Investigating EEG-Based Motor Imagery Decoding in the Continuous Control BCI Paradigm

by

© Mahdi Moeini

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

Department of Engineering and Applied Science
Memorial University of Newfoundland

May 2025

St. John's, Newfoundland and Labrador, Canada

I would like to dedicate this thesis to the pillars of my life, whose love and support have shaped me into who I am today.

To my mother, for her boundless love, countless devotion, and quiet strength—and for her unwavering encouragement that became my foundation through every storm.

To my father, for his selflessness, sacrificing his own comfort to prioritize my future, ensuring I could chase my dreams without a second thought.

To my sister, for her endless kindness and generosity, always offering help without hesitation. Her warmth and support have been a constant light in my life.

And to the one I love, for being my safe haven, my confidant, and my greatest believer. Her unshakable faith in my dreams gave me the courage to keep moving forward.

Together, their support and sacrifices granted me the freedom to focus on my ambitions—a gift I will spend my life striving to repay.

Abstract

Brain-computer interfaces (BCIs) based on motor imagery (MI) hold significant potential for restoring autonomy to individuals with severe motor disabilities. While decoding of short, discrete MI tasks has been extensively studied, the reliability of sustained MI decoding for continuous control paradigms remains under-explored. This study evaluates the feasibility of classifying four MI tasks (Hand, Feet, Tongue, Singing) versus rest over extended intervals (8–20 seconds) using EEG.

Fifteen participants performed cued MI tasks in a simulated continuous control paradigm. A machine learning pipeline—optimized for real-time compatibility—was systematically assessed, focusing on training data volume, epoching strategies, feature selection, and classifiers. Under the optimal configuration (four training blocks, 4-second epochs with 75% overlap, 10 features, SVM), comparable decoding accuracies were obtained on average across participants for the four tasks (Hand: 78.4% ± 9.5, Feet: 74.0% ± 10.0, Tongue: 74.5% ± 6.4, Singing: 74% ± 5.6), with no significant inter-task differences (Friedman test, $p = 0.41$). Within participants, however, there was considerable variability in decoding accuracy among the tasks. Comparative analysis against a shorter-interval MI dataset revealed a significant reduction in accuracy under the continuous paradigm only for Singing MI (Mann-Whitney U, $U = 44$, $p = 0.01$). Subjective feedback

highlighted Singing MI as the participants' most preferred task (53.3% of participants selected it as their favorite), while Tongue MI was the least preferred (33.3% of participants selected it as their least favorite).

These results underscore the viability of continuous MI-based control while emphasizing the need to balance technical performance with user-centric design. The novel task of singing MI emerged as a promising candidate for intuitive BCI applications, warranting further exploration of hybrid paradigms and ecological validation to enhance real-world usability.

Acknowledgment

I extend my deepest gratitude to my supervisor, Dr. Sarah Power, for her unwavering guidance, expertise, and encouragement throughout this research. Her insightful feedback, patience in navigating challenges, and commitment to fostering innovation were instrumental in shaping this work.

I am also thankful to my research group members for their collaboration and support: Mr. Hadi Mohammadpour, Mr. Erfan Mansoori, Mr. Bijan Roudini, Mr. Jacob White, Ms. Faghihe Massaeli, Ms. Mona Hejazi, and Mr. Roy Brushett for their collaboration and support.

Finally, I owe immense appreciation to my family and friends for their endless patience, emotional support, and encouragement during the demanding phases of this research. Their belief in my work sustained my motivation and resilience.

Table of Contents

List of Figures .... **vi**

List of Abbreviations .... **viii**

Introduction **1**

1.1 Problem Statement 1

1.2 Research Objectives 4

1.3 Thesis Organization 5

Literature Review **6**

2.1 Brain-Computer Interfaces (BCI) 6

2.1.1 Active vs. Passive BCI 6

2.2 Active BCI Paradigm 7

2.2.1 Command Generation 8

2.2.2 Data Acquisition 10

2.2.3 Data Windowing/Epoching 14

2.2.4 Data Preprocessing 15

2.2.5 Feature Extraction 16

2.2.6 Feature Selection 19

2.2.7 Classification 20

2.2.8 Feedback 23

2.3 Control Paradigms in Active BCI 23

2.4 Motivation and Research Objectives 29
Methods 32
3.1 Participants...32
3.2 Data Acquisition 33
3.3 Experimental Protocol 33
3.4 Motor Imagery Classification via EEG 38
3.4.1 Continuous MI Task Decoding 38
3.4.2 Comparison of MI Decoding Over Longer and Shorter Task Intervals...42
3.4.3 Subjective Data Analysis...44
Results 45
4.1 Continuous MI Task Decoding...45
4.2 Comparison of MI Decoding Over Longer and Shorter Task Intervals.........52
4.3 Subjective Data Analysis..53
Discussion..55
5.1 Continuous MI Task Classification 55
5.2 Comparison of MI Decoding Over Longer and Shorter Task Intervals.........59
5.3 Study Limitations and Future Work 61
Conclusion...64
6.1 Primary Conclusion 64
Bibliography 66

List of Figures

Figure 1.1: Operational flow of an active BCI paradigm. 2

Figure 2.1: Active BCI Paradigm Overview 8

Figure 2.2: Electrode placement diagrams based on the 10-10 system for EEG recording 14

Figure 2.3: Continuous vs. Switch paradigm...26

Figure 3.1: Summary of the experimental protocol 33

Figure 3.2: Tutorial screen explaining the motor imagery of the feet task 35

Figure 3.3: The difficulty rating scale used by participants to evaluate the imagery tasks. 37

Figure 4.1: Box plot showing the classification accuracies for the four different MI tasks 46

Figure 4.2: Box plot showing classification accuracy for different training set sizes 47

Figure 4.3: Box plots showing classification accuracies for a) different epoching approaches, b) classification algorithms, and c) feature set sizes 48

Figure 4.4: MI task vs. Rest classification accuracies for the four MI tasks for each participant, and on average across participants.. 49

Figure 4.5: Confusion matrices for the classification of each MI task vs. Rest 50

Figure 4.6: Total number of misclassifications per second across 15 participants for the four MI 52

Figure 4.7: Decoding accuracies for each MI task vs. Rest, averaged across participants, for the current dataset (8-20-second intervals) and the MI_{short} dataset (4-second intervals). 52

Figure 4.8: Average participant ratings of task difficulty for the four MI tasks for the seven experimental blocks, as well as the overall average across the seven blocks.. 53

Figure 4.9: Summary of participants' responses to task preference questions posed at the end of the experimental session. 54

Figure 4.10: Participants' highest classification accuracy among the four MI tasks, and their favorite task. 54

List of Abbreviations

BCI	Brain-Computer Interface
MI	Motor Imagery
ALS	Amyotrophic Lateral Sclerosis
EEG	Electroencephalography
fMRI	Functional Magnetic Resonance Imaging
fNIRS	Functional Near-Infrared Spectroscopy
FIR	Finite Impulse Response
PSD	Power Spectral Density
ApEn	Approximate Entropy
SampEn	Sample Entropy
CSP	Common Spatial Pattern
FBCSP	Filter Bank Common Spatial Pattern
mRMR	Minimum Redundancy Maximum Relevance
LDA	Linear Discriminant Analysis
SVM	Support Vector Machine
XGBoost	Extreme Gradient Boosting
KVIQ-10	Kinesthetic and Visual Imagery Questionnaire
ICEHR	Interdisciplinary Committee for Ethics in Human Research

H	Hand Motor Imagery
F	Feet Motor Imagery
T	Tongue Motor Imagery
S	Singing Motor Imagery

Chapter 1

Introduction

1.1 Problem Statement

The goal of brain-computer interfaces (BCIs) is to establish a direct connection between the human brain and a computer. These systems monitor the user's brain activity and from it derive meaningful information to use as input to an external device. In *active BCI*, the brain activity is intentionally generated by the user for the explicit purpose of directly controlling the external device [1] (e.g., wheelchair navigation), while in *passive BCI*, spontaneous brain activity is used to gain insights into the cognitive state of the user in order to modify or adapt their ongoing interaction with the device (e.g., adapting video game difficulty based on player engagement) [2].

Active BCIs must be able to identify and distinguish a set of distinct, reproducible brain activity patterns that are linked to specific output commands. Users intentionally generate these patterns for the explicit purpose of directly controlling the connected device [1]. The BCI records brain activity using electroencephalography (EEG) [3] or another functional neuroimaging method, analyzes the signals using machine learning algorithms to predict the user's intended command, and sends the command to the device. Typically, users generate the distinct brain activity patterns through performing mental tasks, the most common of which are motor imagery (MI) tasks [4]. Different MI tasks involve imagining the movement of specific body parts, such as the hands, feet, or tongue, which generate distinct neural patterns in the brain [5]. For example, in a BCI used to control a computer, the user may perform imagery of the right hand to move the cursor right, imagery of the left hand to move the cursor left, and imagery of the feet to click the mouse. Figure 1.1 illustrates the block diagram of how an active BCI operates.

Figure 1.1: Block diagram illustrating the operational flow of an active BCI from the pattern generation stage to the command stage.

BCI systems can be categorized as either "synchronous (system-paced)" or "asynchronous (self-paced)". In a synchronous BCI, the system provides external cues and allows the user to issue commands only at specific, predefined intervals determined by the system (e.g., [6], [7]). On the other hand, an asynchronous BCI enables the user to issue commands at any time, without being restricted by external timing cues (e.g.,[8], [9]). The terms synchronous and asynchronous describe BCI control paradigms in terms of how *the BCI controls the external device:* in the former, the BCI reads and decodes the user's brain activity and issues commands only at specific times (e.g., [6], [7], [10], [11]), whereas in the latter, the BCI reads and decodes the brain activity and issues commands continuously (e.g., [8], [9], [12], [13]). BCI systems can also be categorized based on *how the user controls the device*: switch control and continuous control [14]**.** Switch control involves the user performing distinct mental tasks for short durations to issue discrete commands. This approach works well for tasks requiring brief, separate actions, such as selecting an item on a screen or turning a device on and off (e.g., [15], [16]). Continuous control, in contrast, allows the user to sustain an action by maintaining a mental task for as long as the command needs to be executed. For example, a user could perform an MI task to keep a video game avatar moving in a particular and stop performing the task when they wish to stop the avatar's movement (e.g., [17]). While switch control can be implemented in both synchronous and asynchronous BCIs, continuous control is more suited to asynchronous systems due to the need for constant, real-time monitoring and interpretation of brain signals.

Many studies over decades of BCI research have shown that MI tasks can be detected from EEG signals. A recent review of 89 studies showed that reported classification accuracies vary across different cases. For binary classification, performance in the upper quartile falls within the 85%–100% range, while for multi-class classification, the upper quartile ranges from 83% to 93% [18]. However, a vast majority of MI-BCI studies have focused on short, discrete trials where users perform MI for just a few seconds before resting (e.g., [9], [19], [20]). Indeed, another review of 25 public MI datasets found that the average trial length used in these datasets is only 4.26 seconds, with most lasting between 1 and 10 seconds [21]. This is very different from continuous control, where users must sustain detectable mental states for longer periods, and the system must accurately decode these states continually, at regular intervals. Studies investigating continuous control BCIs typically test system performance online in real-world applications (e.g., control of wheelchair (see [22] for a recent review), drone (see [23] for a recent review), or avatar (e.g., [24], [25]), and thus report only task-based performance metrics like completion time or error rates rather than directly assessing how well the system decodes the mental states (e.g., [26], [27]). This of course makes sense since in real-world applications, the actual mental states of the user are unknown. However, without studies directly assessing decoding accuracy over longer intervals, it is unclear whether performance issues in continuous control BCIs come from the user's ability to sustain MI, the system's ability to decode it, or other aspects of the control strategy (e.g., compensatory mechanisms, shared control, etc.).

1.2 Research Objectives

This research sought to fill this gap in the existing MI-BCI literature by evaluating the decoding of multiple MI tasks vs Rest over longer trials with extended durations (8–20 seconds) of both conditions, which reflect a continuous control scenario. It investigated both traditional MI tasks, such as hand, foot, and tongue imagery, along with a novel singing motor imagery task proposed in [28]. The evaluation focused on two key aspects: mental state decoding accuracy, and subjective measures of user difficulty rating and general task preference. A comprehensive analysis was performed, exploring multiple analysis methods to maximize decoding accuracy. Furthermore, classification outcomes from the longer-duration trials were compared with those obtained from shorter-duration trials using a previously collected dataset reported in [28].

Specifically, the research objectives of this thesis were to:

1) Assess the continuous classification of different MI tasks versus a rest state over longer trials with extended durations (8–20 seconds) of each condition, simulating real-world continuous control conditions.
2) Compare different MI tasks—hand, foot, tongue, and singing motor imagery—in terms of mental state decoding accuracy and user preference in a continuous control-like scenario.

3) Compare MI decoding accuracy in a continuous control-like scenario with results obtained from a previously reported switch control-like dataset [28].

The aim of this thesis was to better understand how reliably mental states can be generated and decoded over longer intervals than are typically studied in BCI research, which is important for the design and evaluation of BCIs for continuous-control scenarios.

1.3 Thesis Organization

The remainder of this thesis is organized as follows:

Chapter 2 provides a comprehensive background on BCI systems, covering all necessary concepts to understand the methods and objectives of this study. It includes a review of relevant literature, discussing previous research on MI-BCI and continuous control paradigms. Building on this foundation, **Chapter 3** describes the experimental and analytical methods used in this study, detailing data collection, preprocessing, classification techniques, and the evaluation of different MI tasks. The results of these analyses are presented in **Chapter 4**, where key findings are outlined. These findings are then discussed in **Chapter 5**. Finally, **Chapter 6** summarizes the thesis contributions, addressing its limitations, and providing recommendations for future research.

Please note that portions of Chapters 1, 2, and 6, and the entirety of Chapters 3 to 5, are taken directly from a manuscript titled *"Investigating EEG-Based Motor Imagery Decoding in the Continuous Control BCI Paradigm",* by M. Moeini and S. D. Power, which will be submitted to IEEE Access.

Chapter 2

Literature Review

2.1 Brain-Computer Interfaces (BCI)

Brain-computer interface (BCI) systems establish a direct communication pathway between the human brain and a computer, enabling interaction without physical movement. They work by deriving useful information from the user's brain activity that can then be used as input to control an external device. The potential applications of BCI technology have significantly expanded in recent years. BCIs are now being explored in a wide range of fields, including external device control (e.g., [29], [30]), neurorehabilitation (e.g., [31], [32]), brain-controlled gaming (e.g., [33], [34]), and cognitive state analysis, such as fatigue monitoring (e.g., [35], [36]) and sleep pattern assessment (e.g., [37], [38]).

2.1.1 Active vs. Passive BCI

BCI systems can be categorized into active and passive BCIs based on *whether the user's control over the device is direct or indirect*. Active BCI systems use brain activity generated intentionally by the user for the specific purpose of directly controlling an external device or system [1]. For example, a user may generate specific patterns of brain activity in order to control a robotic arm (e.g., [27]) or wheelchair (e.g., [39]). In contrast, passive BCI systems monitor spontaneously occurring brain activity to gain insights into the user's cognitive or emotional state - without any intentionality on their part - and adjust an ongoing human-computer interaction accordingly [2]. For example, a passive BCI system could monitor a user's brain activity to detect signs of fatigue during a driving simulation (e.g., [40]) and provide real-time adjustments, such as reducing the complexity of tasks or triggering alerts to maintain safety.

2.2 Active BCI Paradigm

The structure of a typical active BCI is shown in Figure 2.1. It begins with the *command generation* stage, where the user generates the specific brain activity patterns that are associated with the different commands they may want to send to the device. Next is *data acquisition*, where the user's brain activity is measured and recorded using some functional imaging technology, and *data pre-processing*, where artifacts are removed from the collected data, and it is otherwise prepared for analysis. *Feature extraction* and *feature selection* involve calculating a pool of potentially meaningful measures from the data and choosing a subset of the most informative ones. In *classification,* machine learning algorithms are used to determine the participant's intended command based on the selected features of the data. Finally, based on the classification result, a corresponding command is issued to the device. The user then observes if their command was decoded correctly or not. This final step constitutes the feedback, closing the loop of the BCI system. Each of these stages will be discussed in more detail in the following sections.

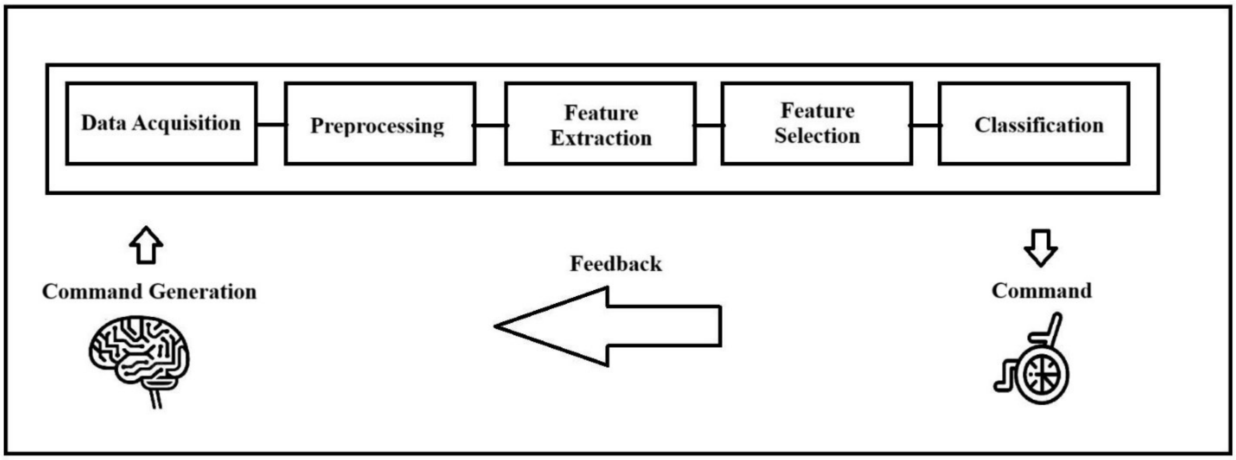


Figure 2.1: Active BCI Paradigm Overview.

2.2.1 Command Generation

There are several well-established methods used to generate distinct brain activity patterns in active BCIs. One popular method is based on event-related potentials (ERPs), which are involuntary time-locked neuroelectric responses to specific sensory, cognitive, or motor events [41]. The most common ERPs used in BCIs are the steady-state visually evoked potential (SSVEP) and the P300 response [41]. The SSVEP occurs when focusing on a flickering visual stimulus, which generates rhythmic brain activity in the visual cortex corresponding to the frequency of the stimulus. SSVEP-based BCIs take advantage of this by presenting multiple flickering targets, each associated with a distinct frequency, allowing users to select commands by directing their visual attention to a specific target (e.g., [42], [43]). The P300 response, on the other hand, is elicited by sudden, attention-capturing stimuli and is often used in tasks requiring target detection. P300-based BCIs utilize this mechanism in applications such as speller systems, where a user identifies a desired letter or symbol among a series of flashing options, and the system detects the P300 response to determine the selection (see [44] for a recent review). Because ERP-based BCIs rely on brain activity generated as an unconscious reaction to a stimulus, they are actually more precisely classified as “reactive BCIs” [45]. However, since the user intentionally controls their focus on the stimuli, these systems still involve direct, intentional control over the device, much like active BCIs (see [46]).

Another popular method is using mental tasks (see a recent review [47]). The goal is to use different mental tasks that produce unique brain activity patterns and develop decoding algorithms that can distinguish the tasks from one another based on EEG. The different tasks are then associated with different output commands to send to the external device. While many tasks have been explored, such as arithmetic calculations (e.g., [48], [49], [50]) and mental rotation (e.g., [51], [52], [53]), the most widely used are motor imagery (MI) tasks (e.g., [6], [13], [34], [54]). MI involves participants imagining the movement of body parts, such as hand or foot movements, without physically executing the movement [55]. This mental simulation leads to detectable changes in the sensorimotor cortex [55]. MI can be further divided into visual MI and kinesthetic MI [56]. Visual MI involves the user imagining seeing themselves perform a movement, while kinesthetic MI involves mentally simulating the sensation of performing the movement [56]. Kinesthetic MI often results in stronger and more distinguishable brain activity patterns due to increased activation of the motor cortex, which leads to more reliable BCI systems [57].

In BCI systems based on mental tasks, the more distinct and reliable patterns a user can generate—and the BCI can detect—the greater the number of distinct commands the BCI can send, which leads to increased functionality, speed, and precision in BCI systems [28]. For example, if a BCI can only decode MI vs. Rest, it can issue just one command (e.g., "activate" or "do nothing"). However, if the BCI can distinguish between left hand (LH), right hand (RH), foot (F), tongue (T), and Rest, it can achieve four commands, enabling more complex control. For instance, a BCI could be designed so that RH-MI corresponds to moving a cursor or wheelchair to the right, LH-MI moves it to the left, F-MI moves it down/backward, and T-MI moves it up/forward, and Rest is "do nothing". This would allow the user to control the device with high precision, enabling more sophisticated interactions with external devices (see [58], [59]).

2.2.2 Data Acquisition

2.2.2.1 Functional Imaging Modalities in BCI

Brain activity can be recorded using various functional neuroimaging techniques, each with unique underlying principles of operation, advantages, and limitations. Different methods capture brain activity by measuring different physiological responses related to neural activity, such as electrical signals, magnetic fields, or changes in blood flow and oxygenation levels. In order for a functional imaging technology to be practical for use in real-world BCI systems, it must be portable, easy to use, and relatively inexpensive. This eliminates techniques like positron emission tomography (PET) [60] and functional magnetic resonance imaging (fMRI) [61], despite their good signal quality. In BCIs, brain activity can be recorded **non-invasively,** from outside the scalp without penetrating the skull, or **invasively**, which requires the placement of electrodes directly on/in the brain's cortical surface. Invasive methods offer superior signal quality and much better BCI control is possible with such methods, however, they also pose significantly greater medical risks [62]. The techniques most often explored in BCI research include:

- **Electrocorticography (ECoG):** An invasive neuroimaging technique that records brain activity through electrodes placed directly on the surface of the cerebral cortex. It measures the summed electrical potential generated by the population of neurons directly beneath each electrode [63]. ECoG provides excellent spatial and temporal resolution, making it highly effective for identifying specific brain regions involved in various tasks. However, it requires surgical implantation of electrodes, which limits its use to clinical or research settings and poses risks associated with invasive procedures (see [64] for a recent review).

- **Intracortical Electrode Arrays:** An invasive technique that involves inserting microelectrodes directly into the cortical tissue to measure the electrical activity of individual neurons or small groups of neurons [65]. This method provides extremely high spatial and temporal resolution, enabling detailed analysis of neural circuits and single-cell activity [65]. However, like ECoG, it is invasive, requiring surgical implantation, and it carries surgical risks [66].
- **Electroencephalography (EEG)** A non-invasive neuroimaging technique that records electrical activity generated by neurons in the brain using electrodes placed on the scalp [67]. EEG offers excellent temporal resolution, making it ideal for real-time applications [68]. However, its spatial resolution is relatively low due to the difficulty in localizing the exact source of brain activity [69]. The effect is that EEG-based BCIs cannot achieve the same precise control over external devices that those based on invasive methods can. However, since EEG does not carry the same risks as invasive methods, EEG-BCIs are suitable for a wider range of applications and user populations.

Functional Near-Infrared Spectroscopy (fNIRS): An optical imaging technique that measures brain activity based on changes in blood oxygenation that result from the increased metabolic demand of active neural tissue [70]. It works by emitting near-infrared light into the scalp and detecting how it scatters as it interacts with brain tissue. Some advantages of fNIRS are that it is non-invasive, portable, suitable for long-term monitoring, and relatively inexpensive. A significant limitation is its relatively low temporal resolution [71].

EEG is used in this study due to its non-invasiveness, high temporal resolution, and cost-effectiveness.

2.2.2.2 Electroencephalography (EEG)

EEG captures voltage fluctuations produced by synchronized neural firing in the cerebral cortex using specialized electrodes placed on the scalp [68]. These signals primarily reflect the postsynaptic potentials of pyramidal neurons in the neocortex, which are aligned perpendicularly to the cortical surface [72]. This alignment allows their collective electrical activity to summate, making it detectable at the scalp. Due to volume conduction, electrical signals spread through the scalp and surrounding tissues, causing each electrode to capture activity from a broad region of the underlying cortex rather than a localized source [73]. Because EEG measures the aggregate activity of large populations of neurons rather than individual neurons, its spatial resolution is relatively low.

The analog signals detected at the EEG electrodes, which are extremely weak, are amplified and passed to an Analog-to-Digital Converter (ADC). The ADC samples these signals at a specific sampling frequency, typically ranging from 256 Hz to 1024 Hz, to ensure adequate resolution for capturing the rapid fluctuations in brain activity. The resulting digital signals are stored for offline analysis and/or processed online in real time. EEG signals can be collected using either wet or dry electrode systems. Wet electrode systems require the application of a conductive gel or paste on the scalp to reduce impedance and improve signal quality. By lowering impedance, which is often kept below 5 kΩ, the signal-to-noise ratio (SNR) improves, resulting in cleaner data acquisition [74]. However, wet electrodes can be time-consuming to set up and may dry out during long recording sessions, reducing signal quality. Dry electrode systems do not require a conductive medium; instead, they rely on specialized materials that can detect brain activity directly from the scalp. While dry electrodes simplify setup and improve user comfort, they generally exhibit higher impedance and lower signal quality compared to wet electrodes, making them less suitable for high-precision recordings [75].The number of electrodes used during EEG acquisition can vary significantly depending on the system and application, typically ranging from just a few in basic setups to 32, 64, 128, or even 256 channels in advanced systems, where higher channel counts offer better spatial resolution but also increase data complexity. In BCI applications, minimizing the number of electrodes while maintaining the ability to accurately detect the mental states of interest is desirable for reasons of practicality and cost.

Electrode placement plays a critical role in EEG acquisition as it determines which brain regions are being monitored. The 10-20 International System is the most widely adopted standard for positioning electrodes [76]. It ensures consistent electrode placement based on the measured distances between key anatomical landmarks on the scalp, specifically the nasion**,** inion, and preauricular points. The system divides the scalp into 10% and 20% intervals for electrode placement, providing a balance between spatial coverage and signal quality [77].

Electrode positions in the 10-20 system are labeled with a combination of letters and numbers, where the letters (F, T, C, P, O) represent the brain regions (Frontal, Temporal, Central, Parietal, Occipital), and the numbers indicate the relative position on the left (odd numbers) and right (even numbers) hemispheres [77]. While the 10-20 system allows for the placement of 32 EEG electrodes, extensions of this system allow for higher density arrangements (e.g., the 10-10 or 10-5 systems) [78]. Figure 2.2 illustrates the standard electrode placement map based on the 10-10 system for 64 electrodes, showing how electrodes are positioned across the scalp for optimal brain activity recording.

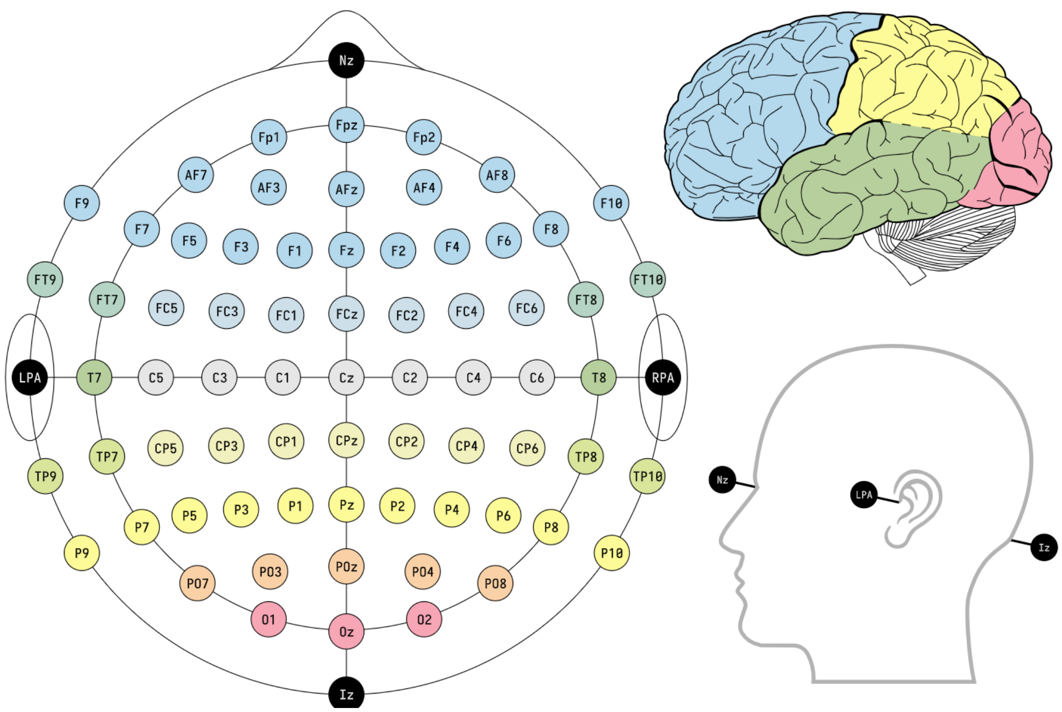


Figure 2.2: Electrode placement diagrams based on the 10-10 system for EEG recording, illustrating standardized scalp positions and corresponding brain regions.

2.2.3 Data Windowing/Epoching

As the EEG data is acquired, short windows (known as “epochs”) are isolated and input to the analysis pipeline (preprocessing, feature extraction/selection, classification) [80] sequentially, in a “sliding window” approach. Epochs typically range from 1 to 4 seconds in length, and adjacent epochs can be either overlapping or non-overlapping [81]. The choice of epoch length and amount of overlap can significantly impact decoding accuracy [81]. Longer epochs provide more data, potentially allowing the model to better capture the underlying brain activity patterns, but they may also introduce delays in real-time applications. Shorter epochs offer faster processing but may contain insufficient information for robust classification.

2.2.4 Data Preprocessing

In EEG-based BCI research, data preprocessing is an essential step to prepare the data for further analysis by machine learning models. Since EEG has a low SNR, this stage can be very important [82]. However, the decision to apply artifact removal techniques depends on the specific application. In some cases, aggressive artifact removal may eliminate useful EEG signals along with noise, potentially impacting the quality of the data [83].

Some common potential sources of EEG artifacts include:

- **Electrode Noise:** Electrode noise occurs when the conductive gel applied to the electrodes dries out, leading to an increase in impedance between the electrodes and the scalp. This elevated impedance reduces signal quality and increases noise interference, making it harder to accurately capture brain activity data [84]. Instead of artifact removal, this issue is best mitigated through artifact prevention, such as ensuring proper electrode placement and maintaining good conductivity with conductive gel. Regular monitoring of impedance levels during experiments can help maintain signal quality.
- **60 Hz Electrical Interference:** This noise originates from electrical devices in the surroundings of the recording setup, and occurs at the frequency of the standard power

supply, which is 60 Hz in Canada [85]. If frequencies above 60 Hz are relevant for a particular application, a sharp notch filter [86] is commonly applied to remove this specific interference while preserving other frequency components [87]. Alternatively, if higher frequencies are not of interest, a low-pass filter can be used to eliminate all frequencies beyond the required range [82], [88].

- **Participant-Generated Artifacts:** These artifacts arise from involuntary physiological activities such as eye blinks, swallowing, or by minor body movements during data recording [88]. One common method for removing such artifacts is by Independent Component Analysis (ICA), a statistical technique that decomposes EEG signals into independent components representing different sources of brain activity and noise [89]. Artifacts like eye blink and muscle movement can be isolated as separate components and removed, while brain-related components are preserved, ensuring a cleaner EEG signal for analysis [88]. ICA is popular due to its ability to separate signals without relying on predefined patterns, making it useful for EEG artifact correction. However, as mentioned above, there is a risk of eliminating the essential components of the data and potentially affecting the integrity of the signal.

2.2.5 Feature Extraction

Raw EEG data is typically represented as a 2D matrix, where the columns represent the different electrodes, and the rows correspond to samples from each electrode channel at specific time points. Feeding this high-dimensional, complex data directly into a machine-learning model can lead to poor performance due to the complex nature of the signals [90] and the substantial computational resources required for processing this volume of data [91]. Feature extraction addresses these challenges by distilling the essential information from the raw EEG signals into a set of meaningful, lower-dimensional features. This process not only simplifies the data, making it more interpretable for machine learning algorithms but also reduces the computational burden, thereby conserving hardware and software resources [90].

Features commonly used in EEG analysis include:

- **Time-domain Features** are extracted directly from the EEG signal by analyzing the amplitude variations over time. These features describe the statistical properties of the signal and are often helpful for initial data exploration. For instance, the *mean* represents the average amplitude of the EEG signal, providing a measure of central tendency, and the *z-score* standardizes the signal by measuring how many standard deviations a data point deviates from the mean, aiding in normalization and outlier detection [90].
- **Frequency-domain Features** describe the signal's frequency content, providing insights into the oscillatory activity in the brain [92]. For example, *power spectral density* (PSD) describes the distribution of power across different frequency bands, indicating the brain's rhythmic activity [93]. *Band power* measures the signal's power in specific frequency ranges, such as delta (0.5–4 Hz; deep sleep and unconscious states [94]), theta (4–8 Hz; memory and light meditation [95]), alpha (8–13 Hz; relaxation and calmness [74]), beta (13–30 Hz; active thinking and problem-solving [74]), and gamma (>30 Hz; higher cognitive functions and perception [96]), each associated with different cognitive or physiological states.
- **Entropy-based Features** measure the complexity and unpredictability of the EEG signal, providing insight into the non-linear dynamics of brain activity [97]. For example,

approximate entropy (ApEn) estimates the regularity and complexity of a time series, with lower ApEn values indicating a more regular signal and higher values indicating more complexity [98]. *Sample entropy* (SampEn) is similar to ApEn but with a reduced bias for smaller datasets, making it more reliable for EEG data analysis [97].

- **Statistical Features** summarize the distribution and shape of the EEG signal. These features are often useful for capturing patterns and variations in the signal [92]. For example, *skewness* measures the asymmetry of the signal distribution around the mean, with positive skewness indicating a longer right tail and negative skewness indicating a longer left tail [99]. *Kurtosis* measures how extreme the peaks and tails of the signal distribution are, with higher values indicating more extreme deviations from the mean [99].

See [92] for a recent review of feature extraction methods for EEG-based BCI.

2.2.5.1 Filter Bank Common Spatial Pattern (FBCSP)

The Common Spatial Pattern (CSP) algorithm is a widely used feature extraction technique in BCI studies, particularly for MI-BCI [100]. CSP enhances class discrimination in binary classification problems by extracting spatial features from EEG signals that maximize variance for one class while minimizing it for the other. Essentially, it finds optimal spatial filters that project EEG signals into a new space where classes are more separable [101]. The power of the projected/filtered EEG signals is calculated and used as features for classification.

The performance of CSP heavily depends on the frequency range of the input EEG data. If the EEG data fed into CSP is either unfiltered or filtered within an inappropriate frequency band, the extracted features may not be useful for MI task discrimination, leading to poor classification accuracy [102]. To address this limitation, Ang et al. [103] introduced Filter Bank Common Spatial Pattern (FBCSP), an enhanced version of CSP that incorporates frequency decomposition.

FBCSP first decomposes the EEG signal into multiple frequency sub-bands using a filter bank, ensuring that relevant frequency components are retained. CSP is then applied separately to each sub-band, extracting spatial features for each [103]. Finally, the extracted features from all sub-bands are concatenated, creating a comprehensive feature set that captures both spatial and frequency-specific information. This enriched representation significantly improves the classification performance of EEG-based BCI systems [103].

2.2.6 Feature Selection

Feature selection occurs after feature extraction. This phase is not always required, as sometimes the full extracted feature matrix can be directly fed into the machine learning model. However, feature selection is commonly used to identify the most relevant features and reducing data dimensionality. This reduction helps the ML algorithms perform better due to the reduction of noise and irrelevant patterns, which may add bias to the classifier [104]. Feature selection is done during the training phase based on the training data, and in the testing or "use" phase only the identified subset of features is used as input to the classification stage.

The Minimum Redundancy Maximum Relevance (mRMR) algorithm is a feature selection technique that identifies the most relevant features for distinguishing between classes while minimizing redundancy among them. It achieves this by using mutual information as a measure of both relevance and redundancy [105]. Relevance is assessed by evaluating how strongly a feature correlates with the target variable, ensuring that only the most informative features are selected. Redundancy is minimized by measuring the dependency between features, ensuring that selected features provide complementary rather than overlapping information [105]. Due to its ability to balance these two criteria, mRMR enables the selection of highly discriminative features, making it particularly effective for MI-BCI applications (e.g., [106], [107], [108]).

2.2.7 Classification

In machine learning, classification is the task of predicting the label or category of a given sample of input data based on patterns learned from training data with known labels. For instance, a classifier may be trained to predict whether a new image belongs to the category of "dog" or "cat". This is an example of binary, or “2-class”, classification. Multi-class classification involves the classification of three or more classes.

In MI-BCI research, classification algorithms are used to predict which of a predefined set of MI tasks the user is performing at any given time (and thus what command they wish to send to the external device) based on the features extracted from short segments of EEG data. To do this, the classifier must first be trained on data with known class labels. For practical BCI systems, the training data must be collected in a “calibration session” before the BCI can actually be used for real-time control of the device. In some BCI studies where the goal is to explore the decoding of novel mental states, or develop and validate new signal processing or machine learning algorithms, it is more efficient to collect a full dataset and subsequently perform analysis offline (e.g., [109], [110]). In these cases, a portion of the data is used to train the model, while the remainder is used as a “test set” to evaluate the classifier's performance on unseen data.

Many different machine learning algorithms have been explored for use in MI-BCIs (see a recent review [111]). Among them, linear classifiers like linear discriminant analysis (LDA) and support vector machines (SVM) have proven particularly effective and have become very popular in MI-BCI studies [112]. Extreme gradient boosting (XGBoost or XGB) is a non-linear classifier that has also been effective in many studies (e.g., [113], [114], [115]). These classifiers are discussed in more detail in the following sections.

2.2.7.1 Linear Discriminant Analysis (LDA)

LDA is a linear classifier widely used for both dimensionality reduction and classification tasks [116]. It works by projecting data onto a lower-dimensional space while maximizing the separation between multiple classes. LDA minimizes intra-class variance while maximizing inter-class variance, making it effective for classification tasks where classes are linearly separable [117].

LDA has demonstrated effectiveness in EEG-based BCI studies (e.g., [118], [119], [120]), and has become a widely used technique due to its simplicity, interpretability, and computational efficiency.

2.2.7.2 Support Vector Machines (SVM)

SVM is a powerful machine learning algorithm commonly used for both classification and regression tasks [121]. It is particularly effective in high-dimensional spaces and for datasets where a clear margin of separation exists between classes [122]. SVM works by finding the best possible boundary, called a hyperplane, that separates data points belonging to different classes. The goal is to create a boundary where the gap (or margin) between the two classes is as wide as possible [123]. The data points that are closest to this boundary are called support vectors. These support vectors define the position of the hyperplane and directly influence how the boundary is drawn [123]. SVM aims to maximize the margin between the support vectors of each class. A wider margin helps the model make better predictions on new, unseen data, improving its ability to generalize [124]. Since EEG data typically has high-dimensional features, SVM is a powerful tool as it effectively handles complex datasets [125].

2.2.7.3 Extreme Gradient Boosting (XGBoost)

XGBoost is a powerful and efficient machine learning algorithm based on gradient boosting, used for both classification and regression tasks [126]. It works by building multiple decision trees sequentially, where each tree attempts to correct the errors made by the previous ones. This iterative process improves the model's accuracy over time [127]. It uses gradient descent to reduce errors in its predictions by adjusting how the decision trees split the data based on mistakes made earlier [127]. Since XGBoost has built-in regularization (which penalizes model complexity to prevent overfitting) and can handle noisy, complex datasets effectively [127], it makes an excellent choice for EEG classification. This effectiveness has been demonstrated in numerous research studies, highlighting its reliability in MI-BCI tasks (e.g., [113], [114], [115]).

2.2.8 Feedback

Once an EEG data epoch is classified as one of the predefined set of MI tasks, the BCI sends the corresponding command to the connected device. The user observes whether or not their intended command was decoded correctly, and this feedback allows them to adjust their mental strategies in real time to adapt to the BCI algorithm, effectively "learning to use" the BCI. Ideally, the BCI algorithm would also adapt to the user over time, as it is well known that EEG signals are highly non-stationary [128].

As mentioned, since it is an efficient way to develop and validate new signal processing or machine learning algorithms, BCI studies often involve offline data classification. Of course, feedback is not possible in offline analysis. Since the feedback can have a significant effect on mental state generation and decoding, it is very important to validate the results of the offline analysis in online scenarios, where real-time classification is performed, and feedback is provided to the participant.

2.3 Control Paradigms in Active BCI

BCIs can be classified as either "synchronous/system-paced" or "asynchronous/self-paced". In synchronous BCIs, users must follow external cues and issue commands (i.e., generate defined brain activity patterns/perform mental tasks) only at pre-defined intervals dictated by the system [129], whereas in asynchronous BCIs users can issue commands whenever they choose [130]. Asynchronous control is, of course, more natural and user-friendly but also more technically challenging to implement since the periods of intentional control need to be identified from the highly variable "no-control" or baseline state [54]. The terms synchronous and asynchronous describe BCI control paradigms in terms of how *the BCI controls the external device:* in the former, the BCI reads and decodes the user's brain activity and issues commands only at specific times (e.g., [6], [7], [10], [11]), whereas in the latter, the BCI reads and decodes the brain activity and issues commands continuously (e.g., [8], [9], [12], [13]).

BCI control can also be categorized based on how *the user controls the BCI,* as either "switch" or "continuous" control [14]. In the switch control paradigm, users perform different tasks for short durations (typically a few seconds) to issue discrete commands (e.g., [15], [16]). This paradigm is very well-suited to control scenarios where the resulting action of the device is inherently discrete, like selecting options on a computer screen or turning the device on or off. Some common BCI control scenarios are inherently continuous, however, like navigating a wheelchair (see [22] for a recent review), drone (see [23] for a recent review), or video game avatar (e.g., [25],[24]). In these cases, the action of the device (i.e., movement) must be maintained for longer periods of time. Switch control can certainly be useful even in these scenarios. For example, discrete start/stop/turn left/turn right commands could be issued through the brief performance of different associated mental tasks, with the continuous movement of the device occurring automatically in between (termed low-level control) (e.g., [58]). Alternatively, the different mental tasks could be mapped to different routes/destinations rather than different directions, with the device autonomously navigating a full pre-defined path based on the single discrete command issued (termed high-level control) (e.g., [131], [132]). While these strategies can be very effective in many situations, they may not always be the best solution, for instance, in applications where the start and stop commands would have to be issued frequently or in quick succession (e.g., cursor control, scrolling through a webpage or document), or when continuous control simply feels more natural to the user.

In the continuous control paradigm, on the other hand, the user can issue prolonged commands by performing the MI task for as long as they wish the action to be maintained. For example, they could perform an MI task to start moving an avatar in a particular direction and continue to do so until they wish to stop it. This paradigm allows continuous, fluid control and may be more natural in some applications (e.g., [14], [17], [133], [134]). To illustrate the difference between discrete and continuous control, Figure 2.3 depicts how a user would send commands to move a video game avatar in one direction in each of these paradigms. Note that switch control can be implemented in either synchronous or asynchronous systems, but continuous control is better suited to asynchronous systems due to the need for constant, real-time monitoring and interpretation of the brain activity.

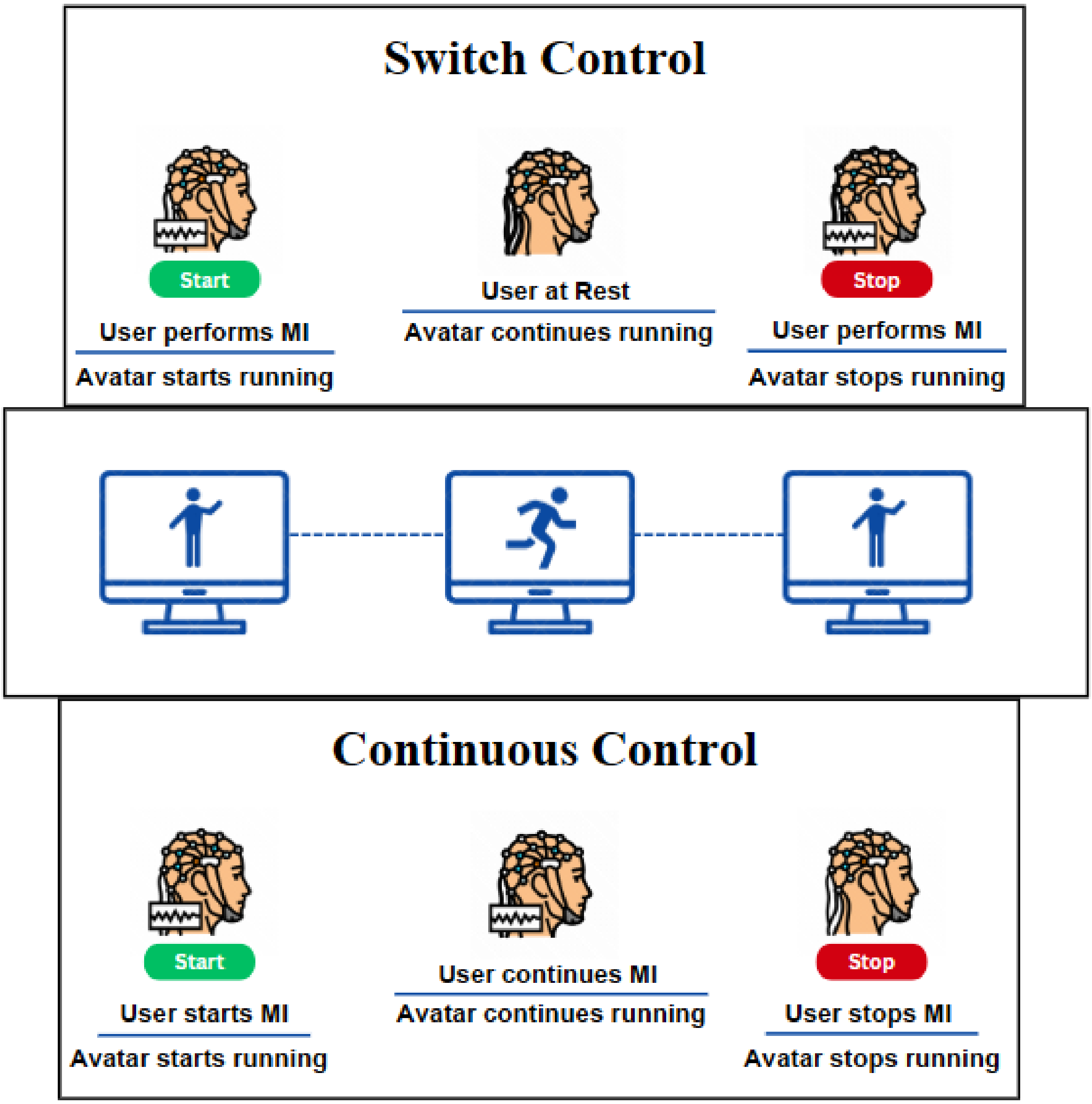


Figure 2.3: Example of how a user would control the one-directional movement of a video game avatar using MI tasks under the switch and continuous control paradigms. In the switch paradigm (top), the user performs MI at the start position to initiate movement of the avatar, remains at rest as the avatar moves automatically forward, and performs MI to stop movement of the avatar at the desired stop position. In the continuous paradigm (bottom), the user performs MI to initiate movement at the start position and maintains the MI continuously until the avatar reaches the desired stop position, at which point the user stops performing MI to stop the avatar.

While the use of the discrete control paradigm is more prevalent in BCI research, studies in many different application areas have used the continuous control paradigm, with the control either based entirely on the user's neural activity (e.g., [12], [14], [135]) or incorporating a complementary control strategy (termed "shared control"), for example using environmental sensors to detect obstacles or assist with navigation (e.g., [39], [136], [17]). For example, Velasco-Álvarez et al. [135] implemented continuous control for navigating a virtual wheelchair. Using a graphical interface, users could select among three possible commands (turn left, turn right, move forward) by performing a single MI task (right-hand motor imagery). Once initiated, the wheelchair continued the selected action until the MI task ceased, allowing participants to sustain fluid movement. In LaFleur et al. [137], a continuous control paradigm was used to control a quadcopter in 3D space, with MI of the left hand, right hand, both hands, and an idle state used to turn left, turn right, ascend, and descend, respectively. Forward motion was automatically controlled by the system, reducing the cognitive load on participants and allowing them to focus solely on directional and altitude adjustments. Using a similar paradigm, An et al. [138] developed a BCI-controlled Unmanned Aerial Vehicle (UAV) for navigating a 2D indoor space. Participants used MI tasks and an idle state for directional and altitude adjustments (left-hand MI to move upward or turn left, right-hand MI to move downward or turn right, and idle to maintain current direction), while forward motion was automatically controlled. More recently, Zhang et al. [17] utilized a continuous control paradigm for robotic arm operation, where left-hand, right-hand and tongue MI tasks corresponded to directional commands for guiding the arm in real time. When no MI commands were detected, the system relied on autonomous grasping control, which used vision feedback for path and grasp planning. Further supporting the practical application of continuous control in rehabilitation, Choi et al. [12] applied continuous control to lower-limb exoskeletons.

Users maintained the continued forward movement of the exoskeleton by performing sustained “gait MI” (mental imagery of walking). The system halted when the user stopped performing MI, employing a decision buffer for reliability in a “shared control” paradigm. Some studies have even combined the continuous control paradigm with switch control to harness the unique strengths of each. For example, Shi et al. [139] used a hybrid control paradigm for UAV navigation, combining switch control for decision-making with continuous control for manual navigation. During the decision-making phase, participants used MI tasks such as left-hand MI to select "Yes" (accepting a suggested direction) or right-hand MI to select "No" (rejecting the suggested direction). These binary choices were presented by the semi-autonomous navigation system, which provided feasible directions while ensuring obstacle avoidance. If the user did not choose "Yes" or "No" within the allotted time, the UAV hovered momentarily, and the participants were required to take control manually. In manual control mode, participants guided the UAV continuously using MI tasks: left-hand MI directed the UAV to turn left, right-hand MI to turn right, and idle imagination maintained forward motion. Similarly, He et al. [140] developed an asynchronous BCI system for human-computer interaction where switch control was used for discrete actions like button selection while continuous control allowed smooth, real-time cursor movement using MI tasks. This hybrid approach balanced precision and adaptability, making it effective for complex tasks like web browsing, text input, and file management. Fernandez-Rodríguez et al. [14] integrated continuous and switch control paradigms in a comparative study to examine the effectiveness, efficiency, and user satisfaction of each for wheelchair control. Their findings showed that while switch control generally required fewer commands and was simpler for users when moving the wheelchair, continuous control was generally more effective when the user needed to maintain the stop position. Neither paradigm was superior overall in terms of perceived mental workload or user satisfaction, with the best choice being highly user-dependent.

2.4 Motivation and Research Objectives

Key to the effectiveness of continuous BCI control is the ability of users to reliably generate the distinct patterns of brain activity required to issue a command over prolonged intervals (i.e., longer than just a few seconds) and the ability for the BCI to continuously (or near-continuously, for example, every 1 second) decode this activity. While there is abundant evidence over decades of BCI research that MI tasks can be decoded from EEG signals, a vast majority of these studies use or mimic the discrete control paradigm. That is, the user performs the MI task(s) over brief intervals (typically 3-5 seconds) which are both preceded and followed by rest. Indeed, a recent review of 25 publicly available MI datasets spanning the past 13 years [21] highlights this prevalence: the average MI duration during trials was just 4.26 seconds, with durations ranging from 1 to 10 seconds. This is very different from having to maintain task performance and, importantly, the stability of the associated neural activity patterns that the BCI recognizes, for longer periods of time. Of course, studies using the continuous control paradigm, like those cited above, do perform decoding over longer intervals of task performance. However, in terms of system evaluation, most such studies report only task-specific performance metrics like task completion time, error rate, success/failure rate, etc. rather than reporting the accuracy of mental state decoding itself (i.e., the accuracy with which the user successfully generates the desired command and the BCI correctly predicts it) (e.g., [26], [27], [135], [137]). The few such studies that do report mental state decoding accuracy during continuous control scenarios tend not to provide sufficient details regarding how the accuracies were obtained to allow reliable interpretation of the results (e.g., [12], [17]). This is not a criticism of these continuous control BCI studies - the focus on performance metrics over mental state decoding accuracy in such studies is necessary given that the "true labels" for the mental states are inherently unavailable during realistic operation scenarios designed to evaluate the real-world usability of the developed systems. Furthermore, these performance metrics provide a valuable measure of overall system usability, incorporating the effect of both the mental state decoding and any complementary control mechanisms in the system. That said, understanding the reliability of mental state decoding in continuous control scenarios and its potential contribution to system performance is crucial for researchers to effectively identify the source of performance deficiencies and improve the overall design.

Since it has not been specifically investigated, it is unclear how well MI tasks can be decoded when users must perform them for extended intervals, as required in the continuous control paradigm. It is important to understand whether users can sustain task performance over extended periods and, if so, whether the brain activity patterns associated with the tasks remain stable enough for accurate decoding. Furthermore, no studies have systematically compared different MI tasks to determine which ones are most intuitive, and can be most reliably decoded, in continuous control scenarios.

This study aimed to address these gaps by examining the continuous and asynchronous classification of different MI tasks versus a rest state over longer trials including multiple extended intervals (8–20 seconds) of each condition, simulating conditions typically encountered during the continuous control of external devices. Different MI tasks – the conventional tasks of hand, foot and tongue imagery, as well as a novel singing motor imagery task proposed in [28]– were compared in terms of both mental state decoding accuracy, and user preference. For mental state decoding, different analysis approaches were considered. Furthermore, the MI decoding accuracies obtained in the experimental continuous control-like scenario were compared to those achieved with the same tasks in a switch control-like scenario using the dataset from [28]. To allow comparison of different analysis methods, mental state decoding was performed offline. However, an online BCI system was simulated using methods entirely compatible with online implementation.

Chapter 3

Methods

3.1 Participants

Seventeen healthy participants (mean age 29.3 ± 3.9 years; 16 right-handed; 10 male, 7 Female) volunteered based on the inclusion criteria stated in the recruitment documents. The inclusion criteria required that participants be between 18 and 65 years old, have normal or corrected-to-normal vision and hearing, and have no history of neurological disease, disorder, or injury, as well as no cognitive impairment. Four of the participants had prior experience volunteering for BCI studies. Participants were asked to avoid smoking, drinking alcohol, consuming caffeine, and exercising for at least four hours before the experiment. They were also asked to wash their hair and avoid using hair products before the session in order to maximize EEG signal quality. Participation was entirely voluntary, and all participants provided written informed consent before the experiment. The Interdisciplinary Committee on Ethics in Human Research (ICEHR) at Memorial University of Newfoundland, NL, Canada, approved the experimental protocol (ICEHR #20240813-EN). Data from one male participant were excluded as it was discovered at the end of the session that he did not meet the inclusion criteria (had a diagnosis of ADHD). Data from another male participant was corrupted and also excluded from the analysis due to excessive noise.

3.2 Data Acquisition

A 64-channel ActiCHamp EEG system (Brain Products, GmbH), with active electrodes placed according to the International 10-10 System via a flexible cap, was used for data acquisition. Data were recorded using BrainVision Recorder at a sampling rate of 500 Hz. Electrode impedances were reduced through the injection of electrolyte gel, monitored, and kept below 10 KΩ throughout the session to maintain signal quality.

3.3 Experimental Protocol

Participants were asked to complete a single experimental session. The experimental protocol is described in detail below and summarized in Figure 3.1.

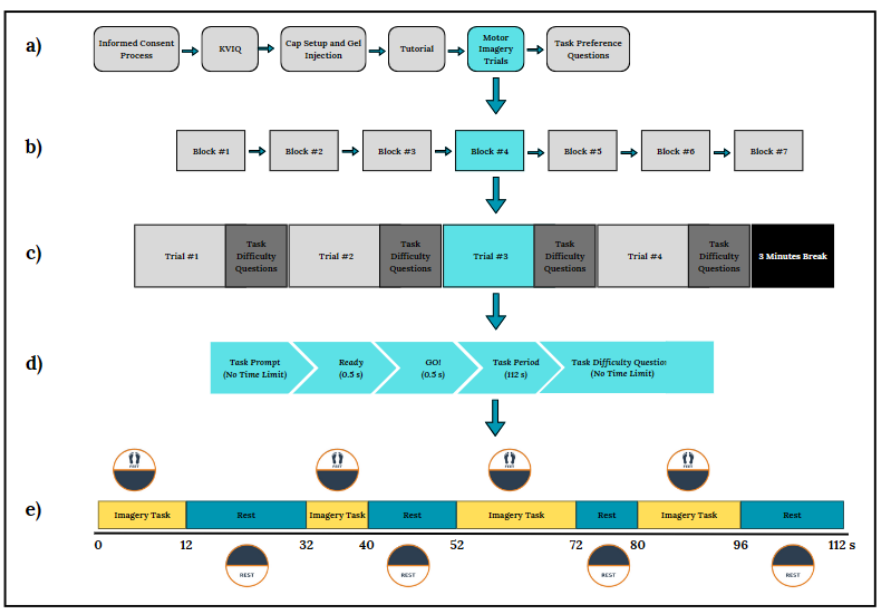


Figure 3.1: Summary of the experimental protocol, including an example trial.

The participant was seated comfortably in a quiet room, positioned approximately 40 cm in front of a computer screen. Following the informed consent process, the participant completed the Kinesthetic and Visual Imagery Questionnaire (KVIQ-10) [141], which is commonly used in BCI studies to familiarize participants with the concepts of visual and kinesthetic motor imagery and to assess their ability to perform these tasks. The participant was then fitted with the EEG cap, and electrolyte gel was applied to reduce electrode impedances below 10 KΩ. Next, they were given a tutorial outlining the experimental protocol and explaining the tasks they would be asked to perform. These tasks are described below. Note that each of the four MI tasks involved kinesthetic motor imagery, meaning that participants were specifically instructed to try to imagine the sensation associated with the specific movement rather than just visualizing the performance of the movement [55].

- **Motor Imagery of Hand (H):** Imagine tapping the dominant hand, bending at the wrist, at a rate of approximately once per second.
- **Motor Imagery of Feet (F):** Imagine tapping both feet on the ground, bending at the ankle, at a rate of approximately once per second.
- **Motor Imagery of Tongue (T):** Imagine sticking the tongue out of the mouth and pulling it back, at the rate of approximately once per second.
- **Motor Imagery of Singing (S):** Imagine singing a song, including the associated jaw, tongue, and lip movements. Participants were asked to choose four different songs that they were very familiar with from the following list of 10 widely known English songs with lyrics:

- “Happy Birthday to You”
- “The Alphabet Song”
- “Jingle Bells”
- “We Wish You a Merry Christmas”
- “The Wheels on the Bus”
- “Barbie Girl” by Aqua
- “I’ll be There for You (“Friends” theme song)” by The Rembrandts
- “We Will Rock You” by Queen
- “Wellerman (Sea Shanty)” by Nathan Evans
- “Who Let the Dogs Out” by Baha Men

The name of the song that the participant was to imagine for a particular trial was randomly selected from their chosen list and displayed on the screen prior to the start of the trial.

- **Rest (R):** Do not think about anything in particular; let your mind wander normally, but try not to perform the other motor imagery tasks; focus the gaze on the screen with eyes open, blinking normally.

The code was developed using PsychToolbox in MATLAB to present the tutorial and the cues for the experimental trials on the screen. As an example, Figure 3.2 depicts the section of the tutorial explaining the feet MI task and its associated cue. During the tutorial, the experimenter also verbally explained the task and answered any questions the participant had.

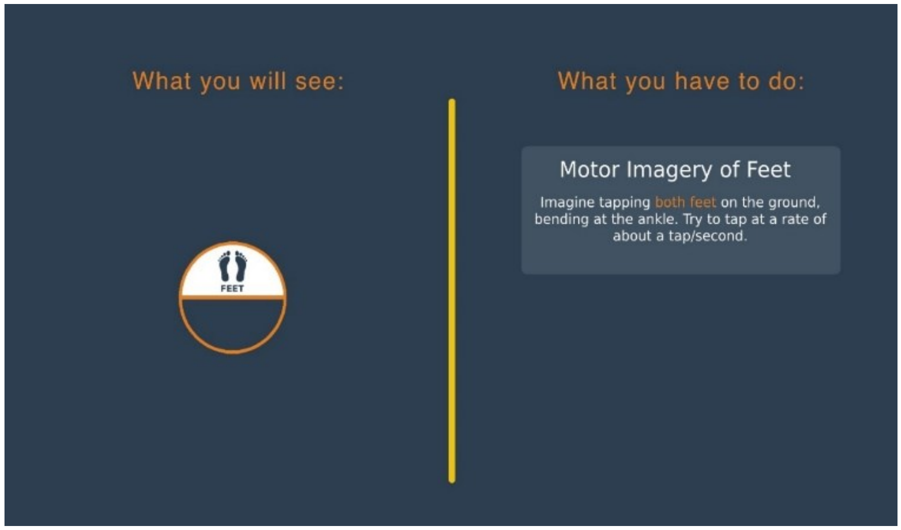


Figure 3.2: Tutorial screen explaining the motor imagery of the feet task.

Next, the participants completed the main part of the experiment, the MI trials. There was a total of 28 trials (seven trials per MI task), divided into seven blocks of four trials (one trial for each MI task per block). The task order was randomized for each block.

Figure 3.1 d shows the timing structure of each trial. Initially, the imagery task to be performed in the upcoming trial was indicated on the screen. When the participant was ready, they pressed the space bar to start the trial, and the words "Ready" and "GO!" appeared for 0.5 seconds each. This was followed by the active task period in which the participant alternated between intervals of MI task performance and rest, following visual cues shown on the screen.

Each trial had eight intervals: four for both the MI task and Rest conditions. The four intervals of each condition were of lengths 8, 12, 16, and 20 seconds, for a total trial duration of 112 seconds (i.e., (8+12+16+20) x 2 = 112). The order of the interval durations was randomized for each trial. All trials began with the MI task and ended with Rest. The paradigm was meant to simulate a continuous-control scenario where the user would control the BCI output by performing the MI task for relatively long intervals, interspersed with relatively long intervals of the rest condition.

Having the participants switch between the conditions based on cues rather than at their own discretion allowed us to know the true label of each EEG segment (either MI or Rest), which was necessary to accurately determine the decoding accuracy. Figure 3.1e shows the timing for the task period for an example trial.

After each trial, participants were asked to rate the difficulty of the imagery task on a scale of 1 to 5 (from least to most difficult). This scale is shown in Figure 3.3.

Figure 3.3: The 5-point difficulty rating scale used by participants to evaluate the imagery tasks after each trial.

After all seven blocks of MI trials were completed, participants were asked the following three questions regarding their preference for the tasks:

1. What was your favorite task?
    a. Hand MI
    b. Feet MI
    c. Tongue MI
    d. Singing MI

2. What was your least favorite task?
    a. Hand MI
    b. Feet MI
    c. Tongue MI
    d. Singing MI

3. Which task did you prefer to do?
    a. Tongue MI
    b. Singing MI

The third question was included because Tongue MI and Singing MI involve similar motor imagery processes related to movements of the mouth and tongue. The aim was to directly compare participants' preferences between these two tasks to determine which they found more intuitive, engaging, or easier to perform.

3.4 Motor Imagery Classification via EEG

3.4.1 Continuous MI Task Decoding

The primary objective of this study was to determine how accurately different MI tasks can be continuously classified from the rest state over longer intervals of these conditions than are typically investigated in BCI studies. Binary classifiers (MI vs Rest) were trained individually for each of the four MI tasks (i.e., H, F, T, S).

To simulate an online BCI scenario, the first *n* blocks of trials were used for training, while the final two blocks were used for testing. To investigate the effect of different amounts of training data on decoding accuracy, values of n=1, 2, 3, 4, and 5 (which corresponded to 56 s, 112 s, 168 s, 224 s, and 280 s of EEG data per condition, respectively) were considered. So that the classification results for all values of n training blocks could be directly compared, the final two blocks were used for testing in all cases. The amount of training data is an important factor to explore because, in real-world scenarios, collecting training data presents a burden for the BCI user. Minimizing the amount of training data needed while achieving acceptable performance is crucial for maximizing the practicality of the BCI.

The EEG data analysis pipeline is described in the following subsections. Data pre-processing was done using the EEGLAB toolbox in MATLAB, while the remainder of the analysis was done in Python using NumPy, Pandas, and Scikit-Learn. Different approaches were explored for several parts of the analysis, specifically data epoching, number of EEG features, and classification algorithm. The effects of these factors on the classification accuracy of MI tasks vs. Rest were investigated by statistical analysis. Though data was pre-recorded and the analysis was performed offline, real-time mental state decoding was simulated by ensuring all steps in the pipeline were entirely compatible with online implementation.

3.4.1.1 EEG Data Preprocessing

First, the EEG trial data were re-referenced from the FCz electrode to the average of all channels. Next, the data was down-sampled from 500 Hz to 250 Hz.

Short EEG epochs were extracted from each trial for classification into MI or Rest. Five different data windowing approaches were investigated:

1. Epochs of 4 s with no overlap (prediction every 4 seconds)
2. Epochs of 2 s with no overlap (prediction every 2 seconds)
3. Epochs of 4 s with 75% overlap (prediction every 1 second)
4. Epochs of 2 s with 50% overlap (prediction every 1 second)
5. Epochs of 1 s with no overlap (prediction every 1 second)

Each approach represented a different balance among (i) length of EEG data used to model the MI and Rest patterns, (ii) prediction frequency, and (iii) number of samples (epochs) available for classification. In all five scenarios, epochs were extracted in the same way for both the training and test trials.

3.4.1.2 Feature Extraction and Selection

Filter Bank Common Spatial Pattern (FBCSP) [103] is a widely used feature extraction method in MI-BCIs. In this study, FBCSP was employed across all classification scenarios. The process began with using linear-phase FIR bandpass filters to filter the 64-channel EEG signals into nine non-overlapping frequency bands between 4 Hz and 40 Hz, each with a 4 Hz bandwidth. The Common Spatial Pattern (CSP) algorithm was then applied to the signals within each frequency band. CSP enhances the discriminative power of EEG signals by linearly transforming them from the original channel space into spatially-filtered signals. The spatial filters used are optimized, based on the training data, to maximize variance for one class while minimizing it for another. The power of each of the spatially filtered signals was then computed to serve as features for classification. For this study, CSP was configured to produce 10 spatial filters (5 pairs) per frequency band, resulting in a total of 90 features (9 frequency bands × 10 spatially filtered signals per frequency band).

The Maximum Relevance Minimum Redundancy (mRMR) algorithm [105] was then applied to select an optimal subset of 10 features. The number of CSP components and selected features were chosen based on preliminary analysis of the data, as well as prior work reported in [28].

Both the full feature set of 90 features and the reduced 10-dimensional feature set were investigated for classification.

3.4.1.3 Classification Algorithms

Three different traditional machine learning algorithms were used for classification: linear discriminant analysis (LDA), linear support vector machines (SVM), and extreme gradient boosting (XGBoost or XGB). LDA and SVM are simple but powerful linear classifiers that have been shown to be effective for EEG-based MI classification [112]. XGBoost is a tree-based method that uses the ensemble technique to increase the classification accuracy and shows higher accuracy compared to other classifiers in many studies (e.g., [113], [114], [115]).

While deep learning methods are being increasingly explored in EEG-based BCI research [142], they were not considered in this work due to insufficient data volume.

3.4.1.4 Statistical Comparison of Four MI Tasks

For each of the four MI tasks, a total of 150 classification scenarios (5 training set sizes x 5 epoching approaches x 2 feature set sizes x 3 classifiers) were investigated per participant. Rather than base the comparison among the four MI tasks on a single scenario, to improve generalizability, they were compared based on the average accuracy across all 150 scenarios for each of the 15 participants. When the (non-parametric) Friedman test indicated that a significant difference existed among the four tasks, the Nemenyi post-hoc test was used to identify specific differences between pairs of tasks.

3.4.1.5 Statistical Comparison of Different Analysis Approaches

Friedman test and post-hoc Nemenyi test were similarly used to determine the effects of i) training set size, ii) epoching approach, iii) classification algorithm on MI classification accuracy. Wilcoxon test were used to determine the effect of number of features. For each factor, the different levels of that factor were compared based on the average accuracy across all scenarios (all combinations of the other factors) for each of the 15 participants.

3.4.2 Comparison of MI Decoding Over Longer and Shorter Task Intervals

To gain insight into how the classification of MI tasks from Rest compares when individuals must maintain these conditions over longer intervals vs. the more typical short intervals, the results obtained in this experiment were compared to a previously collected dataset, originally described in [28]. In that dataset, a separate group of 14 participants completed 70 trials each of the same four motor imagery tasks as in this experiment (i.e., H, F, T S), as well as a Rest condition. The overall experimental procedure (e.g., KVIQ and tutorial followed by MI trials, recording of task difficulty rating after each trial, posing of task preference questions at end of session), the visual cues provided to participants, and the task instructions were the same as for this experiment. The major difference compared to this experiment was the trial structure: the trials had just a single 4-second MI period that was both preceded and followed by rest, which is very typical in the EEG-based MI-BCI literature [81]. Also, two additional tasks were included in the previous experiment, which were not analyzed in the present analysis. This dataset will henceforth be referred to as the MI_{short} dataset.

As before, the classification of MI vs. Rest was done for each of the four tasks individually. The analysis pipeline for the current dataset followed the same steps as described in *section 3.4.1* but with only one option explored for each factor, specifically:

- Train/test split: Trained on first four blocks (i.e., 224 sec of data per condition), tested on remaining three blocks (i.e., 168 sec of data per condition)
- Epoching approach: 4-second epochs with 0% overlap, resulting in 56 training samples per condition (note that this was chosen to match the 4-second imagery intervals of the previous dataset)
- Number of Features: 10
- Classifier: SVM

To be comparable to the new dataset, the MI_{short} dataset was pre-processed and analyzed in exactly the same way. The first 56 trials for each of the MI tasks and the Rest were used for training, and the remaining samples were used for testing.

The classification accuracies were compared between the longer interval data (the current dataset) and the short interval data (the MI_{short} dataset [28]) using the (non-parametric) Mann-Whitney U test. Comparisons were done separately for each MI task.

3.4.3 Subjective Data Analysis

For each MI task, the average task difficulty ratings (across participants) were calculated per block, and overall. For each of the three task preference questions posed at the end of the experiment, the percentage of participants who selected each possible response was calculated.

The Pearson correlation coefficient between the average participant rating for each block and the block number was calculated to determine if participants' perceptions of difficulty for each task changed throughout the experimental session. Also, the overall average task difficulty ratings (across all blocks) for the four MI tasks were compared using the Friedman test.

Chapter 4

Results

4.1 Continuous MI Task Decoding

Figure 4.1 depicts the classification accuracies obtained for the different MI tasks (Hand, Feet, Tongue, Singing) versus Rest from the current study data (i.e., in the continuous control-like paradigm). The plot depicts the distribution, over the 15 participants, of the overall average classification accuracy across all 150 classification scenarios for each task. The Friedman test indicated that there was no significant difference in classification accuracy among the four MI tasks (χ^2=2.28, p = 0.52).

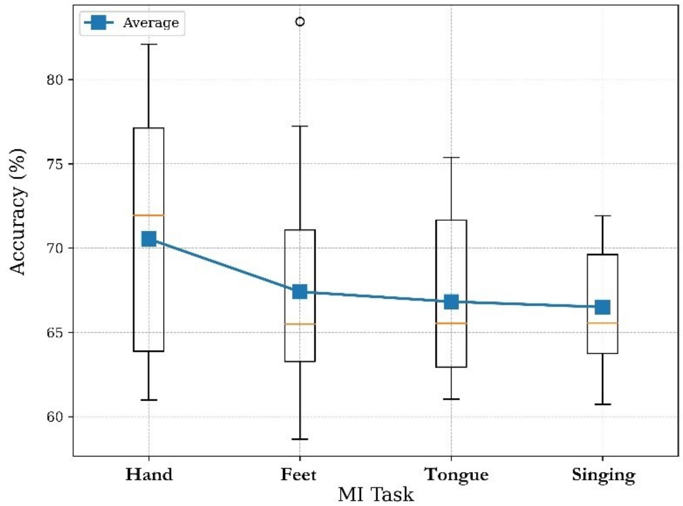


Figure 4.1: Box plot showing the MI task vs. Rest classification accuracies for the four different motor imagery tasks (Hand, Feet, Tongue, Singing) in the continuous control paradigm. There were no significant differences among the different tasks. The average classification accuracy across participants is indicated by a blue square.

Figure 4.2 depicts the distribution over participants of the overall average classification accuracies achieved (across all classification scenarios, including all four MI tasks) for the different training set sizes. An increase in accuracy is seen for each increase in training set size. The Friedman test indicated significant differences in classification accuracy among the five different training set sizes (χ^2=58.6; p<0.001). A post-hoc Nemenyi test revealed that training on five blocks produced significantly higher accuracies than training on one (p=0.001), two (p=0.001), or three (p=0.01) blocks, but there was no significant increase over using four training blocks (p=0.68). The significance of results of all pairwise comparisons, determined by post-hoc Nemenyi tests, are indicated in Figure 4.2.

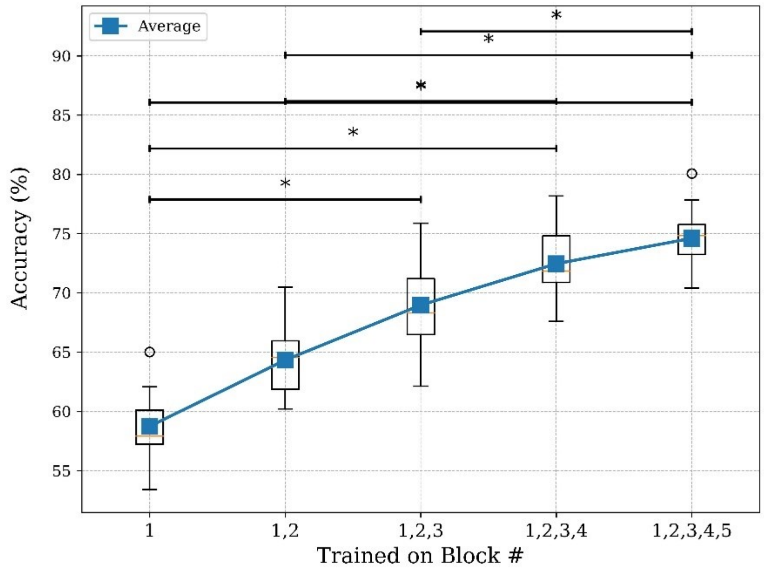


Figure 4.2: Box plot showing classification accuracy for different training set sizes. Significant differences (at $\alpha = 0.05$) between pairs of training set sizes are represented by lines connecting them. The average classification accuracy across participants is indicated by a blue square.

Figure 4.3 depicts the distribution over the 15 participants of the overall average classification accuracies achieved for the different a) epoching approaches, b) classification algorithms, and c) feature set sizes across all classification scenarios explored for each. Friedman tests indicated significant differences among the different epoching approaches (χ^2=41.23; p<0.001) and classifiers (χ^2=25.20; p<0.001); significance results for all pairwise comparisons, determined by post-hoc Nemenyi tests, are indicated in Figures 4.3a and 4.3b. A Wilcoxon test indicated a significant difference for the two feature set sizes as well, with the reduced 10-dimensional feature set producing better accuracies than the full 90-dimensional feature set (W=0; p<0.001).

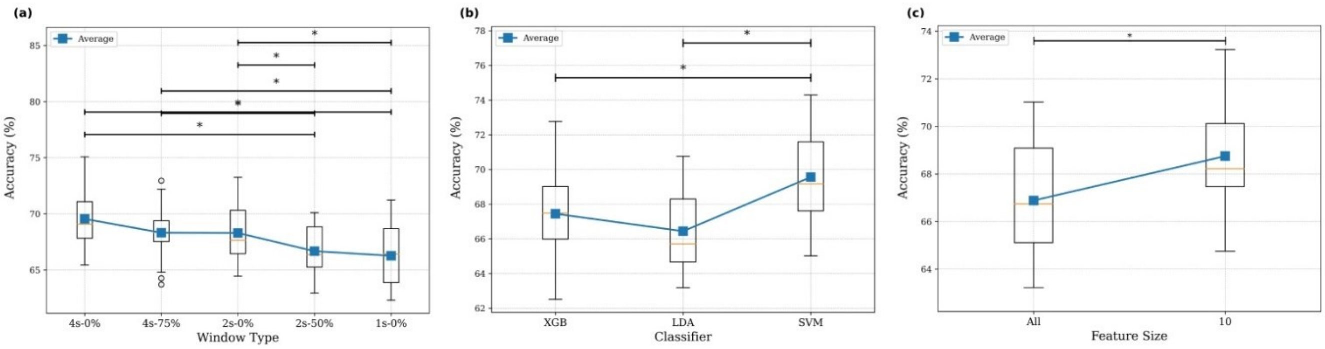


Figure 4.3: Box plots showing classification accuracies for a) different epoching approaches, b) classification algorithms, and c) feature set sizes. Significant differences between pairs, as determined by post-hoc Nemenyi tests, are indicated in Figures 4.3a and 4.3b. The average classification accuracy across participants is represented by a blue square.

The results from the analysis of each factor were used to select the best analysis pipeline with which to implement a real, online MI-BCI, considering a balance between classification accuracy and practicality in terms of the amount of time needed for classifier training, computational complexity, and prediction frequency. The selected pipeline was: four training blocks, 4-second epochs with 75% overlap, 10 features, and an SVM classifier. Figure 4.4 shows the per-participant MI task vs. Rest classification accuracies for the four MI tasks corresponding to this pipeline, as well as the overall average accuracies (Hand: 78.4% ± 9.5, Feet: 74.0% ± 10.0, Tongue: 74.5% ± 6.4, Singing: 74% ± 5.6). A Friedman test was conducted for this specific scenario, and no significant differences in decoding accuracy were found among the MI tasks (χ^2=2.90; p=0.41).

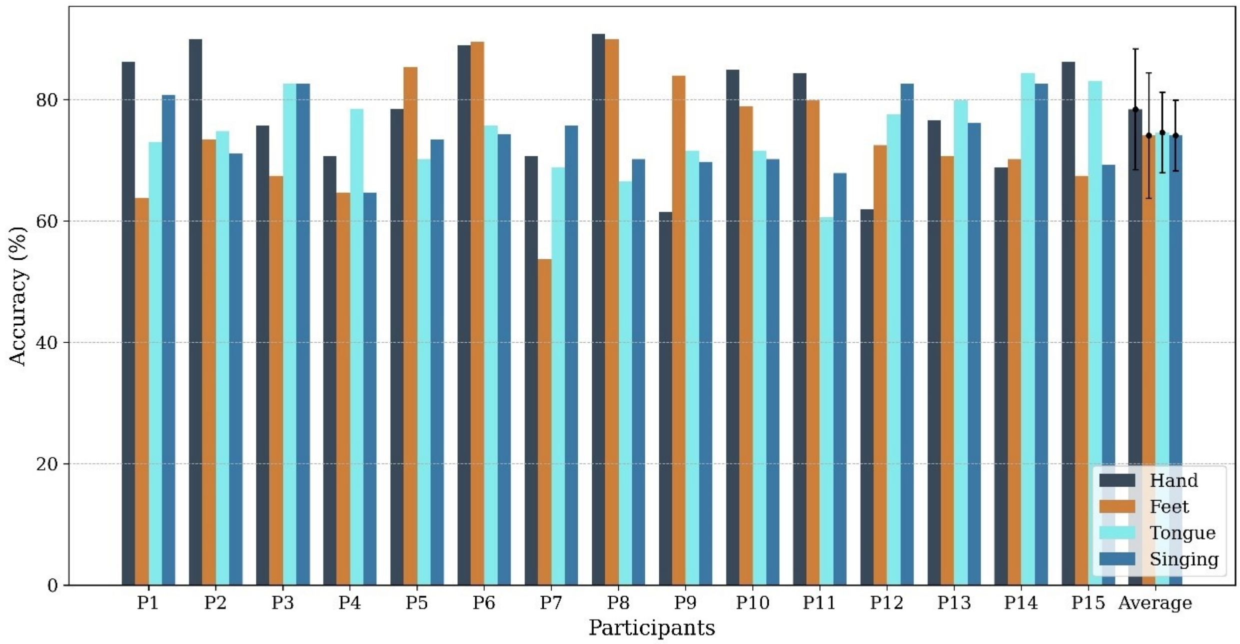


Figure 4.4: MI task vs. Rest classification accuracies for the four MI tasks (Hand, Feet, Tongue, Singing) for each participant, and on average across participants. The classification was based on using four training blocks, 4-second epochs with 75% overlap, 10 selected features, and an SVM classifier. Across participants, the Friedman test revealed no significant differences in MI task vs. Rest classification accuracies among the four MI tasks (p=0.41).

Figure 4.5 shows the confusion matrices for the classification of each of the four MI tasks vs. Rest, using the selected analysis pipeline of four training blocks, 4-second epochs with 75% overlap, 10 selected features, and an SVM classifier. Each confusion matrix provides a detailed breakdown of the classifier’s performance, with diagonal elements indicating correct classifications and off-diagonal elements representing misclassifications. The classification and misclassification counts are summed across all 15 participants, providing an aggregated view of task performance.

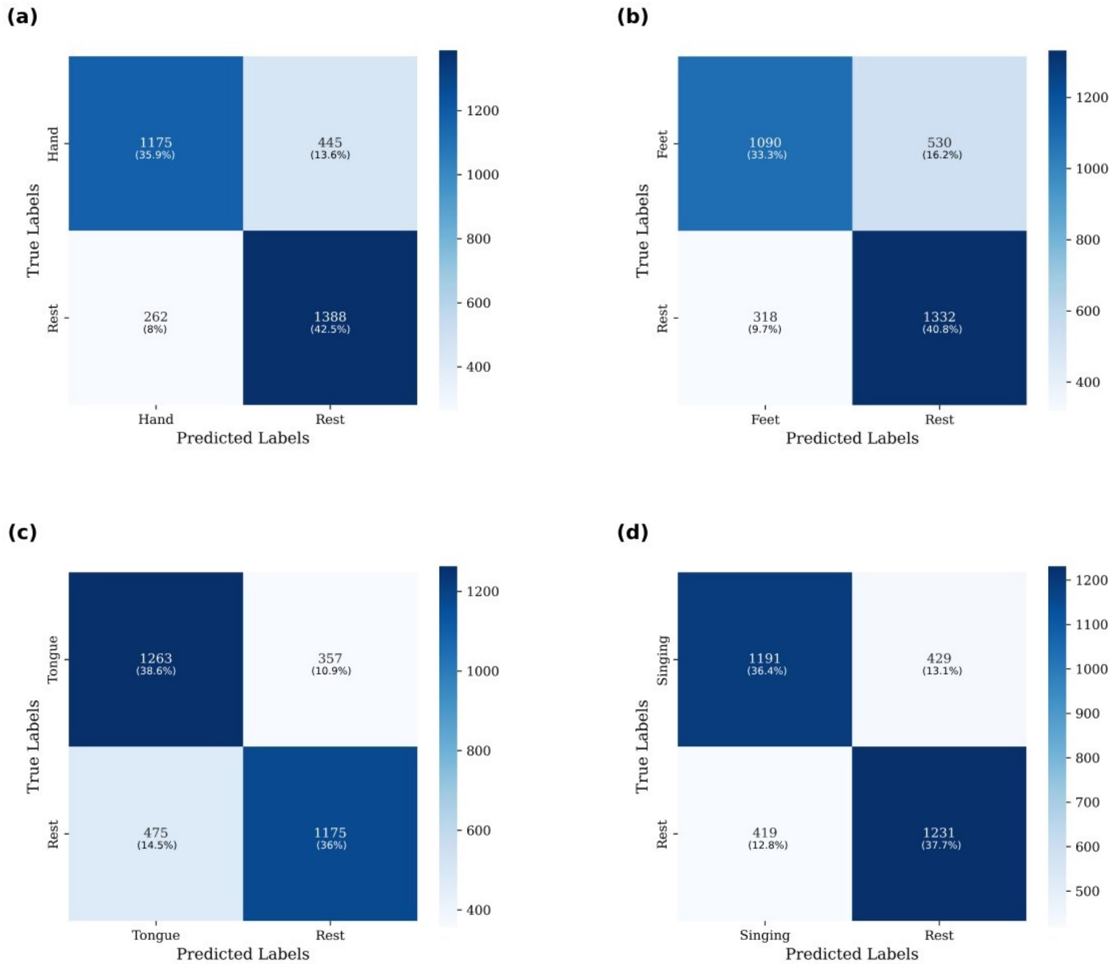
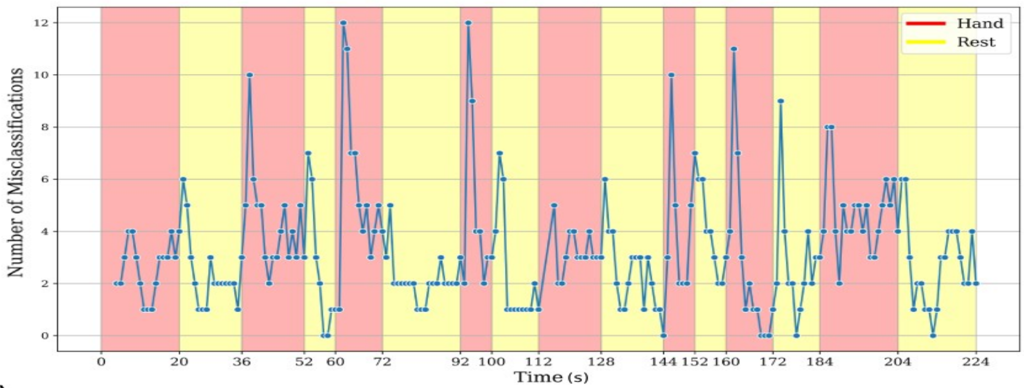


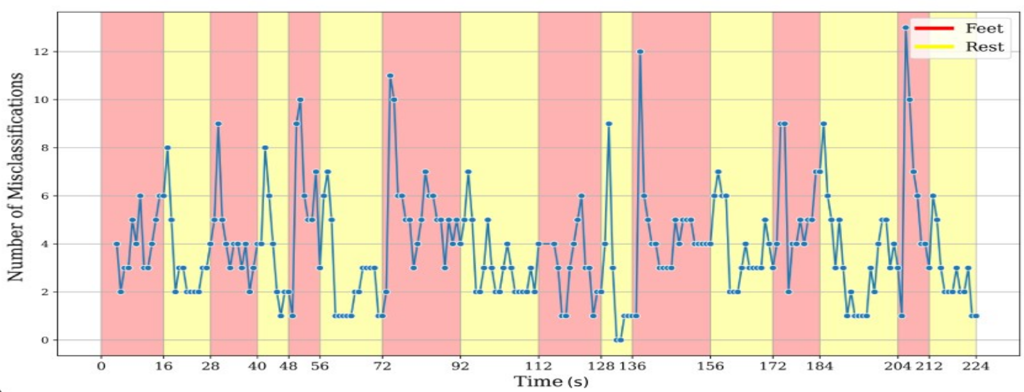
Figure 4.5: Confusion matrices for the classification of each MI task vs. Rest for a) Hand, b) Feet, c) Tongue, d) Singing. The classification was based on the selected EEG pipeline of four training blocks, 4-second epochs with 75% overlap, 10 selected features, and the SVM classifier. Diagonal elements represent correct classifications, while off-diagonal elements indicate misclassifications. These matrices summarize classification results across 15 participants, with classification and misclassification counts summed over all participants.

To understand the distribution of misclassifications over time across the task and rest intervals, the number of misclassifications, summed over the 15 participants, were plotted for each second of the two test trials (see Figure 4.6). This plot is again based on the selected analysis pipeline of four training blocks, 4-second epochs with 75% overlap, 10 selected features, and the SVM classifier.

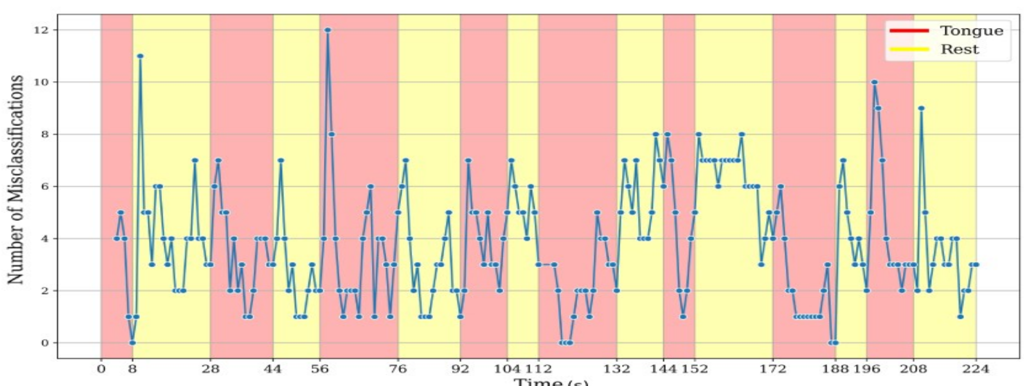
(a)


Hand
Rest
Number of Misclassifications
Time (s)
0 20 36 52 60 72 92 100 112 128 144 152 160 172 184 204 224

(b)


Feet
Rest
Number of Misclassifications
Time (s)
0 16 28 40 48 56 72 92 112 128 136 156 172 184 204 212 224

(c)


Tongue
Rest
Number of Misclassifications
Time (s)
0 8 28 44 56 76 92 104 112 132 144 152 172 188 196 208 224

(d)

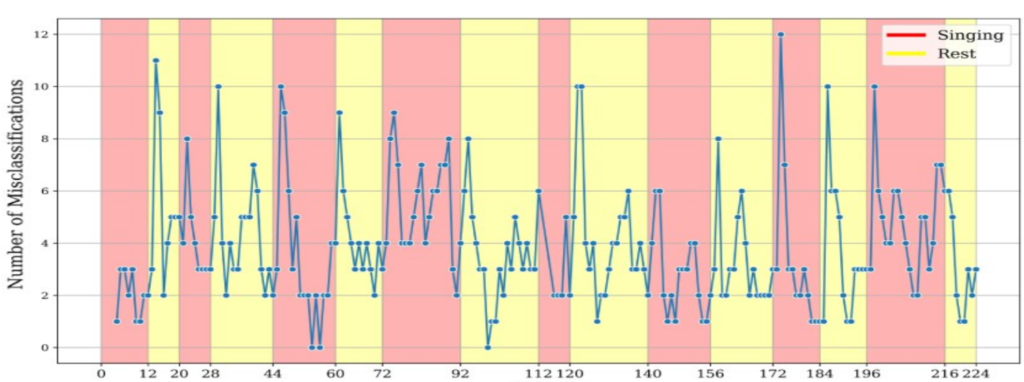

Singing
Rest
Number of Misclassifications
Time (s)
0 12 20 28 44 60 72 92 112 120 140 156 172 184 196 216 224

Figure 4.6: Total number of misclassifications per second across 15 participants for the four MI tasks: (a) Hand, (b) Feet, (c) Tongue, and (d) Singing. For each task, the predictions for the two 112-second test trials have been concatenated for a total of 224 seconds per plot. The analysis was performed using the selected analysis pipeline (four training blocks, 4-second epochs with 75% overlap, 10 selected features, and the SVM classifier). The red background represents active MI intervals, while yellow indicates Rest intervals.

4.2 Comparison of MI Decoding Over Longer and Shorter Task Intervals

Figure 4.7 shows the average MI task vs. Rest decoding accuracies across participants for each of the four MI tasks in the current dataset (with 8 to 20-second intervals) and the MI_{short} dataset (with 4-second intervals). While the accuracies for the short intervals are higher for all four tasks, the Mann-Whitney U tests indicated that the difference is significant only for the singing motor imagery task (U = 44; p=0.01).

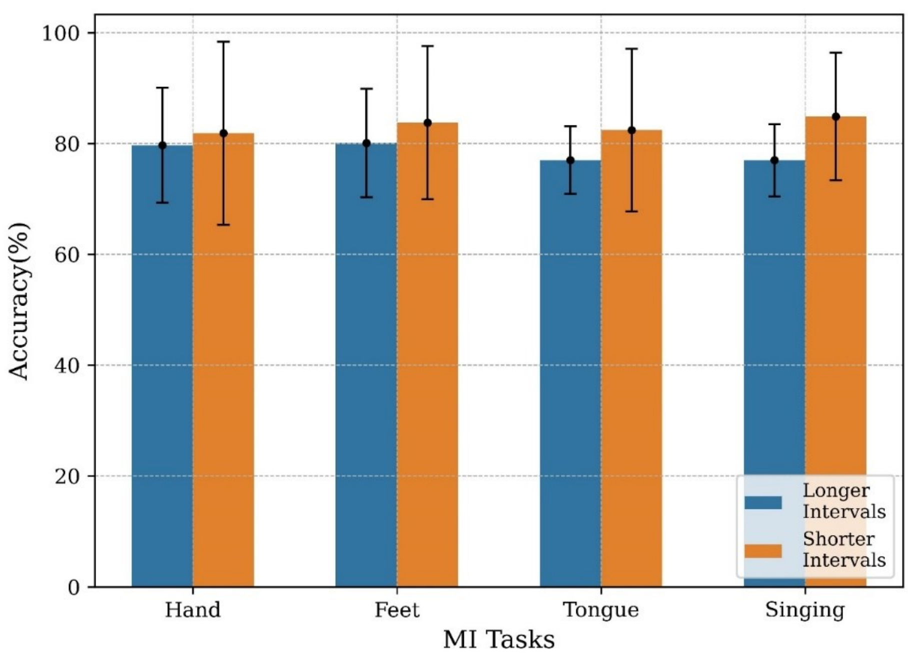


Figure 4.7: Decoding accuracies for each MI task vs. Rest, averaged across participants, for the current dataset (8-20-second intervals) and the MI_{short} dataset (4-second intervals).

4.3 Subjective Data Analysis

Figure 4.8 presents, for each MI task, the average difficulty ratings across participants for each of the seven blocks of the experiment, as well as the overall average across all blocks. The Pearson correlation analysis revealed no statistically significant correlations between block number and difficulty ratings for any of the tasks (p>0.12), indicating that participants' perception of task difficulty did not change throughout the experimental session.

The overall average participant ratings across all seven blocks for Hand, Feet, Tongue and Singing (shown in Figure 4.8) were 2.18 ± 1.05, 2.29 ± 0.90, 2.46 ± 0.85, and 2.51 ± 1.07, respectively. The Friedman test revealed no significant differences among the four tasks (χ^2=3.76, p=0.29).

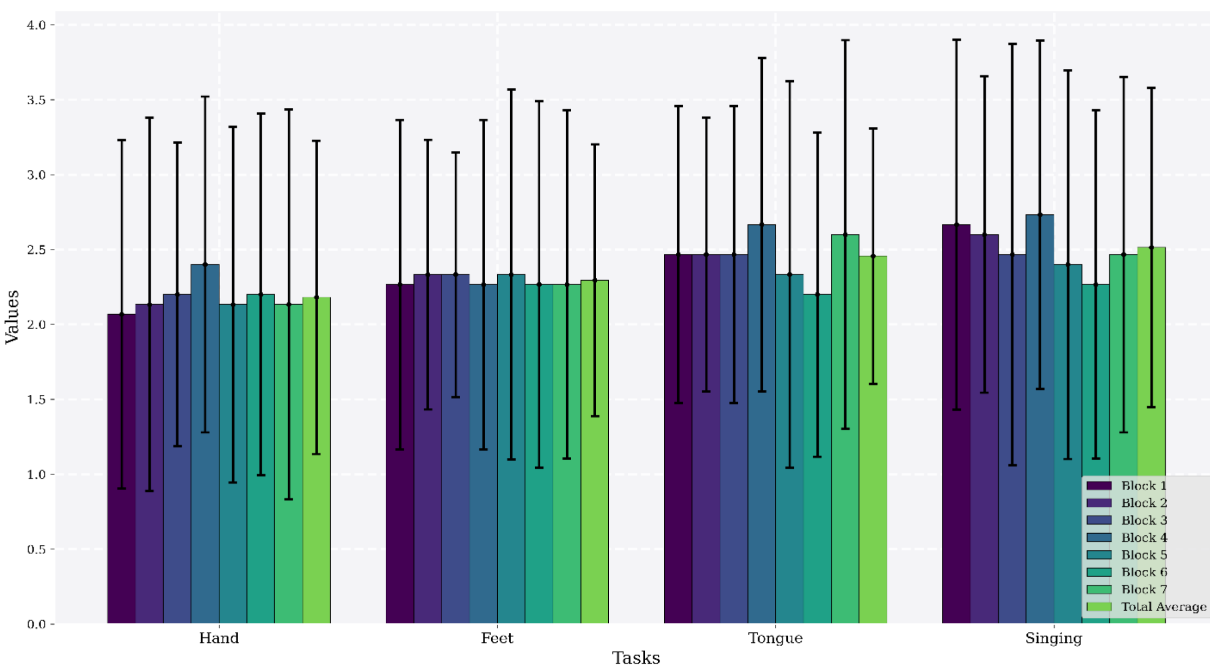


Figure 4.8: Average participant ratings of task difficulty for the four MI tasks for the seven experimental blocks, as well as the overall average across the seven blocks. Pearson correlation analysis revealed no statistically significant correlation between block number and difficulty ratings for any of the tasks (p>0.12). The Friedman test showed no significant differences among the four tasks (p = 0.29).

Figure 4.9 summarizes the participants' responses to the three task preference questions posed at the end of the experimental session. Notably, motor imagery of singing was rated most often as the favorite task (53.3% of participants) while tongue imagery was rated most often as the least favorite task (33.3% of participants). Also, 80% of participants preferred singing imagery over tongue imagery.

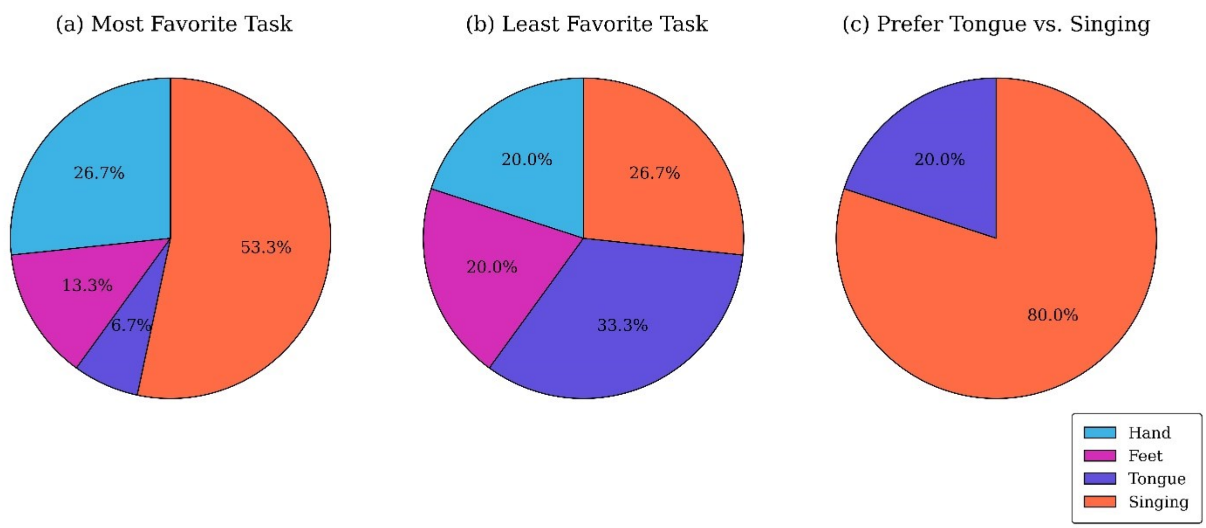


Figure 4.9: Summary of participants' responses to task preference questions posed at the end of the experimental session.

Figure 4.10 illustrates each participant's highest classification accuracy among the 4 MI tasks (top row) and their most preferred MI task (bottom row) based on subjective feedback.

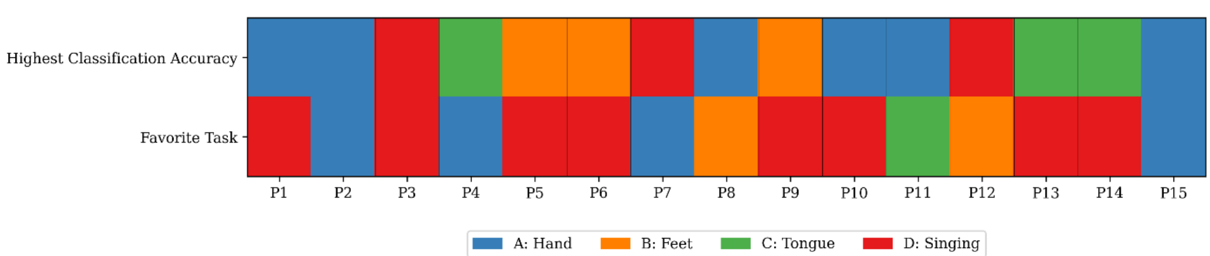


Figure 4.10: Participants' highest classification accuracy among the four MI tasks, and their favorite task.

Chapter 5

Discussion

This study explored the use of motor imagery tasks for continuous control BCI applications, focusing on the decoding accuracy of MI tasks vs. Rest over longer intervals than are typically investigated in the BCI literature, using a traditional machine learning pipeline. For different parameters (specifically, EEG window size/overlap, amount of training data, feature set size, and classifier type), different values were explored to identify the "best" configuration to balance decoding ability and practical considerations. Based on the chosen configuration, four MI tasks—Hand, Feet, Tongue, and Singing—were compared in terms of decoding accuracy and user preference. Decoding accuracy when tasks are performed over longer intervals vs. shorter intervals was compared using a previously collected dataset.

5.1 Continuous MI Task Classification

A key component of MI-BCI system design is the amount of training data required to achieve reliable classification. Unsurprisingly, in general across all classification scenarios investigated, increasing the amount of training data led to improved classification accuracy, with a marked improvement between one (112 seconds of training data) and four (448 seconds of training data) blocks (see Figure 4.2). Beyond four blocks, the accuracy gains were minimal as compared to the additional training time that would be required in a real system. This suggests that, for this study, using four blocks strikes an optimal balance between performance and practical constraints such as participant fatigue and training time. However, it is important to note that in general, other factors—such as the specific application, participant characteristics, or system complexity—may influence the ideal amount of training data. These factors should be carefully considered in different scenarios.

Figure 4.3 shows the effects on decoding accuracy of the different epoching approaches, feature set sizes, and classifiers explored. In terms of epoching, while in general the highest accuracies were achieved via a 4 second window with 0% overlap, there was no significant difference between this configuration and either a 4 second window with 75% overlap ($p=0.28$) or a 2 second window with 0% overlap ($p=0.23$). Ultimately, a 4 second window with 75% overlap was chosen as optimal for this dataset, as it provides a balance between accuracy and prediction frequency - predictions could be made every second as opposed to every 4 seconds or every 2 seconds in the other scenarios, respectively.

In terms of the feature set sizes explored, using 10 optimal features resulted in higher accuracy compared to using the full 90-dimensional feature set. This improvement is likely due to the reduction in noise and decreased likelihood of overfitting, emphasizing the importance of selecting a well-optimized subset of features for BCI applications. In terms of classifiers, SVM significantly outperformed both LDA ($p=0.001$) and XGBoost ($p=0.003$), possibly due to its robustness with small datasets and ability to handle high-dimensional data. Based on these findings, an optimal configuration comprising four training blocks, 4-second epochs with 75% overlap, 10 features, and an SVM classifier was selected and used for the remaining analysis.

Figure 4.4 shows that, using the selected analysis pipeline, comparable classification accuracies were achieved on average for the four different MI tasks versus Rest. While Hand MI had the highest average accuracy at 78.4% ± 9.5%, this result was not significantly different (χ^2=2.90, p=0.41) from that of the Feet, Tongue, and Singing tasks, which all yielded average accuracies of approximately 74-75%. The per participant results, however, show substantial inter-subject variability, in terms of accuracies for a given task. For instance, Participants 6 and 8 had the highest accuracies for the Hand and Feet tasks and the lowest for the Tongue and Singing tasks while both Participants 12 and 14 had the opposite results, achieving the highest accuracies for the Singing and Tongue tasks with the lowest for the Hand and Feet tasks. While the averages across participants suggest no statistical differences among the four tasks, these examples illustrate how factors (e.g., participant-specific neural patterns, task difficulty/preference) can significantly influence performance, even when the group averages appear comparable. From the confusion matrices (Figure 4.5), more misclassifications occurred within task states than rest states, except for Tongue MI where errors were more frequent in the rest state. This suggests that the Tongue MI task may be less distinct from the rest state, potentially due to weaker neural activity, overlapping features, or inconsistencies in user performance during motor imagery. Interestingly, reported participant preferences aligned with these observations—Tongue MI was selected most often as participants' least favorite task (33.3%) and least often as the most favorite task (6.7%) (Figure 4.9). Despite its overall average classification accuracy being comparable to other tasks, these findings highlight the importance of considering both objective performance metrics and subjective user feedback in task evaluation.

Singing MI emerged as a promising task overall with an average decoding accuracy of 74%, and with 53.3% of participants identifying it as their favorite task. Participants preferred it 4:1 to Tongue MI. The Hand MI task, which is perhaps the most commonly used task in BCI research, achieved the highest average classification accuracy (78.4%), was the most preferred task for 26% of participants, and the least preferred for only 20%. Among all participants, only three individuals who achieved particularly high accuracy in a specific task identified that task as their favorite (see Figure 4.10), which underscores the need to balance technical performance with user satisfaction.

The overall average participant difficulty ratings across all seven blocks for Hand, Feet, Tongue, and Singing tasks were 2.18 ± 1.05, 2.29 ± 0.90, 2.46 ± 0.85, and 2.51 ± 1.07, respectively, with Singing and Tongue tasks receiving the highest ratings (see Figure 4.8). The Feet and Tongue tasks had the lowest standard deviations, reflecting higher consistency in difficulty ratings across participants. Despite these differences in ratings and variability, the Friedman test revealed no significant differences in overall task difficulty ratings across tasks ($p=0.29$), indicating that on average participants did not perceive any task as significantly harder or easier than others. Additionally, Pearson correlation analysis showed no significant trends in difficulty ratings across experimental blocks ($p>0.12$), suggesting that perceptions of task difficulty remained stable throughout the experiment.

A closer examination of classification errors across the course of the MI trials provides additional insights. Figure 4.6 shows that errors were predominantly concentrated near transitions between task and rest states. This is unsurprising given the use of a 4-second window with a 75% overlap, where predictions during the first three seconds after a change of interval are based on data that still includes information from the previous condition. To address this issue, a majority voting approach was implemented, where predictions for a 4-second window (with 75% overlap) were based on votes from the past two prediction points and the current point. However, this approach did not result in any improvement in classification performance. This suggests the need for robust algorithms such as adaptive filtering methods or dynamic windowing techniques to handle transitional phases in continuous control scenarios, as these boundary areas are particularly prone to misclassification. On the other hand, this effect could be amplified due to the experimental conditions in which participants were following task cues, and there would be a slight delay between the cue changing and the participant actually switching from task to rest, or vice versa. Because the classification accuracy was calculated using true labels based on the task cues rather than what the participant was actually doing at each second during this transition period, this could contribute to some points being counted as misclassified when they actually weren't, but this issue would not exist in real, online BCI operation. Interestingly, the temporal distribution of errors did not show any notable imbalance across longer intervals, such as the 20-second durations, indicating that task-specific brain activity patterns remained stable over extended periods. This stability reinforces the feasibility of using MI tasks for sustained control in real-world BCI applications.

5.2 Comparison of MI Decoding Over Longer and Shorter Task Intervals

Figure 4.7 shows that overall, comparable decoding accuracies were achieved over longer trials (112 seconds) with alternating 8 to 20 seconds MI and rest intervals, as were achieved with short trials with discrete, 4-second intervals. Mann-Whitney U tests revealed no significant differences between the two conditions for Hand, Feet and Tongue MI, however, decoding of Singing MI vs.

Rest was significantly higher ($U=44$; $p=0.01$) for the MI_{short} dataset ($84.8\% \pm 11.5$) as compared to the current dataset ($77.0\% \pm 6.5$). The fact that the results are so close between the two datasets is encouraging for the use of continuous control in BCIs. One might have expected the accuracies to be lower in the current dataset due to several factors. Firstly, longer intervals of MI performance can lead to fatigue, which may negatively impact participants' focus and the quality of their MI signals. Additionally, longer trials, in general, introduce greater opportunities for movement artifacts and eye blinks, which can interfere with signal quality. Furthermore, extended trial durations increase the likelihood of external distractions or lapses in concentration, which can further degrade performance. In addition, trials in the MI_{Short} dataset contained a single interval of a single condition (MI or rest), meaning classification was not performed sequentially. This key difference eliminated the potential for transitional misclassifications, which are more prominent in our dataset due to the continuous analysis approach. Despite these additional challenges, the comparable accuracy levels observed in our dataset are particularly encouraging. They reinforce the robustness of the system and its ability to decode MI tasks versus rest effectively, even in continuous control-like scenarios.

Note that the differences in results reported in Figure 4.7 as compared to Figure 4.4 are attributable to slight variations in the analysis pipelines used. The results in Figure 4.4 were obtained through training and testing on the first four and last two blocks of trials, respectively, using a 4-second window with 75% overlap. The results in Figure 4.7 were obtained through training and testing on the first four and last three blocks, respectively, using a 4-second window with 0% overlap. In both cases, an SVM classifier and an optimized 10-dimensional feature set were used. The results in Figure 4.4 were based on the “optimal” pipeline selected after preliminary analysis of the factors of amount of training data, epoching approach, number of features, and classifier, while for the Figure 4.7 results the analysis was chosen to align with the MI_{short} dataset for a fair comparison.

5.3 Study Limitations and Future Work

While this study provides valuable insights into MI task vs Rest decoding accuracy and user experience over longer intervals of both conditions, several limitations must be acknowledged. EEG signals, which form the basis of classification, are inherently prone to artifacts such as muscle movements, eye blinks, and electrical interference. These factors can negatively affect classification accuracy, and the limited spatial resolution of EEG further constrains the precision of decoding brain activity. Exploring complementary modalities or advanced artifact removal techniques could improve the robustness of these systems.

Although the results across the two datasets are encouraging, it is important to note that they were derived from different experimental protocols. While the key aspects of the protocols were consistent, differences such as the involvement of different experimenters and the inclusion of additional tasks in the MI_{Short} dataset may have introduced subtle variations that limit direct comparability.

This study focused solely on the binary classification of MI task versus Rest. While this approach is practical and effective for many applications, it’s not clear how the results would translate to binary scenarios involving two MI tasks (e.g., Hand MI vs. Foot MI), or to multi-class scenarios involving three or more conditions. Multi-class systems offer greater degrees of freedom and enable more complex and versatile BCI applications. Future studies should explore these scenarios to extend our understanding of continuous mental state decoding.

Another limitation is the offline nature of the study. Although online conditions were simulated by using algorithms entirely suitable for online implementation, participants did not receive real-time feedback. Feedback mechanisms could significantly impact brain activity, either enhancing or impairing the ability to decode MI versus Rest. This effect remains unexplored and should be further investigated in a true online setting to better understand the interaction between feedback, brain activity, and classification performance. Given the real-time, continuous control goal of this study, an important next step would be testing the system in a real-world application, such as actual wheelchair control or interaction with other assistive devices, with the user in the loop.

Furthermore, the controlled laboratory setting of this study minimized distractions such as environmental noise and the multitasking demands that are typically present in real-world BCI control scenarios. While this helped isolate and focus on MI tasks, it does not fully replicate the challenges faced in practical applications. Real-world conditions may introduce additional factors that affect brain activity, user concentration, and performance, emphasizing the need for future studies in more ecologically valid settings.

Additionally, the study relied on single-session data collection, which limits insights into inter-session variability and long-term task performance. Prolonged exposure to MI tasks may reveal changes in user preferences, learning effects, or adaptation, which were not captured here. Longitudinal studies would help address these gaps and provide a more comprehensive understanding of MI-BCI usability over time.

A further direction for future research is inter-subject classification. While this study focused on intra-subject decoding for personalized modeling, inter-subject testing—where models are trained on data from multiple users and evaluated on new, unseen subjects—could support the development of subject-independent MI-BCI systems

Finally, intervals of 8–20 seconds were examined to simulate the sustained task durations most likely required for real-world applications. It is unclear how the results would translate to longer interval durations, and this should be explored in future research to better understand the durability of mental state generation and decoding for continuous BCI control.

Chapter 6

Conclusion

6.1 Primary Conclusion

This study demonstrated that all MI tasks investigated —Hand, Feet, Tongue, and Singing—could be effectively distinguished from Rest over longer intervals (8–20 seconds) than are typically investigated in BCI research studies (about 4 seconds). Under the identified optimal configuration identified for this dataset —four training blocks, 4-second epochs with 75% overlap, 10 selected features, and an SVM classifier—all tasks achieved classification accuracies in the range of 74-78%, which were comparable to those obtained in the MI_{Short} dataset. Importantly, there were no statistically significant differences among the tasks in terms of classification accuracy or task difficulty ratings, indicating that all tasks are viable options for MI-BCI systems from both technical and usability perspectives, however there is significant variability in terms of user preference. Singing MI emerged as a particularly promising task, favored by a majority of participants for its intuitiveness and ease of use.

This finding is particularly encouraging for continuous control applications, as it suggests that reliable mental state generation and decoding is possible over durations suitable for real-world scenarios. Future research should focus on exploring multi-class classification scenarios, incorporating real-time feedback mechanisms, and testing these systems in ecologically valid environments to further enhance their usability and robustness.

Bibliography

[1] K. Douibi, S. Le Bars, A. Lemontey, L. Nag, R. Balp, and G. Breda, “Toward EEG-based BCI applications for industry 4.0: Challenges and possible applications,” *Front Hum Neurosci*, vol. 15, p. 705064, 2021.

[2] T. O. Zander and C. Kothe, “Towards passive brain–computer interfaces: applying brain–computer interface technology to human–machine systems in general,” *J Neural Eng*, vol. 8, no. 2, p. 025005, 2011.

[3] G. A. Light *et al.*, “Electroencephalography (EEG) and event-related potentials (ERPs) with human participants,” *Curr Protoc Neurosci*, vol. 52, no. 1, pp. 6–25, 2010.

[4] A. M. Batula, J. A. Mark, Y. E. Kim, and H. Ayaz, “Comparison of brain activation during motor imagery and motor movement using fNIRS,” *Comput Intell Neurosci*, vol. 2017, no. 1, p. 5491296, 2017.

[5] J. Decety, “The neurophysiological basis of motor imagery,” *Behavioural brain research*, vol. 77, no. 1–2, pp. 45–52, 1996.

[6] Y. Cui, S. Xie, Y. Fu, and X. Xie, “Predicting motor imagery BCI performance based on EEG microstate analysis,” *Brain Sci*, vol. 13, no. 9, p. 1288, 2023.

[7] L. Tong, Y. Qian, L. Peng, C. Wang, and Z.-G. Hou, "A learnable EEG channel selection method for MI-BCI using efficient channel attention," *Front Neurosci*, vol. 17, p. 1276067, 2023.

[8] Y. Zhou, S. He, Q. Huang, and Y. Li, "A hybrid asynchronous brain-computer interface combining SSVEP and EOG signals," *IEEE Trans Biomed Eng*, vol. 67, no. 10, pp. 2881–2892, 2020.

[9] Y. Yu, Y. Liu, J. Jiang, E. Yin, Z. Zhou, and D. Hu, "An Asynchronous Control Paradigm Based on Sequential Motor Imagery and Its Application in Wheelchair Navigation," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 12, pp. 2367–2375, Dec. 2018, doi: 10.1109/TNSRE.2018.2881215.

[10] P. Gaur, H. Gupta, A. Chowdhury, K. McCreadie, R. B. Pachori, and H. Wang, "A sliding window common spatial pattern for enhancing motor imagery classification in EEG-BCI," *IEEE Trans Instrum Meas*, vol. 70, pp. 1–9, 2021.

[11] Z. Khademi, F. Ebrahimi, and H. M. Kordy, "A transfer learning-based CNN and LSTM hybrid deep learning model to classify motor imagery EEG signals," *Comput Biol Med*, vol. 143, p. 105288, 2022.

[12] J. Choi, K. T. Kim, J. H. Jeong, L. Kim, S. J. Lee, and H. Kim, "Developing a motor imagery-based real-time asynchronous hybrid BCI controller for a lower-limb exoskeleton," *Sensors*, vol. 20, no. 24, p. 7309, 2020.

[13] J. W. Choi, B. H. Kim, and S. Jo, "Asynchronous motor imagery brain-computer interface for simulated drone control," in *2021 9th International Winter Conference on Brain-Computer Interface (BCI)*, IEEE, 2021, pp. 1–5.

[14] Á. Fernández-Rodríguez, F. Velasco-Álvarez, M. Bonnet-Save, and R. Ron-Angevin, "Evaluation of Switch and Continuous Navigation Paradigms to Command a Brain-Controlled Wheelchair," *Front Neurosci*, vol. 12, Jun. 2018, doi: 10.3389/fnins.2018.00438.

[15] N. S. Holm and S. Puthusserypady, "An improved five class MI based BCI scheme for drone control using filter bank CSP," in *2019 7th International Winter Conference on Brain-Computer Interface (BCI)*, IEEE, 2019, pp. 1–6.

[16] Y.-H. Liu, S. Huang, H.-C. Huang, and W.-H. Peng, "Novel motor imagery-based brain switch for patients with amyotrophic lateral sclerosis: a case study using two-channel electroencephalography," *IEEE Consumer Electronics Magazine*, vol. 8, no. 2, pp. 72–77, 2019.

[17] W. Zhang, F. Sun, H. Wu, C. Tan, and Y. Ma, "Asynchronous brain-computer interface shared control of robotic grasping," *Tsinghua Sci Technol*, vol. 24, no. 3, pp. 360–370, 2019.

[18] P. Arpaia, A. Esposito, A. Natalizio, and M. Parvis, "How to successfully classify EEG in motor imagery BCI: a metrological analysis of the state of the art," *J Neural Eng*, vol. 19, no. 3, p. 031002, 2022.

[19] D. H. Lee, J. H. Jeong, H. J. Ahn, and S. W. Lee, "Design of an EEG-based Drone Swarm Control System using Endogenous BCI Paradigms," in *9th IEEE International Winter Conference on Brain-Computer Interface, BCI 2021*, Institute of Electrical and Electronics Engineers Inc., Feb. 2021. doi: 10.1109/BCI51272.2021.9385356.

[20] Y. Yu *et al.*, "Self-paced operation of a wheelchair based on a hybrid brain-computer interface combining motor imagery and P300 potential," *IEEE Transactions on Neural*

Systems and Rehabilitation Engineering, vol. 25, no. 12, pp. 2516–2526, Dec. 2017, doi: 10.1109/TNSRE.2017.2766365.

[21] D. Gwon, K. Won, M. Song, C. S. Nam, S. C. Jun, and M. Ahn, "Review of public motor imagery and execution datasets in brain-computer interfaces," *Front Hum Neurosci*, vol. 17, p. 1134869, 2023.

[22] M. Y. M. Naser and S. Bhattacharya, "Towards practical BCI-driven wheelchairs: A systematic review study," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, pp. 1030–1044, 2023.

[23] K. Glavas, K. D. Tzimourta, P. Angelidis, S. Bibi, and M. G. Tsipouras, "Brain-Computer Interface Controlled Drones: A Systematic Review," *IEEE Access*, 2024.

[24] H.-S. An, J.-W. Kim, and S.-W. Lee, "Design of an asynchronous brain-computer interface for control of a virtual avatar," in *2016 4th International Winter Conference on Brain-Computer Interface (BCI)*, IEEE, 2016, pp. 1–2.

[25] G. Prapas, K. Glavas, K. D. Tzimourta, A. T. Tzallas, and M. G. Tsipouras, "Mind the move: Developing a brain-computer interface game with left-right motor imagery," *Information*, vol. 14, no. 7, p. 354, 2023.

[26] D. Forenzo, Y. Liu, J. Kim, Y. Ding, T. Yoon, and B. He, "Integrating simultaneous motor imagery and spatial attention for EEG-BCI control," *IEEE Trans Biomed Eng*, 2023.

[27] Q. Huang, Z. Zhang, T. Yu, S. He, and Y. Li, "An EEG-/EOG-based hybrid brain-computer interface: Application on controlling an integrated wheelchair robotic arm system," *Front Neurosci*, vol. 13, p. 1243, 2019.

[28] H. Mohammadpour, "Investigating the possibilities of using singing imagery to enhance EEG-based active BCIs," M.S. thesis, Memorial Univ. Newfoundland, St. John's, NL, Canada, 2023.

[29] Y. Zhang, H. Zhang, X. Gao, S. Zhang, and C. Yang, "Uav target detection for iot via enhancing erp component by brain–computer interface system," *IEEE Internet Things J*, vol. 10, no. 19, pp. 17243–17253, 2023.

[30] S. Rajarajan, T. Kowsalya, N. S. Gupta, P. M. Suresh, P. Ilampiray, and S. Murugan, "IoT in Brain-Computer Interfaces for Enabling Communication and Control for the Disabled," in *2024 10th International Conference on Communication and Signal Processing (ICCSP)*, IEEE, 2024, pp. 502–507.

[31] A. Vavoulis, P. Figueiredo, and A. Vourvopoulos, "A review of online classification performance in motor imagery-based brain–computer interfaces for stroke neurorehabilitation," *Signals*, vol. 4, no. 1, pp. 73–86, 2023.

[32] W. Liao, J. Li, X. Zhang, and C. Li, "Motor imagery brain–computer interface rehabilitation system enhances upper limb performance and improves brain activity in stroke patients: a clinical study," *Front Hum Neurosci*, vol. 17, p. 1117670, 2023.

[33] P. Arpaia *et al.*, "Endless Runner Game in Virtual Reality Controlled by a Self-paced Brain-Computer Interface Based on EEG and Motor Imagery," in *International Conference on Extended Reality*, Springer, 2024, pp. 208–225.

[34] G. Prapas, K. Glavas, A. T. Tzallas, K. D. Tzimourta, N. Giannakeas, and M. G. Tsipouras, "Motor imagery approach for BCI game development," in *2022 7th South-East Europe*

Design Automation, Computer Engineering, Computer Networks and Social Media Conference (SEEDA-CECNSM), IEEE, 2022, pp. 1–5.

[35] X. Li, P. Chen, X. Yu, and N. Jiang, "Analysis of the relationship between motor imagery and age-related fatigue for cnn classification of the eeg data," *Front Aging Neurosci*, vol. 14, p. 909571, 2022.

[36] J. R. G. Keough *et al.*, "Fatigue in children using motor imagery and P300 brain-computer interfaces," *J Neuroeng Rehabil*, vol. 21, no. 1, p. 61, 2024.

[37] S. Arif, S. Munawar, R. R. Marie, and S. A. Shah, "Leveraging Wavelets and Deep CNN for Sleep Pattern Recognition in Road Safety: An EEG Study," in *International Conference on Recent Trends in Image Processing and Pattern Recognition*, Springer, 2023, pp. 227–241.

[38] L. A. Moctezuma, T. Abe, and M. Molinas, "EEG-based 5-and 2-class CNN for sleep stage classification," *IFAC-PapersOnLine*, vol. 56, no. 2, pp. 3211–3216, 2023.

[39] J. W. Choi, J. Park, S. Huh, and S. Jo, "Asynchronous Motor Imagery BCI and LiDAR-Based Shared Control System for Intuitive Wheelchair Navigation," *IEEE Sens J*, vol. 23, no. 14, pp. 16252–16263, 2023.

[40] W. Nie, S. Mai, Z. Yuan, J. Hu, Q. Zhang, and J. Yang, "Design of a Fatigue Driving Warning System Utilizing Brain-Computer Interface (BCI) Technology," in *2024 3rd International Symposium on Sensor Technology and Control (ISSTC)*, IEEE, 2024, pp. 190–193.

[41] S. McWeeny and E. S. Norton, "Understanding event-related potentials (ERPs) in clinical and basic language and communication disorders research: A tutorial," *Int J Lang Commun Disord*, vol. 55, no. 4, pp. 445–457, 2020.

[42] X. Li *et al.*, "Evaluation of an online SSVEP-BCI with fast system setup," *Journal of Neurorestoratology*, vol. 12, no. 2, p. 100122, 2024.

[43] P. Zeng, L. Fan, Y. Luo, H. Shen, and D. Hu, "Task-oriented EEG denoising generative adversarial network for enhancing SSVEP-BCI performance," *J Neural Eng*, vol. 21, no. 6, p. 066003, 2024.

[44] J. Pan, X. Chen, N. Ban, J. He, J. Chen, and H. Huang, "Advances in P300 brain–computer interface spellers: toward paradigm design and performance evaluation," *Front Hum Neurosci*, vol. 16, p. 1077717, 2022.

[45] P. Tufvesson and F. Heskebeck, "Automatic control of reactive brain computer interfaces," *IFAC Journal of Systems and Control*, vol. 27, p. 100251, 2024.

[46] J. H. Teo, N. S. Ahmad, and P. Goh, "Visual stimuli-based dynamic commands with intelligent control for reactive BCI applications," *IEEE Sens J*, vol. 22, no. 2, pp. 1435–1448, 2021.

[47] M. Saini and U. Satija, "State-of-the-art mental tasks classification based on electroencephalograms: a review," *Physiol Meas*, vol. 44, no. 6, p. 06TR01, 2023.

[48] H. Huang *et al.*, "Real-Time Attention Regulation and Cognitive Monitoring Using a Wearable EEG-based BCI," *IEEE Trans Biomed Eng*, 2024.

[49] A. Varshney, S. K. Ghosh, S. Padhy, R. K. Tripathy, and U. R. Acharya, "Automated classification of mental arithmetic tasks using recurrent neural network and entropy features obtained from multi-channel EEG signals," *Electronics (Basel)*, vol. 10, no. 9, p. 1079, 2021.

[50] E. Bojorges-Valdez, J. C. Echeverría, and O. Yanez-Suarez, "Evaluation of the continuous detection of mental calculation episodes as a BCI control input," *Comput Biol Med*, vol. 64, pp. 155–162, 2015.

[51] T. A. Izzuddin, N. M. Safri, and M. A. Othman, "Mental imagery classification using one-dimensional convolutional neural network for target selection in single-channel BCI-controlled mobile robot," *Neural Comput Appl*, vol. 33, no. 11, pp. 6233–6246, 2021.

[52] N. E. Anton, M. C. Ziliak, and D. Stefanidis, "Augmenting mental imagery for robotic surgery using neurofeedback: results of a randomized controlled trial," *Global Surgical Education-Journal of the Association for Surgical Education*, vol. 2, no. 1, p. 62, 2023.

[53] G. Mi. Poor, L. M. Leventhal, S. Kelley, J. Ringenberg, and S. D. Jaffee, "Thought cubes: exploring the use of an inexpensive brain-computer interface on a mental rotation task," in *The proceedings of the 13th international ACM SIGACCESS conference on Computers and accessibility*, 2011, pp. 291–292.

[54] H. Wu, S. Li, and D. Wu, "Motor imagery classification for asynchronous EEG-based brain-computer interfaces," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2024.

[55] Y. J. Yang, E. J. Jeon, J. S. Kim, and C. K. Chung, "Characterization of kinesthetic motor imagery compared with visual motor imageries," *Sci Rep*, vol. 11, no. 1, p. 3751, 2021.

[56] A. Guillot and C. Collet, *The neurophysiological foundations of mental and motor imagery*. Oxford University Press, 2010.

[57] S. Rimbert and S. Fleck, "Long-term kinesthetic motor imagery practice with a BCI: Impacts on user experience, motor cortex oscillations and BCI performances," *Comput Human Behav*, vol. 146, p. 107789, 2023.

[58] K. Glavas, K. D. Tzimourta, A. T. Tzallas, N. Giannakeas, and M. G. Tsipouras, "Empowering Individuals With Disabilities: A 4-DoF BCI Wheelchair Using MI and EOG Signals," *IEEE Access*, 2024.

[59] M. Albán-Escobar, P. Navarrete-Arroyo, D. R. De la Cruz-Guevara, and J. Tobar-Quevedo, "Assistance Device Based on SSVEP-BCI Online to Control a 6-DOF Robotic Arm," *Sensors*, vol. 24, no. 6, p. 1922, 2024.

[60] D. L. Bailey, M. N. Maisey, D. W. Townsend, and P. E. Valk, *Positron emission tomography*, vol. 2. Springer, 2005.

[61] P. M. Matthews and P. Jezzard, "Functional magnetic resonance imaging," *J Neurol Neurosurg Psychiatry*, vol. 75, no. 1, pp. 6–12, 2004.

[62] B. J. Edelman *et al.*, "Non-invasive Brain-Computer Interfaces: State of the Art and Trends," *IEEE Rev Biomed Eng*, 2024.

[63] K. J. Miller, D. Hermes, and N. P. Staff, "The current state of electrocorticography-based brain–computer interfaces," *Neurosurg Focus*, vol. 49, no. 1, p. E2, 2020.

[64] M. P. Branco, S. H. Geukes, E. J. Aarnoutse, N. F. Ramsey, and M. J. Vansteensel, "Nine decades of electrocorticography: A comparison between epidural and subdural recordings," *European Journal of Neuroscience*, vol. 57, no. 8, pp. 1260–1288, 2023.

[65] Y. Wang, X. Yang, X. Zhang, Y. Wang, and W. Pei, "Implantable intracortical microelectrodes: reviewing the present with a focus on the future," *Microsyst. Nanoeng.*, vol. 9, no. 1, p. 7, 2023.

[66] E. C. Leuthardt, D. W. Moran, and T. R. Mullen, "Defining surgical terminology and risk for brain computer interface technologies," *Front Neurosci*, vol. 15, p. 599549, 2021.

[67] C. D. Binnie and P. F. Prior, "Electroencephalography.," *J Neurol Neurosurg Psychiatry*, vol. 57, no. 11, pp. 1308–1319, 1994.

[68] R. Abiri, S. Borhani, E. W. Sellers, Y. Jiang, and X. Zhao, "A comprehensive review of EEG-based brain–computer interface paradigms," *J Neural Eng*, vol. 16, no. 1, p. 011001, 2019.

[69] P. L. Nunez, R. B. Silberstein, P. J. Cadusch, R. S. Wijesinghe, A. F. Westdorp, and R. Srinivasan, "A theoretical and experimental study of high resolution EEG based on surface Laplacians and cortical imaging," *Electroencephalogr Clin Neurophysiol*, vol. 90, no. 1, pp. 40–57, 1994.

[70] A. Ortiz-Rosario and H. Adeli, "Brain-computer interface technologies: from signal to action," *Rev Neurosci*, vol. 24, no. 5, pp. 537–552, 2013.

[71] P. Pinti *et al.*, "The present and future use of functional near-infrared spectroscopy (fNIRS) for cognitive neuroscience," *Ann N Y Acad Sci*, vol. 1464, no. 1, pp. 5–29, 2020.

[72] E. K. S. Louis *et al.*, “Electroencephalography (EEG): An Introductory Text and Atlas of Normal and Abnormal Findings in Adults,” *Children, and Infants*, pp. 1–95, 2016.

[73] T. C. Ferree, M. T. Clay, and D. M. Tucker, “The spatial resolution of scalp EEG,” *Neurocomputing*, vol. 38, pp. 1209–1216, 2001.

[74] M. Teplan, “Fundamentals of EEG measurement,” *Measurement science review*, vol. 2, no. 2, pp. 1–11, 2002.

[75] C. Fonseca *et al.*, “A novel dry active electrode for EEG recording,” *IEEE Trans Biomed Eng*, vol. 54, no. 1, pp. 162–165, 2006.

[76] K. B. E. Böcker, J. A. G. van Avermaete, and M. M. C. van den Berg-Lenssen, “The international 10–20 system revisited: Cartesian and spherical co-ordinates,” *Brain Topogr*, vol. 6, pp. 231–235, 1994.

[77] M. Soufineyestani, D. Dowling, and A. Khan, “Electroencephalography (EEG) technology applications and available devices,” *Applied Sciences*, vol. 10, no. 21, p. 7453, 2020.

[78] V. Jurcak, D. Tsuzuki, and I. Dan, “10/20, 10/10, and 10/5 systems revisited: their validity as relative head-surface-based positioning systems,” *Neuroimage*, vol. 34, no. 4, pp. 1600–1611, 2007.

[79] W. Commons, “File:EEG 10-10 system with additional information.svg — Wikimedia Commons, the free media repository,” 2023. [Online]. Available: https://commons.wikimedia.org/w/index.php?title=File:EEG_10-10_system_with_additional_information.svg&oldid=733795147

[80] D. A. Blanco-Mora, A. Aldridge, C. J. Vieira, A. Vourvopoulos, P. Figueiredo, and S. B. i Badia, "Finding the Optimal Time Window for Increased Classification Accuracy during Motor Imagery.," in *BIODEVICES*, 2021, pp. 144–151.

[81] T. Geng and J. Q. Gan, "Towards a virtual 4-class synchronous BCI using motor prediction and one motor imagery," Verlag der Technischen Universität Graz, 2008, pp. 203–207.

[82] J. A. Urigüen and B. Garcia-Zapirain, "EEG artifact removal—state-of-the-art and guidelines," *J Neural Eng*, vol. 12, no. 3, p. 031001, 2015.

[83] A. Delorme, "EEG is better left alone," *Sci Rep*, vol. 13, no. 1, p. 2372, 2023.

[84] S. Kotte and J. R. K. K. Dabbakuti, "Methods for removal of artifacts from EEG signal: A review," in *Journal of Physics: Conference Series*, IOP Publishing, 2020, p. 012093.

[85] S. Leske and S. S. Dalal, "Reducing power line noise in EEG and MEG data via spectrum interpolation," *Neuroimage*, vol. 189, pp. 763–776, 2019.

[86] K. Hirano, S. Nishimura, and S. Mitra, "Design of digital notch filters," *IEEE Transactions on Communications*, vol. 22, no. 7, pp. 964–970, 1974.

[87] N. Retdian and T. Shima, "Power line noise suppression using N-Path Notch Filter for EEG," in *2016 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*, IEEE, 2016, pp. 1–4.

[88] X. Jiang, G.-B. Bian, and Z. Tian, "Removal of artifacts from EEG signals: a review," *Sensors*, vol. 19, no. 5, p. 987, 2019.

[89] S. Makeig, A. Bell, T.-P. Jung, and T. J. Sejnowski, "Independent component analysis of electroencephalographic data," *Adv Neural Inf Process Syst*, vol. 8, 1995.

[90] A. K. Singh and S. Krishnan, "Trends in EEG signal feature extraction applications," *Front Artif Intell*, vol. 5, p. 1072801, 2023.

[91] S. Krishnan, *Biomedical signal analysis for connected healthcare*. Academic Press, 2021.

[92] D. Pawar and S. N. Dhage, "Feature extraction methods for electroencephalography based brain-computer interface: a review," *Entropy*, vol. 1, p. 4, 2020.

[93] O. Dressler, G. Schneider, G. Stockmanns, and E. F. Kochs, "Awareness and the EEG power spectrum: analysis of frequencies," *Br J Anaesth*, vol. 93, no. 6, pp. 806–809, 2004.

[94] S. Long, R. Ding, J. Wang, Y. Yu, J. Lu, and D. Yao, "Sleep Quality and Electroencephalogram Delta Power. Front Neurosci. 2021; 15: 803507."

[95] W. Klimesch, "EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis," *Brain Res Rev*, vol. 29, no. 2–3, pp. 169–195, 1999.

[96] C. S. Herrmann, D. Strüber, R. F. Helfrich, and A. K. Engel, "EEG oscillations: from correlation to causality," *International Journal of Psychophysiology*, vol. 103, pp. 12–21, 2016.

[97] N. Arunkumar, K. R. Kumar, and V. Venkataraman, "Entropy features for focal EEG and non focal EEG," *J Comput Sci*, vol. 27, pp. 440–444, 2018.

[98] S. M. Pincus, "Approximate entropy as a measure of system complexity.," *Proceedings of the National Academy of Sciences*, vol. 88, no. 6, pp. 2297–2301, 1991.

[99] A. R. Hassan and M. I. H. Bhuiyan, "Automatic sleep scoring using statistical features in the EMD domain and ensemble methods," *Biocybern Biomed Eng*, vol. 36, no. 1, pp. 248–255, 2016.

[100] A. Yuksel and T. Olmez, “Filter bank common spatio-spectral patterns for motor imagery classification,” in *Information Technology in Bio-and Medical Informatics: 7th International Conference, ITBAM 2016, Porto, Portugal, September 5-8, 2016, Proceedings 7*, Springer, 2016, pp. 69–84.

[101] K. K. Ang, Z. Y. Chin, C. Wang, C. Guan, and H. Zhang, “Filter bank common spatial pattern algorithm on BCI competition IV datasets 2a and 2b,” *Front Neurosci*, vol. 6, p. 39, 2012.

[102] Q. Novi, C. Guan, T. H. Dat, and P. Xue, “Sub-band common spatial pattern (sbcsp) for brain-computer interface. In 2007 3rd International IEEE,” in *EMBS Conference on Neural Engineering*, pp. 204–207.

[103] K. K. Ang, Z. Y. Chin, H. Zhang, and C. Guan, “Filter bank common spatial pattern (FBCSP) in brain-computer interface,” in *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, IEEE, 2008, pp. 2390–2397.

[104] I. Guyon and A. Elisseeff, “An introduction to variable and feature selection,” *Journal of machine learning research*, vol. 3, no. Mar, pp. 1157–1182, 2003.

[105] H. Peng, F. Long, and C. Ding, “Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy,” *IEEE Trans Pattern Anal Mach Intell*, vol. 27, no. 8, pp. 1226–1238, 2005.

[106] M. Y. Hossain and A. B. M. S. U. Doulah, “Detection of motor imagery (MI) event in Electroencephalogram (EEG) signals using artificial intelligence technique,” in *2020 IEEE East-West Design & Test Symposium (EWDTS)*, IEEE, 2020, pp. 1–6.

[107] M. H. Kabir, S. Mahmood, A. Al Shiam, A. S. Musa Miah, J. Shin, and M. K. I. Molla, "Investigating feature selection techniques to enhance the performance of EEG-based motor imagery tasks classification," *Mathematics*, vol. 11, no. 8, p. 1921, 2023.

[108] M. Radman, A. Chaibakhsh, N. Nariman-Zadeh, and H. He, "Feature fusion for improving performance of motor imagery brain-computer interface system," *Biomed Signal Process Control*, vol. 68, p. 102763, 2021.

[109] Z. Tang, C. Li, and S. Sun, "Single-trial EEG classification of motor imagery using deep convolutional neural networks," *Optik (Stuttg)*, vol. 130, pp. 11–18, 2017.

[110] R. Alazrai, M. Abuhijleh, H. Alwanni, and M. I. Daoud, "A deep learning framework for decoding motor imagery tasks of the same hand using EEG signals," *IEEE Access*, vol. 7, pp. 109612–109627, 2019.

[111] R. Dhiman, "Machine learning techniques for electroencephalogram based brain-computer interface: A systematic literature review," *Measurement: Sensors*, vol. 28, p. 100823, 2023.

[112] N. Vivaldi, M. Caiola, K. Solarana, and M. Ye, "Evaluating performance of eeg data-driven machine learning for traumatic brain injury classification," *IEEE Trans Biomed Eng*, vol. 68, no. 11, pp. 3205–3216, 2021.

[113] A. Tiwari and A. Chaturvedi, "A multiclass EEG signal classification model using spatial feature extraction and XGBoost algorithm," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, IEEE, 2019, pp. 4169–4175.

[114] M. Vijay, A. Kashyap, A. Nagarkatti, S. Mohanty, R. Mohan, and N. Krupa, "Extreme gradient boosting classification of motor imagery using common spatial patterns," in *2020 IEEE 17th India Council International Conference (INDICON)*, IEEE, 2020, pp. 1–5.

[115] T. Thenmozhi and R. Helen, "Feature selection using extreme gradient boosting Bayesian optimization to upgrade the classification performance of motor imagery signals for BCI," *J Neurosci Methods*, vol. 366, p. 109425, 2022.

[116] H. Kim, B. L. Drake, and H. Park, "Multiclass classifiers based on dimension reduction with generalized LDA," *Pattern Recognit*, vol. 40, no. 11, pp. 2939–2945, 2007.

[117] A. J. Izenman, "Linear discriminant analysis," in *Modern multivariate statistical techniques: regression, classification, and manifold learning*, Springer, 2013, pp. 237–280.

[118] A. Subasi and M. I. Gursoy, "EEG signal classification using PCA, ICA, LDA and support vector machines," *Expert Syst Appl*, vol. 37, no. 12, pp. 8659–8666, 2010.

[119] R. Deore, B. Gawali, and S. Mehrotra, "Digit Recognition System Using EEG Signal," *Brain-Computer Interfaces: Current Trends and Applications*, pp. 375–416, 2015.

[120] P. K. Upadhyay and C. Nagpal, "Sleep stage and heat stress classification of rodents undergoing high environmental temperature," in *Computational Methods and Data Engineering: Proceedings of ICMDE 2020, Volume 1*, Springer, 2021, pp. 577–587.

[121] M. Premalatha and C. V Lakshmi, "Machine Learning Techniques for SVM Classification and Regression Loss Function," *Artificial Intelligent Systems and Machine Learning*, vol. 3, pp. 483–490, 2011, [Online]. Available: https://api.semanticscholar.org/CorpusID:123672713

[122] I. H. Sarker, "Machine learning: algorithms, real-world applications and research directions. SN Comput Sci 2: 160," 2021, *ed*.

[123] S. Suthaharan and S. Suthaharan, "Support vector machine," *Machine learning models and algorithms for big data classification: thinking with examples for effective learning*, pp. 207–235, 2016.

[124] T. Hastie, R. Tibshirani, and J. Friedman, "The elements of statistical learning: data mining, inference, and prediction," 2017, *Springer*.

[125] M. Sha'Abani, N. Fuad, N. Jamal, and M. F. Ismail, "kNN and SVM classification for EEG: a review," in *InECCE2019: Proceedings of the 5th International Conference on Electrical, Control & Computer Engineering, Kuantan, Pahang, Malaysia, 29th July 2019*, Springer, 2020, pp. 555–565.

[126] S. Fatima, A. Hussain, S. Bin Amir, S. H. Ahmed, and S. M. H. Aslam, "Xgboost and random forest algorithms: an in depth analysis," *Pakistan Journal of Scientific Research*, vol. 3, no. 1, pp. 26–31, 2023.

[127] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 785–794.

[128] S. R. Liyanage, C. Guan, H. Zhang, K. K. Ang, J. Xu, and T. H. Lee, "Dynamically weighted ensemble classification for non-stationary EEG processing," *J Neural Eng*, vol. 10, no. 3, p. 036007, 2013.

[129] C.-H. Han, K.-R. Müller, and H.-J. Hwang, “Brain-switches for asynchronous brain–computer interfaces: a systematic review,” *Electronics (Basel)*, vol. 9, no. 3, p. 422, 2020.

[130] S. G. Mason and G. E. Birch, “A brain-controlled switch for asynchronous control applications,” *IEEE Trans. Biomed. Eng.*, vol. 47, no. 10, pp. 1297–1307, Oct. 2000.

[131] D. Puanhvuan and Y. Wongsawat, “Semi-automatic P300-based brain-controlled wheelchair, Complex Medical Engineering,” 2012, *IEEE*.

[132] R. Zhang *et al.*, “Control of a wheelchair in an indoor environment based on a brain–computer interface and automated navigation,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, no. 1, pp. 128–139, 2015.

[133] S. Saeedi, R. Chavarriaga, and J. del R. Millán, “Long-term stable control of motor-imagery BCI by a locked-in user through adaptive assistance,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 4, pp. 380–391, 2016.

[134] L. Tonin, F. C. Bauer, and J. del R. Millán, “The role of the control framework for continuous teleoperation of a brain–machine interface-driven mobile robot,” *IEEE Transactions on Robotics*, vol. 36, no. 1, pp. 78–91, 2019.

[135] F. Velasco-Álvarez, R. Ron-Angevin, and M. J. Blanca-Mena, “Free virtual navigation using motor imagery through an asynchronous brain–computer interface,” *Presence*, vol. 19, no. 1, pp. 71–81, 2010.

[136] B. Xu, D. Liu, M. Xue, M. Miao, C. Hu, and A. Song, “Continuous shared control of a mobile robot with brain–computer interface and autonomous navigation for daily assistance,” *Comput Struct Biotechnol J*, vol. 22, pp. 3–16, 2023.

[137] K. LaFleur, K. Cassady, A. Doud, K. Shades, E. Rogin, and B. He, "Quadcopter control in three-dimensional space using a noninvasive motor imagery-based brain–computer interface," *J Neural Eng*, vol. 10, no. 4, p. 046003, 2013.

[138] Y. An, T. Shi, L. Ren, W. Liu, and X. Jiang, "UAV control in 2D space based on brain computer interface," in *2017 4th International Conference on Systems and Informatics (ICSAI)*, IEEE, 2017, pp. 594–598.

[139] T. Shi, H. Wang, and C. Zhang, "Brain Computer Interface system based on indoor semi-autonomous navigation and motor imagery for Unmanned Aerial Vehicle control," *Expert Syst Appl*, vol. 42, no. 9, pp. 4196–4206, 2015.

[140] S. He *et al.*, "EEG- And EOG-Based Asynchronous Hybrid BCI: A System Integrating a Speller, a Web Browser, an E-Mail Client, and a File Explorer," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 2, pp. 519–530, Feb. 2020, doi: 10.1109/TNSRE.2019.2961309.

[141] F. Malouin, C. L. Richards, P. L. Jackson, M. F. Lafleur, A. Durand, and J. Doyon, "The kinesthetic and visual imagery questionnaire (KVIQ) for assessing motor imagery in persons with physical disabilities: A reliability and construct validity study," *Journal of Neurologic Physical Therapy*, 2007, doi: 10.1097/01.NPT.0000260567.24122.64.

[142] K. M. Hossain, M. A. Islam, S. Hossain, A. Nijholt, and M. A. R. Ahad, "Status of deep learning for EEG-based brain–computer interface applications," *Front Comput Neurosci*, vol. 16, p. 1006763, 2023.