

DELFT UNIVERSITY OF TECHNOLOGY

---

# Weight classification during actively assisted elbow flexion and extension

---

BY K.A.F.L. HEYNS

4432754

To obtain the degree of Master of Science  
at the Delft University of Technology,  
to be defended at Friday October 13 2023.

*Thesis committee:*

Prof.dr.ir. J. Harlaar, TU Delft, supervisor, chair

Ir.drs. S.J. Filius, TU Delft, daily supervisor

Dr.ing. L. Marchal Crespo, TU Delft, supervisor

Dr. A. Zgonnikov, external committee member



# Weight classification during actively assisted elbow flexion and extension

**Abstract**—Severe muscle weakness is a symptom appearing in certain neuromuscular diseases (NMDs), such as Duchenne Muscular Dystrophy (DMD), affecting people's daily lives by reducing functionality, decreasing independence, and reducing the ability to perform essential daily activities. This patient group might benefit from using active-assistive devices by having the potential to provide precise support torque counterbalancing the passive forces acting on the arm, the movement intention of the user, and external forces exerted by lifted objects. However, the determination of support to counteract the weight of lifted objects is an ongoing challenge. This research aims to improve the understanding of external forces by using data classification algorithms to distinguish between different lifted weights in a human experiment. Fourteen healthy individuals participated in this experiment, lifting weights ranging from 0 - 1000 grams while an active-assistive device compensated for the passive torques acting on the arm. Data was collected using various sensors: a force sensor, an Inertial Measurement Unit (IMU), a joint encoder, and surface Electromyography (sEMG) electrodes. Subsequently, this data was processed and fed into a K Nearest Neighbour (KNN) classifier and a Support Vector Machine (SVM) classifier to determine the lifted weights during human elbow flexion and extension. The classifier showing the highest performance achieved an accuracy of 39.70% on the test dataset, indicating several misclassifications. However, a recall percentage of 76.95% for the 1000-gram class within the multi-class classification demonstrates the capability to distinguish larger weights. While demonstrating potential in weight discrimination, especially for larger weights, improvements in the compensation strategy, arm support alignment, and experimental design are crucial. Future research on the impact of picking and placing objects, the influence of muscle weakness, and the application of alternative data classification algorithms are essential to further enhance understanding of the interaction with objects and result in more accurate predictions.

**Index Terms**—KNN classifier, SVM classifier, arm support, sensor measurements, weight compensation, human experiment, human-machine interfaces

## 1 INTRODUCTION

Neuromuscular diseases (NMDs) significantly impact individuals by limiting their physical functionality, subsequently affecting their quality of life [1], [2]. Particularly, a group of NMDs, such as Duchenne Muscular Dystrophy (DMD), characterised by severe and progressive muscle weakness, can lead to a reduced ability to perform essential daily activities, such as drinking or picking up objects [3]. These diseases considerably impact an individual's upper limb functionality, restricting autonomy, and ability to fully participate in society [4], [5]. Consequently, this patient group depends on upper limb assistive devices [6].

Various upper limb supports have been developed to assist people experiencing muscle weakness attributed to NMDs in regaining arm functionality [7], [8]. Among all upper limb supports, active-assistive arm devices stand out as particularly promising for individuals with a severely limited upper limb function [8]–[10]. Actively actuated devices have the potential to provide the user with the desired amount of support by counterbalancing arm weight and additional external forces, such as lifted weights [11], [12]. Unlike passively actuated devices, active-assistive devices do not rely on pre-stored mechanical energy, allowing them to exert greater torques on the user or control movements more precisely [6], [7]. Therefore, for patients in advanced stages of muscle weakness, characterised by a severely decreased arm function and increased passive joint stiffness, passive actuation proves insufficient, emphasizing the need for active-assistive devices [13], [14].

Even though active-assistive devices have great potential in restoring the ability to perform daily life activities, they still encounter different obstacles and require improvements before use in real-life settings [6], [12], [15]. One of these challenges is determining the precise amount of support the active-assistive device needs to exert on the user to perform various daily activities, ensuring seamless cooperation between the device and the user [16].

To derive the desired support, all forces acting on the human arm should be identified. These forces include the effect of gravity and passive joint impedance of the human arm, as well as voluntary forces representing the human movement intention and interaction forces with the environment [13]. Currently, several compensation methods have been developed to counteract the passive forces acting on the arm (gravity of the arm and passive joint impedance) [13], [17], and the human movement intention [18]–[21]. Some of these studies show promising outcomes in the compensation of these forces and have a positive influence on regaining functionality and an increased range of motion (ROM) to perform daily tasks [13], [21].

However, understanding the user's intention and differentiating this from interaction with their environment, such as picking up objects, is a complex task, yet crucial for providing the desired support on the user [22]–[24]. Several studies show promising results in compensating external forces using force sensors [25], [26]. However, the force sensor is mounted between the user and the lifted object, measuring

direct gravitational forces by the object. Thus, it hinders the users' ability to perform daily activities. Additionally, several studies use electromyography (EMG) to support the user in load-lifting tasks [27]–[30]. These devices do not provide precise support, therefore only reducing the user's effort to a limited extent, which is insufficient for people suffering from severe muscle weakness. To achieve precise compensation for external forces exerted on the human arm, such as lifted weights, in active-assistive devices, additional research is necessary.

This research aims to improve the understanding of external forces, exerted by lifted objects, on the human arm during elbow flexion and extension. This research focuses on the initial step in solving this problem by looking into four different sensors to capture information about object interaction during elbow flexion and extension tasks. This sensor output is interpreted using simple data classification algorithms to identify and distinguish interaction with objects ranging in weight. Hence, the following research question is answered: *Can we use data classification algorithms to detect and differentiate weights, ranging from 10 to 1000 gram, lifted by an individual during elbow flexion and extension when using an assistive arm support to compensate for arm weight?*

## Outlook

The thesis is structured as follows. Chapter 2 discusses the study design of the human experiment, including a description of the sensing techniques, the active-assistive device and the algorithms for data classification. Chapter 3 contains the results of the experiment and data classification. The discussion and conclusions can be found in chapters 4 and 5.

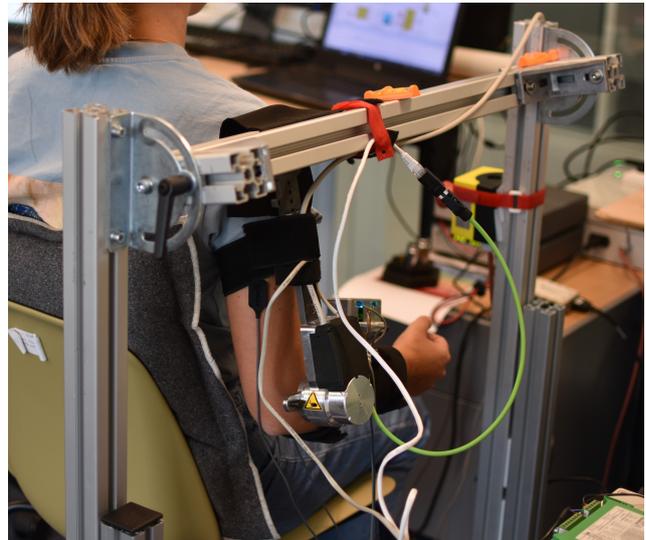
## 2 METHODS

### A. Participants

A human experiment was conducted to gather sensor measurements to gain insight into the impact of lifting weights during elbow flexion and extension. The obtained data is used as input for the data classification algorithms. The experiment involved fourteen participants, eight of whom were women and six men. The Human Research Ethics Committee (HREC) from Delft University of Technology approved the study (ID2972). All the participants provided their explicit written consent to participate in the experiment.

### B. Experimental setup

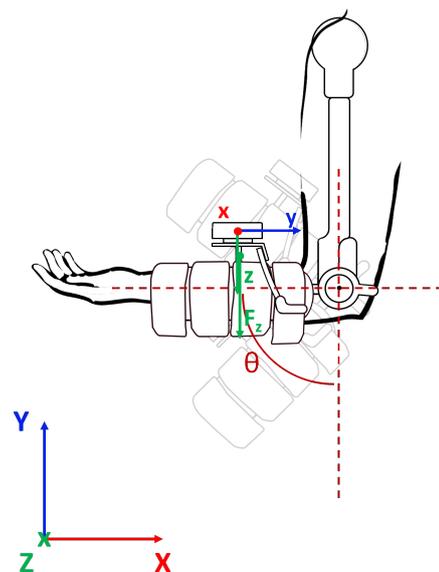
The experimental setup consisted of extrusion profiles to which the arm support was fixed, as illustrated in Figure 1. The extrusion profiles can be adjusted in height to align the joints with the device and enable shoulder abduction/adduction angle modification.



**Fig. 1.** Experimental setup human experiment, showing a participant placed within the active-assistive arm support

Throughout the experiment, the shoulder abduction angle remained fixed at 15 degrees and the shoulder flexion angle at 0 degrees, while the elbow joint angle varied within the participant's ROM.

The one-degree-of-freedom (1DOF) active-assistive device consists of two rigid arms parallel to the human upper- and forearm, connected by a series elastic actuator (SEA) aligned with the elbow joint. The participants were connected to the device by an arm sleeve fixed to the forearm and an elastic band attached to the upper arm. Figure 2 shows all axes and variables of interest within the experimental setup. In Appendix B, a photo of the complete setup and a schematic diagram of the electronics can be found.



**Fig. 2.** Schematic of the experimental setup showing the variables and axes (X,Y,Z - global axes, x,y,z - local IMU axes)

### C. Sensors

The sensors used within the experimental setup include a force sensor, an Inertial Measurement Unit (IMU), a joint encoder, and sEMG electrodes. It is essential to capture information on the kinematics and dynamics of the human arm during these tasks to improve understanding of lifting objects during elbow flexion and extension movements [31]. All sensors, except for the sEMG electrodes, capture direct information on the human arm kinematics and dynamics [32], [33]. While sEMG electrodes do not directly measure the kinematics and dynamics, they can still offer insights into the different forces acting on the arm by capturing muscular activation [34]. In addition, prior research has shown that the included sensors can provide valuable information regarding the interaction with objects [16], [26], [30]. Integrating a diverse group of sensors aims to enhance understanding of their importance in human arm movements, particularly in relation to the interaction with lifted weights.

All signals were acquired with a sampling frequency of 1000 Hz, using two different systems synchronised by a 1 voltage trigger signal. The sEMG signal was collected with a 12-bit resolution ( $\pm 10$  V) using a National Instruments (NI) USB-6008 Data Acquisition (DAQ) device (system noise  $< 5 \mu\text{V}$  (rms)), which was connected to a laptop via a USB connection [35]. The raw signal was collected and stored using MATLAB. The other signals were acquired using a Beckhoff EtherCAT system, interfacing with the different sensors via Ethernet cables. The signals were collected with a 16-bit resolution ( $\pm 10$  V) and stored using TwinCAT, MATLAB, and Simulink [36]. Besides data acquisition of all sensor measurements, TwinCAT was used to synchronise both data acquisition systems by sending a trigger signal and implementing the required control and compensation strategies by sending information to the SEA via Ethernet cables.

#### *Force sensor*

A Miniature S-Beam Jr. Load Cell (Model: LSB200, FUTEK) is used as a force sensor and mounted between the rigid arm of the assistive device and the arm sleeve in which the user was strapped [37]. This sensor was chosen for its capability to measure interaction forces between the active-assistive device and the participants. The sensor was calibrated following a procedure described in Appendix C, providing a measurement range of up to 111.21 N (25 lbs) [37]. The interaction force ( $F_z$ ) is measured over the local z-axis, as shown in Figure 2.

#### *IMU*

Movement data of the arm was obtained using a 9DOF MTi-1 IMU (Xsens), measuring angles ( $\theta_x, \theta_y, \theta_z$ ), angular velocities ( $v_x, v_y, v_z$ ) within a range of

2000 deg/s, and accelerations ( $a_x, a_y, a_z$ ) within a range of 16 g [38]. The IMU was mounted on the rigid bar aligned with the human forearm, as shown in Figure 2, including the local IMU axes. This experiment focuses on three variables: the elbow joint angle ( $\theta_x$ ), angular velocity about the x-axis ( $v_x$ ), and the linear acceleration of the z-axis ( $a_z$ ).

#### *Joint encoder*

The SEA integrated into the elbow joint of the active-assistive device directly measures the elbow joint torque ( $T$ ) by measuring the deflection of the spring within the SEA [39]. In addition, several encoders were integrated into the elbow joint of the device: a rotor encoder, a gearbox encoder, and a spring encoder, measuring the elbow joint angle ( $\theta$ ), and elbow joint velocity ( $v$ ), which is the derivative of the elbow joint angle ( $\theta$ ).

#### *sEMG*

The muscular activity was measured using sEMG of the biceps brachii (short and long head) and triceps brachii (long head) via a Delsys Bagnoli EMG system with DE-2.1 Single Differential Wired sEMG sensors. The system comprises a 16-channel amplifier with an output voltage range of  $\pm 5$  V (system noise (R.T.I)  $< 1.2 \mu\text{V}$  (rms)) [40]–[42]. The amplifier gain was set to 1000 for all participants, maximising the signal amplitude and avoiding amplifier saturation. Electrodes were positioned on the upper arm, according to the SENIAM standards [43].

### D. Study design

#### *Biometric data*

Table 1 provides an overview of the collected biometric data of the participant's group. The mass and center of mass calculation were based on the study of Winter et al. [44].

#### *User calibration*

To align the elbow joint with the SEA, the setup was adjusted in height. Additionally, a ROM measurement was performed to determine the user's maximal elbow joint ROM, and subsequently, the participant was permitted to move within 90% of this ROM during the experiment. Furthermore, the isometric maximum voluntary contraction (MVC) of both the biceps and triceps was measured to normalise the data afterwards. Measurements were performed while bending or stretching the elbow against a tabletop, with the shoulder joint in a neutral position and the elbow at 90 degrees of flexion [45].

**TABLE 1.** Overview of the collected biometric data of the participant group (n = 14)

Variable	Mean (range)
Age [years]	24.64 (22 - 28)
Height [cm]	177.85 (167 - 190)
Weight [kg]	69.14 (61 - 85)
BMI [kg/m <sup>2</sup> ]	21.88 (17.54 - 25.14)
Mass (hand + forearm) [kg]	1.52 (1.34 - 1.87)
Center of mass (hand + forearm - distal) [cm]	14.43 (13.10 - 15.33)
Dominant hand	Right (n = 10), Left (n = 4)
Frequency of sport activity per week (>30 minutes)	3.43 (1 - 6)
Intensity of sport activity (1 - low, 2 - medium, 3 - high)	2.79 (2-3)

### Arm weight compensation

During the experiment, an arm weight compensation strategy was used to counteract the passive torque exerted on the human forearm and hand, including the effect of gravity and internal passive joint impedance [46], [47]. The impedance-based compensation strategy uses position-controlled calibration measurements to identify the passive torques acting on the human forearm and hand. During the position-controlled measurements, the muscular activity of the biceps and triceps was measured using sEMG, ensuring that these muscles were not activated and no other forces than the passive joint torques acting on the forearm and hand were measured. In Appendix D, further explanation is given regarding the arm weight compensation.

### Tasks

The participants were asked to perform repetitive flexion/extension exercises, lifting different weights or no weights, to collect data for classification. The weights were selected based on standard weights encountered in daily tasks, as indicated in the Performance of the Upper Limb (PUL) [48]. The weights included 10, 50, 100, 200, 500, and 1000 grams. Participants performed five different tasks:

- Dynamic task - The participant flexed and extended their elbow joint over the full ROM while lifting weights, completing ten repetitions for each weight. This task was performed maintaining two different elbow joint velocities, representing common ADL elbow flexion/extension velocities, indicated by a metronome [49], [50]:
  - Elbow joint velocity of ca. 45 deg/s
  - Elbow joint velocity of ca. 90 deg/s
- Pick/place task - Participants had to pick up or put down weights from a tripod while flexing and extending their elbow joint over the full ROM, completing ten repetitions for each weight. The tripod's height was adjustable to accommodate two different pickup angles:
  - Elbow joint angle of ca. 45 degrees
  - Elbow joint angle of ca. 90 degrees
- Quasi-static - The participant maintained for three seconds in four different joint angles: maximum elbow extension, ca. 45 degrees and ca. 90 degrees,

and maximum elbow flexion joint angle while lifting weights. The user had to maintain the indicated elbow joint angles themselves after a brief instruction. Hence, there exists a degree of variation in the 45 degrees, 90 degrees, minimum and maximum elbow joint angle.

In Table 2, a complete overview is shown of the different tasks, compensation strategies, and lifted weights used in the experiment, resulting in ten different datasets per subject.

**TABLE 2.** Overview of ten datasets collected during the human experiment by performing the different tasks

Compensation strategy	Tasks	Weights
No compensation	Dynamic task - 45 deg/s	0, 10, 50, 100, 200, 500, and 1000 gram
	Dynamic task - 90 deg/s	0, 10, 50, 100, 200, 500, and 1000 gram
	Pick/place task - 45 degrees	10, 50, 100, 200, 500, and 1000 gram
	Pick/place task - 90 degrees	10, 50, 100, 200, 500, and 1000 gram
	Quasi-static task	0, 10, 50, 100, 200, 500, and 1000 gram
Compensation of forearm and hand weight	Dynamic task - 45 deg/s	0, 10, 50, 100, 200, 500, and 1000 gram
	Dynamic task - 90 deg/s	0, 10, 50, 100, 200, 500, and 1000 gram
	Pick/place task - 45 degrees	10, 50, 100, 200, 500, and 1000 gram
	Pick/place task - 90 degrees	10, 50, 100, 200, 500, and 1000 gram
	Quasi-static task	0, 10, 50, 100, 200, 500, and 1000 gram

### Questionnaire

Participants provided their Rating of Perceived Exertion (RPE) directly after completing each task. The RPE serves as a tool to quantify the effort experienced by the participant during the various tasks in the experiment [51]. The RPE scale ranges from six, which indicates 'No exertion at all', to twenty, which indicates equal to 'Maximal exertion'.

Directly after the experiment, participants filled out two questionnaires: System Usability Scale (SUS) and NASA-TLX. The SUS was used to evaluate the entire arm support's complexity, usability, and comfort [52]. A score above 68 indicates that the arm support is ranked above average compared to other studies using the SUS score. Subsequently, the mean SUS score can be ranked by acceptability and in an adjective scale ('good', 'OK', 'excellent' usability) based on the

study conducted by Bangor et al. [53]. The NASA-TLX compares both mental and physical workload in exercises with and without arm weight compensation during the experiment [54].

## E. Data processing

This section describes how the collected raw sensor signals are processed to be used for analysis and data classification purposes. Figure 3 provides an overview of all the steps taken within this process. Additional information regarding the data processing steps, including visualisations, can be obtained from Appendix E.

### Data filtering

The initial step in processing the data collected during the experiment involved filtering the raw sensor outputs, as illustrated in Figure 3, to improve their Signal-to-Noise Ratio (SNR). The kinematic and dynamic data, including angular velocities, linear acceleration, force, and joint torque, underwent low-pass filtering due to the presence of high-frequency noise in these measurements [49], [55], [56]. The process of filtering the sEMG data is more complex, involving multiple filtering methods, as this is a biological signal influenced by nearby muscles, movement of the skin, and even heart activity [57], [58]. In addition, a trial and error process was used to determine the cut-off frequencies of the filters. Therefore, the following filters were applied to the different raw sensor signals:

- The elbow joint velocity data, as a derivative of the joint angle, provided by the joint encoder, underwent a 2nd order low-pass filter with a cut-off frequency of 2 Hz.
- The elbow joint torque data, based on the deflection of a spring in the SEA, underwent a 1st order low-pass filter with a cut-off frequency of 1.5 Hz.
- The force data provided by the force sensor underwent a 2nd order low-pass filter with a cut-off frequency of 10 Hz.
- The muscular activity of the biceps and triceps, provided by the sEMG sensors, got through several filtering steps: (i) a notch filter of 50 Hz for powerline interference cancellation, (ii) a 3rd order high-pass filter with a cut-off frequency of 10 Hz, (iii) taking the absolute values of the sEMG signal, (iv) a 4 Hz 3rd order low-pass filter, and finally (v) normalisation by the measured MVC values of both biceps and triceps [57], [59].

- The gyroscope velocity and linear acceleration of the accelerometer, all outputted by the IMU, were 2nd order low-pass filtered, with a cut-off frequency of 1 Hz.

### Data preparation

Figure 3 shows four steps in the data processing to prepare the filtered data for analysis, visualisation, and classification of the collected data:

- *Delay removal* - In the experimental setup, measurements were recorded in two different environments. One system captured sEMG measurements, while the other collected real-time data from various sensors. Despite using a synchronization trigger signal, a minor delay was noted between the systems. This delay is attributed to how MATLAB stores data and to the absence of trigger signal information within the sEMG data. The dataset containing more samples was cut from both ends to synchronise the data. The distribution of the synchronisation delay will be shown in the results.
- *End stop removal* - In the active-assistive device, hardware, and software end stops were implemented as a safety measure. To prevent the device from exceeding the software end-stop limits and shutting down, angle guards apply damping with a high value on the user to counteract exceeding the software limits. The damping is applied 2.9 degrees before exceeding the software limit. However, this damping resulted in interaction forces within the joint limits. Therefore, data samples with joint angles being part of the angle guards were removed from the dataset.
- *Data labelling* - Data labelling is essential, assigning each data sample a unique label, representing the lifted weight, for the supervised data classification algorithm. Data labelling was performed manually by selecting the corresponding flexion- and extension cycles for every weight. Therefore, all sensor

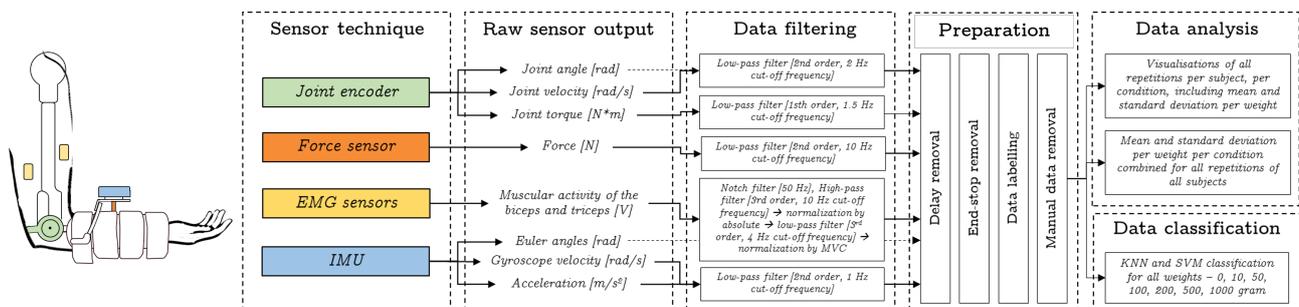


Fig. 3. Workflow including all data processing steps from raw sensor measurements to input for data analysis and data classification

measurements collected simultaneously share the same label.

- *Manual data removal* - Manual removal of data samples was necessary due to artefacts in the data caused by unexpected events during the experiment, such as dropping weights or sensor interaction. Additionally, any time periods during which the weights were altered have been excluded from the data.

**F. Data analysis**

Data was analysed by visualising the force sensor-based joint torque ( $T_z$ ), sEMG biceps activity, sEMG triceps activity, elbow joint torque ( $T$ ), and elbow joint velocity ( $v$ ) measured by the joint encoder, all plotted against the elbow joint angle ( $\theta$ ) measured by the joint encoder. This was done for all data samples per task (dynamic & pick/place) for every participant. Furthermore, the mean and standard deviation for all repetitions per weight were derived per participant. Lastly, the mean and standard deviation for all repetitions per weight were derived for all participants together, combined for both the dynamic tasks and the pick/place tasks. The analysis focused on the data gathered using the arm weight compensation strategy.

**G. Data classification**

Data classification was used to distinguish the weights lifted by the user during elbow flexion and extension due to the categorical nature of both the data and the research objective, aiming to precisely identify the different lifted weights. First, the two used supervised classification algorithms will be introduced in this section. Afterwards, all steps within the classification process will be discussed.

*Classifier*

K Nearest Neighbour (KNN) and Support Vector Machine (SVM) classification algorithms were used to classify the sensor output. The two distinct classifiers were optimised, trained, and tested to evaluate and compare the impact of the classification algorithms on the dataset. KNN classification is a classification

algorithm based on the idea that surrounding data points contain useful information on the corresponding label when assigning a new data sample to a class [60]. In addition, KNN classifiers have the ability to classify large non-parametric datasets accurately [61]. SVM is a classification algorithm that is based on the idea of finding a hyperplane that best divides a dataset into classes. The aim is to maximise the margin between the classes' closest data points, called support vectors. SVM is known to be a robust and powerful classification algorithm, even for high dimensional data and non-linear datasets [62], [63]. Besides that, both classification algorithms show to be robust biosignal classifiers [64]–[66]. Data classification was only performed on the dataset collected during the dynamic tasks, as this is a first step to gaining insights into the effect of lifting objects.

*Classification strategy*

The classification strategy for both classifiers was to look into the effects of feature selection and optimisation of the hyperparameters. Before training the different classifiers, the dataset was downsampled and split into train and test data. In Figure 4, an overview shows all steps taken in the data classification. The following classifiers, for both the KNN and SVM classification algorithms, have been trained, tested and evaluated:

- Classifier with default settings, including all features
- Classifier with default settings, including the features selected via feature selection
- Classifier with optimised hyperparameters, including the features selected via feature selection

In addition, one-to-one classifications will be performed between the class of 0 gram and all other classes (0 gr vs. 10 gr, 0 gr vs. 50 gr, etc.) to check the ability of both classifiers to differentiate between two weights. In both classification algorithms, default hyperparameter settings were used, and all features were included. No feature selection or hyperparameter selection has been performed on these classifiers due

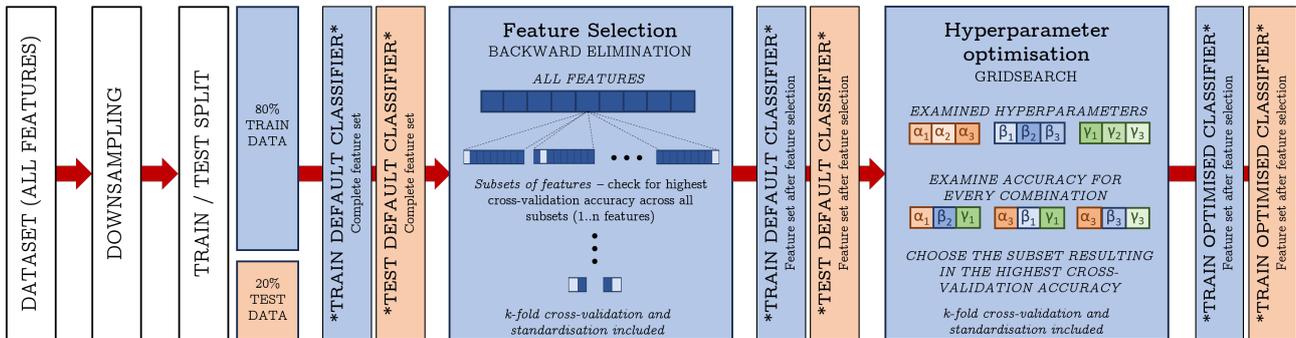


Fig. 4. Workflow data classification for both SVM and KNN classifier (blue = training dataset, orange = test dataset)

to the time-consuming process of performing this for every one-to-one classification algorithm.

#### *Data visualisation*

The first step in data classification is data visualisation, a necessary step to gain more insight into the dataset [67]. To make it easier to understand the distribution of the data, identify patterns and outliers, and make classification decisions accordingly, the following visualisation methods were used:

- Class distribution - Essential visualisation step to check for no bias towards one or several classes [68].
- Descriptive statistics - Using Python's describe() function to receive every feature's mean, standard deviation and percentile values. This will give more insight into the data distribution per feature, especially outliers.
- Box plots - Box plots show the distribution and variability of the data. For every feature, a box plot was made to show the variability of the data over the different classes.
- Correlation - The Pearson's correlation coefficient was used to measure the linear correlation between all features (one-to-one). In addition, the correlation between every feature and the classes is derived. Both were checked with a graphical heatmap containing these Pearson correlation coefficients. The strength of the correlation coefficient will be rated based on values in the book of Swinscow et al. [69].

#### *Down sampling data set*

The sensor measurements were performed with a frequency of 1000 Hz. To reduce the computational load, the dataset is down-sampled to a frequency of 100 Hz. Downsampling is performed by taking the mean of every ten samples per label.

#### *Split train/test set*

To provide a reasonable estimate of the model performance and prevent overfitting, the distribution between the training and test set was chosen as 80% versus 20% [67].

#### *k-fold cross-validation*

k-fold cross-validation was used in both training the classifiers, performing feature selection, and optimising the hyperparameters. In all instances, a value for  $k$  of 5 was used to split the training set into 5 folds, using  $k-1$  folds as the train set and the remaining fold to evaluate the model [70]. Each fold contains a proportional representation of each class. In general, cross-validation is used to gain insight into the model performance across the training dataset, ensuring the classifier models are robust and not overfitted to specific data subsets [67], [70].

#### *Standardisation*

Standardisation is necessary for both KNN and SVM classification to achieve more accurate results [71]. The standardisation method used was the standardisation by the z-score, resulting in features with a mean of zero and a standard deviation of one, fitted on the training set [72], [73]. The dataset was standardised after splitting the dataset into the train and test sets to prevent data leakage to the test set [67], [72]. Within the k-fold cross-validation, the train, and evaluation set were standardised in every computation.

#### *Feature selection*

Feature selection is essential to reduce overfitting, make models easier to interpret and make models less computationally intensive [74]. In this research, a wrapper method was chosen as feature selector. Wrapper methods use the specified classification algorithms to select the optimal subset of features, aiming to optimise the model performance [75]. In this study, a wrapper method known as backward elimination was employed. It was initiated by using the complete feature set and iteratively selecting subsets of features that cause the highest cross-validated accuracy until a single feature was left [76]. The combination of features, executed with the highest cross-validation accuracy, will be used as input for the classifiers [77]. This feature selection process included standardisation and k-fold cross-validation to guarantee the reliability of the feature selection [67]. The feature selection is performed on the training dataset to prevent data leakage to the test set [67].

#### *Hyperparameter optimisation*

To identify the optimal hyperparameters within a specific selection, GridsearchCV was used [78]. Grid search explored all combinations of hyperparameters, to determine the hyperparameters resulting in the highest cross-validated accuracy on this dataset [79]. To prevent data leakage and overfitting on the test set, hyperparameter optimisation was performed on the training set, including features selected via backward elimination [72], [79]. During the grid search, both standardisation and k-fold cross-validation were included in the process to ensure the robustness of the hyperparameter optimisation [67].

#### **KNN classifier**

For KNN, the following hyperparameter variables were evaluated in the grid search:

- **n\_neighbors** - {1, 5, 10} and {20 – 1000} in steps of 20
- **weights** - 'uniform', 'distance'
- **metric** - 'Euclidean', 'Manhattan'

## SVM classifier

For SVM, the following hyperparameter options were evaluated in the grid search:

- **C** - 0.01, 0.1, 1, 10, 100
- **gamma** - 0.01,  $(1/(n\_features * variance))$ , 0.1, 1, 10
- **kernel** - 'linear', 'rbf'

Evaluation of the linear kernel combining the different gamma hyperparameter values will not be performed, as it has no impact on the accuracy of the linear kernel [80].

### Evaluation

To evaluate the data classification, the following metrics were used:

- Confusion matrix for the multi-class classifications
- Accuracy and macro-averaged precision, recall, and f1 score for the multi-class classifications
- Accuracy, precision, recall, and f1 score for the one-to-one classifications

## H. Outcome measurements

The primary outcomes of this research will be the accuracy and macro-averaged f1-score, precision, and recall for all classification algorithms, including the features selected by feature selection. Moreover, the confusion matrix will be shown for the KNN and SVM classifiers with optimised hyperparameters and feature set after performing feature selection. Furthermore, the one-to-one classifications' accuracy, precision, recall, and f1-score will be shown.

As a secondary outcome, a graph showing the mean and standard deviation of the dynamic task over different weights will be discussed.

Lastly, the results of the questionnaires and maximum synchronisation delay will be discussed.

## 3 RESULTS

### A. Participants

After the performance of the human experiment, three participants were excluded. One participant

was excluded because of significant misalignment due to the absence of the upper arm elastic band connecting the human upper arm to the arm support. Two participants have been excluded based on the calculated arm weight compensation. However, the force sensor values during extension exceeded those during flexion, contrary to the principle of hysteresis in the spring of the SEA. Appendix F provides more insight into the exclusion of these participants.

### B. Data classification outcomes

In this section, outcomes concerning the data classification will be shown. However, only a selection of the results from the data classification will be presented in this section. A complete overview of the results, including data visualisations, can be found in Appendix G.

#### Data visualisation

Several data visualisation methods have been used to gain insight into the dataset. Within the dataset, the mean number of samples per class is 546221.43 (standard deviation of 15640.48), with a minimum of 521380 samples in the 0-gram class (13.6% of the complete dataset) and a maximum of 572279 samples in the 1000-gram class (14.9% of the complete dataset).

Looking into the correlation heatmap between the features, five feature combinations show a very strong linear correlation, having a Pearson value above 0.8. Only, the sEMG biceps activity and sEMG triceps activity are very strongly correlated to any other features. Focussing on the linear correlation between features and the classes, the strongest relation compared to the others can be seen in the sEMG biceps activity, having a Pearson correlation coefficient of 0.46. This linear correlation between the classes, representing the weight of the lifted object, and the sEMG activity is confirmed by the data distribution in the box plots, shown in Appendix G.

#### Feature selection

As a feature selection method, backward elimination was performed. This shows that a combination of the

**TABLE 3.** Evaluation of KNN classifiers (accuracy = fraction of total predictions that are correct, precision = fraction of a predicted label that was identified correctly, recall = sensitivity, fraction of the true label that was identified correctly, F1-score = a combination metric of precision and recall, emphasizing the balance between the two)

Classifier	Accuracy	Precision	Recall	F1-score	Time fit	Time predict
Default KNN Classifier (all features)	34.77%	34.63%	34.85%	34.51%	1.37 s	14.73 s
Default KNN Classifier (after feature selection)	36.57%	36.48%	36.80%	36.33%	0.50 s	4.96 s
KNN Classifier with the optimised hyperparameters (after feature selection)	39.70%	39.09%	39.06%	39.42%	0.37 s	77.97 s

**TABLE 4.** Evaluation of SVM classifiers (accuracy = fraction of total predictions that are correct, precision = fraction of a predicted label that was identified correctly, recall = sensitivity, fraction of the true label that was identified correctly, F1-score = a combination metric of precision and recall, emphasizing the balance between the two)

Classifier	Accuracy	Precision	Recall	F1-score	Time fit	Time predict
Default SVM Classifier (all features)	34.27%	32.83%	32.98%	33.94%	4927.41 s	2042.22 s
Default SVM Classifier (after feature selection)	24.72%	20.46%	19.96%	24.22%	5697.67 s	2138.82 s
SVM Classifier with the optimised hyperparameters (after feature selection)	24.63%	19.93%	20.97%	24.14%	10794.78 s	2481.30 s

elbow joint angle ( $\theta$ ), elbow joint velocity ( $v$ ), elbow joint torque ( $T$ ), load cell force ( $F_z$ ), sEMG biceps activity and sEMG triceps activity results in the highest cross-validation accuracy for the KNN classifier. Within SVM classification, a combination of sEMG biceps activity, IMU acceleration ( $a_z$ ), and elbow joint angle ( $\theta$ ) results in the highest cross-validation accuracy. In Appendix G.2, a complete overview of the backward elimination results can be found.

*Evaluation KNN classifier*

First, a default KNN Classifier (hyperparameters: **n\_neighbors** - 5, **weights** - 'uniform', **metric** - 'minkowski') was used to classify the dataset into the different classes, presenting the lifted weight, to get more insight into the response of the classifier on the dataset. Table 3 shows all values for the performance metrics evaluating the KNN classifiers.

After feature selection, KNN classification has been performed on the feature set containing the six selected features. This results in an increase of ca. 2% for accuracy, precision, recall, and F1-score compared to the default classification performed on the entire feature set. Furthermore, the computation time for fitting the classifier decreased by 0.87 seconds, and the time for predicting all new data samples decreased by 9.77 seconds.

Lastly, a KNN classification was conducted after performing a Grid search to optimise the hyperparameters. Grid search found that the KNN classifier with the following hyperparameters results in the highest cross-validation accuracy:

- **n\_neighbors** - 280

- **weights** - 'distance'
- **metric** - 'Euclidean'

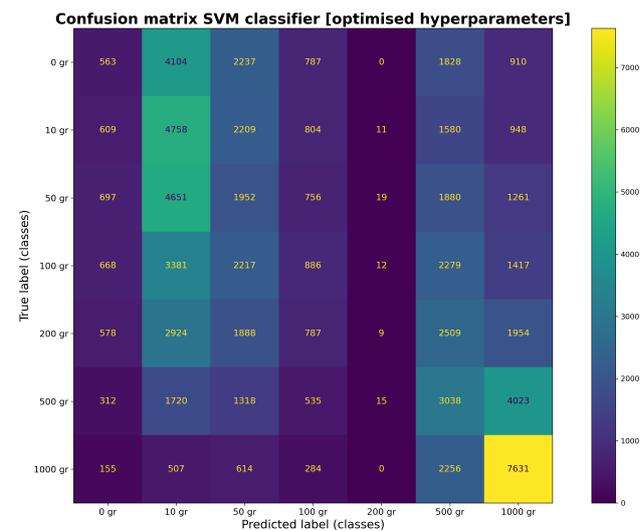
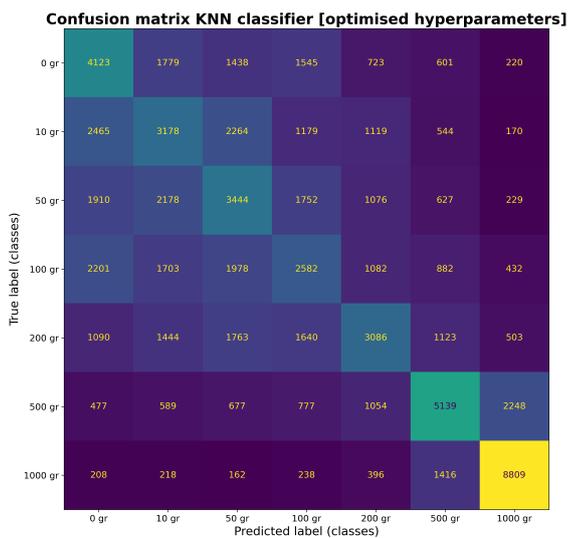
Furthermore, the values for the performance metrics to evaluate the KNN classifier with the optimised hyperparameters are listed in Table 3. Optimisation of the hyperparameters results in an increase in the performance metrics, ranging from 2.26% to 3.13% compared to the default classification on the same feature set. The confusion matrix of this classification in Figure 5a shows the amount of data samples predicted as a specific label compared to their true label. Among the classes, the 1000-gram class shows the highest recall percentage. This is the percentage of data samples from this class that is correctly identified as this class. For the 1000-gram class is this a percentage of 76.95%. On the other hand, the class with the lowest recall percentage is the 100-gram class, having a percentage of 23.78%.

*Evaluation SVM classifier*

First, a default SVM Classifier (hyperparameters: **C** - 1.0, **kernel** - 'rbf', **gamma** - 1 / (n\_features\*variance)) classified the test dataset into the different classes. In Table 4, the values of all performance metrics are shown for this classifier.

Feature selection shows a decrease of 9.55% in accuracy, 12.37% in precision, 13.03% in recall, and 9.72% in the F1-score, compared to the classifier with default hyperparameters on the entire feature set.

Finally, SVM classification was done after a grid search to find the best hyperparameters. Grid search found that the following combination of hyperparameters results in the highest accuracy:



(a) Confusion matrix of KNN classifier after feature selection and optimisation of the hyperparameters

(b) Confusion matrix of SVM classifier after feature selection and optimisation of the hyperparameters

**Fig. 5.** Confusion matrices KNN and SVM classifier with optimised hyperparameters (The diagonal values represent the number of correct classifications for each class. All other values indicate the number of misclassifications.)

- **C** - 100
- **gamma** -  $1/(n\_features * variance)$
- **kernel** - 'rbf'

All values for the performance metrics evaluating the SVM classifier with optimised hyperparameters can be found in Table 4. The values for the performance metrics are almost equal to the SVM classifier using the default hyperparameters tested on the same feature set. The values show a maximum decrease in precision of 0.53% and a maximum increase in recall of 1.01% compared to the default SVM classifier. In Figure 5b, the confusion matrix of this classifier is shown. The ability of this SVM classifier to predict the label of a specific class correctly, being the recall, differs a lot between the classes. The lowest recall percentages were shown in the 0-gram, 100-gram and 200-gram classes, having percentages of 0.05%, 0.08% and 0.0008%, meaning there were almost no correct predictions for these classes. On the other hand, the 10-gram and 1000-gram classes show recall percentages of 43.58% and 66.66%. Therefore, the difference in the ability to perform correct predictions differs a lot between the classes, having a bias towards the 10-gram, and 1000-gram classes, and almost no correct predictions in the 0-gram, 100-gram, and 200-gram classes, as shown in the confusion matrix.

### C. One-to-one classification

#### KNN classifier

In Table 5, the performance metrics for evaluating the one-to-one KNN classification, comparing all classes with the class lifting no weight (0 gram), is shown. The table shows that all performance metrics increase when weight increases compared to 0 gram. The only exception to this is the class of 100 gram. Focussing on the accuracy values, they range from 59.48% for 10 gram to 91.49% for 1000 gram. Among these accuracy values, the class of 100 gram, decreased in the accuracy value by 2.01% compared to the accuracy of the one-to-one classification between 0 and 50 gram.

**TABLE 5.** Evaluation of KNN one-to-one classifiers (accuracy = fraction of total predictions that are correct, precision = fraction of a predicted label that was identified correctly, recall = sensitivity, fraction of the true label that was identified correctly, F1-score = a combination metric of precision and recall, emphasizing the balance between the two) [gr = gram]

One-to-one classifier	Accuracy	Precision	Recall	F1-score
0 gr vs. 10 gr	59.48%	61.07%	57.31%	59.13%
0 gr vs. 50 gr	63.12%	65.21%	61.82%	63.47%
0 gr vs. 100 gr	61.11%	62.69%	58.66%	60.61%
0 gr vs. 200 gr	69.79%	71.42%	67.01%	69.15%
0 gr vs. 500 gr	81.60%	81.25%	83.31%	82.27%
0 gr vs. 1000 gr	91.49%	91.69%	93.03%	92.36%

#### SVM classifier

Table 6 shows the performance metrics for evaluating the one-to-one SVM classification, comparing all classes with the class lifting no weight (0 gram). It can

be seen that all performance metrics increase when the lifted weight increases compared to 0 gram. The accuracy values range from 57.27% for 10 gram to 90.88% for 1000 gram.

**TABLE 6.** Evaluation of SVM one-to-one classifiers (accuracy = fraction of total predictions that are correct, precision = fraction of a predicted label that was identified correctly, recall = sensitivity, fraction of the true label that was identified correctly, F1-score = a combination metric of precision and recall, emphasizing the balance between the two) [gr = gram]

One-to-one classifier	Accuracy	Precision	Recall	F1-score
0 gr vs. 10 gr	57.27%	59.07%	53.56%	56.18%
0 gr vs. 50 gr	60.65%	62.48%	60.22%	61.33%
0 gr vs. 100 gr	62.82%	64.61%	59.94%	62.19%
0 gr vs. 200 gr	68.23%	70.50%	63.84%	67.00%
0 gr vs. 500 gr	80.58%	81.62%	80.15%	80.88%
0 gr vs. 1000 gr	90.88%	91.95%	90.50%	91.22%

### D. Visualisation dynamic task

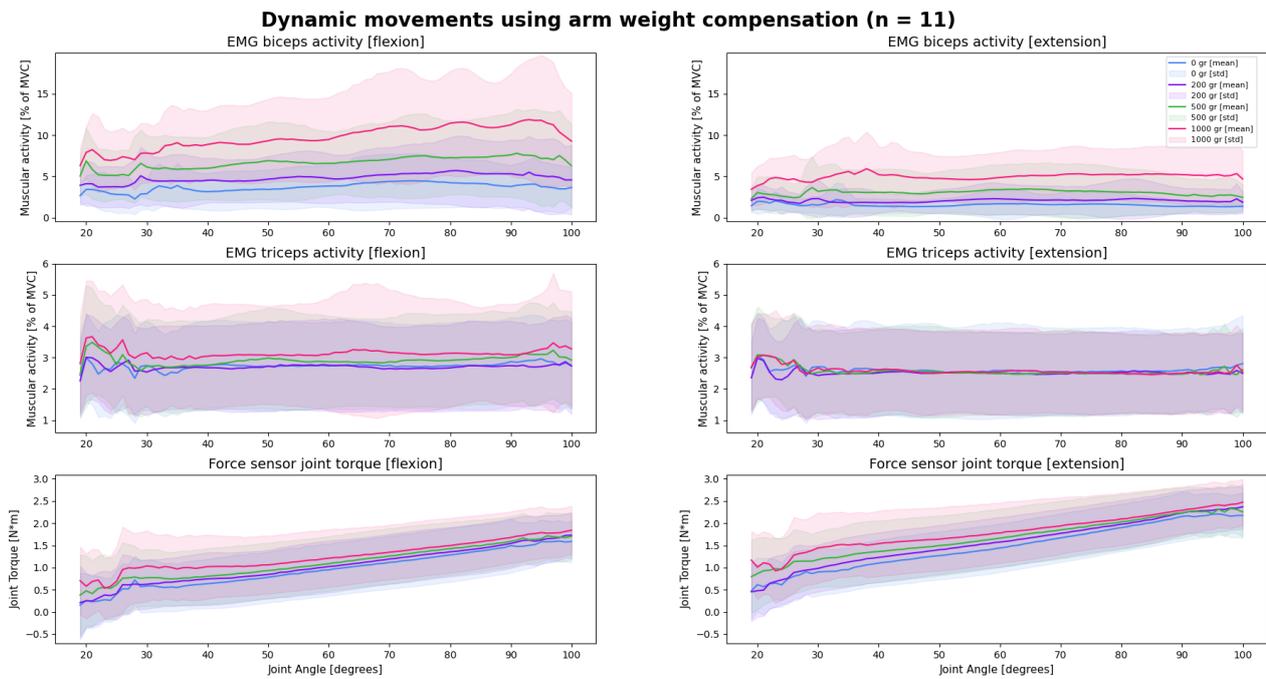
Figure 6 shows the mean and standard deviation for the sEMG biceps activity (% of MVC), sEMG triceps activity (% of MVC), and the force sensor joint torque ( $Nm$ ), against the joint angle in degrees, distinguished for the flexion and extension, during a dynamic task for both elbow joint velocities (ca. 45 deg/s, ca. 90 deg/s). Each graph showcased a specific weight selection, comprising zero, 200, 500, and 1000 grams, to facilitate a clear visualisation. Except for the triceps activity during extension, the graphs for biceps activity, triceps activity during flexion, and force sensor joint torque visually show a clear difference in the mean value while lifting the different weights. For sEMG biceps activity and force sensor joint torque, the following values apply in a flexion and extension cycle during the 90-degree elbow joint angle, supporting this clear difference. The sEMG biceps activity during the flexion cycle equals the mean values of 3.41 % (no weight), 4.59 % (200 gr), 6.87 % (500 gr), and 10.60 % (1000 gr). Comparing this to the extension cycle, having mean values of 1.41 % (no weight), 1.97 % (200 gr), 2.63 % (500 gr), and 5.00 % (1000 gr). The mean values for the force sensor joint torque at a joint flexion angle of 90 degrees are 1.51  $Nm$  (no weight), 1.53  $Nm$  (200 gr), 1.62  $Nm$  (500 gr), and 1.65  $Nm$  (1000 gr). Comparing this to the extension cycle, the mean values at a joint extension angle of 90 degrees are 2.08  $Nm$  (no weight), 2.08  $Nm$  (200 gr), 2.16  $Nm$  (500 gr), and 2.18  $Nm$  (1000 gr).

Only a selection of the gathered data from the human experiment was shown in this section. The remaining data visualisations, including repetitions per person, can be found in Appendix H.

### E. Questionnaires

#### SUS

For the SUS questionnaire, a mean SUS score of 71.73 was found over the participant group, with a standard



**Fig. 6.** Features plotted against the joint angle during dynamic task (mean - dark line showing the mean of all repetitions for all participants in dynamic tasks per weight, std = transparent area surrounding the mean showing the standard deviation of all repetitions for all participants in dynamic tasks per weight ( $1 \cdot \sigma$ ) [shared ROM over all participants ( $n = 11$ )]

deviation of 13.67. Therefore, the mean SUS score is above 68 and rated above average compared to other devices. This is equal to being an ‘acceptable’ score and can be classified as ‘OK’ usability following the study of Bangor et al. [53]. The individual SUS scores range from 45 to 87.5.

**NASA-TLX**

The NASA-TLX questionnaire compares the user situation with and without arm weight compensation. The weighted rating of performing the tasks with no arm weight compensation is 44.12, compared to a rating of 37.36 for task performance with arm weight compensation. Therefore, the mental and physical workload for exercises with arm weight compensation was rated lower than a situation with no arm weight compensation.

**RPE**

In Table 7, the mean and standard deviation of the RPE for every task are shown. The task with the lowest RPE is the quasi-static task, with compensation for the weight of the forearm and hand. On the other hand, the dynamic task at a velocity of 90 deg/s, having no compensation, has the highest RPE of 11.64. When comparing identical tasks with and without compensation, it can be observed that all tasks executed with the arm weight compensated, resulted in a lower RPE than those performed without compensation.

**TABLE 7.** Overview of the RPE per task over all subjects

Compensation strategy	Tasks	Mean RPE (standard deviation)
Compensation of forearm and hand weight	Dynamic task - 45 deg/s	9.91 (1.64)
	Dynamic task - 90 deg/s	9.82 (2.04)
	Pick/place task - 45 degrees	9.64 (1.63)
	Pick/place task - 90 degrees	10.00 (1.26)
No compensation	Quasi-static task	8.73 (1.10)
	Dynamic task - 45 deg/s	10.91 (2.12)
	Dynamic task - 90 deg/s	11.64 (2.29)
	Pick/place task - 45 degrees	10.18 (2.27)
	Pick/place task - 90 degrees	10.18 (1.94)
	Quasi-static task	9.36 (1.29)

**F. Maximum synchronization delay**

The number of removed data samples ranges from 0 to 181 data samples due to the delay. Resulting in a minimum delay of 0 seconds to a maximum delay of 0.181 seconds. The mean value for the removed data samples is 88.09, with a standard deviation of 46.08, equal to a mean delay of 0.088 s.

**4 DISCUSSION**

This research conducted an exploratory study to gain insights into the capability to detect and distinguish lifted objects while flexing and extending the elbow joint by using various sensor measurements coupled with straightforward data classification algorithms. In this discussion, the main findings of this research will be discussed along with the limitations, followed by recommendations for further research.

**4.1 Data classification accuracy**

A KNN classifier demonstrated the highest accuracy value of 39.70% after feature selection and hyperparameter optimisation. SVM classification only reaches

a maximum accuracy value of 34.27%. The accuracy values are higher than random guessing, equal to a value of 14.90% based on the proportion of samples in the most frequent class (1000-gram). However, the performance indicates a considerable number of misclassifications [81].

Possible reasons for the difference in performance between the KNN and SVM classifiers could be that this dataset benefits from a local rather than a global classification method [82]. KNN classification is a local classification method, that looks into data samples near newly presented data to determine the class. On the other hand, global methods, such as SVM classification, attempt to make a single classification rule that is suitable for the whole dataset [82]. Furthermore, SVM classification is more sensitive to outliers, making it challenging to distinguish weights based on biological signals as sEMG due to their low SNR [83].

The confusion matrix, belonging to the best KNN classifier, shows that misclassifications even occur between the 0 and 1000-gram classes. Considering the future application, this indicates an overestimation, resulting in an extensive support torque. On the other hand, when underestimating, assuming the user has severe muscle weakness, the user will no longer be able to lift an object or even drop it. For this vulnerable patient group, it is essential that no major misclassifications, and thus over- or underestimations, are made.

The KNN classifier exhibits the highest recall for the 1000-gram class. This recall value indicates greater accuracy in identifying this weight [84]. Additionally, one-to-one classifications were performed to see what amount of weight the classifiers would be able to identify compared to lifting no weight. One-to-one classifications reaffirmed the highest accuracy values for the 1000-gram class, with 91.49% for KNN classification and 90.88% for SVM classification. The lower the weight, the closer the accuracy values are to random guessing.

In conclusion, the results indicate that data classification can, to an extent, differentiate weights of lifted objects, especially 500- and 1000-gram. The ability to distinguish weights increases when the difference with no weight, or the closest weight, is larger.

#### 4.2 Interpretation of the dynamic task analysis for data classification insights

The dataset, consisting of sensor measurements collected during the dynamic task, was used as input for the data classification algorithms. During this task, participants had to flex and extend their elbow joint while lifting different weights at a constant joint velocity, guided by a metronome. Having almost no changes in velocity and acceleration due to the constant

velocity, results in the outcome that these features do not contain much useful information on the amount of weight lifted within this experiment. However, the elbow joint velocity does indicate flexion or extension of the elbow joint.

The combined mean and standard deviation for all participants shows that the sEMG biceps activity is the clearest indicator of the lifted weights while showing the most noticeable differences between the classes. Slight differences can be seen in the joint torque based on the force sensor, and the sEMG triceps activity during flexion. Furthermore, the Pearson correlation coefficient between the features and classes, shows the highest value for the sEMG biceps activity, followed by the force sensor force ( $F_z$ ) and the sEMG triceps activity.

Having healthy people as participants explains these observations. While healthy people can fully compensate for lifting weights using muscle strength, the sEMG biceps activity is expected to be the clearest indicator. The correlation of the sEMG triceps activity with the classes is less strong than the sEMG biceps activity because its main function is to extend the arm. Therefore, the lifted weight did not lead to major differences in the force sensor measurements or joint torque measurements of the SEA. By way of contrast, people suffering from severe muscle weakness do not have the muscle strength to lift those objects. Therefore, including this patient group might result in different findings.

Even though there is a constant velocity and healthy participants were included, features providing information on the force, torque, velocity, or acceleration were part of the selected features after performing backward elimination on the feature set. However, looking at the cross-validated accuracies in Appendix G.2, it is shown that including these features does not have a lot of influence on the accuracy values. For the KNN classifier, an increase of 3.04% in accuracy value was shown when including the selected features compared to only sEMG biceps activity. In the SVM classifier, the difference is only 0.02%. In conclusion, due to the constant velocity and including healthy participants, sEMG biceps activity is, in theory, the only feature providing information on the lifted weight. However, sEMG is known for its low SNR and sensitivity to the placements of the electrodes [85]. This low SNR is reflected in this study, as the sEMG signal has a large standard deviation. Eventually, this leads to low accuracy values, necessitating the incorporation of other sensors or the utilization of alternative object detection techniques to improve the effectiveness of weight determination through data classification.

### 4.3 Arm weight compensation

Throughout the experiment, the passive forces of the arm were compensated. Providing precise arm weight compensation is essential to compensate for the passive forces and purely measure the movement intention forces and interaction forces with objects within the experiment. However, only 80% of the passive forces were compensated due to deviating values from the gravity model. The possible reasons for the observed deviations from the gravity model could be due to inaccurately collected biometric data, or interaction forces interfering with the passive force measurements. These disturbances could be caused by the participant not being at rest during the position-controlled calibration measurement, or joint misalignment that caused interaction forces in the joint limits. An inaccurate support torque calculation could lead to overestimating the support torque. Consequently, influencing the results due to interaction forces with the active-assistive device. Therefore, it was chosen to only compensate for 80% of the measured passive forces with the chance of having non-compensated passive forces influencing the results. In future experiments, a more consistent way of determining this compensation should be used to have the full support of the passive torques acting on the arm.

### 4.4 Experimental setup

Several refinements could be made to the experimental setup. First, the sensors influenced the outcomes of the human experiment. The IMU was mounted such that it restricted the ROM of the forearm. This led to a maximum flexion angle of around 100 degrees in all experiments. In addition, the IMU sometimes influenced the sEMG measurements in a maximum flexion angle by having contact with the sEMG electrodes, resulting in artefacts in the data. Additionally, in the experiment, wired sEMG electrodes were used, causing power line interference. Using a notch filter of 50 Hz reduced the effects of power line interference in the data. However, wired electrodes might be affected by other environmental influences, such as contact with the setup, resulting in noise or other artefacts [86]. Further, the force sensor was mounted between the arm support and arm sleeve with just one screw on both sides. Although this assembly was robust initially, this diminished over time.

Secondly, by using extrusion profiles in the experiment setup, it was possible to adjust the height of the arm support and the length of the rigid bar parallel to the upper arm. However, the length of the rigid bar parallel to the forearm could not be adjusted, and the arm sleeve used was one size. The setup and assistive device could not be perfectly adapted to the user, leading to possible misalignment. In addition, participants indicated that they sometimes felt discomfort during maximum extension and maximum flexion.

However, this misalignment only influences the data in maximum flexion and extension.

### 4.5 Recommendations

Based on the findings and limitations of this study, several recommendations can be made.

First, a follow-up study should examine the data of the pick/place task. While performing this task, it was decided to standardize the angle at which an object is picked up, but there are no other restrictions. Because no constant velocity is maintained, velocities and accelerations can vary, which may provide useful information on the picked-up weight.

The logical follow-up to the limitation of including healthy participants is to allow people with muscle weakness to participate. Another method would be to find a way in which muscle weakness can be mimicked in the experiment, e.g. by using larger weights or creating muscle fatigue. Muscle weakness may affect the sensor measurements and, thus, the accuracy of the data classification. Research indicates that sEMG characteristics vary in people suffering from muscle weakness, compared to healthy patients [87]–[89]. Specifically, in patients suffering from DMD, the overall level of muscular activity decreases. However, more high-frequency activity is observed [87]. In addition, the progress of the disease affects sEMG measurements, which means that the accuracy of data classification may vary over time based on the progression [90]. Nonetheless, these findings emphasize the urge to acquire sEMG measurements from people suffering from several NMDs at different progress stages to validate the data classification method. In addition, people suffering from severe muscle weakness, cannot always lift objects occurring in daily life tasks [18], [91]. Contrary to healthy participants, this group does not have the muscle strength to lift all objects. Therefore, an additional weight might be indicated by the force or torque sensors when muscle strength cannot compensate for these weights. Further research is necessary to understand how muscle weakness affects sensor measurements to be able to determine the precise support in active-assistive devices.

Another suggestion is to investigate other data classification algorithms. In this exploratory study, simple classifiers were used to check whether these algorithms would be sufficient to recognise a weight. For now, the accuracy values do not represent reliable classification, and modifications in the data classification algorithms could increase the accuracy. Currently, a single time point is used as input to the data classifier. Using this method means that no additional information can be extracted from adjacent time points. Combining time points as input for the classifier could be beneficial to recognise patterns and see longer trends in the data [79], [92]. Therefore, enhances the ability of a

classifier to observe dependencies and underlying relations among features to improve the accuracy of the classifier [92], [93]. In addition, using multiple time points as input of the classifier decreases the sensitivity for outliers, which is especially important when using noisy signals, such as sEMG [94]. Lastly, in future applications, the classification algorithm should perform as an object detector, directly recognising an object, determining its weight and acting on it by applying the appropriate support torque to the user. Therefore, further research should be conducted using regression models, which can identify and predict continuous variables, such as various weights, instead of being limited to certain classes. However, the dataset must be expanded to include a wider range of lifted weights.

The current dataset, derived from 11 healthy participants performing flexion and extension cycles while lifting several weights, is still limited. During the experiment, the user was limited in ROM by safety regulations and geometric constraints due to the IMU and arm support. Additionally, the study focused on a 1DOF situation of elbow flexion and extension. However, a comprehensive active-assistive device should include all shoulder degrees of freedom, and elbow pronation and supination. Prior research has shown the importance of freedom of movement in daily activities [95]. The performance of simple flexion and extension tasks in a controlled setting is a good starting point to gain insights into the influence of lifting weights on the different sensors. However, with the prospect of using the device in a real-life setting, the capabilities of the active-assistive device should be extended so that the device can respond appropriately to interaction with the environment in any situation.

## 5 CONCLUSION

In this research, an exploratory study has been performed to obtain insights into the effect of holding objects during elbow flexion and extension, as well as the effectiveness of data classification algorithms to distinguish the amount of weight lifted. The methodology involved conducting a human experiment to gather data from fourteen healthy participants. During the experiment, four different sensors were used to gather information about the interaction with the objects. Sensor measurements were used as input to a KNN and SVM classifier to classify the data into seven distinct labels, each representing a lifted weight. The best-performing classifier demonstrated an accuracy of 39.70%, indicating that there were many misclassifications. Within the KNN multi-class classifier, the highest recall percentage was shown in the 1000-gram class. Furthermore, accurate values were observed in the one-to-one classifications of 500- and 1000-gram. In conclusion, it can be argued that the used data classification algorithms can, to an extent,

identify the weights of lifted objects, particularly those weighing 500 and 1000 grams.

To effectively implement data classification, it is crucial that the accuracy is improved and that there are no misclassifications to protect this vulnerable patient group. Further investigation is required to examine the impact of picking up or placing an object, the impact of muscle weakness, the application of alternative data classification algorithms, and the expansion of the current degrees of freedom and movements in the dataset.

## ACKNOWLEDGMENTS

I'd like to thank my supervisors, Suzanne Filius, Laura Marchal Crespo, and Jaap Harlaar, for their guidance during my thesis project. A special thanks to the participants for dedicating their time to my experiment. Moreover, I would like to thank all students at the Clinical Biomechanics department for their assistance and all the motivating meetings we had together. Your collective contributions were of great value in delivering my thesis research.

## REFERENCES

- [1] T. M. Burns, C. D. Graham, M. R. Rose, and Z. Simmons, "Quality of life and measures of quality of life in patients with neuromuscular disorders," pp. 9–25, 7 2012.
- [2] M. de Visser, M. Vermeulen, J. Wokke, I. de Groot, and A. Wintzen, *Neuromusculaire ziekten*. Nederlands Tijdschrift Geneeskunde, 9 1999.
- [3] M. A. Grootenhuys, J. de Boone, and A. J. van der Kooij, "Living with muscular dystrophy: Health related quality of life consequences for children and adults," *Health and Quality of Life Outcomes*, vol. 5, 6 2007.
- [4] C. M. McDonald, "Physical activity, health impairments, and disability in neuromuscular disease," *American journal of physical medicine & rehabilitation*, vol. 81, no. 11 Suppl, 11 2002. [Online]. Available: <https://pubmed.ncbi.nlm.nih.gov/12409816/>
- [5] A. Bergsma, M. M. Janssen, A. C. Geurts, E. H. Cup, and I. J. de Groot, "Different profiles of upper limb function in four types of neuromuscular disorders," *Neuromuscular Disorders*, vol. 27, no. 12, pp. 1115–1122, 12 2017.
- [6] M. Gandolla, A. Antonietti, V. Longatelli, and A. Pedrocchi, "The Effectiveness of Wearable Upper Limb Assistive Devices in Degenerative Neuromuscular Diseases: A Systematic Review and Meta-Analysis," 1 2020.
- [7] L. A. Van Der Heide, B. Van Nijhuijs, A. Bergsma, G. J. Gelderblom, D. J. Van Der Pijl, and L. P. De Witte, "An overview and categorization of dynamic arm supports for people with decreased arm function," *Prosthetics and Orthotics International*, vol. 38, no. 4, pp. 287–302, 8 2014. [Online]. Available: [https://journals-sagepub-com.tudelft.idm.oclc.org/doi/10.1177/0309364613498538?url\\_ver=Z39.88-2003&rft\\_id=ori%3Arid%3Acrossref.org&rft\\_dat=cr\\_pub++0pubmed](https://journals-sagepub-com.tudelft.idm.oclc.org/doi/10.1177/0309364613498538?url_ver=Z39.88-2003&rft_id=ori%3Arid%3Acrossref.org&rft_dat=cr_pub++0pubmed)
- [8] P. T. Straathof, J. Lobo-Prat, F. Schilder, P. N. Kooren, M. I. Paalman, A. H. Stienen, and B. F. Koopman, "Design and control of the A-Arm: An active planar arm support for adults with Duchenne muscular dystrophy," in *Proceedings of the IEEE RAS and EMBS International Conference on Biomedical Robotics and Biomechatronics*, vol. 2016-July. IEEE Computer Society, 7 2016, pp. 1242–1247.

- [9] P. N. Kooren, J. Lobo-Prat, A. Q. Keemink, M. M. Janssen, A. H. Stienen, I. J. De Groot, M. I. Paalman, R. Verdaasdonk, and B. F. Koopman, "Design and control of the Active A-Gear: A wearable 5 DOF arm exoskeleton for adults with Duchenne muscular dystrophy," *Proceedings of the IEEE RAS and EMBS International Conference on Biomedical Robotics and Biomechatronics*, vol. 2016-July, pp. 637–642, 7 2016.
- [10] L. Van Der Heide and L. De Witte, "The perceived functional benefit of dynamic arm supports in daily life," *Journal of Rehabilitation Research and Development*, vol. 53, no. 6, pp. 1139–1150, 2016.
- [11] S. Dalla Gasperina, L. Roveda, A. Pedrocchi, F. Braghin, and M. Gandolla, "Review on Patient-Cooperative Control Strategies for Upper-Limb Rehabilitation Exoskeletons," *Frontiers in robotics and AI*, vol. 8, 12 2021. [Online]. Available: <https://pubmed.ncbi.nlm.nih.gov/tudelft.idm.oclc.org/34950707/>
- [12] M. A. Gull, S. Bai, and T. Bak, "A Review on Design of Upper Limb Exoskeletons," *Robotics 2020*, Vol. 9, Page 16, vol. 9, no. 1, p. 16, 3 2020. [Online]. Available: [https://www.mdpi.com/2218-6581/9/1/16](https://www.mdpi.com/2218-6581/9/1/16/htmlhttps://www.mdpi.com/2218-6581/9/1/16)
- [13] J. Lobo-Prat, A. Q. Keemink, B. F. Koopman, A. H. Stienen, and P. H. Veltink, "Adaptive Gravity and Joint Stiffness Compensation Methods," *IEEE International Conference on Rehabilitation Robotics (ICORR)*, 2015.
- [14] T. Rahman, R. Ramanathan, S. Strouw, W. Sample, R. Seliktar, W. Harwin, M. Alexander, and M. Scavina, "Towards the control of a powered orthosis for people with muscular dystrophy," *Journal of Engineering in Medicine*, vol. 215, no. 3, pp. 267–274, 3 2001.
- [15] R. A. Gopura, D. S. Bandara, K. Kiguchi, and G. K. Mann, "Developments in hardware systems of active upper-limb exoskeleton robots: A review," *Robotics and Autonomous Systems*, vol. 75, pp. 203–220, 1 2016.
- [16] J. Lee, M. Kim, and K. Kim, "A control scheme to minimize muscle energy for power assistant robotic systems under unknown external perturbation," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 12, pp. 2313–2327, 12 2017.
- [17] D. Ragonesi, S. K. Agrawal, W. Sample, and T. Rahman, "Quantifying anti-gravity torques for the design of a powered exoskeleton," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 21, no. 2, pp. 283–288, 2013.
- [18] J. Lobo Prat, "Control interfaces to actively support the arm function of men with duchenne muscular dystrophy," Ph.D. dissertation, University of Twente, Enschede, 9 2016.
- [19] K. Nizamis, J. Lobo-Prat, A. Q. Keemink, R. Carloni, A. H. Stienen, and B. F. Koopman, "Switching Proportional EMG Control of a 3DEndpoint Arm Support for People with Duchenne Muscular Dystrophy," pp. 235–240, 2015.
- [20] E. Bardi, M. Gandolla, F. Braghin, F. Resta, A. L. Pedrocchi, and E. Ambrosini, "Upper limb soft robotic wearable devices: a systematic review," 12 2022.
- [21] Q. Wu, B. Chen, and H. Wu, "Neural-network-enhanced torque estimation control of a soft wearable exoskeleton for elbow assistance," *Mechatronics*, vol. 63, 11 2019.
- [22] Y. Li and S. S. Ge, "Human-robot collaboration based on motion intention estimation," *IEEE/ASME Transactions on Mechatronics*, vol. 19, no. 3, pp. 1007–1014, 2014.
- [23] R. Tibold and A. J. Fuglevand, "Prediction of muscle activity during loaded movements of the upper limb," *Journal of NeuroEngineering and Rehabilitation*, vol. 12, no. 1, 2015.
- [24] H. M. Qassim and W. Z. Wan Hasan, "A review on upper limb rehabilitation robots," pp. 1–18, 10 2020.
- [25] H. D. Lee, B. K. Lee, W. S. Kim, J. S. Han, K. S. Shin, and C. S. Han, "Human-robot cooperation control based on a dynamic model of an upper limb exoskeleton for human power amplification," *Mechatronics*, vol. 24, no. 2, pp. 168–176, 3 2014.
- [26] X. Wang, Q. Song, S. Zhou, J. Tang, K. Chen, and H. Cao, "Multi-connection load compensation and load information calculation for an upper-limb exoskeleton based on a six-axis force/torque sensor," *International Journal of Advanced Robotic Systems*, vol. 16, no. 4, 7 2019.
- [27] M. Hosseini, R. Meattini, A. San-Millan, G. Palli, C. Melchiorri, and J. Paik, "A sEMG-Driven Soft ExoSuit Based on Twisted String Actuators for Elbow Assistive Applications," *IEEE Robotics and Automation Letters*, vol. 5, no. 3, pp. 4094–4101, 7 2020.
- [28] B. Treussart, F. Geffard, N. Vignais, and F. Marin, "Controlling an upper-limb exoskeleton by EMG signal while carrying unknown load," *Proceedings - IEEE International Conference on Robotics and Automation*, pp. 9107–9113, 5 2020.
- [29] R. Meattini, G. Palli, and C. Melchiorri, "Experimental evaluation of a sEMG-based control for elbow wearable assistive devices during load lifting tasks," *IEEE International Conference on Rehabilitation Robotics*, pp. 140–145, 8 2017.
- [30] N. Yang, J. Li, P. Xu, Z. Zeng, S. Cai, and L. Xie, "Design of Elbow Rehabilitation Exoskeleton Robot with sEMG-based Torque Estimation Control Strategy," *2022 6th International Conference on Robotics and Automation Sciences, ICRAS 2022*, pp. 105–113, 2022.
- [31] J. Rosen, J. C. Perry, N. Manning, S. Burns, and B. Hannaford, "The human arm kinematics and dynamics during daily activities - Toward a 7 DOF upper limb powered exoskeleton," *2005 International Conference on Advanced Robotics, ICAR '05, Proceedings*, vol. 2005, pp. 532–539, 2005.
- [32] L. Ricci, F. Taffoni, and D. Formica, "On the orientation error of IMU: Investigating static and dynamic accuracy targeting human motion," *PLoS ONE*, vol. 11, no. 9, 9 2016.
- [33] P. Herbin and M. Pajor, "Human-robot cooperative control system based on serial elastic actuator bowden cable drive in ExoArm 7-DOF upper extremity exoskeleton," *Mechanism and Machine Theory*, vol. 163, p. 104372, 9 2021.
- [34] L. Zhang, G. Liu, B. Han, Z. Wang, and T. Zhang, "sEMG Based Human Motion Intention Recognition," 2019.
- [35] National Instruments, "USB-6008 Specifications," Tech. Rep., 2023.
- [36] Beckhoff, "EL3104 — EtherCAT Terminal, 4-channel analog input — Product information Technical data," Tech. Rep., 6 2023.
- [37] FUTEK, "Miniature S-Beam Jr. Load Cell (LSB200)," [Online]. Available: <https://www.futek.com/store/legacy-sensors-and-instruments/miniature-s-beam-lsb200/FSH00105>
- [38] Movella, "MTi-1," Tech. Rep., 5 2023.
- [39] O. S. Al-Dahiree, R. A. R. Ghazilla, M. O. Tokhi, H. J. Yap, and M. Gul, "Design and Characterization of a Low-Cost and Efficient Torsional Spring for ES-RSEA," *Sensors*, vol. 23, no. 7, 4 2023.
- [40] B. Boehler, "Electromyographic analysis of the triceps brachii muscle during a variety of triceps exercises," 5 2011. [Online]. Available: <https://minds.wisconsin.edu/handle/1793/53487>
- [41] E. Kholinne, R. F. Zulkarnain, Y. C. Sun, S. J. Lim, J. M. Chun, and I. H. Jeon, "The different role of each head of the triceps brachii muscle in elbow extension," *Acta Orthopaedica et Traumatologica Turcica*, vol. 52, no. 3, p. 201, 5 2018. [Online]. Available: <https://pubmed.ncbi.nlm.nih.gov/tudelft.idm.oclc.org/pmc/articles/PMC6136322/>

- [42] Delsys Incorporated, "DELSYS ® Bagnoli TM EMG System User's Guide," Tech. Rep.
- [43] H. J. Hermens, B. Freriks, R. Merletti, D. Stegeman, J. Blok, G. Rau, C. Disselhorst-Klug, and G. Hägg, "European Recommendations for Surface Electromyography Results of the SENIAM project," *SENIAM Standards*, 1999.
- [44] D. A. Winter, *Biomechanics and motor control of human movement*, 4th ed. Waterloo: John Wiley & Sons, inc., 2009.
- [45] D. Meldrum, E. Cahalane, R. Conroy, D. Fitzgerald, and O. Hardiman, "Maximum voluntary isometric contraction: Reference values and clinical application," vol. 8, no. 1, pp. 47–55, 2009. [Online]. Available: <https://www.tandfonline.com/tudelft.idm.oclc.org/doi/abs/10.1080/17482960601012491>
- [46] T. Wei, "Validation of a weight compensation for wearable upper extremity support system," Ph.D. dissertation, Delft University of Technology, Delft, 5 2023. [Online]. Available: <http://resolver.tudelft.nl/uuid:6d5f541c-4666-479c-a1e5-e364f06af39a>
- [47] K. Papa, "Passive Elbow Joint Impedance Identification with the use of an Elbow Device," Ph.D. dissertation, Delft University of Technology, Delft, 8 2023. [Online]. Available: <http://resolver.tudelft.nl/uuid:1f6e5d56-66e1-4cea-b626-266c9aded162>
- [48] "PUL 2.0 original scoresheet." [Online]. Available: [https://plos.figshare.com/articles/figure/Upper\\_limb\\_function\\_in\\_Duchenne\\_muscular\\_dystrophy\\_24\\_month\\_longitudinal\\_data/6624272](https://plos.figshare.com/articles/figure/Upper_limb_function_in_Duchenne_muscular_dystrophy_24_month_longitudinal_data/6624272)
- [49] M. Xiloyannis, D. Chiaradia, A. Frisoli, and L. Masia, "Physiological and kinematic effects of a soft exosuit on arm movements," *Journal of NeuroEngineering and Rehabilitation*, vol. 16, no. 1, pp. 1–15, 2 2019. [Online]. Available: <https://jneuroengrehab.biomedcentral.com/articles/10.1186/s12984-019-0495-y>
- [50] M. A. Buckley, A. Yardley, G. R. Johnson, and D. A. Carus, "Dynamics of the upper limb during performance of the tasks of everyday living - A review of the current knowledge base," *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, vol. 210, no. 4, pp. 241–245, 1996.
- [51] N. Williams, "The Borg Rating of Perceived Exertion (RPE) scale," *Occupational Medicine*, vol. 67, no. 5, pp. 404–405, 7 2017.
- [52] John Brooke, "System Usability Scale (SUS)," *Usability Eval. ind.*, vol. 189, 11 1995.
- [53] A. Bangor, P. T. Kortum, and J. T. Miller, "An empirical evaluation of the system usability scale," *International Journal of Human-Computer Interaction*, vol. 24, no. 6, pp. 574–594, 8 2008.
- [54] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research," *Elsevier Science Publishers B.V.*, vol. 52, pp. 139–183, 1 1988.
- [55] G. Wu and P. R. Cavanagh, "ISB RECOMMENDATIONS FOR STANDARDIZATION IN THE REPORTING OF KINEMATIC DATA," Tech. Rep. 10, 1995.
- [56] B. Ugurlu, M. Nishimura, K. Hyodo, M. Kawanishi, and T. Narikiyo, "A framework for sensorless torque estimation and control in wearable exoskeletons," *International Workshop on Advanced Motion Control, AMC*, 2012.
- [57] J. Wang, L. Tang, and J. E Bronlund, "Surface EMG Signal Amplification and Filtering," *International Journal of Computer Applications*, vol. 82, no. 1, pp. 15–22, 11 2013.
- [58] M. Irfan, K. Ullah, F. Muhammad, S. Khan, F. Althobiani, M. Usman, M. Alshareef, S. Alghaffari, and S. Rahman, "Automatic Detection of Outliers in Multi-Channel EMG Signals Using MFCC and SVM," *Intelligent Automation and Soft Computing*, vol. 36, no. 1, pp. 169–181, 2023.
- [59] N. Lotti, M. Xiloyannis, G. Durandau, E. Galofaro, V. Sanguineti, L. Masia, and M. Sartori, "Adaptive model-based myoelectric control for a soft wearable arm exosuit: A new generation of wearable robot control," *IEEE Robotics and Automation Magazine*, vol. 27, no. 1, pp. 43–53, 3 2020.
- [60] O. Kramer, "Dimensionality Reduction with Unsupervised Nearest Neighbors 123," Tech. Rep. [Online]. Available: <http://www.springer.com/series/8578>
- [61] T. Liu, A. W. Moore, and A. Gray, "New Algorithms for Efficient High-Dimensional Nonparametric Classification," Tech. Rep., 2006.
- [62] J. Cervantes, F. Garcia-Lamont, L. Rodríguez-Mazahua, and A. Lopez, "A comprehensive survey on support vector machine classification: Applications, challenges and trends," *Neurocomputing*, vol. 408, pp. 189–215, 9 2020.
- [63] A. A. Soofi and A. Awan, "Classification Techniques in Machine Learning: Applications and Issues," *Journal of Basic & Applied Sciences*, vol. 13, pp. 459–465, 2017.
- [64] A. Sikidar, K. E. C. Vidyasagar, M. Gupta, B. Garg, and D. Kalyanasundaram, "Classification of mild and severe adolescent idiopathic scoliosis (AIS) from healthy subjects via a supervised learning model based on electromyogram and ground reaction force data during gait," *Biocybernetics and Biomedical Engineering*, vol. 42, no. 3, pp. 870–887, 7 2022.
- [65] A. Dellacasa Bellingegni, E. Gruppioni, G. Colazzo, A. Davalli, R. Sacchetti, E. Guglielmelli, and L. Zollo, "NLR, MLP, SVM, and LDA: A comparative analysis on EMG data from people with trans-radial amputation," *Journal of NeuroEngineering and Rehabilitation*, vol. 14, no. 1, 8 2017.
- [66] M. U. Khan, H. Khan, M. Muneeb, Z. Abbasi, U. B. Abbasi, and N. K. Baloch, "Supervised machine learning based fast hand gesture recognition and classification using electromyography (EMG) signals," in *ICAEM 2021 - 2021 International Conference on Applied and Engineering Mathematics, Proceedings*. Institute of Electrical and Electronics Engineers Inc., 8 2021, pp. 81–86.
- [67] A. Géron, "Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow SECOND EDITION Concepts, Tools, and Techniques to Build Intelligent Systems," Tech. Rep., 2019. [Online]. Available: <http://oreilly.com>
- [68] H. M and S. M.N, "A Review on Evaluation Metrics for Data Classification Evaluations," *International Journal of Data Mining & Knowledge Management Process*, vol. 5, no. 2, pp. 01–11, 3 2015.
- [69] T. Swinscow and M. J. Campbell, *Statistics at square one*, 9th ed. BMJ Publishing Group, 1996.
- [70] I. Farkaš, P. Masulli, S. Otte, and S. Wermter, "Artificial Neural Networks and Machine Learning-ICANN 2021," Bratislava, Tech. Rep., 9 2021. [Online]. Available: <http://www.springer.com/series/7407>
- [71] N. G. Raju, K. P. Lakshmi, V. M. J. Scholar, A. Kalidindi, and V. Padma, "Study the Influence of Normalization/Transformation process on the Accuracy of Supervised Classification," *Third International Conference on Smart Systems and Inventive Technology (ICSSIT 2020)*, 2020.
- [72] C. Huyen, *Designing Machine Learning Systems*. O'Reilly Media, 2022, vol. 1.
- [73] S. Das and U. M. Cakmak, "Hands-On Automated Machine Learning : A Beginner's guide to Building Automated Machine Learning Systems using AutoML and Python." Tech. Rep., 2018.
- [74] U. Stańczyk and L. C. Jain, *Feature Selection for Data and Pattern Recognition*, 2015, vol. 584. [Online]. Available: <http://www.springer.com/series/7092>
- [75] Z. H. Zhou, *Machine Learning*. Springer Nature, 1 2021.

- [76] F. Maulidina, Z. Rustam, S. Hartini, V. V. Wibowo, I. Wirasati, and W. Sadewo, "Feature optimization using Backward Elimination and Support Vector Machines (SVM) algorithm for diabetes classification," in *Journal of Physics: Conference Series*, vol. 1821, no. 1. IOP Publishing Ltd, 3 2021.
- [77] H. Raza, H. Cecotti, and G. Prasad, "Optimising Frequency Band Selection with Forward-Addition and Backward-Elimination Algorithms in EEG-based Brain-Computer Interfaces," *International Joint Conference on Neural Networks (IJCNN)*, pp. 1–7, 2015.
- [78] scikit-learn developers, "sklearn.model\_selection.GridSearchCV." [Online]. Available: [https://scikit-learn.org/stable/modules/generated/sklearn.model\\_selection.GridSearchCV.html](https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html)
- [79] E. Elgeldawi, A. Sayed, A. R. Galal, and A. M. Zaki, "Hyperparameter tuning for machine learning algorithms used for arabic sentiment analysis," *Informatics*, vol. 8, no. 4, 12 2021.
- [80] scikit-learn developers, "sklearn.svm.SVC." [Online]. Available: <https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html>
- [81] R. Vohra, A. Hussain, A. K. Dudyala, J. Pahareeya, and W. Khan, "Multi-class classification algorithms for the diagnosis of anemia in an outpatient clinical setting," *PLoS ONE*, vol. 17, no. 7 July, 7 2022.
- [82] B. Bischl, J. Schiffner, and C. Weihs, "Benchmarking local classification methods," *Computational Statistics*, vol. 28, no. 6, pp. 2599–2619, 12 2013.
- [83] P. Borah and D. Gupta, "Affinity and transformed class probability-based fuzzy least squares support vector machines," *Fuzzy Sets and Systems*, vol. 443, pp. 203–235, 8 2022.
- [84] M. Grandini, E. Bagli, and G. Visani, "Metrics for Multi-Class Classification: an Overview," 8 2020. [Online]. Available: <http://arxiv.org/abs/2008.05756>
- [85] J. G. Kreifeldt, "Signal Versus Noise Characteristics of Filtered EMG Used as a Control Source," *Tech. Rep.* 1, 1971.
- [86] B. Chan, I. Saad, N. Bolong, and K. E. Siew, "A Review of Surface EMG in Clinical Rehabilitation Care Systems Design," in *2021 IEEE 19th Student Conference on Research and Development (SCoReD)*, 11 2021, pp. 371–376.
- [87] M. Frascarelli, L. Rocchi, and I. M. D. Feola, "EMG COMPUTERIZED ANALYSIS OF LOCALIZED FATIGUE IN DUCHENNE MUSCULAR DYSTROPHY," *Tech. Rep.*
- [88] E. Lindeman, F. Spaans, J. P. H. Reulen, P. Leffers, and J. Drukker, "Surface EMG of proximal leg muscles in neuromuscular patients and in healthy controls. Relations to force and fatigue," *Tech. Rep.*, 1999. [Online]. Available: [www.elsevier.com/locate/jelekin](http://www.elsevier.com/locate/jelekin)
- [89] I. Elamvazuthi, N. H. Duy, Z. Ali, S. W. Su, M. K. Khan, and S. Parasuraman, "Electromyography (EMG) based Classification of Neuromuscular Disorders using Multi-Layer Perceptron," in *Procedia Computer Science*, vol. 76. Elsevier B.V., 2015, pp. 223–228.
- [90] S. Koçer, "Classification of emg signals using neuro-fuzzy system and diagnosis of neuromuscular diseases," *Journal of Medical Systems*, vol. 34, no. 3, pp. 321–329, 6 2010.
- [91] M. B. Wagner, P. J. Vignos, C. Carozzi, and A. L. Hull, "Assessment of Hand Function in Duchenne Muscular Dystrophy," *Tech. Rep.*, 1993.
- [92] P. Geurts, "Pattern Extraction for Time Series Classification," *Tech. Rep.*, 2001.
- [93] D. H. Kisa, M. A. Ozdemir, O. Guren, and A. Akan, "EMG based Hand Gesture Classification using Empirical Mode Decomposition Time-Series and Deep Learning," in *TIPTEKNO 2020 - Tip Teknolojileri Kongresi - 2020 Medical Technologies Congress, TIPTEKNO 2020*. Institute of Electrical and Electronics Engineers Inc., 11 2020.
- [94] F. Lotte, L. Bougrain, A. Cichocki, M. Clerc, M. Congedo, A. Rakotomamonjy, and F. Yger, "A review of classification algorithms for EEG-based brain-computer interfaces: A 10 year update," 4 2018.
- [95] K. Bienias, J. Ścibek, J. Cegielska, and J. Kochanowski, "Evaluation of activities of daily living in patients with slowly progressive neuromuscular diseases," *Neurologia i Neurochirurgia Polska*, vol. 52, no. 2, pp. 222–227, 3 2018.

## APPENDIX A

### NOMENCLATURE

Symbol	Definition	Unit
<i>1DOF</i>	1-Degree-Of-Freedom	
<i>ADL</i>	Activities of Daily Living	
<i>DMD</i>	Duchenne Muscular Dystrophy	
<i>HREC</i>	Human Research Ethics Committee	
<i>IMU</i>	Inertial Measurement Unit	
<i>KNN</i>	K Nearest Neighbour	
<i>MVC</i>	Maximum Voluntary Contraction	
<i>NMDs</i>	Neuromuscular Diseases	
<i>PUL</i>	Performance of the Upper Limb	
<i>RF</i>	Random Forest Classifier	
<i>ROM</i>	Range of Motion	
<i>RPE</i>	Rating of Perceived Exertion	
<i>SEA</i>	Series Elastic Actuator	
<i>sEMG</i>	surface Electromyography	
<i>SNR</i>	Signal-to-Noise Ratio	
<i>SUS</i>	System Usability Scale	
<i>SVM</i>	Support Vector Machine	
$\alpha$	shoulder abduction & adduction angle	deg
$\beta$	shoulder joint flexion & extension angle	deg
$\theta$	elbow joint flexion & extension angle	deg
$v$	elbow joint velocity	rad/s
$\theta_x, \theta_y, \theta_z$	IMU angles about the three principal axes	deg
$v_x, v_y, v_z$	IMU angular velocity about the three principal axes	rad/s
$a_x, a_y, a_z$	IMU linear acceleration in three dimensions	$m/s^2$
$F_z$	interaction force	N
$T_z$	force sensor joint torque	N*m
$T$	elbow joint torque	N*m

## APPENDIX B SET-UP EXPERIMENT

### B.1 Overview setup experiment

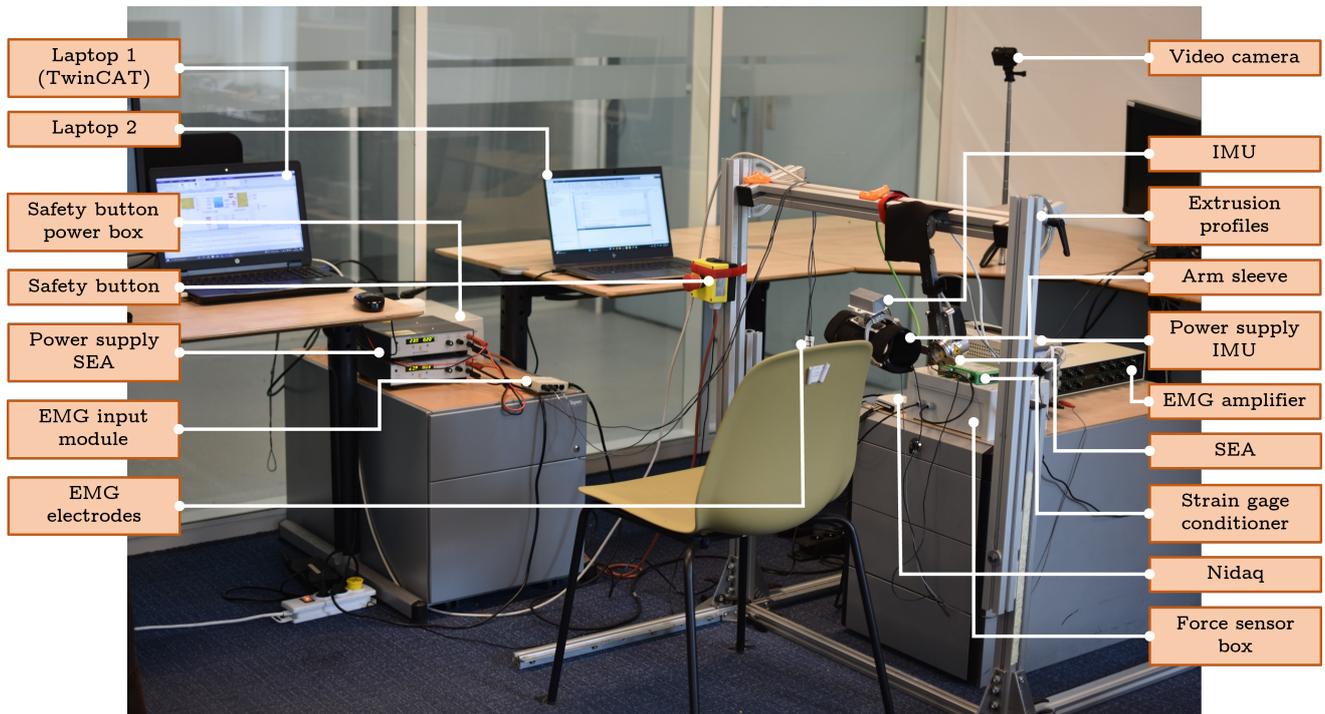


Fig. 7. Overview setup experiment

### B.2 Technical diagram setup experiment

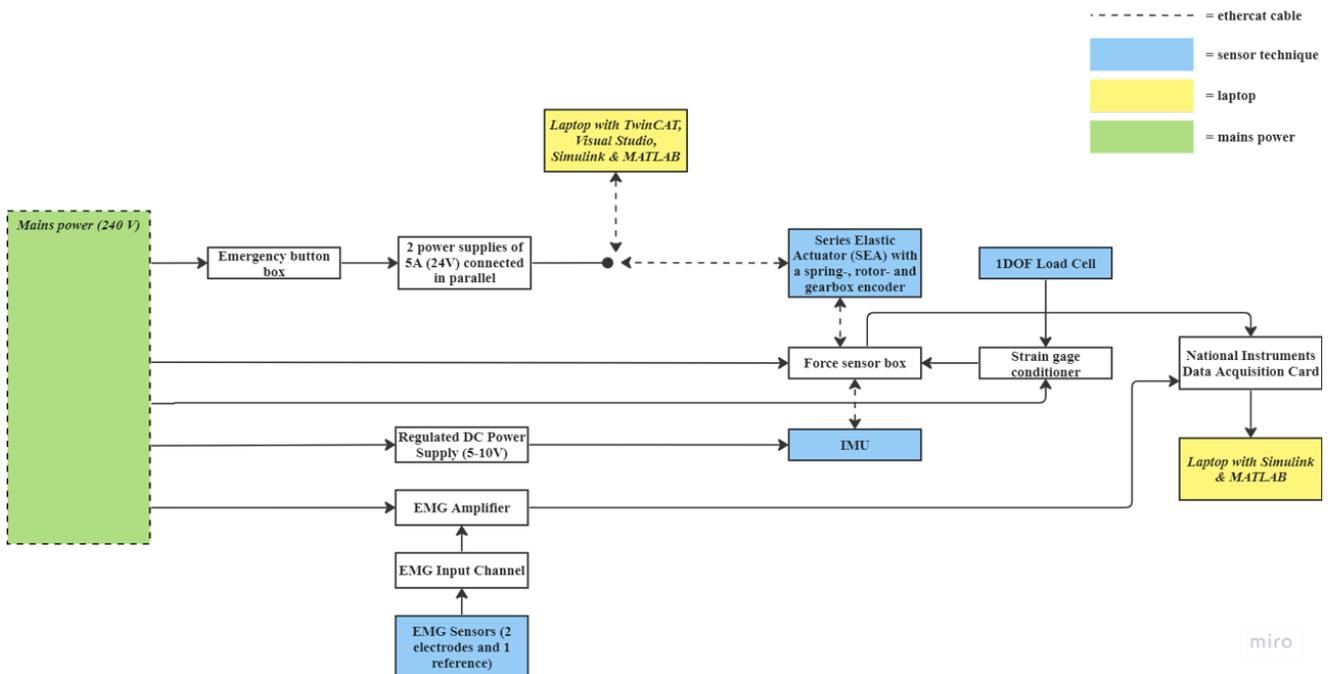


Fig. 8. Schematic diagram of the electronics

## APPENDIX C

### SENSORS

Within this appendix, only the FUTEK Load Cell will be discussed due to its specific calibration method.

#### C.1 FUTEK Load Cell

This section focuses on the calibration process of the FUTEK Load Cell, measuring the interaction force ( $F_z$ ) between the human arm and the active-assistive device.

Calibration is necessary to convert the raw output of the force sensor, given in bits ( $2^{16} - 1$  bits,  $\pm 10$  V), into accurate force values (Newton). The calibration employs two position-controlled measurements, moving in four cycles from -68 degrees elbow extension to 165 degrees elbow flexion, and back again at a constant angular velocity of  $0.10 \text{ rad/s}$ . The difference between the position-controlled measurements is the known weight attached to the force sensor. During one measurement, only the arm sleeve is connected to the force sensor. During the other measurement, a deadweight of 1.078 kg is placed within the arm sleeve. While performing the position-controlled measurements, the elbow joint flexion/extension angle ( $\theta$ ) is captured by the joint encoder, and the interaction force ( $F_z$ ) by the force sensor.

Due to the linear characteristic of the force sensor, a linear function ( $y = a * x + b$ ) can be determined for the conversion of bits to Newton, based on the calibration measurements. These variables will be included in the linear function:

- $x$  is equal to the raw output of the force sensor
- $a$  is the slope, equal to the calibration factor. The calibration factor is based on the ratio between the model-based gravity acting on the deadweight ( $m_{dw} \cdot g_{acc} \cdot \cos \alpha$ ), and the raw force measurements of the measurement with deadweight, and without deadweight. The measurements used for this force ratio are measured at the position shown in Figure 9b. In this position, pure gravitational forces ( $F_{grav\_dw\_sleeve}$  &  $F_{grav\_sleeve}$ ) are measured while the interaction forces ( $F_z$ ) are aligned with the global gravity axes ( $Y$ ).
- $b$  is the offset, equal to the measured absolute offset multiplied by the calibration factor to convert it into Newton. The offset measurement uses the force information in a 0-degree elbow joint angle without deadweight, as shown in Figure 9c. Now, the measured interaction force ( $F_z$ ) is orthogonal to the global gravity axis ( $Y$ ), having no influence of the gravity, and therefore only the absolute offset is measured.

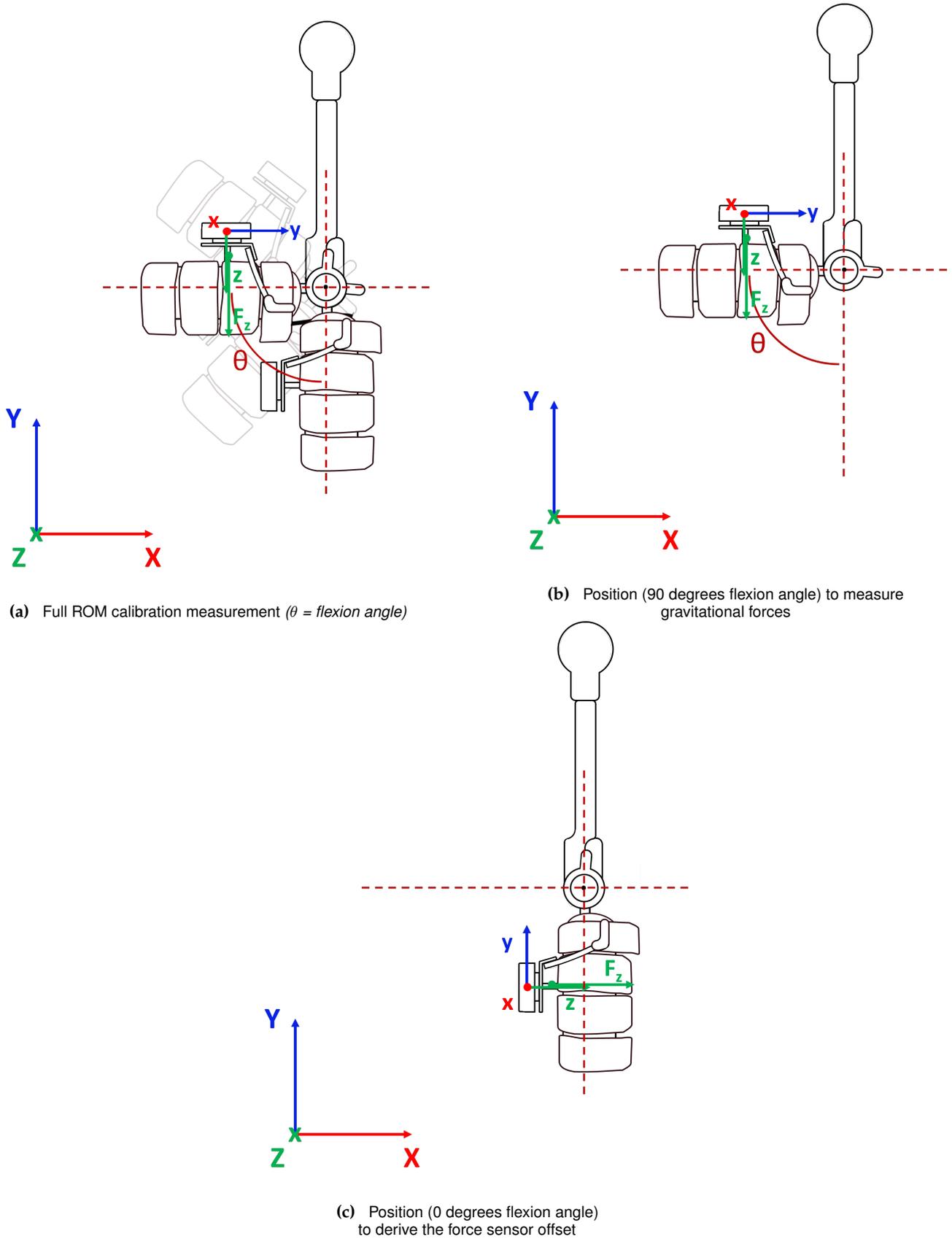
Combining all this information results in the following linear calibration function:

$$\begin{aligned}
 y &= a \cdot x + b \\
 F_N &= a \cdot F_{\text{BITS}} + b \\
 F_N &= \text{cal\_factor} \cdot F_{\text{BITS}} - (\text{offset} \cdot \text{cal\_factor}) \\
 &= \left( \frac{m_{dw} \cdot g_{acc} \cdot \cos \alpha}{F_{grav\_dw\_sleeve} - F_{grav\_sleeve}} \right) \cdot F_{\text{BITS}} - \left( \text{offset} \cdot \frac{m_{dw\_sleeve} \cdot g_{acc} \cdot \cos \alpha}{F_{grav\_dw\_sleeve} - F_{grav\_sleeve}} \right)
 \end{aligned} \tag{1}$$

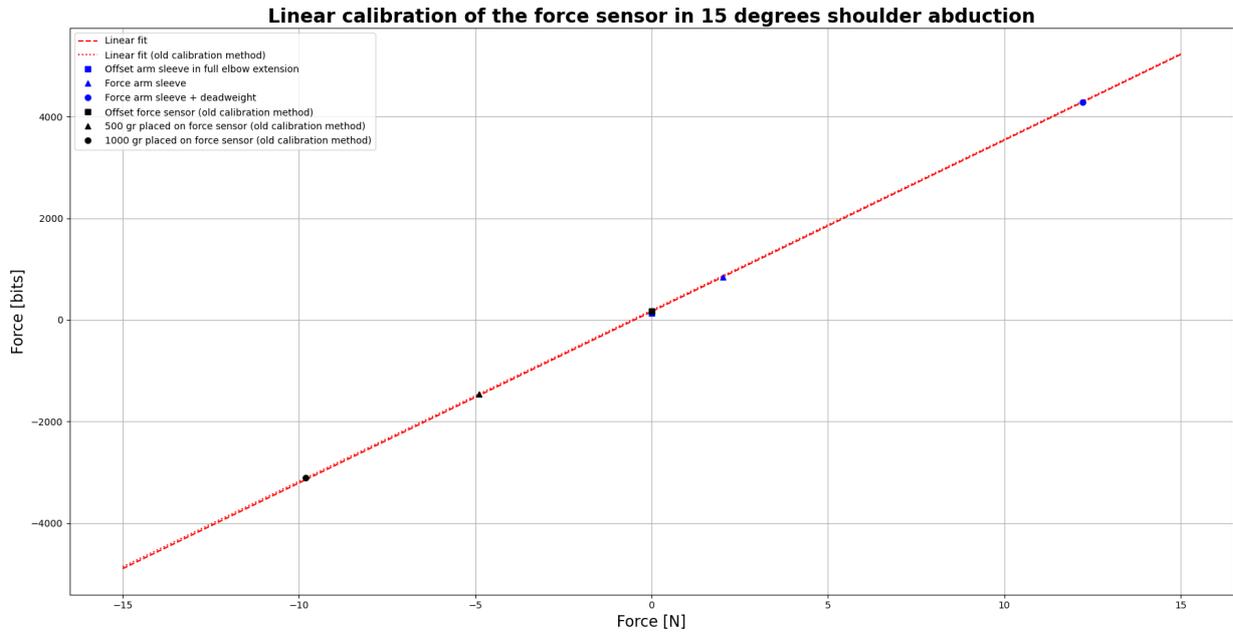
Daily calibration was performed due to changing environmental conditions and changes in the experimental setup. All measurements are performed in a 15-degrees shoulder abduction angle.

##### C.1.1 Exception calibration measurements (subject S26 & S63)

A different calibration setup was used during the first two experiments (S26 & S63). During this calibration measurement, the force sensor was not integrated in the exoskeleton but was taken out of the setup and placed on a table. During the calibration, an offset was measured when the force sensor was placed on a flat table, having no other interactions with the environment. Afterwards, several weights were put on top of the force sensor to fit a linear function through these data points. The linear fit applied for both calibration methods can be seen in Figure 10, showing almost no differences. Therefore, subjects S26 & S63 have been included in the experiment.



**Fig. 9.** Force sensor calibration measurements (green arrow =  $F_z$ ,  $X, Y, Z$  - global axes,  $x, y, z$  - local IMU axes)



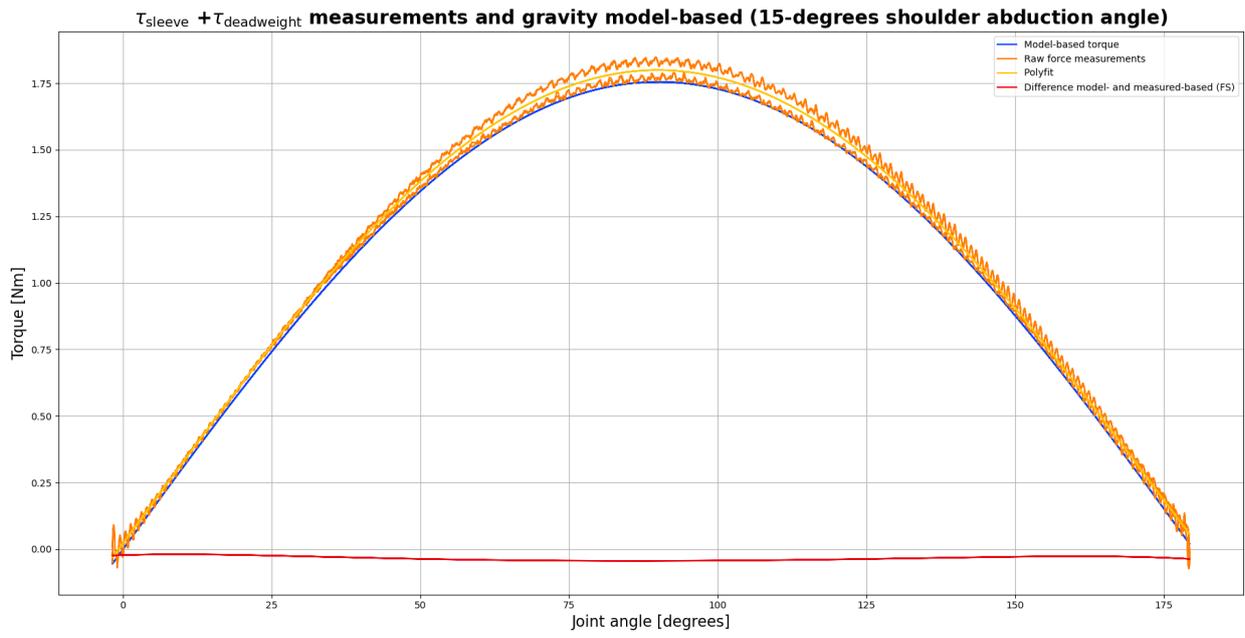
**Fig. 10.** Linear fit for two different force sensor calibration methods. One calibration method is based on the force sensor incorporated within the experimental setup (red dashed line), and another method is based on the force sensor calibration out of the experimental setup (red dotted line)

### C.1.2 Results calibration

The accuracy of the force sensor calibration is tested by comparing the results to a gravity model. The model-based torque was calculated by using the mass of both the deadweight ( $m_{deadweight}$ ) and the sleeve ( $m_{sleeve}$ ), and multiplying this with the gravity acceleration ( $g_{acc}$ ), the distance between the elbow joint and the force sensor ( $l_{FS}$ ), the shoulder abduction angle ( $\cos \alpha$ ), and the elbow flexion angle ( $\sin \theta$ ). This equation was used to derive the model-based torque of the arm sleeve ( $\tau_{sleeve}$ ), and the deadweight ( $\tau_{deadweight}$ ):

$$\tau_{deadweight} + \tau_{sleeve} = (m_{deadweight} + m_{sleeve}) \cdot g_{acc} \cdot \cos(\alpha) \cdot \sin(\theta) \cdot l_{FS} \quad (2)$$

The only variable that changes over time is  $\theta$ , the elbow flexion angle. In Figure 11, the orange line represents the combined torque of the sleeve and deadweight based on force sensor measurements, on which a polynomial fit is performed (yellow). The blue line represents the gravity-model-based values. Comparing these lines, only small differences between the measured-based and model-based torque occur. The maximum difference in torque between the measured- and model-based values is  $0.044 \text{ N} \cdot \text{m}$ . The difference in values of the force sensor measurements is striking, but this can be attributed to known hysteresis in the force sensor, having a maximum value of 0.1, meaning a maximum difference of 0.1% between an increasing and decreasing load [37].



**Fig. 11.** Measured- and model-based torque measurements, based on the force sensor

## APPENDIX D

### ARM WEIGHT COMPENSATION

In this research, a measured-based arm weight compensation strategy was used based on precise force measurements, which are recorded by the FUTEK load cell, of the passive forces acting on the forearm and hand, including the effect of gravity and internal passive joint impedance. The following steps were taken within this compensation strategy:

- Position-controlled dynamic calibration measurement - To identify the passive forces acting on the human forearm and hand, a position-controlled dynamic measurement is used over 95% of the participant's maximum ROM in elbow flexion and extension.
- sEMG inclusion - The force sensor measurements will be included or excluded for torque calculation based on the sEMG activity of the biceps and triceps.
- Elbow joint torque calculation - Based on the included force measurements, the support torque will be calculated.
- Torque control - 80% of the support torque is exerted on the user via torque control

#### D.1 Position controlled measurement

To identify the passive forces acting on the human forearm and the hand, position-controlled measurements are done. To perform these measurements, the actuator is moving within 95% of the participant's maximum ROM, with an elbow joint velocity of  $0.10 \text{ rad/s}$ . This is repeated for 8 cycles. During the identification, measurements of the joint encoder (joint angle [ $\text{rad}$ ], joint torque [ $N * m$ ]), force sensor (force [ $N$ ]), and sEMG (sEMG activity of the biceps and triceps [ $V$ ]) were taken.

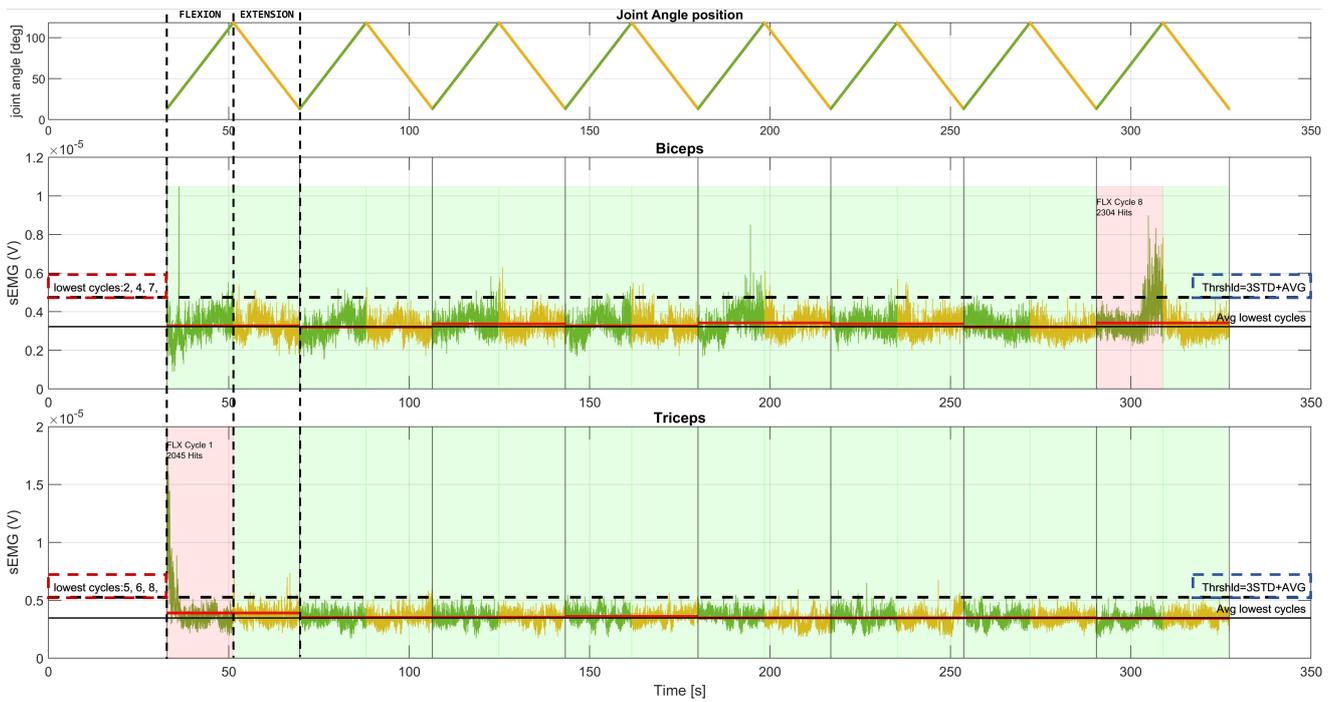
#### D.2 sEMG inclusion

Based on the sEMG activity of the biceps, and triceps, the force measurements within a flexion or extension cycle were included or excluded. This step is essential to ensure that the user was at rest during the measurement and that no other forces will be measured than the passive forces. This process consists of the following steps:

- The collected data was split into sixteen subparts, to divide every cycle into flexion and extension, based on the joint angle. In Figure 12, the separated flexion and extension cycles are shown as the green part (flexion) and the yellow part (extension) of the graphs.
- The average sEMG activity, combined for every flexion and extension cycle, is derived, shown as red solid lines in the sEMG biceps activity and sEMG triceps activity graphs in Figure 12.
- An average sEMG activity is derived, based on the three flexion/extension cycles showing the lowest sEMG activity. This was performed separately for the biceps and triceps activity. In addition, the standard deviation for these cycles was derived, multiplied by three and added to the mean value, representing the noise threshold of the sEMG signal [?]. This value represents the upper threshold, shown in the graph as the horizontal black dashed line in the sEMG activity graphs, which preferably will not be exceeded to make sure the participants is at rest during the measurements. In Figure 12, the cycles with the lowest activity on which the average is based are indicated on the left, within the dashed red coloured block.
- When more than 3.5% of the data samples in a flexion or extension cycle exceed the upper threshold, the force sensor measurements will be excluded. In a flexion cycle, this concerns the sEMG biceps activity, and in the extension cycle, this concerns the sEMG triceps activity. The first and last (eight) flexion and extension cycles were always excluded.
- The force measurements, collected by the force sensor, from the included cycles will be used for the elbow joint torque support calculation.

#### D.3 Elbow joint support torque calculation

Some processing steps should be taken to go from force measurements to elbow joint support torque. First, the force sensor data will be filtered. Subsequently, to obtain a torque value, the force sensor measurements must be multiplied by the distance of the elbow joint to the force sensor ( $l_{FS}$ ). Moreover, the SEA exerts the support torque ( $\tau_{support}$ ). Consequently, it is essential to incorporate the additional torques ( $\tau_{add}$ ) created by the active-assistive device itself. These additional torques ( $\tau_{add}$ ) consist of the torque exerted by the IMU, the rigid bar aligned with the forearm, and the force sensor itself. The additional torque ( $\tau_{add}$ ) is based on the combined mass of these parts ( $m_{add}$ ), and the combined center of mass ( $COM_{add}$ ). To derive the support torque ( $\tau_{support}$ ), the following equation must be incorporated:



**Fig. 12.** Example of sEMG inclusion, showing raw sEMG measurements for the biceps and triceps during the position-controlled measurements. A flexion or extension cycle will be excluded when 3.5% of the data samples exceed the upper threshold, meaning that the user is not relaxing the arm

$$\tau_{support}(\theta) = F(\theta) \cdot l_{FS} \cdot \tau_{add}(\theta) \quad (3)$$

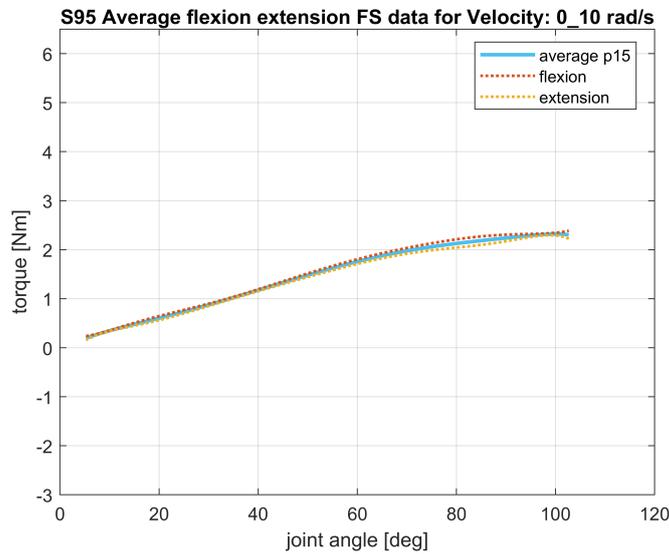
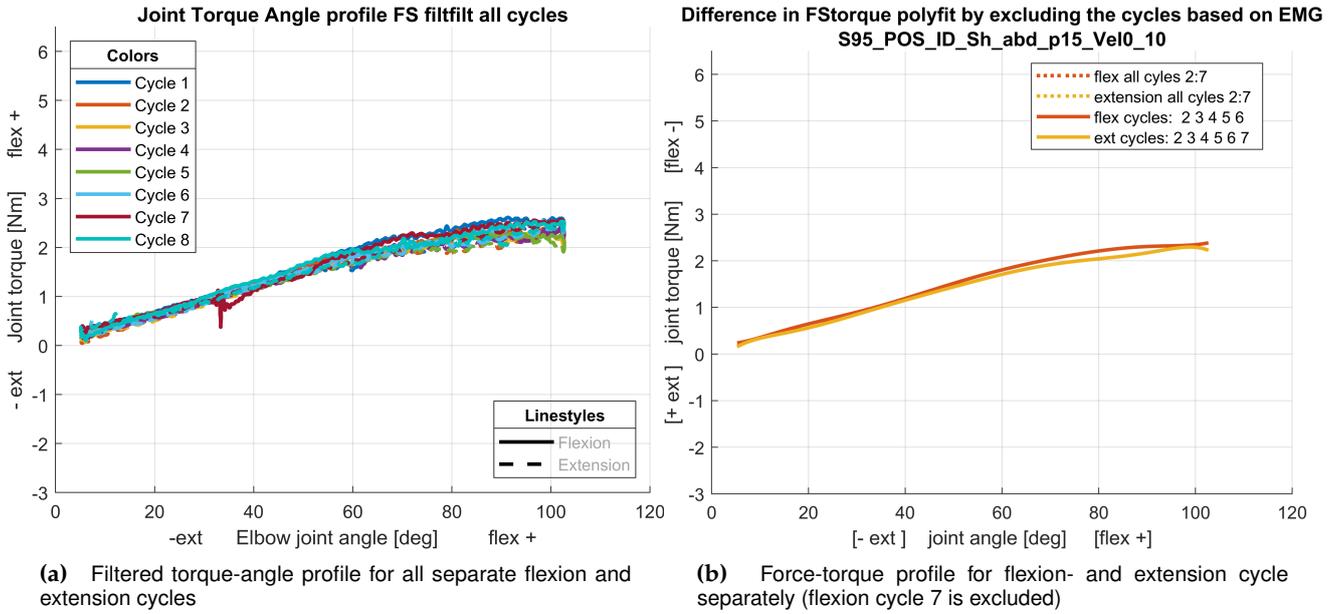
$$\tau_{add}(\theta) = m_{add} \cdot COM_{add} \cdot g_{acc} \cdot \sin(\theta) \cdot \cos(\alpha) \quad (4)$$

In Figure 13a, the torque-angle profile is shown for every flexion/extension cycle, after filtering and applying the equations to go from force measurements to torque. These torque measurements are interpolated, and the average of all flexion and all extension cycles over the joint angles is derived. Afterwards, a 7th-order polynomial fit is applied to this profile to smoothen the data, resulting in a combined torque-angle profile for the flexion and extension cycle separately, as shown in Figure 13b. Lastly, an average torque-angle profile is calculated for both flexion- and extension cycle, as shown in Figure 13c. The coefficients of this polynomial fit will be saved within a look-up table and used as support torque within the active-assistive device.

#### D.4 Torque control

Throughout the experiment, only 80% of the derived support torque was exerted on the user. Due to the sensitivity of the force sensor was chosen to decrease the amount of support. On the one hand, this increased the impact of passive forces on the measurements, while on the other hand, it ensured that there were no interaction forces between the user and the device in a situation where the support is overestimated.

Within the active-assistive device, the look-up table containing the coefficients of the torque-angle profile is used to determine the support torque based on the elbow joint angle ( $\theta$ ). This torque is exerted on the user, considering the current support torque and the torque safety limits.



**(c)** Combined force-torque profile including both flexion- and extension cycle

**Fig. 13.** Calculation of the support torque-angle profile based on force measurements of a FUTEK load cell

## APPENDIX E DATA PREPARATION

After filtering the data, several data preparation steps were performed before analysing and classifying the data. These steps will be elaborated on in this appendix.

### E.1 Trigger signal

As described in the Methods section, a trigger signal synchronises the Simulink/TwinCAT environment and the sEMG system. Nonetheless, certain deficiencies persist in the sEMG system’s data acquisition script on how MATLAB stores NiDAQ data. Additionally, the trigger signal was absent in the collected sEMG data. Therefore, this resulted in a difference in the amount of samples of the sEMG data and data captured in the TwinCAT environment.

Additional measurements revealed that extra samples were recorded at the beginning and end of the measurement. This happened even though the trigger signal stopped, and the trigger signal value was set back to 0 Volt. In Figure 14, an example of this discrepancy in gathered data samples is shown. Before initiating the trigger signal, 71 samples were already recorded. Furthermore, at the end of the trigger signal, 43 additional measurements were recorded. Therefore, the following strategy has been adopted to manage the issue of the trigger sample inconsistency. Due to the storage of the supplementary sEMG samples, the total sample number is always equivalent to or exceeds the number of TwinCAT samples. To address this problem, the sample counts from both systems are compared, resulting in the identification of redundant samples. As shown in Figure 14, both at the beginning and end of a measurement, redundant samples are taken. Therefore, the number of redundant samples will be halved. At the start and end of an sEMG measurement, this quantity of samples is removed.

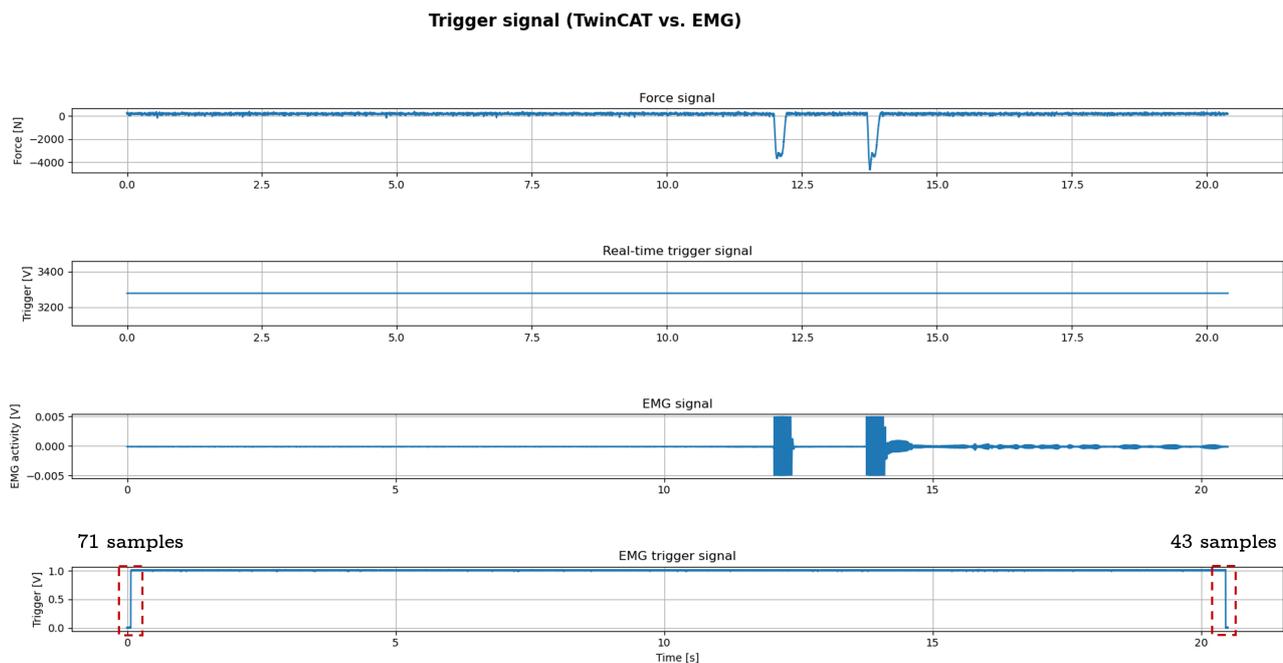


Fig. 14. Sample difference in trigger signal

### E.2 Delay removal

One of the safety regulations is the implementation of a software end-stop, based on the ROM of every participant, to ensure that this is not exceeded. While the actuator is powered, the participant can move freely within 90% of the range of motion. When the participant exceeds 90% of the ROM, the actuator will give an error and switch off. Angle guards have been implemented 2.9 degrees (0.05 radians) before exceeding this 90% ROM, to ensure that the actuator will not be switched off during active device use. However, when entering these angle guards, damping is applied to the user, influencing the gathered data in the experiments. Therefore, data samples collected within these angle guards will be removed. In conclusion, a ROM measurement per

subject is performed to find and use the maximum ROM values. During the position-controlled calibration measurements, the user moved within 95% of this ROM. During the performance of the tasks, the user was able to move within 90% of this ROM. Subsequently, angle guards applied damping with large values at 2.9 degrees before reaching maximum and minimum joint angles. In Figure 15, the effect of the end-stop removal is shown for the joint angle ( $\theta$ ), the joint torque based on the force sensor measurements, sEMG biceps activity, and sEMG triceps activity. The red marked lines show the removed samples of the dataset due to the angle guards.

### End-stop Removal in a Dynamic Task at 45 deg/s (subject S98)

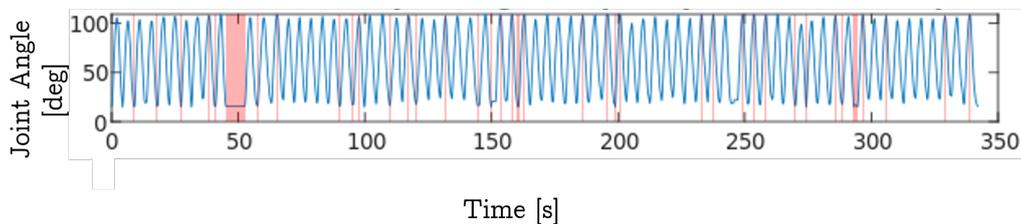


Fig. 15. Effect of end-stop removal on the dynamic task dataset of participant S98, during a velocity of ca. 45 deg/s

### E.3 Data labelling

Data labelling is an essential step towards data classification. Every data sample in time, measured with a frequency of 1000 Hz, is assigned a label. This label represents the lifted weight at that moment. Data labelling was performed manually by selecting the corresponding flexion- and extension cycles for every weight. The result of data labelling for one of the participants during a dynamic task at a velocity of 45 deg/s can be seen in Figure 16.

### Data Labelling of Dynamic Task at 45 deg/s (subject S98)

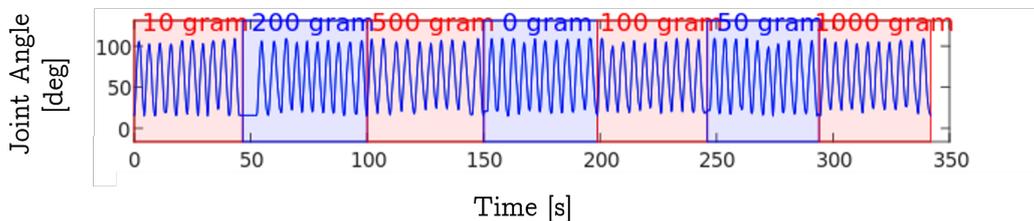
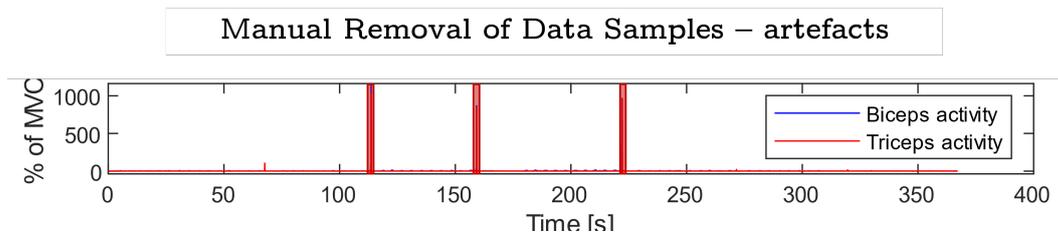


Fig. 16. Data labelling the dynamic task dataset of participant S98, during a velocity of ca. 45 deg/s

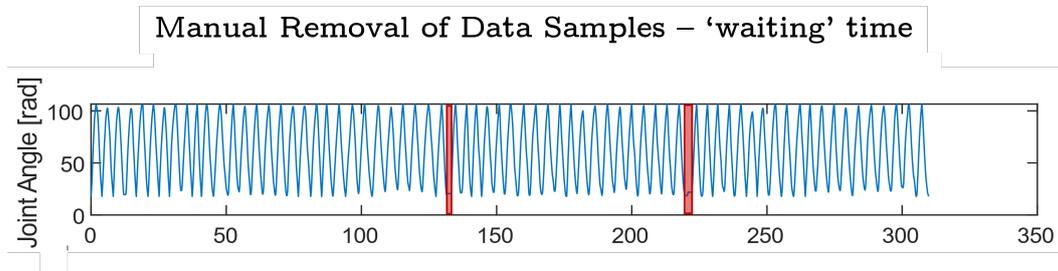
### E.4 Manual removal of data

Although all previous data filtering and data preparation steps were performed, not all redundant information was removed from the dataset. Two types of data were removed from the dataset:

- Artefacts in data due to unexpected events during the experiment, such as dropping weights or sensor interaction. Figure 17 shows an example in which the sEMG electrodes were not properly connected to the skin any more due to interaction with the environment and therefore show values of around 1000% of MVC. To prevent manipulation of the results, these peaks have been removed manually.
- After performing all repetitions while a weight was lifted, there was some 'waiting' time when the weight was switched. This data has also been removed from the dataset. In Figure 18, an example of the removal of this 'waiting' time is shown.



**Fig. 17.** Manual removal of an artefact by disconnecting the electrode from the skin through interaction with the environment.



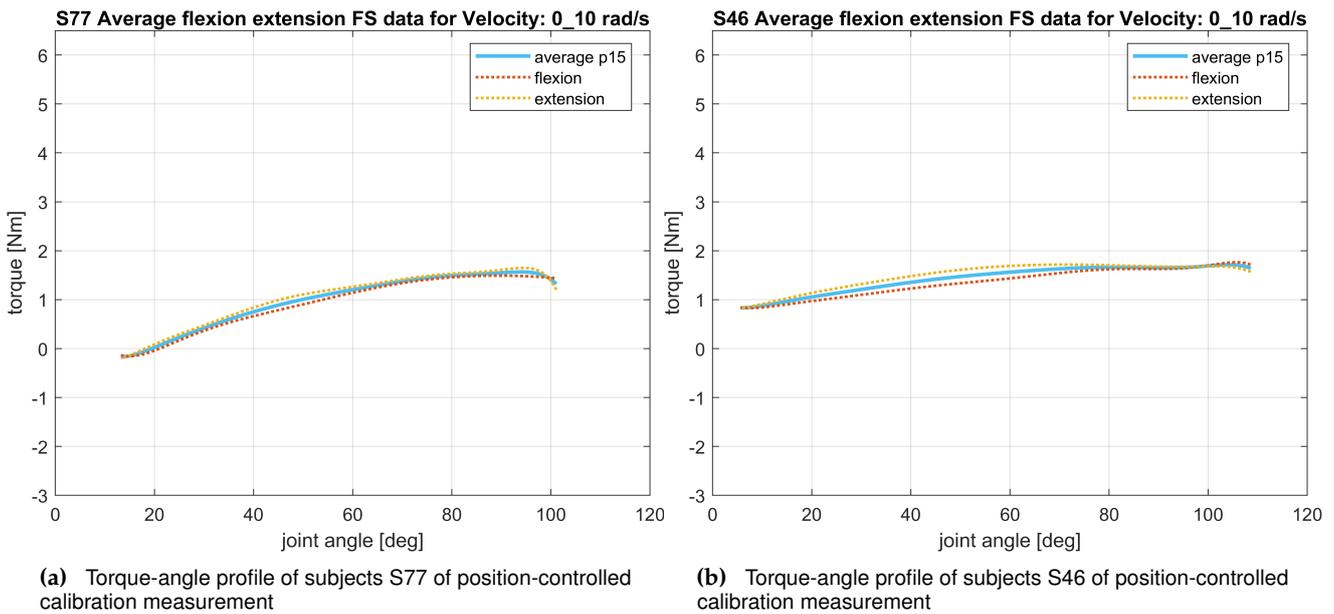
**Fig. 18.** Manual removal of the changing time between different weights in the task

## APPENDIX F PARTICIPANT EXCLUSION

Three participants have been excluded from the data analysis and data classification. Two participants were excluded based on the torque-angle profile, containing the support torque used to compensate for the passive forces acting on the human arm. One participant has been excluded based on joint misalignment.

### F.1 Position-controlled arm weight compensation measurement

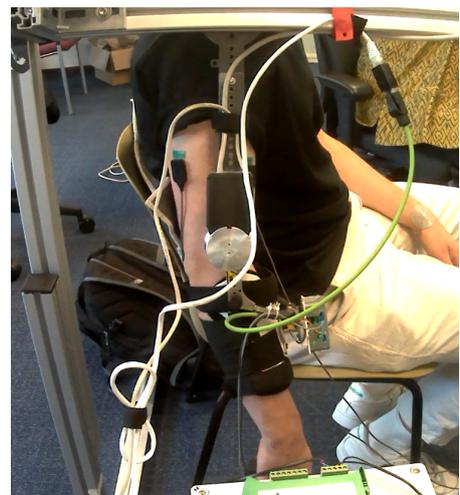
Due to hysteresis in the force sensor, the measurements show a difference in the flexion and extension cycle when measuring the passive forces during the position-controlled measurement [37]. Therefore, hysteresis would result in exceeding values for flexion above extension. However, participants S46 and S77 both show torque-angle profiles with higher torque values for the extension, than the flexion cycle. Hence, it is possible that factors other than passive forces may have had an impact on the force measurements, and the participants were excluded.



**Fig. 19.** Exclusion based on torque-angle profile during position-controlled measurements (torque in extension cycle exceeds torque in flexion cycle)

### F.2 Joint misalignment

To align the participant with the active-assistive device, the height of the extrusion profiles and the length of the upper arm could be adjusted. Furthermore, the participant was placed within an arm sleeve that was subsequently attached to the device, and an elastic band was placed around the upper arm connecting the participant to the device. The elastic band was forgotten during the experiments with one participant, and significant misalignments were shown in the experiment videos. One shot of the experiment is shown in Figure 20 in which the misalignment is shown. It is evident that the elbow joint is not aligned with the SEA and that the upper arm is positioned further back than the device.



**Fig. 20.** Significant joint misalignment in participant S17 due to the absence of the elastic upper band, which ensures stabilization and better alignment of the upper arm and elbow joint with the device

## APPENDIX G DATA CLASSIFICATION

This appendix provides all background information and extensive results of the data classification. Only the data of the dynamic tasks is considered in the data classification.

### G.1 Data Visualization

#### G.1.1 Class Distribution

In Figure 21 the distribution of the classes, the used weights in the human experiment, is shown. With a minimum of 521380 samples in the 0-gram class (13.6% of complete dataset), and a maximum of 572279 samples in the 1000-gram class (14.9% of complete dataset).

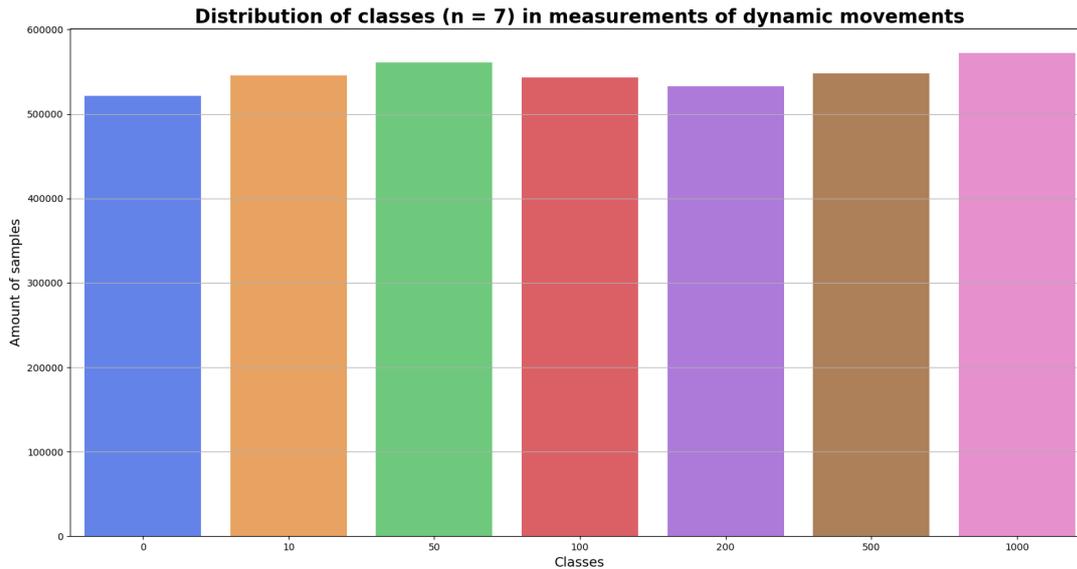


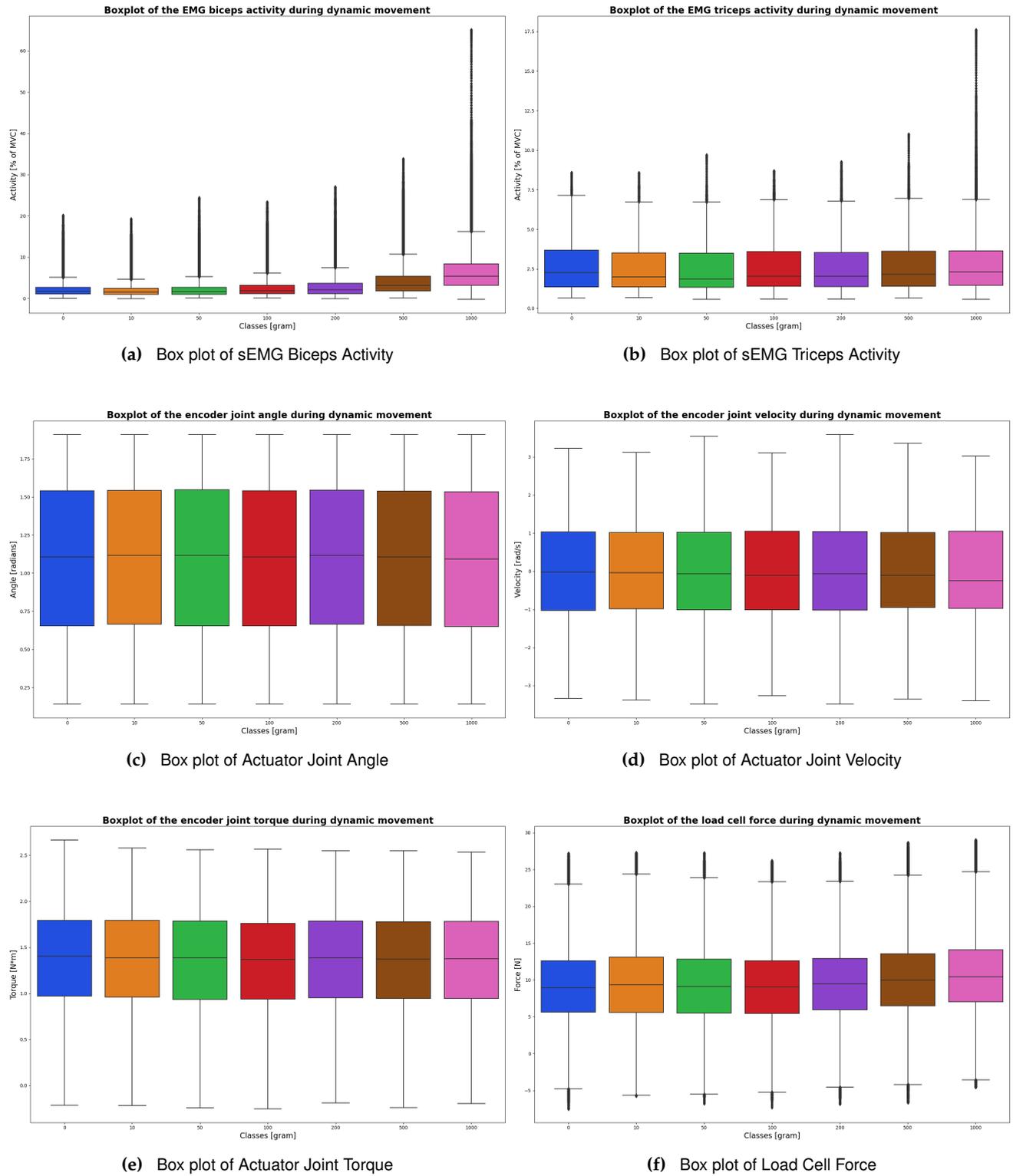
Fig. 21. Class distribution in dynamic movements' dataset, with the classes representing the lifted weights

#### G.1.2 Descriptive statistics

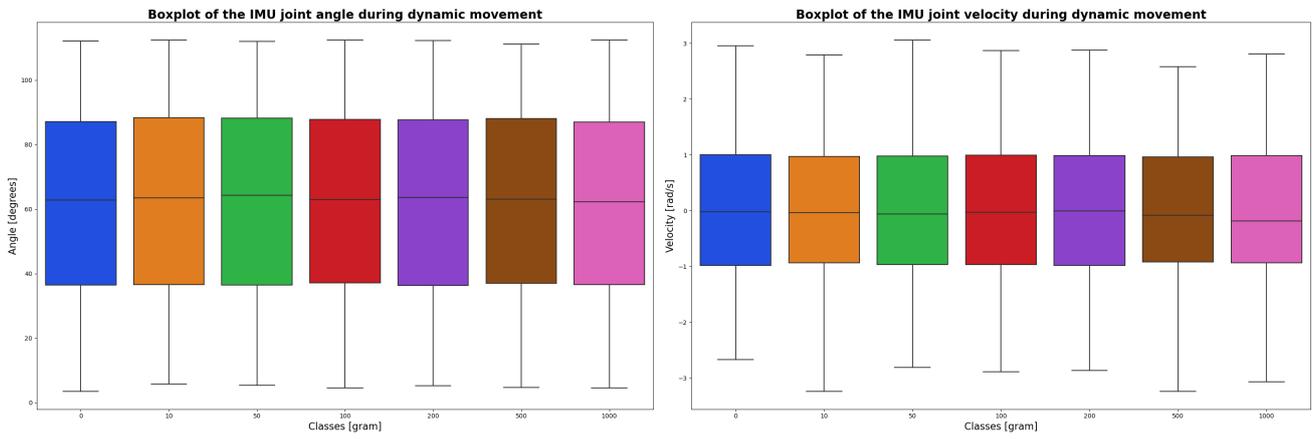
TABLE 9. Descriptive statistics of the dynamic movement data set (total samples = 3823550)

Feature	Mean	Standard Deviation	Minimum value	Percentile (25%)	Percentile (50%)	Percentile (75%)	Maximum value
Encoder Joint Angle [rad]	1.09	0.48	0.14	0.66	1.11	1.54	1.91
Encoder Joint Velocity [rad/s]	-0.004	1.26	-3.49	-0.99	-0.09	1.03	3.59
SEA Joint Torque [N*m]	1.34	0.58	-0.26	0.95	1.38	1.78	2.67
Load Cell Force [N]	9.81	5.80	-7.52	5.92	9.48	13.13	29.01
sEMG activity biceps [% of MVC]	3.25	3.19	-0.26	1.24	2.14	4.07	65.11
sEMG activity triceps [% of MVC]	2.49	1.29	0.56	1.37	2.09	3.58	17.61
IMU Joint Angle [rad]	61.93	28.25	3.42	36.70	63.19	87.78	112.46
IMU Joint Velocity [rad/s]	-0.0012	1.16	-3.24	-0.96	-0.06	0.99	3.06
IMU Joint Acceleration [m/s <sup>2</sup> ]	7.25	1.93	-2.14	5.84	7.85	8.91	11.81

G.1.3 Box plot per feature

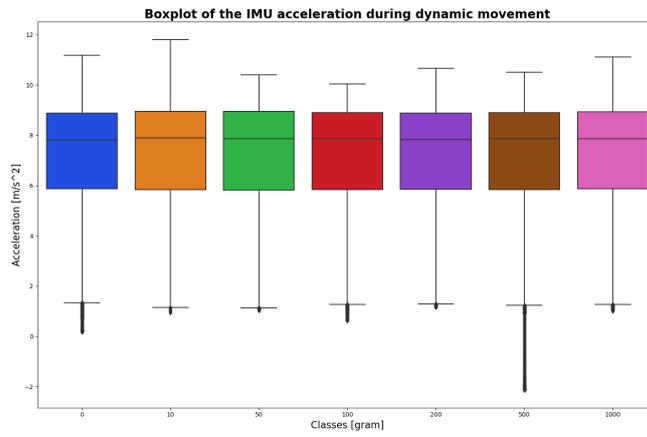


**Fig. 22.** Box plots of set of features (n = 6) showing distribution over classes - part 1



(a) Box plot of IMU Joint Angle

(b) Box plot of IMU Joint Velocity



(c) Box plot of IMU Joint Acceleration

Fig. 23. Box plots of set of features (n = 3) showing distribution over classes - part 2

### G.1.4 Correlation heatmap

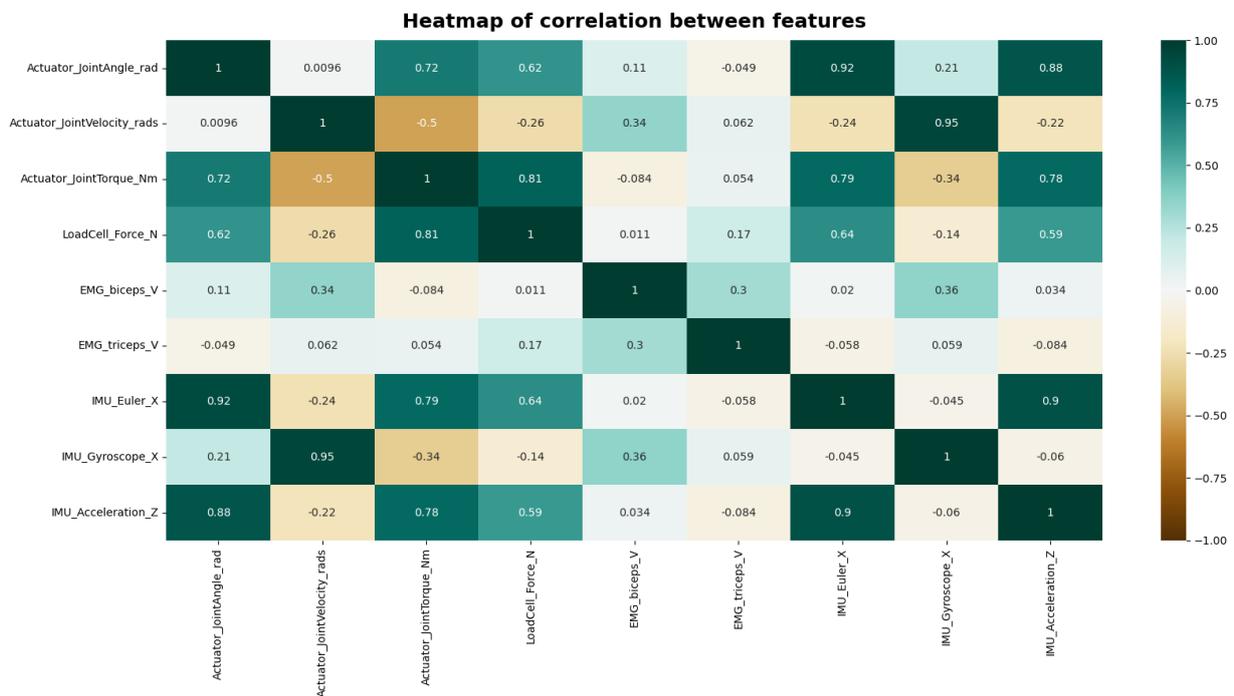
#### G.1.4.1 Correlation between features

In Figure 24 the correlation between all features is shown in the dynamic movements' dataset. In the book of Swinscow et al. [69] the following strengths of the correlation coefficient have been described:

- 0 - 0.19: very weak
- 0.2 - 0.39: weak
- 0.4 - 0.59: moderate
- 0.6 - 0.79: strong
- 0.8 - 1.0: very strong

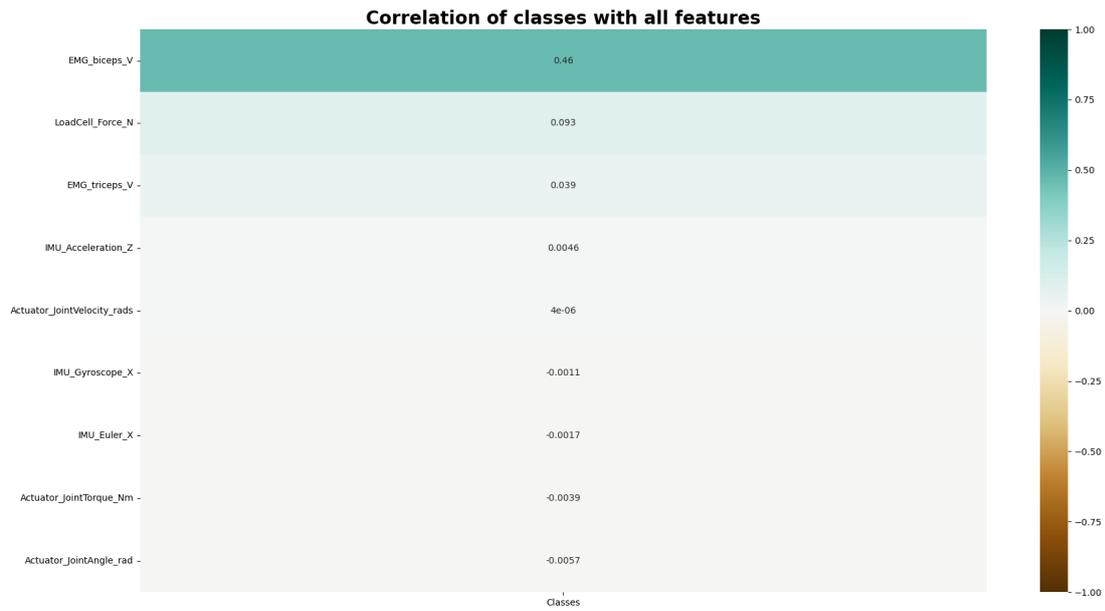
Therefore, the following features have a very strong correlation:

- Actuator Joint Velocity & IMU Joint Velocity - very strong (0.95)
- Actuator Joint Angle & IMU Joint Angle - very strong (0.92)
- IMU Joint Angle & IMU Joint Acceleration - very strong (0.9)
- Actuator Joint Angle & IMU Joint Acceleration - very strong (0.88)
- Actuator Joint Torque & Load Cell Force - very strong (0.81)



**Fig. 24.** Heatmap of the correlation coefficients between features in the dynamic movement dataset. The numbers within the heatmap represent the Pearson correlation coefficients between the corresponding features.

### G.1.4.2 Correlation between features and classes



**Fig. 25.** Heatmap of the correlation coefficient between features and the classes in the dynamic movement dataset. The numbers within the heatmap represent the Pearson correlation coefficients between the features and the classes.

## G.2 Feature selection - backward elimination

Backward elimination was used as a feature selection method. Within the backward elimination, one feature is constantly removed from the dataset until one feature is left. This is done using cross-validation accuracies for every subset of features, checking for the highest value. The mean of the cross-validated accuracies for the best-performing feature subset is shown in Table 10 and Table 11. In addition, the feature excluded based on backward elimination in every step is shown.

### G.2.1 KNN Classification

**TABLE 10.** Backward elimination feature selection for the KNN classification algorithm. The mean cross-validation accuracy score, including standard deviation, is shown for the best-performing feature subset. The excluded feature based on the highest performance is displayed.

Amount of features	Mean cross-validation accuracy (standard deviation)	Excluded feature in feature subset
9 features	21.46% (2.96%)	-
8 features	21.74% (2.51%)	IMU Joint Angle
7 features	21.94% (2.68%)	IMU Acceleration
6 features	21.97% (2.65%)	IMU Joint Velocity
5 features	21.60% (2.51%)	Elbow Joint Torque
4 features	20.76% (1.14%)	Elbow Joint Angle (encoder)
3 features	19.81% (0.98%)	Load Cell Force
2 features	19.68% (1.49%)	sEMG Triceps Activity
1 feature	18.94% (1.29%)	Elbow Joint Velocity (encoder)

The backward elimination resulted in a subset of six features having the highest mean cross-validation accuracy. This feature subset consists of the sEMG biceps activity, the elbow joint velocity by the joint encoder ( $v$ ), the sEMG triceps activity, the load cell force ( $F_z$ ), the elbow joint angle by the encoder ( $\theta$ ), and the elbow joint torque by the encoder ( $T$ ).

### G.2.2 SVM Classification

**TABLE 11.** Backward elimination feature selection for the SVM classification algorithm. The mean cross-validation accuracy score, including standard deviation, is shown for the best-performing feature subset. The excluded feature based on the highest performance is displayed

Amount of features	Mean cross-validation accuracy (standard deviation)	Excluded feature in feature subset
9 features	22.11% (2.70%)	-
8 features	22.47% (2.49%)	sEMG Triceps Activity
7 features	22.59% (1.06%)	Elbow Joint Torque
6 features	22.60% (1.22%)	IMU Joint Velocity
5 features	22.76% (1.47%)	Load Cell Force
4 features	23.39% (1.44%)	Elbow Joint Velocity (encoder)
3 features	23.52% (1.39%)	IMU joint angle
2 features	23.36% (1.24%)	IMU Acceleration
1 feature	23.50% (1.15%)	Elbow Joint Angle (encoder)

The backward elimination resulted in a subset of three features having the highest mean cross-validation accuracy. This feature subset consists of the sEMG biceps activity, the elbow joint angle by the encoder ( $\theta$ ), and the IMU acceleration in the z-axis ( $a_z$ ).

### G.3 Grid Search

Grid Search is a method to tune the hyperparameters by checking the mean cross-validation accuracy of all combinations of hyperparameters. This is for both KNN and SVM classification performed on a specific subset of hyperparameters.

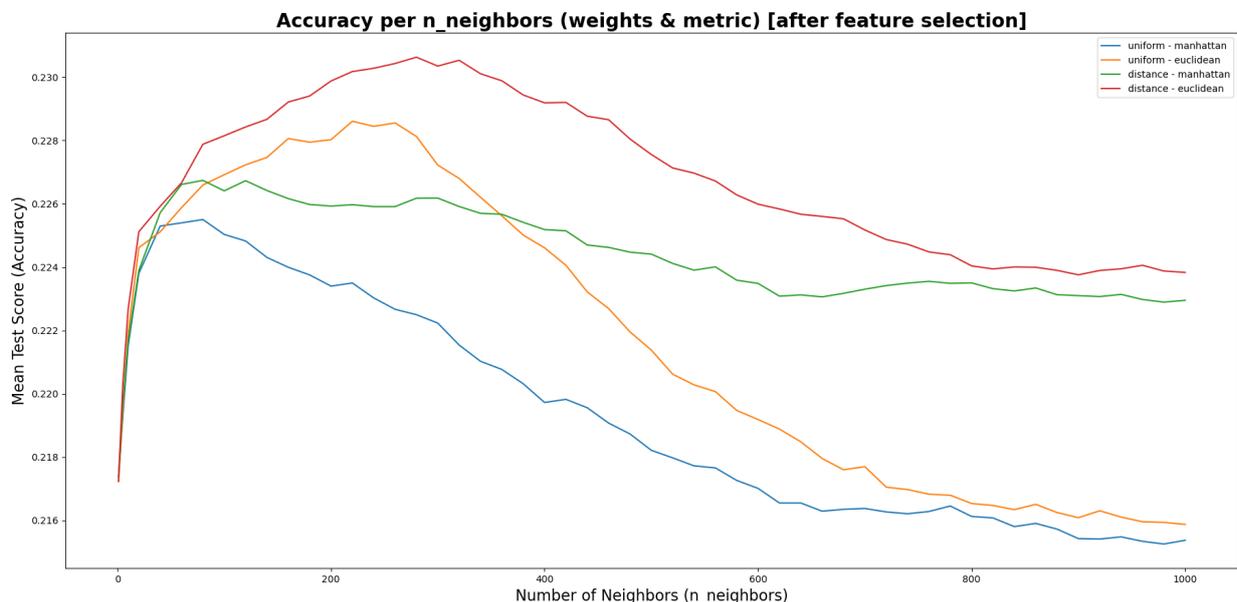
#### G.3.1 KNN Classification

The following hyperparameters were evaluated in the grid search:

- **n\_neighbors** - {1, 5, 10} and 20 – 1000 in steps of 20
- **weights** - 'uniform', 'distance'
- **metric** - 'Euclidean', 'Manhattan'

Figure 27 shows the output of the grid search resulting in the highest mean cross-validation accuracy for the following combination of hyperparameters:

- **n\_neighbors** - 280
- **weights** - 'distance'
- **metric** - 'Euclidean'



**Fig. 26.** Performed grid search in the knn classification for the hyperparameters neighbours, weights, and metric. The different graphs represent how a combination of weights and metrics changes over increasing neighbors.

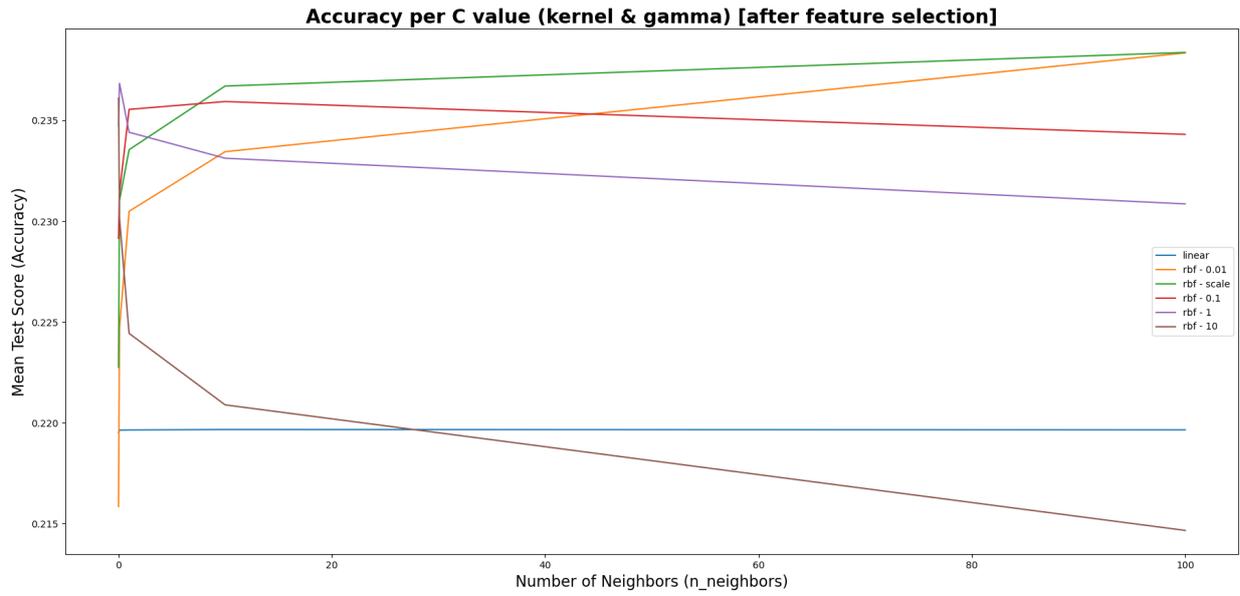
#### G.3.2 SVM Classification

The following hyperparameters were evaluated in the grid search:

- **C** - 0.01, 0.1, 1, 10, 100
- **gamma** - 0.01,  $(1/(n\_features * variance))$ , 0.1, 1, 10
- **kernel** - 'linear', 'rbf'

Gamma does not affect the linear kernel because this hyperparameter represents the curvature of a decision boundary. Therefore, it only has influence on the 'rbf' kernel. Figure 27 shows the output of the grid search resulting in the highest mean cross-validation accuracy for the following combination of hyperparameters:

- **C** - 100
- **gamma** - 'scale'
- **kernel** - 'rbf'



**Fig. 27.** Performed grid search in the SVM classification for the hyperparameters C, gamma, and the kernel. The different graphs represent how a combination of the kernel and gamma changes over an increasing value for C.

### G.4 Evaluation

Most of the performance metrics, evaluating both classifiers, are shown in the Results section. However, additional results are shown in this section. The additional results in this section are the mean and standard deviation of the cross-validation accuracies and the confusion matrices for all classifiers.

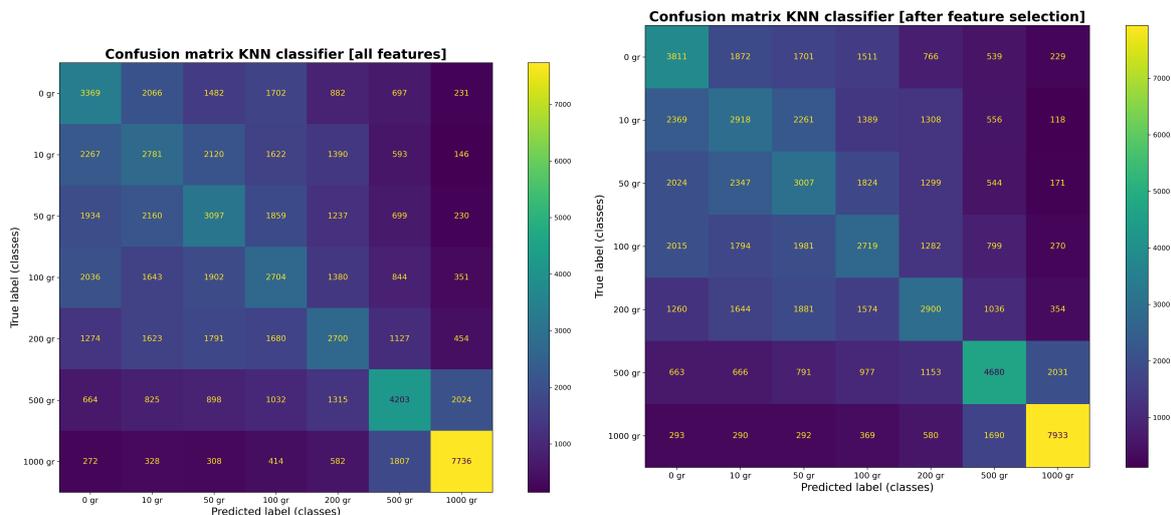
#### G.4.1 KNN classifier

##### G.4.1.1 k-fold cross-validation (k=5)

**TABLE 12.** Evaluation of KNN classifiers (accuracy = fraction of total predictions that are correct, precision = fraction of a predicted label that was identified correctly, recall = sensitivity, fraction of the true label that was identified correctly, F1-score = a combination metric of precision and recall, emphasizing the balance between the two,  $\mu$  = mean cross-validation accuracy, SD = standard deviation of the cross-validation accuracies)

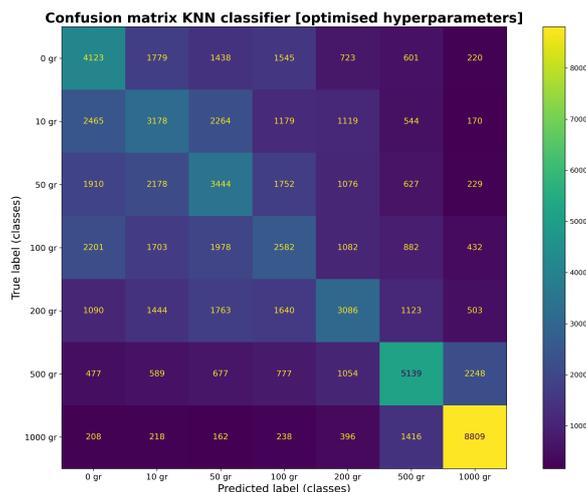
Classifier	Accuracy	Precision	Recall	F1-score	Time fit	Time predict	$\mu$ (SD)
Default KNN Classifier (all features)	34.77%	34.63%	34.85%	34.51%	1.37 s	14.73 s	21.46% (2.41%)
Default KNN Classifier (after feature selection)	36.57%	36.48%	36.80%	36.33%	0.50 s	4.96 s	21.94% (2.40%)
KNN Classifier with the optimised hyperparameters	39.70%	39.09%	39.06%	39.42%	0.37 s	77.97 s	23.06% (1.88%)

##### G.4.1.2 Confusion Matrix



(a) Confusion matrix KNN classifier with default settings, including all features

(b) Confusion matrix KNN classifier with default settings, including the features selected by feature selection in the feature set



(c) Confusion matrix KNN classifier with optimised hyperparameter setting, including the features selected by feature selection in the feature set

**Fig. 28.** Confusion matrices KNN classifier (The diagonal values represent the number of correct classifications for each class. All other values indicate the number of misclassifications.)

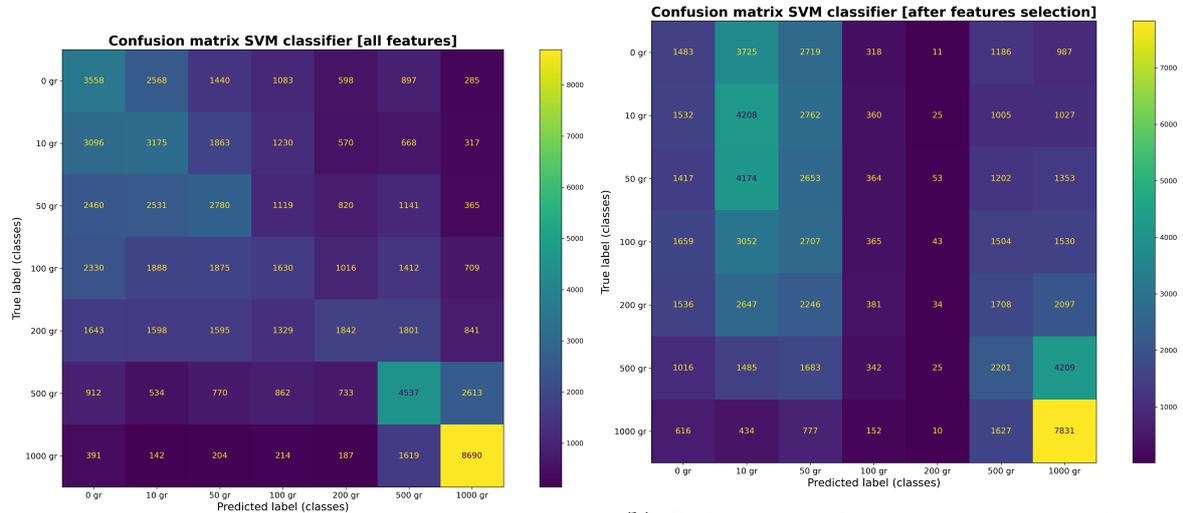
### G.4.2 SVM classifier

#### G.4.2.1 k-fold cross-validation (k=5)

**TABLE 13.** Evaluation of SVM classifiers (accuracy = fraction of total predictions that are correct, precision = fraction of a predicted label that was identified correctly, recall = sensitivity, fraction of the true label that was identified correctly, F1-score = a combination metric of precision and recall, emphasizing the balance between the two,  $\mu$  = mean cross-validation accuracy, SD = standard deviation of the cross-validation accuracies)

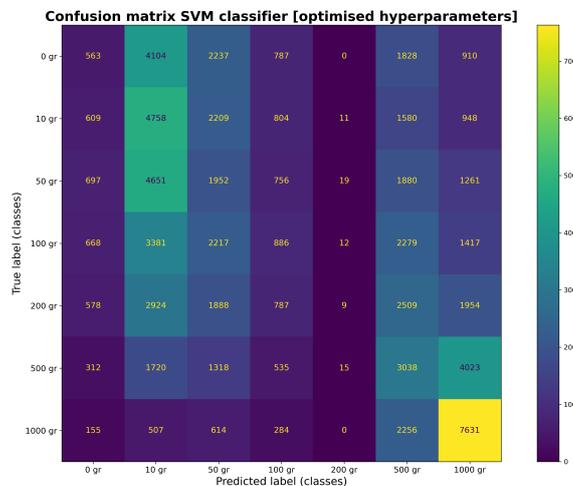
Classifier	Accuracy	Precision	Recall	F1-score	Time fit	Time predict	$\mu$ (SD)
Default SVM Classifier (all features)	34.27%	32.83%	32.98%	33.94%	4927.41 s	2042.22 s	22.11% (2.7%)
Default SVM Classifier (after feature selection)	24.72%	20.46%	19.96%	24.22%	5697.67 s	2138.82 s	24.55% (1.39%)
SVM Classifier with the optimised hyperparameters (after feature selection)	24.63%	19.93%	20.97%	24.14%	10794.78 s	2481.30 s	23.84% (1.41%)

#### G.4.2.2 Confusion Matrix



(a) Confusion matrix SVM classifier with default settings, including all features

(b) Confusion matrix SVM classifier with default settings, including the features selected by feature selection in the feature set



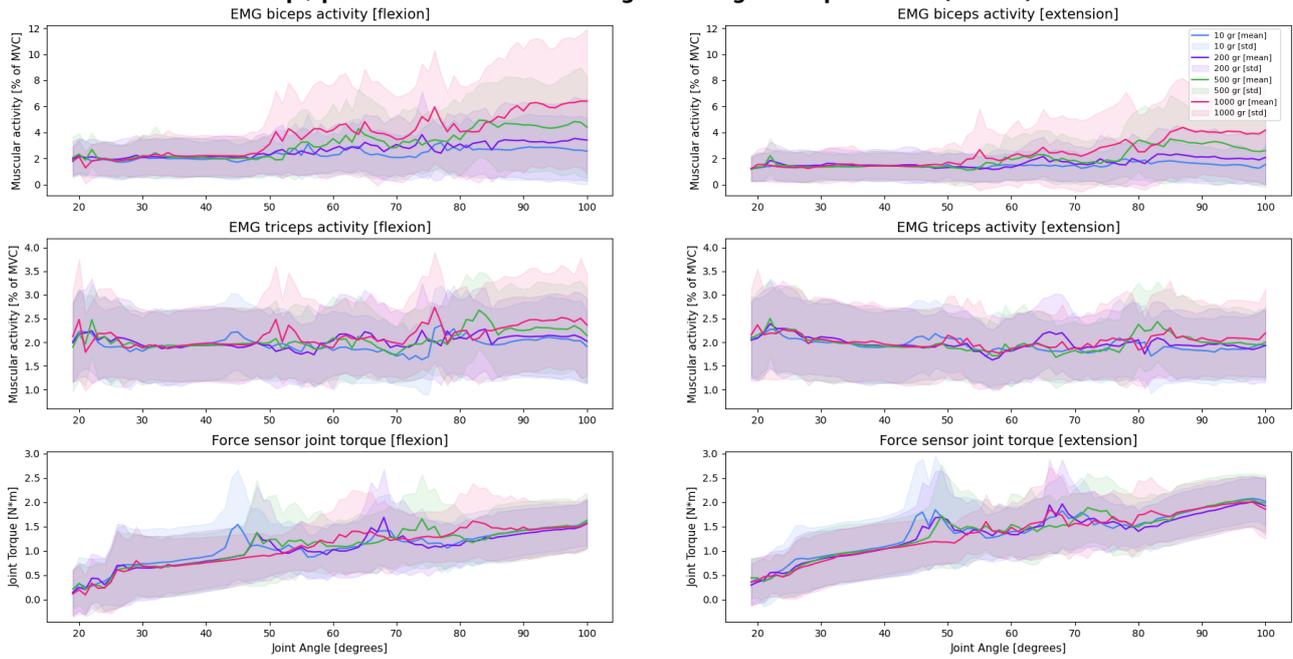
(c) Confusion matrix SVM classifier with optimised hyperparameter setting, including the features selected by feature selection in the feature set

**Fig. 29.** Confusion matrices SVM classifier (The diagonal values represent the number of correct classifications for each class. All other values indicate the number of misclassifications.)

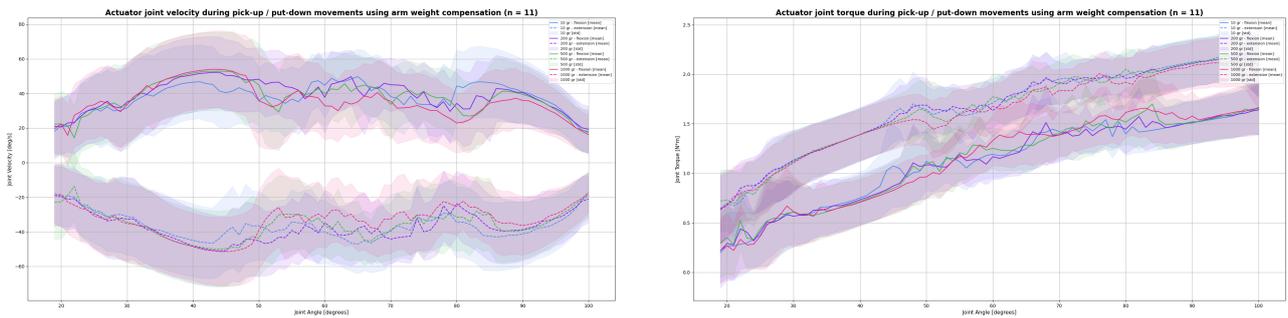


## H.2 Pick/place task, combining the pickup angle of 45 degrees and 90 degrees for all subjects (n = 11)

### Pick-up / put-down movements using arm weight compensation (n = 11)



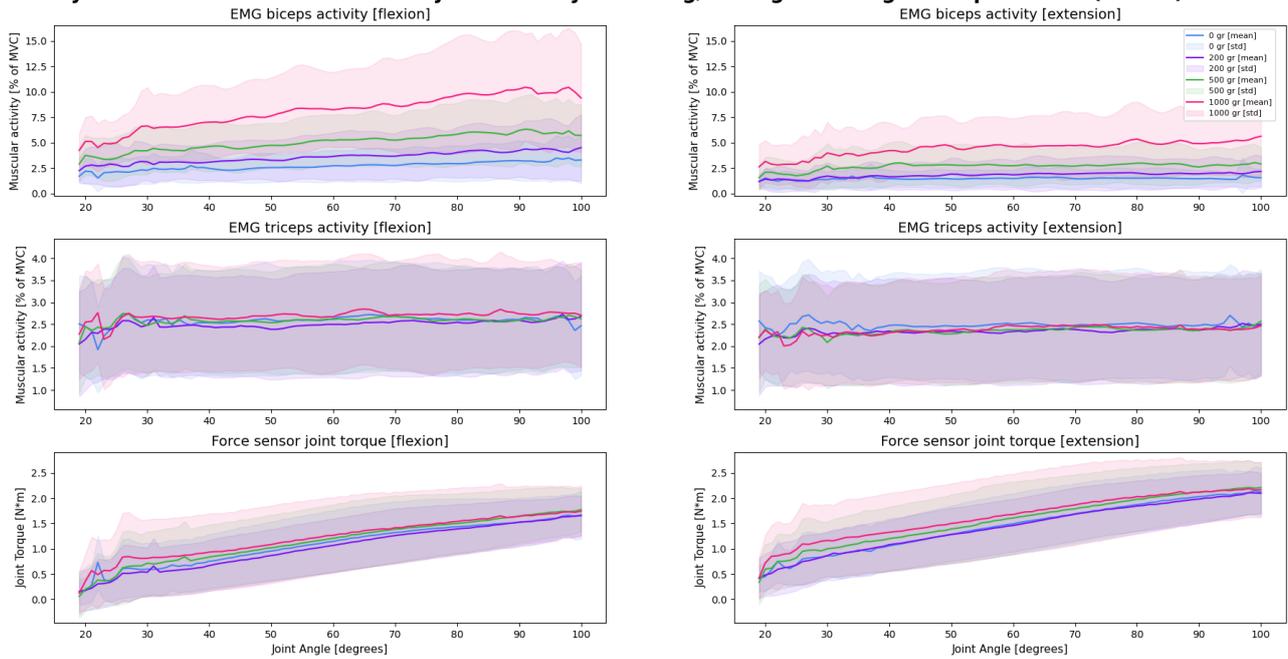
**Fig. 32.** Features during pick/place task [joint angle - 45 degrees and 90 degrees]



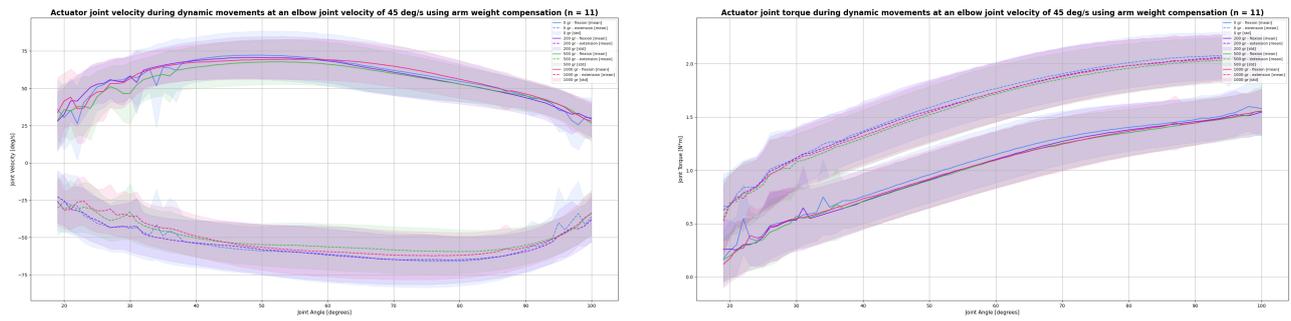
**(a)** Actuator Joint Velocity **(b)** Actuator Joint Torque  
**Fig. 33.** Actuator joint velocity, and joint torque during pick/place task [joint angle - 45 degrees and 90 degrees]

### H.3 Dynamic task, at the joint velocity of 45 deg/s for all subjects (n = 11)

#### Dynamic movements at an elbow joint velocity of 45 deg/s using arm weight compensation (n = 11)



**Fig. 34.** Features during dynamic task [joint velocity - 45 deg/s]



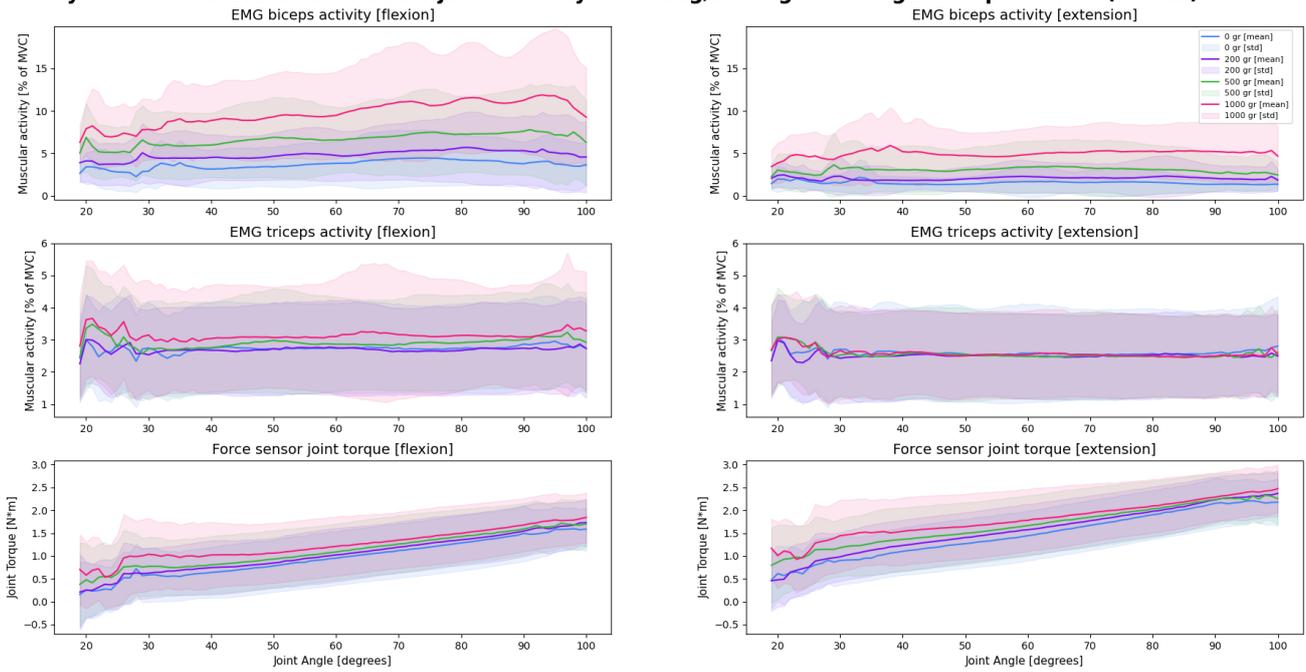
**(a)** Actuator Joint Velocity

**(b)** Actuator Joint Torque

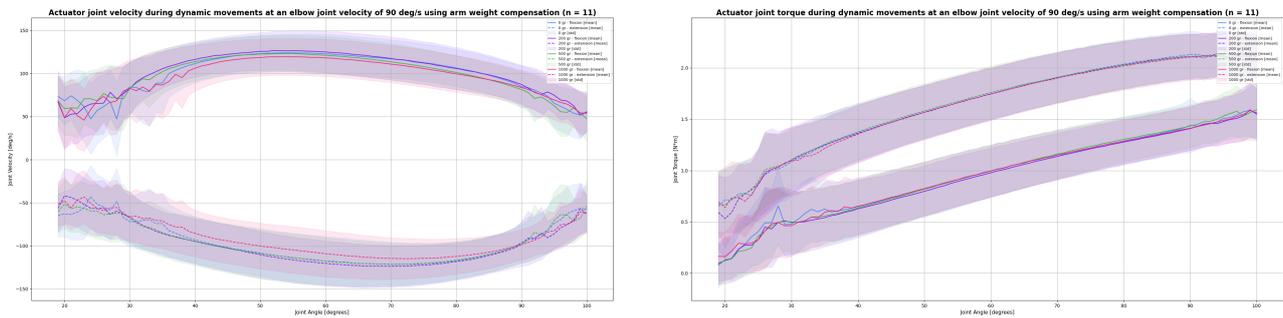
**Fig. 35.** Actuator joint velocity, and joint torque during dynamic task [joint velocity - 45 deg/s]

### H.4 Dynamic task, at the joint velocity of 90 deg/s for all subjects (n = 11)

#### Dynamic movements at an elbow joint velocity of 90 deg/s using arm weight compensation (n = 11)



**Fig. 36.** Features during dynamic task [joint velocity - 90 deg/s]



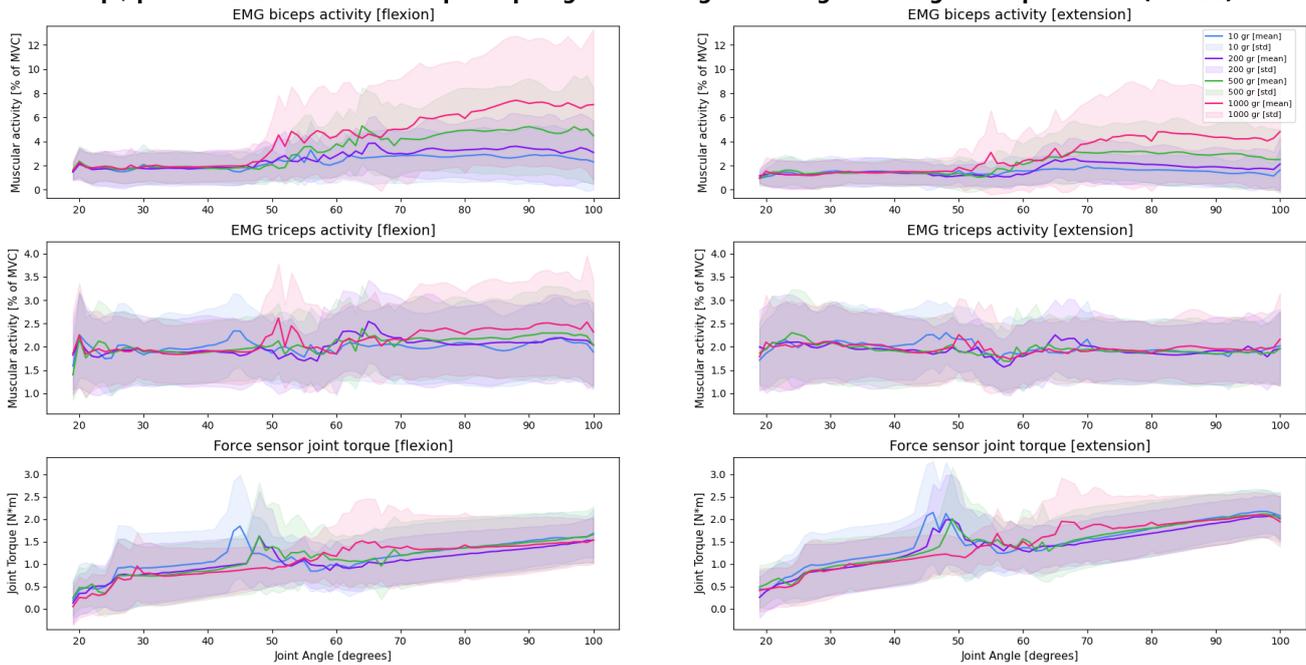
**(a)** Actuator Joint Velocity

**(b)** Actuator Joint Torque

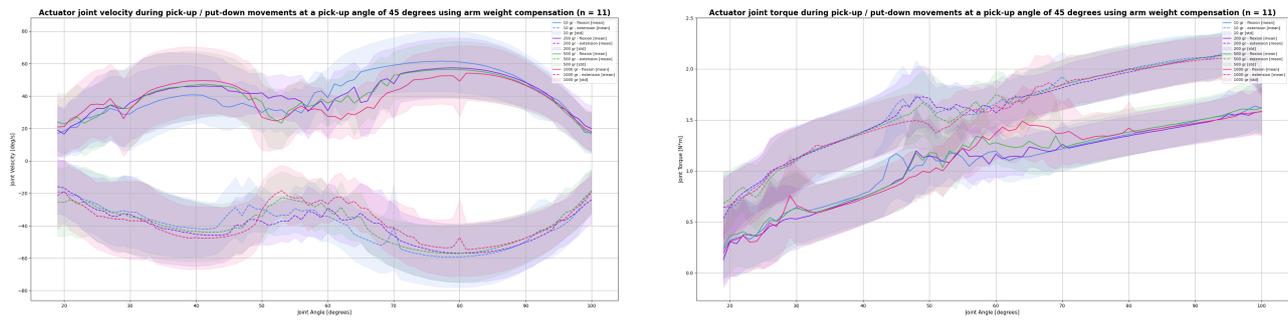
**Fig. 37.** Actuator joint velocity, and joint torque during dynamic task [joint velocity - 90 deg/s]

### H.5 Pick/place task, with a pickup angle of 45 degrees for all subjects (n = 11)

#### Pick-up / put-down movements at a pick-up angle of 45 degrees using arm weight compensation (n = 11)



**Fig. 38.** Features during pick/place task [joint angle - 45 degrees]



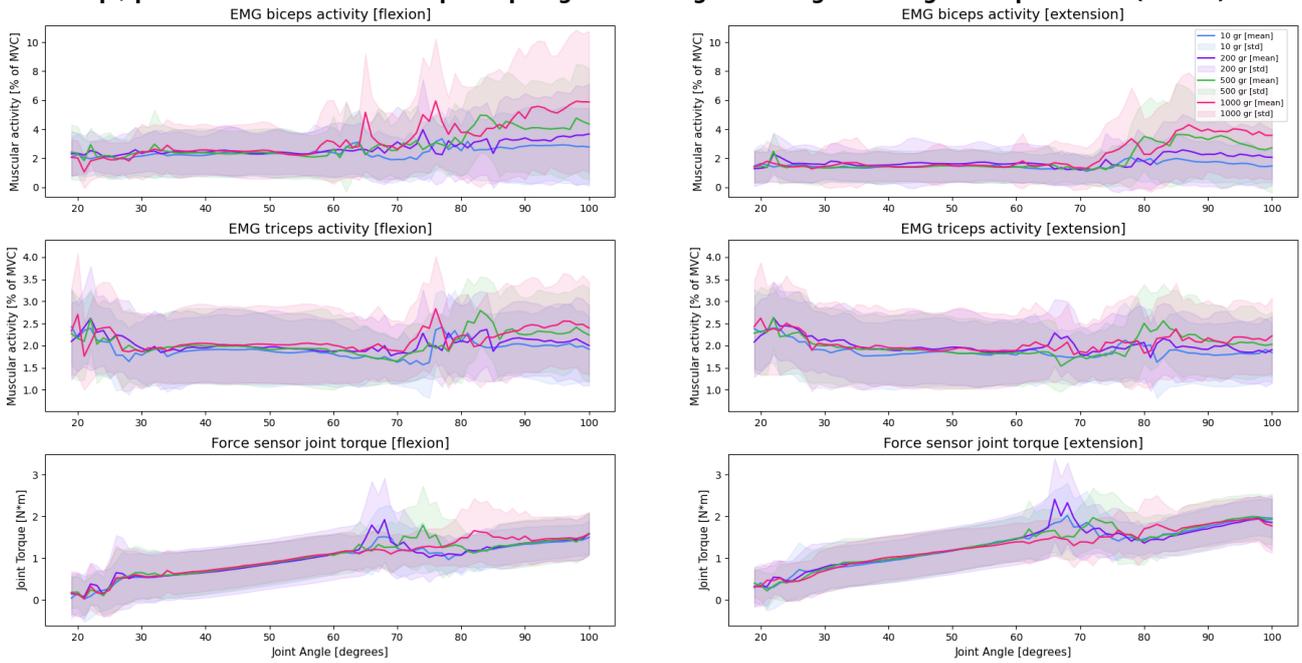
**(a)** Actuator Joint Velocity

**(b)** Actuator Joint Torque

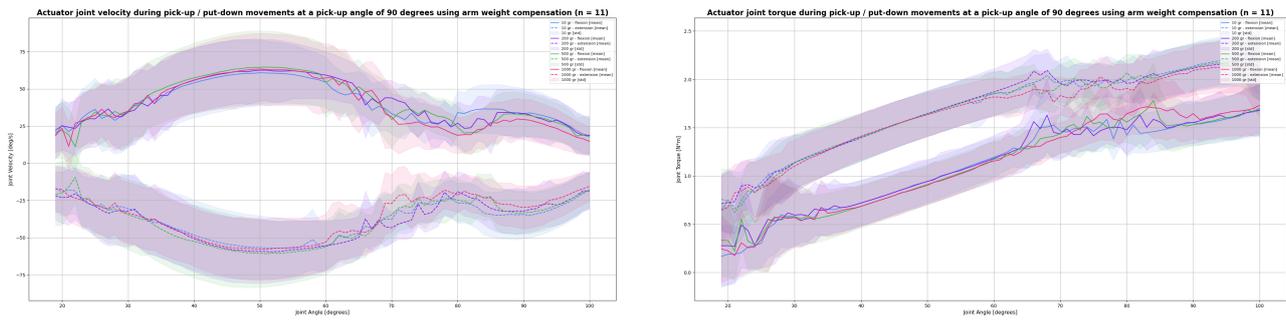
**Fig. 39.** Actuator joint velocity, and joint torque during pick/place task [joint angle - 45 degrees]

### H.6 Pick/place task, with a pickup angle of 90 degrees for all subjects (n = 11)

#### Pick-up / put-down movements at a pick-up angle of 90 degrees using arm weight compensation (n = 11)



**Fig. 40.** Features during pick/place task [joint angle - 90 degrees]



(a) Actuator Joint Velocity

(b) Actuator Joint Torque

**Fig. 41.** Actuator joint velocity, and joint torque during pick/place task [joint angle - 90 degrees]

### H.7 Overview measurements - subject S04

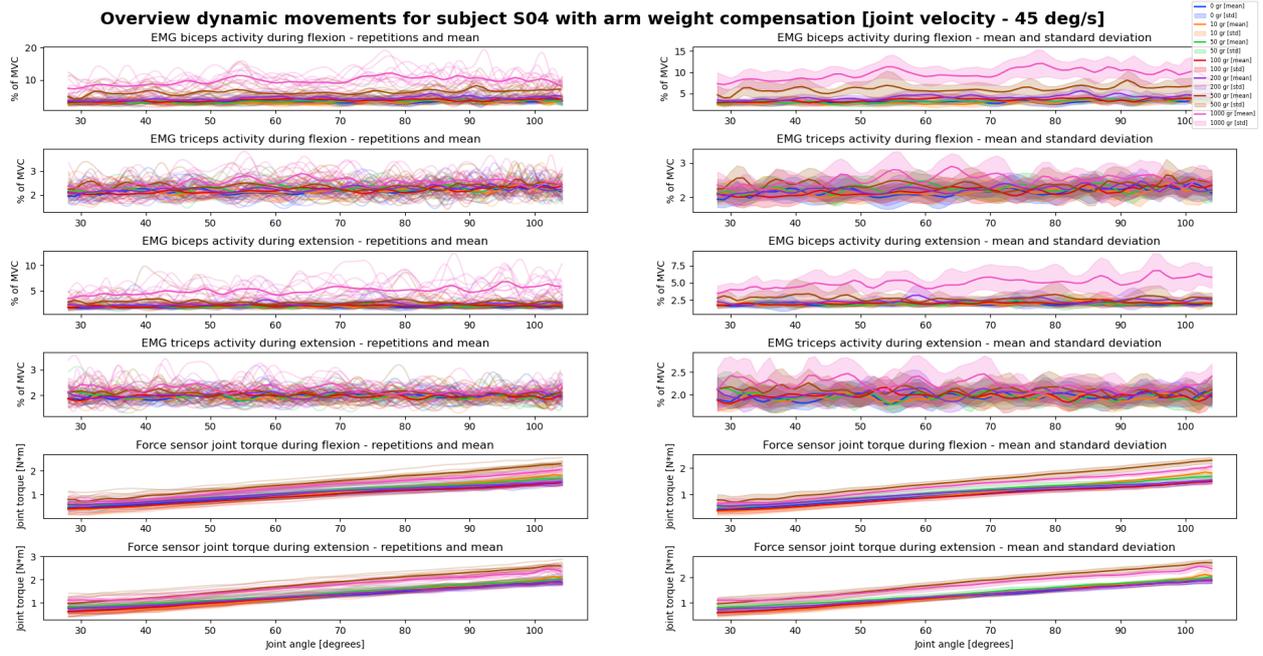


Fig. 42. Features during dynamic task [joint velocity - 45 deg/s]

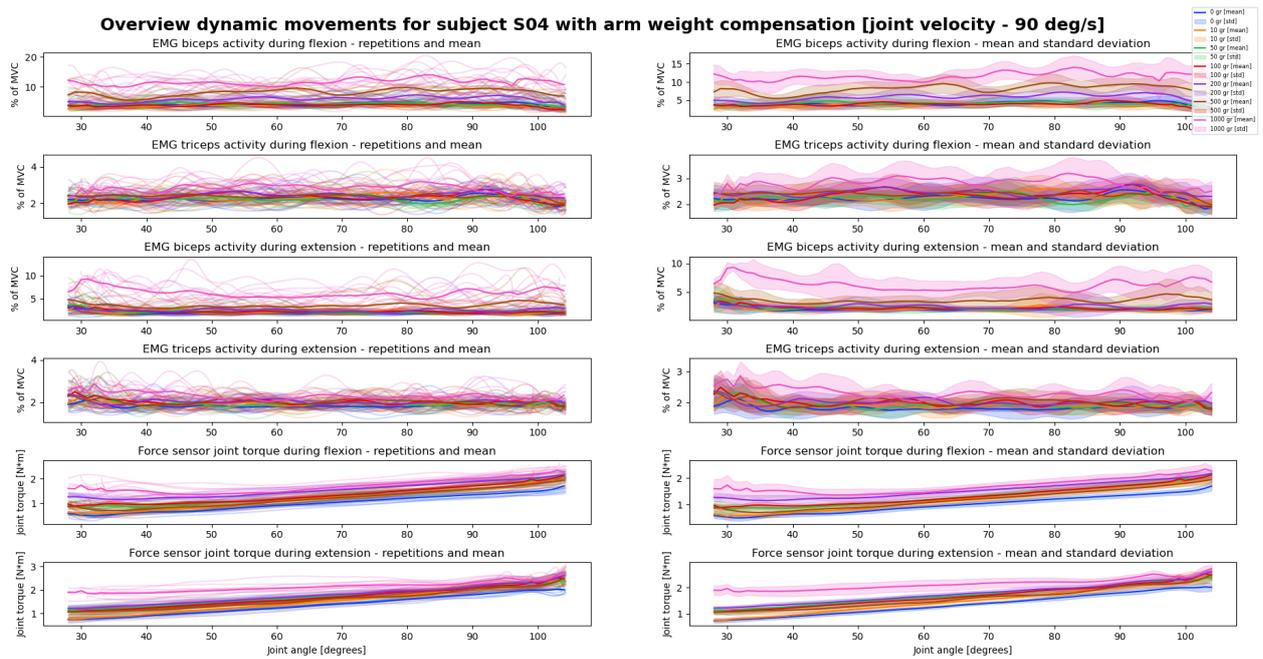


Fig. 43. Features during dynamic task [joint velocity - 90 deg/s]

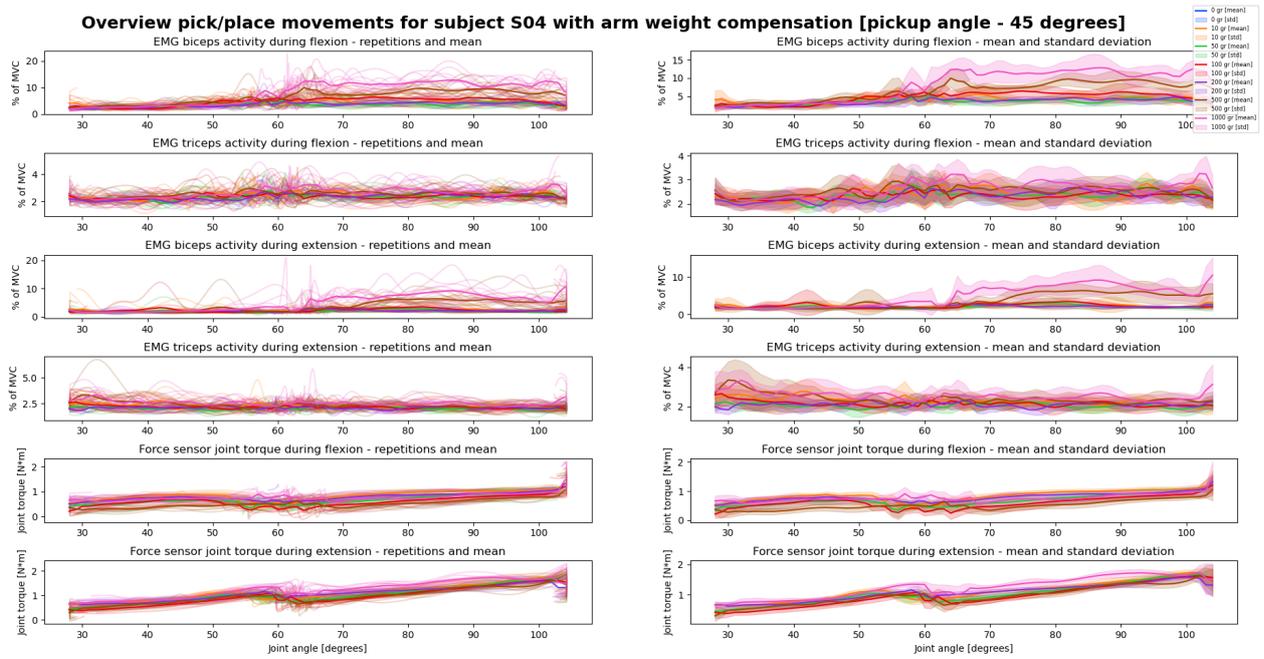


Fig. 44. Features during pick/place task [joint angle - 45 degrees]

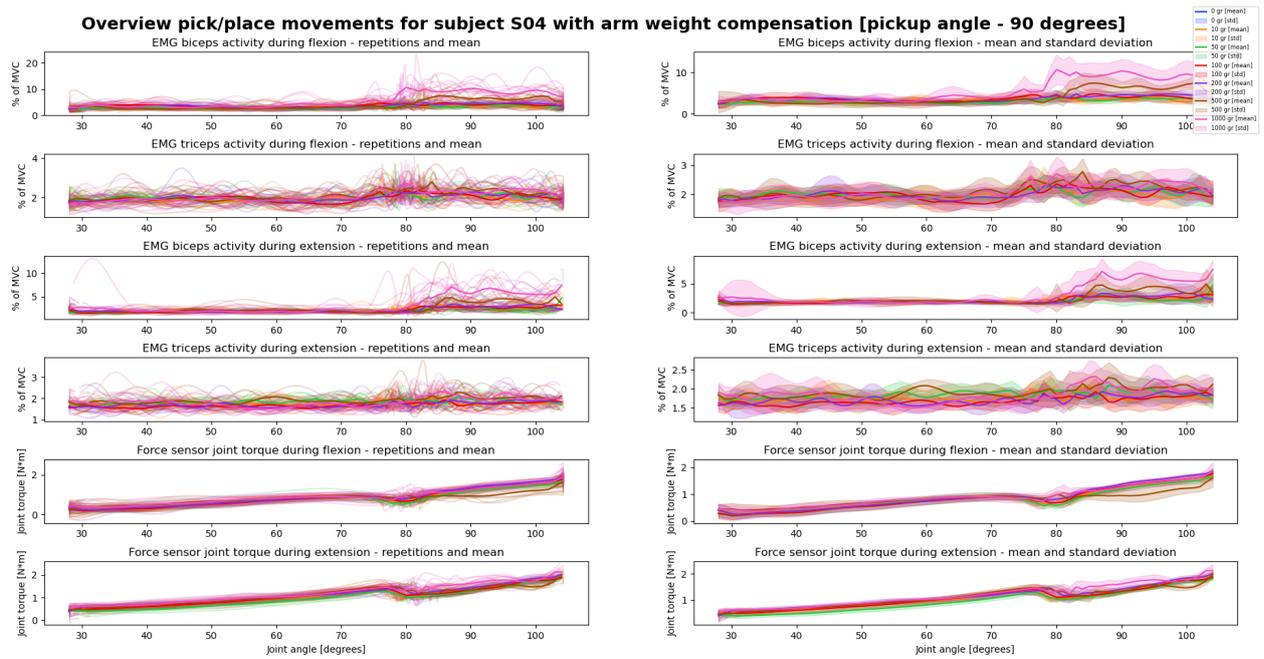


Fig. 45. Features during pick/place task [joint angle - 90 degrees]

### H.8 Overview measurements - subject S06

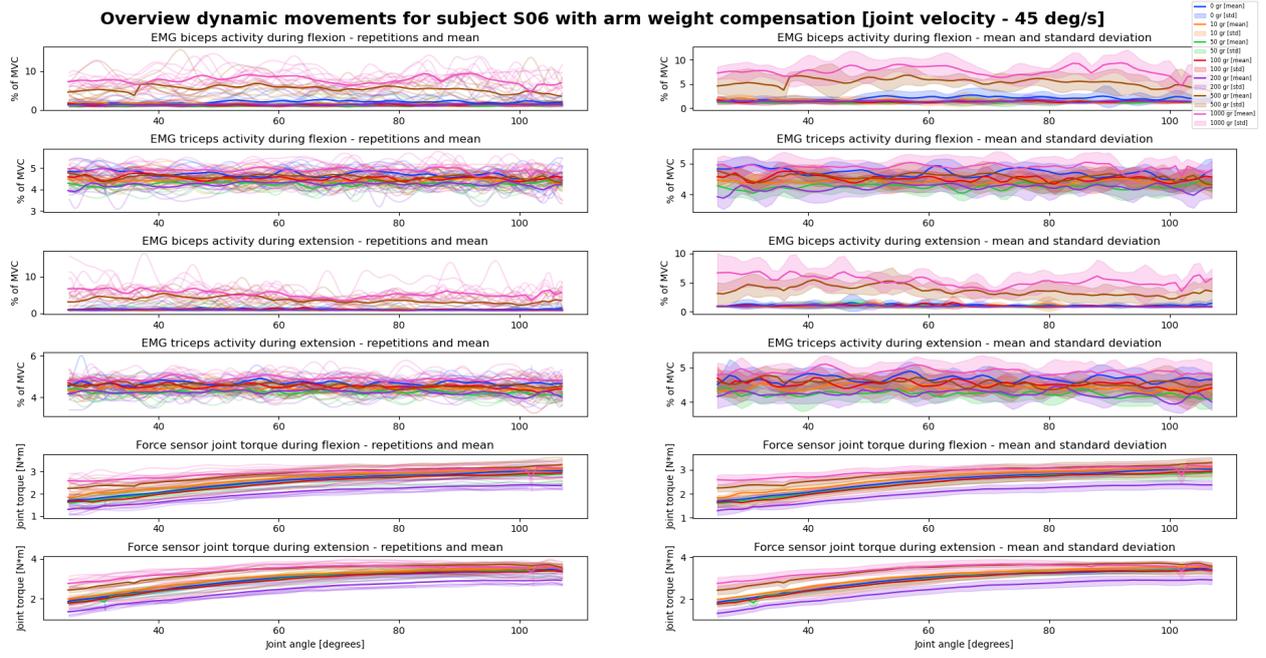


Fig. 46. Features during dynamic task [joint velocity - 45 deg/s]

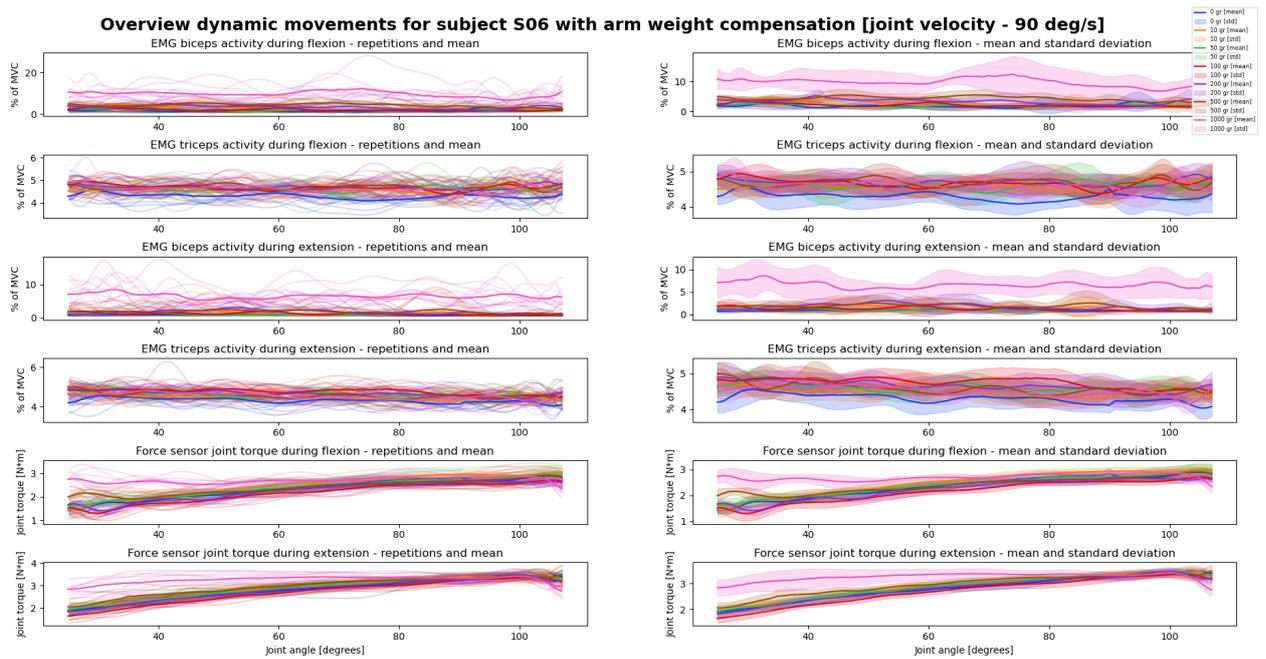


Fig. 47. Features during dynamic task [joint velocity - 90 deg/s]

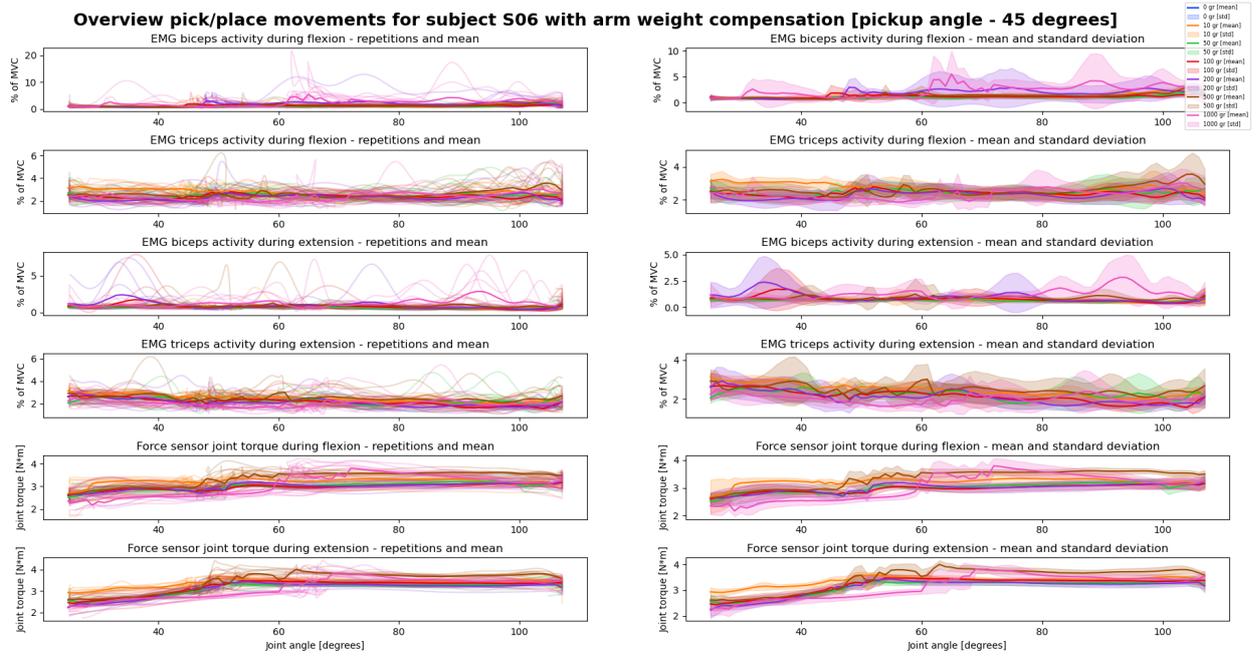


Fig. 48. Features during pick/place task [joint angle - 45 degrees]

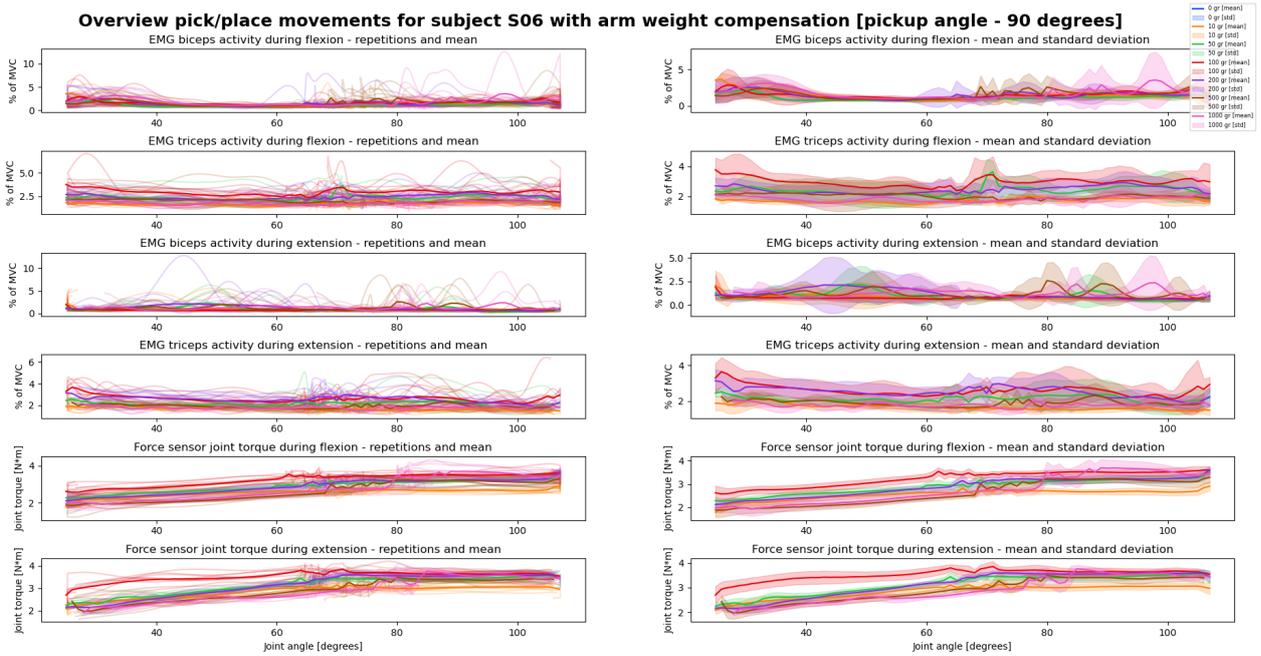


Fig. 49. Features during pick/place task [joint angle - 90 degrees]

### H.9 Overview measurements - subject S26

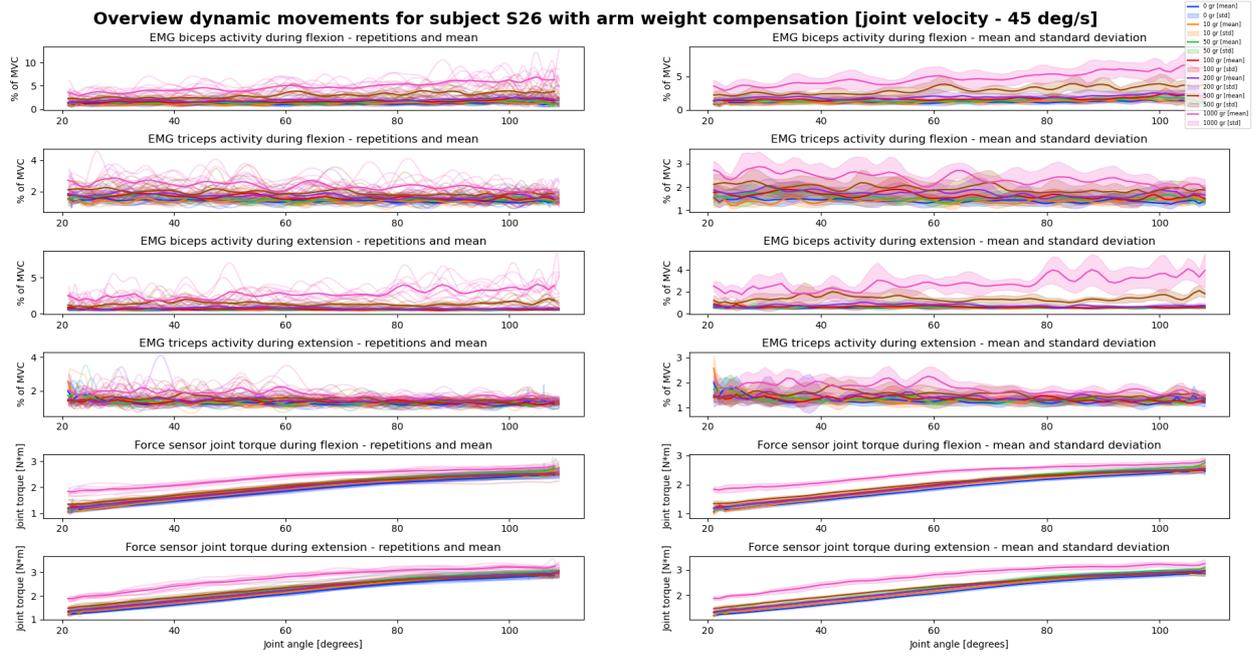


Fig. 50. Features during dynamic task [joint velocity - 45 deg/s]

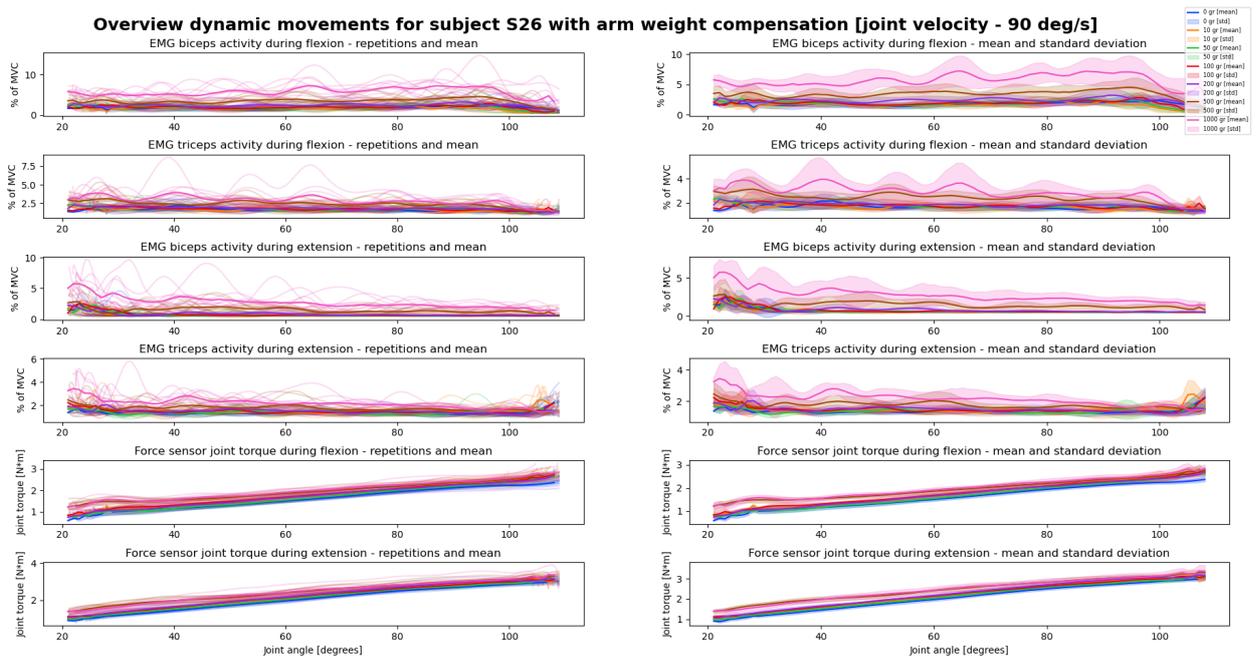


Fig. 51. Features during dynamic task [joint velocity - 90 deg/s]

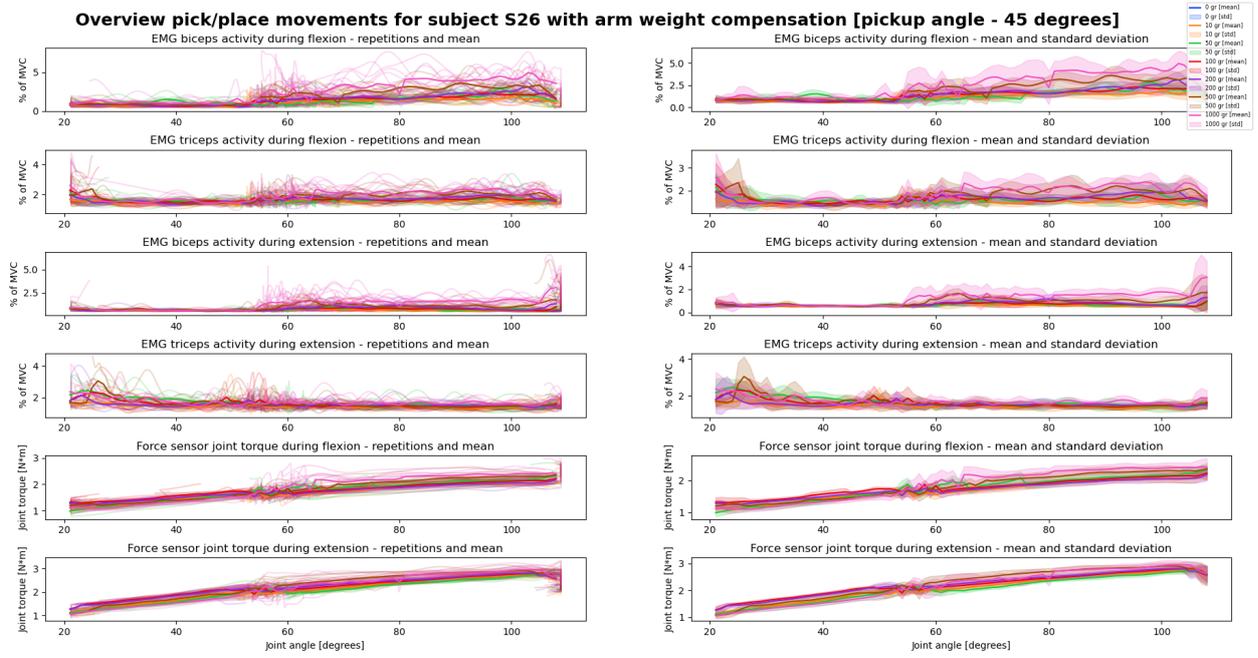


Fig. 52. Features during pick/place task [joint angle - 45 degrees]

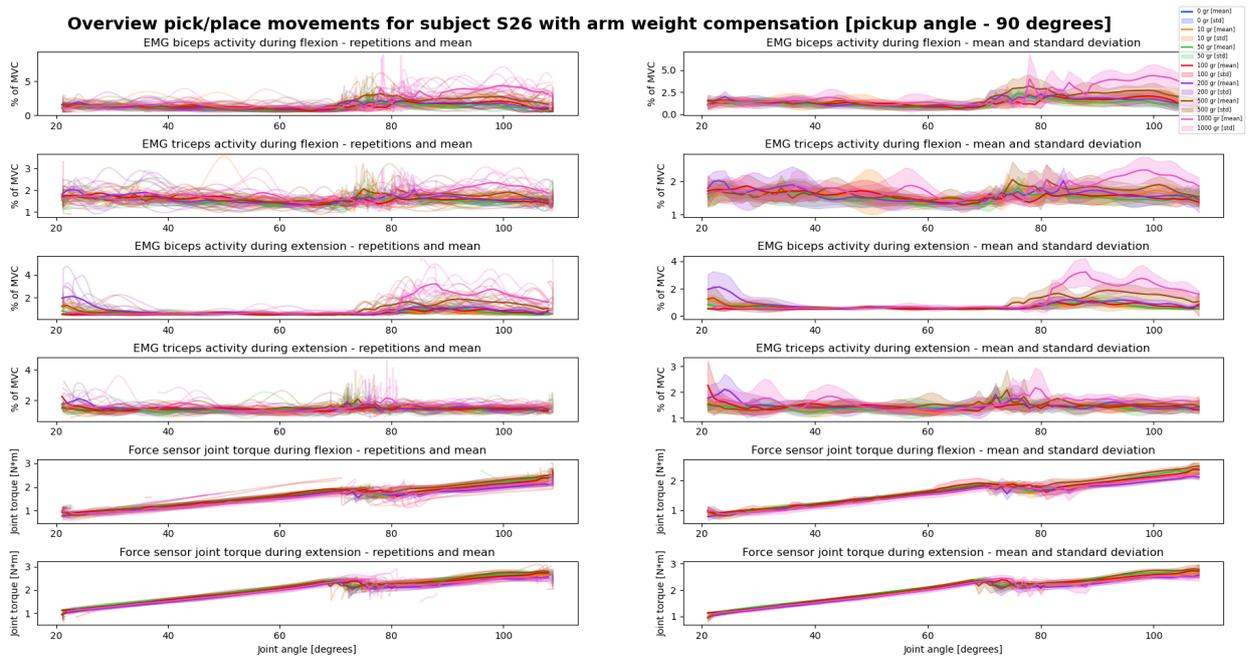


Fig. 53. Features during pick/place task [joint angle - 90 degrees]

### H.10 Overview measurements - subject S31

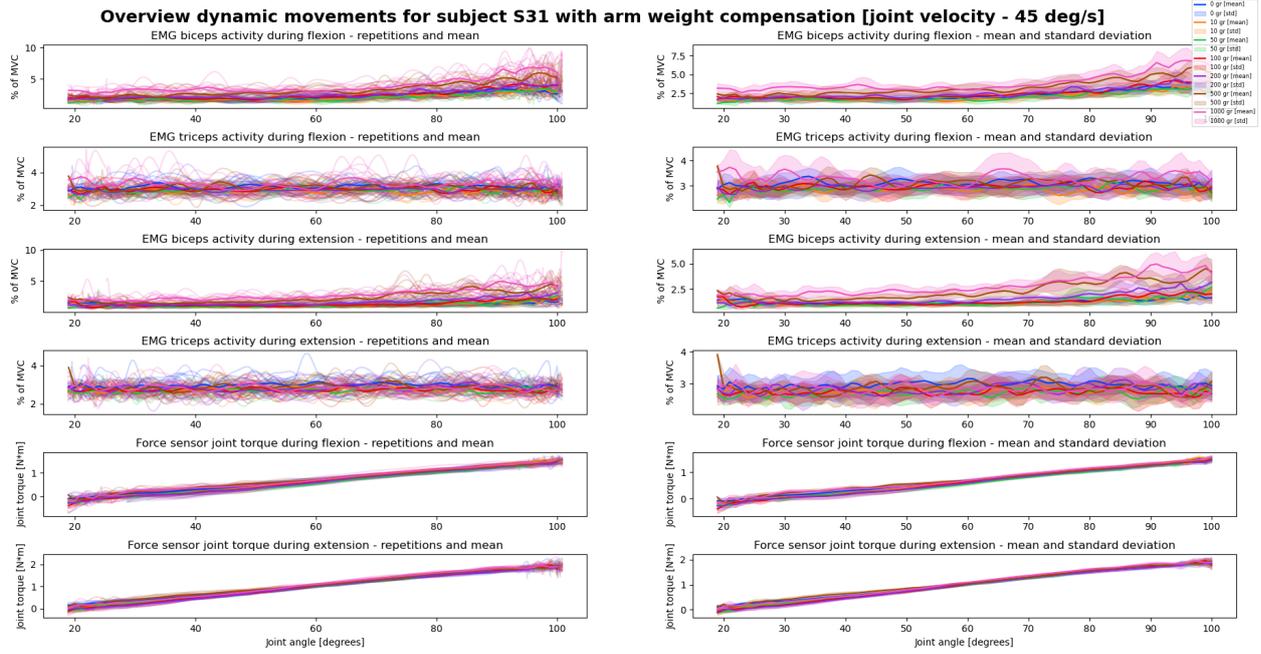


Fig. 54. Features during dynamic task [joint velocity - 45 deg/s]

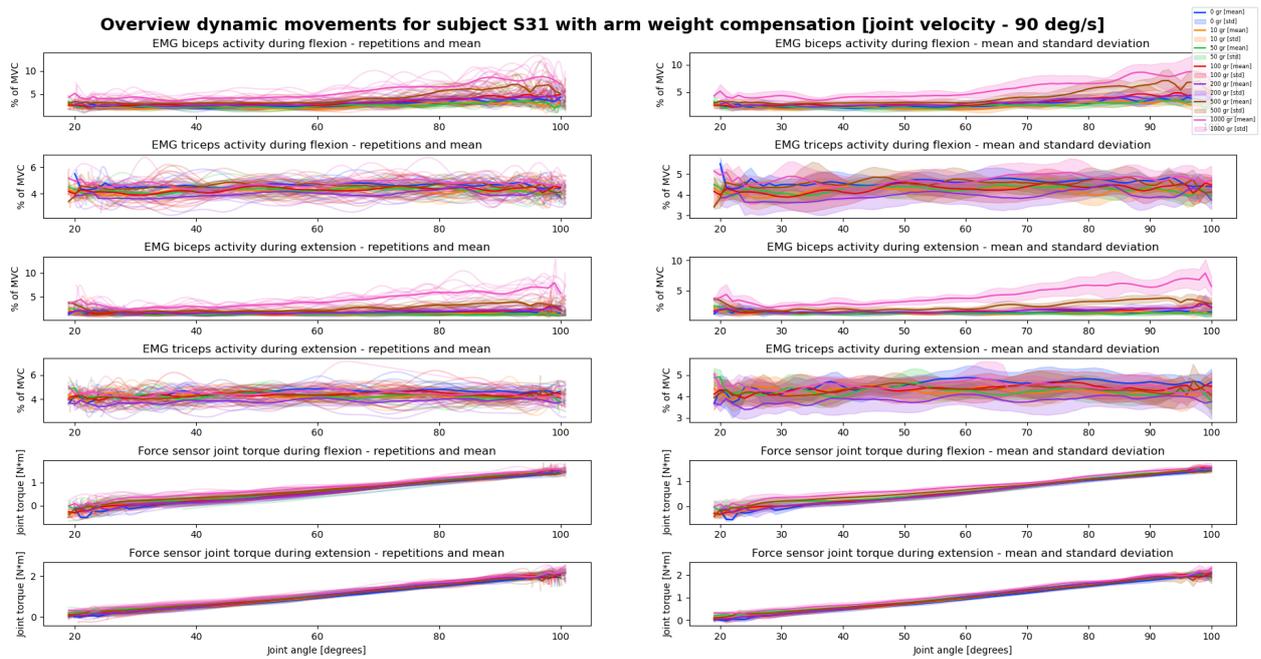


Fig. 55. Features during dynamic task [joint velocity - 90 deg/s]

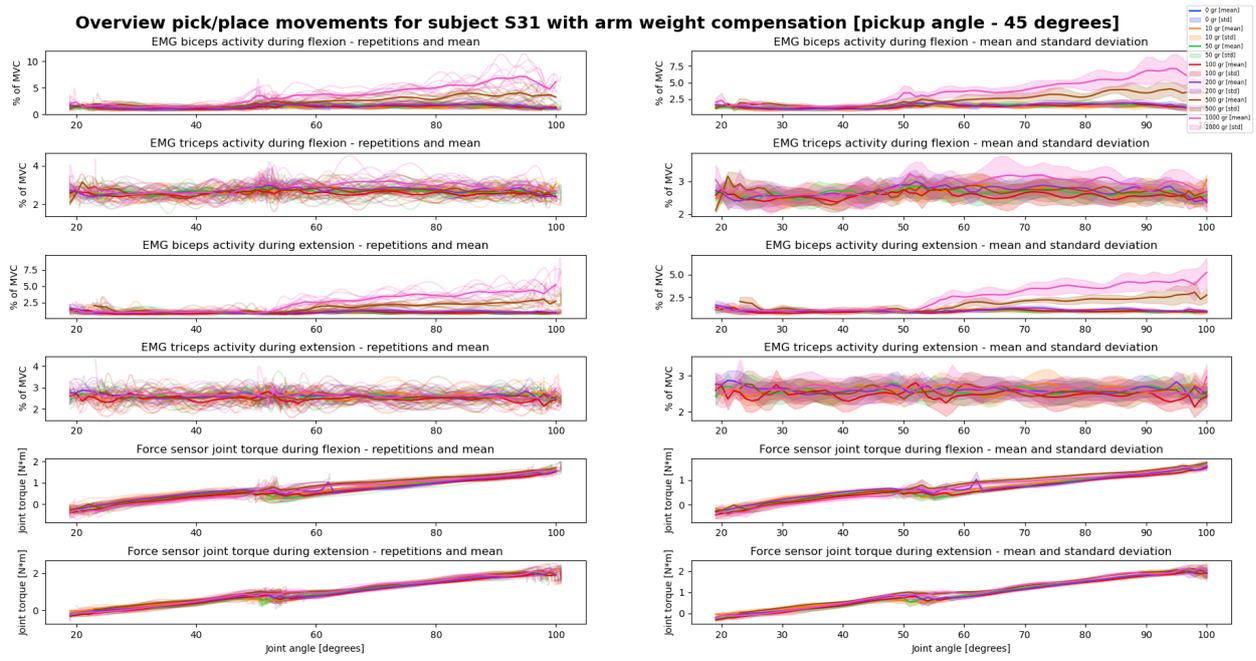


Fig. 56. Features during pick/place task [joint angle - 45 degrees]

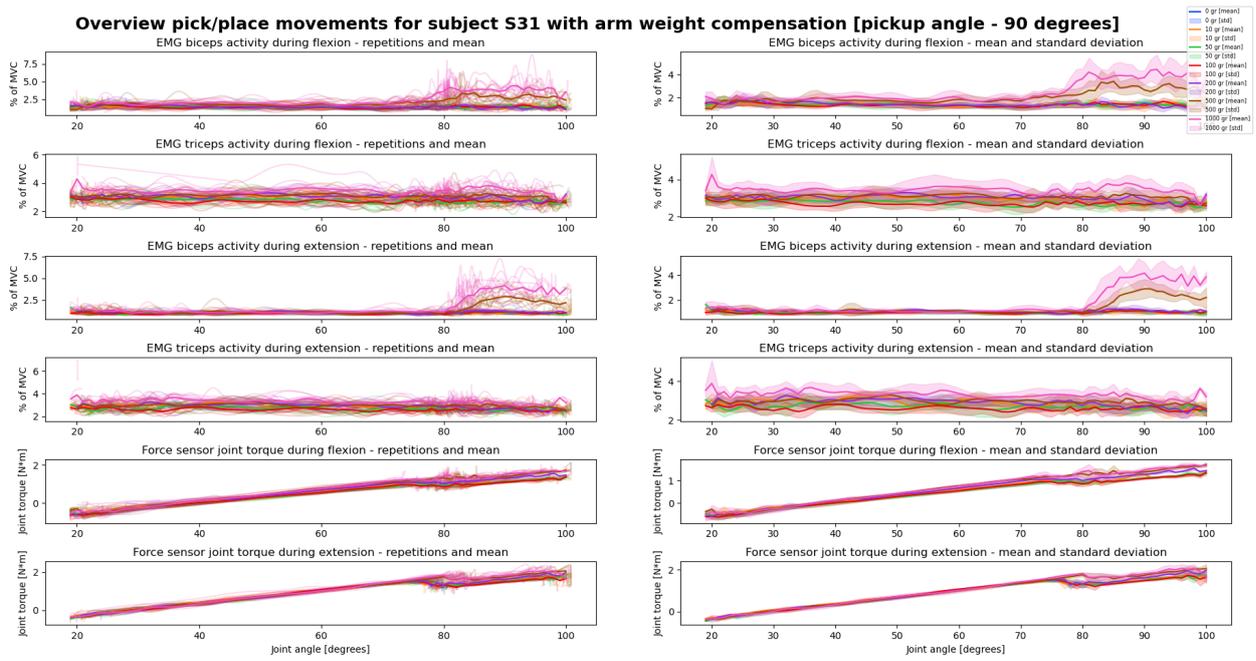


Fig. 57. Features during pick/place task [joint angle - 90 degrees]

### H.11 Overview measurements - subject S42

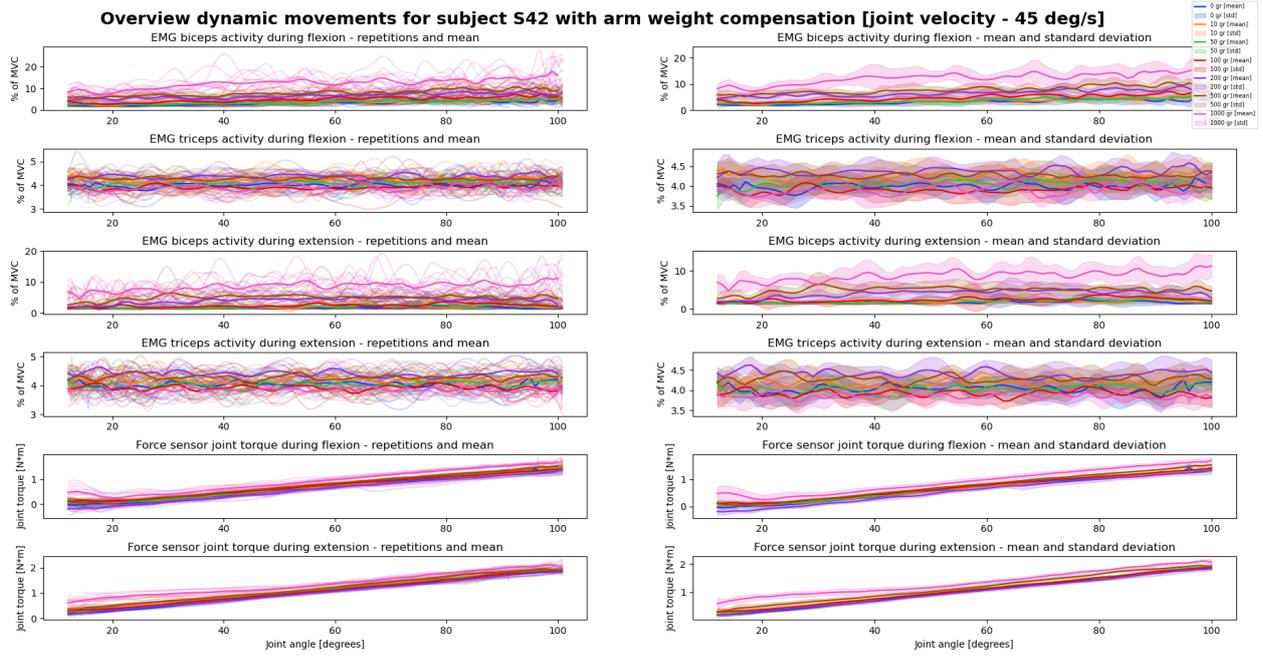


Fig. 58. Features during dynamic task [joint velocity - 45 deg/s]

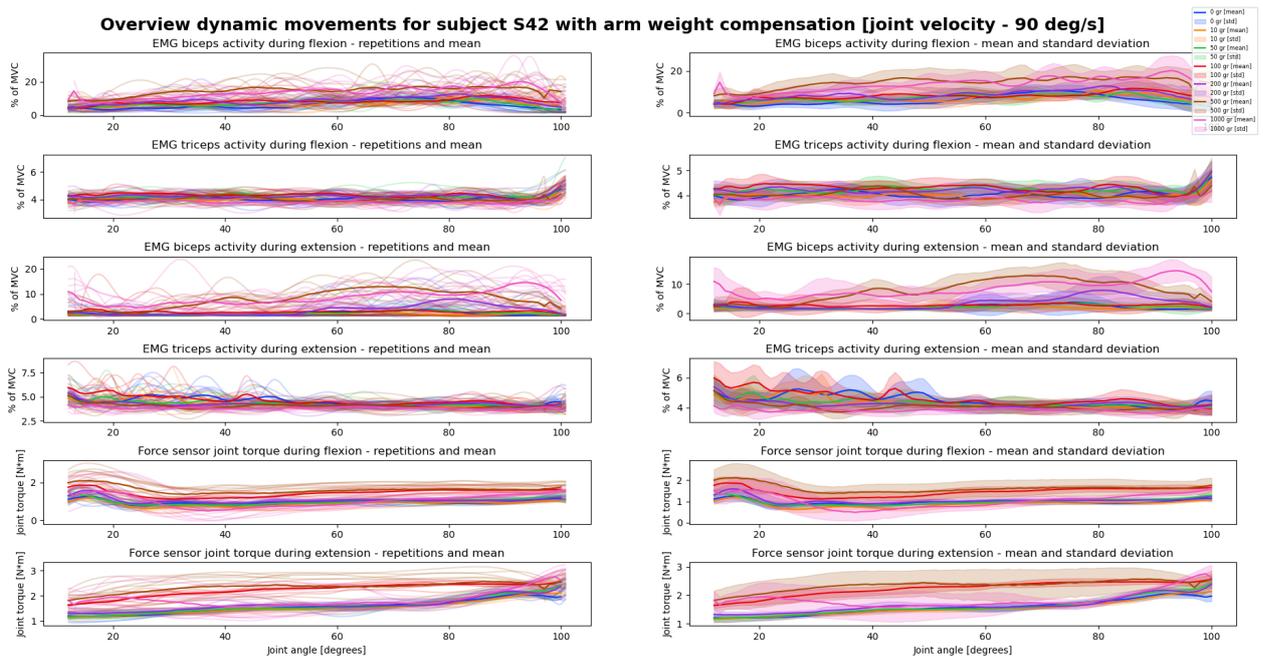


Fig. 59. Features during dynamic task [joint velocity - 90 deg/s]

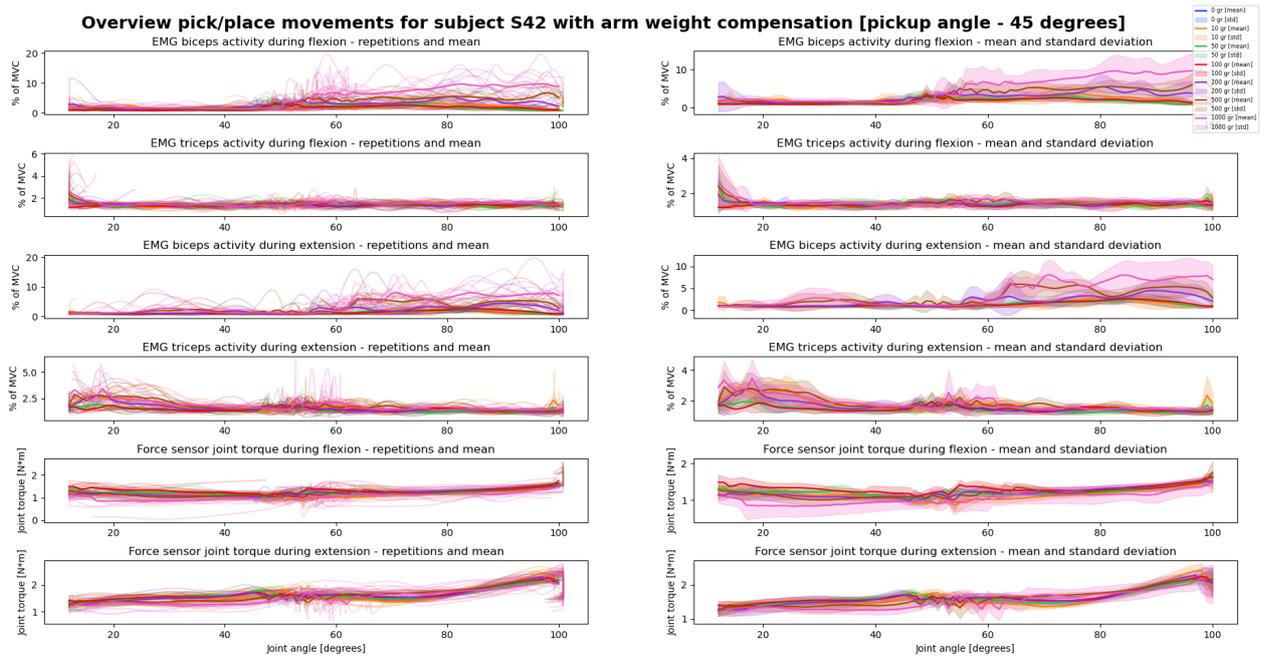


Fig. 60. Features during pick/place task [joint angle - 45 degrees]

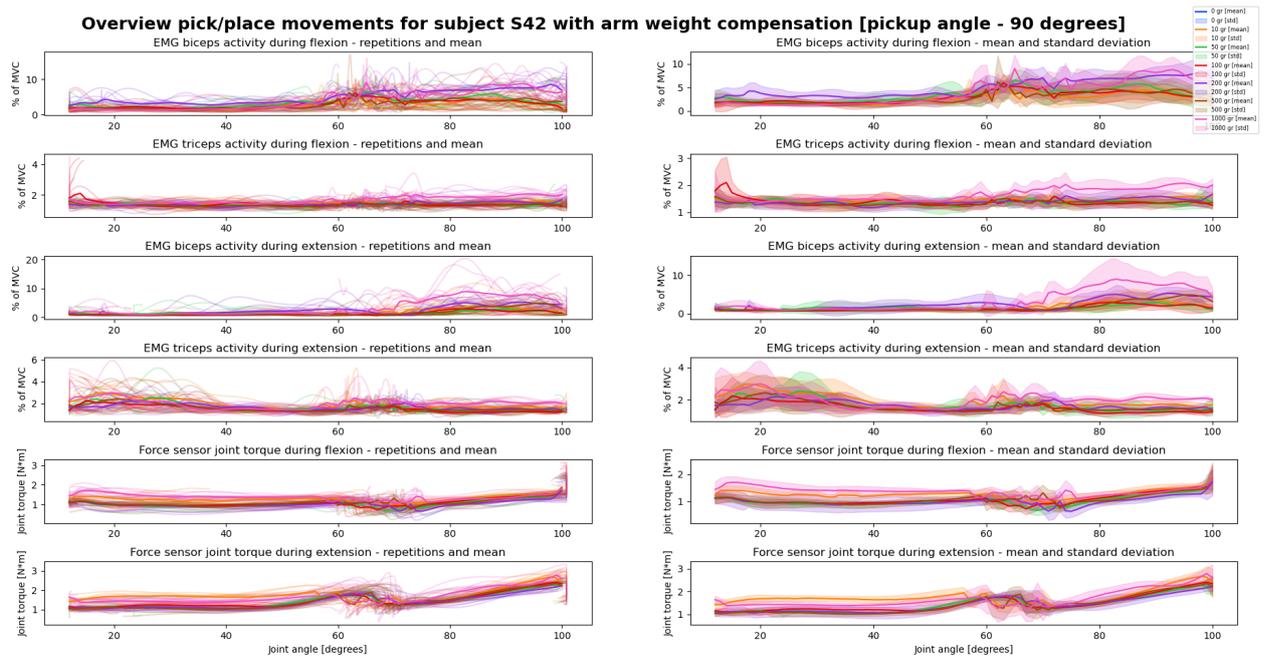


Fig. 61. Features during pick/place task [joint angle - 90 degrees]

### H.12 Overview measurements - subject S59

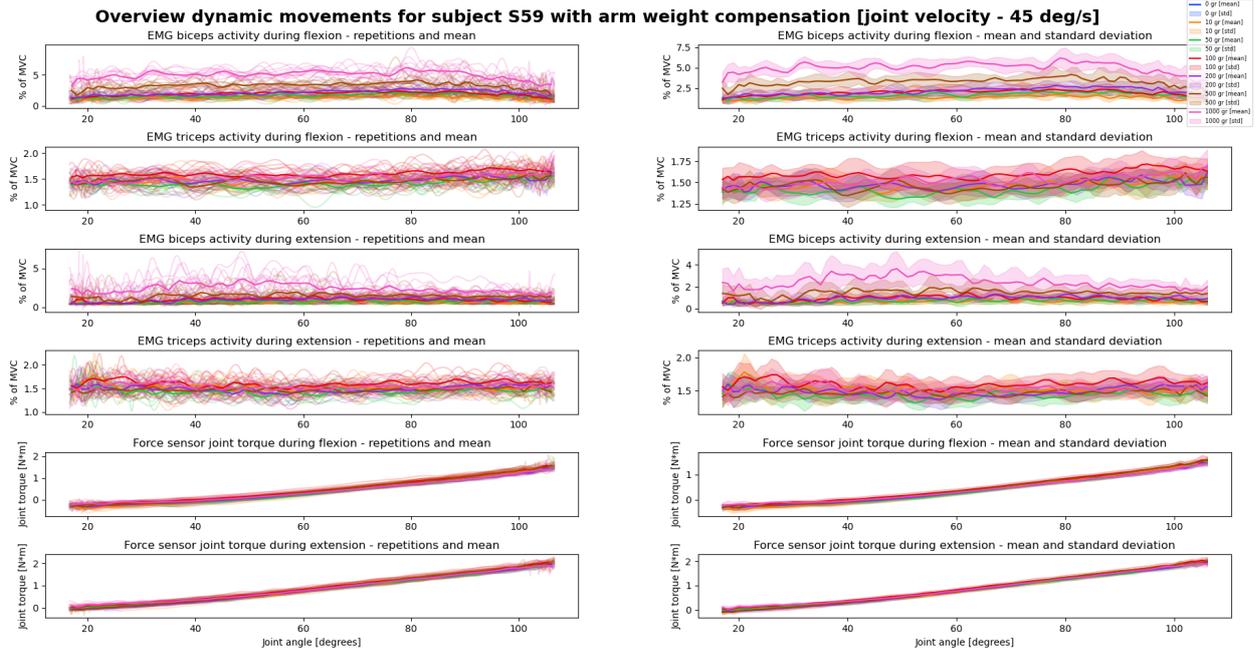


Fig. 62. Features during dynamic task [joint velocity - 45 deg/s]

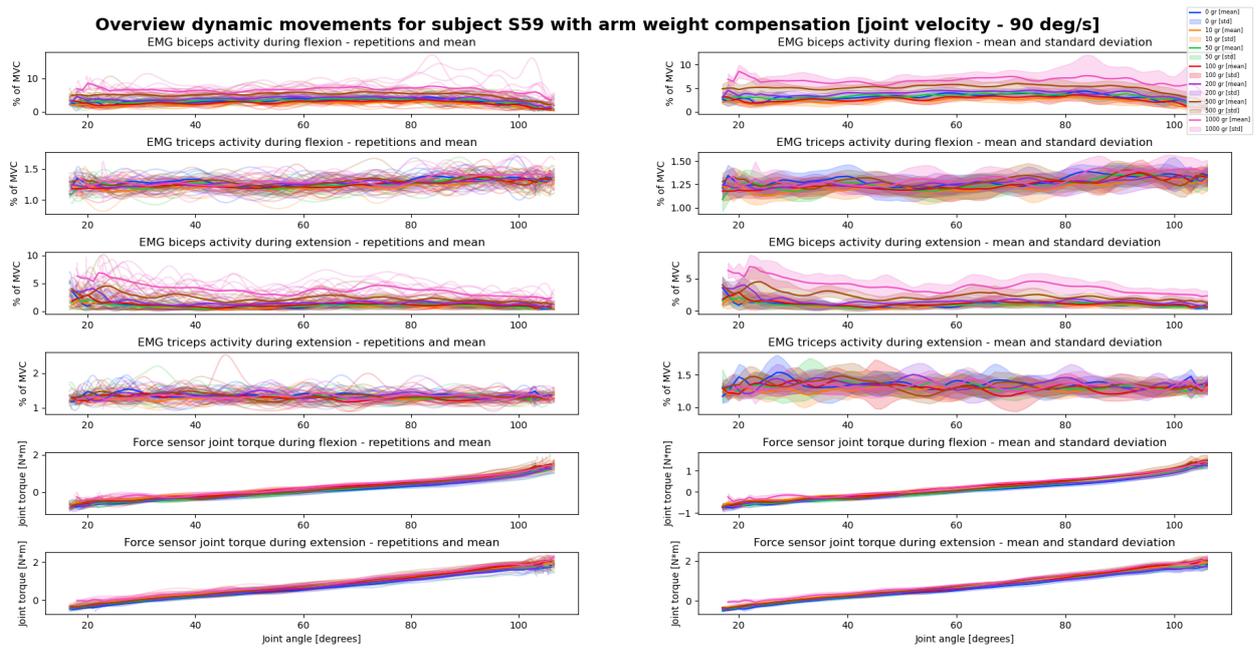


Fig. 63. Features during dynamic task [joint velocity - 90 deg/s]



Fig. 64. Features during pick/place task [joint angle - 45 degrees]

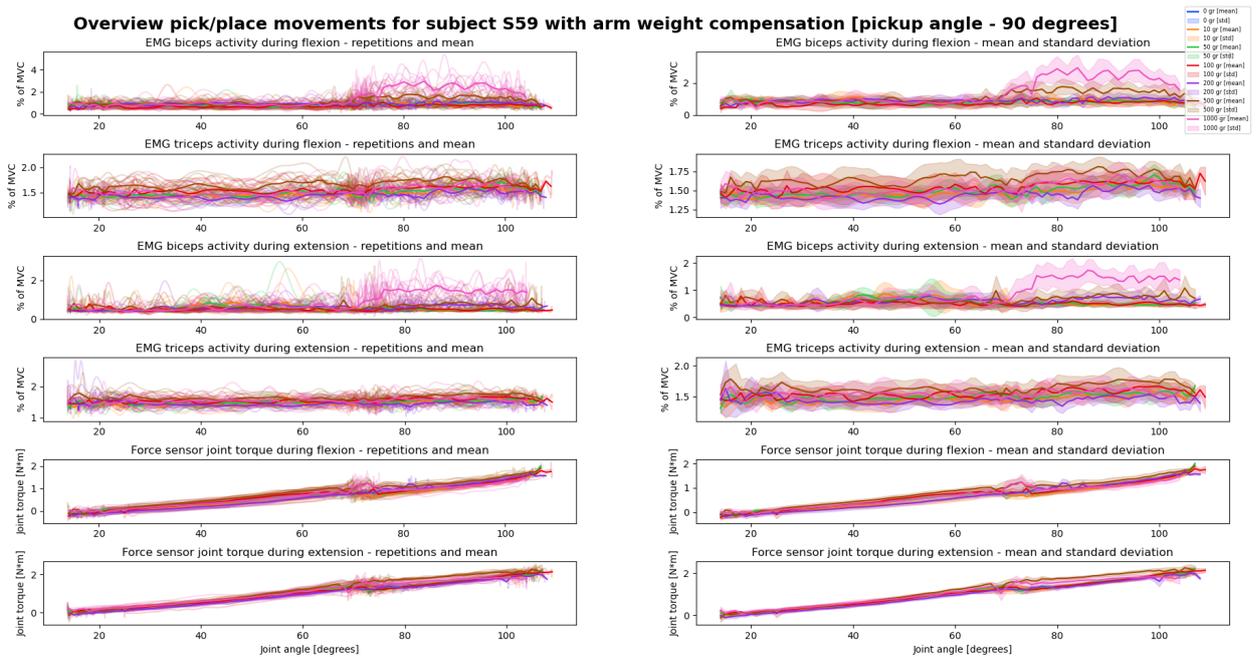


Fig. 65. Features during pick/place task [joint angle - 90 degrees]

### H.13 Overview measurements - subject S63

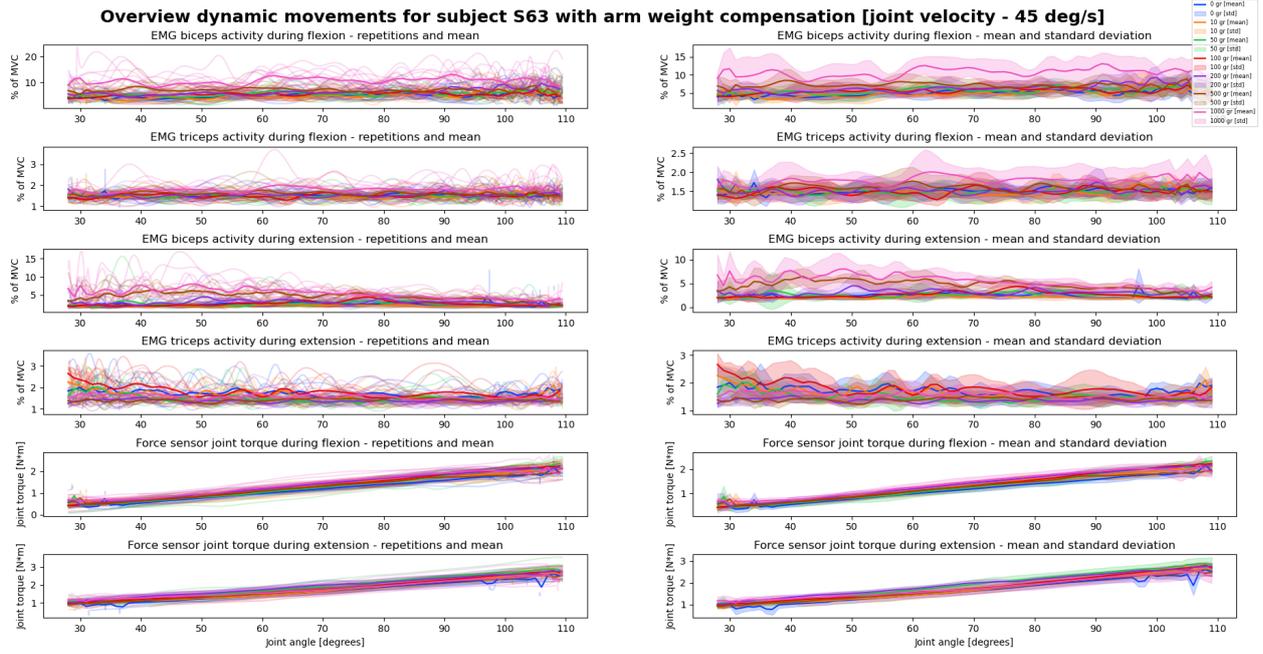


Fig. 66. Features during dynamic task [joint velocity - 45 deg/s]

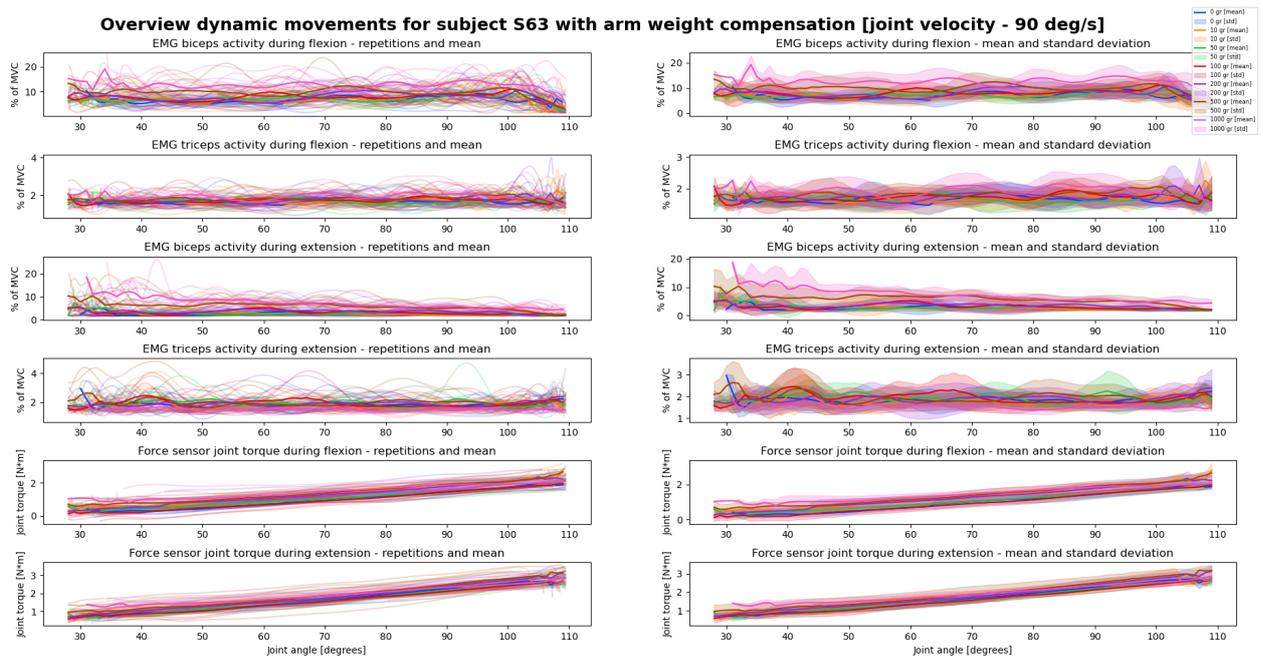


Fig. 67. Features during dynamic task [joint velocity - 90 deg/s]

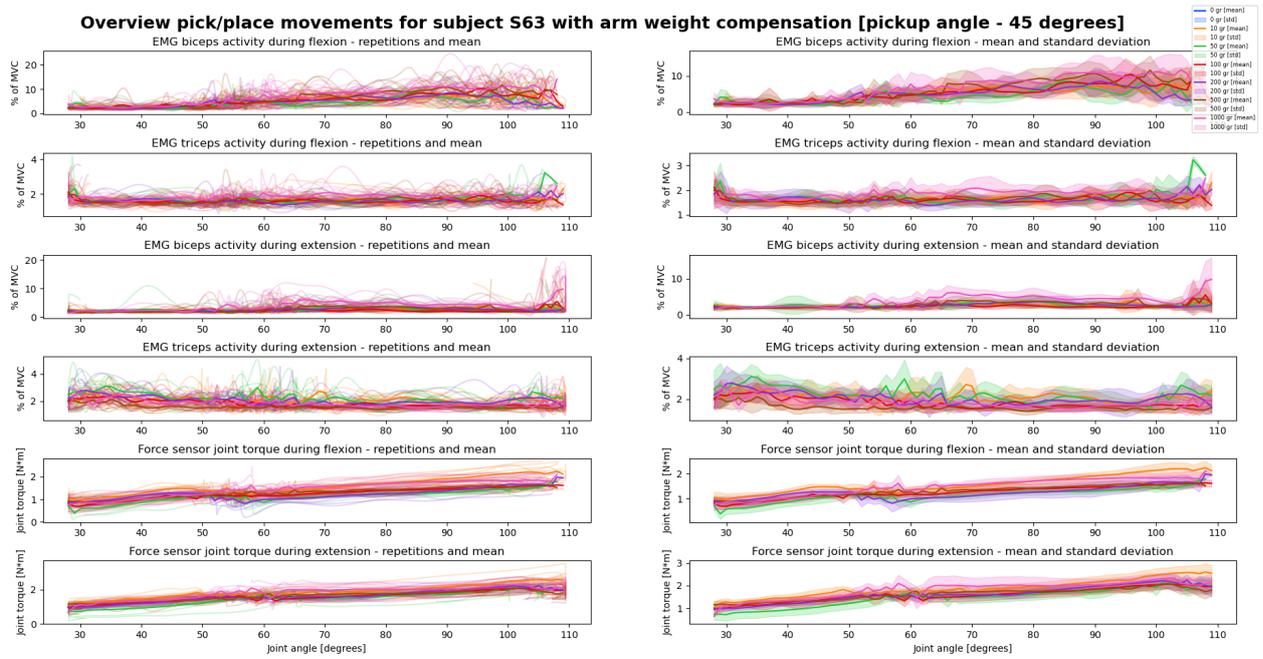


Fig. 68. Features during pick/place task [joint angle - 45 degrees]

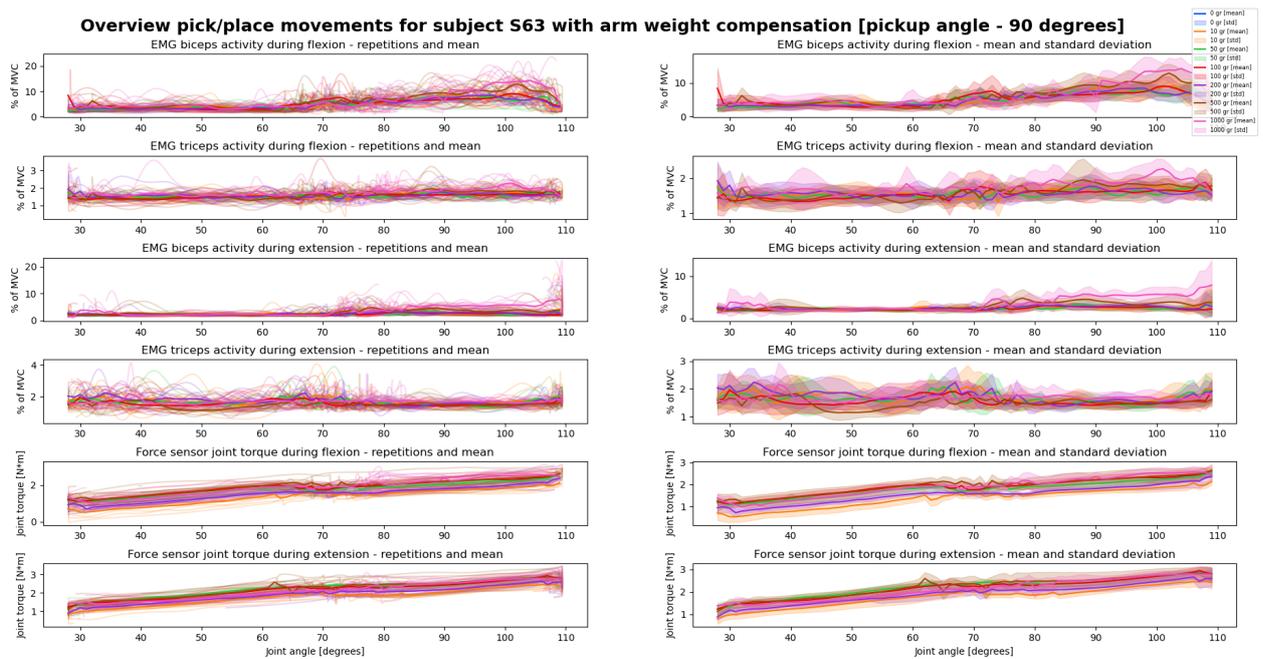


Fig. 69. Features during pick/place task [joint angle - 90 degrees]

### H.14 Overview measurements - subject S68

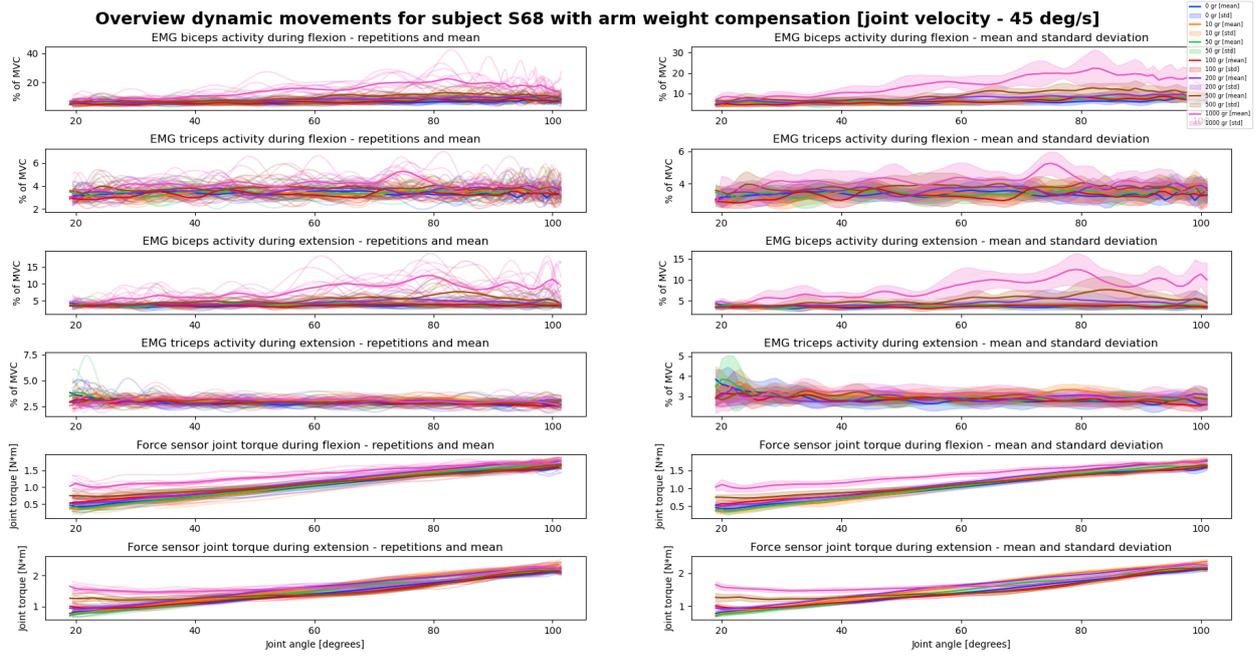


Fig. 70. Features during dynamic task [joint velocity - 45 deg/s]

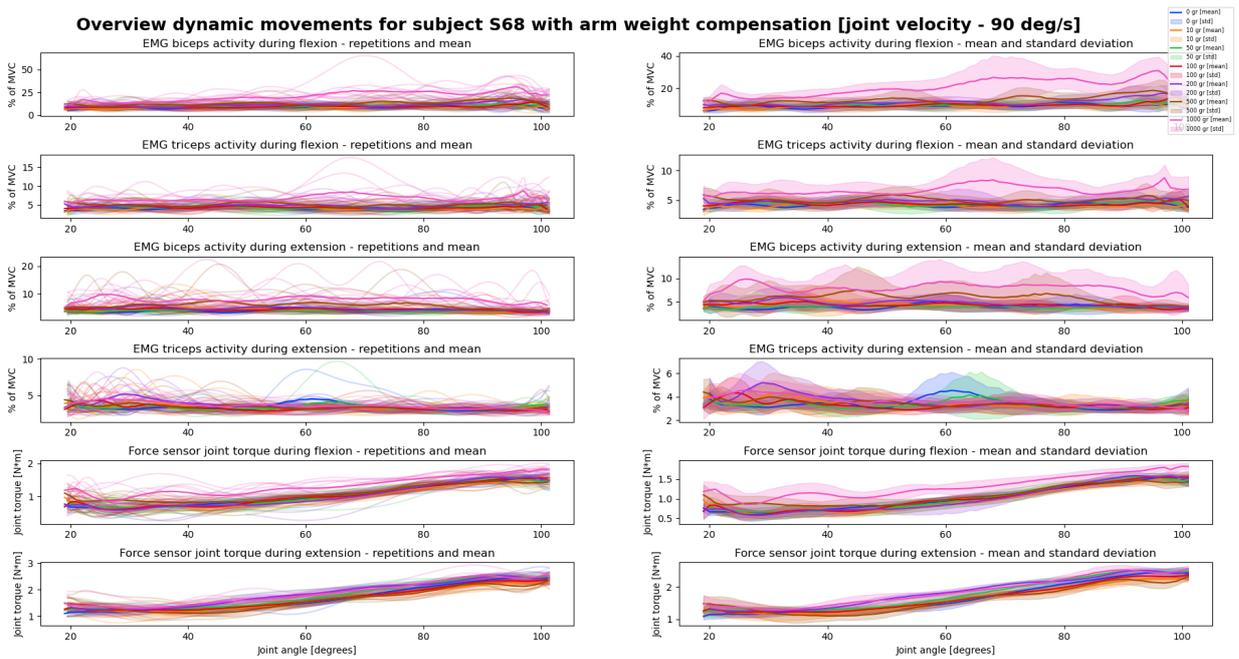


Fig. 71. Features during dynamic task [joint velocity - 90 deg/s]

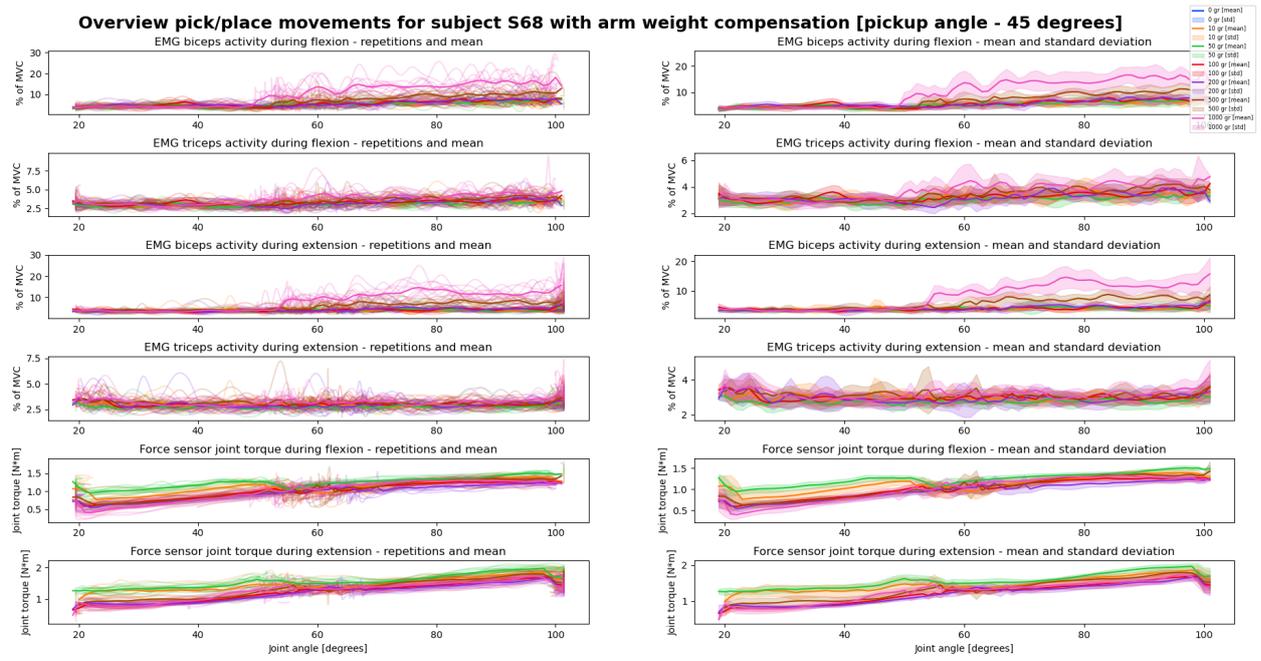


Fig. 72. Features during pick/place task [joint angle - 45 degrees]

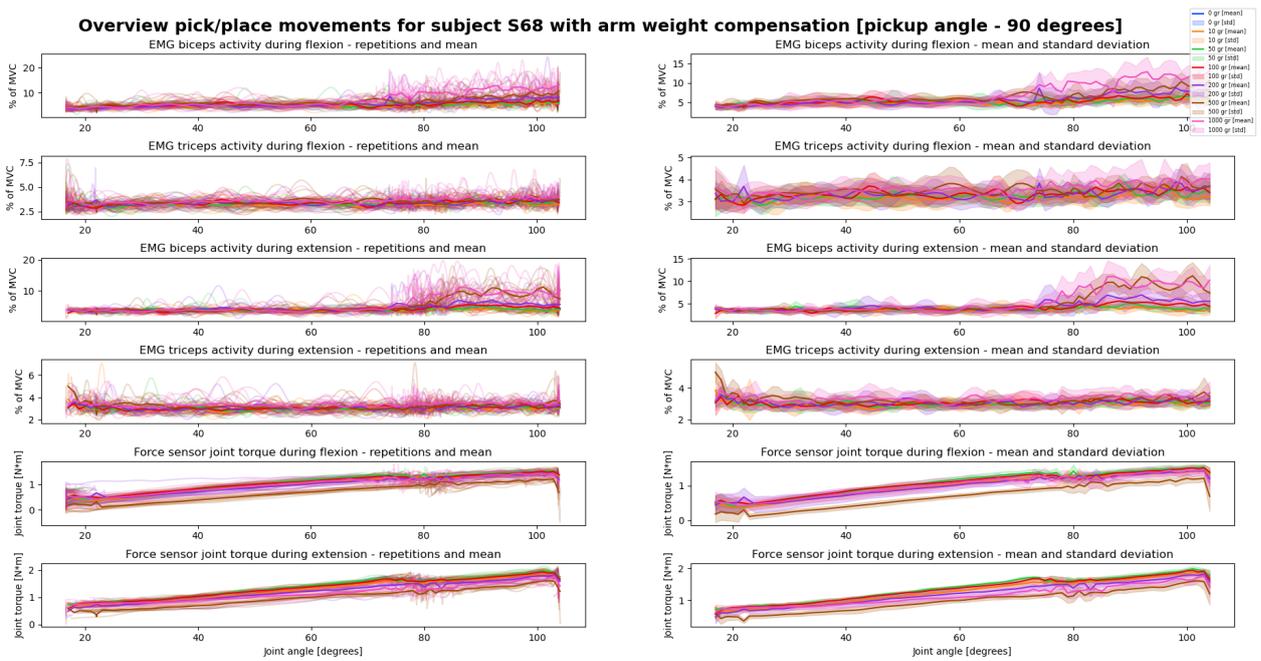


Fig. 73. Features during pick/place task [joint angle - 90 degrees]

### H.15 Overview measurements - subject S80

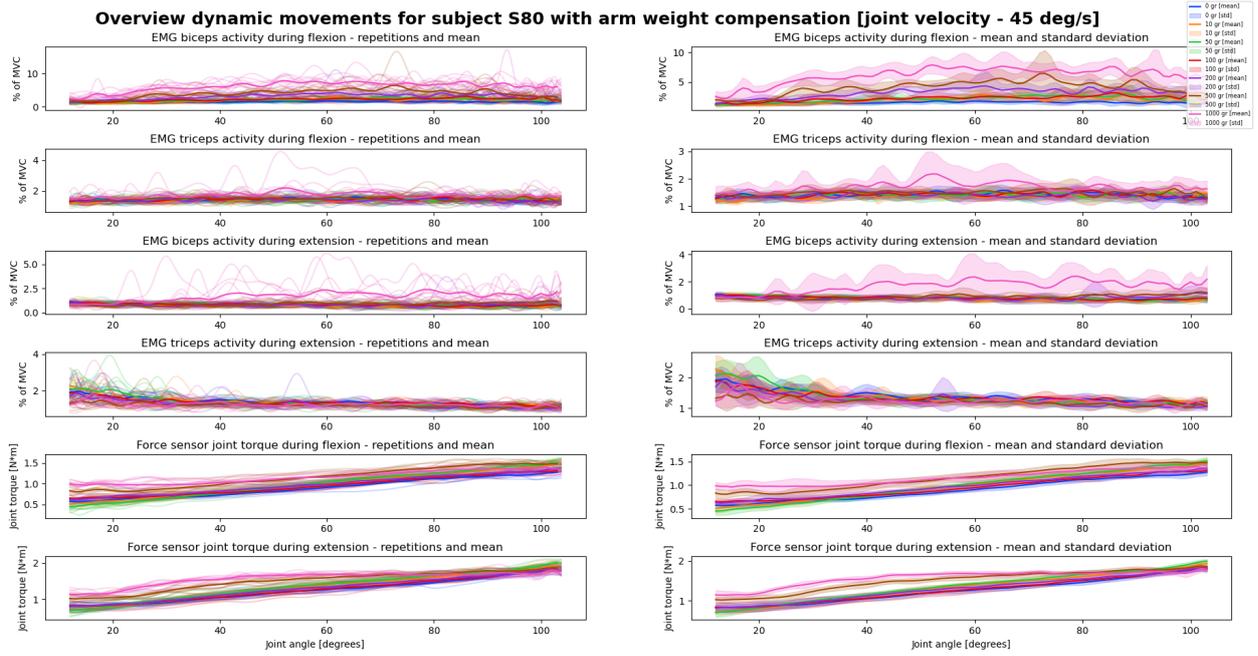


Fig. 74. Features during dynamic task [joint velocity - 45 deg/s]

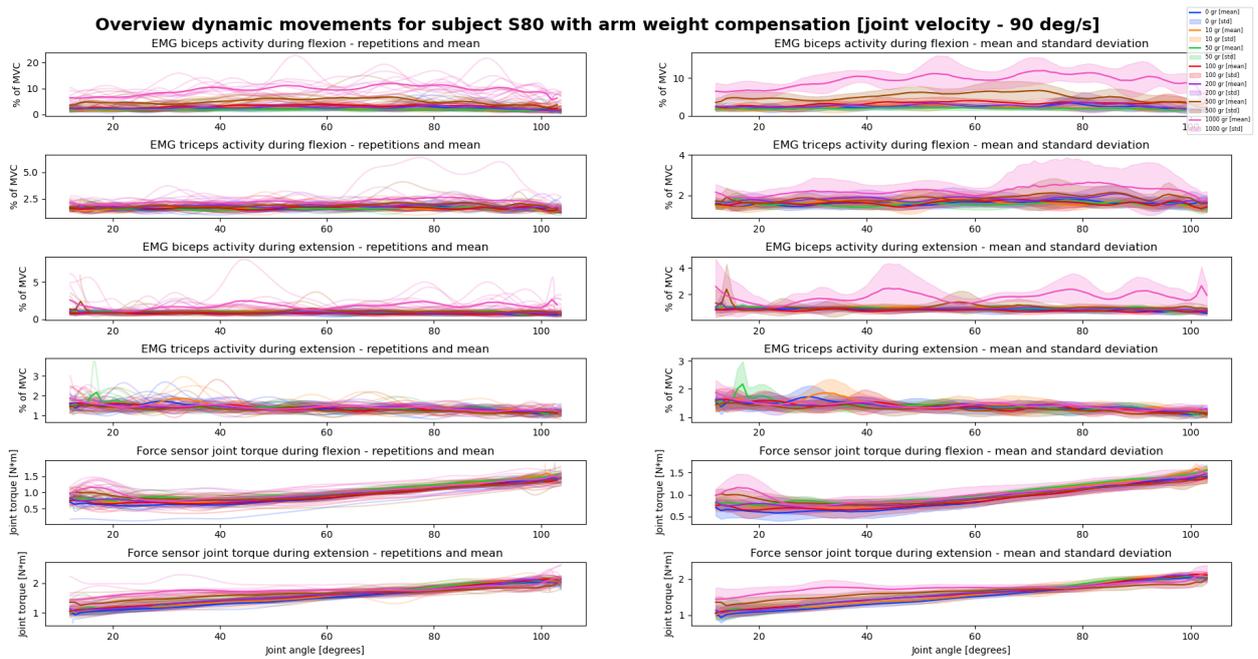


Fig. 75. Features during dynamic task [joint velocity - 90 deg/s]

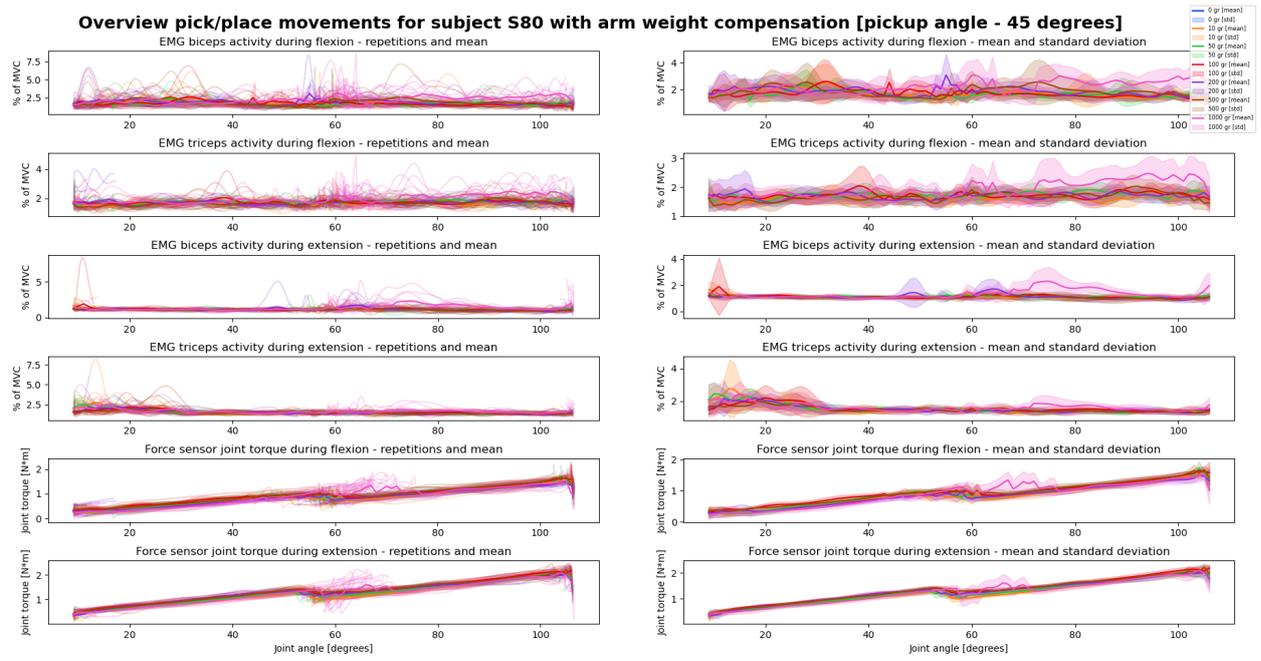


Fig. 76. Features during pick/place task [joint angle - 45 degrees]

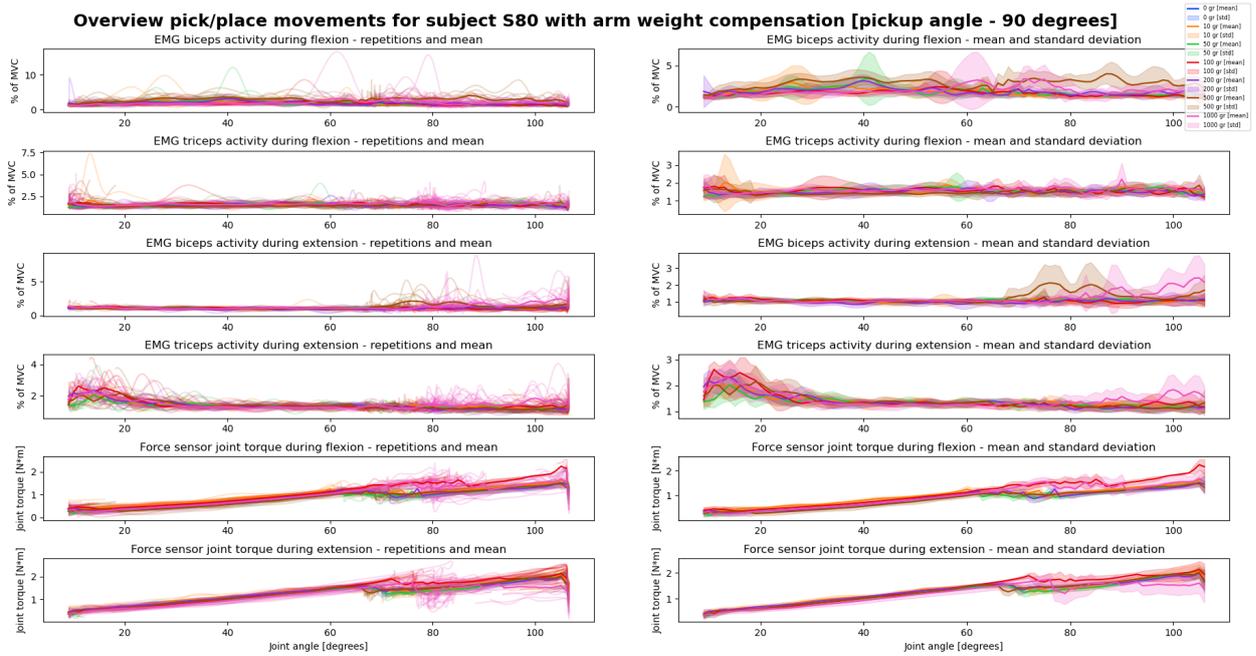


Fig. 77. Features during pick/place task [joint angle - 90 degrees]

### H.16 Overview measurements - subject S95

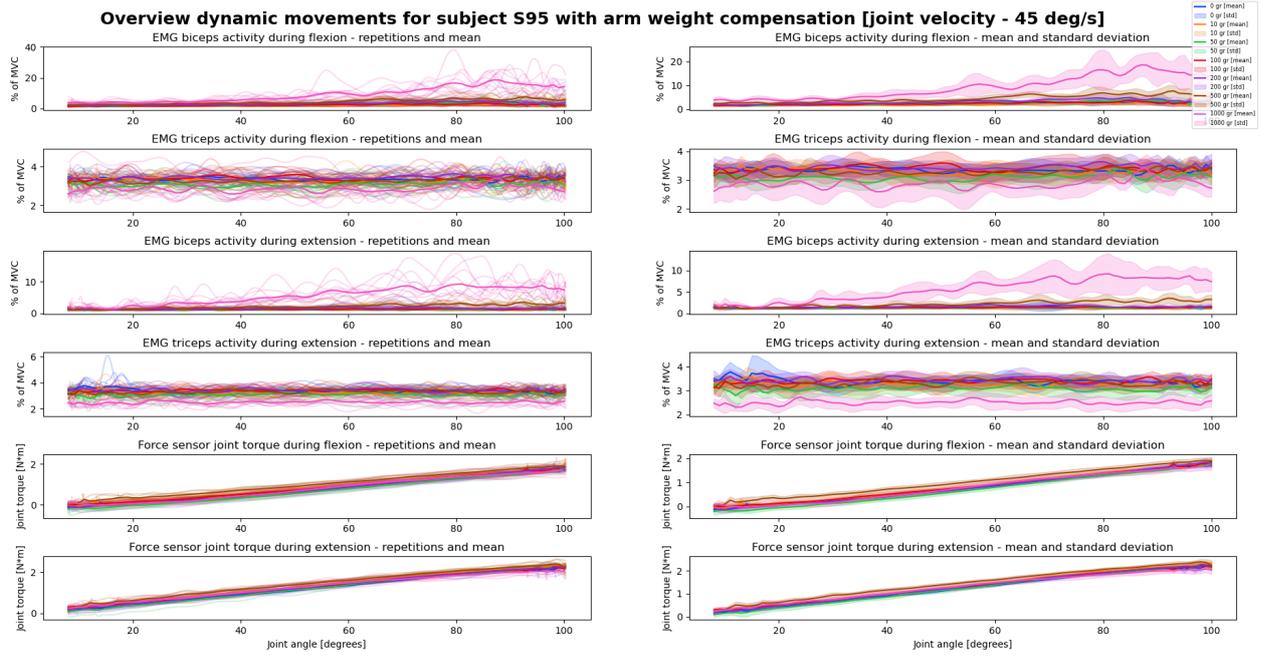


Fig. 78. Features during dynamic task [joint velocity - 45 deg/s]

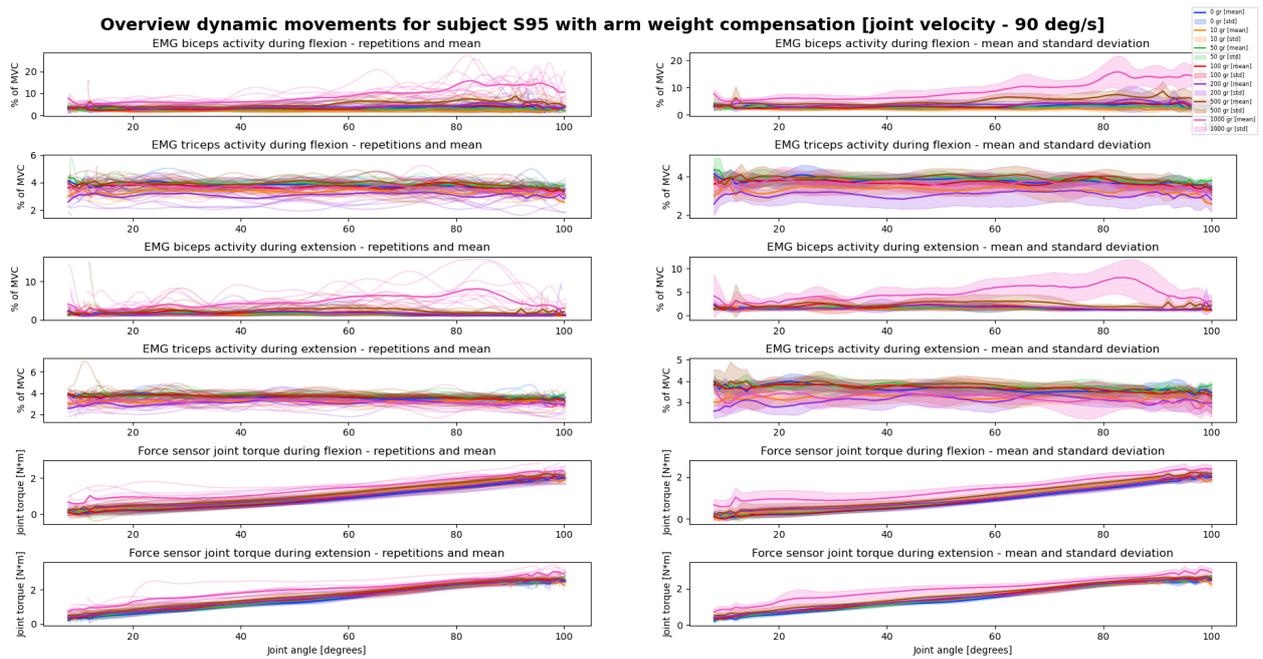


Fig. 79. Features during dynamic task [joint velocity - 90 deg/s]

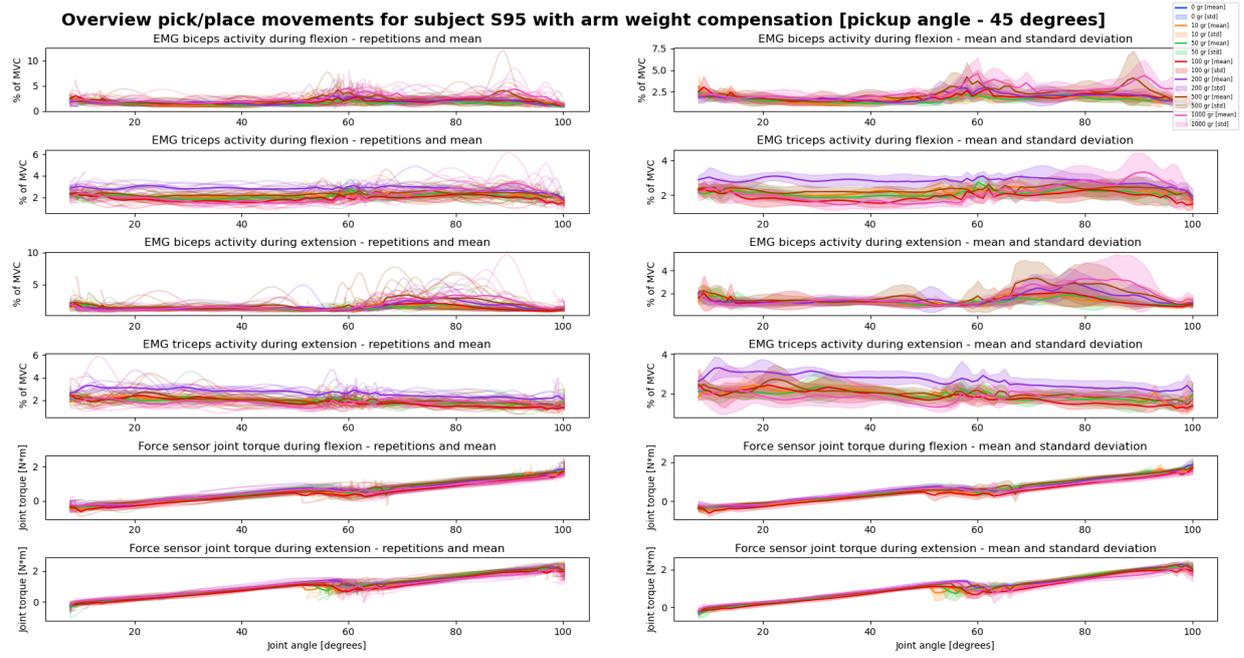


Fig. 80. Features during pick/place task [joint angle - 45 degrees]

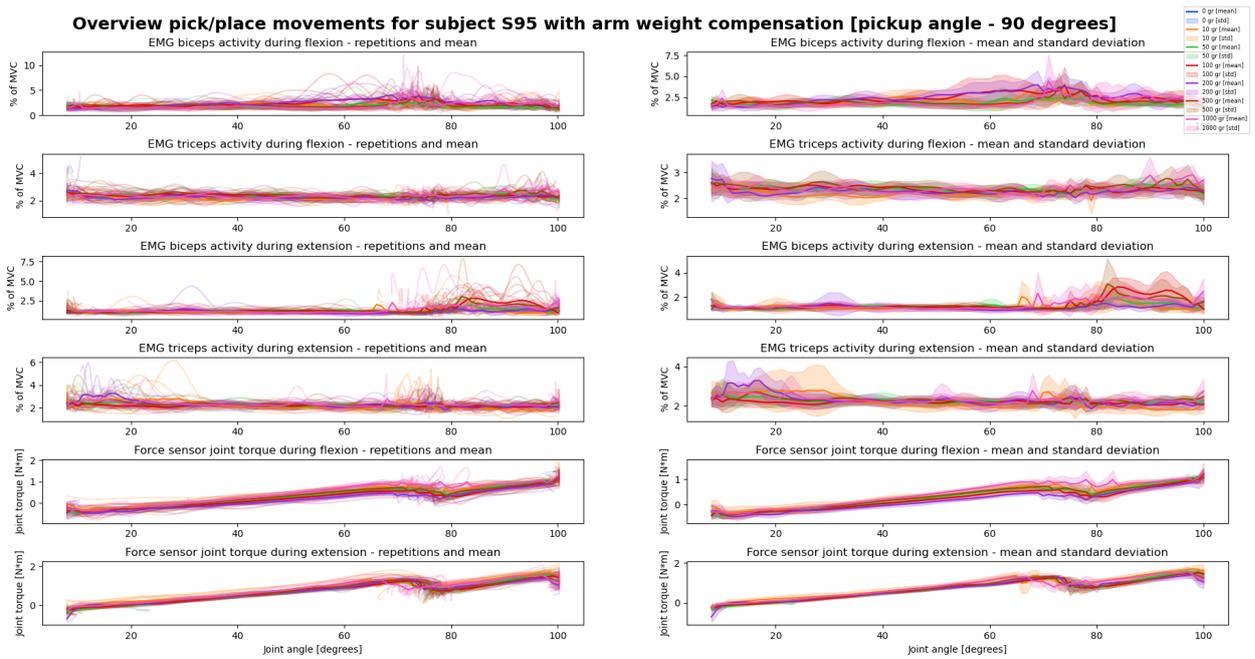


Fig. 81. Features during pick/place task [joint angle - 90 degrees]

### H.17 Overview measurements - subject S98

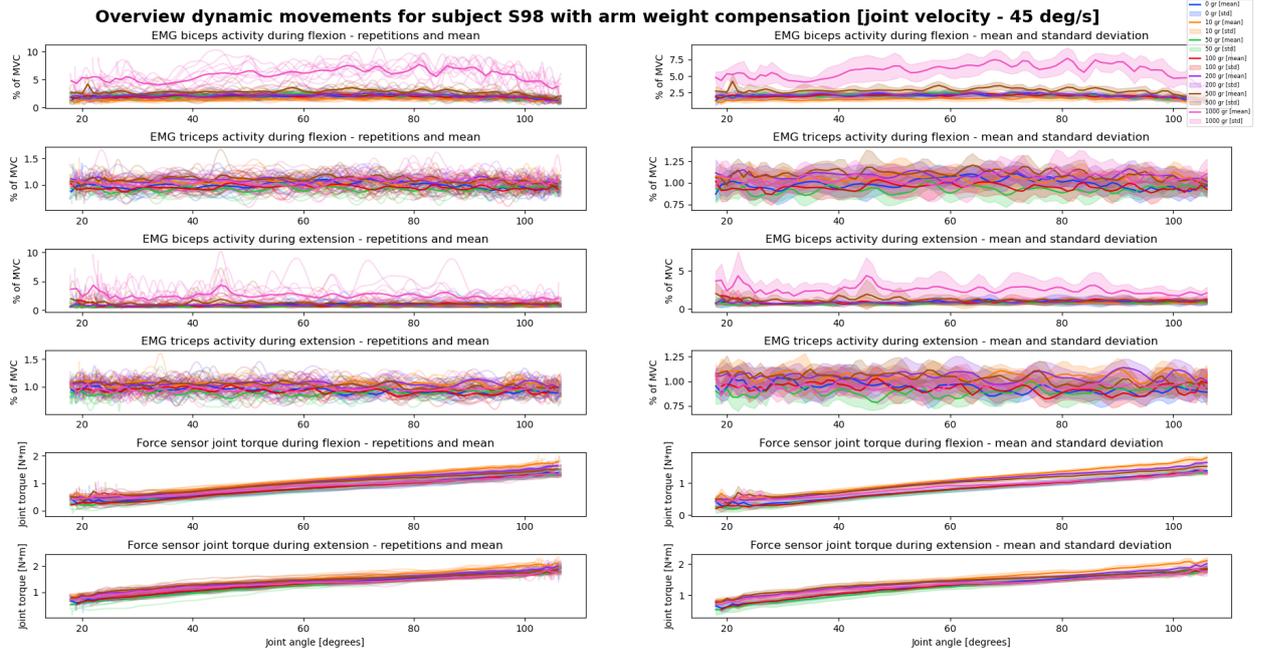


Fig. 82. Features during dynamic task [joint velocity - 45 deg/s]

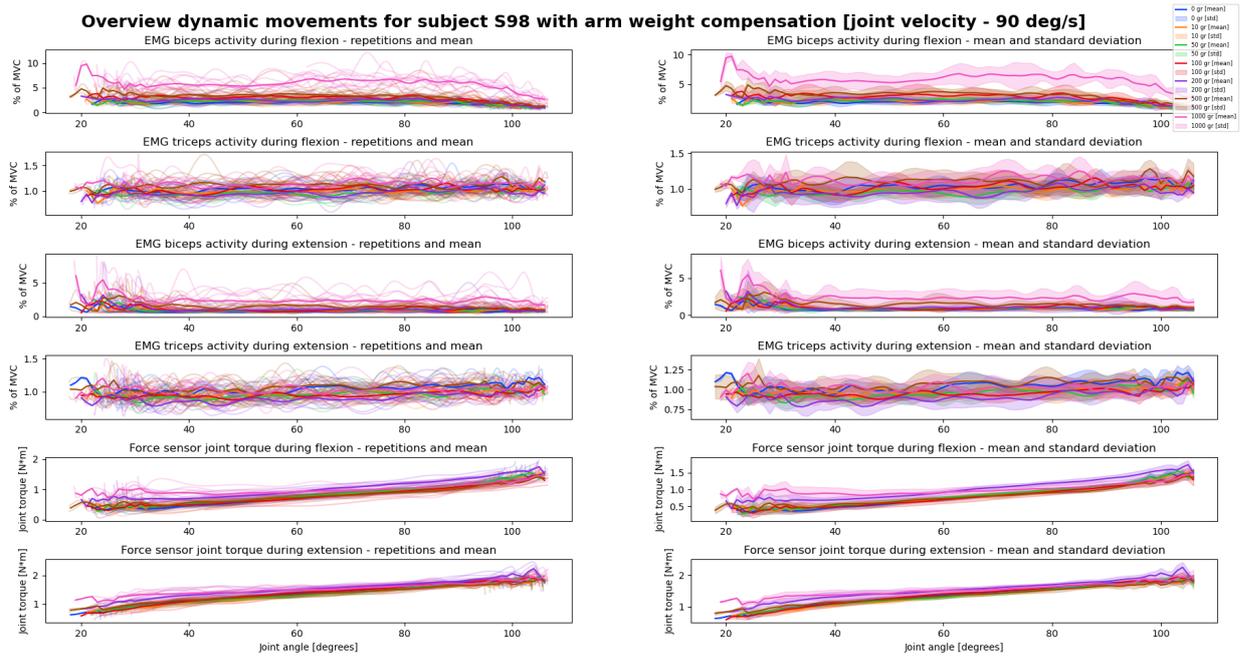
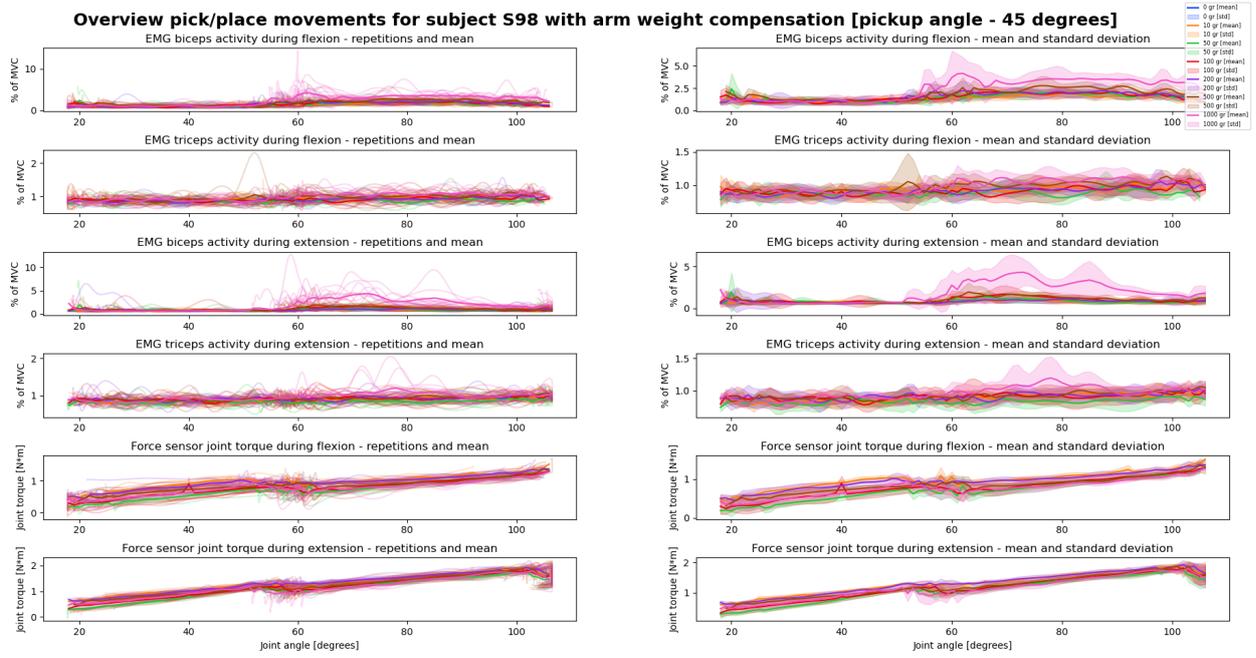
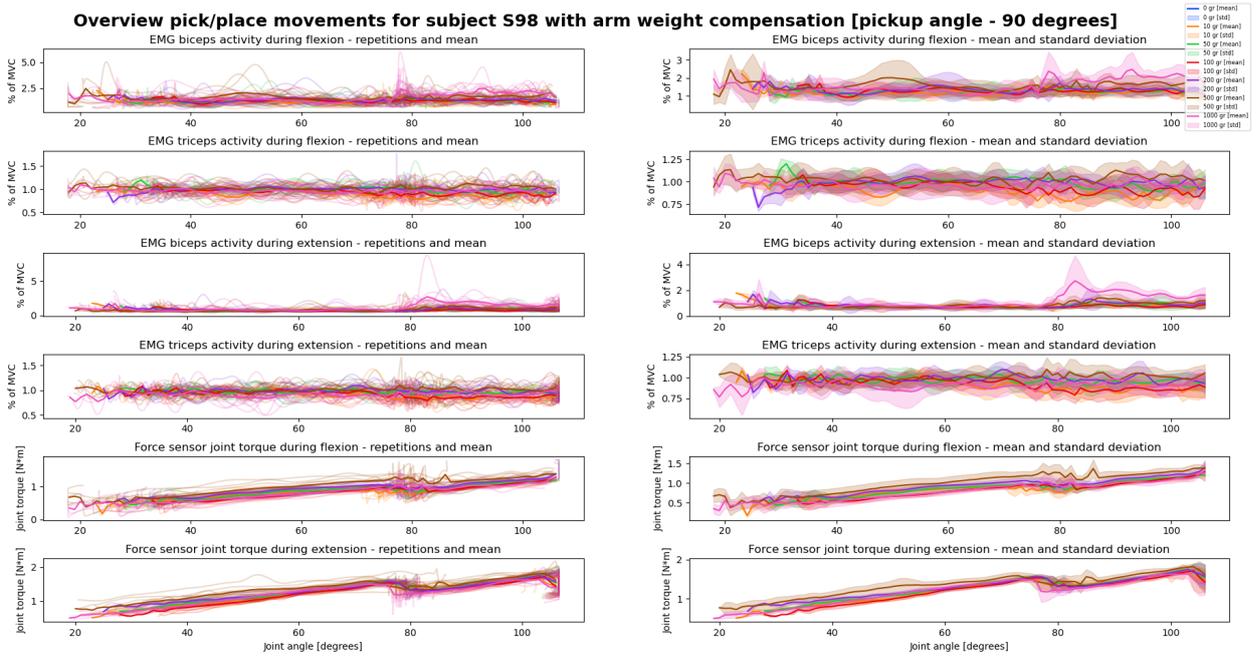


Fig. 83. Features during dynamic task [joint velocity - 90 deg/s]



**Fig. 84.** Features during pick/place task [joint angle - 45 degrees]



**Fig. 85.** Features during pick/place task [joint angle - 90 degrees]