Decoding grasping-related muscle activations using deep and shallow machine learning models

Tommaso Portesan

Master Thesis – Applied Cognitive Neuroscience

S6209785
June 2025
Department of Psychology
University of Groningen
Examiner/Daily supervisor:
A.I. Andreea Sburlea, Dr

ABSTRACT:

Grasping is a critical motor function that involves complex neuromuscular interactions, with significant potential to advance fields such as prosthetics and rehabilitation. This master's thesis explores the decoding of grasping-related muscle activations through the application of deep and shallow machine learning models. The main objectives were to elucidate the muscular mechanisms underlying grasping and to assess the distinguishability of grasp types and object properties using electromyography (EMG) signals. EMG data were previously collected from 16 healthy participants performing tasks with three grasp types (power grasp, five-finger precision, two-finger precision) on four objects (large/small cylinders and spheres). The EMG analysis was based on five distinct phases of the grasping movement: fixation, observation, planning/execution, holding, and releasing. The methodology employed Representational Dissimilarity Matrices (RDMs) to evaluate the distinctiveness of muscle activation patterns, Linear Discriminant Analysis (LDA) for shallow classification, and Convolutional Neural Networks (CNNs) for deep learning-based classification. Key findings indicate that muscle activity patterns were most distinct during the holding, planning/execution, and releasing phases, with CNNs achieving up to 88.77% accuracy (chance level 33%) in classifying grasp types. However, object type classification was less successful (peak accuracy of 43.29%, chance level 25%), suggesting that EMG signals better reflect hand configurations than object characteristics. These results highlight the efficacy of machine learning in interpreting grasping dynamics and point to promising applications in developing advanced prosthetic systems, neuromotor rehabilitation, and brain-computer interfaces, offering new avenues to enhance motor control technologies.

1. Introduction

Grasping, the act of gripping objects, is a fundamental motor skill that requires complex interaction between the central nervous system and the muscular system (Castiello, 2005). This action, essential in daily life and clinical fields such as rehabilitation and prosthesis development, involves the coordination of sensory, cognitive, and motor processes (Jeannerod et al., 1994). The importance of studying grasping lies not only in understanding the underlying neuromuscular mechanisms but also in its potential practical applications, such as improving brain-computer interfaces and motor control systems (Omedes et al., 2018), (Spataro et al., 2017), (Zhang et al., 2023). Notably, Napier's research revealed that, despite significant variations in movement elements like force, posture, duration, and speed, the control principles guiding grasping remain strikingly elegant. His work emphasized that the intended purpose of an action dictates the grip type, as seen in the distinct grasps used for writing with a pen versus placing it into a box (Napier, 1956).

Building on this foundation, past studies have demonstrated high accuracy in classifying different hand grips through the analysis of electromyographic (EMG) activity, as evidenced by research such as Miften et al. (2021), who developed a framework for multi-category hand grasp classification, Wang et al. (2022), who focused on phase-based grasp classification for prosthetic control, and Batzianoulis et al. (2017), who decoded grasp gestures in reaching-to-grasping motions using EMG signals. This classification accuracy aligns with literature that shows the differentiation of EMG activity across various types of grasping.

While previous studies have successfully classified different hand grips using EMG signals, there remains a need to understand the temporal evolution of muscle activity patterns across the various phases of grasping and how these patterns are influenced by specific object properties and grasp types, which this thesis aims to address through a detailed analysis of EMG data using advanced machine learning techniques.

To analyze in detail the muscular dynamics associated with grasping, electromyography (EMG) data is a crucial tool (Grosse et al., 2002). EMG allows for the direct measurement of muscle electrical activity during movement, providing insights into muscle activation patterns (Day, 2002).

In this work, we analyzed the EMG data collected during a grasping experiment previously reported in Sburlea et al. (2021), where participants performed different types of grasps on various objects. Specifically, 12 experimental conditions were considered, resulting from the combination of 4 object types (large cylinder - LC, large sphere - LS, small cylinder -

SC, small sphere - SS) and 3 grasp types (power grasp, five-finger precision grasp, two-finger precision grasp). The EMG data, recorded from 8 channels placed on the arm, were acquired at a sampling rate of 200 Hz for a duration of 15 seconds per trial, but only 12 seconds were analysed, since the “rest” phase was excluded, resulting in roughly approximately 2400 samples. These samples were divided into 5 distinct phases: Fixation (0-2 s), Observation (2-5 s), Planning_Execution (5-8 s), Holding (8-10 s), and Releasing (10-12 s). This temporal structure allows for the exploration of muscle activity evolution through the different phases of the movement.

The study of Sburlea et al. (2021) multimodal data was simultaneously collected, containing EEG, EMG, and kinematic activity associated with the grasping conditions. Their results showed that EEG signals enable the decoding of both the type of grasp (power, five- and two-finger precision grasps) and the intrinsic properties of the object (shape and size) in the observation, planning/execution and release phases.

In this study, we aim to gain a better understanding of the muscular mechanisms underlying grasping, finding out whether various grip and object types can also be distinguished based on data from electromyography EMG signals, and potentially inform future applications such as prosthetic control about the features that can lead to good discriminability among objects and among grasp types at different movement stages. We set three main research goals. First, we will use Representational Dissimilarity Matrices (RDMs) (Kriegeskorte et al., 2008) to assess the dissimilarity between the 12 grasping conditions, based on the activity patterns of the 8 EMG channels. This analysis provides a detailed picture of the relationships between the different combinations of grasp and object. Next, we will employ Linear Discriminant Analysis (LDA) (Fisher, 1936), a shallow model, to classify the different grasping conditions. The LDA will be trained on a 70-30 split of the data, demonstrating its ability to linearly discriminate between classes in multivariate contexts. And, third, we applied Convolutional Neural Networks (CNNs) (Krizhevsky et al., 2017), a deep machine learning model, to identify more complex and non-linear patterns in the EMG data. CNNs are particularly suitable for analyzing temporal and multi-channel signals such as EMG. We hypothesize that during the execution and holding stages, decoding performance for grasp types will be higher due to distinct muscle activation patterns tied to different hand configurations. Additionally, we expect that object properties influencing force or precision requirements will be reflected in the EMG signals, particularly during active grasping phases.

My contribution:

In summary, since in the research of Sburlea et al. (2021) they focused exclusively on the EEG data collected, in this thesis, we analyzed the EMG data collected in their experiment, performing the following 3 analyses:

- Representational Dissimilarity Matrix (RDM), which quantifies how different patterns of data, in this case, muscle activity, are represented across the 12 grasping conditions and stages of movement.
- Linear Discriminant Analysis (LDA), a shallow machine learning method used in this study to classify both grasp types (power, five-finger precision, two-finger precision) and object types (large cylinder, small cylinder, large sphere, small sphere) based on electromyographic (EMG) data.
- Convolutional Neural Networks (CNNs), deep learning models that capture complex, non-linear patterns in temporal EMG signals for more robust classification of grasp and object types.

The main objective of this research is to decode the muscle activations related to grasping and to understand how different grasp types and object properties influence muscle activity. Specifically, we aim to: evaluate the performance of shallow and deep machine learning models in discriminating among grasp types and objects based on EMG data, identify the temporal phases of the movement where muscle activity is most informative for decoding, and analyze the relationships between grasping conditions through the representational dissimilarity approach.

The results of this study can have significant implications in areas such as the development of advanced prosthetics, neuromotor rehabilitation, and brain-computer interfaces, contributing to improving the quality of life for individuals with motor disabilities (Müller-Putz et al., 2019). In summary, this thesis combines shallow and deep machine learning techniques with multivariate analysis to offer an innovative perspective on decoding grasping EMG data, collected in Sburlea et al. (2021), opening new possibilities for research and practical application in the biomedical field.

2. Dataset and Methods

2.1 Participants:

The electromyography (EMG) data analyzed in this study were originally collected by (Sburlea et al., 2021) from a group of 16 healthy individuals. These participants, aged between

20 and 32 years with a mean age of 25.3 years and a standard deviation of 3.5 years, included 9 females. All were right-handed, as confirmed by an adapted version of the Edinburgh Handedness Inventory (Oldfield, 1971). They reported no neurological disorders and had normal or corrected-to-normal vision.

2.2 Experimental Paradigm:

The experimental design centered on capturing muscle activity during grasping tasks, requiring participants to observe and replicate specific hand movements shown on a computer screen. Participants sat facing the screen, with wooden objects positioned centrally between themselves and the display. The task was programmed using Psychtoolbox (version 3) in MATLAB, and a brief practice session preceded the experiment to ensure familiarity with the procedure.

Each trial unfolded over multiple phases, beginning with a fixation period lasting 3 seconds, during which participants maintained their gaze on a central cross while resting their right hand on a mousepad. This was followed by a 4-second observation phase, where an image of a grasp involving one of four objects appeared, allowing participants to study the posture and object characteristics without gaze restrictions. Next, an execution phase began when an ‘x’ symbol prompted participants to replicate the observed grasp within 4 seconds, shifting their focus to the object. After completing the movement, they held the grasp until a “Relax” instruction appeared, marking the holding phase. The trial concluded with a 2-second relaxation phase, followed by a “Get ready” cue to prepare for the next trial.

The experiment consisted of 24 runs, each approximately 7 minutes long, totaling about 3 hours of data collection with short breaks between runs. Participants performed three grasp types: power grasp, a strong grip involving the whole hand to hold objects firmly, five-finger precision, a precise grip using all five fingers, often for manipulating objects requiring detailed control and two-finger precision, a precise grip using only two fingers (the thumb and index finger) with four different objects: a large cylinder (5 cm diameter, 24 cm length), a small cylinder (3 cm diameter, 24 cm length), a large sphere (8 cm radius), or a small sphere (5 cm radius). Each run featured 27 trials with a single object, and every trial lasted 15 seconds. Across the experiment, each combination of grasp type and object was repeated 54 times, presented in random order.

2.3 Data Acquisition:

Muscle activity data were recorded using a Myo armband, a device manufactured by Thalmic Labs Inc. (Ontario, Canada), equipped with eight evenly spaced EMG sensors. The armband was positioned on the right forearm, just below the elbow, targeting the extrinsic hand muscles involved in grasping. Prior to data collection, the device was calibrated for each participant to optimize signal quality, and it transmitted data wirelessly via Bluetooth. Synchronization with other data streams, such as visual stimuli timing, was achieved using the Lab Streaming Layer (LSL), with a photodiode providing precise alignment between EMG signals and screen events.

2.4 Data Preprocessing:

The EMG data underwent preprocessing by Sburlea et al. (2021) using Matlab R2016b (Mathworks, Inc. USA) to ensure suitability for analysis. EMG data were epoched in fourteen-second-long segments relative to the beginning of the trial. The EMG datasets were structured as channels x trials x time samples. The sampling rate is 200Hz, so for 14 seconds, 2800 samples were recorded in theory, but due to some artifacts, the last samples were removed, without creating any problems since the last 2 seconds corresponded to the resting stage. So the total of the samples to be analysed were 2400, resulting in a total of 12 seconds. Trials were then segmented into distinct stages: fixation phase (0–2 s), where participants focused on a fixation point observation phase (2–5 s), where participants looked at the object they were going to grasp, paying attention to details like the object's size and shape, which helped them get ready for the next step, planning and execution phase (5–8 s), in which participants first planned how they would grasp the object based on what they observed and then they carried out the action by reaching for and grasping the object, holding phase (8–10 s), where participants held the object steady for a short time and releasing phase (10–12 s), in which participants let go of the object. These cleaned and segmented EMG data formed the basis for the analyses in this study. All the trials were reordered according to a common order of grasping conditions among subjects. From all types of data, the trials in which the task had been incorrectly executed were rejected. (e.g., movement execution during the observation phase).

The eight EMG data channels were processed using Hilbert transform, standardized using z-score, and, finally, the envelope (power) of the data was computed.

2.5 Representational Dissimilarity Matrix (RDM):

A Representational Dissimilarity Matrix (RDM) is a tool used in neuroscience and related fields to measure and visualize how different patterns of data (e.g., brain activity, muscle signals, or behavioral responses) are across a set of conditions or stimuli. It creates a square matrix where each cell represents the dissimilarity (or distance) between the data patterns of two conditions, typically calculated using metrics like 1 minus Pearson correlation or Euclidean distance. Higher values indicate greater differences between conditions. The matrix provides a comprehensive view of the relationships among conditions, often visualized as a heatmap, and can be further analyzed with techniques like Multidimensional Scaling (MDS) to map these dissimilarities into a lower-dimensional space for easier interpretation. RDMs are valuable for understanding how distinct or similar representations are within a dataset, such as comparing muscle responses across different grasping conditions (Figure 1).

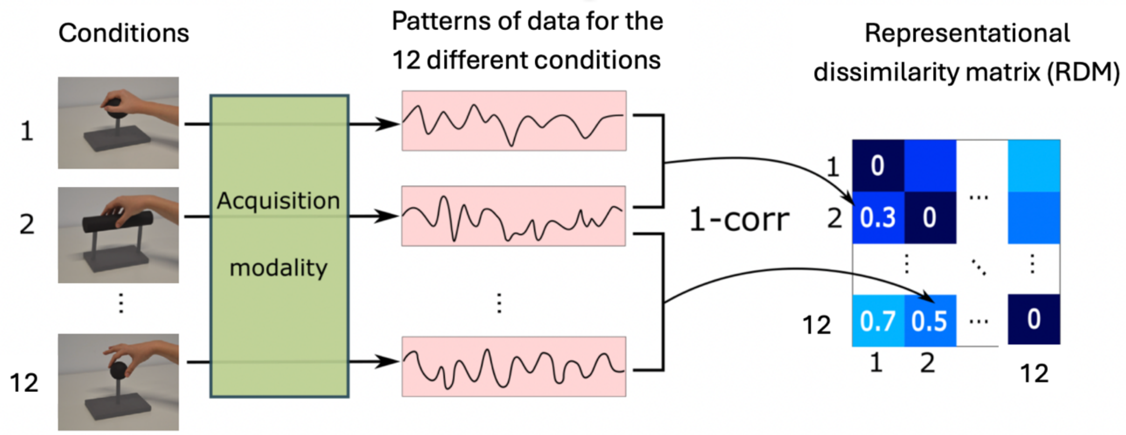

Figure 1: Illustration of the Representational Dissimilarity Matrix (RDM) process, showing 12 different grasping conditions (e.g., various grasp types and objects) with associated EMG data patterns acquired through an acquisition modality. The dissimilarity between patterns is calculated using 1 minus the Pearson correlation coefficient (1-corr), resulting in a 12x12 RDM matrix where higher values (e.g., 0.7) indicate greater dissimilarity between conditions. Figure adapted from Sburlea et al., 2021

The RDM analysis (Kriegeskorte et al., 2008) was performed on electromyographic (EMG) data collected from 16 subjects during grasping tasks involving four objects—Large Cylinder (LC), Small Cylinder (SC), Large Sphere (LS), and Small Sphere (SS)—each grasped with three techniques (power, five-finger precision, two-finger precision), yielding 12 conditions (4 objects × 3 grasp types). EMG signals, recorded across eight channels at 200 Hz, were standardized to a uniform length of up to 2400 samples (12 seconds) based on the minimum sample count across subjects and objects (2393 samples). The data were segmented into five movement phases: fixation (0-2 s), observation (2-5 s), planning/execution (5-8 s), holding (8-10 s), and releasing (10-11.965 s). Data were organized by object, and condition-specific trials were identified using indices from a separate file (indexes_trials.mat). Because trial indices were global while data were stored per object, a cumulative offset method was applied to correctly map indices to conditions.

For each condition, the EMG signals were averaged across trials for each of the 8 channels, resulting in an average time series per channel per condition. These average time series from all 8 channels were then concatenated to form a single vector for each condition. Pairwise dissimilarities between these vectors were computed using 1 minus the Pearson correlation coefficient, generating a 12 × 12 Representational Dissimilarity Matrix (RDM) for each subject. The RDMs were subsequently averaged across subjects to produce a group-level RDM, which was visualized as a heatmap. Furthermore, the group-level RDM was subjected to Multidimensional Scaling (MDS) analysis to assess the distinctiveness of the conditions, with the choice between 2D and 3D representations guided by stress values.

This analysis allowed for a quantitative assessment of how distinct muscle activation patterns are across different grasping conditions.

2.6 Linear Discriminant Analysis (LDA):

Linear Discriminant Analysis (LDA) is a statistical technique employed for classification and dimensionality reduction. Originally introduced by Fisher (1936), LDA assumes that the data within each class follows a Gaussian distribution and that all classes share the same covariance matrix. The primary objective of the method is to identify a linear projection of the data that maximizes class separation, achieved by optimizing the ratio between the between-class variance and the within-class variance.

To investigate the classification of grasp types across different movement phases, we performed a standard Linear Discriminant Analysis (LDA) on electromyographic (EMG) data collected from 16 subjects performing grasping tasks. For each phase, EMG data were loaded from subject-specific files (e.g., EMG_Data_LC_S3.mat), and approximately 54 trials per condition per subject were extracted using predefined indices from indexes_trials.mat, mapped via a cumulative offset system to align global indices with object-specific data. Features were derived by averaging the EMG signals across the temporal dimension for each trial and channel, yielding a feature matrix of size [number of trials × 8 channels] per phase, with corresponding grasp type labels (power, five-finger precision, two-finger precision) assigned to each trial. The data were then split into training (70%) and testing (30%) sets using a non-stratified holdout method with a fixed random seed for reproducibility. An LDA model was trained on the training set to classify grasp types based on the extracted features, and its performance was evaluated on the test set by calculating the classification accuracy as the proportion of correctly predicted grasp types, reported as a percentage for each phase. This analysis allowed us to assess the discriminability of grasp types and object types across the different phases of the grasping movement, providing insights into how muscle activity patterns vary with grasp and object type during distinct temporal stages.

2.7 Convolutional Neural Networks (CNNs):

Convolutional Neural Networks (CNNs) are a specialized class of deep learning models designed to process structured data, such as images or time-series signals, by leveraging convolutional layers that detect local patterns through the application of filters. Introduced by LeCun et al. (1989), CNNs utilize hierarchical feature extraction, where initial layers capture low-level features (e.g., edges or temporal variations) and deeper layers combine these into more complex representations. This architecture, combined with pooling layers to reduce spatial dimensions and fully connected layers for classification, enables efficient learning and high performance in tasks like pattern recognition, making CNNs a cornerstone of modern machine learning.

To further explore the classification of grasp types across different movement phases, we employed a Convolutional Neural Network (CNN) analysis on electromyographic (EMG) data collected from the subjects performing grasping tasks. The data were then split into training (70%) and testing (30%) sets using a non-stratified holdout method with a fixed random seed for reproducibility. A 1D CNN architecture was designed to process the temporal EMG sequences, featuring a sequence input layer accommodating the 8 channels with a minimum sequence length determined dynamically, followed by two convolutional layers (with 16 and 32 filters of size 5, respectively, and 'same' padding to preserve temporal dimensions), each paired with batch normalization and ReLU activation to enhance training stability and introduce non-linearity, a global max-pooling layer to reduce temporal dimensions, two fully connected layers (64 neurons and 3 neurons for the 3 grasp classes), a softmax layer, and a classification layer. The CNN was trained using the Adam optimizer over 20 epochs with a mini-batch size of 32, shuffling the data every epoch, and validating performance on the test set every 30 iterations, with training progress visualized via accuracy and loss plots. The model's performance was evaluated by classifying grasp types on the test set, with accuracy computed as the proportion of correctly predicted labels, reported as a percentage for each phase, providing a robust assessment of grasp type discriminability across the temporal stages of movement and offering insights into the neural encoding of grasping actions.

3. Results

3.1 Representational Dissimilarity Matrix (RDM):

The Representational Dissimilarity Matrix (RDM) analysis was employed to evaluate the distinctiveness of neural representations, derived from electromyographic (EMG) signals, across 12 grasping conditions during five phases of movement: fixation, observation, planning/execution, holding, and releasing. For each phase, a 12×12 RDM was computed, with dissimilarity values reflecting the differentiation between condition pairs, and Multidimensional Scaling (MDS) was applied to assess how well these dissimilarities could be represented in a 2D space, with stress values indicating fit quality. The holding phase demonstrated the most pronounced dissimilarities, ranging from 0.05 to 1.22 (Figure 5 and 10), alongside the lowest MDS stress value of 0.0204, suggesting highly distinct and easily separable neural representations when the subject actively maintains the grasp. Similarly, the planning/execution (Figure 4 and 9) and releasing phases (Figure 6 and 11) exhibited notable dissimilarities, ranging from 0.03 to 0.43 and 0.03 to 0.42 respectively, with low stress values of 0.0491 and 0.0479, indicating clear differentiation during the transition into and out of the grasp. In contrast, the fixation (Figure 2 and 7) and observation phases (Figure 3 and 8) showed lower dissimilarities, ranging from 0.04 to 0.17 and 0.07 to 0.23, with higher stress values of 0.1668 and 0.1811, reflecting more overlapping or less distinct neural representations during these preparatory stages. These results suggest that muscle activity patterns, as captured by EMG, are most differentiated during active grasping and its transitions, while preparatory phases involve more similar activation patterns across conditions.

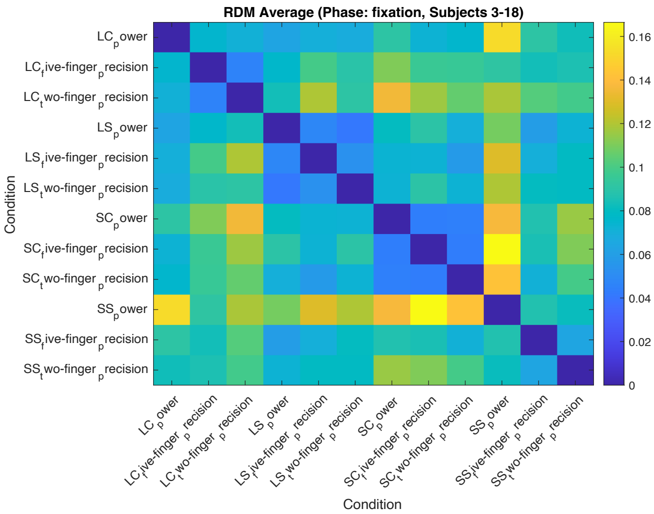

Figure 2: Heatmap of average RDM for fixation phase across conditions and grasp types (0-0.16)

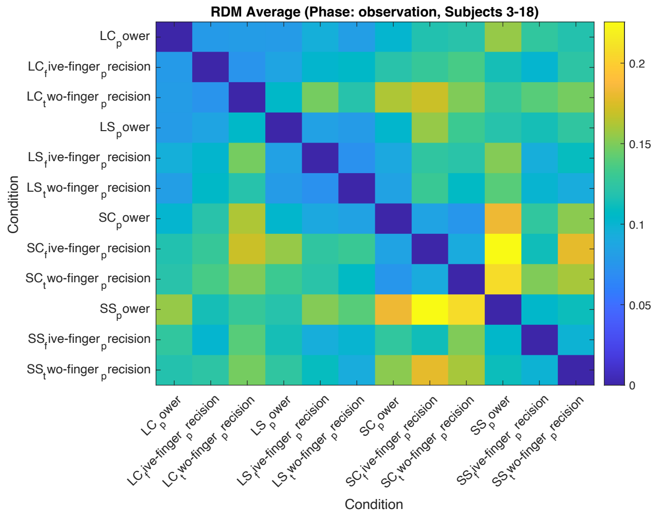

Figure 3: Heatmap of average RDM for observation phase across conditions and grasp types (0-0.16)

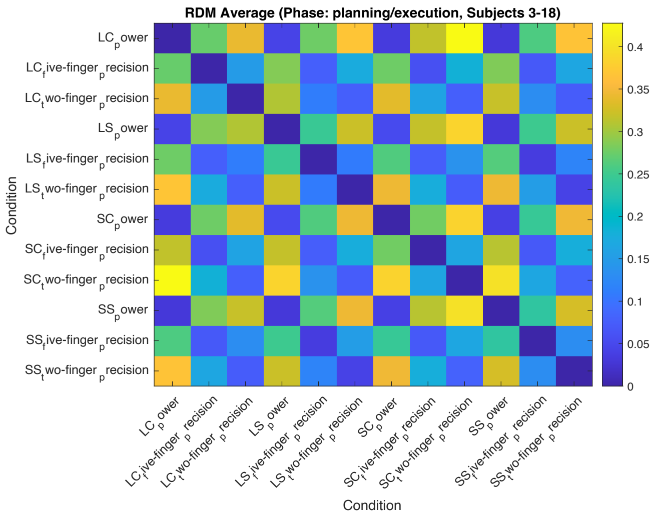

Figure 4: Heatmap of average RDM for planning/execution phase across conditions and grasp types (0-0.4)

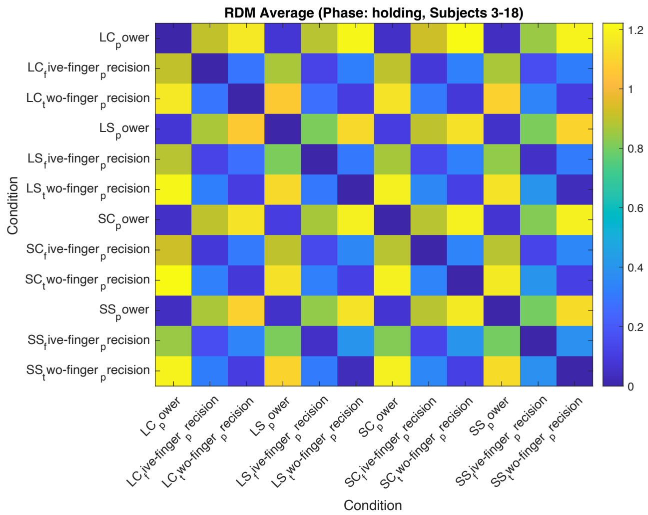

Figure 5: Heatmap of average RDM for holding phase across conditions and grasp types (0-1.2)

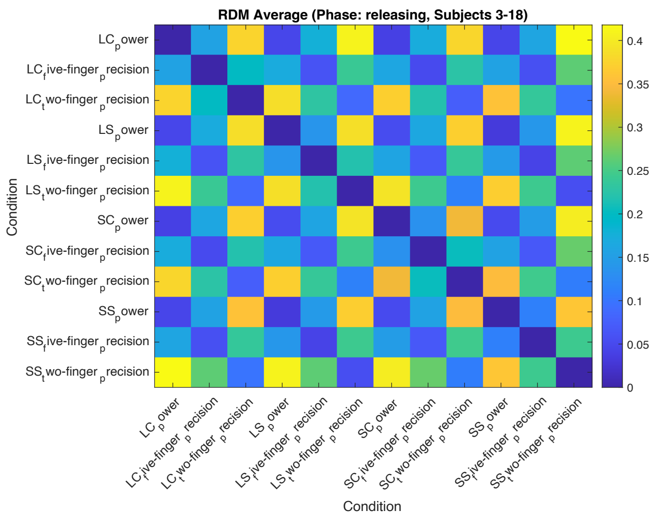

Figure 6: Heatmap of average RDM for releasing phase across conditions and grasp types (0-0.4)

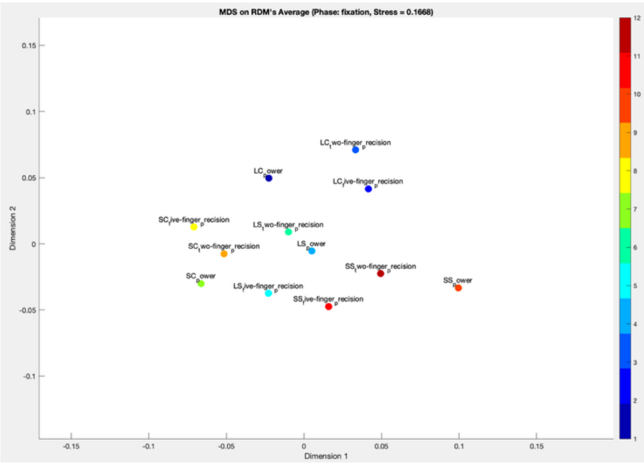

Figure 7: MDS plot, with a stress value of 0.1668, visualizes the clustering of different conditions based on the average RDM during the fixation phase, with colors indicating condition values from 1 to 12

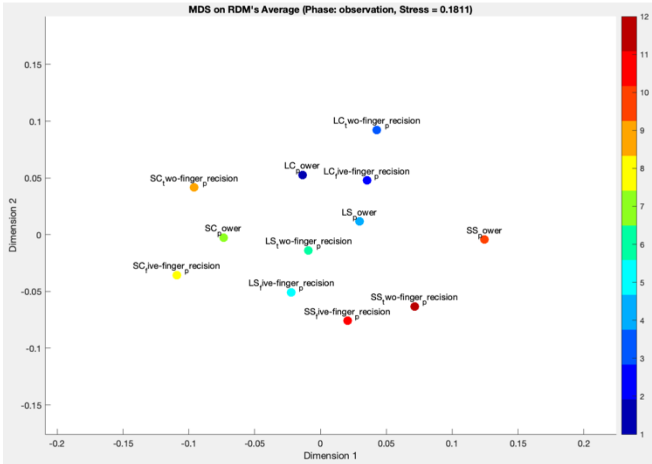

Figure 8: MDS plot, with a stress value of 0.1811, visualizes the clustering of different conditions based on the average RDM during the observation phase, with colors indicating condition values from 1 to 12

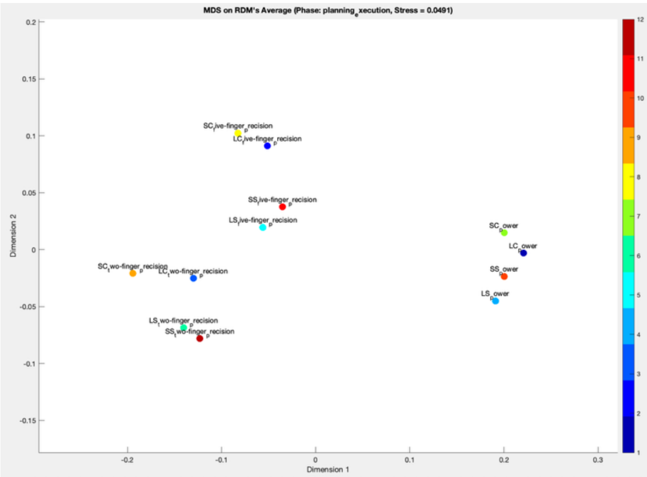

Figure 9: MDS plot, with a stress value of 0.0491, visualizes the clustering of different conditions based on the average RDM during the planning execution phase, with colors indicating condition values from 1 to 12

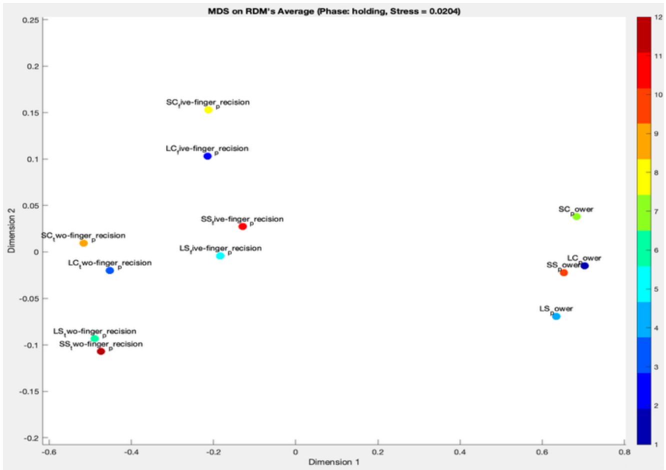

Figure 10: MDS plot, with a stress value of 0.0204, visualizes the clustering of different conditions based on the average RDM during the holding phase, with colors indicating condition values from 1 to 12

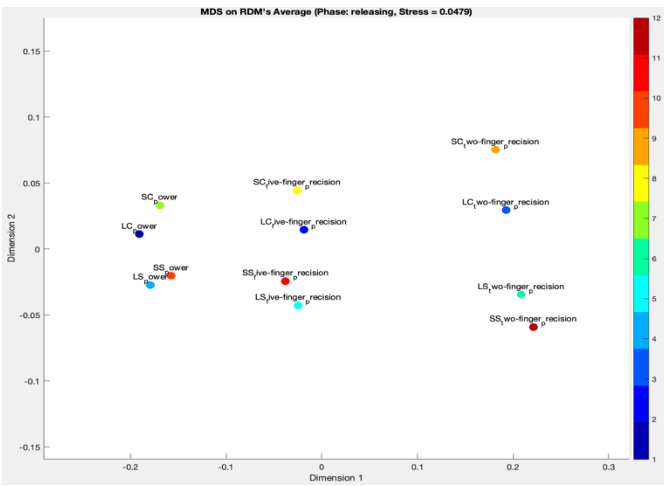

Figure 11: MDS plot, with a stress value of 0.0479, visualizes the clustering of different conditions based on the average RDM during the releasing phase, with colors indicating condition values from 1 to 12

3.2 Linear Discriminant Analysis (LDA):

Linear Discriminant Analysis (LDA) was applied to classify grasp types, namely power, five-finger precision, and two-finger precision, using electromyographic (EMG) data collected from 16 subjects, resulting in a dataset of 2393 samples. Given the three grasp types, the chance level for random classification was 33%. The classification accuracy demonstrated notable variation across the five distinct phases of the grasping movement. A binomial test was used to assess the statistical significance of the classification accuracies against the chance level. The holding phase yielded the highest accuracy at 73.86% (number of successes: 2283 out of 3091, $p<.001$), substantially surpassing the chance level, followed by the planning/execution phase with an accuracy of 65.51% (number of successes: 2025 out of 3091, $p<.001$) and the releasing phase at 57.13% (number of successes: 1766 out of 3091, $p<.001$), both of which also performed significantly above chance. In contrast, the preparatory phases, fixation and observation, produced accuracies of 34.26% (number of successes: 1059 out of 3091, $p=0.1413$) and 34.16% (number of successes: 1056 out of 3091, $p=0.1684$), respectively, which were only slightly above the chance level of 33% and did not achieve statistical significance ($p>0.05$). These results suggest that EMG signals exhibit greater distinctiveness during the active and transitional phases of the movement, indicating more pronounced differentiation in the muscle representations of grasp types during these stages, while the lower accuracies in the initial phases are due to reduced variability or specificity in EMG signals before movement execution. This analysis underscores the effectiveness of LDA in distinguishing grasp types, particularly during dynamic phases, and provides valuable insights into the temporal dynamics of motor control in grasping tasks.

Additionally, LDA was utilized to classify object types, large cylinder (LC), small cylinder (SC), large sphere (LS), and small sphere (SS), using the same EMG dataset from 16 subjects, totaling 2393 samples. With four object types, the chance level for random classification was 25%. A binomial test was employed to assess the statistical significance of the classification accuracies against this chance level. The classification accuracies across the five movement phases were 29.91% for the fixation phase (number of successes: 847 out of 2832, $p<.001$), 29.80% for the observation phase (number of successes: 844 out of 2832, $p<.001$), 30.58% for the planning/execution phase (number of successes: 866 out of 2832, $p<.001$), 31.53% for the holding phase (number of successes: 893 out of 2832, $p<.001$), and 32.49% for the releasing phase (number of successes: 920 out of 2832, $p<.001$). All accuracies significantly exceeded the chance level (p<0.001), demonstrating consistent performance above random guessing across all phases. Despite this, the accuracies remained modest, suggesting that the EMG signals offer limited differentiation for distinguishing between object types.

To further analyse these findings, confusion matrices were computed for each phase (Figures 12 and 13), providing per-condition classification accuracies. For grasp types, the holding phase showed the strongest performance, with power grasp correctly classified at 91.3%, five-finger precision at 69.1%, and two-finger precision at 61.7%, though two-finger precision was often misclassified as five-finger precision (34.3%) and vice versa (26.7%). The planning/execution phase yielded correct predictions of 79% for power, 65.5% for five-finger precision, and 52.6% for two-finger precision, with notable confusion between two-finger and five-finger precision (36.0%). The releasing phase achieved 64.2% for power, 39.1% for five-finger precision, and 67.9% for two-finger precision, while fixation and observation phases showed lower performance (41.5%, 32.7% and 28.8% in fixation and 39.7%, 39.4% and 23.8% in observation). For object types, the holding phase had the highest correct classification for LC at 37.2%, with rates ranging from 21.1% (SS) to 32% (LS), while LC was frequently misclassified as SC (31.8%). The fixation phase showed correct rates from 21.1 % (LS) to 35.8% (LC), with LC confused with LS (38.8%). The planning/execution phase yielded correct predictions of 34.1% for LS, 32.6% for SC and 32.4% for LC, with notable confusion between LS and SS (36.1%). The releasing phase achieved 32% for LC, 31.9% for LS, and 31.2% for SC.

These detailed insights confirm that grasp type discrimination is most robust in active phases, driven by power grasp accuracy, while object type classification remains challenging due to overlapping EMG patterns, particularly between similar objects like LS and SS.

Confusion Matrices for Grasp Type Classification

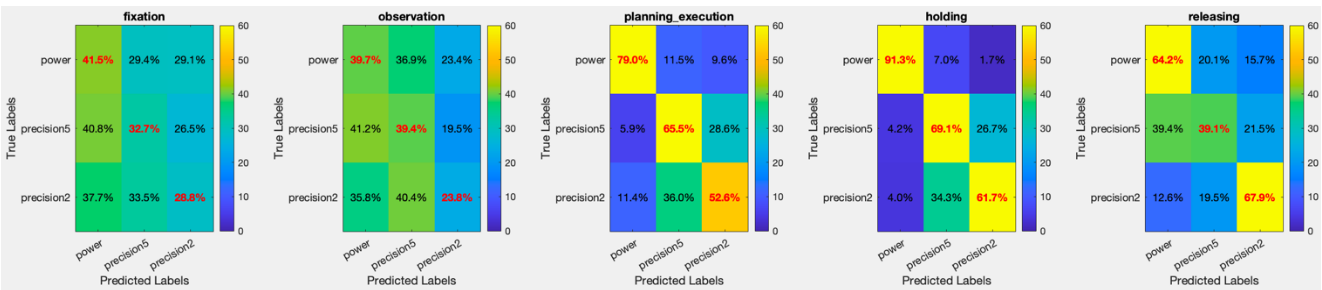

Figure 12: Confusion matrix illustrating the classification of grasp types (power, five-finger precision, two-finger precision) across five phases of the grasping task: fixation, observation, planning/execution, holding, and releasing

Confusion Matrices for Object Type Classification

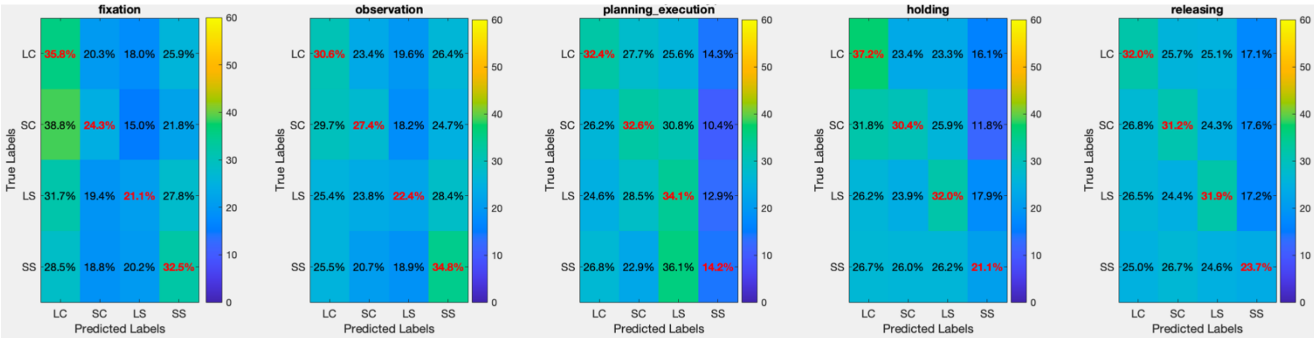

Figure 13: Confusion matrix illustrating the classification of object types (large cylinder, small cylinder, large sphere, small sphere) across five phases of the grasping task: fixation, observation, planning/execution, holding, and releasing

3.3 Convolutional Neural Network (CNN):

The CNN models were trained using a constant learning rate of 0.001, with 225 iterations per epoch and a total of 4,500 iterations, on a single CPU. The learning rate schedule was kept constant throughout all training sessions.

The Convolutional Neural Network (CNN) was employed to classify grasp types across five distinct phases of the task. During the Fixation phase, the model reached a final validation accuracy of 34.62% (Figure 14). In the Observation phase, the final validation accuracy was 34.52% (Figure 15). For the Planning-Execution phase, the network achieved a significantly higher accuracy of 87.64% (Figure 16). Similarly, the Holding phase yielded a final accuracy of 88.77% (Figure 17). Lastly, during the Releasing phase, the CNN attained a validation accuracy of 87.32% (Figure 18). These results indicate notably stronger classification performance in the motor execution-related phases (Planning, Holding, Releasing) compared to the initial perceptual stages (Fixation, Observation).

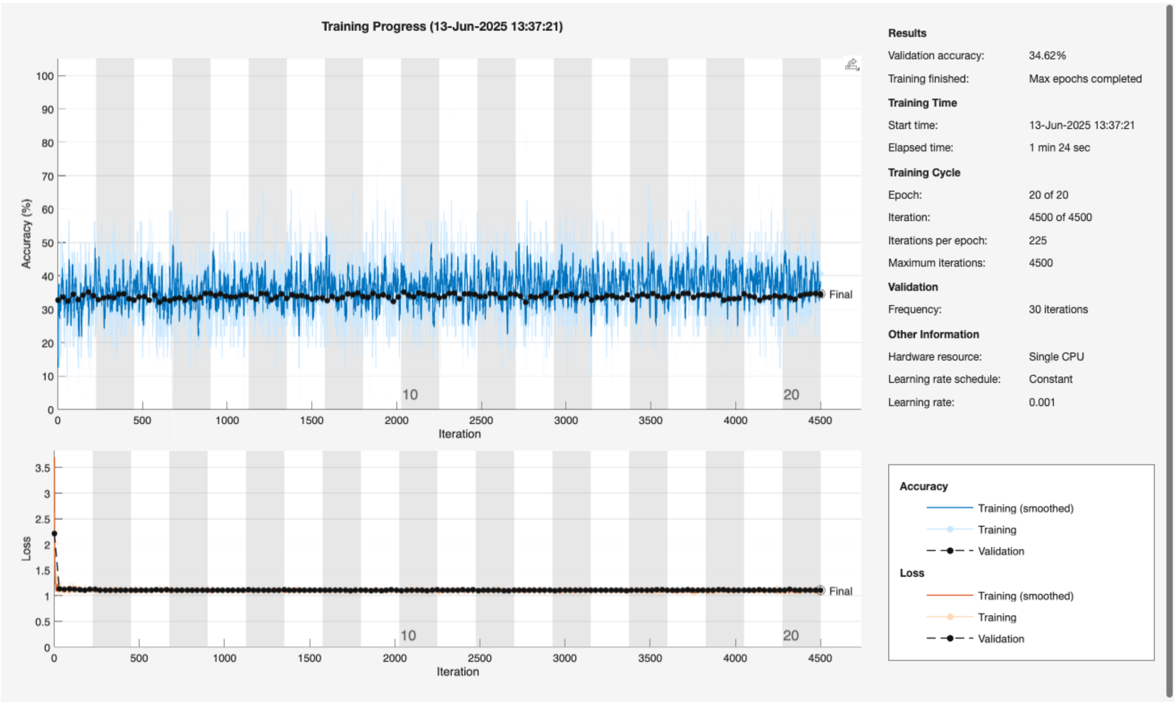

Figure 14: Chart illustrating training progress, showing accuracy and loss over 4500 iterations across 20 epochs using Convolutional Neural Networks for grasp type classification in the fixation phase, with a validation accuracy of 34.62%, trained on a single CPU with a constant learning rate of 0.001

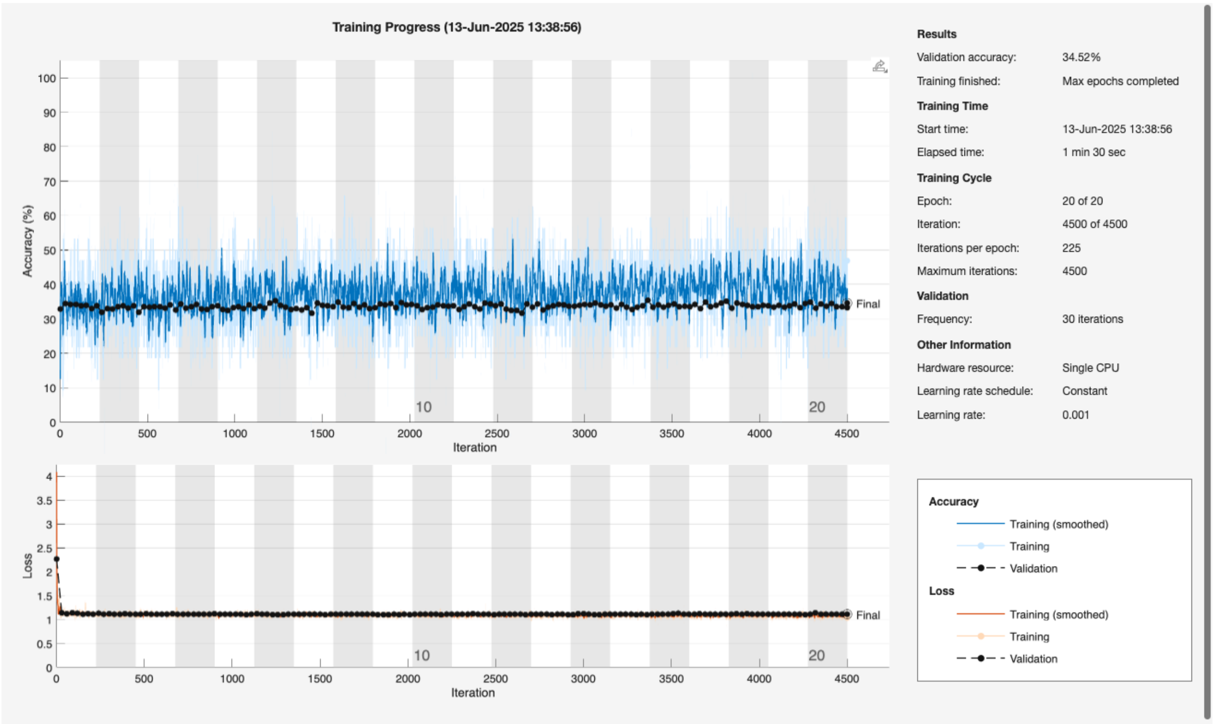

Figure 15: Chart illustrating training progress, showing accuracy and loss over 4500 iterations across 20 epochs using Convolutional Neural Networks for grasp type classification in the observation phase, with a validation accuracy of 34.52%, trained on a single CPU with a constant learning rate of 0.001

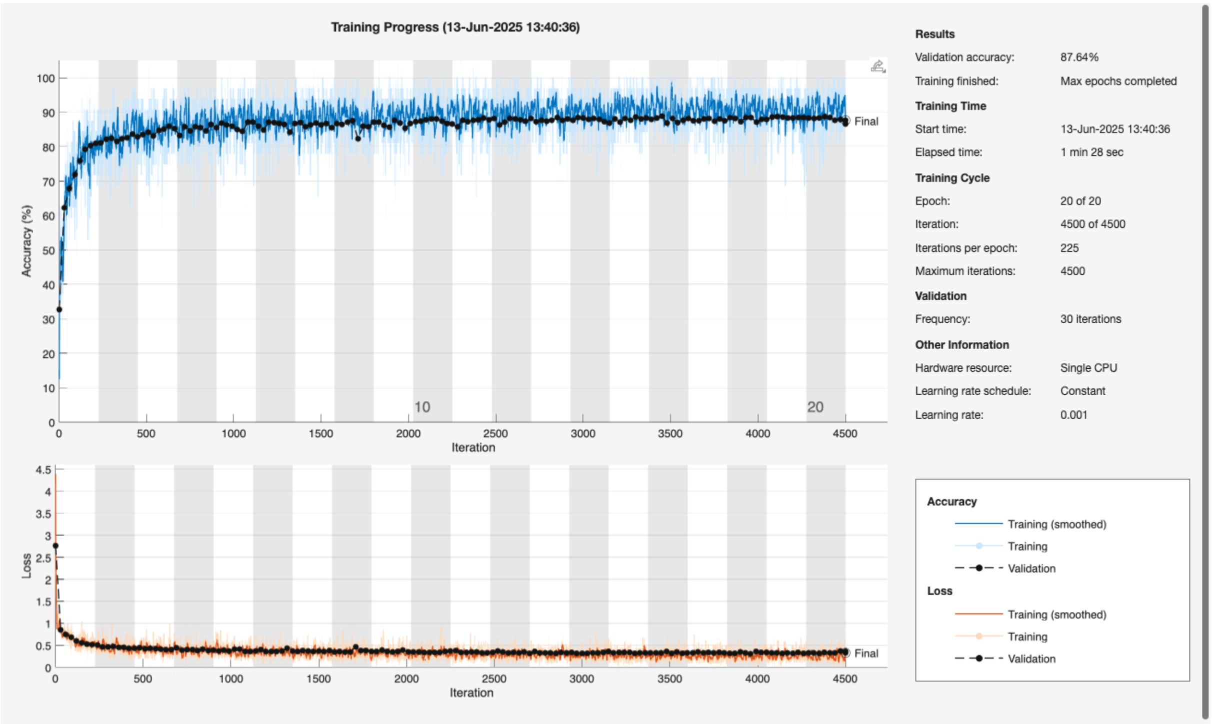

Figure 16: Chart illustrating training progress, showing accuracy and loss over 4500 iterations across 20 epochs using Convolutional Neural Networks for grasp type classification in the planning/execution phase, with a validation accuracy of 87.64%, trained on a single CPU with a constant learning rate of 0.001

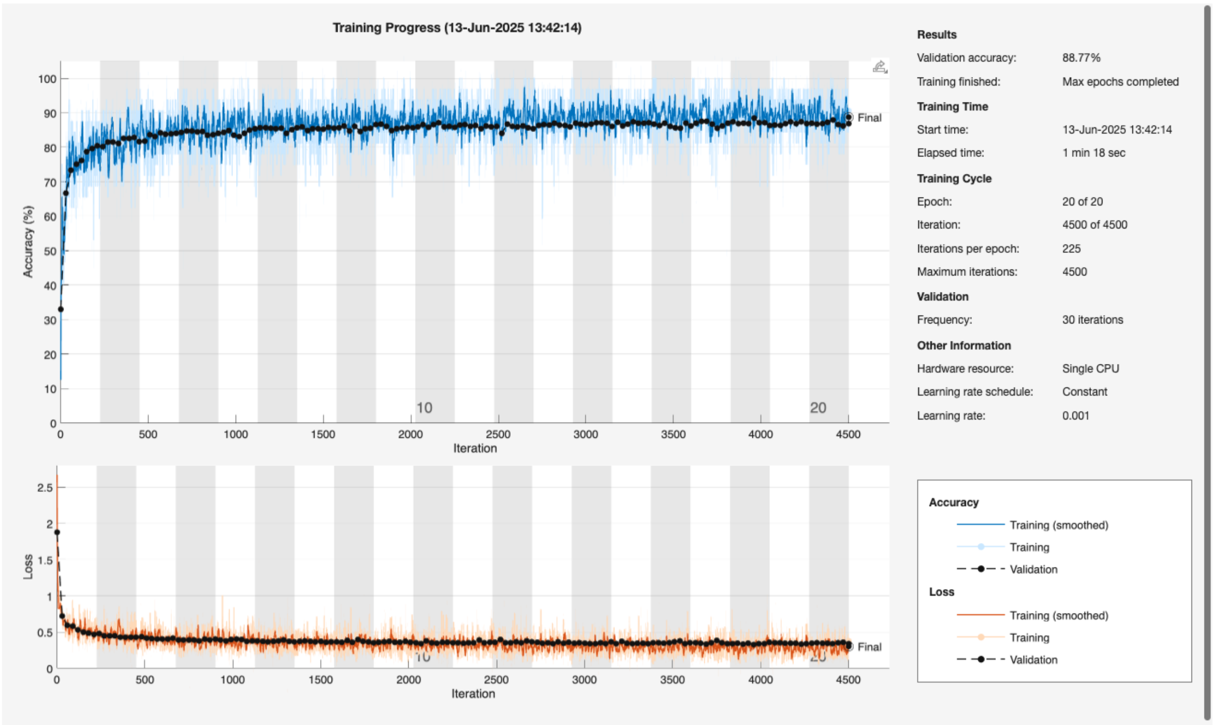

Figure 17: Chart illustrating training progress, showing accuracy and loss over 4500 iterations across 20 epochs using Convolutional Neural Networks for grasp type classification in the holding phase, with a validation accuracy of 88.77%, trained on a single CPU with a constant learning rate of 0.001

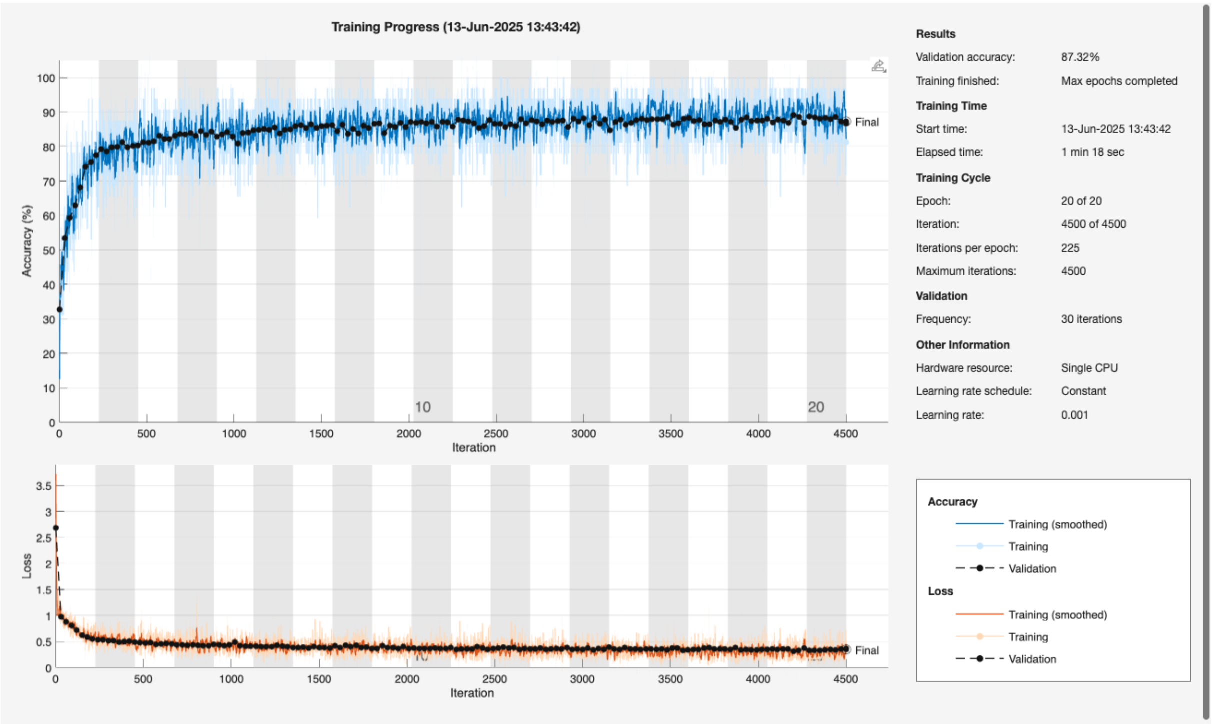

Figure 18: Chart illustrating training progress, showing accuracy and loss over 4500 iterations across 20 epochs using Convolutional Neural Networks for grasp type classification in the releasing phase, with a validation accuracy of 87.32%, trained on a single CPU with a constant learning rate of 0.001

In addition to grasp classification, the CNN was applied to distinguish between different object types throughout the five phases of the task. Performance during the early perceptual stages was modest, with validation accuracies of 30.60% in the Fixation phase and 30.77% in the Observation phase (Figures 19 and 20). A noticeable improvement was observed during the Planning-Execution phase, where the model reached 43.29% accuracy (Figure 21). Classification performance remained higher during the action-oriented phases, with 37.37% in the Holding phase (Figure 22) and 39.63% in the Releasing phase (Figure 23). Although all values are above chance level (25%), they remain considerably lower than the accuracies obtained in the grasp classification, highlighting a stronger discriminability of hand configuration over object identity in this context.

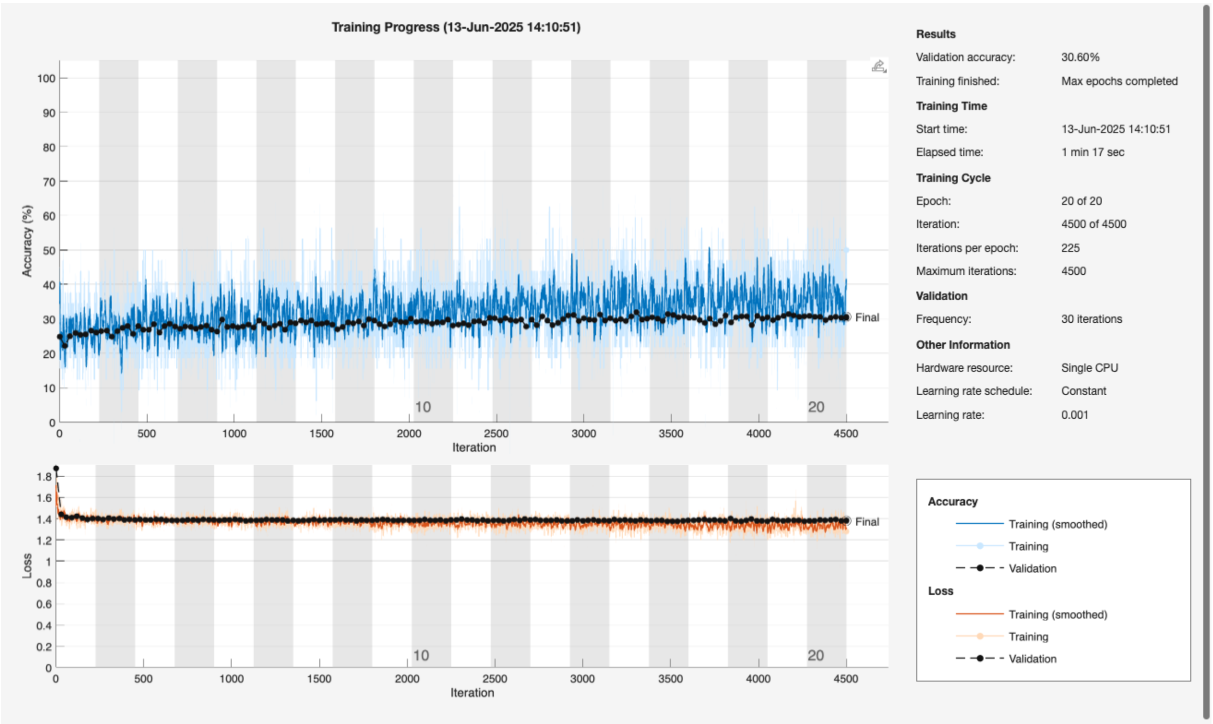

Figure 19: Chart illustrating training progress, showing accuracy and loss over 4500 iterations across 20 epochs using Convolutional Neural Networks for object type classification in the fixation phase, with a validation accuracy of 30.60%, trained on a single CPU with a constant learning rate of 0.001

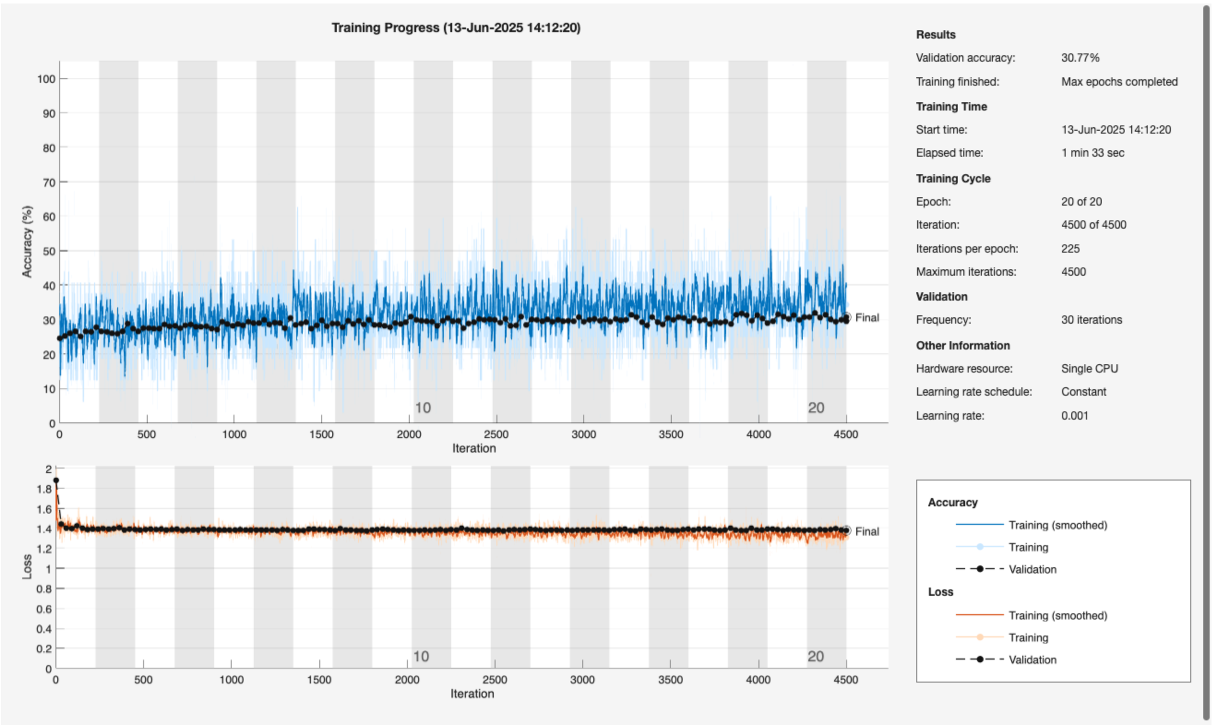

Figure 20: Chart illustrating training progress, showing accuracy and loss over 4500 iterations across 20 epochs using Convolutional Neural Networks for object type classification in the observation phase, with a validation accuracy of 30.77%, trained on a single CPU with a constant learning rate of 0.001

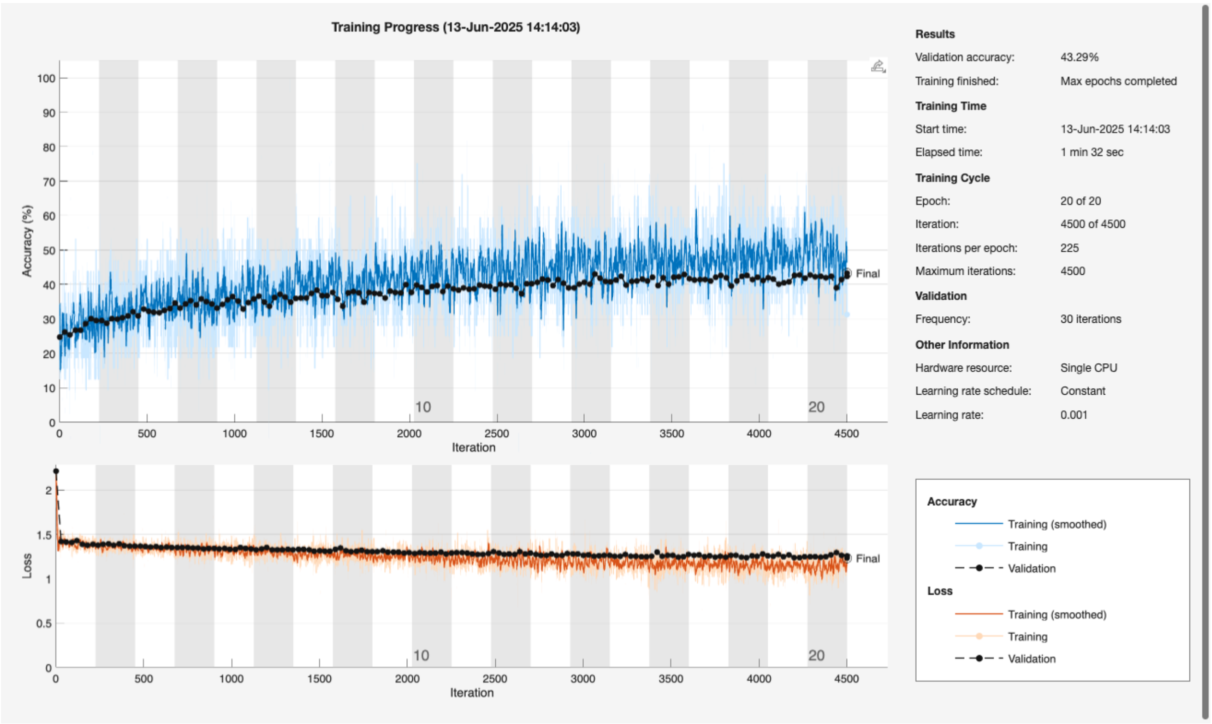

Figure 21: Chart illustrating training progress, showing accuracy and loss over 4500 iterations across 20 epochs using Convolutional Neural Networks for object type classification in the planning/execution phase, with a validation accuracy of 43.29%, trained on a single CPU with a constant learning rate of 0.001

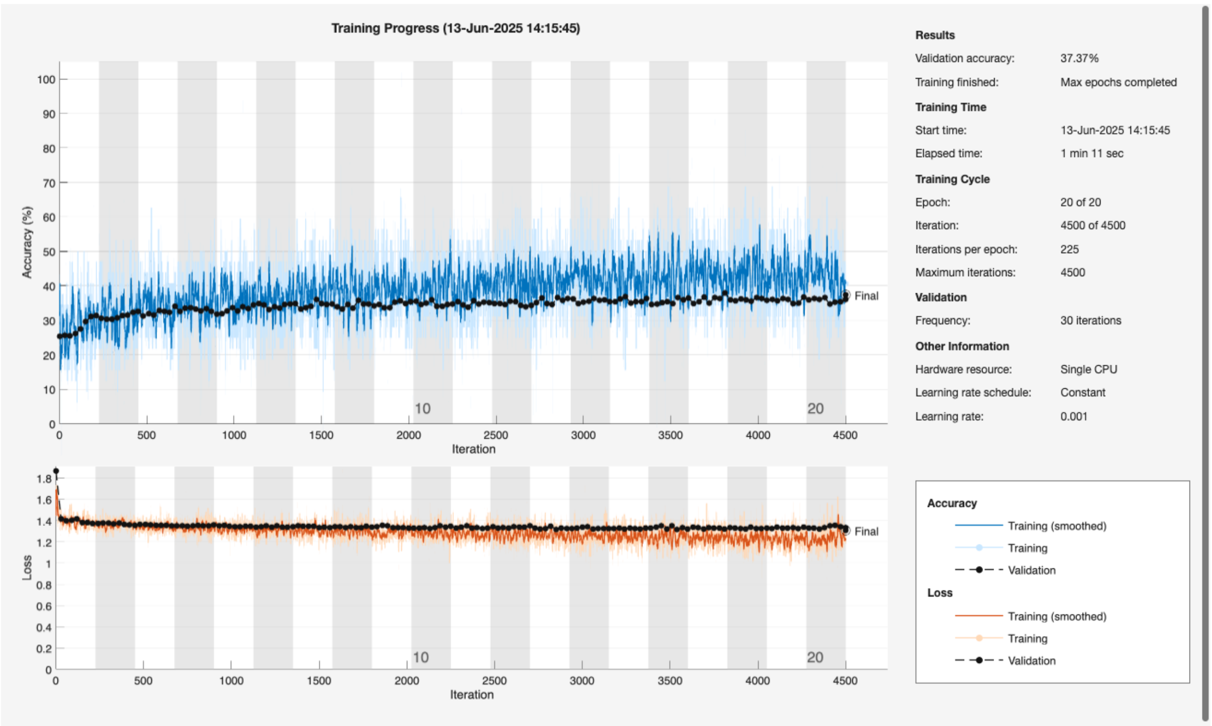

Figure 22: Chart illustrating training progress, showing accuracy and loss over 4500 iterations across 20 epochs using Convolutional Neural Networks for object type classification in the holding phase, with a validation accuracy of 37.37%, trained on a single CPU with a constant learning rate of 0.001

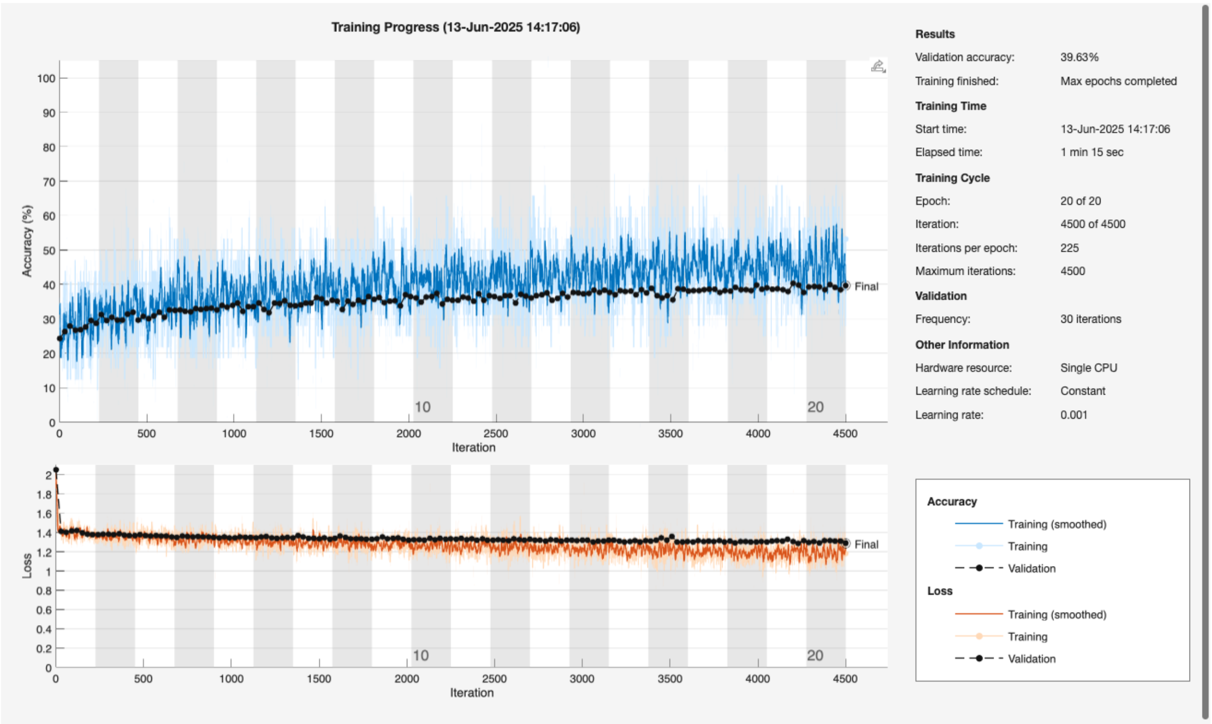

Figure 23: Chart illustrating training progress, showing accuracy and loss over 4500 iterations across 20 epochs using Convolutional Neural Networks for object type classification in the releasing phase, with a validation accuracy of 39.63%, trained on a single CPU with a constant learning rate of 0.001

4. Discussion

The research presented in this thesis has explored the decoding of muscle activations during grasping tasks, leveraging both shallow and deep machine learning techniques alongside multivariate analysis. By applying Representational Dissimilarity Matrices (RDMs), Linear Discriminant Analysis (LDA), and Convolutional Neural Networks (CNNs) to electromyography (EMG) data, this study has shed light on how muscle activity patterns differ across various phases of grasping and how they relate to specific grasp types and object properties.

The Representational Dissimilarity Matrices (RDMs) revealed that muscle activation patterns were most distinct during the holding, planning/execution, and releasing phases of grasping, as opposed to the fixation and observation phases. This suggests that active movement stages involve unique neuromuscular engagements, likely due to the specific demands of initiating, maintaining, and terminating a grasp, which require varied hand configurations and force applications.

Using Linear Discriminant Analysis (LDA), a shallow machine learning model, the study achieved a peak accuracy of 73.86% in classifying grasp types (power, five-finger precision, two-finger precision) during the holding phase. Convolutional Neural Networks (CNNs), a deep learning approach, further improved this to an impressive 88.77% accuracy in the same phase. This high performance indicates that EMG signals robustly capture the distinct muscle activation patterns associated with different grasp types, particularly when the hand is actively engaged in holding an object. For example, a power grasp likely involves broader muscle activation compared to the finer control of a two-finger precision grasp.

In contrast, classifying object types (large cylinder, small cylinder, large sphere, small sphere) proved more challenging, with LDA reaching a maximum accuracy of 32.49% (releasing phase) and CNNs peaking at 43.29% (planning/execution phase). These modest accuracies, though above chance level (25%), suggest that EMG signals are less sensitive to object properties like size and shape. This could be because muscle activation is primarily dictated by the grasp type rather than the object itself; grasping a large or small cylinder elicits similar patterns if the same grip is applied.

The high accuracy of 88.77% achieved in classifying grasp types on EMG data aligns closely with previous studies, such as Miften et al. (2021) and Wang et al. (2022), who also reported robust performance in distinguishing hand grips using EMG signals. This consistency underscores the reliability of EMG data for decoding grasp types across different research efforts. However, unlike these studies, which primarily focused on grasp classification, this research extends the analysis to object type classification, revealing a key difference: the modest accuracies of 43.29% (CNNs) and 32.49% (LDA) suggest that EMG signals are less effective at capturing object-specific properties like size and shape, a challenge less emphasized in prior work and highlighting a potential limitation or unique contribution of this study.

The CNNs consistently outperformed LDA, especially in grasp type classification, highlighting their ability to detect complex, non-linear patterns in EMG data that simpler models like LDA might miss. This advantage is particularly evident during active phases, where the intricate temporal dynamics of muscle activity are more pronounced.

4.1 Limitations

One of the main limitations of this study is the size and diversity of the participant group. The EMG data were collected from 16 healthy, right-handed individuals, which provided a solid starting point for analyzing muscle activity patterns. However, this relatively small sample may not fully represent the broader population. With only 16 participants, the findings might not generalize well to larger or more varied groups. Additionally, all participants were right-handed and free of motor impairments, meaning the results may not be generalized to left-handed individuals or those with conditions like amputations or neuromuscular disorders. This homogeneity limits the practical relevance of the findings for applications like prosthetic control, where users often have diverse physical characteristics. To address this in future work, it would be valuable to expand the sample size and include participants with a wider range of traits, such as left-handed individuals, people with motor disabilities, or even older adults, to see if the observed muscle patterns hold across these groups. A larger and more diverse sample could strengthen the reliability of the models and make the results more applicable to real-world scenarios.

An other limitation of this study could be the use of Representational Dissimilarity Matrices (RDMs) with 1 minus Pearson correlation as the dissimilarity metric may limit the analysis by only capturing pairwise differences, potentially missing complex, higher-order interactions among grasping conditions. The Linear Discriminant Analysis (LDA) relied on features extracted by averaging EMG signals over time, which might oversimplify the temporal dynamics of muscle activity and discard critical pattern variations. Similarly, the Convolutional Neural Network (CNN) approach used a fixed architecture and hyperparameters, which may not be fully optimized for EMG data, possibly constraining its ability to detect subtle temporal features. Additionally, concatenating channel data into a single vector for RDMs and averaging trials for LDA could reduce the richness of the multi-channel EMG signals, limiting the depth of the analysis.

A promising direction for future research would also be to combine EMG with brain signals, such as those from electroencephalography (EEG), which could capture planning or intent before the muscles kick in. Testing this hybrid approach could reveal whether pairing these signals improves decoding across all phases, especially the early ones.

The scope of the experimental design also presents some limitations. The study focused on a controlled set of four objects, large and small cylinders and spheres, and three grasp types: power, five-finger precision, and two-finger precision. While this setup allowed for a clear and manageable analysis, it doesn’t reflect the full range of grasping tasks people encounter in daily life. Real-world grasping involves objects with all sorts of shapes, textures, weights, and sizes, like picking up a soft cloth versus a heavy book, and these differences could influence muscle activity in ways this study didn't explore. For example, grasping a slippery glass might require more precise muscle adjustments than holding a rough wooden block, but such variations weren't tested here. Future studies could broaden the range of objects and grasp types, perhaps including items with different textures or weights, to see how these factors affect muscle patterns. It might also be interesting to design more naturalistic tasks, like grasping objects from a cluttered table or performing a sequence of actions, to better mimic everyday situations and test how adaptable the muscle activation patterns are.

One of the most exciting possibilities for future work is applying these findings to real-world challenges, like prosthetic control or rehabilitation. This study focused on healthy participants, but the long-term goal is to help people with motor impairments, such as amputees or those recovering from strokes. The next step could be to test these models on data from these groups to see how well they perform outside the lab. For prosthetics, this might mean adapting the CNNs to decode muscle signals in real time, allowing a device to respond instantly to a user's grasp intentions. This would require overcoming practical hurdles, like reducing delays in signal processing and ensuring the models work consistently across different users. It could also involve tailoring the system to individual needs, say, training it on a specific amputee's residual muscle signals, to make it more personalized and effective.

5. Conclusion

In summary, this thesis has shown that shallow and deep machine learning can effectively decode muscle activations during grasping, with the clearest patterns emerging during active movement phases. The high accuracy for grasp types and the temporal insights from RDMs, LDA, and CNNs lay a strong foundation for understanding grasping dynamics. However, limitations like the small and uniform sample, reliance on surface EMG, and a narrow set of objects and grasps mean there's still work to do. By expanding the participant pool, integrating additional data sources, testing new models, and linking findings to practical applications and neuroscience, future research can take these results further. This study is a stepping stone toward unraveling the complexities of muscle activity in grasping and creating solutions that enhance lives in the biomedical field.

6. References

Batzianoulis, I., El-Khoury, S., Pirondini, E., Coscia, M., Micera, S., & Billard, A. (2017). EMG-based decoding of grasp gestures in reaching-to-grasping motions. *Robotics and Autonomous Systems*, *91*, 59–70. https://doi.org/10.1016/j.robot.2016.12.014

Castiello, U. (2005). The neuroscience of grasping. *Nature Reviews Neuroscience*, *6*(9), 726–736. https://doi.org/10.1038/nrn1744

Day, S. (2002). *Important Factors in surface EMG measurement By Dr*. https://www.semanticscholar.org/paper/Important-Factors-in-surface-EMG-measurement-By-Dr-Day/d7059f6f1f180a80f7365b3cdb162df3243e2dbe

Fisher, R. A. (1936). The Use of Multiple Measurements in Taxonomic Problems. *Annals of Eugenics*, *7*(2), 179–188. https://doi.org/10.1111/j.1469-1809.1936.tb02137.x

Grosse, P., Cassidy, M. J., & Brown, P. (2002). EEG–EMG, MEG–EMG and EMG–EMG frequency analysis: Physiological principles and clinical applications. *Clinical Neurophysiology*, *113*(10), 1523–1531. https://doi.org/10.1016/S1388-2457(02)00223-7

Jeannerod, M., Decety, J., & Michel, F. (1994). Impairment of grasping movements following a bilateral posterior parietal lesion. *Neuropsychologia*, *32*(4), 369–380. https://doi.org/10.1016/0028-3932(94)90084-1

Kriegeskorte, N., Mur, M., & Bandettini, P. A. (2008). Representational similarity analysis—Connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, *2*. https://doi.org/10.3389/neuro.06.004.2008

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet Classification with Deep Convolutional Neural Networks. *Communications of the ACM*, *60*(6), 84–90. https://doi.org/10.1145/3065386

LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation Applied to Handwritten Zip Code Recognition. *Neural Computation*, *1*(4), 541–551. https://doi.org/10.1162/neco.1989.1.4.541

Miften, F. S., Diykh, M., Abdulla, S., Siuly, S., Green, J. H., & Deo, R. C. (2021). A new framework for classification of multi-category hand grasps using EMG signals. *Artificial Intelligence in Medicine*, *112*, 102005. https://doi.org/10.1016/j.artmed.2020.102005

Müller-Putz, G. R., Ofner, P., Pereira, J., Pinegger, A., Schwarz, A., Zube, M., Eck, U., Hessing, B., Schneiders, M., & Rupp, R. (2019). Applying intuitive EEG-controlled grasp neuroprostheses in individuals with spinal cord injury: Preliminary results from the MoreGrasp clinical feasibility study. *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 5949–5955. https://doi.org/10.1109/EMBC.2019.8856491

Napier, J. R. (1956). The prehensile movements of the human hand. *The Journal of Bone and Joint Surgery. British Volume*, *38-B*(4), 902–913. https://doi.org/10.1302/0301-620X.38B4.902

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

Omedes, J., Schwarz, A., Müller-Putz, G. R., & Montesano, L. (2018). Factors that affect error potentials during a grasping task: Toward a hybrid natural movement decoding BCI. *Journal of Neural Engineering*, *15*(4), 046023. https://doi.org/10.1088/1741-2552/aac1a1

Sburlea, A. I., Wilding, M., & Müller-Putz, G. R. (2021). Disentangling human grasping type from the object’s intrinsic properties using low-frequency EEG signals. *Neuroimage: Reports*, *1*(2), 100012. https://doi.org/10.1016/j.ynirp.2021.100012

Spataro, R., Chella, A., Allison, B., Giardina, M., Sorbello, R., Tramonte, S., Guger, C., & La Bella, V. (2017). Reaching and Grasping a Glass of Water by Locked-In ALS Patients through a BCI-Controlled Humanoid Robot. *Frontiers in Human Neuroscience*, *11*. https://doi.org/10.3389/fnhum.2017.00068

Wang, S., Zheng, J., Zheng, B., & Jiang, X. (2022). Phase-Based Grasp Classification for Prosthetic Hand Control Using sEMG. *Biosensors*, *12*(2), Article 2. https://doi.org/10.3390/bios12020057

Zhang, S., Chen, Y., Zhang, L., Gao, X., & Chen, X. (2023). Study on Robot Grasping System of SSVEP-BCI Based on Augmented Reality Stimulus. *Tsinghua Science and Technology*, *28*(2), 322–329. https://doi.org/10.26599/TST.2021.9010085