategy for myoelectric prosthetic hand application in real-life conditions. Not only because the sweet period during the Early Grasping

phase is suitable for prosthesis control, as discussed before, but this strategy can also increase the recognition accuracy compared to other strategies. The possible reason for this strategy's higher accuracy could be that more variation data were included in the model training. From cases 3 and 6 in Table 4.2, we can see that if we remove the Firm Grasping Phase from the training set, the accuracy decreases from 85.5% to 81.01%. The Firm Grasping Phase is essential for training data because it may contain information about the final target gesture. In cases 3 and 7, we find that if we remove the Reaching phase from the training set, the accuracy decreases from 85.5% to 82.51%. This means that the Reaching phase is also vital for training data because it is the progress in which the gesture is formed.

For case 5, the accuracy is only 60.80% when only using the Firm Grasping Phase for training because this period lost much information about gesture formation in Reaching and Early Grasping Phases. For case 8, the accuracy reaches 74.99% only using the sweet period for training because this training data also lost the part of information about the gesture in the Reaching Phase and the Firm Grasping Phase. However, using all phase data for training and the sweet period data for testing achieved the best accuracy, which can be the common practice in real-life situations where training a model is not time-sensitive.

Besides using better strategies for sEMG-based grasp classification, integrating vision is another method to improve recognition accuracy. As we mentioned in the Section 4.3.3, there is an approximately 800 ms time difference between the sweet periods of sEMG and visual information. Therefore, in the real-life application, we can obtain the vision recognition result before processing sEMG classification and integrate these two classification outcomes without causing a delay in myoprosthetic hand

control in the real-life application. After the integration, the accuracy increases from 85.50% to 90.06%, from which we can find that the visual information can effectively increase the overall gesture classification accuracy, thus increasing the performance of the sEMG prosthetic hand.

5.4 Conclusion

In order to reduce the delay of myoprosthetic hand control in a real-life situation while maintaining a high recognition accuracy, we investigated the grasp classification performance during three grasping phases to identify the sweet period. We found that the sEMG sweet period is located between 1.1 s and 1.4 s from the start of the hand grasping, which happens in the Early Grasping phase before the hand is firmly grasped.

Furthermore, we found that using sEMG from all three grasping phases (Reaching, Early Grasping, and Firm Grasping phases) for grasp classification model training achieved the best accuracy. Together with the identified sweet period for testing, the grasp classification accuracy and the response speed of the prosthetic hand can be balanced to achieve high performance.

In order to increase the performance of myoprosthetic hand control in real-life situations by integrating visual information to sEMG, we investigated the object and gesture recognition performance during the entire natural grasp process to identify the sweet period for grasp classification. We found that the vision sweet period is between 0 s and 3.2 s from the start of the hand grasping, which happens in the Reaching phase. Furthermore, we found that visual information can yield a higher

classification accuracy. When integrating gesture recognition and sEMG classification, we improved myoprosthetic hand control performance.

5.5 Limitation and Future Work

This study has some limitations. First, only one target object was chosen by the RetinaNet model to simplify the experiment. In future, we could use gaze tracking technology to identify the target object in prosthetic hand control. Second, we only used one head-mounted camera for capturing the visual information. With the fast developing ubiquitous computing technology, multiple cameras will be available on the prosthetic hand or the ambient environment. Thus we will conduct the actual task, such as data collection, to further improve and validate the system, such as employing multiple channels' visual information to enhance prosthesis control and explore the sweet period/periods. Last, considering this paper focuses on finding the best duration for vision grasp classification and its integration strategies with sEMG, we have not validated our results on the unseen objects in our study. The recognition accuracy might be affected when dealing with unseen objects.

Bibliography

[1] Amputation level. https://www.ottobock.africa. Accessed: 2022-07-13.

[2] Myoelectric prosthetic hand from coapt. https://coaptengineering.com. Accessed: 2022-07-13.

[3] M. Asfour, C. Menon, and X. Jiang. A machine learning processing pipeline for reliable hand gesture classification of fmg signals with stochastic variance. *Sensors*, 21(4):1504, 2021.

[4] D. J. Atkins, D. C. Heard, and W. H. Donovan. Epidemiologic overview of individuals with upper-limb loss and their reported research priorities. *Jpo: Journal of prosthetics and orthotics*, 8(1):2–11, 1996.

[5] J. R. Augustine. *Human Neuroanatomy*. John Wiley & Sons, Incorporated, Hoboken, UNITED STATES, 2008.

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

[7] J. T. Belter, J. L. Segil, and B. SM. Mechanical design and performance specifications of anthropomorphic prosthetic hands: a review. *Journal of rehabilitation research and development*, 50(5):599, 2013.

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

[9] S. Boudaoud, H. Rix, M. Al Harrach, and F. Marin. Robust functional statistics applied to probability density function shape screening of semg data. In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 2213–2216. IEEE, 2014.

[10] H. Bouwsema. Learning to handle a myoelectric upper-limb prosthesis: The development of an evidence-based guideline for training. 2014.

[11] I. M. Bullock, J. Z. Zheng, S. De La Rosa, C. Guertler, and A. M. Dollar. Grasp frequency and usage in daily household and machine shop tasks. *IEEE transactions on haptics*, 6(3):296–308, 2013.

[12] H. Burger and Č. Marinček. Return to work after lower limb amputation. *Disability and rehabilitation*, 29(17):1323–1329, 2007.

[13] S. L. Carey, D. J. Lura, and M. J. Highsmith. Differences in myoelectric and body-powered upper-limb prostheses: Systematic literature review. *Journal of Rehabilitation Research & Development*, 52(3), 2015.

[14] C. Castellini, P. Artemiadis, M. Wininger, A. Ajoudani, M. Alimusaj, A. Bicchi, B. Caputo, W. Craelius, S. Dosen, K. Englehart, et al. Proceedings of the first workshop on peripheral machine interfaces: Going beyond traditional surface electromyography. *Frontiers in neurorobotics*, 8:22, 2014.

[15] C. Castellini and P. Van Der Smagt. Surface emg in advanced hand prosthetics. *Biological cybernetics*, 100(1):35–47, 2009.

[16] J.-T. Chen, S.-M. Chen, T.-S. Kuan, K.-C. Chung, and C.-Z. Hong. Phentolamine effect on the spontaneous electrical activity of active loci in a myofascial trigger spot of rabbit skeletal muscle. *Archives of physical medicine and rehabilitation*, 79(7):790–794, 1998.

[17] X. Chen, X. Zhang, Z.-Y. Zhao, J.-H. Yang, V. Lantz, and K.-Q. Wang. Hand gesture recognition research based on surface emg sensors and 2d-accelerometers. In *2007 11th IEEE International Symposium on Wearable Computers*, pages 11–14. IEEE, 2007.

[18] X. Chen, X. Zhang, Z.-Y. Zhao, J.-H. Yang, V. Lantz, and K.-Q. Wang. Multiple hand gesture recognition based on surface emg signal. In *2007 1st International conference on Bioinformatics and Biomedical Engineering*, pages 506–509. IEEE, 2007.

[19] M. Cognolato, A. Gijsberts, V. Gregori, G. Saetta, K. Giacomino, A.-G. M. Hager, A. Gigli, D. Faccio, C. Tiengo, F. Bassetto, et al. Gaze, visual, myoelectric, and inertial data of grasps for intelligent prosthetics. *Scientific data*, 7(1):1–15, 2020.

[20] B. Crawford, K. Miller, P. Shenoy, and R. Rao. Real-time classification of electromyographic signals for robotic control. In *AAAI*, volume 5, pages 523–528, 2005.

[21] E. Criswell and J. R. Cram. *Cram's introduction to surface electromyography*. Jones and Bartlett, Sudbury, MA, 2nd ed. edition, 2011.

[22] M. R. Cutkosky et al. On grasp choice, grasp models, and the design of hands for manufacturing tasks. *IEEE Transactions on robotics and automation*, 5(3):269–279, 1989.

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

[24] D. Farina, C. Cescon, and R. Merletti. Influence of anatomical, physical, and detection-system parameters on surface emg. *Biological cybernetics*, 86(6):445–456, 2002.

[25] D. Farina, N. Jiang, H. Rehbaum, A. Holobar, B. Graimann, H. Dietl, and O. C. Aszmann. The extraction of neural information from the surface emg for the control of upper-limb prostheses: emerging avenues and challenges. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 22(4):797–809, 2014.

[26] T. Feix, J. Romero, H.-B. Schmiedmayer, A. M. Dollar, and D. Kragic. The grasp taxonomy of human grasp types. *IEEE Transactions on human-machine systems*, 46(1):66–77, 2015.

[27] A. L. Fougner. Proportional myoelectric control of a multifunction upper-limb prosthesis. Master's thesis, Institutt for teknisk kybernetikk, 2007.

[28] J. Friedman, T. Hastie, R. Tibshirani, et al. *The elements of statistical learning*, volume 1. Springer series in statistics New York, 2001.

[29] P. Geethanjali. Myoelectric control of prosthetic hands: state-of-the-art review. *Medical Devices (Auckland, NZ)*, 9:247, 2016.

[30] G. Ghazaei, A. Alameer, P. Degenaar, G. Morgan, and K. Nazarpour. Deep learning-based artificial vision for grasp classification in myoelectric hands. *Journal of neural engineering*, 14(3):036025, 2017.

[31] C. L. Gray, R. C. Lyle, R. J. McGuire, and D. F. Peck. Electrode placement, emg feedback, and relaxation for tension headaches. *Behaviour Research and Therapy*, 18(1):19–23, 1980.

[32] V. Gregori. *An Analysis of the Visuomotor Behavior of Upper Limb Amputees to Improve Prosthetic Control*. PhD thesis, Ph. D. thesis). University of Rome,"La Sapienza, 2019.

[33] Y. Hao, M. Controzzi, C. Cipriani, D. B. Popovic, X. Yang, W. Chen, X. Zheng, and M. C. Carrozza. Controlling hand-assistive devices: utilizing electrooculography as a substitute for vision. *IEEE Robotics & Automation Magazine*, 20(1):40–52, 2013.

[34] J. S. Hebert, Q. A. Boser, A. M. Valevicius, H. Tanikawa, E. B. Lavoie, A. H. Vette, P. M. Pilarski, and C. S. Chapman. Quantitative eye gaze and movement differences in visuomotor adaptations to varying task demands among upper-extremity prosthesis users. *JAMA network open*, 2(9):e1911197–e1911197, 2019.

[35] H. Hermens, S. Stramigioli, H. Rietman, P. Veltink, and S. Misra. Myoelectric forearm prostheses: State of the art from a user-centered perspective. *J. Rehabil. Res. Dev.*, 48(6):719, 2011.

[36] H. J. Hermens, B. Freriks, C. Disselhorst-Klug, and G. Rau. Development of recommendations for semg sensors and sensor placement procedures. *Journal of electromyography and Kinesiology*, 10(5):361–374, 2000.

[37] I. Herrera-Luna, E. J. Rechy-Ramirez, H. V. Rios-Figueroa, and A. Marin-Hernandez. Sensor fusion used in applications for hand rehabilitation: A systematic review. *IEEE Sensors Journal*, 19(10):3581–3592, 2019.

[38] H.-P. Huang and C.-Y. Chen. Development of a myoelectric discrimination system for a multi-degree prosthetic hand. In *Proceedings 1999 IEEE International Conference on Robotics and Automation (Cat. No. 99CH36288C)*, volume 3, pages 2392–2397. IEEE, 1999.

[39] M. E. Huang, C. E. Levy, and J. B. Webster. Acquired limb deficiencies. 3. prosthetic components, prescriptions, and indications. *Archives of physical medicine and rehabilitation*, 82(3):S17–S24, 2001.

[40] B. Hudgins, P. Parker, and R. N. Scott. A new strategy for multifunction myoelectric control. *IEEE transactions on biomedical engineering*, 40(1):82–94, 1993.

[41] F. Hundhausen, D. Megerle, and T. Asfour. Resource-aware object classification and segmentation for semi-autonomous grasping with prosthetic hands. In *2019 IEEE-RAS 19th International Conference on Humanoid Robots (Humanoids)*, pages 215–221. IEEE, 2019.

[42] G. James, D. Witten, T. Hastie, and R. Tibshirani. *An introduction to statistical learning*, volume 112. Springer, 2013.

[43] C. H. Jang, H. S. Yang, H. E. Yang, S. Y. Lee, J. W. Kwon, B. D. Yun, J. Y. Choi, S. N. Kim, and H. W. Jeong. A survey on activities of daily living and occupations of upper extremity amputees. *Annals of rehabilitation medicine*, 35(6):907–921, 2011.

[44] X. Jiang, L.-K. Merhi, and C. Menon. Force exertion affects grasp classification using force myography. *IEEE Transactions on Human-Machine Systems*, 48(2):219–226, 2017.

[45] X. Jiang, L.-K. Merhi, Z. G. Xiao, and C. Menon. Exploration of force myography and surface electromyography in hand gesture classification. *Medical engineering & physics*, 41:63–73, 2017.

[46] R. S. Johansson, G. Westling, A. Bäckström, and J. R. Flanagan. Eye–hand coordination in object manipulation. *Journal of neuroscience*, 21(17):6917–6932, 2001.

[47] J. J. A. M. Junior, M. L. Freitas, H. V. Siqueira, A. E. Lazzaretti, S. F. Pichorim, and S. L. Stevan Jr. Feature selection and dimensionality reduction: An extensive comparison in hand gesture classification by semg in eight channels armband approach. *Biomedical Signal Processing and Control*, 59:101920, 2020.

[48] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu. Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*, 30, 2017.

[49] K. S. Kim, H. H. Choi, C. S. Moon, and C. W. Mun. Comparison of k-nearest neighbor, quadratic discriminant and linear discriminant analysis in classification of electromyogram signals based on the wrist-motion directions. *Current applied physics*, 11(3):740–745, 2011.

[50] I. Kuzborskij, A. Gijsberts, and B. Caputo. On the challenge of classifying 52 hand movements from surface electromyography. In *2012 annual international conference of the IEEE engineering in medicine and biology society*, pages 4931–4937. IEEE, 2012.

[51] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár. Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision*, pages 2980–2988, 2017.

[52] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick. Microsoft coco: Common objects in context. In *European conference on computer vision*, pages 740–755. Springer, 2014.

[53] B. Maat, G. Smit, D. Plettenburg, and P. Breedveld. Passive prosthetic hands and tools: A literature review. *Prosthetics and orthotics international*, 42(1):66–74, 2018.

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

[55] M. Markovic, S. Dosen, C. Cipriani, D. Popovic, and D. Farina. Stereovision and augmented reality for closed-loop control of grasping in hand prostheses. *Journal of neural engineering*, 11(4):046001, 2014.

[56] M. Markovic, S. Dosen, D. Popovic, B. Graimann, and D. Farina. Sensor fusion and computer vision for context-aware control of a multi degree-of-freedom prosthesis. *Journal of neural engineering*, 12(6):066022, 2015.

[57] C. Mason, J. Gomez, and T. J. Ebner. Primary motor cortex neuronal discharge during reach-to-grasp: controlling the hand as a unit. *Archives italiennes de biologie*, 140(3):229–236, 2002.

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

[59] N. Miljković and M. S. Isaković. Effect of the semg electrode (re) placement and feature set size on the hand movement recognition. *Biomedical Signal Processing and Control*, 64:102292, 2021.

[60] T. Mitchell, B. Buchanan, G. DeJong, T. Dietterich, P. Rosenbloom, and A. Waibel. Machine learning. *Annual review of computer science*, 4(1):417–433, 1990.

[61] M. Mohri, A. Rostamizadeh, and A. Talwalkar. *Foundations of machine learning*. MIT press, 2018.

[62] D. Nam, J. M. Cha, and K. Park. Next-generation wearable biosensors developed with flexible bio-chips. *Micromachines*, 12(1):64, 2021.

[63] J. G. Ngeo, T. Tamei, and T. Shibata. Continuous and simultaneous estimation of finger kinematics using inputs from an emg-to-muscle activation model. *Journal of neuroengineering and rehabilitation*, 11(1):1–14, 2014.

[64] M. Niedernhuber, D. G. Barone, and B. Lenggenhager. Prostheses as extensions of the body: Progress and challenges. *Neuroscience & Biobehavioral Reviews*, 92:1–6, 2018.

[65] P. O'Neill, E. L. Morin, and R. N. Scott. Myoelectric signal characteristics from muscles in residual upper limbs. *IEEE Transactions on Rehabilitation Engineering*, 2(4):266–270, 1994.

[66] A. Phinyomark, P. Phukpattaranont, and C. Limsakul. Feature reduction and selection for emg signal classification. *Expert systems with applications*, 39(8):7420–7431, 2012.

[67] A. Phinyomark, P. Phukpattaranont, and C. Limsakul. Feature reduction and selection for emg signal classification. *Expert systems with applications*, 39(8):7420–7431, 2012.

[68] L. Resnik, H. H. Huang, A. Winslow, D. L. Crouch, F. Zhang, and N. Wolk. Evaluation of EMG pattern recognition for upper limb prosthesis control: a case study in comparison with direct myoelectric control. *Journal of NeuroEngineering and Rehabilitation*, 15(1):23, Mar. 2018.

[69] L. Resnik, M. R. Meucci, S. Lieberman-Klinger, C. Fantini, D. L. Kelty, R. Disla, and N. Sasson. Advanced upper limb prosthetic devices: implications for upper limb prosthetic rehabilitation. *Archives of physical medicine and rehabilitation*, 93(4):710–717, 2012.

[70] G. Rizzolatti and G. Luppino. The cortical motor system. *Neuron*, 31(6):889–901, 2001.

[71] A. D. Roche, H. Rehbaum, D. Farina, and O. C. Aszmann. Prosthetic myoelectric control strategies: a clinical perspective. *Current Surgery Reports*, 2(3):1–11, 2014.

[72] F. C. Sebelius, B. N. Rosen, and G. N. Lundborg. Refined myoelectric control in below-elbow amputees using artificial neural networks and a data glove. *The Journal of hand surgery*, 30(4):780–789, 2005.

[73] S. Shalev-Shwartz and S. Ben-David. *Understanding machine learning: From theory to algorithms*. Cambridge university press, 2014.

[74] A. M. Simon, K. L. Turner, L. A. Miller, L. J. Hargrove, and T. A. Kuiken. Pattern recognition and direct control home use of a multi-articulating hand prosthesis. In *2019 IEEE 16th International Conference on Rehabilitation Robotics (ICORR)*, pages 386–391. IEEE, 2019.

[75] D. G. Simons and J. R. Dexter. Comparison of local twitch responses elicited by palpitation and needling of myofascial trigger points. *Journal of Musculoskeletal Pain*, 3(1):49–61, 1995.

[76] A. Soares, A. Andrade, E. Lamounier, and R. Carrijo. The development of a virtual myoelectric prosthesis controlled by an emg pattern recognition system based on neural networks. *Journal of Intelligent Information Systems*, 21(2):127–141, 2003.

[77] M. Sobuh, L. P. Kenney, A. J. Galpin, S. B. Thies, J. McLaughlin, J. Kulkarni, and P. Kyberd. Visuomotor behaviours when using a myoelectric prosthesis. *Journal of neuroengineering and rehabilitation*, 11(1):1–11, 2014.

[78] A. Stango, F. Negro, and D. Farina. Spatial correlation of high density emg signals provides features robust to electrode number and shift in pattern recognition for myocontrol. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 23(2):189–198, 2014.

[79] T. Supuk, T. Bajd, and G. Kurillo. Assessment of reach-to-grasp trajectories toward stationary objects. *Clinical biomechanics*, 26(8):811–818, 2011.

[80] L. T. Taverne, M. Cognolato, T. Bützer, R. Gassert, and O. Hilliges. Video-based prediction of hand-grasp preshaping with application to prosthesis control. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 4975–4982. IEEE, 2019.

[81] F. Ummar A. and I. Nisheena V. The upper limb invariant myoelectric prosthetic control: A review. In *2020 International Conference on Futuristic Technologies in Control Systems Renewable Energy (ICFCR)*, pages 1–4, 2020.

[82] D. Van Der Riet, R. Stopforth, G. Bright, and O. Diegel. An overview and comparison of upper limb prosthetics. *2013 Africon*, pages 1–8, 2013.

[83] S. Wang, J. Zheng, B. Zheng, and X. Jiang. Phase-based grasp classification for prosthetic hand control using semg. *Biosensors*, 12(2):57, 2022.

[84] J. Wu, X. Li, W. Liu, and Z. J. Wang. semg signal processing methods: A review. In *Journal of Physics: Conference Series*, volume 1237, page 032008. IOP Publishing, 2019.

[85] A. J. Young, L. J. Hargrove, and T. A. Kuiken. The effects of electrode size and orientation on the sensitivity of myoelectric pattern recognition systems to electrode shift. *IEEE Transactions on Biomedical Engineering*, 58(9):2537–2544, 2011.

[86] M. Zardoshti-Kermani, B. C. Wheeler, K. Badie, and R. M. Hashemi. Emg feature evaluation for movement control of upper extremity prostheses. *IEEE Transactions on Rehabilitation Engineering*, 3(4):324–333, 1995.

[87] M. Zecca, S. Micera, M. C. Carrozza, and P. Dario. Control of multifunctional prosthetic hands by processing the electromyographic signal. *Critical Reviews^{TM} in Biomedical Engineering*, 30(4-6), 2002.

[88] C. Zhang, X. Chen, S. Cao, X. Zhang, and X. Chen. A novel hd-semg preprocessing method integrating muscle activation heterogeneity analysis and kurtosis-guided filtering for high-accuracy joint force estimation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 27(9):1920–1930, 2019.

[89] X. Zhang, Z. Huang, J. Zheng, S. Wang, and X. Jiang. Dcnngrasp: Towards accurate grasp pattern recognition with adaptive regularizer learning. *arXiv preprint arXiv:2205.05218*, 2022.