Pseudo-online Detection and Classification for Upper-limb Movements from Scalp Electroencephalogram

by

Jiansheng Niu

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Systems Design Engineering

Waterloo, Ontario, Canada, 2021

AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Stroke has been a significant healthcare issue worldwide, leading to motor impairment and complicated rehabilitation procedures, which often last for years after lesion. In recent years, brain-computer interface (BCI) research shed some light on new approaches for motor ability recovery and potential neural plasticity inducement for stroke patients. Electroencephalogram (EEG) is widely used in BCI to measure brain activity. In this thesis study, nine healthy participants were recruited to perform four movements in a self-initiated manner, including left wrist extension (WE_L), right wrist extension (WE_R), left index finger extension (IE_L), and right index finger extension (IE_R). A hierarchical structure was proposed first to detect movement intentions from the rest state and then classify different movement types. Movement-related cortical potential (MRCP) and sensorimotor rhythm (SMR) were believed to associate with movement intention generation in human EEG. Thus, three frequency bands of EEG (0.05-5Hz, 5-40Hz, 0.05-40Hz) containing MRCP or SMR were investigated to provide features for detection and classification algorithms. In detection, a majority voting-based ensemble learning method was proposed to integrate the strongness of three algorithms, including support vector machine (SVM), EEGNET, and Riemannian feature-based SVM. The proposed method achieved an average true positive rate (TPR) of 79.6% ± 8.8%, false positives per minute (FPs/min) as 3.1 ± 1.2 within a latency of 91.4 ± 111.9ms. For classification, an adaptive boosting-based ensemble learning algorithm was proposed to classify movement pairs and four movements in pseudo-online and time-locked analyses. As a result, It proved the feasibility of classifying movements in different arms with higher than significant chance level accuracy. In summary, the proposed system offered a novel solution to decode upper-limb movements for rehabilitation-aimed BCI.

Acknowledgements

I would like to express my sincere gratitude to my supervisor Professor Ning Jiang for his limitless patience and guidance throughout my master's degree journey. This thesis work encountered countless obstacles during the pandemic, and it would be impossible to finish this study without his generous help and continuous support. Furthermore, I have been lucky to be a student of Professor Ning Jiang to learn and explore this interdisciplinary field between engineering and neuroscience.

I would also like to thank my thesis reader Professor Shi Cao and Professor Aimee Nelson, for their valuable suggestions and feedback.

I would like to thank my lab colleagues for the help they provided during data collection. I appreciate the supportive and creative academic environment of the lab.

I would like to thank my parents and my girlfriend for their relentless love and encouragement in pursuing this degree.

Table of Contents

AUTHOR'S DECLARATION ii
Abstract iii
Acknowledgements iv
Table of Contents v
List of Figures vii
List of Tables viii
Chapter 1 Introduction 1
Chapter 2 Background 2
2.1 Stroke, Neural Plasticity, and Rehabilitation 2
2.2 EEG-based Brain-Computer Interface 5
2.2.1 Signal Acquisition 5
2.2.2 Feature Extraction 6
2.2.3 Feature Selection 7
2.2.4 Classification Algorithms 8
2.3 Movement-related BCI Features 10
Chapter 3 Methodology 13
3.1 Expriment and preprocessing 13
3.1.1 Participants 13
3.1.2 Instrumentation 14
3.1.3 Experiment Protocol 15
3.1.4 Preprocessing 16
3.2 Algorithms 18
3.2.1 EEGNET 18
3.2.2 Support Vector Machine 19
3.2.3 Adaptive Boosted Logistic Regression 20
3.2.4 Riemannian Geometry 21
3.2.5 Detection Ensemble Learning 23
3.3 Analyses 24

3.3.1 Pseudo-online Analysis 25
3.3.2 Time-locked Classification Analysis 29
Chapter 4 Results 30
4.1 MRCP 30
4.2 ERD/ERS 32
4.3 Pseudo-online Detection 34
4.3.1 Relationship between Detection Algorithms and Frequency Bands 34
4.3.2 Impact of Window Number on Detection Performance 36
4.3.3 Detection Performance with One Window 39
4.3.4 Influence of Four Movements on Detection Performance 41
4.3.5 Trial Distribution of Detection Latency 43
4.4 Pseudo-online Classification 45
4.5 Time-locked Classification 47
Chapter 5 Discussion 50
5.1 Influence of Frequency Band on Detection and Classification 50
5.2 Feasibility of Ensemble Learning 52
5.3 Detection 53
5.4 Classification 55
5.5 Study Limitation and Future investigation 57
Chapter 6 Conclusion 58
References 59

List of Figures

Figure 1. Stroke rehabilitation with time 3
Figure 2. MRCP and various sub-components for real and imaginary movement. 10
Figure 3. EEG and EMG electrodes position 14
Figure 4. Experiment protocol 15
Figure 5. EEG and EMG preprocessing 16
Figure 6. EEGNET architecture 18
Figure 7. SVM illustration 19
Figure 8. AdaBoost diagram 20
Figure 9. Riemannian manifold and tangent space 21
Figure 10. Ensemble learning structure 23
Figure 11. Training and testing procedures for pseudo-online and time-locked analysis 24
Figure 12. MRCP of four movements 30
Figure 13. ERD/ERS of high band EEG for four movements. 32
Figure 14. Violin plots of F1 score for four algorithms in three frequency bands. 34
Figure 15. ROC curves of four algorithms 36
Figure 16. Detection latency variation with window number change 37
Figure 17. Boxplots of TPR, FPs/min and latency of 9 subjects for 4 algorithms with statistical grouping. 40
Figure 18. Boxplots of TPR and latency using the ensemble method for four classes with statistical grouping. 41
Figure 19. Histograms of detection latency for nine subjects and their mean. 43
Figure 20. Accurcay and confusion matrix of three types of classification 45
Figure 21. 4-class time-locked classification using R_Ada_Lgr and sLDA 47
Figure 22. Pairwise time-locked classification using R_Ada_Lgr. 48

List of Tables

Table 1. Detection performance of nine subjects using Ensemble method, EEGNT, R_SVM and SVM 39

Table 2. TPR and latency of WE_L, WE_R, IE_L, IE_R using Ensemble method……....41

Chapter 1

Introduction

Stroke rehabilitation is a global healthcare challenge that requires efforts and resources from all societies, including organized inpatient care and therapy-based rehabilitation at home [1]. Brain-computer interface (BCI) provides a potential solution for motor training after stroke by measuring movement intentions from brain through electroencephalogram (EEG) and translates them to control commands for assistive devices. Two types of EEG modalities have been related to movement planning, execution, and imagination: movement-related cortical potential (MRCP) and sensorimotor rhythm (SMR), reflecting different frequency band features of EEG [2][3]. The respective state-of-the-art decoding algorithms of these two types of EEG modalities were investigated to design a robust BCI system capable of detecting movement intention from rest state and distinguishing multiple types of movements. In the current research, I targeted on decoding of four distal simple movements, including left wrist extension (WE_L), right wrist extension (WE_R), left index finger extension (IE_L), and right index finger extension (IE_L), which were regarded as more challenging than proximal compound movements in BCI field [4][5]. To the author's best knowledge, these four movements were never systematically compared in the literature despite their popularity in stroke rehabilitation therapy [1]. In order to validate the feasibility of real-time BCI targeted for rehabilitation applications, this research conducted analysis in pseudo-online and time-locked fashion with data acquired from healthy participants. There are two primary objectives:

1. Detect intentions of the four movements from rest state in a pseudo-online manner and investigate the optimal way of combining algorithms and frequency band features (MRCP and SMR) to achieve a high detection rate with low latency.
2. Evaluate the possibility of classifying the four movements against each other following the detection in a pseudo-online manner and understand the contribution of MRCP and SMR during classification.

This research provided a novel EEG-based BCI solution for distal simple movements decoding by investigating the ensemble of both algorithms and frequency bands in pseudo-online fashion.

Chapter 2

Background

2.1 Stroke, Neural Plasticity, and Rehabilitation

Stroke is a neurological deficit led to an acute focal injury of the central nervous system caused by vascular problems such as cerebral infarction and intracerebral hemorrhage[1]. The elderly population with hypertension, diabetes mellitus, cardiovascular diseases have more risk of this disease [7]. It is recognized as one of the most serious global health issues, ranked the second most common cause of death worldwide, with a mortality rate of around 5.5 million deaths per year [7]. Apart from the high mortality rate, it is also the first cause of adult-onset disability in developed countries, which leads to 50% of survivors being disabled [7][8]. After half a year, only 60% of those patients who received the inpatient rehabilitation can achieve simple activities of daily living (ADL) such as toileting and walking short distances, and only half of the stroke survivors could resume their professional careers [9][10][11].

It is proved that the mechanism for the acquisition, retention, and retrieval of information in the stroke patients' brain is basically the same as that in a healthy brain apart from partially damaged neural pathways in stroke survivors [12]. Some studies showed that it is possible to rebuild neuropathways by taking advantage of the residual functioning part of the injured brain through training [12][13][14]. The rebuilding of neuropathway is based on neural plasticity, which is defined as "the ability of the nervous system to respond to intrinsic or extrinsic stimuli by reorganizing its structure, connections, and function." [15]. Different types of plasticity, including Hebbian plasticity and homeostatic plasticity, help the brain quickly adapt to changes in a stabilized process. Hebbian plasticity influences the synaptic strength and efficacy by varying the number of dendritic spines and synapses in fast response to a stimulus at a seconds

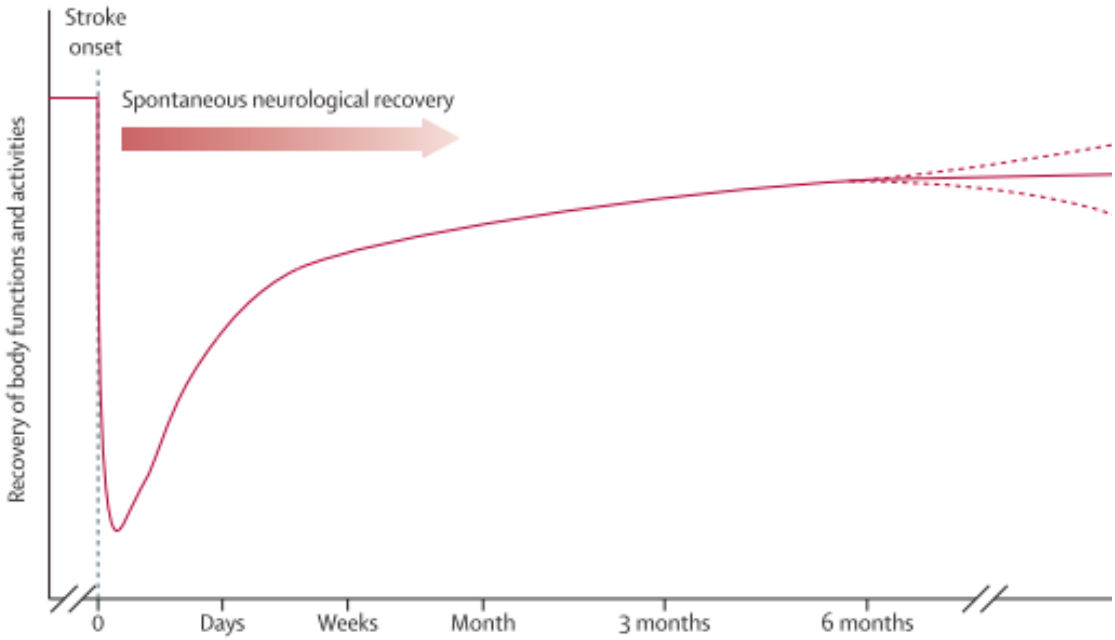

Figure 1. Stroke rehabilitation with time. [1] This figure illustrates the recovery of body functions and activities in terms of time elapsed from the stroke onset. The vertical dashed line at 0 stands for the stroke onset. The red curve represents the recovery extent. The x-axis shows the timeline associated with strode onset, with primary focus on 0 to 6 months. After 6 months, the recovery extent may vary according to the maintenance of physical conditions and monitoring quality of life, shown as dashed red curves.

or minutes level. On the contrary, homeostatic plasticity stabilizes and balances the changes during the Hebbian plasticity process at a slower pace (hours to days) to provide more robustness [16].

It was believed that a plateau of functional gains would be reached after the first 3- 6 months post-lesion, as shown in Figure 1, which means gaining functional recovery after this plateau is significantly difficult, if not impossible [1][12]. However, some studies argued that Hebbian plasticity could be induced any time after stroke with properly designed therapy [12]. The method proposed in this thesis research also falls into this category to provide continuous recovery after the plateau.

With the aim to induce Hebbian plasticity, various stroke rehabilitation interventions were designed for motor treatment. Task-specific and context-specific training are two common

standards in rehabilitation training, which indicates the training should suit the patients' needs in their preferred environment [1]. For arm recovery, constraint-induced movement therapy (CIMT) and robot-assisted training were found to be beneficial. For leg recovery, several methods were suggested, such as electromechanical-assisted gait training, task-oriented physical fitness training for walking, high-intensity therapy, and speed-dependent treadmill training [1].

However, methods such as CIMT require residual functionality of upper limbs, which limited its target population since around 20% to 30% of stroke patients cannot meet the standard of residual movement. Motor imagery, mirror therapy, and brain-computer interface provide an alternative solution for these patients [17].

2.2 EEG-based Brain-Computer Interface

Electroencephalography (EEG) is an electrophysiological signal measuring the electrical activity on the scalp, elicited by the neural activities inside the brain. Brain-computer interface (BCI) is a system extracting intentions and thoughts from the brain and translating them into command signals for external devices. EEG was widely applied as an input signal to BCI and has been applied in various areas, including communication, health care, entertainment, and marketing. For example, multiple industrial products have been available to provide mental health monitoring, sleep stage recognition, and focus level detection in the health care sector. Recent research further demonstrates the potential of BCI by decoding speech, handwriting, and movement intentions simultaneously from the human brain [18][19]. BCI applications usually require two phases: 1) an offline training process to calibrate the decoding models and algorithms and 2) an online phase to analyze real-time signals and translate them into output commands [20]. A typical BCI system consists of several components: signal acquisition, feature extraction, feature selection, and classification.

2.2.1 Signal Acquisition

Based on the types of signals BCI extracts, it can be roughly divided into invasive and non-invasive kinds [21]. Invasive BCI plants electrodes directly inside the brain, either on the surface of the cortex like Electrocorticography (ECoG) or even deeper inside the cortex in the form of intra-cortical electrodes, such as the Utah array [22][23]. Non-invasive methods include electroencephalography (EEG), magnetoencephalography (MEG), positron emission tomography (PET), functional magnetic resonance imaging (fMRI), and near-infrared spectroscopy (fNIRS). EEG and MEG measure surface electric and magnetic fields caused by neurons' electrical discharges (e.g., the firing of neurons) inside the brain, respectively. Medical imaging techniques such as PET, fNIRS, and fMRI measure brain activity through blood flow and oxygenation changes. Compared to non-invasive methods, invasive BCIs have a natural advantage of better signal bandwidth (i.e., information content) and quality (i.e., robustness against external artifacts). However, intrinsic limitations such as risky surgical procedures and high cost of maintenance limit its popularity in severe-conditioned patients.

Among non-invasive approaches, each solution has its pros and cons regarding temporal resolution, spatial resolution, and cost [21]. For example, fMRI and fNIRS give proper spatial resolution while poor temporal resolution. EEG and MEG both provide excellent temporal resolution and relatively good spatial information. However, MEG is not portable, much more complex in instrumentation, and significantly more expensive than EEG in operating and maintenance, making EEG a popular choice in the non-invasive BCI field, and this thesis study focuses exclusively on EEG.

2.2.2 Feature Extraction

There are mainly two types of features in EEG-based BCI: spatial feature and temporal feature. This kind of BCI usually uses multiple electrodes to capture scalp-level electrical activities caused by groups of cortical neurons firing together. Different BCI paradigms would elicit different patterns of EEG spatially. For example, motor imagery tasks would elicit oscillatory activities in the motor and premotor cortex, which can be measured as event-related desynchronization/synchronization (ERD/ERS) on frontal and central channels according to 10-20 international system [24][25]. Due to the homunculus structure of the primary motor and sensory cortex, such patterns usually have a contra-lateral activation, meaning the left arm movement leads to activation in the right hemisphere and vice versa [3]. In another scenario, watching flickering objects in a certain frequency would elicit steady-state visual evoked potential (SSVEP) in the occipital area, manifesting in occipital channels on the surface [26]. Hence, each BCI paradigm has its unique spatial features shown as different locations of elicited EEG. The electrical activities elicited by exogenous stimulus or endogenous intentions usually occur at a specific time to stimulus, called the temporal feature of EEG. For instance, an event-related potential (ERP) named P300 appears 300ms after the subject is exposed to a surprising stimulus [27]. Movement-related cortical potential (MRCP), another ERP, is also time-locked to the movement event or movement cue in movement execution or imagery tasks [28]. Therefore, both spatial and temporal features are important for EEG decoding. Spatial filters are used to enhance signal-to-noise ratios (SNR) of certain channels, which can be achieved by either data-driven or data-independent approaches [20]. Data independent spatial

filters include the Laplacian filter, which improves the central channel SNR by subtracting its surroundings [29]. Data-driven methods such as common spatial filter (CSP), xDAWN proved their efficacy in motor imagery and ERP paradigms [30][31].

Temporal features are usually demonstrated in two forms: time point feature and band power feature [20]. The time point feature is continuous EEG voltage values in certain periods. This feature usually requires bandpass or low pass filtering to get representable time points of a certain EEG pattern. ERP classification usually relies on this feature since they are mostly low-frequency signals with specific shapes or trends, time-locking to events. Band power feature stands for the energy of EEG in selected channels and time intervals. Classification information of EEG for various paradigms is contained in different bands. For instance, alpha band (8-12Hz) and beta band (13-32Hz) are specifically useful for motor imagery as the ERD and ERS are prominent in these frequency bands [24]. Delta band (0-4Hz) is also proved to be useful for deep sleep stage detection and movement intention detection [28][32]. Gamma band (above 33Hz) is responsible for the conscious state of mind and maybe a valid biomarker for Parkinson's disease [33]. In order to extract the band power feature, a bandpass filter is required to localize the interested bands, and then some statistics of power can be computed. Other features have also been investigated other than time point and band power feature. For example, connectivity features such as phase-locking value and coherence measure the correlation between channels and frequency bands [20]. The covariance matrix was also constructed as a feature to embed relevant classification information for certain algorithms like the Riemannian geometry method.

2.2.3 Feature Selection

EEG provides time-series features in multiple channels, which naturally lead to high-dimensional data. However, machine learning classifiers normally cannot perform well in high dimensional features, especially if the training samples are inadequate, called curse-of-dimensionality [20]. Feature selection is necessary to prevent such circumstances by selecting the most valuable features and discarding redundant features. Principle component analysis (PCA) is a dimensionality reduction method to project data to its several largest variance

dimensions, successfully applied in the BCI field [34]. Local preserving projection (LPP) is another feature selection method by keeping the data structure of neighboring trials from high dimensions and projecting it to lower dimensions [35]. Channel selection methods like canonical correlation analysis (CCA) and frequency band selection like maximal mutual information can also be classified to feature selection approach [36][37].

2.2.4 Classification Algorithms

In the early days of BCI development, classification algorithms were mainly derived from general machine learning methods such as linear classifiers, neural networks, nonlinear Bayesian classifiers, nearest-neighbor classifiers, and a combination of classifiers [20]. Among these algorithms, linear discriminant analysis (LDA) and support vector machine (SVM) were the most popular models thanks to their good generalization properties and low computational cost. Various general neural networks were also applied to decode EEG, including multilayer perceptron (MLP), Bayesian logistic regression neural network (BLRNN), adaptive logic network (ALN). However, these attempts were not successful enough to outperform simple linear classifiers at that time. The best algorithms were usually obtained by combining multiple models in ways of boosting, voting, or stacking, by which the individual classifiers complemented together to achieve better performance [38].

These algorithms in early days encountered the same challenges in the BCI field: 1) low SNR of EEG signals, 2) non-stationarity of EEG between users and sessions, and 3) limited training data. More recent algorithms aimed to overcome one or more of these challenges by various approaches such as adaptive classifiers, matrix and tensor classifiers, deep learning, and transfer learning [20].

Adaptive classifiers dynamically learn the parameters based on incoming data to overcome the non-stationarity of EEG. They proved their superiority in both offline and online studies. For example, adaptive SVM, adaptive LDA, and adaptive Bayesian classifiers gave a good performance for both asynchronous and passive P300 spellers in an offline study [39]. Online adaptive LDA and quadratic discriminant analysis (QDA) were applied successfully in motor imagery [40]. Adaptive algorithms take the learning process of humans in the loop to enhance

the overall system performance over time. However, the adaptation rate is hard to control as too slow or too fast changes in algorithms' responses might confuse humans and decrease classification results [20].

Matrix and tensor classifiers refer to algorithms directly learning from the high-dimensional EEG array, which is normally 3d matrix in the shape of trial, channel, and time. The Riemannian geometry approach calculates the covariance matrix of EEG and projects the matrix to a high-dimensional space called the Riemannian manifold. Each trial can be recognized as one point on the manifold, and they should cluster together if useful features are embedded inside the covariance matrix [41]. Thanks to the properties of Riemannian geometry, mean and distance could be calculated between points on the manifold, which could then form a minimum distance to mean (MDM) classifier. Otherwise, the high-dimensional points could be projected to a tangent space to utilize standard classifiers such as SVM or logistic regression. This family of methods has achieved state-of-the-art performances in multiple paradigms, including motor imagery and ERP. Another way to work with high-dimensional data is to redesign classical algorithms like LDA or SVM to generalize their decision rules for tensors, from which tensor support machine (TSM) and tensor fisher discriminant analysis (TFDA) were proposed [42][43]. This method usually requires fewer parameters to tune, like the Riemannian geometry approach is parameter-free and friendly to a smaller training dataset. However, they can also be restrained by expensive computational costs due to high dimensional matrix calculation.

Deep learning algorithms were applied successfully in the BCI field in various architectures. Convolutional neural network (CNN), deep belief network (DBN), and recurrent neural network (RNN) are the most popular ones, which constitute 43%, 18%, and 10% of all deep learning algorithms used in this area, respectively [44]. Recent deep learning algorithms were explicitly designed to focus on spatial and temporal features inside the data. For example, a CNN called EEGNET emulates the filter bank common spatial filter (FBCSP) to decode temporal features in the first convolutional layer and learns weights for channels in the second convolutional layer, which achieved outstanding performance in motor imagery and ERP

classification [45]. Deep neural networks can train on raw EEG data, which integrates feature extraction, selection, and classification to minimize the burdens of tunning parameters.

Transfer learning stands for applying the knowledge learned from a different but related area to a new field, which is especially useful in real-time BCI design. Since EEG is non-stationary over time and subjects, sometimes the algorithm trained on one subject or one run of the same subject cannot achieve similar performance in another subject or another run of the same subject. Several successful attempts were made to transfer session-to-session information in motor imagery BCI and P300 speller [49][50].

2.3 Movement-related BCI Features

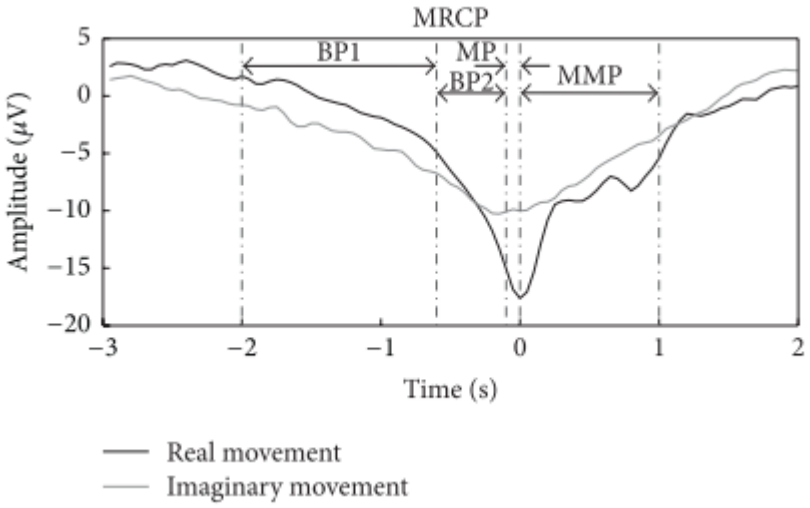

Figure 2. MRCP and various sub-components for real and imaginary movement. X-axis ranges from -3s to 2s with 0s as the movement onset. Y-axis ranges from -20μV to 5μV. Different parts of MRCP are illustrated in the figure as BP1, BP2, MP and MMP chronically. The grey line represents imaginary movement, and the black line stands for real movement. [2]

Aimed for stroke motor recovery, the extraction of movement intention would be the first step to designing a BCI system. MRCP and SMR are two major categories of EEG modalities that are believed to be reflective of cortical activations for movement preparation, planning, execution, and imagination [2][3]. MRCP is a low-frequency (0.05 - 5Hz) low-amplitude (5 – 30μV) EEG generated around 1.5s to 2s before movement onset and has the negative peak

aligned with the onset timing as shown in Figure 2. MRCP is a broad term for both Bereitschaftspotential (BP) and contingent negative variation (CNV). BP is generated by volitional movement, while CNV is elicited in response to external cues. MRCP consists of 4 sub-components: BP1, BP2, motor potential (MP), and movement-monitoring potential (MMP). BP1 is a slow negative cortical potential starting 1.5s to 2s before the movement onset. BP2 decreases sharply following BP1 from approximately 400ms – 500ms before movement and is considered the most prominent feature of MRCP. MP is the short period around movement onset, while MMP is typically a rebound after movement with around 1s length [2]. The generating sources of MRCP have been investigated extensively, and several implications were made. One study inferred the BP might originate from subcortical structures such as basal ganglia and thalamus since they are involved in the motor control loop inside the brain called basal ganglia-thalamocortical circuits [51]. Others suggested the early BP might be recorded by both sensorimotor areas, which was also proved by another work reporting that potentials were generated in line with BP1 from both the ipsilateral and contralateral supplementary motor areas (SMAs) [52][53].

SMR is a brain oscillation recorded during movement execution and imagination above the sensorimotor cortex, shown as power decrease (ERD) or increase (ERS) on alpha and beta band [3]. Spatial information is the most distinguishable feature when classifying different movements using SMR [54]. The physiological source of SMR was found to be coincident with bold-oxygen-level-dependent (BOLD) fMRI at the primary visual cortex and the thalamus during rest [3]. According to Homunculus, the active SMR is generated in a somatotopic way, which maps the movement of body parts with the corresponding brain regions.Both MRCP and SMR are applied to either detecting movement intentions from idle periods or classifying different movements. The experimental protocol to elicit these two modalities are not the same. MRCP is always elicited by ballistic movement, which is a sudden contraction of muscle or a one-time imagination of movement, while SMR is mainly induced by repetitively movement executions or imaginations. Furthermore, MRCP can predict movements due to its prominent BP component before onset, while ERD/ERS always manifest during or after movement onset. In conclusion, MRCP and SMR both contain critical classification information for movement

and can be used in multiple applications. A combination of these two sometimes could provide a performance boost in specific scenarios [4][55].

Chapter 3

Methodology

3.1 Expriment and preprocessing

This research collected EEG and electromyography (EMG) data from nine participants to investigate the brain response to four movements: WE_L, WE_R, IE_L, IE_R. The details of the experiment requirements and the preprocessing techniques are illustrated in the following parts.

3.1.1 Participants

This study recruited nine healthy right-handed participants aged between 18 and 45, including two females and seven males. No gender preference was considered. All participants reported that they did not have any known neuromuscular disease, brain-related conditions, cardiac conditions (e.g., pacemakers, arrhythmias, and cardiac conduction disturbances), peripheral neuropathy, history of seizure, or memory disorder. Waterloo ethics committee approved this study under application 43069. All participants were informed of the procedure and potential risk of the experiment and signed a consent form prior to participation of the experiment.

3.1.2 Instrumentation

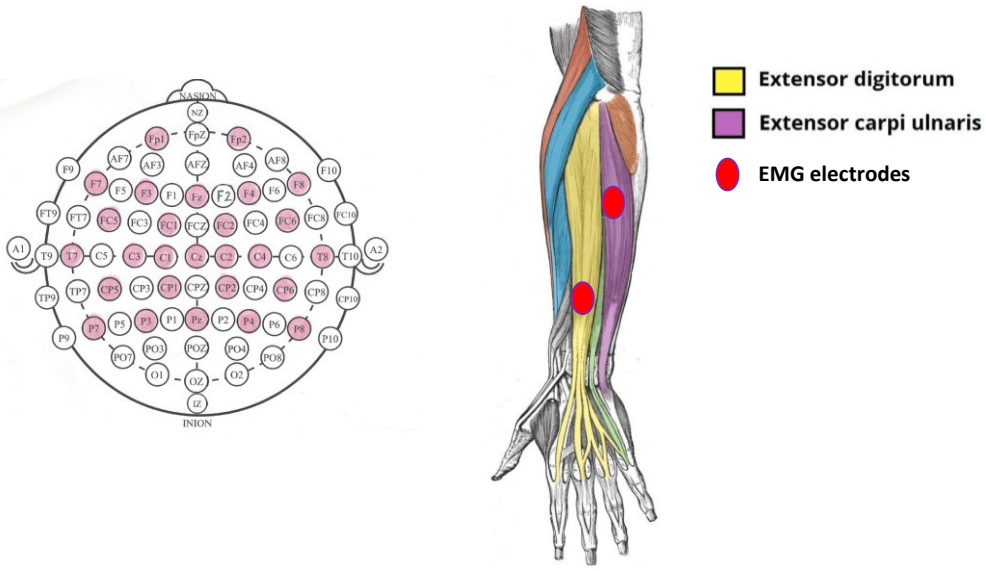

Figure 3. EEG and EMG electrodes position [56] The left graph shows an international 10-20 EEG electrode setup. The colored channels are used in this research. The right graph represents the different muscles of an arm in different colors. The yellow part is the extensor digitorum muscle and the purple part is the extensor carpi ulnaris muscle. The red circles indicate the EMG sensor position during the experiments.

EEG and EMG were recorded using Enobio 32 (NeuroElectric) system with common ground and reference on the earlobe. According to the international 10-20 system, 27 channels of EEG were recorded (Fp1, Fp2, F7, F3, Fz, F4, F8, FC5, FC1, FC2, FC6, T7, C3, C1, Cz, C2, C4, T8, CP5, CP1, CP2, CP6, P7, P3, Pz, P4, P8) with 500Hz sampling frequency as shown in left of Figure 3. Four EMG electrodes were placed on the extensor digitorum and extensor carpi ulnaris muscles (red locations on right of Figure 3) to measure the index finger extension and wrist extension activities, respectively.

3.1.3 Experiment Protocol

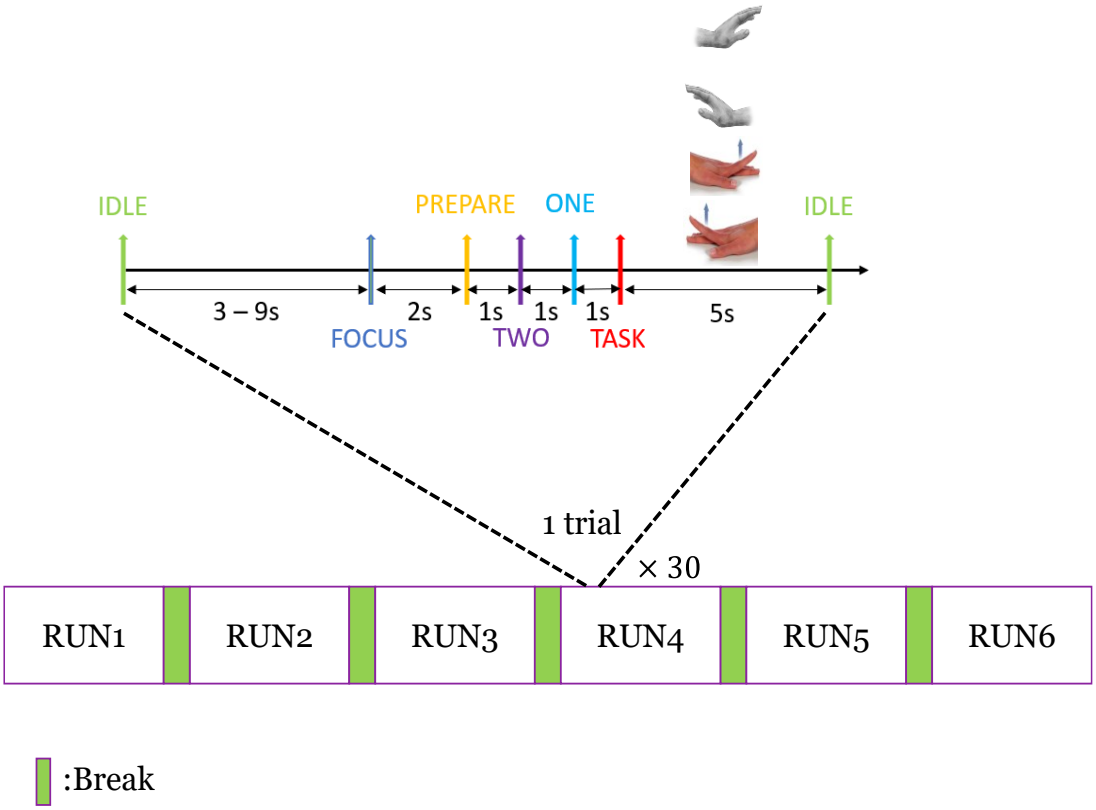

Figure 4. Experiment protocol

The participants were seated comfortably in a chair and placed both their forearms on a table in front of them. A computer monitor was in front of participants at around half a meter distance. They were asked to perform WE_L, WE_R, IE_L, or IE_R according to the instructions. The instruction includes a computer voice saying the name of the task and a corresponding task image. Inside each trial, six cues were displayed on the monitor below the task image in the following order: idle, focus, prepare, two, one, task. Participants can relax and blink their eyes during the idle period, but talking and movement were not allowed. Instead, participants were kept vigilant during preparation periods (focus, prepare, two, and one) with minimal eye movement. Once the task cue showed up on the screen, participants were required to wait for around 1s to perform the required task, as shown in Figure 3. This 1s waiting period was set to

ensure the participants' intention induced the movement rather than a subconscious response to the cue. The trial duration varied from 13 to 19 seconds by randomly assigning 3 to 9 seconds for the idle period. This was set to prevent participants from predicting the cue time. For each movement type, 45 trials were recorded in a random order to generate 180 trials in total. The trials were separated into six runs evenly, with 30 trials in each run. Every run took around 10 minutes, and short breaks were given after each run. In total, the whole session took approximately 1.5 hours.

3.1.4 Preprocessing

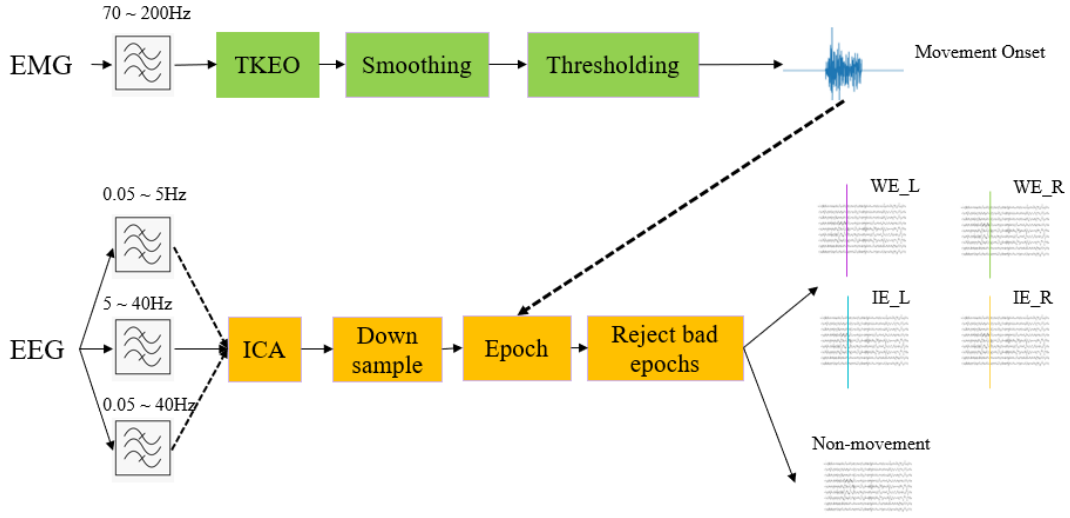

Figure 5. EEG and EMG preprocessing

This study preprocessed EEG and EMG to prepare the data for further feature extraction and classifier training. For EMG preprocessing, a 6th order Butterworth bandpass filter with a cutoff frequency of 70Hz and 200Hz was applied to 4 channels of EMG on six runs. After the filtering, most electrocardiogram (ECG) related artifacts were removed. Then, the Teager-Kaiser energy operator (TKEO) was applied to enhance the detection accuracy of EMG burst boundaries. Next, the enhanced EMG was smoothed by the moving average technique. Finally, a threshold was chosen for every EMG onset by an expert.

For EEG preprocessing, three frequency bands were extracted for further analysis: 0.05Hz to 5Hz as low band (i.e., MRCP band), 5Hz to 40Hz as high band, and 0.05Hz to 40Hz as full band. They were filtered by a 4^{th} order Butterworth bandpass filter. First, bad portions of data caused by head movements and unstable connections of electrodes were examined and removed based on the morphology of EEG by an expert. Next, independent component analysis (ICA) algorithm was used to compute sources of EEG, from which an expert removed independent components responsible for eye movements. Then the cleaned components were projected back to the sensor space to reconstruct the EEG signal. The eye-blink removed data was down sampled to 100Hz to save computational power for later processing. Then, movement epochs were extracted as -1.6s to 0.6s segments, and non-movement epochs were extracted as -8s to -2s and 2s to 8s segments with EMG labeled onsets as 0s. Longer non-movement period was chosen because the rest state is longer than the movement state. Finally, noise-contaminated epochs were identified and removed by an expert.

In order to investigate SMR, alpha and beta power were computed by Morlet wavelets with five cycles. ERD/ERS were computed on these two bands by subtracting the baseline power (-3s to -2s) with power from -2s to 4s and then divided by the baseline power.

3.2 Algorithms

3.2.1 EEGNET

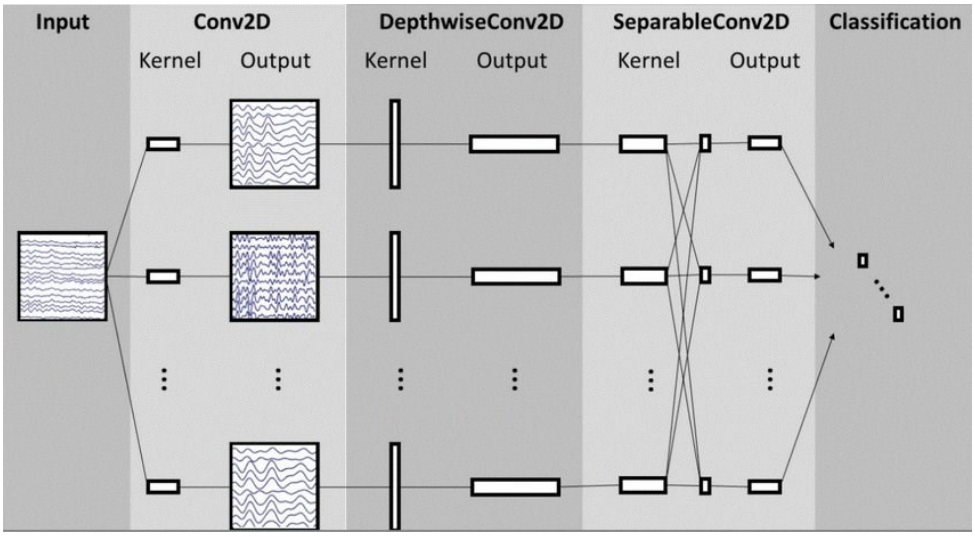

Figure 6. EEGNET architecture [45]

EEGNET is a compact deep learning architecture designed specifically for EEG signal analysis. It has been successfully applied in the BCI field in a number of studies [45]-[48]. Two layers of convolutions were designed to emulate temporal and spatial filters as FBCSP. In the first layer, the kernel size was set to be (1, 50) by following the suggestion in the original paper, in which the kernel was set to be half of the sampling frequency. F1, the number of convolutional kernels in this layer was set to be 8. In the second layer, a depth-wise convolutional layer with kernel size (n_ch, 1) decodes the spatial information from band-passed frequency features produced in the first layer. D, the number of spatial filters of each output in the first layer, was chosen to be 6. After obtaining the features containing both temporal and spatial information, a separable convolution layer consisting of depth and pointwise convolution layers was applied to integrate and optimize features learned in the previous layers. The number of pointwise convolutional filters F2 was selected to be 48 as a multiplication result of F1 and D to pass all learned features into the dense layer. Dropout regularization with a 0.5 drop rate was applied after spatial filtering and separable convolution. Finally, the SoftMax algorithm performed classification based on output features from the dense layer. All hyperparameters, including

F1, F2, D, and dropout rate, were selected by grid search, in which multiple combinations of values were tested.

EEGNET used Adam as the optimizer and categorical cross-entropy as the loss function. The model was trained for 300 iterations and was evaluated on the validation set. The best model with the highest accuracy was reserved for the testing set.

3.2.2 Support Vector Machine

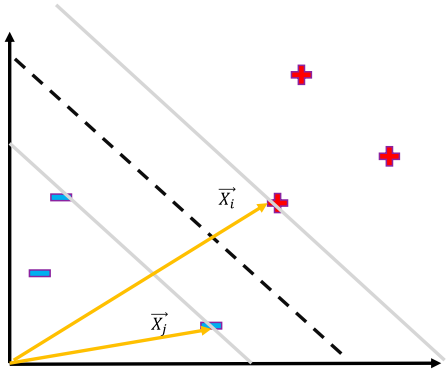

Figure 7. SVM illustration

Support vector machine (SVM) is a well-known machine learning algorithm, which has been applied in various fields. In Figure 7, red and blue samples represent positive and negative classes, respectively. The dashed black line is the decision boundary. Two gray parallel lines are margins beside the boundary. SVM finds the largest margin between two classes and solves an optimization problem based on two vectors $\overrightarrow{X_i}$ and $\overrightarrow{X_j}$, shown as yellow vectors in Figure 7. Quadratic programming could solve this problem as it is proved to be a convex function [57]. The solvers are beyond the scope of this thesis. Since the optimization just depends on samples, it gives great convenience to utilize different kernels to transform non-linearly separatable samples to other spaces, such as fisher kernel, polynomial kernel, radial basis function (RBF) kernel [58]. In this study, we used RBF kernel to get optimal performance.

3.2.3 Adaptive Boosted Logistic Regression

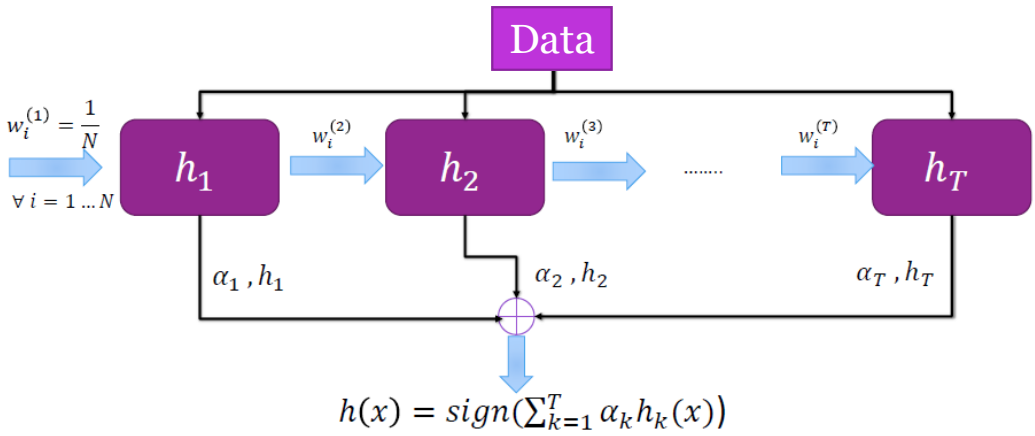

Figure 8. AdaBoost diagram [60]

Adaptive boosting (AdaBoost) is an ensemble learning method to sequentially train classifiers on continuously re-weighted datasets and usually could achieve better results than a single classifier. Weak learners (i.e., classifiers with slightly above chance level accuracy) such as logistic regression and decision tree are commonly used to construct the AdaBoost algorithm. As Figure 8 demonstrates, the EEG data firstly goes into first weak learner h_1 with a unity weight w_i on all trials. Once h_1 is trained, the trials are re-weighted such that misclassified trials gain more weights while correctly classified trials reduce their weights. In this way, the misclassified trial by h_1 will obtain more attention in the next classifier. In addition, h_1 is also given a weight (i.e., α_1)based on its performance, the decision function of $h_1$is saved for later combination. Then next classifier takes the outputs of the previous one and repeats the procedure. Finally, a certain number of weak learners are trained and are integrated by a summation function $h(x)$ to generate the final prediction, as shown in Figure 8. In this study, AdaBoost used logistic regression as the weak leaner and trained 10 of them to convert it into a strong learner. Logistic regression was chosen due to its fast-computing speed and robust performance in Riemannian geometry methods in EEG fields [61]

3.2.4 Riemannian Geometry

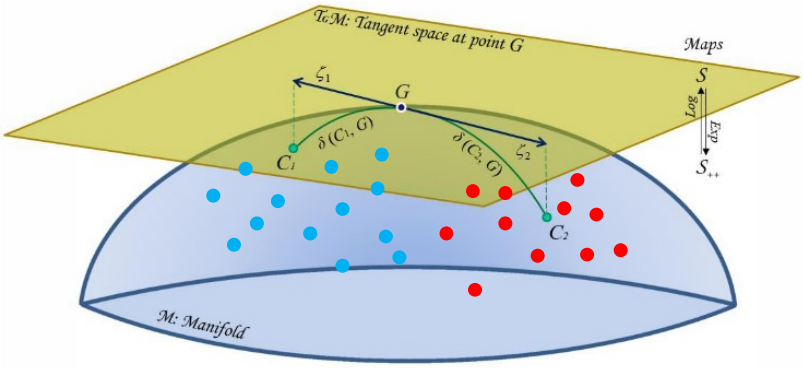

Figure 9. Riemannian manifold and tangent space [61]

On the Riemannian manifold, each point, illustrated as green and red points in Figure 9, is a symmetric positive definite (SPD) matrix. We need to transfer EEG data into this form to utilize the properties of Riemannian geometry. A unique covariance matrix was constructed for this purpose to capture both temporal and spatial information. Each class C_k for $k \in \{1, \dots, K\}$ contains all trials of one task, which is in size of $(n_{trial},\ n_{ch}, n_t)$. K is the class number; n_{trial} is the number of trials; n_{ch}stands for the number of channels; n_t is the number of time samples. As shown in Figure 9, blue and red points belong to C_1 and C_2 classes, respectively. We first calculated the grand averages over n_{trial} for each class $\overline{C_k}$ to get K matrices with a dimension of (n_{ch}, n_t). Then, for each trial X_i, which has the same dimension as $\overline{C_k}$, we constructed a super trial by:

$$X_i^{super} = \begin{pmatrix} \overline{C_1} \\ \cdot \\ \cdot \\ \cdot \\ \overline{C_K} \\ X_i \end{pmatrix} \tag{1}$$

X_i^{super} is the newly constructed super trial with the shape of $((K+1)\times n_{ch},\ n_t)$. The SPD matrix was built up in this form as a covariance matrix:

$$Cov_i = \frac{1}{n_t - 1} \times X_i^{super} \times \left(X_i^{super}\right)^T = \begin{pmatrix} \overline{C_1} * \overline{C_1}^T & \cdots & \overline{C_1} * \overline{C_K}^T & \overline{C_1} * X_i^T \\ \vdots & & \vdots & \vdots \\ \overline{C_K} * \overline{C_1}^T & \cdots & \overline{C_K} * \overline{C_K}^T & \overline{C_K} * X_i^T \\ X_i * \overline{C_1}^T & \cdots & X_i * \overline{C_K}^T & X_i * X_i^T \end{pmatrix} \quad (2)$$

Cov_i is the sample covariance matrix of the super trial in the shape of $((K+1)\times n_{ch})$ by $((K+1)\times n_{ch})$. Cov_i mainly contains three parts,

1. The covariance matrices among class grand averages, shown in red;
2. The covariance matrix of the current trial, shown in blue;
3. The covariance matrix between class grand averages and the current trial, shown in green.

The red part will always be the same for each trial since it does not include any current trial information, which is not helpful for classification. However, the blue and green parts hold the classification-related information. The spatial information of the current trial is encoded inside the sample covariance matrix (blue part) since its off-diagonal elements are the covariance between pairs of electrodes. It is a classical form of covariance matrix which has been used for motor imagery study [59]. Besides, the temporal information of the current trial is embedded inside the cross-covariances between the current trial and class grand averages (green parts) because the cross covariances could only be large if the grand average of one class matches the current trial. For example, if the morphology of X_i is similar to $\overline{C_1}$, then only $\overline{C_1} * Xi^T$ and $X_i * \overline{C_1}^T$ would be large.

Next, the covariance matrices of all trials were computed, representing points on the Riemannian manifold. Then, a tangent space was constructed at the geometric mean, the average of all trials, shown as point G in Figure 9. Finally, points on the Riemannian manifold were projected to the tangent space by logarithmic mapping. Since the tangent space is a Euclidean space, we could use standard classifiers for further processing. This study

constructed two classifiers based on Riemannian features: Riemannian-based support vector machine (R_SVM) and Riemannian-based AdaBoost logistic regression (R_Ada_Lgr).

3.2.5 Detection Ensemble Learning

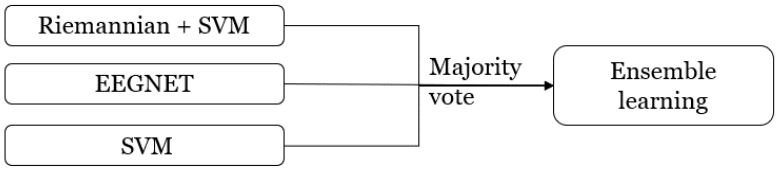

Figure 10. Ensemble learning structure

This study proposed an ensemble learning method for detecting movement intention, combining R_SVM, EEGNET, and SVM by majority voting. First, these three classifiers were trained individually on low, high and full bands EEG. Then, their results were combined by a simple hard majority voting rule such that the result was dependent on two or more identical outputs of three classifiers.

3.3 Analyses

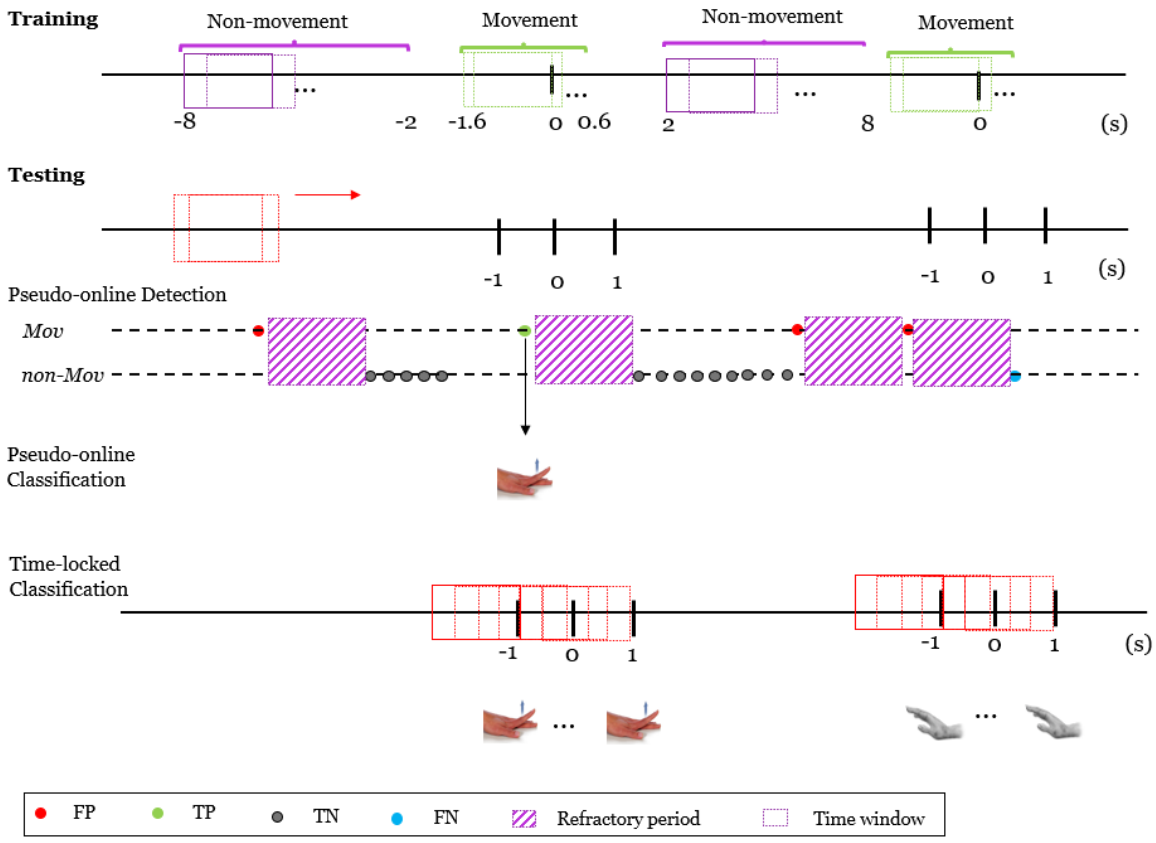

Figure 11. Training and testing procedures for pseudo-online and time-locked analysis. A timeline including two random trials are illustrated to demonstrate the training and testing details of both pseudo-online and time-locked analysis. The top timeline shows the training window extraction for both analyses. The movement and non-movement windows are shown in green and purple rectangles, respectively. The movement windows are extracted between -1.6s and 0.6s with a stride of 0.01s and the non-movement windows are extracted from -8s to -2s and 2s to 8s with a stride of 0.05s, with respective to movement onset as 0s. In the testing timeline, a window (red rectangle) moves at a constant rate (0.1s) to scan the testing run, which could be predicted as either Mov or non-Mov during pseudo-online detection. Detection period is between -1s to 1s of each trial. FP, TP, TN and FN of detection are shown in red, green, gray and blue, respectively. A refractory period is shown in purple rectangle with italic lines. In pseudo-online classification, only the TP of detection would be classified to one movement type (IE_L for the first trial). In time-locked classification, movement types would be predicted for all windows (red rectangles) extracted by a stride of 0.1s within the detection period.

EEG could be analyzed offline, pseudo-online, or online based on the different data collection approaches and experiment protocols. The offline analysis is done after the data collection based on trials, which is the most prevalent method of EEG analysis in literature due to its simplicity. In contrast, online analysis processes data in real-time and sometimes provide feedback to make a closed-loop BCI. Obviously, online analysis is more preferable for most practical situations. When data are acquired continuously throughout the experimental with data between the trials also recorded, Pseudo-online analysis can be used, as an approximation of online analysis. It is a useful technique during the algorithm development, before applying newly development algorithms in online analysis.

3.3.1 Pseudo-online Analysis

In an online BCI scenario, the system analyzes input data and constantly makes predictions based on the real-time incoming EEG stream. We implemented the sliding time window on preprocessed epochs to simulate the online scenario after the data collection to analyze classifier performance on these windows pseudo-online. In this analysis, detection classifiers first predicted time windows as movement or non-movement. Next, the classification algorithm gave predictions on movement windows to distinguish movement types.

The pseudo-online analysis adopted 6-fold cross-validation to evaluate classifiers based on runs. Training, validation, and testing sets were separated in terms of classifiers' characteristics. For EEGNET, one run was pre-selected for testing; one run was randomly chosen from the left runs for validation; four runs were reserved for training. However, for SVM, R_SVM, and R_Ada_Lgr, five runs were used for training, one run was used for testing. No validation set was assigned for them because no hyperparameter tuning was needed.

3.3.1.1 Detection

Detection is essentially a binary classification between non-movement and movement. The ensemble learning method, consisting of SVM, R_SVM, and EEGNET, conducted the detection task using low-band, high-band, and full-band EEG. Firstly, the individual classifiers were trained using all bands. The training, validation, and testing set of one algorithm used the

same frequency band. Next, each classifier selected its best-performed band. Finally, three classifiers constitute the ensemble learning method.

- Training

As shown in Figure 11, a 2s time window scanned movement epochs (-1.6s to 0.6s) and non-movement epochs (-8s to -2s and 2s to 8s) with a 0.01s and 0.05s stride, respectively. Shorter stride on movement epochs obtained a more detailed description of temporal information, while a longer stride on non-movement epochs saved computational power since there should not be any time-sensitive information during the rest state. As a result, one trial generated 21movement windows and 162 non-movement windows. In order to balance two classes, 21 windows were randomly selected from 162 non-movement windows for every trial. SVM, R_SVM, and EEGNET were trained using these windows in their own training sets.

- Testing

After training, EEGNET obtained the best model by tuning parameters in the validation set while SVM and R_SVM directly got the optimal parameters in the training set. A 2s time window scanned the testing run with 0.1s stride to obtain testing windows, as shown in Figure 11. For each time window, EEGNET, SVM, and R_SVM first predicted whether it was movement or non-movement, then their predictions were majority voted to give a final label. If a movement window was found, all classifiers would take a break for 2s without giving new predictions. This 2s break is the refractory period meaning no second movement intention could be detected inside this duration because MRCP length, which is 2s, restrained movement resolution. The overlapping of MRCPs for successive intentions is beyond the scope of the current study.

Unanimous voting of multiple consecutive windows was also investigated, which means a detection could only be found if several successive windows reported movement consistently. This measure would restrict the non-movement window being mispredicted as a movement but would likely miss some of the movement windows on the other hand.

- Evaluation

$$TPR = \frac{TP}{TP + FN} \tag{3}$$

$$FPs\ per\ \text{min} = \frac{FP}{non - mov\ duration(\text{min})} \tag{4}$$

$$F1 = \frac{TP}{TP + 0.5 * (FP + FN)} \tag{5}$$

$$DL = t_{detected\ window} - t_{movement\ onset} \tag{6}$$

In order to measure the performance of algorithms, four basic metrics were introduced: true positive number (TP), false positive number (FP), true negative number (TN), and false negative number (FN). A true positive happens if detection was within ±1s of movement onset, shown as the green point in Figure 11. In contrast, any detection beyond this range was a false positive, shown as the red points in Figure 11. False negatives occurred if any window in this range was predicted as non-movement, shown as the blue point in Figure 11. True negatives are correctly predicted non-movement windows, which are grey points in Figure 11.Based on these four metrics, true positive rate (TPR), false positives per minute (FPs/min), and F1 scores were computed to evaluate the system according to equation (3)(4)(5), respectively. TPR was the rate of correctly detected trials; FPs/min was the number of false positives that happened per minute in the rest state. TPR and FPs/min were drawn in the receiver operator characteristic (ROC) curve to demonstrate the performance change by varying unanimous voting window numbers. F1 score was used to quantify the behavior of the classifier in the unbalanced scenario in pseudo-online testing, for it has more false samples (non-movement window) than true samples (movement window).

Detection latency (DL) is the time difference between the end of the true positive window and its corresponding movement onset according to equation (6), representing the responsive speed of the detector.

3.3.1.2 Classification

Once a movement window was detected, the classification algorithm (R_Ada_Lgr) predicted its movement type. In order to investigate the influence of lateralization on classification, we

performed two binary classifications: left vs. right and WE vs. IE. Finally, we performed a 4-class classification (WE_L vs. WE_R vs. IE_L vs. IE_R). In addition, three frequency bands were investigated to determine the optimal frequency band in each classification.

- Training

R_Ada_Lgr was trained on movement windows, extracted from -1.6s to 0.6s of each trial using a 2s time window scanning at a stride of 0.01s. For the 4-class classification, four types of movement windows constitute four classes. For binary classifications, the windows were combined based on their labels. For example, for WE-vs.-IE classification, WE_L and WE_R windows were combined as the WE class, IE_L and IE_R windows were combined as the IE class. A similar procedure was applied for left-vs.-right classification as well.

- Testing

Since the classification followed the detection, the trained R_Ada_Lgr only gave predictions to previously detected windows in the testing run.

- Evaluation

The classification performance was measured only on true positive detection windows, false positives were not considered because assigning movement type to non-movement windows is unnecessary. As a result, one trial obtained one classification label, so the classification was conducted on the trial level. Accuracy and confusion matrices were used to present the classification results. The accuracy was chosen due to the trial numbers in each classification condition being nearly balanced. In addition, the confusion matrix was selected to show the classification details. The diagonal elements of a confusion matrix gave the number of correctly classified trials, while the off-diagonal elements represented misclassified trials. In addition, the confusion matrix was normalized to get the percentage of predicted categories in true classes by computing the division between each element and the sum of its corresponding row.

3.3.2 Time-locked Classification Analysis

Unlike the pseudo-online classification, the time-locked classification did not depend on detected windows. Instead, it classified all windows in a short period (-1s to 1s of movement onset) to investigate the influence of time on classification performance. R_Ada_Lgr and sLDA classified low-band EEG to provide a comparison between the proposed algorithm and the conventional algorithm. Two types of classifications were carried out in this analysis: 1) 4-class classification (WE_L vs. WE_R vs. IE_L vs. IE_R), 2) pairwise classification (WE_L vs. WE_R, WE_L vs.IE_L, WE_L vs. IE_R, WE_R vs. IE_R, WE_R vs. IE_L, IE_L vs. IE_R).

- Training

The training process was precisely the same as pseudo-online classification. We extracted 21 windows from movement epochs (-1.6s to 0.6s of movement onset) from the training runs using a 2s time window moving at a stride of 0.01s. R_Ada_Lgr and sLDA were trained for different classifications by combining windows accordingly.

- Testing

A 2s time window scanned every trial from -1s to 1s in the testing run with a stride of 0.1s, as Figure 11 shows. R_Ada_Lgr and sLDA predicted all extracted testing windows for both pairwise and 4-class classifications.

- Evaluation

Accuracy was computed for each testing window in a trial-based manner, meaning the proportion of correctly classified trials was the accuracy for that window. As a result, accuracies of all windows between -1s and 1s were obtained using R_Ada_Lgr and sLDA.

Chapter 4

Results

4.1 MRCP

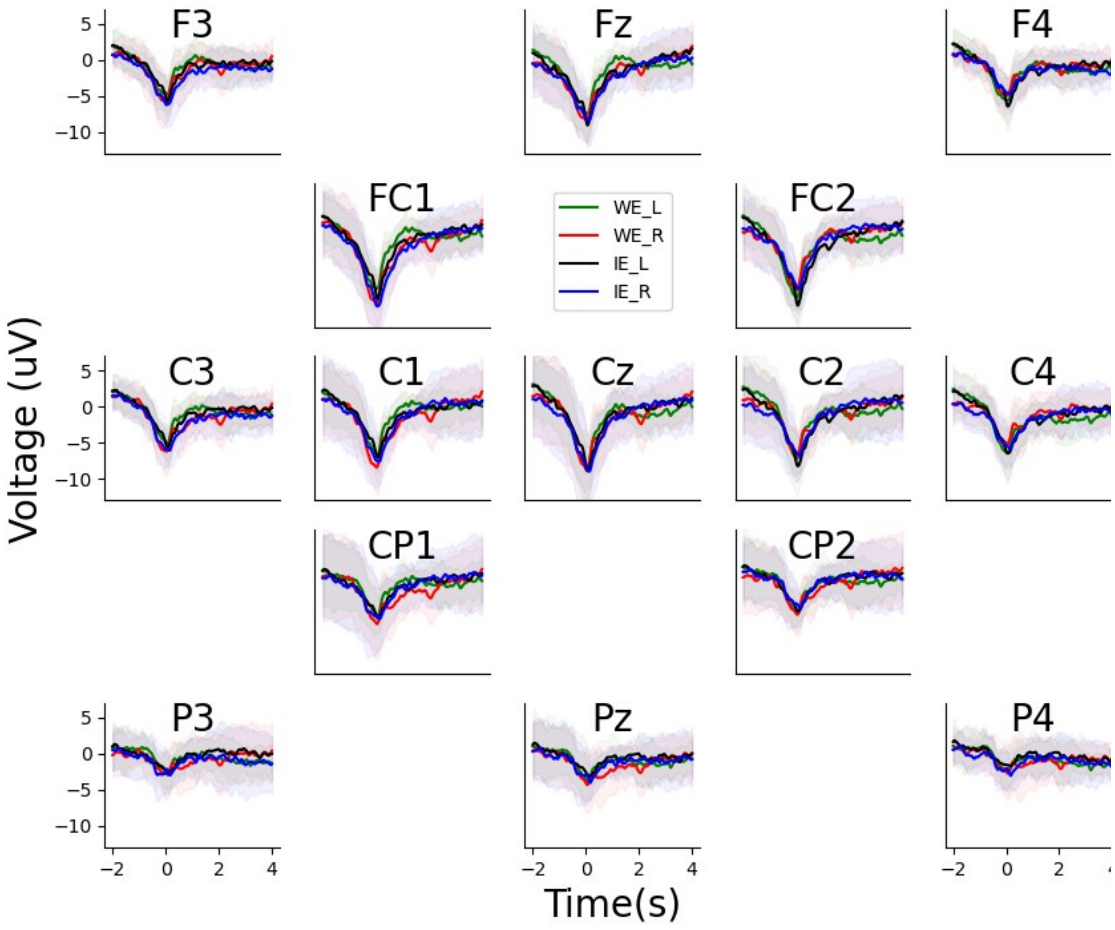

Figure 12. MRCP of four movements. MRCP of WE_L, WE_R, IE_L and IE_R are shown in green, red, black and blue curves, respectively, in 15 channels covering frontal, central and parietal lobes. The shaded areas are the standard deviations of MRCP among nine subjects. The x-axis is from -2s to 4s with respective to movement onset as 0s. The y-axis ranges from -13μV to 7μV.

The MRCPs shown in Figure 12 were calculated by averaging nine subjects' low-band EEG (0.05 - 5Hz). MRCPs decreased from -2s to 0s and rebounded from 0s to around 1s in all

channels with varied peak-to-peak amplitudes. From direct observation, the peak-to-peak amplitude of MRCPs decrease from central channels to surrounding channels. FC1, FC2, and Cz have the largest MRCPs peak-to-peak amplitude at approximately 12μV. Conversely, P3, Pz, and P4 have the smallest MRCPs with approximately 4μV peak-to-peak amplitude. The varied MRCP amplitude on these channels suggests that the upper-limb MRCPs were generated below FC1, FC2, and C2 with a diffuse distribution over the premotor and motor cortex. The morphology characteristics of MRCP observed in the current study are in agreement with other studies of upper-limb MRCP in the literature [5][62][77].

The variability of MRCPs might relate to time and channel locations. For example, in Fz, FC1, FC2, C1, Cz, and C2, the shaded areas became narrower when approaching the movement onsets. However, this trend is unclear in other channels, presenting as approximately equal variability from -2s to 4s.

The MRCP differences among the four movements are generally subtle across all channels. However, slight contra-lateralization could be noticed in two pairs of channels (FC1 vs. FC2 and C1 vs. C2). Close observation shows that the peak-to-peak amplitudes of green curves (WE_L) are smaller than the red curves (WE_R) in FC1 and C1 and vice versa in FC2 and C2. Similar patterns are found for IE_L and IE_R as well. Nevertheless, no MRCP difference between two movements on the same limb could be observed.

4.2 ERD/ERS

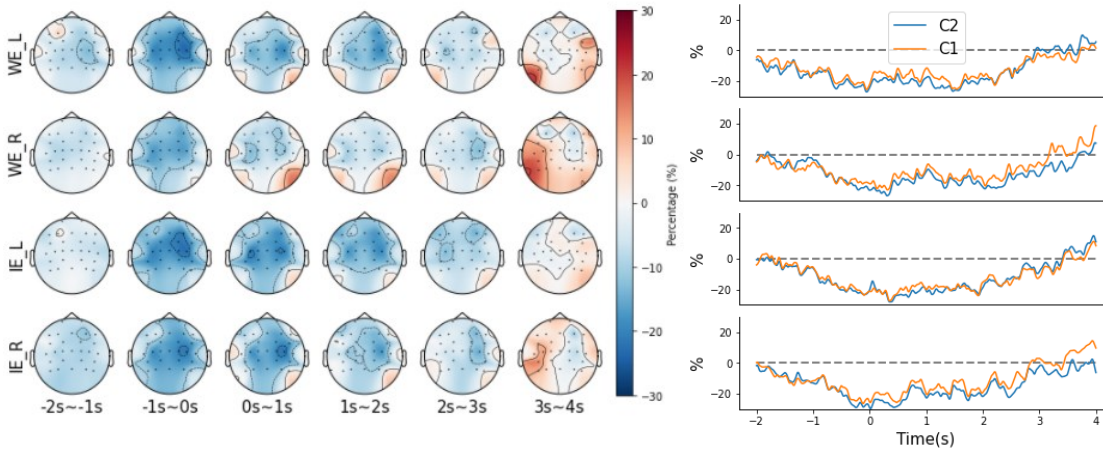

Figure 13. ERD/ERS of high-band EEG for four movements. The left topographical maps show the average ERD/ERS of four movements in six 1s intervals from -2s to 4s with baseline from -3s to -2s. The percentage change is indicated by color variation from red (30%) to blue (-30%). The right figures present the continuous power changes of four movements on C1 and C2 in orange and blue, respectively, from -2s to 4s with the range from -30% to 30%. The dashed gray lines indicate the baseline power. ERD occurs from around -2s to 3s when the power changes are below the baseline, while ERS occurs from approximately 3s to 4s when the power changes are above the baseline.

In order to calculate ERD/ERS, the baseline power was first computed by averaging the power between -3s and -2s. Then the percentage change was calculated by the following power from -2s to 4s subtracting and dividing by the baseline power. If the following power is lower than baseline power, ERD occurs. Conversely, ERS occurs. By observing Figure 13, the topographical maps shows that the power decrease is manifested as ERD dropping from -2s and hitting the bottom between -1s and 0s. Then, the power increases and surpasses the baseline at around 3s, showing as ERS. A similar trend shows up on the continuous power change for C1 and C2 on the right-side figure of Figure 13. In addition, ERD topographical maps show contra-lateralization on WE_L and IE_L between -1s and 0s, and the activation region is around FC2, C2, and C4. However, this pattern is not apparent in the right-side figure, which suggests the contra-lateralization of ERD is too subtle to be captured in continuous power

change. Although the ERD of left movements indicates contra-lateralization in topographical maps, ERD of the right movements does not give clear activation regions.

Unlike ERD, ERS gives pronounced contra-lateralization on both left and right movements. In topographical maps from 3s to 4s, ERS appears in the left hemisphere for the right-hand movements and the right hemisphere for the left-hand movements. The distribution of ERS is more widespread and not as localized as ERD. The contra-lateralization of ERS is also evident in the right-side figure from 3s to 4s, which might infer that the ERS could be a better indicator to differentiate the sidedness of movements.

Apart from contra-lateralization, ERD and ERS of WE and IE on the same limb also manifest subtle differences. For example, from the topographical maps, the ERD of IE_R is slightly lower than WE_R on C2 between -1s and 0s. Conversely, the ERS of IE_R is slightly larger than WE_R in the left hemisphere from 3s to 4s.

Overall, the ERD/ERS might be a good indicator for classification between two sides of movements. Nevertheless, it does not give an indicator as time-locked as MRCP for detection.

4.3 Pseudo-online Detection

4.3.1 Relationship between Detection Algorithms and Frequency Bands

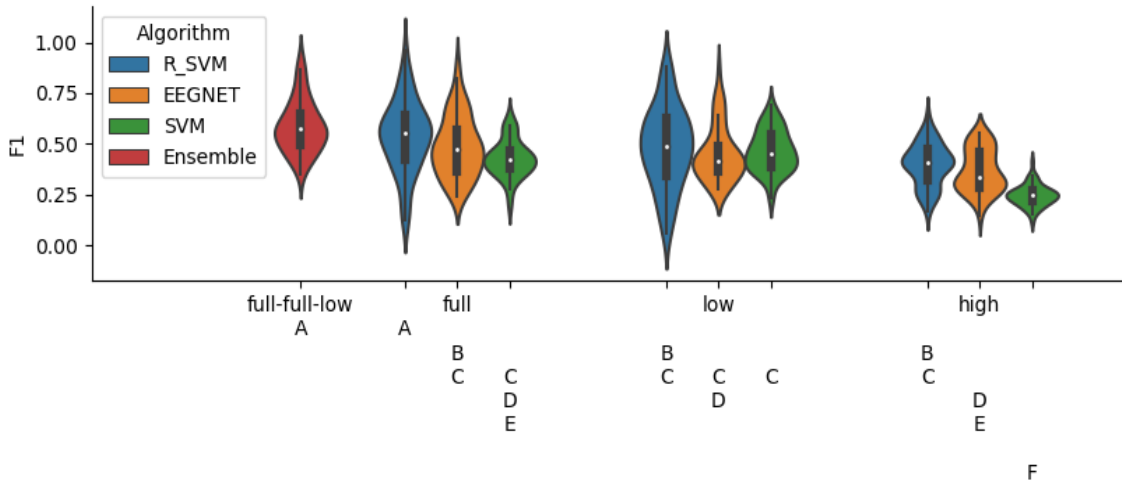

Figure 14. Violin plots of F1 score for four algorithms in three frequency bands. The violin plots of four detection algorithms, including R_SVM, EEGNET, SVM, Ensemble, are shown in blue, yellow, green, and red, respectively. Four groups of frequency bands (full-full-low, full, low and high) are shown in x-axis from left to right. Ensemble method consists of full-band trained R_SVM, full-band trained EEGNET and low-band trained SVM. Hence, its frequency band is shown as full-full-low. The other three algorithms were trained with full, low and high bands. Inside each violin plot, the white point stands for the median value, the black bar is the interquartile range, and the curve outside represents the kernel density of the data distribution. The statistical groupings calculated by two-way ANOVA with alpha as 0.05 are presented as letters from A to F below the plots.

As mentioned in Section 3.3.1.1, the F1 score gives a balanced measurement for different combinations between detection algorithms and frequency bands. We conducted a statistical analysis using two-way mixed model analysis of variance (ANOVA) with subject as a random factor, algorithm and band as fixed factors. According to equation (5), the F1 score was calculated for every fold. Each subject got six F1 scores due to 6-fold cross-validation. As a result, fifty-four F1 scores were obtained from nine subjects. Each violin plot in Figure 14 represents the data distribution of 54 F1 scores for each combination.

We can see from direct observation that classifiers trained on the full and the low band generally have better F1 scores than classifiers trained on the high band. The statistical

grouping further proves this phenomenon. For example, the SVMs trained by the full and low bands are significantly better than SVMs trained using the high band since the high-band SVM is in group F while the other two share group C, D, and E. A slowly decreasing performance shows up on EEGNET by observing the change of statistical groupings from full band to high band. (Group B, C for the full band; group C, D for the low band and group D, E for the high band)

Slight differences also exist between full and low bands for all algorithms. For example, R_SVM performs significantly better in the full band, and EEGNET performs slightly better in the full band than the low band by observing their statistical groupings. However, SVM trained by the full band is a bit worse than that trained by the low band because its F1 scores in the full band cover broader statistical groupings, which are group D and E, than its correspondent.

Three individual classifiers trained on their preferred bands were combined to establish the ensemble method. According to the statistical groupings in Figure 14, the ensemble method significantly outperforms nearly all individual classifiers except the R_SVM trained on the full band. Nevertheless, the F1 score variability of the ensemble method is smaller than full-band trained R_SVM, presenting the superiority of the ensemble method in another perspective.

4.3.2 Impact of Window Number on Detection Performance

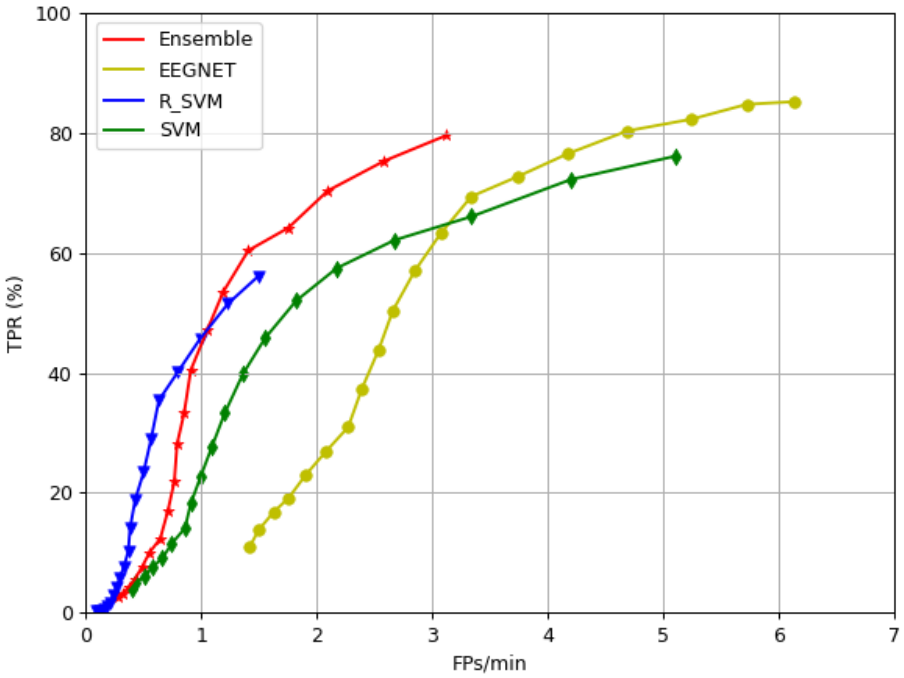

Figure 15. Modified ROC curves of four algorithms. The modified ROC curves of full-band trained EEGNET, full-band trained R_SVM, low-band trained SVM and full-full-low trained ensemble method are shown as yellow, blue, green, and red, respectively. On each curve, the TPR and FPs/min are reported for all points from upper-right to lower-left, representing the number of windows needed for unanimous voting from 1 to 19.

As introduced in 3.3.1.1, the detection performance of four algorithms trained on their optimal bands was investigated by varying the number of windows needed for unanimous voting from 1 to 19. TPR and FPs/min were calculated according to equations (12) and (13) for each window number, respectively. Modified ROC curves were obtained by connecting 19 TPR-FPs/min points. From Figure 15, TPR and FPs/min drop down from the right-upper side to the left-lower side in each ROC curve with the window number increasing. As the window increases, it is less likely to get FP and TP since more windows contribute to the unanimous voting. By observing the characteristics of the four curves, we can find that the algorithms act in different ways. For example, EEGNET gives the best TPR with the worst performance on FPs/min, while R_SVM shows a contrary result with the worst TPR and best FP/min. SVM

has a middle performance between EEGNET and R_SVM. The ensemble method outperforms the others as its ROC locates on the left-upper side compared with other methods, which means

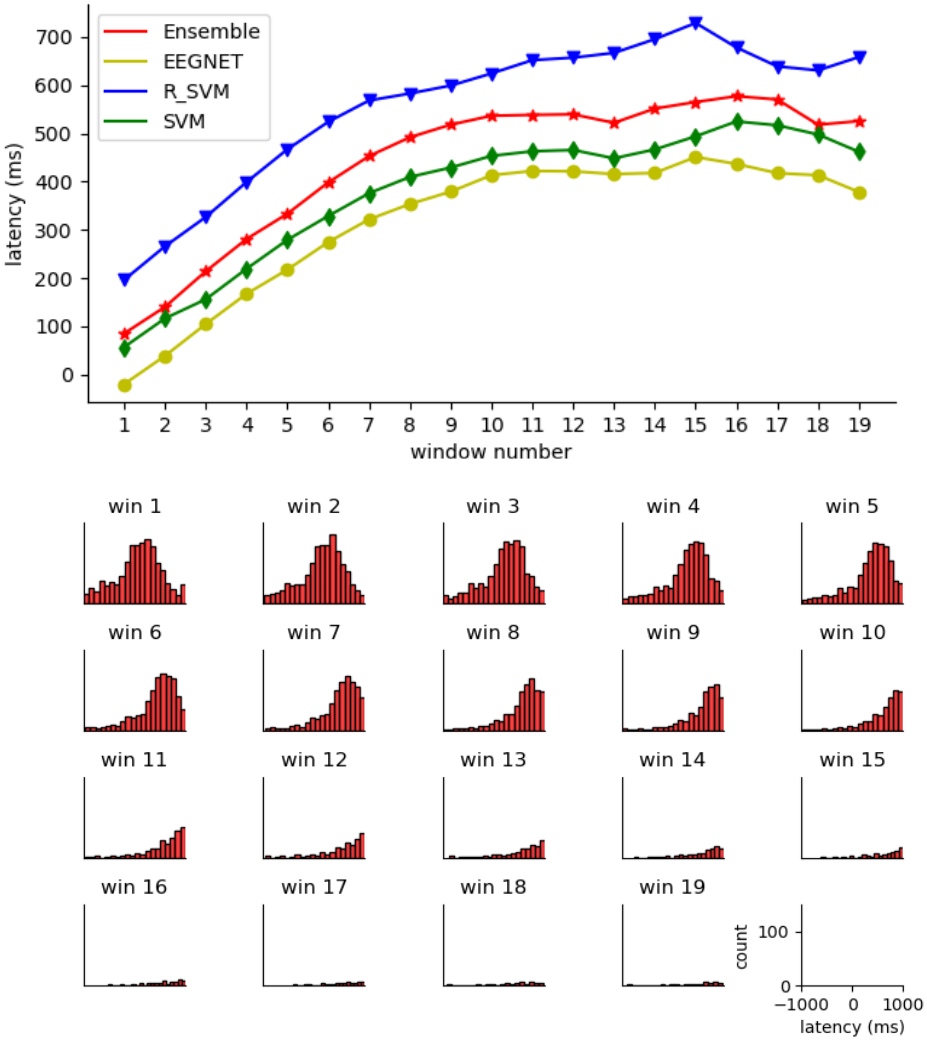

Figure 16. Detection latency variation with window number change. In the upper figure, average detection latencies of all TPs in nine subjects are shown for all algorithms by varying the window number from 1 to 19. The ensemble method, EEGNET, R_SVM and SVM are plotted in red, yellow, blue, and green, respectively. The latency is given in millisecond unit ranging from -50ms to 750ms. In the lower figure, the ensemble method detection latency are plotted as histograms from window number 1 to 19. The units and labels of the histogram show in the lower right plot, x-axis stands for the latency time ranging from -1000ms to 1000ms and y-axis is the trial count number ranging from 0 to 150. In each histogram, 20 bars are plotted with 100ms interval from -1000ms to 1000ms.

better TPR and FPs/min.

The detection latencies of correctly detected trials are calculated by equation (6). The upper graph of Figure 16 shows the averaged latencies among all trials for each window amount. In this graph, the latency increases steadily from window number 1 to around window number 8 for all algorithms. Then it continues climbing at a slower rate until around window number 13. Finally, it fluctuates to the end. By comparing four algorithms, the average latency from the shortest to the longest appears in the following order for all windows: EEGNET < SVM < Ensemble < R_SVM. The latency of the ensemble method is approximately the average of the others.

The lower graph of Figure 16 illustrates the detection latency of all trials using the ensemble method in histograms for every window number. The ensemble method is selected to represent the general trending of latency distribution since the histograms of all methods look similar. The histograms give another perspective on latency based on individual trials compared to the averaged latency of all trials in the upper graph. Observing the histogram when the window number is one shows that the distribution is approximately Gaussian. With the window number increasing from 1 to 8, the mean of the histogram keeps shifting to the right, leading to postponed detection time. Besides, the total TP count decreases slowly, consistent with the TPR drop shown in Figure 15. With the window number increasing from 9 to 19, the distribution keeps right-skewed, and the total TP count decreases dramatically. As a result, more randomness was added to the latency, leading to a slower rate of increase and fluctuation.

In conclusion, more unanimous voting windows lead to better FPs/min but worse TPR and detection latency. As a result, the system with one window was adopted, giving high TPR, short latency, and high FPs/min by satisfying two metrics out of three. More detailed results of this choice of are presented next.

4.3.3 Detection Performance with One Window

Subject	Ensemble			EEGNET			R_SVM			SVM		
	TPR (%)	FPs/ min	Latency (ms)	TPR (%)	FPs/ min	Latency (ms)	TPR (%)	FPs/ min	Latency (ms)	TPR (%)	FPs/ min	Latency (ms)
1	85.5	3.3	37.0	90.2	6.3	-95.0	62.4	1.5	154.7	87.3	4.4	28.6
2	84.8	3.4	60.0	84.8	7.0	-44.0	45.7	1.2	52.5	81.9	4.8	75.3
3	75.2	4.4	-16.0	75.8	7.4	-105.0	49.7	1.6	138.5	77.1	5.6	-25.9
4	85.0	3.0	32.0	90.3	7.8	-120.0	79.2	3.0	72.4	72.0	5.2	63.7
5	96.0	1.8	47.0	95.4	3.2	-23.0	80.6	1.0	264.9	91.4	3.9	34.0
6	77.3	2.8	136.0	92.0	4.9	-37.0	56.4	1.7	231.3	69.9	6.4	78.4
7	68.2	4.2	11.0	74.8	9.8	-35.0	19.2	1.4	254.7	66.9	6.2	-23.9
8	73.4	0.9	345.0	83.8	2.1	309.0	66.2	0.9	391.9	72.7	3.1	287.0
9	71.2	4.3	171.0	80.1	6.8	26.8	45.5	1.2	248.6	66.0	6.3	6.4
Mean ± SD	79.6 ± 8.8	3.1 ± 1.2	91.4 ± 111.9	85.2 ± 6.9	6.1 ± 2.2	-13.8 ± 129.4	56.1 ± 17.9	1.5 ± 0.6	201.1 ± 106.8	76.1 ± 8.5	5.1 ± 1.1	58.2 ± 94.2

Table 1. Detection performance of nine subjects using Ensemble method, EEGNT, R_SVM and SVM. TPR, FPs/min and latency are listed for every subject and every algorithm. The last row shows the mean and standard deviation of nine subjects for each metric.

As discussed in the previous section, one window was adopted for the system. The detection performance with one window was investigated in detail for subjects and algorithms in Table 1. The results in this table correspond to the one-window condition in Figure 15 and Figure 16. The mean TPRs and FPs/min in the last row of Table 1 are equivalent to the right-most points of the ROC curves in Figure 15, and the mean latencies are equivalent to the left-most points in Figure 16. Furthermore, this table reveals the standard deviation in nine subjects. Two-way mixed-factor ANOVA tested the statistical significance of four algorithms in each metric by setting subjects as random factor and algorithms as fixed factor with the alpha of 0.05.

Subjects play an essential role in validating the system's adaptation ability. Regarding TPR, subject 5 gave the highest TPR in all algorithms, while subject 7 has the worst TPR in three classifiers. However, TPR has no significant difference among subjects. FPs/min does not show significant difference among subjects either, although subject 8 has the fewest FPs/min in all algorithms. However, latency shows a significant difference in subject 8 with a p-value

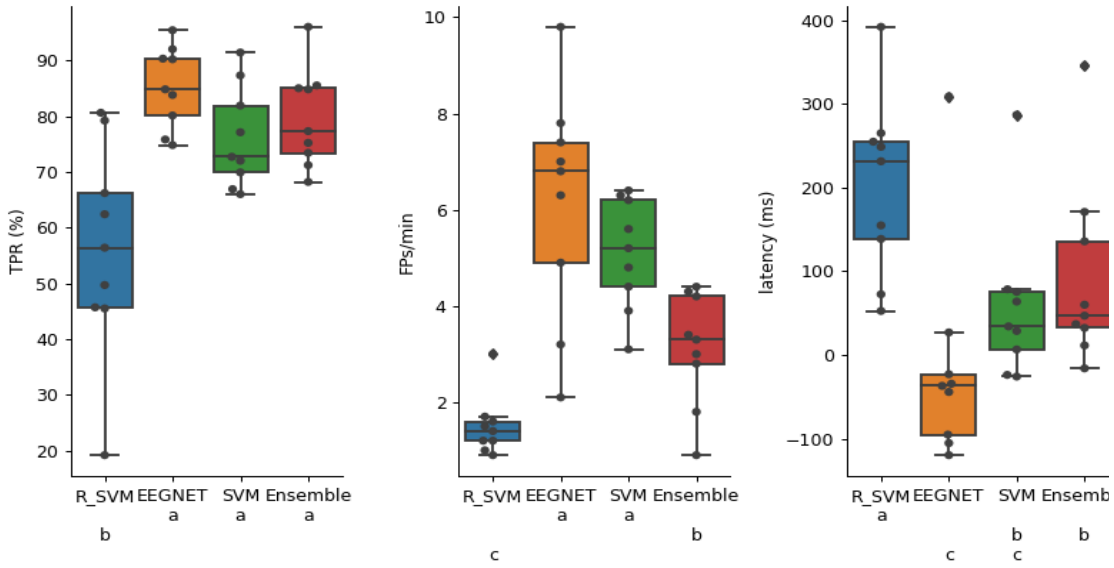

Figure 17. Boxplots of TPR, FPs/min and latency of nine subjects for four algorithms with statistical groupings. TPR, FPs/min and latency are shown from left to right in three graphs. In each graph, R_SVM, EEGNET, SVM and ensemble method are shown as blue, yellow, green, and red, respectively. Boxplot of each algorithm is plotted. The individual subject values are shown as black points. The statistical groupings are shown below each graph.

of 0.004. Overall, the system performs steadily on most tested subjects. Four algorithms are further compared regarding TPR, FPs/min, and latency in one window scenario, as shown in Figure 17. R_SVM has the worst TPR and latency but the best FPs/min in four algorithms based on its statistical grouping. The unbalanced performance of R_SVM in three metrics also shows up in Figure 15 and Figure 16 in multi-window scenarios. Like R_SVM, EEGNET is unbalanced in three metrics: it has the first-tier TPR and latency while the worst level of FPs/min in terms of statistical grouping. Besides, these two algorithms have unstable performance across nine subjects, revealed as the large variability of FPs/min in EEGNET and large variability of TPR and latency in R_SVM. Compared to R_SVM and EEGNET, SVM has more balanced performance across three metrics and relatively stable variability among nine subjects. As its statistical grouping suggests, the ensemble method has the first-level TPR, middle-level FPs/min, and middle-level latency. In addition, it also performs stably in nine subjects with low variability in three metrics.

4.3.4 Influence of Four Movements on Detection Performance

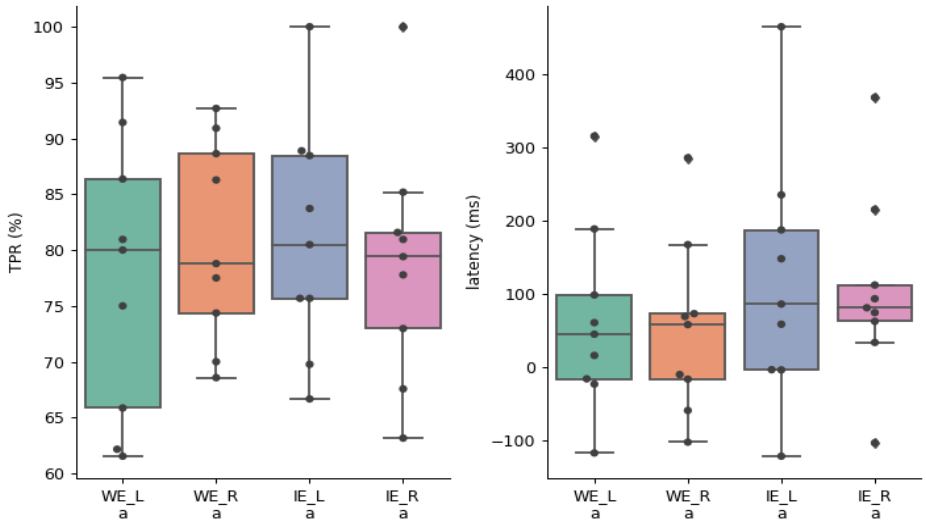

Figure 19. Boxplots of TPR and latency using the ensemble method for four classes with statistical grouping. TPR and latency are shown from left to right in two graphs. In each graph, WE_L, WE_R, IE_L and IE_R are shown as green, orange, purple, and pink respectively. Boxplot of each movement class is plotted with the individual subject value as black point. The statistical grouping are shown below each plot.

Subject	WE_L		WE_R		IE_L		IE_R	
	TPR (%)	Latency (ms)	TPR (%)	Latency (ms)	TPR (%)	Latency (ms)	TPR (%)	Latency (ms)
1	86.4	-16.6	90.9	-17.1	83.7	85.5	81.0	33.1
2	91.4	44.3	78.8	-59.8	88.9	186.8	79.4	80.4
3	62.2	-117.9	92.7	57.4	66.7	-122.3	77.8	-104.1
4	80.0	-23.8	86.3	72.4	88.5	-4.2	85.2	61.9
5	95.5	60.1	88.6	68.6	100.0	58.0	100.0	111.5
6	81.0	15.3	77.5	166.7	69.8	147.4	81.6	92.8
7	65.9	97.9	68.6	-10.8	75.7	-4.4	63.2	73.8
8	75.0	315.0	70.0	285.3	75.7	464.4	73.0	367.8
9	61.5	188.2	74.4	-103.1	80.5	234.6	67.6	214.7
Mean ± Std	77.6±11.8	62.5±120.0	80.9±8.5	51.1±112.3	81.0±9.9	116.2±160.4	78.7±10.0	103.6±121.8

Table 2. TPR and latency of WE_L, WE_R, IE_L, IE_R using Ensemble method. TPR and latency are listed for every subject and every movement. The last row shows the mean and standard deviation of nine subjects for each metric.

Based on one-window detection model, the impact of movement on detection performance was investigated. Figure 18 demonstrates the TPR and latency of four movements using the ensemble method. The performance details for each subject could be found in Table 2. FPs/min was not analyzed because an FP does not have a movement label by default. Two-way mixed-factor ANOVA tested the significant effects of movements with subjects as random factor with alpha of 0.05.

No significant difference was found in movements or subjects. However, the TPR of WE_R and IE_L are generally better than WE_L and IE_R by observing the left graph of Figure 18. Besides, the latency of IE_L and IE_R are slightly higher than WE_L and WE_R by observing the right graph of Figure 18. Subject 5 has significantly better TPR than others in IE_L and IE_R. In addition, subject 8 has a significantly longer latency than others, consistent with the previous finding.

4.3.5 Trial Distribution of Detection Latency

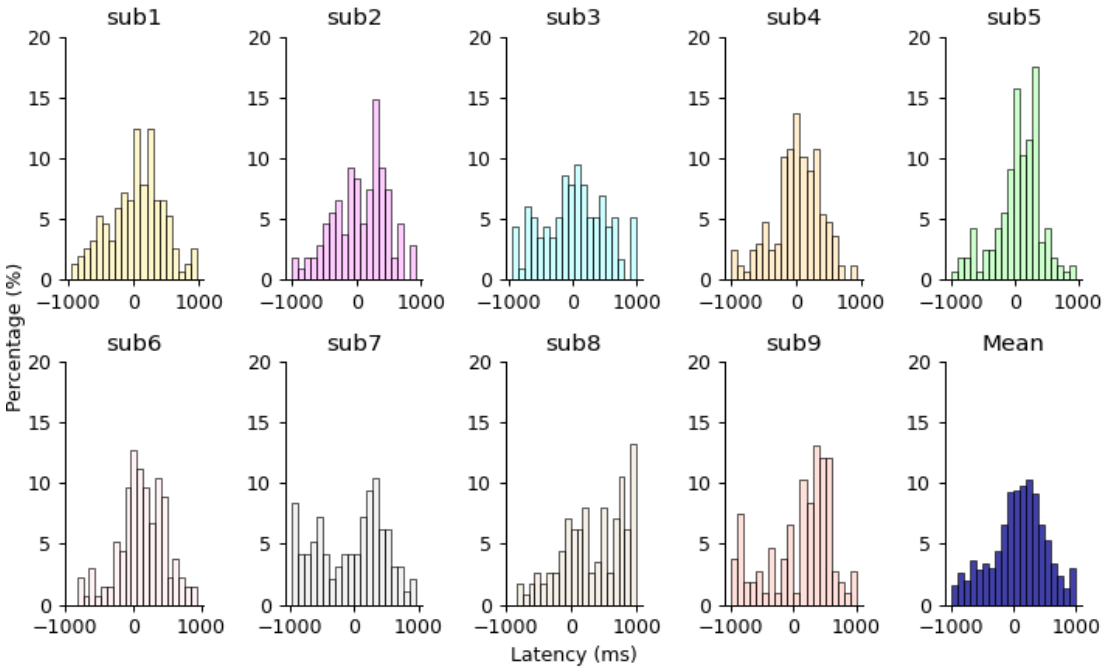

Figure 20. Histograms of detection latency for nine subjects and their mean. The histograms of detection latency using ensemble method are plotted for nine subjects and their mean. In each histogram, x-axis stands for the latency time ranging from -1000ms to 1000ms and y-axis is the trial percentage ranging from 0 to 20%. 20 bins are plotted for each histogram by calculating the percentage of trials fallen into every 100ms period.

Detection latency has been presented as the average of subjects in the previous sections. However, the trial distribution of detection latency inside each subject might reveal more information for the detection performance. Figure 19 shows the histograms of latency for nine subjects and the mean of them. Twenty bins were obtained in each histogram from -1000ms to 1000ms since the testing window gives prediction every 100ms in pseudo-online detection. First, we calculated the number of trials that happened inside each 100ms. Then, they were divided by the total number of trials to obtain the percentage value for each bin in histograms. The mean histogram represents the average latency of nine subjects.

The average value of the mean histogram is around 300ms to 500ms, inferring that most of the detections occurred in this interval. Most subjects' latency histograms are similar to the mean histogram and follow the normal distribution, such as subject 1, subject 2, subject 3, subject 4, subject 5, subject 6, and subject 9. However, the left tail of subject 7's histogram is more like uniform distribution. Besides, subject 8's histogram is right-skewed, with most detections happening in the last 200ms.

Due to the sweet period of inducing Hebbian plasticity being within 400ms to 500ms of the movement onset, the histograms of latency provide another perspective to evaluate the system's efficiency for plasticity induction [82]. From the mean histogram, around 75% of the trials are detected from -500ms to 500ms of movement onset, proving most of the trials would be valid for rehabilitation.

4.4 Pseudo-online Classification

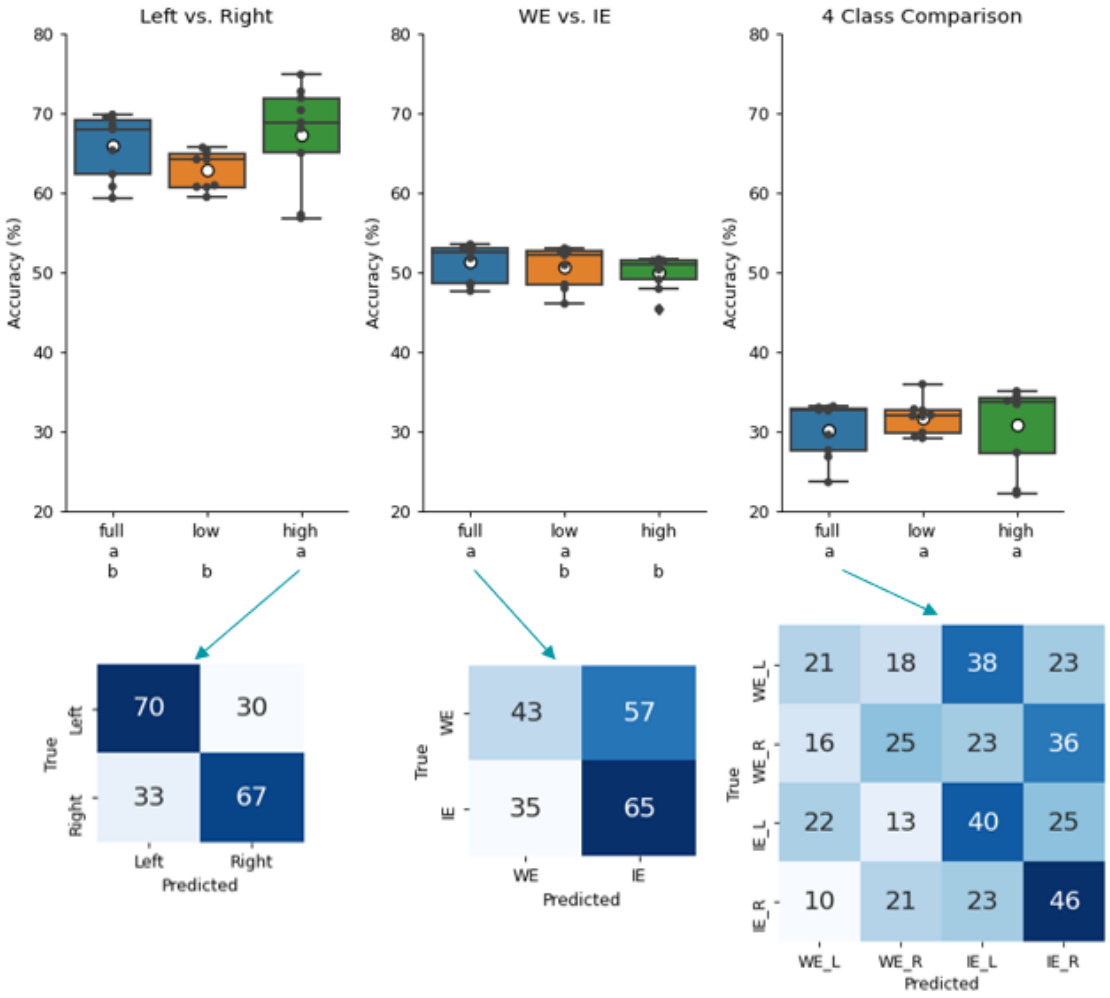

Figure 21. Accuracy and confusion matrix of three classifications. Three different classification results are shown as left vs. right, WE vs. IE and 4-class comparisons in left, middle and right columns, respectively. The upper row shows the classification accuracy boxplots of three frequency bands. For each boxplot, the individual points of nine subjects are shown in black point and the mean accuracy of nine subjects is shown as white point. The statistical groupings are listed below on x-axis. In the lower row, the confusion matrices are shown for the optimal frequency band in each classification case, which are high band, full band, and low band for left vs. right, WE vs. IE and 4 class comparison, respectively. The arrows indicate the match between optimal frequency band in accuracy plots and the normalized confusion matrices. The confusion matrices are normalized so that the sum of each row is 100.

As introduced in 3.3.1.2, three types of classification, including left vs. right, WE vs. IE, and 4-class comparison, are conducted and showed their accuracy and confusion matrices results in Figure 20. In addition, two-way mixed-factor ANOVA was applied to test the statistical differences among three frequency bands in each classification type with subjects as the random factor and the frequency band as the fixed factor. Their statistical groupings are shown as letters below the boxplots.

Observing the accuracy boxplots of left vs. right shows that the high band gives the best performance in terms of mean accuracy and is significantly better than the low band. The normalized confusion matrix was drawn for the high band classification, showing that around 70% of the left and 67% of the right movements are correctly classified. The classification between WE and IE is shown in the middle column of Figure 20, the full-band trained classifier performed the best and is significantly better than the high-band trained classifier. The normalized confusion matrix of full band classification shows that 65% of IE and 43% of WE were classified correctly. As a result, high band contributes most to left vs. right classification, while full band gives the best classification between WE and IE. In addition, it can be seen that the classification between left and right is easier than the classification between WE and IE by comparing these two types of classification,.

The 4-class classification is presented in the right column of Figure 20. There is no significant difference found among the three frequency bands. However, the mean accuracy of the low band is slightly higher than the other two. Observing the confusion matrix shows that IE_L and IE_R have 40% and 46% accuracy, respectively, as shown in the diagonal, above chance level. However, WE_L and WE_R have 21% and 25% accuracy, respectively, around the chance level. In addition, WE is more likely to be predicted as IE. 38% of WE_L were predicted as IE_L and 36% of WE_R were predicted as IE_R by observing the 4-class normalized confusion matrix.

4.5 Time-locked Classification

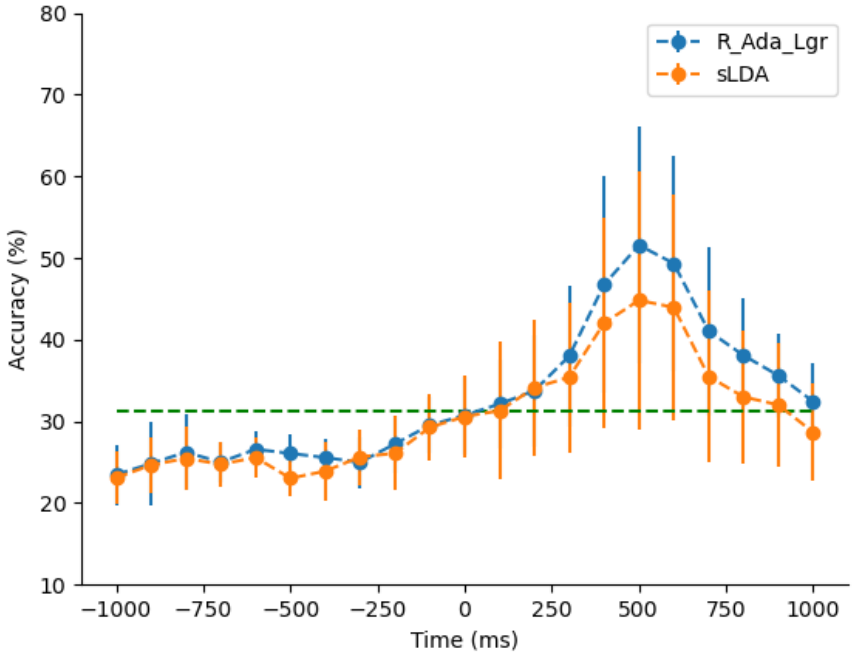

Figure 22. 4-class time-locked classification using R_Ada_Lgr and sLDA. The accuracies of 4-class classification are shown in blue and orange for R_Ada_Lgr and sLDA, respectively. 21 accuracy points are reported for each window with interval of 100ms ranging from -1000ms to 1000ms. The standard deviation among subjects of each accuracy point is shown as bar with corresponding color. The green line is the significant chance level (31.26%) calculated by adjusted Wald interval with alpha of 0.05.

As introduced in 3.3.2, the effect of time on classification performance was investigated in a time-locked analysis, the 4-class time-locked classification was conducted on low band EEG using two algorithms, including R_Ada_Lgr and sLDA. By observing Figure 21, it can be found that the mean accuracies of two classifiers fluctuate from -1000ms to -300ms below the significant chance level, and the standard deviation among subjects is small in this range, suggesting the classification information is not sufficient for all subjects in this period. Then they slowly increase from approximately 25% at -300ms to around 35% at 200ms and surpass the significant chance level (31.26%) at 100ms. The sharp increases happen from 200ms to 500ms with around 27% accuracy boost. R_Ada_Lgr and sLDA achieved 51.8% and 45.2% mean accuracy at 500ms, respectively. Finally, they both decrease to around the significant

chance level at 1000ms. Mean accuracies of the two algorithms are above the significant chance level between 100ms and 900ms, indicating the most relevant information for classification is in this range. The standard deviation in this period is much larger than before, suggesting considerable classification accuracy variability among subjects. Overall, the trends of the two algorithms are similar and show a time-locked peak at 500ms. However, R_Ada_Lgr is better than sLDA from 200ms to 1000ms.

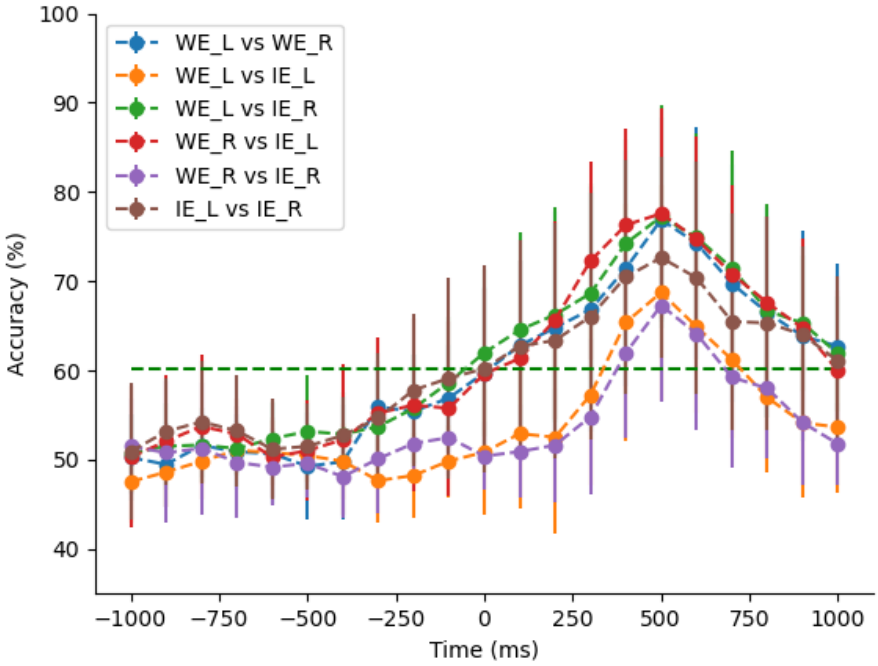

Figure 23. Pairwise time-locked classification using R_Ada_Lgr. The accuracies of 6-pair binary classification with two movements are shown in different colors. 21 accuracy points are reported for each binary classification with interval of 100ms ranging from -1000ms to 1000ms. The standard deviation of each accuracy point is shown as bar with corresponding color. The green line is the significant chance level (60.1%) calculated by adjusted Wald interval with alpha of 0.05.

R_Ada_Lgr conducts the classifications between pairs of movements to analyze the difference of movements. From Figure 22, we can see that the accuracy curves of 6 pairwise classifications have a similar trend, increasing from -1000ms to 500ms and dropping from 500ms to 1000ms. All pairs are time-locked at 500ms to achieve peak accuracy as the 4-class classification shown in Figure 21. However, performance difference exists among pairs. For

example, the orange and purple are lower than others from -500ms to 1000ms curves, representing WE_L vs. IE_L and WE_R vs. IE_R, respectively. In addition, they cross the significant chance level later than other curves by approximately 400ms. The lower accuracy of these two curves indicates it is difficult to differentiate the same-limb movements, consistent with the subtle difference of ipsilateral MRCP shown in Figure 12. The red and green curves representing WE_R vs. IE_L and WE_L vs. IE_R respectively are constantly higher than others between 200ms and 700ms, suggesting that movement type and sidedness contribute to the classification. The other two pairs, including WE_L vs. WE_R and IE_L vs. IE_R, perform between the two groups above.

In conclusion, both 4-class and pairwise classifications show the time-locked phenomenon, the peak accuracy in 500ms, inferring that the most relevant information for classification resides in the period between -1.5s and 0.5s. Furthermore, contralateral movements are easier classified than ipsilateral movements, aligning with the physiological results shown in 4.1.

Chapter 5

Discussion

5.1 Influence of Frequency Band on Detection and Classification

This study trained three individual classifiers, including SVM, R_SVM, and EEGNET, on low, high, and full bands EEG for pseudo-online detection. F1 score provided a unified metric when comparing the combinations between frequency bands and algorithms.

Frequency band plays an important role in detection. As discussed in 4.3, full-band and low-band feature generally perform better than high-band feature. When SVM and EEGNET were used as detector, statistically significant differences between the high band and others two were found, indicating that the high band EEG from 5Hz to 40Hz are not as valuable as the full or low band for the detection task. The supporting evidence could be found in different morphologies in MRCP and SMR from Figure 12 and Figure 13. MRCP has a more prominent negative peak time-locked with movement onset compared to ERD/ERS. By comparing the full band and the low band, it was found that the full band is mostly better than the low band, which suggests the combination of SMR and MRCP would give better results than solely MRCP features. Similar findings were also reported in other studies. Slow cortical potentials (SCP), filtered between 0.1Hz and 6Hz, were found to be more synchronized with the movement onsets than ERD, which might be due to the phase-locked and time-locked nature of SCP [63], A self-initiated walking study also proved that MRCP is more valuable than ERD to predict gait premovement. Furthermore, they provided methods to combine MRCP and ERD features in concatenated and meta ways, which showed better performance than either MRCP or ERD individually [64].

Detection algorithms have their own preference for frequency bands. For example, although the full band is regarded in general as a valuable frequency band for detection using R_SVM and EEGNET, the low band provides the best F1 score for SVM as shown in Figure 14. We can find R_SVM performs significantly better in the full band than the low band, EEGNET prefers the full band, and SVM prefers the low band by observing the median values of their violin plots. The covariance matrix design might explain why R_SVM prefers the full band.

As introduced in 3.2.4, the covariance between a class template and a trial can only be large if the morphology of the trial matches one of the class templates, in which way the temporal information is decoded. MRCP provides such information since the low-frequency signal manifests the shape as their main discriminant feature. However, another part of the covariance matrix (sample covariance of the trial) contains the spatial information, in which ERD shows better contra-lateralization than MRCP. As a result, the full band consisting of both MRCP and ERD performs best for R_SVM. EEGNET prefers full band because it emulates the FBCSP, which requires learning power features by bandpass filtering over a broad bandwidth range in the first convolutional layer. The theoretical reason for SVM preferring low band is difficult to explain. One possible inference could be that it is more sensible for low-frequency features. Empirically, SVM was tested in various MRCP studies and proved its robustness [65][66].

Feature extraction algorithms could also have their preference over frequency bands. For instance, one study reported that the discriminative canonical pattern matching (DCPM) technique performed best in the delta band while CSP performed best in the beta band [67].

The classification was influenced by the frequency band as well. As shown in Figure 20, high-band EEG is more helpful to classify between left and right movements than both low-band and full-band EEG, which could be explained by MRCP hardly showing any noticeable difference between left and right movement, while the topographical maps of ERD showing the contra lateralization in FC2, C2, and C4. One study classified left and right keyboard typing using features from delta band to beta band and reported that the best frequency band is dependent on subjects [67]. The classification between WE and IE shows a different scenario. Full-band EEG provides significantly more discriminant information than the high-band EEG, but it is not significantly better than the low-band EEG. Since neither ERD nor MRCP shows noticeable spatial or morphology features for ipsilateral movements, the full band might combine the information from these two bands and provide better classification performance. A study investigating lateral and palmar grasp supports this finding since they claimed the combination of alpha-band feature and MRCP is better than only using MRCP for classification between two grasps [68]. For the classification among four movements, three frequency bands do not show a significant difference.

5.2 Feasibility of Ensemble Learning

Wisdom of crowd is a vivid explanation for ensemble learning, which strategically aggregates multiple classifiers to make wiser decisions. Standard methods of aggregating classifiers include bagging, boosting, voting, and stacking [69]. The current thesis study constructed the detection model using hard majority voting among three different classifiers: EEGNET, R_SVM, and SVM. The majority voting method did not provide further training to existing models for detection ensemble learning but rather combined their results. This approach is usually suitable for integrating strong learners, classifiers with good performance with strengths and weaknesses towards different metrics. For instance, EEGNET gave the highest TPR but also the highest FPs/min, R_SVM performs the opposite of EEGNET and SVM is between them. In this circumstance, majority voting successfully combined three classifiers in a complementary way to provide balanced TPR and FPs/min.

However, hard majority voting is not the only method to ensemble strong learners. Soft majority voting assigns different weights to individual classifiers and votes based on their performance. However, looking for appropriate weights could take a long time by emulating the possible combinations of weights. Stacking solves this parameter searching problem by training a separate classifier to learn the characteristics of individual classifiers and combine them in a more sophisticated manner. Nevertheless, training such a classifier would require more training data and computational time to generate sufficient samples. For example, we need to divide the dataset into three parts: a training set for individual classifiers, a training set for the stacking classifier, and a testing set for the ensemble classifier. Lesser training data would add more difficulties to getting a good individual classifier. In BCI studies, the ensemble of algorithms often happens in both the feature extraction and the classification stage. One study computed four detection models by combining MRCP and ERD features in different ways: MRCP model, ERD model, concatenated model, and META model. The MRCP and ERD models are individual classifiers based on MRCP and ERD features, respectively. The concatenated model combines MRCP and ERD features to feed into a classifier. The META model is similar to a stacking method, which interprets the output of MRCP and ERD models using LDA. Their results showed that META model gives the best F 0.5 score and accuracy

performance [64]. Another study found the optimal frequency band for two feature extraction methods (DCPM and CSP) and combined their decision using LDA in both feature and decision phases [67].

In the classification phase of the current study, ensemble learning was applied by adaptive boosting multiple logistic regression models using Riemannian features. Classification ensemble learning also outperformed sLDA, which was used by plenty of studies to obtain optimal classification results [70]-[74]. AdaBoost could be used to train other classifiers as well. One study used AdaBoost extreme learning machine (ELM) for motor imagery and showed better classification accuracy than individual ELM, SVM, and LDA [75]. Another study applied AdaBoost SVM with regularized CSP (RCSP) to achieve better performance than previous methods and found AdaBoost may be a good match for RCSP features [76].

5.3 Detection

Decoding movement intention from rest state EEG has always been a challenge in the BCI field for rehabilitation. Such a system requires high TPR and low FPs/min to provide an acceptable detection rate and a short latency for inducing Hebbian plasticity. In terms of the developing phase of such a system, efforts have been put into researching in offline, pseudo-online and online BCI.

Due to the simplicity of offline analysis, tons of studies have been conducted to develop and test every chain of the BCI loop, including preprocessing, feature extraction, and classification algorithms [55][77][78]. Both movement execution and motor imagery have been investigated for detection [75]. Offline studies usually extract the equal length of non-movement and movement epochs as two classes to provide a balanced dataset to a classifier. However, rehabilitation therapy patients spend most of their time resting and are only required to think about moving during short intervals, which could potentially nullify the hypothesis in balanced dataset. Besides, participants were usually asked to perform rest as a task in a constrained condition by keeping the body still and not blinking eyes. The non-movement epochs collected this way can not represent the humans' normal rest state, which will add another layer of difficulty transferring offline to online BCI. Compared to the constrained conditions adopted

by offline studies, our experimental settings allow user to blink. Last but not least, the movement epochs of offline studies always cover long periods after movement onset, which contains the most discriminative motor features. Nevertheless, rehabilitation-aimed detection requires predicting movement intentions as early as possible, without access to after-onset information. Overall, offline analysis is not ideal for practical usage, but it does provide a stable and unified platform for different studies to compare their processing methods.

In order to tackle the overlooked aspects in offline analysis, pseudo-online processing was proposed to process data after data collection on time windows. The current study successfully detected WE and IE movements with an average TPR of 79.6% ± 8.8% and 3.1 ± 1.2 FPs per minute with an average latency of 91.4 ± 111.9ms. Plenty of research has adopted TPR and FPs/min as an evaluation metric for the detection model. One study investigated ballistic ankle dorsiflexion using MRCP in healthy participants and obtained 82.5% ± 7.8% TPR, 6.9 ± 7.41 FPs in 5 minutes, and -66.6 ± 121ms latency [79]. Another study of pedaling intent achieved 76.7% TPR with 4.94 FPs/min on five healthy subjects [80]. Direct comparison with our study is not easy since their targeting movements are lower limb, known to have more profound MRCP features than upper-limb movements. Studies working on upper-limb movement often report false positive rate (FPR) rather than FPs/min. They often conduct window prediction within pre-selected trials like offline studies, which might not represent the non-movement state well if their epoch length is short. For example, one study detected palmar and pincer grasps in asynchronous mode and reported a TPR of 83.49% and FPR of 13.84%. Their FP was regarded as detection made from -2s to -0.3s despite the actual non-movement period being around 5s per trial. This partially clipped rest period could lead to fewer FPs [63]. Compared to them, the rest period of the current study was chosen to be -8s to -2s and 2s to 8s relative to movement onset for each trial, which covered most of the non-movement period and would provide a more representable false positive performance.

Although most studies are still in offline or pseudo-online stages, some attempts were made to detect both lower-limb and upper-limb movements in real-time. One study proposed a manifold-based method called LPP to efficiently reduce EEG dimensionality and preserve intrinsic features in the meantime. They asked healthy participants to execute ankle

dorsiflexion and provided real-time visual feedback. Their results were statistically better than the conventional method, with a TPR of 79% ± 11%, 1.4 ± 0.8 FPs/min, and 315 ± 165ms latency [35]. A following study from the same group demonstrated the feasibility of provoking neural plasticity in healthy participants by using the same algorithm to detect ankle dorsiflexion with a BCI-driven motorized ankle-foot orthosis as feedback. They measured motor evoked potential (MEP) before and after the intervention of transcranial magnetic stimulation (TMS) and found significantly increased activity in MEP, which suggested the occurrence of neural plasticity. They reported TPR as 73% ± 10.3% and FPs/min as 1.3 ± 0.5 [81]. Compared to their previous study, the slight drop in performance might be due to the motor imagery paradigm adopted this time, which has less significant MRCP features than the execution of movement. A preliminary study was conducted for upper-limb movement detection as well, in which they attempted to detect hand open and palmar grasp using MRCP on one spinal cord injury (SCI) patient. In the best session, they obtained 36.9% TPR and 3.6 FPs/min with 2.2s latency [73]. Their following study adopted a similar hierarchical approach to first detect reach-and-grasp movement from rest state and then classify 'move & look' against 'look' on 20 healthy participants. They obtained 54.1% TPR, 1.2 FPs/min with 1.2s latency in online detecting scenario [82]. Although improvements were achieved in all metrics compared to their preliminary study, there is still a performance gap with lower-limb studies, which further proved the challenges in upper-limb movement detection.

Short latency is a critical factor for inducing neural plasticity. Too short or too long latency beyond 400ms to 500ms range is considered not helpful [82]. The current thesis study reports average latency of around 91.4 ± 111.9ms, which is inside this range. Furthermore, we presented the trial distribution of latency in Figure 19, which shows that around 75% of trials were detected within -500ms to 500ms. However, few studies disclosed the trial distribution of latency, which makes it difficult to imply their system's efficiency.

5.4 Classification

The current thesis study investigated classification among four movements in both time-locked and pseudo-online ways and showed the feasibility of differentiating movements of two sides

and of two types on the same side. Time-locked analysis showed better accuracy than pseudo-online analysis for all classification scenarios. For example, the maximum classification accuracy between left and right was around 75% at 500ms after movement onset in time-locked mode. However, the average classification accuracy between left and right using the optimal band in pseudo-online mode was approximately 66%. Similar effects were found for WE vs. IE and 4 class classification as well. The time-locking effect of classification may contribute to the accuracy discrepancy between the two analyses. During pseudo-online classification, the timing of the window is not fixed but depends on the detection latency. By observing Figure 19, we can find the distribution of detection latency is approximately normal, with the mean at around 300ms to 500ms for the average of nine subjects. This means that although most of the windows were detected in the best period for classification around 500ms, considerable amounts of trials were detected out of this range, which leads to the potential accuracy drop in pseudo-online mode. This phenomenon indicates that the pseudo-online analysis considers more complicated and practical aspects, which reiterates the importance of analyzing algorithms in pseudo-online mode.

Sides of the movement are easier classified than movement types from the same limb in both pseudo-online and time-locked scenarios. During time-locked classification between pairs of movements, we can find the peak accuracy of WE_L vs. IE_L and WE_R vs. IE_R are lower than other pairs containing both left and right movements. The accuracy of pseudo-online left vs. right classification (around 67%) is also higher than the classification accuracy between WE and IE (around 52%). This might imply the spatial information is more critical in classification since the temporal information is similar for all movements. Contra-lateralization was observed in both MRCP and ERD for two-side movements, while WE and IE did not show such differences on the same side. The cortical homunculus could partially explain this because wrist and index finger mapping areas are close. This phenomenon was also reported in previous studies. One study experimented with palmar and lateral grasp from both left and right arms offline. In their confusion matrix reported on the evaluation set, the palmar grasp and lateral grasp were more frequently misclassified than left and right grasps [83].

The discrepancy of classification performance also exists between WE and IE. The classifier performs better when recognizing IE by observing WE vs. IE and 4-class confusion matrices in Figure 20. The worse classification of WE may be due to they have shorter detection latency (around 50ms) than IE (around 110ms), as shown in Figure 18, which are further away from the optimal classification time (500ms) as shown in time-locked analysis.

5.5 Study Limitation and Future investigation

The current thesis study adopted a hierarchical pipeline to first detect movements from the rest state and then classify them against each other, giving a proof-of-concept solution for training multiple movements during BCI-assisted rehabilitation. Although detection gave promising results, the classification was still not ideal. There could be several factors leading to the current limitation in hierarchical pipeline design. First, the amount training windows for detection is much more than that for classification due to the experiment setup. Second, the classifications were made based on the detected windows with varied latencies, which might not favor classification since a time-locking effect exists. Future work would focus on designing better classification strategies such as dynamically looking for the best classification timing after detection rather than relying on the same detection window. Another major limitation of this study is that it is still not online, although the processing was considered more practical aspects than traditional offline study. There could be some foreseeable challenges between pseudo-online and online study, including real-time preprocessing and classification algorithm design, algorithm and platform latency, and appropriate plasticity measurement.

Chapter 6

Conclusion

The objectives were as follows:

1. Detect intentions of the four movements from rest state in a pseudo-online manner and investigate the optimal way of combining algorithms and frequency band features (MRCP and SMR) to achieve a high detection rate with low latency.
2. Evaluate the possibility of classifying the four movements against each other following the detection in a pseudo-online manner and understand the contribution of MRCP and SMR during classification.

Objective 1 was complete. We investigated three EEG frequency bands, including low band, high band, and full band, and found out the primary contribution of MRCP during detection. An ensemble learning algorithm was proposed using majority voting among SVM, EEGNET, and R_SVM, achieving state-of-the-art performance for distal simple upper-limb movement detection with an average TPR of 79.6% ± 8.8%, 3.1 ± 1.2 FPs per minutes and an average latency of 91.4 ± 111.9ms. Further explorations were conducted in the effect of window number and movement types on detection performance. Objective 2 was also complete. A novel algorithm R_Ada_Lgr was proposed and was proved to have better performance over the current state-of-the-art algorithm sLDA. Analysis on frequency bands was performed for three types of classifications in a pseudo-online manner: 1) left vs. right, 2) WE vs. IE, 3) WE_L vs. WE_R vs. IE_L vs. IE_R. The left vs. right classification achieved above chance level performance, while the distinction between WE and IE on the same limb was not significantly better than chance. Further investigations were carried out in time-locked analysis and found out the time-locking effect of classification could be the major limitation for classification.

Overall, this study proved the feasibility to detect upper-limb distal simple movements with a high detection rate in short-latency and the possibility to classify left against right movements during pseudo-online analysis.

References

[1] Langhorne P, Bernhardt J, Kwakkel G. Stroke rehabilitation. The Lancet. 2011 May 14;377(9778):1693-702.

[2] Shakeel A, Navid MS, Anwar MN, Mazhar S, Jochumsen M, Niazi IK. A review of techniques for detection of movement intention using movement-related cortical potentials. Computational and mathematical methods in medicine. 2015 Oct;2015.

[3] Yuan H, He B. Brain–computer interfaces using sensorimotor rhythms: current state and future perspectives. IEEE Transactions on Biomedical Engineering. 2014 Mar 19;61(5):1425-35.

[4] López-Larraz E, Montesano L, Gil-Agudo Á, Minguez J. Continuous decoding of movement intention of upper limb self-initiated analytic movements from pre-movement EEG correlates. Journal of neuroengineering and rehabilitation. 2014 Dec;11(1):1-5.

[5] Jankelowitz S, Colebatch J. Movement-related potentials associated with self-paced, cued and imagined arm movements. Experimental brain research. 2002 Nov;147(1):98-107.

[6] Sacco RL, Kasner SE, Broderick JP, Caplan LR, Connors JJ, Culebras A, Elkind MS, George MG, Hamdan AD, Higashida RT, Hoh BL. An updated definition of stroke for the 21st century: a statement for healthcare professionals from the American Heart Association/American Stroke Association. Stroke. 2013 Jul;44(7):2064-89.

[7] Donkor ES. Stroke in the century: a snapshot of the burden, epidemiology, and quality of life. Stroke research and treatment. 2018;2018.

[8] Langhorne P, Coupar F, Pollock A. Motor recovery after stroke: a systematic review. The Lancet Neurology. 2009 Aug 1;8(8):741-54.

[9] Patel AT, Duncan PW, Lai SM, Studenski S. The relation between impairments and functional outcomes poststroke. Archives of physical medicine and rehabilitation. 2000 Oct 1;81(10):1357-63.

[10] Dobkin BH. Focused stroke rehabilitation programs do not improve outcome. Archives of neurology. 1989 Jun 1;46(6):701-3.

[11] Tanaka H, Toyonaga T, Hashimoto H. Functional and occupational characteristics predictive of a return to work within 18 months after stroke in Japan: implications for rehabilitation. International archives of occupational and environmental health. 2014 May;87(4):445-53.

[12] Dobkin BH. Strategies for stroke rehabilitation. The Lancet Neurology. 2004 Sep 1;3(9):528-36.

[13] Liu W, Waller SM, Kepple T, Whitall J. Compensatory arm reaching strategies after stroke: induced position analysis. Journal of rehabilitation research and development. 2013;50(1):71.

[14] Gracies JM, Pradines M, Ghédira M, Loche CM, Mardale V, Hennegrave C, Gault-Colas C, Audureau E, Hutin E, Baude M, Bayle N. Guided Self-rehabilitation Contract vs conventional therapy in chronic stroke-induced hemiparesis: NEURORESTORE, a multicenter randomized controlled trial. BMC neurology. 2019 Dec;19(1):1-1.

[15] Cramer SC, Sur M, Dobkin BH, O'brien C, Sanger TD, Trojanowski JQ, Rumsey JM, Hicks R, Cameron J, Chen D, Chen WG. Harnessing neuroplasticity for clinical applications. Brain. 2011 Jun 1;134(6):1591-609.

[16] von Bernhardi R, Eugenín-von Bernhardi L, Eugenín J. What is neural plasticity?. The plastic brain. 2017:1-5.

[17] Cervera MA, Soekadar SR, Ushiba J, Millán JD, Liu M, Birbaumer N, Garipelli G. Brain-computer interfaces for post-stroke motor rehabilitation: a meta-analysis. Annals of clinical and translational neurology. 2018 May;5(5):651-63.

[18] Willett FR, Avansino DT, Hochberg LR, Henderson JM, Shenoy KV. High-performance brain-to-text communication via handwriting. Nature. 2021 May;593(7858):249-54.

[19] Anumanchipalli GK, Chartier J, Chang EF. Speech synthesis from neural decoding of spoken sentences. Nature. 2019 Apr;568(7753):493-8.

[20] Lotte F, Bougrain L, Cichocki A, Clerc M, Congedo M, Rakotomamonjy A, Yger F. A review of classification algorithms for EEG-based brain–computer interfaces: a 10 year update. Journal of neural engineering. 2018 Apr 16;15(3):031005.

[21] Ramadan RA, Vasilakos AV. Brain computer interface: control signals review. Neurocomputing. 2017 Feb 5;223:26-44.

[22] Waldert S. Invasive vs. non-invasive neuronal signals for brain-machine interfaces: will one prevail?. Frontiers in neuroscience. 2016 Jun 27;10:295.

[23] Wang Y, Yan J, Wen J, Yu T, Li X. An intracranial electroencephalography (iEEG) brain function mapping tool with an application to epilepsy surgery evaluation. Frontiers in neuroinformatics. 2016 Apr 25;10:15.

[24] Pfurtscheller G, Aranibar A. Evaluation of event-related desynchronization (ERD) preceding and following voluntary self-paced movement. Electroencephalography and clinical neurophysiology. 1979 Feb 1;46(2):138-46.

[25] Homan RW, Herman J, Purdy P. Cerebral location of international 10–20 system electrode placement. Electroencephalography and clinical neurophysiology. 1987 Apr 1;66(4):376-82.
[26] Wang Y, Jung TP. Visual stimulus design for high-rate SSVEP BCI. Electronics letters. 2010 Jul 26;46(15):1057-8.
[27] Picton TW. The P300 wave of the human event-related potential. Journal of clinical neurophysiology. 1992 Oct 1;9(4):456-79.
[28] Shibasaki H, Barrett G, Halliday E, Halliday AM. Components of the movement-related cortical potential and their scalp topography. Electroencephalography and clinical neurophysiology. 1980 Aug 1;49(3-4):213-26.
[29] McFarland DJ, McCane LM, David SV, Wolpaw JR. Spatial filter selection for EEG-based communication. Electroencephalography and clinical Neurophysiology. 1997 Sep 1;103(3):386-94.
[30] Ramoser H, Muller-Gerking J, Pfurtscheller G. Optimal spatial filtering of single trial EEG during imagined hand movement. IEEE transactions on rehabilitation engineering. 2000 Dec;8(4):441-6.
[31] Rivet B, Souloumiac A, Attina V, Gibert G. xDAWN algorithm to enhance evoked potentials: application to brain–computer interface. IEEE Transactions on Biomedical Engineering. 2009 Jan 23;56(8):2035-43.
[32] Aboalayon KA, Faezipour M, Almuhammadi WS, Moslehpour S. Sleep stage classification using EEG signal analysis: a comprehensive survey and new investigation. Entropy. 2016 Sep;18(9):272.
[33] Karimi F, Niu J, Almeida Q, Jiang N. Movement Related Cortical Potentials in Parkinson's Disease Patients with Freezing of Gait. In2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) 2020 Jul 20 (pp. 2857-2860). IEEE.
[34] Subasi A, Gursoy MI. EEG signal classification using PCA, ICA, LDA and support vector machines. Expert systems with applications. 2010 Dec 1;37(12):8659-66.
[35] Xu R, Jiang N, Lin C, Mrachacz-Kersting N, Dremstrup K, Farina D. Enhanced low-latency detection of motor intention from EEG for closed-loop brain-computer interface applications. IEEE Transactions on Biomedical Engineering. 2013 Dec 5;61(2):288-96.
[36] Ravi A, Pearce S, Zhang X, Jiang N. User-specific channel selection method to improve SSVEP BCI decoding robustness against variable inter-stimulus distance. In2019 9th International IEEE/EMBS Conference on Neural Engineering (NER) 2019 Mar 20 (pp. 283-286). IEEE.
[37] Ang KK, Chin ZY, Wang C, Guan C, Zhang H. Filter bank common spatial pattern algorithm on BCI competition IV datasets 2a and 2b. Frontiers in neuroscience. 2012 Mar 29;6:39.
[38] Lotte F, Congedo M, Lécuyer A, Lamarche F, Arnaldi B. A review of classification algorithms for EEG-based brain–computer interfaces. Journal of neural engineering. 2007 Jan 31;4(2):R1.
[39] Woehrle H, Krell MM, Straube S, Kim SK, Kirchner EA, Kirchner F. An adaptive spatial filter for user-independent single trial detection of event-related potentials. IEEE Transactions on Biomedical Engineering. 2015 Feb 10;62(7):1696-705.
[40] Vidaurre C, Schlogl A, Cabeza R, Scherer R, Pfurtscheller G. Study of on-line adaptive discriminant analysis for EEG-based brain computer interfaces. IEEE transactions on biomedical engineering. 2007 Feb 20;54(3):550-6.
[41] Barachant A, Bonnet S, Congedo M, Jutten C. Riemannian geometry applied to BCI classification. InInternational conference on latent variable analysis and signal separation 2010 Sep 27 (pp. 629-636). Springer, Berlin, Heidelberg.
[42] Cichocki A, Phan AH, Zhao Q, Lee N, Oseledets IV, Sugiyama M, Mandic D. Tensor networks for dimensionality reduction and large-scale optimizations. part 2 applications and future perspectives. arXiv preprint arXiv:1708.09165. 2017 Aug 30.
[43] Phan AH, Cichocki A. Tensor decompositions for feature extraction and classification of high dimensional datasets. Nonlinear theory and its applications, IEICE. 2010;1(1):37-68.
[44] Craik A, He Y, Contreras-Vidal JL. Deep learning for electroencephalogram (EEG) classification tasks: a review. Journal of neural engineering. 2019 Apr 9;16(3):031001.
[45] Lawhern VJ, Solon AJ, Waytowich NR, Gordon SM, Hung CP, Lance BJ. EEGNet: a compact convolutional neural network for EEG-based brain–computer interfaces. Journal of neural engineering. 2018 Jul 27;15(5):056013.
[46] Wang X, Hersche M, Tömekce B, Kaya B, Magno M, Benini L. An accurate eegnet-based motor-imagery brain–computer interface for low-power edge computing. In2020 IEEE International Symposium on Medical Measurements and Applications (MeMeA) 2020 Jun 1 (pp. 1-6). IEEE.
[47] Khatwani M, Hairston WD, Waytowich N, Mohsenin T. A low complexity automated multi-channel EEG artifact detection using EEGNet. In2019 IEEE EMBS Conference on Neural Engineering 2019 Mar.
[48] Riyad M, Khalil M, Adib A. MI-EEGNET: A novel convolutional neural network for motor imagery classification. Journal of Neuroscience Methods. 2021 Apr 1;353:109037.

[49] Arvaneh M, Guan C, Ang KK, Quek C. EEG data space adaptation to reduce intersession nonstationarity in brain-computer interface. Neural computation. 2013 Aug 1;25(8):2146-71.
[50] Gayraud NT, Rakotomamonjy A, Clerc M. Optimal transport applied to transfer learning for P300 detection. InBCI 2017-7th Graz Brain-Computer Interface Conference 2017 Sep 18 (p. 6).
[51] Shibasaki H, Hallett M. What is the Bereitschaftspotential?. Clinical neurophysiology. 2006 Nov 1;117(11):2341-56.
[52] Neshige R, Lüders H, Shibasaki H. Recording of movement-related potentials from scalp and cortex in man. Brain: a journal of neurology. 1988 Jun 1;111:719-36.
[53] Ikeda A, Lüders HO, Burgess RC, Shibasaki H. Movement-related potentials recorded from supplementary motor area and primary motor area: role of supplementary motor area in voluntary movements. Brain. 1992 Aug 1;115(4):1017-43.
[54] Wang T, Deng J, He B. Classifying EEG-based motor imagery tasks by means of time–frequency synthesized spatial patterns. Clinical Neurophysiology. 2004 Dec 1;115(12):2744-53.
[55] Bai O, Lin P, Vorbach S, Li J, Furlani S, Hallett M. Exploration of computational methods for classification of movement intention during human voluntary movement from single trial EEG. Clinical Neurophysiology. 2007 Dec 1;118(12):2637-55.
[56] Oliver Jones. Muscles in the Posterior Compartment of the Forearm [Internet]. [place unknown] [publisher unknown] 2020 [updated 2020; cited 2021 Spet. 21]. Available from: https://teachmeanatomy.info/upper-limb/muscles/posterior-forearm/
[57] Geron A. Hands-on machine learning with scikit-learn and TensorFlow. Sebastopol, CA: O'Reilly Media; 2017
[58] Patrick Winston, Learning Support Vector Machines. 2014 Jan 10 [cited 2021 Sept 24]. Avaiablable from https://www.youtube.com/watch?v=_PwhiWxHK8o
[59] Guan S, Zhao K, Yang S. Motor imagery EEG classification based on decision tree framework and Riemannian geometry. Computational intelligence and neuroscience. 2019 Jan 21;2019.
[60] Mohammad Javad Shafiee, Ensemble Learning. SYDE 675, Pattern Recognition, University of Waterloo. 2020 Mar 17.
[61] Congedo M, Barachant A, Bhatia R. Riemannian geometry for EEG-based brain-computer interfaces; a primer and a review. Brain-Computer Interfaces. 2017 Jul 3;4(3):155-74.
[62] Jochumsen M, Niazi IK, Taylor D, Farina D, Dremstrup K. Detecting and classifying movement-related cortical potentials associated with hand movements in healthy subjects and stroke patients from single-electrode, single-trial EEG. Journal of neural engineering. 2015 Aug 26;12(5):056013.
[63] Omedes J, Schwarz A, Montesano L, Müller-Putz G. Hierarchical decoding of grasping commands from EEG. In2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) 2017 Jul 11 (pp. 2085-2088). IEEE.
[64] Sburlea AI, Montesano L, Minguez J. Continuous detection of the self-initiated walking pre-movement state from EEG correlates without session-to-session recalibration. Journal of neural engineering. 2015 Apr 27;12(3):036007.
[65] Jochumsen M, Niazi IK, Mrachacz-Kersting N, Farina D, Dremstrup K. Detection and classification of movement-related cortical potentials associated with task force and speed. Journal of neural engineering. 2013 Aug 28;10(5):056015.
[66] Wang J, Bi L, Fei W, Guan C. Decoding Single-Hand and Both-Hand Movement Directions From Noninvasive Neural Signals. IEEE Transactions on Biomedical Engineering. 2020 Oct 27;68(6):1932-40.
[67] Wang K, Xu M, Wang Y, Zhang S, Chen L, Ming D. Enhance decoding of pre-movement EEG patterns for brain–computer interfaces. Journal of neural engineering. 2020 Jan 24;17(1):016033.
[68] Schwarz A, Pereira J, Lindner L, Müller-Putz GR. Combining frequency and time-domain EEG features for classification of self-paced reach-and-grasp actions. In2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) 2019 Jul 23 (pp. 3036-3041). IEEE.
[69] Géron A. Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems. O'Reilly Media; 2019 Sep 5.
[70] Pereira J, Ofner P, Schwarz A, Sburlea AI, Müller-Putz GR. EEG neural correlates of goal-directed movement intention. Neuroimage. 2017 Apr 1;149:129-40.
[71] Ofner P, Schwarz A, Pereira J, Müller-Putz GR. Upper limb movements can be decoded from the time-domain of low-frequency EEG. PloS one. 2017 Aug 10;12(8):e0182578.
[72] Schwarz A, Ofner P, Pereira J, Sburlea AI, Müller-Putz GR. Decoding natural reach-and-grasp actions from human EEG. Journal of neural engineering. 2017 Dec 6;15(1):016005.

[73] Ofner P, Pereira J, Schwarz A, Müller-Putz GR. Online detection of hand open vs palmar grasp attempts in a person with spinal cord injury. InGBCIC 2019 Sep.
[74] Schwarz A, Höller MK, Pereira J, Ofner P, Müller-Putz GR. Decoding hand movements from human EEG to control a robotic arm in a simulation environment. Journal of neural engineering. 2020 May 28;17(3):036010.
[75] Gao L, Cheng W, Zhang J, Wang J. EEG classification for motor imagery and resting state in BCI applications using multi-class Adaboost extreme learning machine. Review of scientific instruments. 2016 Aug 11;87(8):085110.
[76] Miao Y, Yin F, Zuo C, Wang X, Jin J. Improved RCSP and AdaBoost-based classification for motor-imagery BCI. In2019 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA) 2019 Jun 14 (pp. 1-5). IEEE.
[77] Yom-Tov E, Inbar GF. Detection of movement-related potentials from the electro-encephalogram for possible use in a brain-computer interface. Medical and Biological Engineering and Computing. 2003 Jan;41(1):85-93.
[78] Kato YX, Yonemura T, Samejima K, Maeda T, Ando H. Development of a BCI master switch based on single-trial detection of contingent negative variation related potentials. In2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society 2011 Sep (pp. 4629-4632). IEEE.
[79] Niazi IK, Jiang N, Tiberghien O, Nielsen JF, Dremstrup K, Farina D. Detection of movement intention from single-trial movement-related cortical potentials. Journal of neural engineering. 2011 Oct 26;8(6):066009.
[80] Rodríguez-Ugarte M, Iáñez E, Ortíz M, Azorín JM. Personalized offline and pseudo-online BCI models to detect pedaling intent. Frontiers in neuroinformatics. 2017 Jul 11;11:45.
[81] Xu R, Jiang N, Mrachacz-Kersting N, Lin C, Prieto GA, Moreno JC, Pons JL, Dremstrup K, Farina D. A closed-loop brain–computer interface triggering an active ankle–foot orthosis for inducing cortical neural plasticity. IEEE Transactions on Biomedical Engineering. 2014 Mar 26;61(7):2092-101.
[82] Pereira J, Kobler RJ, Ofner P, Schwarz A, Müller-Putz GR. Online detection of movement during natural and self-initiated reach-and-grasp actions from EEG signals. Journal of Neural Engineering. 2021 Jun 15.
[83] Schwarz A, Pereira J, Kobler R, Müller-Putz GR. Unimanual and bimanual reach-and-grasp actions can be decoded from human EEG. IEEE transactions on biomedical engineering. 2019 Sep 23;67(6):1684-95.