university of groningen

faculty of science and engineering

Brain Computer Interfaces for Robust Continuous Control

Lüke Luna van den Wittenboer

university of groningen

faculty of science and engineering

University of Groningen

Brain Computer Interfaces for Robust Continuous Control

Master's Thesis

To fulfill the requirements for the degree of
Master of Science in Artificial Intelligence
at University of Groningen under the supervision of
Dr. A.I. Sburlea (Artificial Intelligence, University of Groningen)
and
Drs. I.P. de Jong (Artificial Intelligence, University of Groningen)

Lüke Luna van den Wittenboer (s3480569)

April 26, 2025

Contents

1 Introduction

The field of Brain-Computer Interfaces (BCI) aims to create systems that facilitate direct communication between the brain and an external device. The core function of a BCI is to detect and decode brain signals and translate them into commands that a computer can interpret. This type of technology can make a fundamental difference for people who suffer from paralysis or amputation, as BCI can facilitate the use of prosthetic limbs, wheelchairs, or cursors, thus offering a new method of interaction with the world.

1.1 State of the Art

One established BCI method involves interpreting Motor Imagery (MI) and/or execution (ME) through the detection of sensorimotor rhythms (SMRs) [1, 2]. One of these SMRs, the mu rhythm, is typically seen as alpha-band oscillations in the motor cortex at rest. During imagined or physical movement when the motor cortex is engaged, the mu rhythm contralateral to the moved or imagined body part is decreased. This process is known as event-related desynchronisation (ERD). Both ERD and its opposite, event-related synchronisation (ERS), can be used for controlling BCIs [3]. SMR-based BCIs have been used for user interaction with virtual objects in 2D [4] and physical objects in 3D space [5, 6], for wheelchairs and robotic arms [7].

Machine Learning (ML) and Deep Learning (DL) methods have proved promising in the field of BCI due to their adeptness in detecting non-linear patterns in the data, a skill that is vital in interpreting brain signals. Multiple studies have shown high accuracy in classifying motor attempts and executions based on EEG signals [8].

It is however important to point out that most of these studies utilise the same publicly available EEG data sets provided by BCI competitions. The use of these datasets has some major caveats. Firstly, they are relatively small, making the generalisability of any ML method trained on them questionable [9].

Furthermore, the data sets often rely on a uniform trial structure, which does not translate into continuous control settings. Not only does this create a staggered, unintuitive movement, but this more complicated device interaction might also lead to an overall decrease in user experience and enhanced sensory load, which can also negatively affect the performance of the user [7]. Some major problems can be identified with transitioning a BCI implementation from a trial-based task to a real-world environment: Firstly, a cued BCI implementation is devoid of any of the planning and anticipation effects one might experience in an uncued setting. Secondly, a trial-based setting creates clearly isolated instances of each different state, whilst this is not the case in an online setting where states might switch continuously. Lastly, cued control will have a transient phase between the presentation of a cue and the execution which is often omitted from analysis [10]. Therefore, many of these models will not be adept at detecting the onset of a state change with the low latency that an online BCI application often requires.

1.2 Background

1.2.1 Brain-Computer Interfaces and Neuroimaging

BCIs can be broadly categorised into two types depending on the method of signal acquisition: invasive and non-invasive. Invasive BCIs involve implanting electrodes directly into the brain tissue, offering high spatial and temporal resolution due to the direct access to neural activity. This approach, however, carries significant medical risks, including infection and long-term signal degradation, limiting its practical application outside clinical settings.

Non-invasive BCIs, in contrast, utilise external sensors to record neural signals without the need for surgery. Among these, electroencephalography (EEG) [11] is a widely used technique. EEG measures electrical activity with electrodes placed on the scalp, providing a safe, cost-effective and portable means of capturing brain signals. EEG has a lower spatial resolution compared to invasive methods, but its high temporal solution makes it a good method of neuroimaging for reactive BCIs.

The use of different types of neuro-imaging simultaneously or in conjunction with other physiological trackers such as eye-tracking or EMG can constitute a multi-modal BCI. Using multi-modal BCIs can increase the accuracy and robustness of a BCI system by compensating for weaknesses in one modality by using the strengths of another. For example, fMRI can be measured next to EEG in order to extract both temporal and spatial information at a high resolution [12]. Hybrid solutions such as using EEG and an eye-tracker to better analyse neural data based on the visual focus of the participant have also been proposed [13].

1.2.2 Motor Imagery, Attempt and Execution

Within Motor-Based BCIs, we make a distinction between three types of tasks: Motor Imagery (MI), Motor Attempt (MA) and Motor Execution (ME).

MI involves the internal rehearsal of a movement without performing any of the associated muscle activity. It relies on the activation of motor and premotor areas, resembling—but not completely mirroring—the neural patterns of actual movement.

MA refers to the conscious effort to perform a movement without the ability to actually perform it. This is often the case in clinical populations where physical movement is impaired or absent, for example, due to stroke or spinal cord injury. MA typically results in a stronger and more widespread cortical activation than MI, often including some degree of residual EMG activity [14].

Finally, ME refers to the actual performance of a physical movement, involving full activation of both central and peripheral motor pathways.

Whilst being different in terms of output, all three processes consistently activate key regions within the primary motor cortex (M1), supplementary motor area (SMA), premotor cortex (PMC), and posterior parietal cortex (PPC) [15, 16, 14].

From a practical perspective, experiments utilising ME are often easier to design and validate compared to MI tasks. Instructions for physical movement are typically more intuitive which can reduce cognitive load for participants. Furthermore, ME tasks provide observable behavioural feedback

which allows the researcher to verify that the participant is performing the intended movement. This makes data collection more efficient and reliable, especially in early-stage experiments or with participants not experienced with using MI-based BCIs. ME data therefore can be a valuable proxy for modelling motor-related brain activity.

1.2.3 Neural Correlates of Motor Control

During motor control tasks, various neurophysiological features emerge in EEG that reflect movement planning, execution, and inhibition. These include event-related desynchronisation (ERD) and synchronisation (ERS) in the mu and beta frequency bands, which indicate activation and deactivation of sensorimotor areas as described in the introduction. Slow cortical potentials, such as movement-related cortical potentials (MRCPs), precede movement and provide insight into the brain's preparation for motor movement [17]. These signals originate primarily from the supplementary motor area (SMA), premotor cortex (PMC), and primary motor cortex (M1) [18]. These features can offer recognisable markers of motor-related activity that can be the basis for decoding in a BCI [19].

1.2.4 Common Spatial Filters (CSPs)

Numerous techniques have been proposed for classifying MI/ME in brain activity measured by EEG, traditionally employing hand-crafted feature extraction methods in the time, frequency, spatial and spatio-temporal domains. A well-known and effective technique for feature extraction is Common Spatial Patterns (CSP) [20]. These are a supervised algorithm that projects EEG signals into a spatial domain where the variance between different conditions, such as ME with the left or right hand, is maximised for one class whilst being minimised for the other. Spatial features are particularly well-suited for EEG analysis because brain activity is distributed across multiple regions and specific regions are responsible for certain tasks, like movement in a specific arm. Some of the downsides of using CSPs are their sensitivity to artefacts produced by e.g. muscles in the face or eye movements, as the algorithm might find maximising variance difference more manageable between non-neural noise compared to actual brain activity.

1.2.5 Shallow Convolutional Neural Networks (SCNN)

In the field of Deep Learning (DL), other methods have been developed for MI classification. One point of focus is Convolutional Neural Networks (CNNs). These models have gained popularity in EEG-based BCI applications due to their ability to automatically extract spatial and temporal features without requiring handcrafted feature selection. However, EEG datasets are typically small compared to those used in conventional deep learning tasks, making deep architectures prone to overfitting due to their higher number of adjustable weights [21].

To mitigate this issue, Shallow Convolutional Neural Networks (SCNNs) have been proposed as an alternative to deeper networks [22]. SCNNs consist of a limited number of convolutional layers, which reduces the number of trainable parameters while maintaining the ability to capture meaningful information. SCNNs operate by applying convolutional filters to the EEG signal to learn spatial and/or features, followed by pooling layers that reduce dimensionality while preserving key information. The extracted features are then passed through fully connected layers for classification. Studies have shown that SCNNs can achieve comparable or even superior performance to traditional feature extraction methods like CSP when applied to MI/ME classification. Furthermore, their ability to generalise across subjects makes them a promising approach for real-time BCI applications [22].

1.2.6 Ridge Regression Models

Ridge Regression [23] is a regularised linear regression technique that incorporates an L2 penalty term to reduce overfitting and improve generalisation, particularly when working with high-dimensional and noisy data. Unlike classification-based methods that predict discrete movement states (e.g., left movement, right movement, or rest), Ridge Regression allows for the estimation of continuous control signals by learning a direct mapping between EEG features and output variables. This makes it particularly suitable for applications such as brain-controlled robotics and driving simulations.

1.3 Research aim

Through this research, we aim to assess the efficacy of using BCI in dynamic environments without predefined cues and delve into the methods that would make the development of such a BCI possible by answering the question:

'Can a motor execution Brain-Computer Interface achieve robust continuous control in a driving task?'

In order to answer this research question, we developed an experimental paradigm that requires real-time control: driving a simulated vehicle around a track. This provides an environment where the navigation of the vehicle is determined by exact timing, direction and duration of control over a BCI. Such a scenario will provide much more realistic challenges in the implementation of a BCI and the analysis of brain activity during the task.

A fundamental challenge in developing such BCIs lies in a paradox common to data-driven ML and DL approaches: to build a reliable model, high-quality labelled data is essential. However, in the early stages of BCI development, such data is either not available or of poor quality due to low classification performance. Poor performance in early sessions can also lead to user frustration, which might further decrease the quality of the data collected.

To address this, we propose a novel paradigm for BCI data collection and model development using ME. With this approach, we employ muscle activity, captured via electromyography (EMG), to train a highly accurate model. This EMG classifier substitutes for the BCI during the first experimental sessions, enabling participants to perform the task effectively and without frustration.

This approach offers two key benefits. First, it provides a reliable and interpretable ground truth for subsequent neural analysis, as muscle activity during ME is well-defined and easily measurable. Second, the resulting dataset, comprising both brain signals and corresponding motor outputs, can be used to train and refine models that aim to achieve true BCI-based control. In this way, EMG-guided data collection can be used as a foundational step that might support the eventual development of robust BCIs capable of continuous control in dynamic environments.

2 Methods

The experimental session consists of a calibration and driving task. The EMG data collected during calibration is used to train a classifier, an ML model that can extract right, left and rest commands based on the EMG data. This EMG classifier is then used to operate a simulated vehicle in a driving task based on real-time EMG input. Both EMG and EEG data are collected simultaneously to form the driving dataset. A visualization of this experimental design can be found in Figure 1.

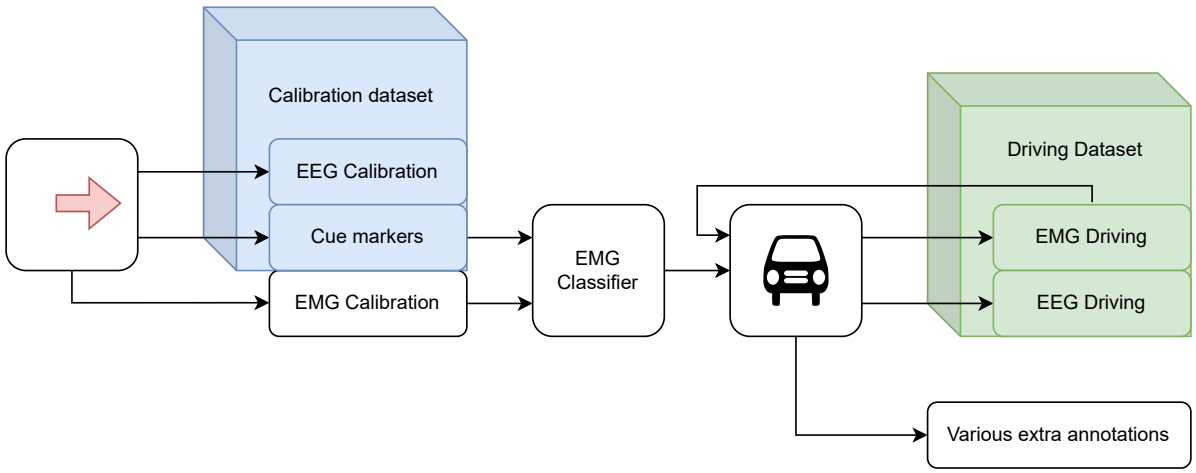


Figure 1: Design of the data acquisition. First participants do a calibration session following the Graz-BCI Motor Imagery paradigm. The EMG data collected with this is used to develop an EMG classifier, which the users then use to control a simulated car. The EMG, EEG and markers are recorded in both sessions resulting in a driving dataset and a calibration dataset.

This study was conducted with 20 healthy participants aged 19-45 years old ($\mu = 26, \sigma^2 = 40$). The data from one participant was excluded from the analysis because a section of the driving was not recorded properly. None of the participants had prior experience using a BCI.
EMG, EOG and EEG signals were recorded using the Biosemi ActiveTwo. 32 EEG channels were collected following the 10-20 system. Two monopolar EMG channels were measured on each forearm to correspond with wrist flexion and wrist extension, measuring the flexor carpi radialis and the extensor carpi radialis longus in each arm, resulting in four total EMG channels. Four EOG channels were measured to capture horizontal and vertical eye movements. All 40 channels were recorded at 2048Hz. Participants were seated in a chair facing a computer with their arms resting on the desk in front of them at a comfortable height. A towel was placed under the forearms to elevate the wrists slightly which allowed for easy and free movement of the wrist without putting tension on the muscles.

Whenever participants needed to perform a 'left' command they flexed their left hand inward. The 'right' command corresponded with flexing the right hand inward. In all other cases, they were instructed to relax their arms with their hands in a relaxed forward position.

2.1 Pre-Experiment Instructions and Questionnaire

Before the start of the data collection and experiment, participants were given an information letter outlining the procedure, as well as a questionnaire to control for handedness, their experience with

using BCIs and their gaming experience. These two documents can be found in the Appendix.

2.2 Resting State Data Collection

Before the calibration, the participant's resting state was recorded with eyes-open/eyes-closed (EO/EC) during resting conditions for two minutes per state. This is done in order to establish a baseline for an individual's neural activity by measuring alpha activity in a resting brain state. This data is available in the dataset but was not used for further analysis within the scope of this project.

2.3 Calibration paradigm

The participants performed a calibration task based on the Graz-BCI calibration paradigm [24]. The calibration consisted of 2x20 shuffled trials of left and right wrist motor execution. The calibration started off with 30 seconds of preparation. Each of the trials consisted of a 3-second preparation cue consisting of a cross, a $t = 1.25s$ cue with either a left- or right-pointing arrow, a $t = 3.75s$ execution cue consisting of a cross and a resting cue consisting of a blank screen. The participants were instructed to pay attention at the cross, remember the direction of the arrow and, once the arrow disappeared, perform the movement instructed by the cue. Participants were instructed to only perform wrist flexion after the cue disappeared to reduce their response time and to allow for a possible preparation effect. We are likely observing a similar effect in the driving task, where participants can see the path ahead and prepare for any necessary movements.
After the execution cross disappeared, the participants were instructed to return their wrists to resting position. A schematic overview of the different cues and movements can be found in Figure 2.

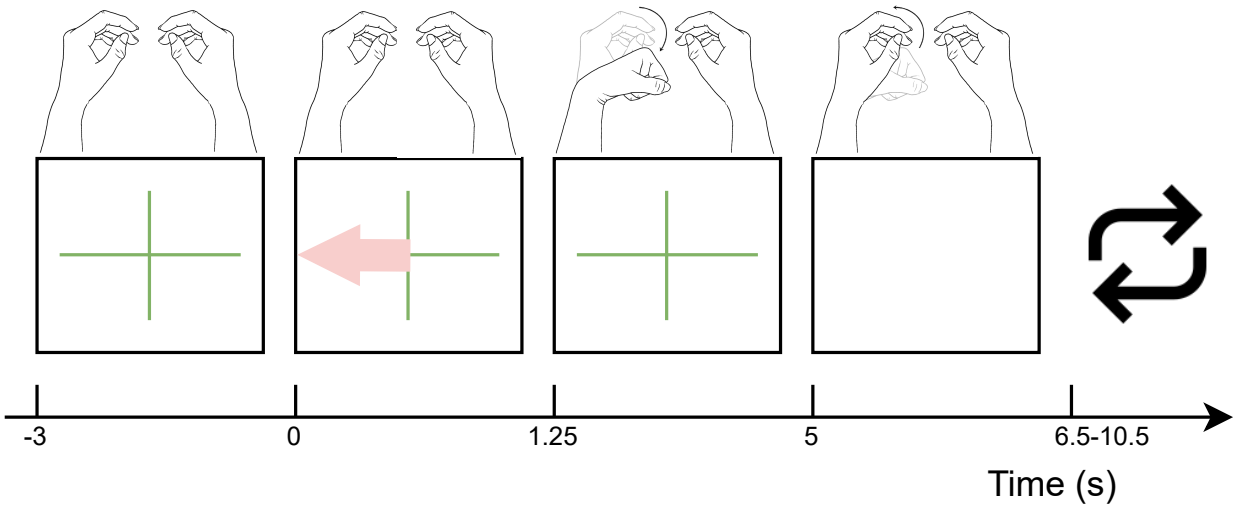


Figure 2: The cues as shown in OpenVibe, their timing in seconds and the expected movements for the calibration paradigm.

2.4 EMG classification

The recorded EMG data from the calibration paradigm was used to train an online EMG classifier to distinguish left wrist flexion, right wrist flexion and rest conditions. This classifier will be used in the experiment to steer a simulated car. This is not a true BCI as it does not interpret brain activity

but rather muscle activity. However, since participants do provide real-time, active control during the driving sessions, they might find the experience similar to using an actual BCI. The EEG data collected throughout this experiment is therefore only used for offline analysis.

The EMG of the 4 channels measured in the calibration is re-referenced using Common Average Reverence (CAR) and averaged. A band-pass filter between 30-500 Hz is applied and notch-filtered at 50 Hz using causal 4th-order Butterworth filters. Since the online classification during the driving session should have low latency to feel as realistic as possible, the trials are cropped to the middle 200ms of each movement $t = 6.125s$ to $t = 6.145s$. The mean powers of the four EMG channels are used as separate features for a Linear Discriminant Analysis (LDA) classifier. Through 10-fold cross-validation, we find a mean per-subject EMG classification performance of 94% with a standard deviation of 5% during pilot studies.

2.5 Experimental paradigm

The control task involves driving around a simulated car in a virtual environment by navigation using left flexion, right flexion and rest for steering left, right and straight respectively. Acceleration and deceleration are not controllable by the participants and are therefore set at a constant forward speed.

The experimental environment for the task was completely created in Unity [25]. The decision to opt for a 3D design as this is closest to any real-life application. The environment does introduce extra challenges, such as eye movement artefacts when the participants plan their route by looking ahead on the track, but these will also be present in any actual implementation of such a BCI. Creating a system that is robust enough to deal with these challenges is vital for drawing any conclusions on the viability of such a system.

The 3D models of the car and the road, as well as the wheel physics, are based on an existing racing game [26].

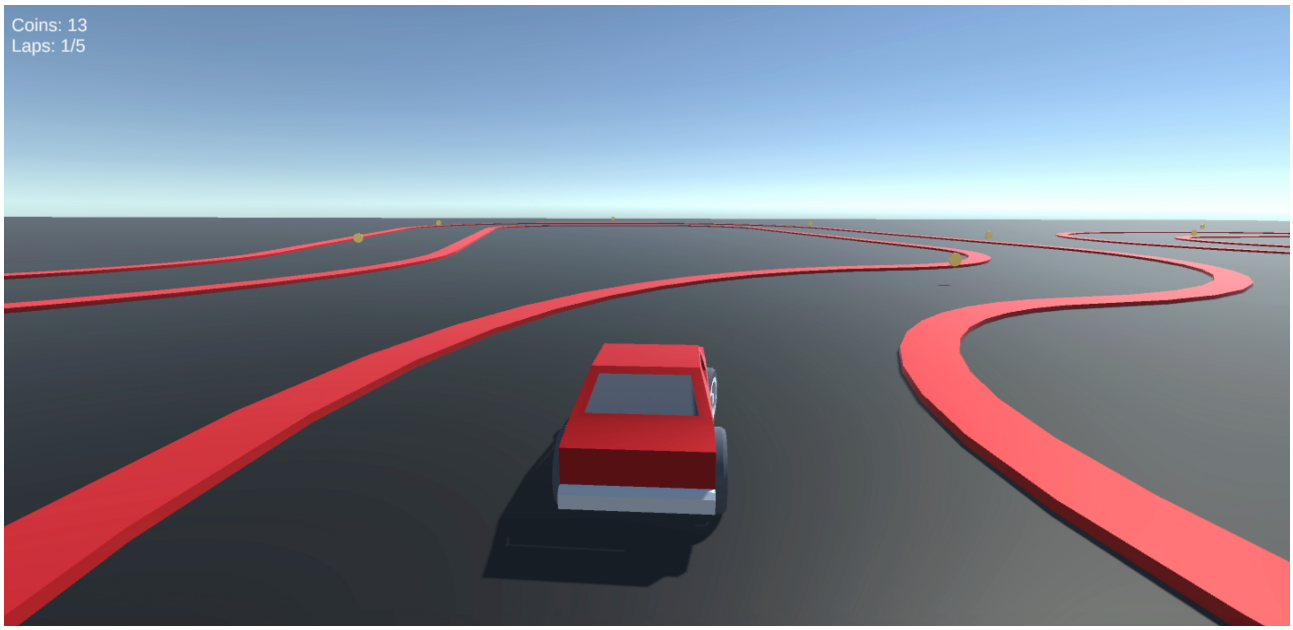


Figure 3: The participant's point of view during the driving task.

2.5.1 Racetrack

A road or racetrack was designed to create more consistency between the experiments among participants in comparison to a randomised route. The shape of the racetrack was designed in order to extract control data over a large number of instances for all three conditions: left turns for left wrist flexion, right turns for right wrist flexion and straight parts for rest. A top-down view of the track can be found in Figure 4. The participants drove a total of five laps per trial and performed three trials per session. The racetrack remained the same throughout all trials for all participants.

The outer bounds of the track only exist to provide guidance to the participant on where to go and will not provide any physical barrier for the participant to steer off-track. Due to the constant forward speed throughout each trial, introducing a hard barrier might create issues with returning back to the course given all the sharp curves included in the track. As we aimed to reduce frustration during the experiment, we ended up not using any hard barriers. The trial was restarted if any participant found themselves in a situation where driving back to the road proved difficult. These tracks were excluded from the analysis.

2.5.2 Navigating the Track

The car's motion is controlled using Unity's `WheelCollider` components, which simulate realistic vehicle physics by applying motor torque for propulsion and steering angles for direction control.

Engine Force and Acceleration

At the start of each trial, a countdown was shown on screen. After the countdown, the car automatically accelerates up to a maximum speed, at which it will stay for the rest of the trial. The experiment was designed to only look at two ME movements for steering to reduce complexity, the forward acceleration was therefore not controlled by the participant. The constant forward speed was determined with pilot studies to be appropriate for both novel as well as more experienced drivers/gamers.
The vehicle is propelled forward by applying torque to the wheels. The torque applied to each wheel is given as follows:

$$\tau = F_{\text{engine}} \cdot v \cdot f_v \tag{1}$$

Where:

- τ is the torque applied to each wheel,
- $F_{\text{engine}} = 900$ is the base engine force,
- v is the vertical input,
- $f_v = 1.3$ is a tuning multiplier to adjust the final torque.

Since the participants do not actually control the car speed themselves, we set $v = 1$, which yields a forward torque as shown below:

$$\tau = 900 \cdot 1 \cdot 1.3 = 1170 \tag{2}$$

Steering through EMG

The EMG classification results were then used to navigate the racetrack, using left wrist flexion, right wrist flexion and rest for steering left, right and straight respectively. Participants were instructed to stay on the track as much as possible and favour longer continuous movements over smaller corrections. This instruction was given with the later analysis in mind, as small corrections are difficult to distinguish from erroneous classifications from the EMG classifier and in some of our further analysis these will therefore be excluded from the training data.

Steering is applied to only the two front wheels to mimic most 'normal' cars. The steering angle is determined by:

$$\theta = f_h \cdot h \cdot \alpha \tag{3}$$

Where:

- θ is the final steer angle,
- $f_h = 0.4$ is a scaling factor tweaked during pilot studies,
- h is the horizontal input from the EMG classifier (ranging from -1 to 1),
- $\alpha = 30^\circ$ is the maximum steering angle.

The maximum steer angle applied during a full turn (i.e., $h = \pm 1$) is then:

$$\theta_{\max} = 0.4 \cdot 1 \cdot 30^\circ = 12^\circ \tag{4}$$

To maintain visual realism, the wheels are updated in each frame using Unity's `GetWorldPose` function. This synchronises the position and rotation of the wheel models with their corresponding in-game shapes.

Task Engagement and Feedback

Small rotating coins were added to the environment that (re)spawned in predefined areas at the start of every lap, and have a randomised chance to spawn in either the centre of the lane, slightly to the left of the centre or slightly to the right of the centre as can be seen in Figure 5. The coins were introduced to increase engagement throughout the experiment, as well as to require the participants to plan their steering trajectory slightly differently every round. However, we did not want to incentivise frequently changing the steering path with small corrections in order to collect the coins. Therefore, the three spawning locations are relatively close together.
A user interface was added to provide the participant with the amount of laps completed and the amount of coins connected. At the end of the experiment, the car would automatically stop and a pop-up with the text 'Finished' appeared on screen.

Information Sharing and Recording through Lab Streaming Layer

The online classification of the EMG classifier is retrieved through Lab Streaming Layer (LSL), an open-source software that is designed to help in the synchronisation and recording of time series data [27]. The EMG classifier calculates the participant's steering attempt every 200 ms based on the input

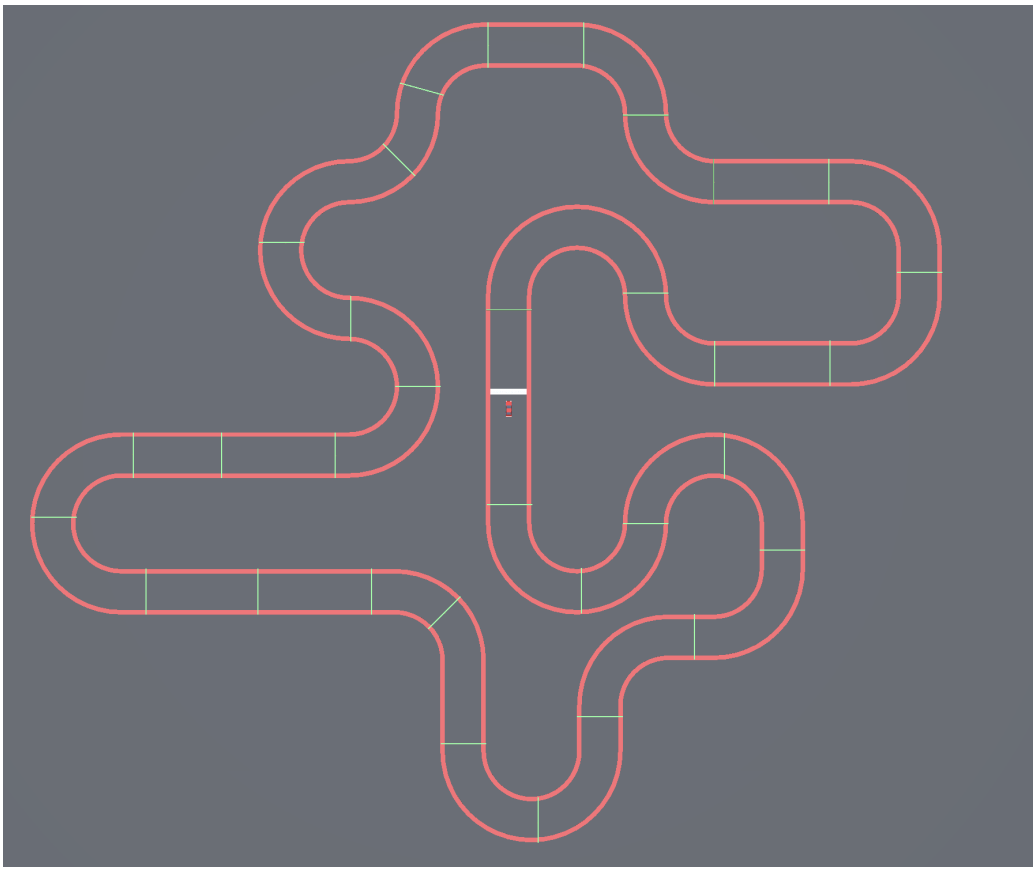


Figure 4: A top-down view of the driving track showing the bounds of the track in red, the starting line in white and the coin spawning areas indicated in green. The latter is not visible during the driving experiment itself.

EMG for left flexion, rest and right flexion which generate a value of 1, 0 or -1 respectively and 'share' this information with Unity through an LSL stream. Functions within the Unity environment are then used to read this stream and interpret these values into rotational translations for the two front wheels of the simulated vehicle. LSL was also used to record additional information about each experimental session and synchronise this information with the EEG recording. Timestamps were made every time a participant crossed the finish line, finished a trial and collected a coin. Furthermore, the location of the car on the track was collected every new frame, in order to be able to reconstruct the driven route for further analysis.

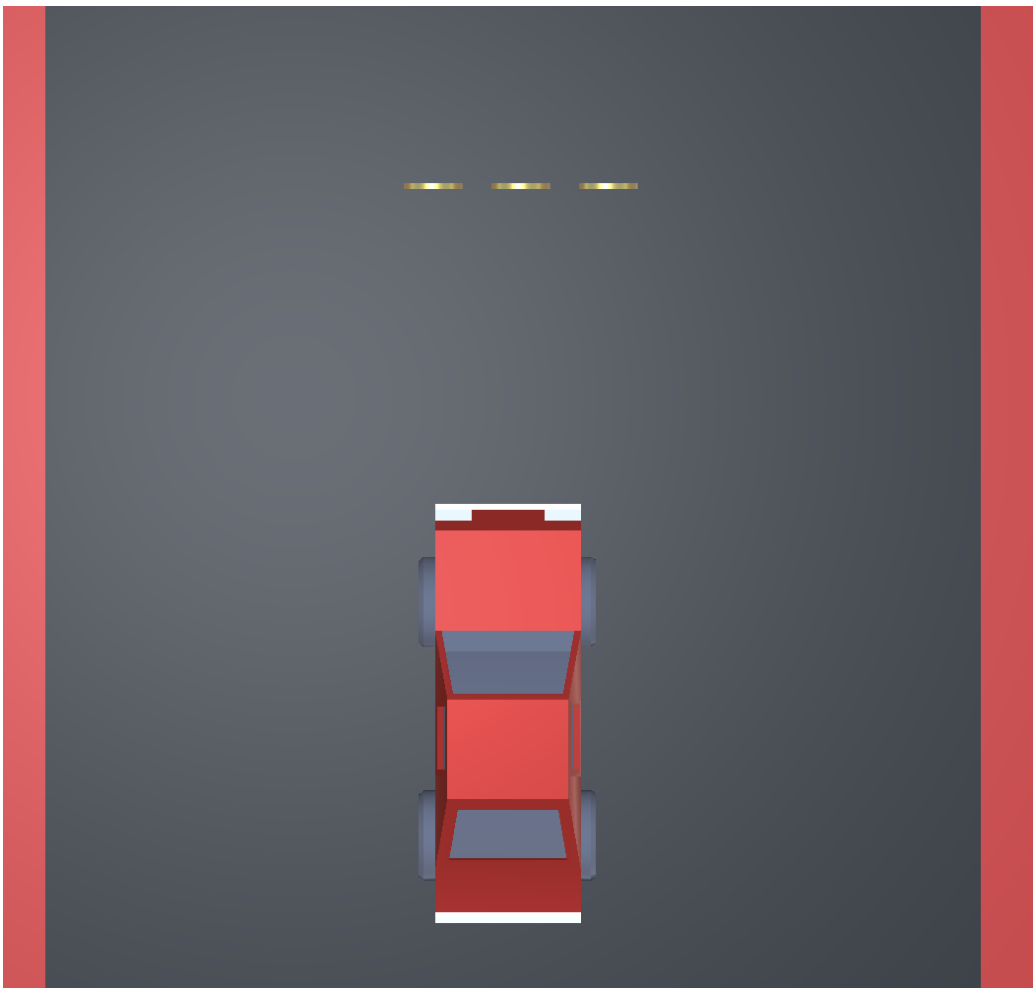


Figure 5: A top-down view of the three different possible positions for the coin. The coin will randomly spawn in only one of these positions.

3 Data Analysis

The data collected throughout the experiment was further analysed using different approaches. The most obvious approach is to use the EMG classification results from the classifier used in steering as a ground truth for training a classifier to learn features from EEG data in order to classify the motor execution into left flexion, right flexion and rest by segmenting the collected data in epochs that contain the same classification result. With this approach, we attempted training both a DL as well as an ML method in order to evaluate the strengths and caveats of both methods.

However, this first approach comes with certain limitations.

First, this method relies on segmentation of the driving data to train and test the classifier. While this allows for structured analysis, it does not accurately reflect how the classifier would perform in a real-time, continuous driving scenario. In segmenting the data, we only consider longer, stable periods where the classification remains unchanged, filtering out micro-adjustments in steering. However, in an actual driving environment, small corrective movements are essential for maintaining precise control, and our approach currently does not account for these finer adjustments.

Second, using a discrete classifier introduces an all-or-nothing distinction between the three movement classes (left, right, and rest). This rigid classification does not allow for gradual transitions or varying intensities of movement, which are crucial in a real-world setting. For instance, slightly steering left and sharply steering left would be treated as the same class, despite their possibly differently intended effects on driving dynamics.

To address these limitations, we explored an alternative approach using regression (Ridge Regression), which allows for a more continuous representation of motor intent, better capturing the nuances of driving control.

3.1 Classification Task

Our first attempt for data analysis involves using the classification results from the EMG classifier used to steer the car in the experiment as ground truth for a classification BCI that can classify between three distinct classes in the data.

For this, we aim to investigate the differences between DL and ML approaches by comparing two different classifiers: a Shallow Convolutional Neural Network (SCNN) and a Logistic Regression model using Common Spatial Patterns (CSP-LR).

3.1.1 Preprocessing

All EEG and EMG signals were preprocessed using MNE-Python to remove noise. Causal, forward-phase filters were used to preserve the temporal dynamics of the data. EMG signals were notch-filtered at 50 Hz up to 500 Hz to remove line noise and bandpass-filtered between 30–500 Hz to capture relevant muscle activity while suppressing noise. EEG signals were bandpass-filtered between 5 and 35 Hz to isolate sensorimotor rhythms.

EEG recordings are often contaminated by non-neural artefacts, which can come from various sources, including eye blinks or movement, muscle activity or line noise. These unwanted signals can obscure neural activity. To address this issue, Independent Component Analysis (ICA) [28] is commonly used for artefact removal. ICA is a blind source separation technique that decomposes the EEG signal into statistically independent components, allowing researchers to identify and exclude these artefacts from analysis whilst preserving neural activity.

For this project, MNE [29] ICA with Infomax decomposition algorithm was used to automatically detect and exclude components associated with muscle activity, eye movements, cardiac signals, line noise, and channel noise. Only EEG channels were included in the decomposition, and ICA was fitted using 32 components.

Epochs were extracted from both the calibration and driving data to segment the continuous EEG data into trials corresponding to distinct motor states: left wrist flexion, right wrist flexion, and rest. Due to differences in how labels were generated in each condition, two separate epoching strategies were applied.

During the calibration paradigm, cues were given to the participant indicating left-hand and right-hand movement. The cue timings were used to define trial onsets. Movement epochs for left and right wrist flexion were extracted from 1.25 to 5 seconds after cue onset.

Resting-state epochs were derived from markers placed at the conclusion of each movement trial. To avoid contamination from residual movement or visual feedback, rest epochs started 0.5 seconds after the marker and extended for 3.75 seconds to match with the left and right wrist flexion epochs. The resulting epochs were time-shifted in metadata (without altering the data) to maintain consistency with movement epochs. All epochs were constructed using MNE's Epochs.

For the driving condition, explicit trial markers were not present, as the task was performed continuously. Instead, EMG-based classifier outputs were used to infer periods of movement and rest. Time stamps and predicted labels from the EMG stream were aligned with the EEG recording and any predicted label (left, right or rest) that persisted for at least 19 consecutive EMG classifier outputs was considered a valid movement trial. Given the EMG classification every 200 ms, this roughly aligns with the length of epochs in the calibration paradigm with 19 $epochs * 200ms = 3.8s$. 200 ms were subtracted from movement onset in order to offset the EMG classification frequency.

Driving epochs were extracted from 0.5 to 3.25 seconds after each inferred movement onset, resulting in 2.75-second trials.

3.1.2 Shallow Convolutional Neural Network

A shallow convolutional neural network (SCNN) was designed in order to extract features from the neural data and classify the EEG signals into three distinct classes: left-wrist flexion, right-wrist flexion, and rest. The architecture was inspired by the principles of filter bank common spatial patterns (FBCSP) [22], focusing on learning time-frequency patterns across EEG channels. The SCNN was implemented in TensorFlow and contains both a temporal convolutional layer followed by a spatial convolutional layer in order to 'mimic' the type of feature extraction one would usually get through bandpass filters and CSPs.

All in all, this model followed the following architecture:

- Input Layer
- Temporal Convolutional Layer, which applies convolution along the temporal dimension using 40 filters of size 25, allowing the network to capture temporal features associated with motor-related activity.
- Spatial Convolutional Layer, which learns spatial patterns across EEG channels with 40 filters, enhancing the network's ability to identify inter-channel relationships.
- Batch Normalization that normalises activations for stability during training and improved generalization.
- Non-linearity, a square activation function was applied to enhance feature separability by emphasizing signal magnitudes.
- Pooling Layer, a mean pooling operation with a window size of 3 and a stride of 1. A logarithmic activation was used to compress the feature representation.,
- Dropout layer. A dropout rate of 50% was used to reduce overfitting by randomly deactivating neurons during training.
- Global average pooling to reduce the spatial and temporal dimension.
- Dense output layer using a softmax activation function to produce class probabilities for the three target classes.

Sparse Categorical Cross-Entropy as a loss function with an Adam Optimizer [30] were used, since this is well suited for a multi-class classification task with mutually exclusive labels.

3.1.3 Logistic Regression with Common Spatial Patterns

In our second model, a Common Spatial Pattern (CSP) transformation was applied before training a Logistic Regression (LR) classifier.

Unlike traditional binary CSP, which optimises spatial filters for two-class problems by maximising variance differences between two conditions, we utilised multiclass CSP. Multiclass CSP extends the method to handle more than two classes by applying one-vs-rest (OVR) decomposition or simultaneous decomposition strategies. In our implementation, multiclass CSP learned spatial filters for three movement conditions: left wrist flexion, right wrist flexion, and rest.

CSP Transformation was performed where the EEG signals were projected onto a set of spatial filters using CSP from the MNE Python Library [29]. The CSP transformation extracted six components, applied shrinkage regularisation, and included a log transformation. A standard logistic regression model was used for classification. Logistic regression was chosen due to its robustness in handling binary classification tasks. The classifier was evaluated using 5-fold cross-validation.

3.1.4 Evaluation

For both models, we evaluate their performance using F1-scores and confusion matrices. Furthermore, we analyse a visual representation of the learned features using a topographical representation of the learned features for the CSP-LR model.

To evaluate model performance across different tasks within the experiment, three experimental conditions were defined. In the calibration condition, models were trained and tested on the same dataset, providing a baseline measure of performance under segmented conditions. The driving condition involved training and testing on data collected during the continuous control task, reflecting the model's ability to learn ME features in a dynamic environment. Finally, the transfer condition tested the model's capacity to generalise by training on calibration data and evaluating on driving data. This last condition mimics most how a real-world application of a continuous control BCI would, in ideal terms, be trained. In the transfer condition, the first 10% of driving data was used to create a customised threshold per participant for differentiation between rest and left/right conditions.

3.2 Regression Task

In this second approach, we aim to predict the continuous EMG activation directly from the EEG data, as shown successfully in prior research [31].

3.2.1 Regression Model

For this task, the EMG signal was extracted as a continuous variable rather than as a discrete class label. This was achieved by computing the envelope of the absolute EMG signal for each arm. The EMG envelope was computed using bandpass filtering, Hilbert transformation, smoothing, and standardization. The filtering step applied a bandpass filter (100-125 Hz) to remove noise and retain movement-related EMG activity. The Hilbert Transform was then used to extract the instantaneous amplitude of the filtered EMG signal. The resulting signal was further smoothed using a convolution window. Finally, z-score normalization was applied to ensure that the resulting envelope signal was standardized across trials and participants. This approach was based on prior research [32].

To construct a single continuous signal representing movement, the envelope of the right arm's EMG signal was negated and combined with the envelope of the left arm's EMG signal. This transformation resulted in a signal where positive values correspond to left wrist flexion, values near zero indicate a resting state, and negative values correspond to right wrist flexion. This approach captures the relative intensity of muscle activation in both arms while maintaining a balanced representation of rest and movement states.

To extract relevant EEG features, data was selected from the 'area of interest': the C3 and C4 electrode locations, along with their adjacent electrodes as these locations correspond to sensorimotor cortex activity involved in wrist movement. The EEG data was subsampled from $fs = 2048Hz$ to $fs = 256Hz$ to reduce computational complexity. Features were computed for the mu (8-12 Hz) and beta (12-30 Hz) frequency bands.

A sliding window with a window size of $200ms$ was selected which allows the model to incorporate past EEG activity when making predictions, rather than treating each time point independently. Only

past neural information was used as this would be employed in an online setting for the model as well.

The extracted EEG features were used to train a Ridge Regression model, which was selected due to its ability to handle high-dimensional feature spaces and reduce the risk of overfitting through L2 regularization. The model's performance was evaluated using Mean Squared Error (MSE), which quantifies the difference between the predicted and actual EMG envelope values. A baseline model was also trained, where Ridge Regression was applied using no EEG features, predicting only based on the statistical properties of the EMG envelope signal. This baseline model ensures that the performance of the EEG-based model was not inflated due to imbalances in the training data.

4 Results

During this research, three different models were developed to analyse dynamic EEG-based BCI data across two tasks: a classification task, in which motor intentions were decoded as discrete labels (left wrist flexion, right wrist flexion, or rest), and a regression task, aimed at continuously predicting muscle activation from EEG.
Each classification model was tested in three experimental configurations: within the calibration data (cross-validation), within the driving data (continuous control), and in a transfer setup, where models trained on calibration data were tested on driving data. The regression task focused specifically on the feasibility of predicting the EMG envelope as a continuous signal from EEG features during the driving task.

4.1 Classification Task: SCNN and CSP-LR

4.1.1 Confusion Matrices

Figure 6 shows confusion matrices comparing the classification performance of the SCNN in the left column and the CSP-LR in the right column across three conditions: calibration-only, driving-only, and transfer (training on calibration, testing on driving data). Overall, both models performed best when trained and tested on the same data type (either calibration or driving), with notable performance drops in the transfer setting.

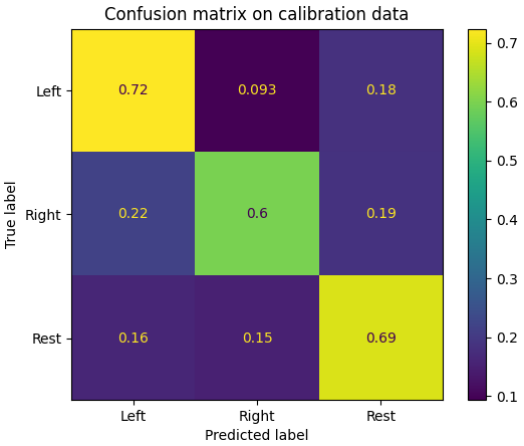
When trained and tested on the calibration data, the SCNN model performed well for left flexion (0.72) and moderately for rest (0.49), but showed confusion in distinguishing right wrist flexion movements. The CSP-LR model, in contrast, showed weaker but more balanced performance across all three classes, with more confusion between movement and rest labels. On the driving data, both models improved, particularly the SCNN, which achieved high classification accuracy for Left (0.89) and Rest (0.71). Predictions for right flexion remained the most difficult for both models, with both frequently misclassifying as either left wrist flexion or rest.

We observed the most striking difference in the transfer condition. When trained on calibration data and tested on driving data, the SCNN model showed a marked tendency to predict the rest class across all inputs, with 64% of left flexions and 63% of right flexions misclassified as rest. This clearly shows a difficulty in generalising movement-related patterns learned during calibration to the more dynamic driving context. Although the overall diagonal values (e.g. rest: 0.69) might appear moderately high, the model's inability to distinguish between movement classes makes it unsuitable for practical BCI control in this condition. The CSP-LR model, while also performing worse in the transfer condition, retained more balanced predictions across all classes (left flexion: 0.49, right flexion: 0.45 and rest: 0.65), suggesting it was better at preserving class discrimination during the shift in the domain.

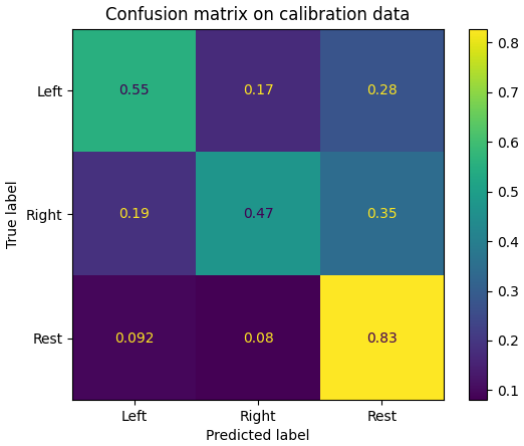
These results highlight that while deep learning models like the SCNN can perform well when trained and tested within the same context, they may be more prone to class bias in real-world transfer scenarios.

4.1.2 F1-Scores

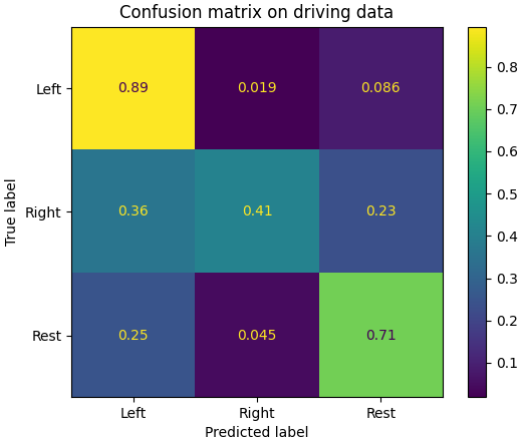
Figure 7 shows the Per-Subject F1-scores across the three different conditions. Overall, the SCNN model exhibited quite high and consistent performance across participants in both the calibration and



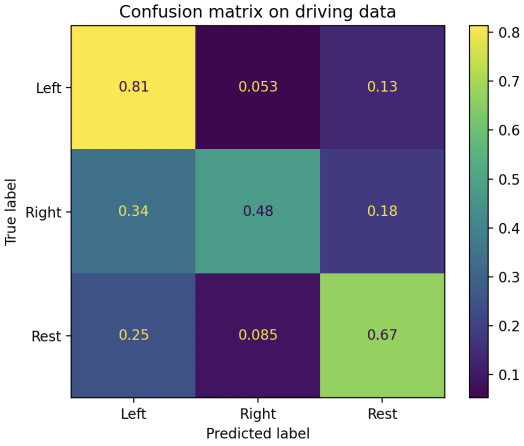
(a) Confusion matrix showing the SCNN model performance trained and tested on the calibration data.



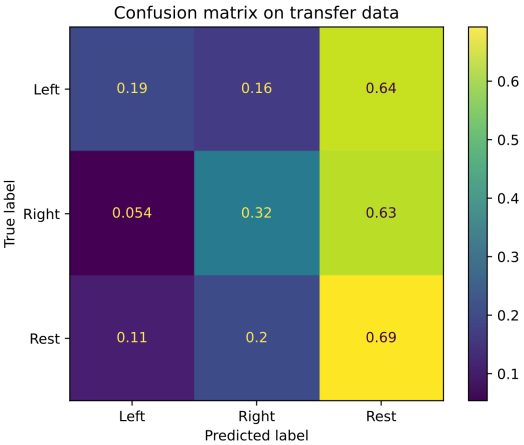
(b) Confusion matrix showing the CSP-LR model performance trained and tested on the calibration data.



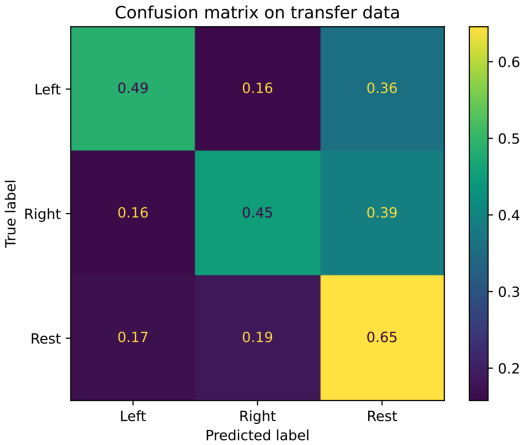
(c) Confusion matrix showing the SCNN model performance trained and tested on the driving data.



(d) Confusion matrix showing the CSP-LR model performance trained and tested on the driving data.

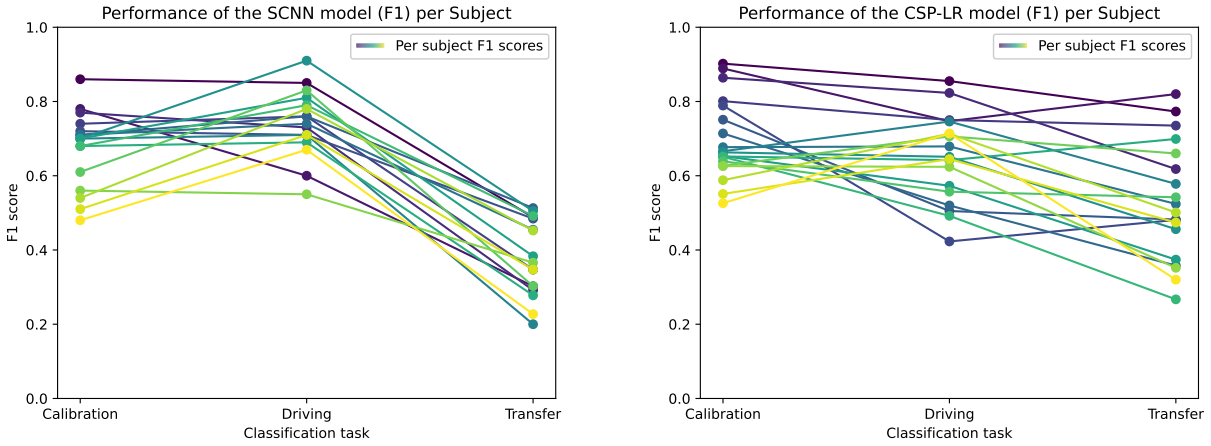


(e) Confusion matrix showing the SCNN model performance trained on the calibration data and tested on the driving data.



(f) Confusion matrix showing the CSP-LR model performance trained on the calibration data and tested on the driving data.

Figure 6: Confusion matrices for both the SCNN and CSP-LR models in different experimental configurations



(a) Per-Subject F1-scores for the SCNN model

(b) Per-Subject F1-scores for the CSP-LR model

Figure 7: Per-Subject F1-scores for the SCNN and CSP-LR models across three conditions: calibration, driving and transfer.

driving tasks. Many subjects achieved F1 scores above 0.7, with several exceeding 0.8, particularly during the driving task. This suggests that the SCNN was not only effective when trained and tested on controlled calibration data, but also adapted well to the dynamic environment of the driving task.

The CSP-LR model showed similar if not better results in the calibration task but performed generally slightly worse in the driving condition.

The difference between models became more pronounced in the transfer condition, where both approaches showed a worse performance compared to the previous conditions, but the CSP-LR model retained moderate performance for some participants, whereas the SCNN model showed a substantial drop in F1 scores across nearly all subjects. From the confusion matrices in the previous section, we know that the model mostly predicts rest here across all data, which might explain this dip in performance.

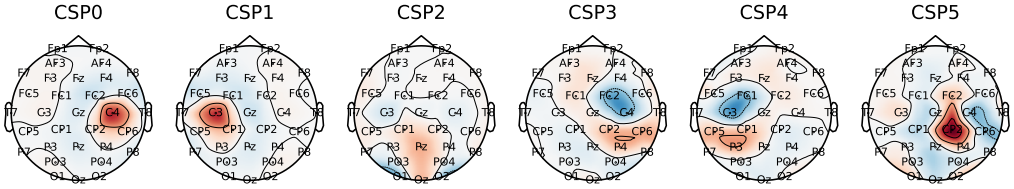
4.1.3 Spatial Features

Figure 8 shows the topographic distributions of the first six CSP components learned during the calibration task (Figure 8a) and the driving task (Figure 8b). These spatial patterns represent the most discriminative features extracted by the CSP algorithm for separating the three motor classes: left flexion, right flexion, and rest.

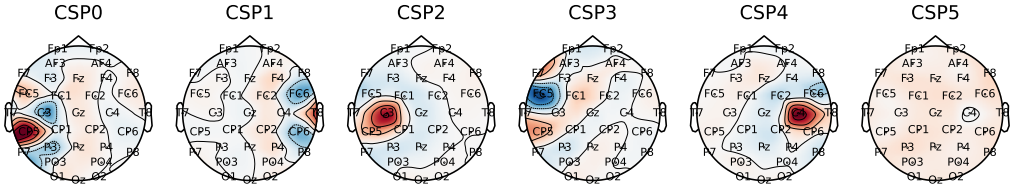
During calibration, the learned filters show strong and well-localised activation over the sensorimotor cortex, particularly around electrodes C3 and C4, which correspond to the left and right motor cortices. CSP0 and CSP1 show clear lateralisation, which indicates differentiation between left and right wrist flexion. CSP2 and CSP5 show more central or parietal activation, which can reflect activity related to the rest condition or general motor inhibition.

In contrast, we can see that the CSP filters learned from the driving task data appear more distributed and less symmetrical. The spatial patterns in CSP2 and CSP4 further reflect again a distinction be-

tween left and right wrist flexion shown in the activation over C3 and C4. However, the learned features appear not only in the sensorimotor cortex but also parietal and frontal regions. For example, CSP1 and CSP3 show more diffuse patterns involving CP5, CP6, and FC6, suggesting additional cognitive or sensorimotor activation given the more complex task.



(a) Learned CSP filters during calibration



(b) Learned CSP filters during driving

Figure 8: Learned CSP filters for both calibration (a) and driving (b) conditions.

4.2 Regression Task: RR

Figure 9 shows the calculated EMG envelope for a portion of the driving data for one participant. The colouring of the background indicates the class labels of the EMG LDA classifier used in the driving task. We can clearly see that the EMG classifier generally matches the trends in the EMG envelope, where a positive EMG envelope matches with a left classification, a negative envelope matches with a right classification and a near-zero envelope matches with a rest classification.

This figure also highlights a notable limitation of using the EMG classifier output as ground truth. Specifically, there is often a slight temporal delay between the onset of muscle activation and the corresponding classifier response. This lag is due to the frequency at which the classifier makes predictions.

Additionally, Figure 9 includes the predicted EMG envelope generated by the Ridge Regression model based on the features selected from EEG data. Visual inspection reveals that the model does not accurately capture the EMG activation patterns. This observation is supported by the comparison of mean squared error (MSE) between the EEG-based regression model (MSE $= 2.32$) and a baseline dummy regression model (MSE $= 2.47$). The dummy model predicts an EMG envelope of 0 each time, given the way the EMG data is normalised. The performance of the RR model, only slightly higher in MSE, shows that the model fails to reliably predict muscle activity from EEG features in this setting.

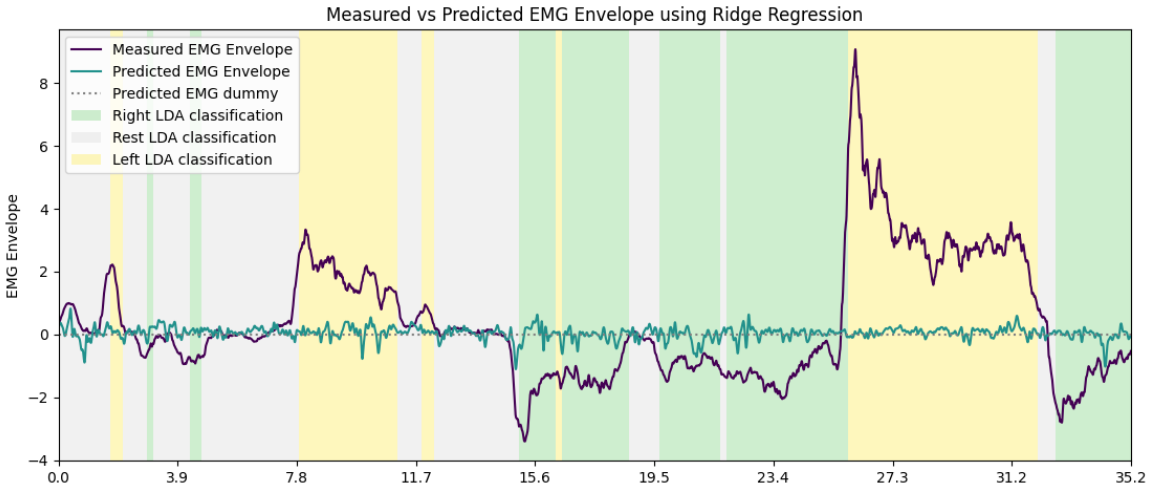


Figure 9: Figure showing the measured EMG envelope, the predicted EMG from EEG data by the RR model and the predicted EMG from a dummy model without learned features. The coloured segments in the background indicate the LDA classification results from the EMG classifier used to steer the car.

5 Discussion

This study set out to evaluate the feasibility of using Brain-Computer Interfaces (BCIs) for robust, continuous control in dynamic environments without predefined cues. By designing a simulated driving task that required dynamic real-time control, we introduced an unconventional paradigm for assessing BCI performance under conditions that more closely resemble practical applications.

We proposed a novel hybrid approach in which a muscle-based EMG classifier is used to bootstrap the BCI training process. This method enabled data collection with limited frustration and produced clean, well-labelled datasets of both neural and muscular activity. The resulting dataset provided a valuable ground truth for EEG analysis and served as a foundation for training multiple BCI models aimed at capturing motor execution signals.

Our findings demonstrate that deep learning models, such as the Shallow Convolutional Neural Network (SCNN), perform well when trained and tested within the same task context, which was shown by the high classification accuracy in both the calibration and driving conditions. However, the SCNN showed considerable degradation in the transfer condition, where models trained on calibration data were tested on driving data, indicating limited generalisation across task contexts. Conversely, while the CSP-LR model showed lower overall accuracy, it retained greater robustness under transfer, suggesting that traditional pipelines may offer more stable performance in cross-domain applications, at least given the training capabilities given the limited data available.

Generating a spatial representation of the SCNN model's learned features might provide valuable insights into the differences between the two approaches. In future research, this might be achieved by employing SHapley Additive exPlanations (SHAP) [33]. SHAP is a model-agnostic interpretability method based on cooperative game theory that assigns each input feature an importance value for a given prediction.

In addition to classifying discrete movement categories, this study explored the feasibility of using EEG to predict continuous EMG activation during a dynamic BCI task. However, results from the regression model trained to predict EMG activity from EEG data indicated poor performance. These findings suggest that the EEG features extracted in this study were insufficient to support accurate continuous prediction of EMG activation under the conditions tested. Interestingly, this outcome contrasts with prior work showing promising results for EMG envelope prediction from EEG [31]. Further research might bring to light some of the underlying issues that resulted in the given performance.

Together, these results highlight both the potentials and the limitations of current approaches to motor execution-based BCIs. Accurate and responsive control can be achieved under ideal conditions. However, ensuring robustness in dynamic environments remains a challenge but seems overall promising.

5.1 Limitations and Future Work

The current approach does not incorporate EEG-based neural signals for online control during the driving task. Instead, only EMG signals were used for steering. While this design simplifies control and ensures reliable performance, it does not allow for an assessment of how well the EEG model would function in a real-world scenario. Future experiments could explore shared control, where both EMG and EEG contribute to the control of the vehicle.

Additionally, shared control could address a current limitation: participants might unknowingly introduce artefacts by moving around or mind-wandering due to the lack of direct neural feedback. Because driving performance remains unchanged regardless of these factors, this could negatively impact subsequent neural signal analysis. In contrast, integrating EEG into control would provide participants with real-time feedback on the quality of their neural signals, making them more aware of the negative effects of artefacts and mental distractions.

One way to improve neural feedback without disrupting the driving experience could be through a non-intrusive visual indicator. For example, a coloured bar (green for alignment, red for mismatch) could be displayed to inform participants when their EEG classification diverges from their EMG classification. This would allow for passive feedback without interfering with the driving task, helping participants maintain neural focus without the need for explicit rewards or punishments. Another option would be to make the speed adjustable to the decodability of the neural signal. Participants would be rewarded for clear neural signals by driving slightly faster. It should be taken into consideration that this might make retrospective analysis more difficult.

Before implementing shared control, preliminary pilot studies could be conducted to determine the appropriate distribution of control weight between the EMG classifier and the EEG model. By optimising these weightings, we can maximise the benefits of both approaches, ensuring that EEG contributes meaningfully without introducing excessive noise or instability.

The participants in this study were novices in using BCIs, and research has shown that performance with BCI systems can improve with training [34, 35, 36]. In real-world applications, users would undergo multiple training sessions to refine their control over the BCI system, leading to better performance and increased adaptability. More training data could also be collected over time to fine-tune the model to individual users. Future research should investigate how performance improves over multiple sessions and whether this training effect facilitates better transferability between calibration and real-time control.

The current experiment relied on Motor Execution, meaning that participants physically moved their muscles while EEG data was recorded. However, the long-term goal of BCI development is to assist users who may lack the capacity for physical movement. While motor execution and motor imagery have been shown to produce similar patterns in EEG signals, there are notable differences in their cortical representations. To ensure the proposed models are viable for their intended application, future experiments should replicate the task using only MI or MA-based control. This would provide crucial insights into the feasibility of transitioning from an ME-based BCI to one that relies solely on MI signals.

One of the challenges identified in this study is the discrepancy between the calibration phase and the driving task, which likely contributes to reduced classification performance when transitioning from one to the other. The calibration task currently does not fully reflect the demands of continuous driving control, potentially leading to a mismatch in the neural representations between the two conditions. To address this, future research should explore alternative calibration paradigms that more closely resemble the driving experience, e.g. by incorporating gradual transitions instead of discrete trials, using dynamic visual stimuli, or providing more real-time feedback.

Another limitation of the current implementation is that it only decodes left and right steering commands. However, real-world continuous control BCIs often require a broader range of functions, such as deceleration and acceleration. To extend the capabilities of the system, future research should introduce additional motor movements corresponding to these commands. One potential approach is to use a similar paradigm where motor execution is first employed (with EMG providing a ground truth), allowing for robust model training before transitioning to a full MI-based system. This would enable a more comprehensive and practical BCI control scheme for driving and other applications.

Lastly, we observed significant variability in performance across participants. Some individuals demonstrated strong classifier/prediction performance, while others struggled to produce decodable signals. Future studies should explore how factors such as fatigue, cognitive load, and individual neurophysiological differences affect BCI performance and whether adaptive models are robust enough to minimise these inconsistencies [37].

5.2 Scientific Contributions

This research contributes to the field of Brain-Computer Interfaces by demonstrating the feasibility and challenges of motor execution-based BCIs for robust, continuous control in dynamic environments.

A key methodological contribution is the introduction of an EMG-guided data collection strategy. This method allows the collection of high-quality BCI-like data in the early stages of BCI development, which can be very helpful in the design of ML/DL-based BCI systems for tasks that have little to no prior data available to be trained on. The study also explored both classification and regression perspectives, highlighting the limitations and potential of EEG-based decoding of continuous muscle activation.

Preliminary findings were published and presented at the 9th International Graz Brain-Computer Interface Conference in 2024, contributing to community discourse on real-time BCI design and model generalisability.

Furthermore, in support of open science, the dataset collected in this study, which includes simultaneous EEG, EOG EMG data for EO/EC, calibration and driving paradigms, will be made publicly available for all participants who provided informed consent. By sharing both the experimental framework and associated multimodal data, this research can be used by any researcher for further analysis, whether as part of training data for transfer learning or large neural networks, or as a basis for exploring alternative approaches to decode the data aside from those tested in this study.

6 Acknowledgements

First and foremost, I would like to thank my supervisors dr. Andreea Sburlea and drs. Ivo de Jong for their invaluable guidance, sharp insights, and endless patience throughout this project. What began as a neatly scoped plan turned into a much more involved journey than anticipated, and I am grateful to everyone who stuck with it until the finish line. Special thanks for the unforgettable times we shared during the conference and our spontaneous detours throughout Austria.

I would also like to extend my appreciation to Rosa, who not only bravely volunteered as the pilot participant but quite literally let me pick her brain for this project (with electrodes, of course). Your enthusiasm and trust made the early testing days far more enjoyable.

Thanks are also due to Kalin for lending his skills and creativity in designing the Unity simulator that brought this experiment to life. Your input helped steer the virtual car — and the project — in the right direction.

I am also grateful to Moritz, Benji, and Lina for keeping my social life alive during the long stretches of thesis work. And to my parents for our weekly Friday crossword calls, which gave me something to look forward to and kept me in a steady rhythm.

Finally, to my partner Tom: thank you for the daily encouragement, emotional support, and occasional physical intervention required to get me out of bed in the morning. This project would have been impossible without your steady presence and tireless motivation.

7 References

[1] Han Yuan et al. “Negative covariation between task-related responses in alpha/beta-band activity and BOLD in human sensorimotor cortex: an EEG and fMRI study of motor imagery and movements”. en. In: *Neuroimage* 49.3 (Feb. 2010), pp. 2596–2606.

[2] Han Yuan and Bin He. “Brain-computer interfaces using sensorimotor rhythms: current state and future perspectives”. en. In: *IEEE Trans. Biomed. Eng.* 61.5 (May 2014), pp. 1425–1435.

[3] G. Pfurtscheller and C. Neuper. “Motor imagery and direct brain-computer communication”. In: *Proc. IEEE Inst. Electr. Electron. Eng.* 89.7 (July 2001), pp. 1123–1134.

[4] José del R Millán and Josep Mouriño. “Asynchronous BCI and local neural classifiers: an overview of the Adaptive Brain Interface project”. en. In: *IEEE Trans. Neural Syst. Rehabil. Eng.* 11.2 (June 2003), pp. 159–161.

[5] Karl LaFleur et al. “Quadcopter control in three-dimensional space using a noninvasive motor imagery-based brain-computer interface”. en. In: *J. Neural Eng.* 10.4 (Aug. 2013), p. 046003.

[6] Devaj Parikh and Kiran George. “Quadcopter control in three-dimensional space using SSVEP and motor imagery-based brain-computer interface”. In: *2020 11th IEEE Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*. Vancouver, BC: IEEE, Nov. 2020.

[7] B. J. Edelman et al. “Noninvasive neuroimaging enhances continuous neural tracking for robotic device control”. en. In: *Sci. Robot.* 4.31 (June 2019), eaaw6844.

[8] James R. Stieger et al. “Benefits of deep learning classification of continuous noninvasive brain-computer interface control”. en. In: *J. Neural Eng.* 18.4 (June 2021).

[9] Michael Tangermann et al. “Review of the BCI competition IV”. en. In: *Front. Neurosci.* 6 (July 2012), p. 55.

[10] Benjamin Blankertz et al. “The non-invasive Berlin brain–computer interface: fast acquisition of effective performance in untrained subjects”. In: *NeuroImage* 37.2 (2007), pp. 539–550.

[11] Hans Berger. “Über das Elektrenkephalogramm des Menschen”. de. In: *Archiv f. Psychiatrie* 87.1 (Dec. 1929), pp. 527–570.

[12] Davide Valeriani et al. “Multimodal collaborative brain-computer interfaces aid human-machine team decision-making in a pandemic scenario”. en. In: *J. Neural Eng.* 19.5 (Oct. 2022), p. 056036.

[13] Eui Chul Lee et al. “A brain-computer interface method combined with eye tracking for 3D interaction”. en. In: *J. Neurosci. Methods* 190.2 (July 2010), pp. 289–298.

[14] Catalina Llanos et al. “Mu-rhythm changes during the planning of motor and motor imagery actions”. en. In: *Neuropsychologia* 51.6 (May 2013), pp. 1019–1026.

[15] Martina Gandola et al. “Functional brain effects of hand disuse in patients with trapeziometacarpal joint osteoarthritis: executed and imagined movements”. en. In: *Exp. Brain Res.* 235.10 (Oct. 2017), pp. 3227–3241.

[16] Sarah Kraeutner et al. “Motor imagery-based brain activity parallels that of motor execution: evidence from magnetic source imaging of cortical oscillations”. en. In: *Brain Res.* 1588 (Nov. 2014), pp. 81–91.

[17] Hiroshi Shibasaki and Mark Hallett. “What is the bereitschaftspotential?” en. In: *Clin. Neurophysiol.* 117.11 (Nov. 2006), pp. 2341–2356.

[18] H Shibasaki et al. “Components of the movement-related cortical potential and their scalp topography”. en. In: *Electroencephalogr. Clin. Neurophysiol.* 49.3-4 (Aug. 1980), pp. 213–226.

[19] Fatemeh Karimi et al. “Detection of movement related cortical potentials from EEG using constrained ICA for brain-computer interface applications”. en. In: *Front. Neurosci.* 11 (June 2017), p. 356.

[20] H Ramoser, J Müller-Gerking, and G Pfurtscheller. “Optimal spatial filtering of single trial EEG during imagined hand movement”. en. In: *IEEE Trans. Rehabil. Eng.* 8.4 (Dec. 2000), pp. 441–446.

[21] Daily Milanés Hermosilla et al. “Shallow Convolutional Network Excel for Classifying Motor Imagery EEG in BCI Applications”. In: *IEEE Access* 9 (June 2021), pp. 98275–98286.

[22] Robin Tibor Schirrmeister et al. “Deep learning with convolutional neural networks for EEG decoding and visualization”. In: *Human Brain Mapping* 38.11 (2017), pp. 5391–5420.

[23] Arthur E Hoerl and Robert W Kennard. “Ridge regression: Biased estimation for nonorthogonal problems”. In: *Technometrics* 42.1 (Feb. 2000), p. 80.

[24] G. Pfurtscheller et al. “Graz-BCI: state of the art and clinical applications”. en. In: *IEEE Trans. Neural Syst. Rehabil. Eng.* 11.2 (June 2003), pp. 177–180.

[25] Unity Technologies. *Unity*. Version 2023.2.3. Game development platform. 2023. URL: https://unity.com/.

[26] SocketWeaver. *Karting*. https://github.com/SocketWeaver/karting. 2019.

[27] C. Kothe et al. *LabStreamingLayer*. Open Source Software Package. 2019. URL: https://github.com/sccn/labstreaminglayer.

[28] B Ans, J Hérault, and C Jutten. “Architectures neuromimétiques adaptatives : Détection de primitives”. In: *Cognitiva* 85 (1985), pp. 593–597.

[29] Alexandre Gramfort et al. “MEG and EEG Data Analysis with MNE-Python”. In: *Frontiers in Neuroscience* 7.267 (2013), pp. 1–13. DOI: 10.3389/fnins.2013.00267.

[30] Diederik P Kingma and Jimmy Ba. “Adam: A method for stochastic optimization”. In: (2014). eprint: 1412.6980 (cs.LG).

[31] A I Sburlea, N Butturini, and G R Müller-Putz. *Predicting emg envelopes of grasping movements from EEG recordings usingunscented Kalman filtering*. 2021.

[32] Andreea Ioana Sburlea, Luis Montesano, and Javier Minguez. “Continuous detection of the self-initiated walking pre-movement state from EEG correlates without session-to-session recalibration”. en. In: *J. Neural Eng.* 12.3 (June 2015), p. 036007.

[33] Scott M Lundberg and Su-In Lee. “A Unified Approach to Interpreting Model Predictions”. In: *Advances in Neural Information Processing Systems 30*. Ed. by I. Guyon et al. Curran Associates, Inc., 2017, pp. 4765–4774.

[34] E Baykara et al. “Effects of training and motivation on auditory P300 brain-computer interface performance”. en. In: *Clin. Neurophysiol.* 127.1 (Jan. 2016), pp. 379–387.

[35] F Pichiorri et al. “Sensorimotor rhythm-based brain-computer interface training: the impact on motor cortical responsiveness”. en. In: *J. Neural Eng.* 8.2 (Apr. 2011), p. 025020.

[36] S Halder, I Käthner, and A Kübler. “Training leads to increased auditory brain-computer interface performance of end-users with motor impairments”. en. In: *Clin. Neurophysiol.* 127.2 (Feb. 2016), pp. 1288–1296.

[37] Fabien Lotte, Florian Larrue, and Christian Mühl. “Flaws in current human training protocols for spontaneous Brain-Computer Interfaces: lessons learned from instructional design”. en. In: *Front. Hum. Neurosci.* 7 (Sept. 2013), p. 568.

A Appendix A

INFORMATION ABOUT THE STUDY: 'Can a motor execution BCI achieve robust continuous control?'

- **Purpose of the research project**

Thank you for your participation in our research study focused on continuous control in Brain-Computer Interfaces (BCIs). In this study, you will use a BCI to control a simulated car. BCIs are systems that enable direct communication between the brain and external devices. This has a wide range of applications, especially for individuals with mobility impairments, where for example paralysed patients can control a wheelchair with their brain

In this experiment, you'll be introduced to two non-invasive methods: EEG (electroencephalogram) and EMG (electromyogram). EEG involves placing electrodes on your scalp to measure the brain's electrical activity. With this we can analyse your brain's signals related to movement. On the other hand, EMG involves placing electrodes on your arms to detect electrical activity generated by muscle movements. The decoded signals from these methods will be used to steer around a car using wrist movements.

The data collected in this study will be used to assess multiple Machine Learning methods that we think might improve the usability of BCI in situations that require continuous control: the ability to analyse ongoing brain signals to smoothly control an external device in real time, such as a wheelchair mentioned in the example earlier.

- **Procedure**

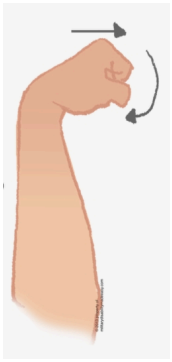
Both EEG and EMG are so called 'non-invasive' methods, meaning that they do not require any implants or injections. We will be placing electrodes on your scalp, near your eyes and on your arms. The electrodes are placed using a special electrolyte gel to make sure the electrical connection is good.

You will sit down in the chair in front of the computer. Place yourself at a comfortable distance from the screen. Put your arms on the arm rests so that your wrists are supported and your arms feel relaxed.

The experiment consists of two parts: the calibration phase and the experimental phase.

The first part of the calibration phase is to measure the effect your eyes have on the signals measured by the electrodes. For this to work well, it is important to minimize movement. You will first be asked to focus on a cross on the screen with your eyes open for two minutes until the cross disappears. You may blink as normal. After a 20 second break you will be asked to close your eyes for two minutes. We will let you know when to open your eyes at the end of the trial.

The second part of the calibration phase is used to practice the movements for the experiment, as well as train the model to detect the movements from the signals measured by the electrodes. For this to work well it is important that you minimise (eye) movement during the trials. At each trial you will see an arrow pointing left or right. After the arrow disappears, you should attempt the corresponding movement:



- Left arrow = Move left wrist inward
- Right arrow = Move right wrist inward

The wrist movement should be **held until the Rest cue appears**. Then, move the hand back into a closed resting position so your arms feel relaxed. You will do this for 1 block of 7 minutes.

In the experimental phase of the study you will drive a simulated car around a track using the movements that you practiced. You will drive several blocks with breaks in between. Try to focus on sitting as still as possible, avoiding eye movements, and performing the movements as you did during the training session.

The control commands are as follows:

- Steer left = Move left wrist = Left arrow
- Steer right = Move right wrist = Right arrow
- Steer straight = Rest both wrists = Rest cue

- **Possible risks or discomfort**

There is no risk associated with the study. We will use a gel in your hair and on your arms, which you will need to wash off after the experiment. For this we have a shower and shampoo/soaps available. You may find the experiment fatiguing.

- **Voluntary nature of participation**

You are free at any time and for any reason to withdraw from the study, without any adverse effects on the study or in any other way. Any reimbursements you are eligible for up to that point will be paid (in proportion to the duration of participation).

- **How we treat your data**

Your personal details will only be stored for the purpose of contacting you with regards to scheduling your session and payment. The personal details are stored on a password-protected server of the University of Groningen. After participation and payment have been completed, any recorded personal data will be permanently deleted.

Participant number: _____

Age: _____

Gender: □ Male □ Female □ Non-binary □ Other, please specify: _____

Do you have a diagnosed neurological disorder? □ Yes □ No

Do you have (corrected-to-)normal vision? □ Yes □ No

Do you have any prior experience with BCI research? □ Yes □ No

How would you describe your sleep quality last night?

□ Very poor □ Poor □ Normal □ Good □ Very good

How would you describe your gaming experience level?

□ Very inexperienced □ Inexperienced □ Normal □ Moderately experienced □ Very experienced

For each of the activities below, please indicate:

Which hand do you prefer for that activity? Do you ever use the other hand for the activity?

		No preference		Do you ever use the other hand?
Writing:	□ Left		□ Right	□ Yes
Drawing	□ Left		□ Right	□ Yes
Throwing	□ Left		□ Right	□ Yes
Using scissors	□ Left		□ Right	□ Yes
Using a toothbrush	□ Left		□ Right	□ Yes
Using a knife (without a fork)	□ Left		□ Right	□ Yes
Using a spoon	□ Left		□ Right	□ Yes
Using a broom (upper hand)	□ Left		□ Right	□ Yes
Striking a match	□ Left		□ Right	□ Yes
Opening a box (holding the lid)	□ Left		□ Right	□ Yes
Holding a computer mouse	□ Left		□ Right	□ Yes
Using a key to unlock a door	□ Left		□ Right	□ Yes
Holding a hammer	□ Left		□ Right	□ Yes
Holding a brush or comb	□ Left		□ Right	□ Yes
Holding a cup while drinking	□ Left		□ Right	□ Yes