From Signal to Control: Real-Time Artifact Handling and the Effect of Preprocessing Strategies in EEG-Based Brain-Computer Interfaces

ADOLF András

PhD Dissertation

Supervisor: Dr. ULBERT István, DSc

Pázmány Péter Catholic University
Roska Tamás Doctoral School of Sciences and Technology

Budapest, 2025

Contents

List of Abbreviations

AR	Artifact Rejection
BCI	Brain-Computer Interface
CNN	Convolutional Neural Network
CSP	Common Spatial Pattern
ECoG	Electrocorticography
EEG	Electroencephalography
EOG	Electrooculogram
ERD	Event Related Desynchronization
ERS	Event Related Synchronization
FASTER	Fully Automated Statistical Thresholding for EEG artifact Rejection
FBCSP	Filter Bank Common Spatial Pattern
FC	Fully Connected
FORCe	Fully Online and Automated Artifact Removal for Brain-Computer Interfacing
ICA	Independent Component Analysis
LDA	Linear Discriminant Analysis
LFP	Local Field Potential
LSTM	Long Short Term Memory
MEG	Magnetoencephalography
MI	Motor Imagery
MST	Motor Simulation Theory
MLP	Multilayer Perceptron
PCA	Principal Component Analysis
RNN	Recurrent Neural Network
SVM	Support Vector Machine
WT	Wavelet Transform

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my supervisor, Dr. István Ulbert, for his continuous support and guidance throughout the entire course of my PhD journey. His expertise, patience, and encouragement were invaluable to me during the most challenging periods of my doctoral studies.

I extend my heartfelt thanks to Dr. Csaba Márton Köllőd, without whom this dissertation would not have been possible. As the leader of the Cybathlon group and the principal developer of the system's underlying codebase, his contributions were fundamental. His willingness to assist whenever I encountered difficulties, both technical and conceptual, was a constant source of motivation and reassurance.

I am also grateful to the other members of the BCI research group, in particular Dr. Gergely Márton and Ward Fadel, for their collaboration and constructive feedback, especially during the preparation of my first-author publication. I would also like to thank Dr. Bálint File for generously sharing his expertise and insights on artifact rejection.

Special thanks are due to the supportive community of the doctoral students in room 240, whose camaraderie and encouragement provided a much-needed sense of belonging throughout the years.

I would also like to express my gratitude to Tivadarné Vida for her patience and invaluable assistance with the administrative aspects of my PhD studies.

I am grateful for the financial and professional support provided by the Doctoral Student Scholarship Program of the Co-operative Doctoral Program (Hungarian abbreviation: KDP) of the Ministry of Innovation and Technology, financed by the National Research, Development and Innovation Fund.

Lastly, I wish to express my sincere appreciation to my family — my parents, my sister, my brother, and my grandparents — for their unwavering emotional support and constant encouragement. Their persistent inquiries into the progress of my work, though sometimes pressing, helped keep me focused and determined to see this journey through to the end.

Chapter 1

Introduction

Brain-Computer Interfaces (BCIs) represent a highly dynamic and promising area of research within neurotechnology and human-computer interaction [1]. Numerous research groups worldwide are actively working on advancing the capabilities and practical applications of these systems. Despite significant progress in recent years, BCIs are still far from being widely accessible or seamlessly integrated into everyday use. Continued research is essential to overcome the existing technical and usability challenges to bring these systems closer to real-world deployment [2].

1.1. Motivation

BCI systems are emerging as transformative tools in assistive technologies, offering individuals with severe motor disabilities a new channel for communication and interaction. Despite decades of research, the performance and reliability of BCI systems are still far from optimal, primarily due to the complexity and variability of EEG signals. These signals are not only inherently low in signal-to-noise ratio but are also highly susceptible to contamination from non-neural artifacts such as eye movements, muscle activity, and environmental noise. Artifact rejection and preprocessing steps play a critical role in determining whether useful neural patterns can be extracted and interpreted accurately [3].

While online artifact rejection methods such as FORCe (Fully Online and Automated Artifact Removal for Brain-Computer Interfacing) [4] and ORICA (Online Recursive Independent Component Analysis) [5] exist, open-source implementations that are easy to modify and integrate are not always available. This dissertation presents an online adaptation of the Fully Automated Statistical Thresholding for EEG artifact Rejection

(FASTER) algorithm [6] that prioritizes transparency, simplicity, and ease of integration, aiming to provide a practical alternative for real-time BCI applications. The system was not only implemented and tested in live settings but was also reviewed and accepted by the Cybathlon organizing committee, highlighting its compliance with technical standards. Its use helped ensure that our BCI-pilots did not rely on artifacts for control, supporting fair competition based on genuine neural activity.

Beyond the development of tools, there is a pressing need to understand the interactions between preprocessing steps, signal representations, and machine learning strategies. Current practices often treat these components in isolation, yet their combined effects significantly shape the final classification outcomes. For example, while artifact rejection generally improves signal quality, its effect is not uniform across subjects - some users benefit, while others experience performance drops. Similarly, transfer learning shows promise but behaves differently depending on the nature of the input data.

Moreover, how EEG data is structured for machine learning - whether using raw time series, frequency-filtered segments, or spatial mappings - affects the ability of neural networks to learn meaningful patterns. Through the development of novel neural architectures and a dense 3D representation, this work also probes the value of incorporating spatial information and the role of frequency bands in classification tasks, especially under cropped training conditions.

In summary, the motivation for this dissertation is twofold:

1. To develop a practical online artifact rejection tool for real-time BCI use, especially in competitive settings.

2. To provide a detailed investigation into how EEG preprocessing, spatial/frequency representations, and learning strategies interact and influence BCI classification accuracy.

By addressing both engineering implementation and theoretical understanding, this work aims to improve the design, reliability, and personalization of future BCI systems.

1.2. Overview of the Thesis

This dissertation is structured around two main groups of thesis points. The first group pertains to the development and implementation of an online version of the FASTER artifact rejection algorithm, as presented in a co-authored publication. This algorithm

was integrated into our Brain-Computer Interface system and employed during the Cybathlon competition. The second group of thesis points is derived from my first-author publication and investigates the impact of various signal processing steps on the classification accuracy of BCI systems.

To establish a comprehensive foundation for these research directions, the dissertation begins with an introduction to Brain-Computer Interfaces, discussing key components such as the types of neural signals used and the paradigms applied for system control. Special emphasis is placed on the motor imagery paradigm, including a review of publicly available datasets designed around it. Significant attention is also given to machine learning methods - particularly neural networks - starting from fundamental principles and progressing toward state-of-the-art models, including transformers.

Chapter 3 presents the design and validation of the proposed Online FASTER algorithm, along with a detailed overview of the Cybathlon competition, where it was deployed. This chapter forms the basis of the first thesis group.

Chapter 4 focuses on the effect of EEG signal preprocessing techniques on classification performance, summarizing the experimental findings reported in my first-author publication. These results underpin the second thesis group.

Finally, Chapter 5 synthesizes the outcomes of the research and formally presents the thesis statements.

Chapter 2

Background and Theory

2.1. Brain-computer interfaces

Brain-computer interfaces provide non-muscular communication between the brain and the external world, using electroencephalographic or other brain-related signals. The main goal of these devices is to help people with severe neuromuscular disorders, such as ALS, brainstem stroke, or spinal cord injury, communicate with the external world and control computer programs or neuroprostheses.

Like every communication or control system, BCIs have an input (for instance, the electroencephalographic activity of the brain), an output (commands to control a device or a computer), an inner system for decision-making, translating the input into output, and a protocol that determines the onset, offset, and timing of operation [1]. The general structure of a BCI system can be observed in Figure 2.1.

The input of the BCI can be categorized into invasive and noninvasive methods and will be discussed in chapters 2.2 to 2.8. The second step of a brain-computer interface is the extraction of features from the input signal. As the signal is usually very noisy, contains many artifacts, and is hard to characterize in its raw form, filtering is essential to get closer to the interpretation. Sections 2.6 and 2.7 discuss these aspects. After filtering, relevant features can be extracted using different methods such as transforming into the frequency domain, using the wavelet transform, extracting common spatial patterns, or adding spatial dimensions to the data. The next step is classification, which is the most discussed part in this dissertation. In this part, a machine learning system classifies the preprocessed input, translating the user's thoughts into control commands.

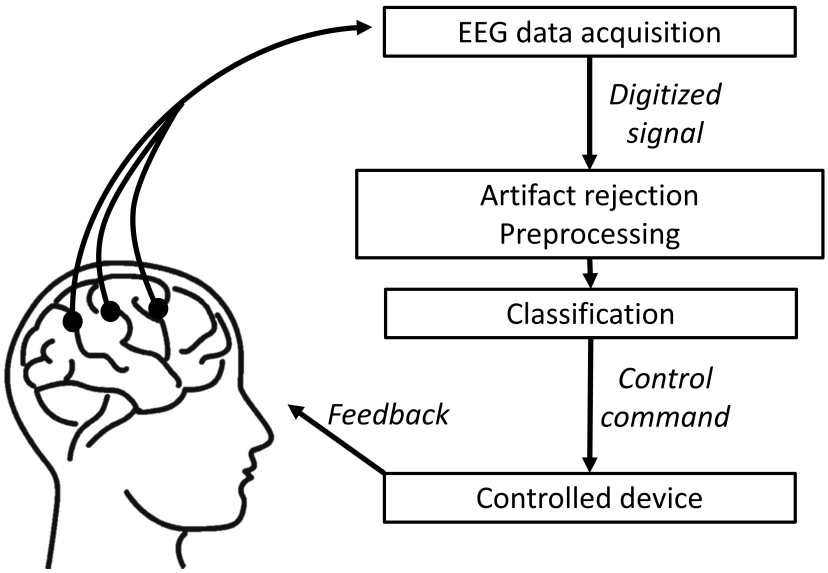

Figure 2.1: The general architecture of a closed-loop BCI system. First, EEG signals are recorded from the scalp and converted into a digitized representation. These data then undergo preprocessing, including filtering, artifact rejection, and other enhancement steps, to obtain a cleaner neural representation. Depending on the application, features may be extracted before a classification algorithm infers the user's intended command, which is subsequently transmitted to an external device. The system delivers feedback to the user, allowing them to refine their mental strategies over time and thereby supporting adaptive BCI operation [J2].

2.2. Electric signals of the brain

2.2.1 The origin of the signals

The human brain is composed of neural cells, or neurons, which possess intrinsic states and communicate through synapses. This intricate system comprises approximately 10–20 billion neurons, each exhibiting an electric potential. Neurons receive input signals via dendrites, which integrate incoming information, and transmit output signals through an axon. When the membrane potential of a neuron surpasses a critical threshold, it generates an action potential, propagating signals to all connected neurons. The movement of ions and the resulting changes in electric potential fields can be recorded using electrodes. Various techniques are available to measure these signals, differing in their level of precision and degree of invasiveness [7].

2.2.2 Invasive methods

Invasive methods of electrical signal acquisition involve procedures where the skull is opened or drilled to create an access point, allowing electrodes to be placed on the brain's surface or inserted into its interior. One notable invasive technique is electrocorticography (ECoG), which has gained popularity for investigating cortical phenomena in clinical contexts. ECoG employs subdural platinum-iridium or stainless-steel electrodes to record electrical activity directly from the cerebral cortex's surface. By bypassing the signal-distorting effects of the skull and intermediate tissues, this method achieves a relatively high signal-to-noise ratio. Spatial resolution can be further enhanced by using flexible, closely spaced subdural grid or strip electrodes [8].

While ECoG and most noninvasive techniques primarily capture electrical activity in the cortex's superficial layers, deeper brain activity can be explored by inserting metal or glass electrodes or silicon probes into the brain. These tools enable the recording of local field potentials (LFPs). Achieving high spatial resolution requires numerous closely spaced observation points and minimal disruption to brain tissue. With a sufficiently high density of recording sites, it is possible to monitor the spiking activity of nearly all neurons within a small volume. Advances in this domain have been driven by the development of micromachined silicon-based probes, which offer an increasing number of recording units and facilitate more comprehensive exploration of neural activity [8].

2.2.3 Non-invasive methods

Noninvasive methods for brain signal acquisition do not require surgical intervention, making them safer and more accessible. However, these methods provide less detailed information about brain activity, as they cannot detect single-neuron or local potentials. Instead, they measure the superposition of extracellular field potentials generated by large neuronal populations, where temporally correlated activity contributes dominantly to the measured signal. Among noninvasive techniques, electroencephalography (EEG) is the most widely used and one of the oldest methods for studying brain electrical activity. Scalp EEG, recorded by individual electrodes, represents a spatiotemporally smoothed version of the LFP, integrated over an area of approximately 10 cm^2 or more. Recent advancements, such as high-density EEG recordings combined with source modeling that incorporates structural MRI-derived anatomical details of gyri and sulci, have significantly improved the spatial resolution of EEG. The distinction between EEG and invasive methods is illustrated in Figure 2.2 [8].

Another prominent noninvasive technique is magnetoencephalography (MEG), which employs superconducting quantum interference devices (SQUIDs) to detect minute magnetic fields generated by neuronal currents outside the skull. MEG offers relatively high spatiotemporal resolution and has the advantage of being less affected by the conductivity properties of the extracellular space compared to EEG [9].

In contrast, functional magnetic resonance imaging (fMRI) represents a fundamentally different noninvasive brain-imaging approach. This method utilizes magnetic resonance imaging to measure the blood oxygenation levels of different brain regions. The BOLD (Blood Oxygen Level Dependent) signal measures the deoxygenated hemoglobin concentration, which is consequent to task-related or spontaneous modulation of neural metabolism. The strengths of this method are the excellent spatial resolution and the imaging of the whole brain, not just the surface of it. At the same time, the drawbacks are the relatively poor temporal resolution (seconds), the price of the machine, and the immense space occupation of the imaging device, which makes it unsuitable for everyday BCI control [10], [11].

Finally, it is important to mention functional near-infrared spectroscopy (fNIRS), a noninvasive and portable brain imaging technique. In fNIRS, emitters project near-infrared light through the scalp and skull, while detectors measure the reflected or transmitted light. Because oxygenated and deoxygenated hemoglobin exhibit different absorption characteristics at near-infrared wavelengths, changes in cerebral hemodynamics

- and consequently, brain activity - can be inferred. Compared to fMRI, fNIRS offers advantages such as portability, lower cost, and greater freedom of movement. Owing to these properties, fNIRS has become an increasingly popular modality for brain–computer interface applications [10], [12].

This study focuses on EEG signals, which will be discussed in greater detail in the following sections.

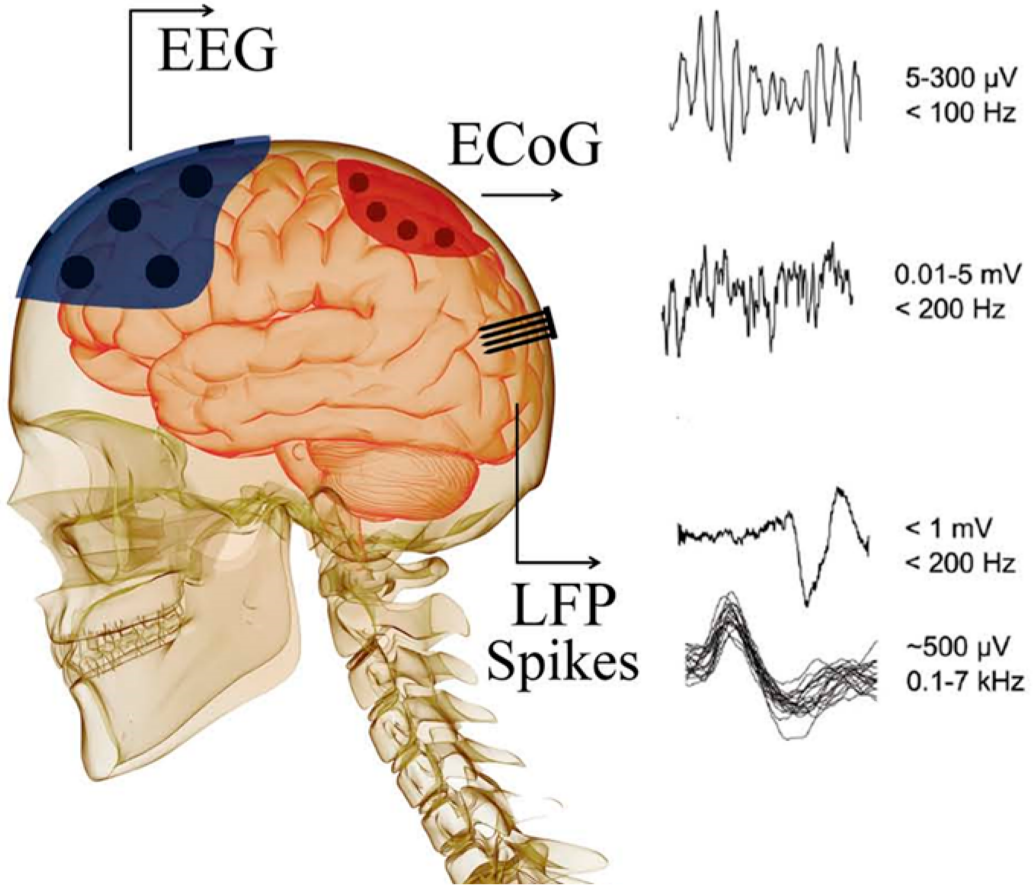

Figure 2.2: Schematic depiction of major neural recording modalities, including EEGs, ECoGs, LFPs, and spikes. Scalp-recorded EEG signals exhibit low amplitudes and contain neural oscillations spanning a wide frequency range, although higher-frequency components typically appear with lower signal-to-noise ratios due to volume conduction and attenuation by the skull. ECoG recordings, obtained directly from the cortical surface, provide higher amplitudes and broader bandwidth. Depth-electrode recordings capture local field potentials, reflecting population-level neural activity, while single-unit spikes represent millisecond-scale action potentials from individual neurons. The right side of the figure illustrates representative waveforms for each modality, together with their typical amplitude and frequency ranges [13].

2.3. EEG signals

The brain's electrical activity, as recorded through surface electrodes, is referred to as the electroencephalogram. EEG signals are typically measured using silver/silver-chloride (Ag/AgCl) electrodes due to their stability, low noise characteristics, and biocompatibility. EEG arises from the synchronized synaptic activity of populations of cortical neurons, which are pyramidal cells organized along cortical columns. The area next to neural dendrites becomes more negative as a result of the post-synaptic activity; therefore, a neuron can be modeled as a dipole. For the electric signals to be measurable from scalp electrodes, the neurons must be aligned in a parallel fashion and be synchronously active [14].

To be detected by an electrode outside the head, neural electrical signals must travel through multiple layers, including the dura mater, skull, and scalp. Once the signal reaches the boundary of a conductive volume, ions can no longer move freely, and capacitance becomes the primary means of propagation. This occurs because insulating layers separate charged regions, forming a stack of capacitors that transmit the signal by repelling charges across each layer. While cerebrospinal fluid and brain tissues are good conductors, they are separated from the electrode by poor conductors like the skull, dead skin, and hair. To achieve less than 5kΩ impedance between the electrode and the scalp, conductive electrolyte gel is commonly applied. Electrode gel improves conductivity by filling air gaps and creating a conductive path from the scalp to the electrode. This gel layer adds another capacitor to the existing stack of insulating and conductive layers. Researchers are developing dry electrodes that do not require gel, but their technology is beyond the scope of this discussion. The EEG signal is largely a result of this capacitive transmission, though some electron flow occurs between the electrode gel and electrode, creating DC potentials. To account for these electrochemical effects, EEG users must allow electrodes to settle before recording. The main advantages of the EEG method are non-invasiveness, relatively easy applicability, and good temporal resolution [14].

2.4. Paradigms

To distinguish EEG signals effectively, it is essential to identify discriminant modalities, which become apparent when specific paradigms are applied or certain tasks are performed. One of the most commonly used paradigms is the kinesthetic imagination of large body movements, such as those involving the hands, legs, feet, and tongue. Imag-

ined movements of these body parts lead to event-related desynchronization (ERD) in the mu and beta frequency bands, while relaxation triggers event-related synchronization (ERS) in the motor cortex. This activity is predominantly observed over the C3 and C4 channels (following the 10/20 electrode placement system) and can be discerned using various feature extraction and classification methods [15]. In this dissertation, I mainly focus on Motor Imagery (MI) tasks, therefore, these types of signals will be discussed in the next chapter.

Steady State Visually Evoked Potentials (SSVEP) are also used for BCI control. This potential is also called photic driving because the origin of this component is the visual cortex. It is evoked by flickering stimuli. A BCI is controlled by the user shifting his attention on parts of the computer monitor that flicker with different frequencies, and the frequency of the observed part is detectable via the EEG signal [15].

Visual P300 paradigms rely on the brain's response to unexpected visual stimuli, commonly known as the Oddball paradigm. The P300 component is characterized by a positive peak, with an amplitude ranging from 5 to 10 microvolts and a latency between 220 ms and 500 ms following the onset of the stimulus. The first Brain-Computer Interface based on this component was developed by Farwell and Donchin in 1988 [16], with the implementation of the P300 Speller, which marked a significant advancement in BCI technology.

There are additional paradigms for BCI control, such as imaginary singing, imaginary rotation of certain objects, imaginary navigation, and performing mathematical operations, such as subtraction [15].

2.5. Motor Imagery

There are multiple explanations of the neural mechanism of motor imagery, with Motor Simulation Theory (MST) being the most widely accepted. MST, first proposed by M. Jeannerod [17], [18], suggests that motor imagery and motor execution share the same neural processes up to the point of muscle activation, where an inhibitory mechanism suppresses the process, at some point between plan encoding and overt action. For this mechanism, M. Jeannerod proposed two possible ways in which it may be realized. The first one is a process inhibiting motor commands from descending past the brainstem or spinal cord, while according to the second proposition, the reason can be a reduced level of motor cortex activation during imagery, resulting in output below the threshold required to activate motor neurons in the spinal cord (and thus below the threshold required for

movement). In either case, MST's core case is the existence of this inhibitory mechanism; otherwise, imaginary movement could not use the same pathways without producing a movement. Supporting evidence includes mental chronometry studies, which show that imagined actions take a similar time as physical ones, and fMRI findings indicating overlapping neural activation in regions such as the central and cingular gyri, supplementary motor area, and inferior parietal lobule during both imagery and execution.

G. Pfurtscheller et al. [19] investigated the reactivity of mu rhythms during motor imagery of the right hand, left hand, foot, and tongue in nine able-bodied subjects using 60 EEG electrodes. Their findings revealed that hand motor imagery led to mu rhythm desynchronization in all subjects, while foot and tongue motor imagery often induced mu rhythm synchronization in the hand area. The most reactive frequency components were centered around 11.7 Hz, with desynchronized components exhibiting broader frequency bands around 10.9 Hz, whereas synchronized components were narrower and displayed slightly higher frequencies at 12.0 Hz. Importantly, classification performance in multi-class BCI improved when ERS patterns were present alongside ERD in at least one or two tasks, suggesting that these EEG phenomena can enhance BCI control through motor imagery alone. Furthermore, the study highlighted both intersubject and intrasubject variability in mu rhythm reactivity during motor imagery, emphasizing the importance of diverse band power changes for optimal task differentiation in single-trial classification. The observed ERD/ERS patterns were specifically elicited by imagined movements, reinforcing their potential utility in motor imagery-based BCI systems. However, the authors also noted a challenge in distinguishing multiple mental states based solely on ERD patterns, as alpha rhythm desynchronization can be influenced by various psychophysiological factors, including perception, memory processes, and task complexity. This underscores the need for incorporating both ERD and ERS features to improve multi-class BCI performance.

2.6. Motor Imagery EEG databases

2.6.1 Physionet

My research was performed on the EEG motor movement/imagery dataset recorded by Schalk et al. as part of the Physionet Database [20]. Data was recorded with a 64-channeled 10-10 EEG system with 160 Hz sampling frequency and by using the BCI2000 framework, without hardware filters [21]. It is one of the largest EEG datasets of motor

imaginary tasks, consisting of recordings from 109 subjects, 14 files for each. In my work, I excluded subjects 88, 92, and 100, due to the sampling frequency and data structure mismatch. I also omitted subject 89, where electrode labels were found to be incorrect. These problems were also reported previously [22], [23].

The dataset contains 5 classes of motor imagery tasks, namely baseline activity and imagined activities of right-hand, left-hand, both-hands, and both-legs movements. Although the database includes EEG signals of tasks, where movements were actually realized, I only used data from imagined movements to train and test the systems, as tetraplegic people are the target patients for BCI research for whom only imagined movements are available. During experiments, I used 4-way classification, for the following classes: right-hand, left-hand, both hands, and both legs. There are approximately 22 samples per subject per class, resulting in 88 samples for each subject.

2.6.2 BCI Competition IV – 2a, 2b

A key benchmark for EEG classification tasks is the BCI Competition IV dataset, particularly datasets 2a and 2b, which were both recorded by C. Brunner, E. Leeb et al. [24]. The number of subjects in these datasets is limited to only nine. The 2a database includes EEG recordings from 22 Ag/AgCl electrodes with an inter-electrode distance of 3.5 cm. All signals were recorded monopolarly, with the left mastoid serving as the reference and the right mastoid as the ground. This dataset contains four motor imagery classes: left hand, right hand, foot, and tongue movements. During data collection, subjects were seated in an armchair in front of a computer screen. Each trial began with the appearance of a fixation cross for 3 seconds, followed by an arrow pointing left, right, up, or down, indicating the specific MI task to perform. Participants were instructed to execute the motor imagery task until the cross disappeared at 6 seconds, resulting in 3 seconds of actual motor imagery per trial.

In contrast, the 2b dataset consists of EEG recordings from only three electrodes (C3, Cz, and C4), sampled at 250 Hz. This experiment involved only two motor imagery classes: lifting the right or left hand. Although both datasets are widely used for BCI performance evaluation, they were not included in this study due to the limited number of subjects.

2.7. EEG Artifacts

Considering the field of EEG, two distinct categories of artifacts are recognized: physiological/biological and non-physiological artifacts. The latter can be caused by the measurement instrument, meaning faulty electrodes, powerline- and environment noise, high impedance of electrodes, cable, or body movement artifacts. These can be mostly avoided by a precise recording system and strict recording procedures [3]. Physiological artifacts may arise from various sources, including cardiac activity, pulse, respiratory patterns, and glossokinetic effects. The two most significant contributors to physiological artifacts are Electrooculograms (EOG), and Electromyograms (EMG). EOGs can come from ocular movements such as eye blinking, eye movement, and eye flatter, while EMGs can result from chewing, clenching, swallowing, sniffing, and talking [25], [26]. The components of the observable EEG signal can be seen with the main artifact sources in Figure 2.3.

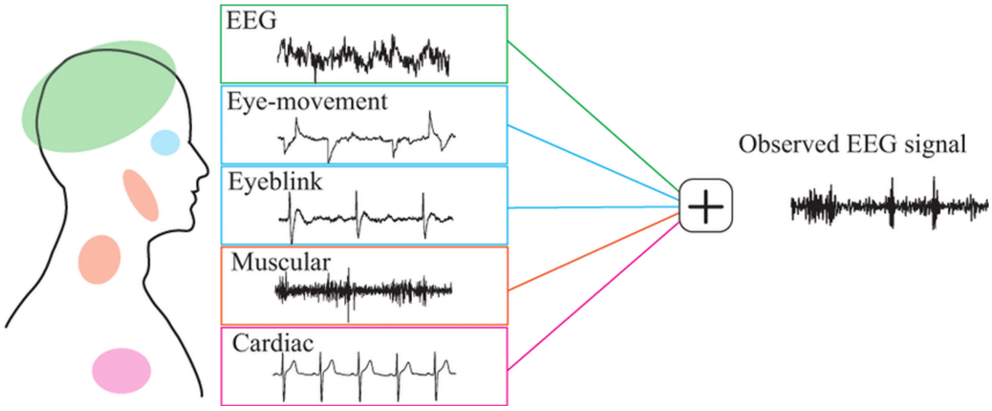

Figure 2.3: Decomposition of the observed EEG signal into its constituent sources. Alongside genuine neural activity, EEG recordings capture non-neural artifacts arising from eye movements, blinks, facial and scalp muscle activity, and cardiac rhythms. These components sum together at the electrodes, producing the measured EEG signal that must be cleaned during preprocessing [27].

It is crucial to remove artifacts from the EEG signal prior to processing, as they can potentially interfere with the interpretation of the original data. One important aspect arising during BCI system development is the fact that we want to use this system, for example, on tetraplegic or locked-in patients, who are unlikely to produce certain types of EEG artifacts; however, we usually develop, test, and compare the BCI systems

dominantly on data derived from healthy subjects. It is clear that healthy subjects will have a different artifact portfolio than patients with severe brain injuries. In addition, nearly all the large databases suitable for neural network-type machine-learning-based BCI system research contain EEG data from healthy subjects, and not from injured patients. Motor imagery tasks in a healthy subject may evoke, for example, covert, unwanted limb movement-related artifacts, which action is evidently missing in tetraplegic patients. In addition, it is also a question of how disabled patients' brain network changes after the injury, and whether their brain processing is comparable with that of normal subjects during motor imagery tasks. In conclusion, at least heuristically, it seems to be a safe strategy to remove artifacts from the EEG to minimize the difference between healthy and diseased populations.

2.8. Artifact rejection

To remove artifact contamination, a filtering algorithm needs to be employed. In the academic literature, several methods have been proposed for this purpose, of which the Independent Component Analysis (ICA) is one of the most frequently utilized mathematical methods [28], [29], [30]. In addition to ICA, other methods such as Wavelet Transformation [31], Canonical Correlation Analysis [32], Empirical-Mode Decomposition [33], and further Hybrid approaches [3] are also commonly employed. One of the widely used algorithms for Artifact Rejection (AR) is the FASTER algorithm [6], also utilized in this work, which, besides using the ICA method, also performs filtering and interpolation over global and epoch-wise bad channels.

2.9. Feature extraction

EEG data is inherently high-dimensional, composed of multiple overlapping neural sources, and is characterized by a low signal-to-noise ratio even after the application of artifact rejection techniques. The extraction of discriminative features - those capable of distinguishing between different cognitive or motor states - is critical for effective Brain-Computer Interface systems. However, designing an appropriate and informative feature set remains a major challenge. The neural signals of interest are often embedded in a noisy background and are temporally and spatially overlapped with activity stemming from unrelated or concurrent brain processes. Consequently, simple signal processing techniques such as band-pass filtering are frequently insufficient to isolate the desired

information, particularly when the target signal is obscured by other concurrent neural sources [34].

To address this, various feature extraction and dimensionality reduction methods have been proposed. Among the most widely used dimensionality reduction methods are Principal Component Analysis (PCA) and the previously mentioned ICA. PCA is a statistical method that applies a linear transformation to convert a set of possibly correlated variables into a set of uncorrelated components, referred to as principal components, which capture the directions of maximum variance in the data. In contrast, ICA attempts to decompose a multivariate signal into statistically independent sources without prior knowledge of their nature, making it especially suitable for blind source separation in EEG signals [34].

Nonlinear approaches for the decomposition of EEG signal include autoencoders, which are inherently neural network–based architectures (discussed in detail in the following section). Conceptually, their objective is similar to certain dimensionality reduction techniques, such as PCA, which can be viewed as a linear analogue. An autoencoder consists of two main components: an encoder, which compresses the input into a lower-dimensional representation, and a decoder, which reconstructs the original input from this latent space. The intermediate (latent) layer generally has fewer dimensions than the input and the output layers. When properly trained, the encoder can serve as an effective feature extractor, while together with the decoder component can even be utilized for artifact rejection [35], [36], [37].

In addition to spatial and statistical methods, time-frequency domain analysis is often employed to capture transient or non-stationary features of EEG signals. Two notable techniques in this domain are Matched Filtering (MF) and Wavelet Transform (WT). Matched Filtering is a pattern recognition approach that detects predefined signal templates within noisy EEG data by measuring the degree of correlation between the incoming signal and a set of reference templates. This method is particularly useful when the temporal profile of the neural response is known in advance [38]. Wavelet Transform, on the other hand, offers a flexible and robust framework for analyzing non-stationary signals. It decomposes the EEG into time-frequency representations, allowing the extraction of localized spectral features that are otherwise difficult to capture using traditional Fourier-based approaches. WT has found broad applications in signal processing, including EEG, audio, and image data [39], [40], [41].

Another important family of methods focuses on spatial feature extraction, with Common Spatial Pattern (CSP) being one of the most prominent. CSP is specifically designed to enhance the discriminability between two classes of multichannel EEG signals by projecting the data into a spatial subspace that maximizes the variance for one class while minimizing it for the other. This transformation significantly improves the signal-to-noise ratio of class-specific patterns, making the subsequent classification more effective. Although originally developed for binary classification problems, extensions of CSP have been proposed to accommodate multiclass BCI paradigms [42], [43]. Furthermore, frequency-aware variants such as Filter Bank CSP (FBCSP) extend the method by applying CSP across multiple frequency bands, enabling the extraction of band-specific spatial features that are particularly effective in motor imagery–based BCIs [44].

The next chapter will delve into classification techniques. Neural networks, particularly transformer-based architectures, represent the current state of the art, while convolutional neural network-based models remain relevant and widely used.

Unlike traditional classifiers, convolutional neural networks have the capacity to learn discriminative features directly from raw EEG data through hierarchical layers of convolution and abstraction. As a result, in many modern BCI applications, explicit feature extraction steps are omitted altogether, and the raw signals are fed directly into the neural networks, which autonomously learn the relevant representations for classification tasks.

2.10. Classification

Machine learning plays a crucial role in the classification of EEG signals, enabling the system to determine the class to which a given EEG epoch belongs. This requires designing and training a classification model using a large number of trials. In this section, I focus on two widely used classifiers, Support Vector Machines and Convolutional Neural Networks, while also highlighting Recurrent Neural Networks and Transformers as a state-of-the-art approach. However, there are further approaches not explained here, such as Linear Discriminant Analysis (LDA), Regularized-LDA, k-nearest-neighbor algorithm (k-NN), Extra Trees algorithm, or Naïve Bayes algorithm. Additionally, there are other deep learning techniques, such as Deep Belief Networks (DBN), which have improved classification by automatically extracting features. Ensemble methods, including LDA ensembles and tree-based classifiers, have also been introduced to address challenges in multi-class classification. Despite these advancements, real-time implementation re-

mains difficult due to calibration constraints, leading to growing interest in alternative approaches such as capsule networks and spiking neural networks [45]. These methods are important to note, but like the previously mentioned approaches, they will not be discussed in detail in this section.

2.10.1 Support Vector Machines

Support Vector Machines (SVMs) were one of the most successful classifiers for motor imagery EEG signal classification [46], [47] before the era of neural networks, and yet there are articles from the near past applying this method [48], [49]. The essence of this algorithm is constructing a hyperplane as the decision surface on the training data, in such a way that the margin of separation between the 2 classes is maximized. Notationally, given a labeled training dataset of $\{x_i, y_i\}_{i=1}^{n}$ where $x \in \mathbb{R}^n$ and y_i is either 1 or -1, depending on the class, Support Vector Machines solves the next optimization problem:

$$\min_{w,\xi_i,b} \left\{ \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n} \xi_i \right\}$$

constrained to:

$$y_i \left(\langle \varphi(x_i), w \rangle + b \right) \geq 1 - \xi_i \quad \forall i = 1, \ldots, n$$

$$\xi_i \geq 0 \quad \forall i = 1, \ldots, n$$

In this notation w and b are the parameters of the corresponding hyperplane, which determines to which class a new x point belongs. It can be shown that minimizing the norm of the weight vector is equivalent to maximizing the margin between the closest points and the hyperplane. φ is a function which transforms the input vectors into a new basis, ξ_i-s are the error values, with the non-zero values belonging to those training points, which are classified inside the margin, or which are on the wrong side of the hyperplane - the value represents the error from the correct classification. These values are important in such cases, where there is no hyperplane that can classify all training points well [50].

2.10.2 Neural Networks

In contemporary BCI systems, neural networks have become the predominant choice for classification, underscoring their critical role in achieving high performance. This section provides a structured overview, beginning with the fundamental components of

neural networks and progressing toward more advanced architectures that are specifically designed or adapted for use in BCI applications.

2.10.2.1 Artificial neuron

The fundamental unit of a neural network is the artificial neuron, a computational model inspired by the biological neuron found in the human nervous system. The earliest and simplest form of this concept is Rosenblatt's perceptron, which computes a weighted sum of its inputs, adds a bias term, and applies a threshold activation to determine the output. Mathematically, this can be expressed as:

$$output = \begin{cases} 0, \ if \ w \cdot x + b \leq 0 \\ 1, \ if \ w \cdot x + b > 0 \end{cases}$$

where w is the weight vector, x is the input vector, and b is the bias term. Modern neural networks extend this model by incorporating nonlinear activation functions, such as the sigmoid, ReLU, or tanh functions, in place of the step function. These nonlinearities enable the network to model complex, continuous relationships between inputs and outputs. Importantly, it has been mathematically established that a neural network composed of a sufficient number of such neurons - organized into at least one hidden layer - can approximate any continuous function to arbitrary precision. This property, known as the universal approximation theorem, underpins the theoretical power of neural networks in a wide range of learning tasks [51].

2.10.2.2 Fully Connected Neural Networks

The fundamental architecture of artificial neural networks is the Fully Connected (FC) neural network, commonly referred to as the Multilayer Perceptron (MLP). This architecture consists of multiple layers composed of varying numbers of artificial neurons, where each neuron in a given layer l_i is connected to every neuron in the preceding layer l_{i-1}. The network begins with an input layer, ends with an output layer, and may include one or more hidden layers situated between them. These hidden layers serve as non-linear feature extractors, enabling the transformation of input data into a feature space where class separability can be enhanced. Through the process of supervised learning, hidden neurons adapt to emphasize discriminative patterns in the training data, facilitating improved classification or regression performance. This hierarchical representation and transformation of data distinguish the MLP from simpler models like the single-layer perceptron proposed by Rosenblatt [52].

A major limitation of fully connected neural networks is their poor scalability with high-dimensional input data. As the input dimension increases, the number of trainable parameters grows rapidly, leading to significant computational overhead, slow convergence during training, and a heightened risk of overfitting due to the model's excessive capacity. This makes pure fully connected architectures impractical for real-world tasks involving high-resolution inputs, such as images or time-series data.

Despite this drawback, fully connected layers are still widely used - particularly in the final stages of neural architectures. In most classification problems, earlier layers (e.g., convolutional or recurrent layers) first extract and condense meaningful features from the input, significantly reducing its dimensionality. These processed feature vectors are then passed to fully connected layers, which act as decision-making units that map the extracted features to class probabilities or regression outputs. This hybrid approach balances expressive power with computational efficiency [53].

2.10.2.3 Convolutional Neural Networks

Convolutional Neural Networks (CNNs) are extensively utilized in image and video processing, as well as in analyzing temporal data sequences, making them particularly advantageous for EEG data processing. In contrast to fully connected neural networks, which often face challenges related to computational complexity and a high susceptibility to overfitting, CNNs offer a more efficient alternative. As mentioned previously, overfitting occurs when a network demonstrates strong performance on the training dataset but exhibits a significant decline in accuracy on validation and test datasets. This phenomenon is often attributed to an excessive number of parameters and suboptimal network architectures [54]. Fully connected networks are also highly sensitive to input shifts; even minor displacements in the input can lead to substantially different outputs. This property is undesirable in signal processing applications, where stability with respect to small input variations is essential. CNNs offer several advantages over fully connected networks. Firstly, they employ local connections, meaning that neurons are connected only to a subset of neurons from the previous layer rather than the entire layer. This architectural design significantly reduces the number of parameters and accelerates convergence. Secondly, CNNs utilize weight sharing, which further decreases the number of parameters, enhancing computational efficiency. Lastly, the incorporation of pooling layers, which will be discussed in detail later, facilitates dimensionality reduction through downsampling, contributing to improved feature extraction and network performance. To

construct a CNN model, four essential components are typically required. The fundamental operation in convolutional neural networks is convolution, which involves applying a kernel - a small matrix with learnable parameters - to the input data. The kernel's weights are convolved with the input, producing a transformed representation. In the case of two-dimensional data, this operation follows the mathematical formulation:

$$Y(x, y) = \sum_{n=-N}^{N} \sum_{m=-M}^{M} X(x - n, y - m) \cdot W(n, m)$$

where X is the input data, and W is the convolutional kernel.

The output of the convolution process is referred to as feature maps, as different kernels capture distinct features within the data. In addition to defining the kernel size, two crucial hyperparameters must be specified: stride and padding. Stride determines the step size of the convolution operation, influencing the resolution of feature extraction. A larger stride results in fewer overlapping operations, reducing computational cost but potentially losing finer details. Padding addresses edge effects by adding additional values (e.g., zero-padding) around the input data. This ensures that border regions are included in the convolution, preserving spatial dimensions when necessary.

The second fundamental component of a CNN is the activation function, similar to its role in fully connected layers. As demonstrated by He et al. [55], if no activation function is applied or only a linear function is used, the network's output remains a linear transformation of the input, regardless of the number of layers. This results in a model equivalent to a primitive perceptron, which has limited learning capacity due to its inability to capture complex representations. To overcome this limitation, nonlinear activation functions are introduced. The most commonly used activation functions include the sigmoid function, the hyperbolic tangent (tanh) function, the Rectified Linear Unit (ReLU), and its various extensions [56].

The third essential component of a CNN is the pooling operator, which is primarily responsible for reducing the spatial dimensions of the input. This operation serves two key purposes: first, it decreases the number of weights and parameters in subsequent layers, enhancing computational efficiency; second, it helps mitigate the risk of overfitting. Several pooling methods exist, with the most common being max pooling, which selects the maximum value within a defined region, and average pooling, which computes the average value. Additionally, more advanced techniques such as mixed pooling, LP pooling, and stochastic pooling have been developed to further optimize feature extraction and model performance [57].

Finally, after multiple convolutional and pooling layers, CNNs can incorporate a fully connected layer to process high-level feature representations. Although alternative approaches, such as adaptive pooling [58] or fully convolutional neural networks [59], have been proposed, motor imagery EEG processing pipelines most commonly employ at least one fully connected layer. Before entering the FC layers, the feature maps are typically flattened into a one-dimensional vector to match the input dimensions required by the final layers. The FC network plays a crucial role in tasks such as classification, where the output is an n×1 vector, with n representing the number of distinct classes. This final stage enables the network to make predictions based on the extracted features.

Convolutional neural networks are widely used as one of the state-of-the-art classifiers for motor imagery signal classification [60], [61], [62]. Among the many CNN architectures proposed for EEG analysis, Shallow ConvNet [63], and EEGNet [64] are particularly prominent and are described in detail later in this dissertation.

2.10.2.4 Recurrent Neural Networks

Recurrent Neural Networks (RNNs) represent a specialized class of deep learning architectures designed to handle sequential data by incorporating a form of internal memory. Unlike traditional feedforward neural networks, which process inputs independently and without regard to order, RNNs are explicitly designed to model temporal or sequential dependencies. This is achieved through a recursive structure, wherein the network maintains a hidden state that is updated at each time step. The output at a given time step is thus a function of both the current input and the hidden state from the previous time step. This makes RNNs particularly well-suited for time-series data, language modeling, and other tasks where the order of inputs carries semantic meaning.

Despite their conceptual strengths, standard RNNs suffer from limitations when dealing with long sequences, primarily due to the problem of vanishing and exploding gradients during backpropagation through time. These issues hinder the network's ability to learn long-range dependencies, effectively limiting its "memory" to only a few previous time steps. To overcome these limitations, the Long Short-Term Memory (LSTM) architecture was proposed as a more robust variant of RNNs. LSTMs are designed to capture long-term dependencies by incorporating a more complex internal structure, consisting of memory cells and a system of gating mechanisms that regulate the flow of information. Each LSTM unit receives the current input and the previous hidden state, and it outputs

both an updated hidden state and a cell state, which serve as memory carriers through the sequence. The core of an LSTM unit lies in its three gates:

1. Input Gate: Controls the extent to which new input values should update the cell state.

2. Forget Gate: Determines which information from the previous cell state should be discarded.

3. Output Gate: Governs how much of the cell state should be exposed as the output at the current time step.

These gates are parameterized using learned weights and are activated using sigmoid and tanh functions, allowing the network to retain, forget, or update information adaptively over long time horizons. In typical classification tasks, particularly in sequence modeling scenarios such as EEG signal classification, the final hidden state at the last time step is often used as a compact representation of the entire input sequence. Overall, LSTM networks have proven to be significantly more effective than simple RNNs in modeling long-range temporal dependencies, making them a valuable tool in EEG-based BCI systems where sequential dependencies and signal dynamics play a critical role [65], [66], [67], [68], [69].

2.10.2.5 Transformers

One of the novel and widely used structures of machine learning is the Transformer model, developed by Vaswani et al. [70]. The original use of this structure was Natural Language Processing (NLP) and translation, but since its release, it has been used for numerous things, such as text generation, image processing, and signal processing as well [71]. The model follows an encoder-decoder structure, such as many of the previous machine translation models, but it is unique in using only the attention mechanism, without recurrence or convolution. Both the encoder and the decoder consist of multiple layers of multi-head attention and pointwise fully connected layers in the feed-forward part. The encoder processes the data, which has been tokenized, embedded, and positional encoded. The authors employed a residual connection around each sub-layer, followed by layer normalization.

Multi-head attention layers are built up of Scaled Dot-Product Attention blocks. There are 3 inputs of this module: a Query, a Key, and a Value vector. The sets of these

vectors are packed together into matrices Q, K, and V, respectively. The attention is computed by the following formula:

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

where d_k is the dimension of the key vector, the square root of which the dot product is scaled. The $softmax$ function serves to obtain the weights on the values. A multi-head attention layer consists of h number of attention blocks, where the input query, key, and value vectors are linearly projected h times with different, learned linear projections. The formula for multi-head-attention is the following:

$$MultiHead\,(Q, K, V) = Concat(head_1,\ \ldots\ head_h)W^O$$

where:

$$head_i = Attention(QW_i^Q,\ KW_i^K,\ VW_i^V)$$

and Ws denote the parameter matrices of the linear projections.

For EEG decoding, transformers also represent state-of-the-art classification performances; however, in most cases, only the Encoder module is used [72], [73], [74], [75], [76], [77].

A representative model within this category is the EEG Conformer, which integrates a convolutional module inspired by the Shallow ConvNet architecture with a self-attention module based on a Transformer encoder. The convolutional component captures low-level local features through one-dimensional temporal and spatial convolution layers, while the Transformer encoder directly models global dependencies among these local temporal representations. The network achieved classification accuracies of 77.66% and 85.87% on the BCI Competition IV 2a and 2b datasets, respectively. Under the same experimental conditions, Shallow ConvNet attained 75.69% and 85.13%, whereas EEGNet achieved 77.39% and 87.71% [78].

Another Transformer-based model, CTNet, follows a similar hybrid design. It first applies a convolutional module, analogous to that of EEGNet, to extract local and spatial features from the EEG time series. Subsequently, a Transformer encoder employing a multi-head attention mechanism captures global dependencies among high-level EEG representations. Finally, a classification module composed of fully connected layers produces the final output. This architecture achieved state-of-the-art performance on the

BCI Competition IV 2a and 2b datasets, with accuracies of 82.52% and 88.49%, respectively [79].

Jin Xie et al. also conducted experiments employing Transformer-based neural network architectures for EEG classification. They proposed five distinct model configurations, varying in the inclusion of convolutional layers and the use of temporal, spatial, or combined (fusion) Transformer modules. Among their analyses, they evaluated the classification performance on 3-second EEG segments from the Physionet database and compared the results with those of EEGNet and Shallow ConvNet. Their findings indicated that the fusion CNN-Transformer model achieved the highest accuracy of 64.22%, outperforming EEGNet (63.16%) and Shallow ConvNet (58.58%) [80].

2.10.2.6 Training a neural network

To train a neural network for classification, it is essential to have a great labeled dataset. To be able to determine the efficiency of a classifier, it is necessary to have a separate set of data on which the system was not trained, and the accuracy achieved on this Test set will be the measure of the designed architecture. In the method of cross-validation, the whole set is iteratively divided into a Train and a Test set, for each item to be once in the set of evaluation. To prevent overfitting, it is a good idea to determine a third set, the Validation dataset, on which the efficiency is tested in each epoch during training. If the accuracy of the network starts to decline, and this decline lasts for a predetermined patience parameter, training stops, and the final weights will be the ones with the highest validation accuracy. The network with these weights will be tested on the Test set, and the accuracy is determined this way [53].

To understand how neural networks weights are updated during training, we have to know about the backpropagation algorithm. Backpropagation is a fundamental algorithm used to train artificial neural networks by minimizing the error between predicted and actual outputs. It works by propagating the error backward through the network, layer by layer, and adjusting the weights using gradient descent. During this process, the algorithm computes the gradient of the loss function with respect to each weight by applying the chain rule of calculus. These gradients indicate how much a small change in each weight would affect the overall error, allowing the model to learn by gradually updating the weights in a way that reduces the loss. Backpropagation enables deep learning models to learn complex patterns from data and is essential to the success of modern machine learning [53].

2.10.2.7 Parameters to tune

The design and training of a neural network involve the careful tuning of numerous hyperparameters, which can be broadly categorized into architectural, training-related, and data-centric parameters. From an architectural perspective, key decisions include the number of layers and neurons per layer, the types of layers used (e.g., convolutional, recurrent, fully connected), and the overall structural layout of the network. Additional parameters such as activation functions, kernel sizes, stride and padding values, pooling types and sizes, as well as input sequence length and window size, also play crucial roles in defining the network's representational capacity and performance. Training-related hyperparameters must also be carefully chosen to ensure effective learning. Among these, the learning rate is perhaps the most critical, as it determines the magnitude of weight updates during training and has a significant impact on convergence. The choice of optimization algorithm (e.g., Adam, Stochastic Gradient Descent) affects the training dynamics and stability, while parameters such as batch size, number of epochs, and early stopping patience influence the generalization ability of the model. The selection of an appropriate loss function is also essential and depends on the specific task - categorical cross-entropy, for instance, is commonly used in classification problems. To prevent overfitting and improve generalization, regularization techniques such as dropout are often employed. Dropout involves randomly deactivating a proportion of neurons during each training iteration, thereby ensuring that the network does not become overly reliant on specific neurons and learns more robust representations. During inference, the full network is utilized, resulting in improved stability and performance. Finally, data-related parameters are also of considerable importance. Data augmentation techniques - such as adding noise, rotating images, or applying temporal shifts - can enhance the diversity of the training set without requiring additional data collection. Additionally, normalization and scaling of input features are crucial preprocessing steps, as they help ensure numerical stability and consistency across input dimensions.

Collectively, these hyperparameters must be thoughtfully configured to develop a neural network model that is both accurate and generalizable. Their appropriate tuning often requires empirical experimentation and may significantly influence the success of the learning process [81], [82].

Chapter 3

Online Artifact rejection algorithm for Cybathlon 2020 and Cybathlon 2024

3.1. The Cybathlon competition

The Cybathlon is a prestigious international competition organized by ETH Zurich, with the primary objective of fostering the development of assistive technologies and computational solutions aimed at improving the quality of life for individuals with physical disabilities. Often referred to as the "Bionics Olympics," the Cybathlon is held every four years, beginning in 2016. The event features various disciplines that simulate real-world challenges faced by people with disabilities. These disciplines initially included six categories: Exoskeletons, Functional Electrical Stimulation (FES), Powered Wheelchairs, Powered Leg Prostheses, Powered Arm Prostheses, and Brain-Computer Interfaces. For the 2024 edition, two additional categories were introduced: the Assistance Robot Race and the Vision Assistance Race.

Our research group, under the leadership of Csaba Márton Köllőd, participated in the Brain-Computer Interface race. In accordance with Cybathlon protocol, participants with severe motor impairments - referred to as pilots - operate assistive systems via non-invasive brain signals. We conducted weekly experimental sessions with two pilots diagnosed with tetraplegia, focusing on improving the reliability and responsiveness of BCI control systems.

Our initial preparations were aimed at participating in Cybathlon 2020. Unfortunately, due to restrictions imposed by the COVID-19 pandemic, we were ultimately un-

able to take part in the event. The 2020 BCI task involved controlling a virtual vehicle navigating a track. Pilots had to issue directional commands (left or right) to guide the vehicle at turns, and trigger a 'lights-on' command when sections of the track became dark. No command was needed when the vehicle was on a straight path. Incorrect or mistimed commands resulted in a speed reduction, and team performance was scored based on the time taken to complete the entire course.

Between 2021 and the release of the new game for Cybathlon 2024, our efforts continued using the 2020 game as a training and evaluation platform. Upon the release of the updated 2024 challenge, we transitioned to the new and significantly more complex game environment. The 2024 BCI race required a broader range of mental commands to complete five distinct and increasingly sophisticated virtual tasks:

1. Wheelchair Navigation - Maneuvering a virtual wheelchair through various obstacles from a starting door to an exit point.

2. Cup-Filling Task - Precisely guiding a cup beneath an automated ice dispenser and holding it still while ice cubes were dispensed.

3. Mouse Control - Moving a cursor to a designated region on a computer screen and executing a virtual mouse click.

4. Key Insertion - Controlling a virtual hand to insert a key into a keyhole and rotate it in the correct direction.

5. Obstacle Avoidance - Navigating a wheelchair through a space populated with moving robot vacuum cleaners, avoiding collisions to successfully complete the task.

Each of the tasks was given 2 times, resulting in 10 tasks for the 8 minutes of the competition.

A key regulatory requirement of the BCI race is the implementation of artifact filtering prior to signal classification, in order to ensure that the systems rely exclusively on EEG-derived information rather than on extraneous signals (e.g., muscle activity or eye movements). In response to this requirement, I extended the original offline FASTER algorithm to support online operation. This enhancement enabled real-time artifact rejection during gameplay, making our BCI system compliant with competition rules while maintaining signal integrity for effective command recognition.

3.2. The FASTER algorithm

The Fully Automated Statistical Thresholding for the EEG artifact Rejection method (FASTER algorithm) was designed by H. Nolan et al [6]. In this chapter, the steps of this method are explained with particular attention to the parts I modified, compared to the original algorithm. I employed a method comprising four sequential steps, with a deviation from the original algorithm by excluding the final artifact detection step across subjects. The algorithm utilizes statistical criteria to identify channels and components that exhibit deviations exceeding three times the standard deviation of the computed parameters.

In preparation for algorithm application, frequency filtering was implemented, employing a 5th-order Butterworth filter. The initial step of the algorithm involves the identification of globally artifactual channels. Channels exceeding predetermined thresholds for variance, the mean of the channel's correlation coefficients with other channels, or Hurst exponent parameters are flagged as faulty. Subsequently, the algorithm proceeds to eliminate epochs containing artifacts. The examined parameters for this step included amplitude range in epoch, deviation from each channel's average value, and variance in each epoch. The third step involves the utilization of ICA to segregate time-dependent data into statistically independent waveforms. During this process, epochs and channels labeled as defective are disregarded. A transformation is performed using the fast-ICA algorithm. Components displaying excessive correlation with the signal of electrodes proximal to the ocular region are omitted in the resultant space if they fail to meet the Z-score criterion for kurtosis, power gradient, Hurst exponent, and median gradient parameters. To revert to the time domain, inverse matrix multiplication is performed. In the last step, defective channels are determined on an epoch-by-epoch basis. Subsequently, both globally and individually impaired channels are replaced through the spherical spline interpolation technique. Ultimately, the data is referenced to the average of all scalp electrodes.

3.3. The Online FASTER algorithm

In order to enable real-time operation, the original FASTER algorithm required specific modifications to adapt it for online use. The resulting implementation, referred to as the Online FASTER algorithm, is designed to perform artifact rejection and signal filtering in a quasi-real-time fashion. The operational pipeline of this algorithm consists

of a preparatory phase followed by an online filtering stage. The preprocessing phase relies on prerecorded EEG data to estimate several parameters that cannot be reliably computed during online execution. Specifically, during this offline *pretraining* stage, the algorithm identifies and marks globally bad channels based on their statistical deviation from normative characteristics. These channels are then excluded from further computation. Additionally, bad epochs are detected within the training data and removed to ensure the integrity of subsequent steps. The most computationally demanding component of the algorithm - the Independent Component Analysis - is also carried out during this phase. The ICA decomposition matrix, derived from the cleaned training data, is preserved and later reused during online execution.

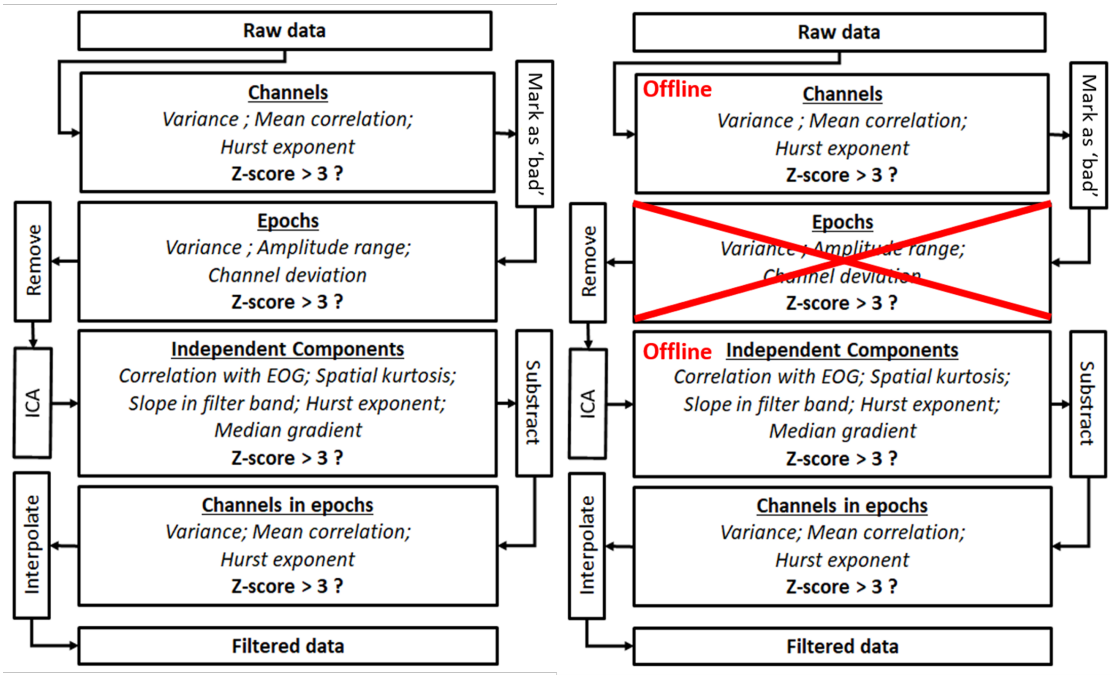

Figure 3.1: The algorithm of the Offline FASTER method (left), with the designed Online FASTER method (right). In the proposed online algorithm, the identification and exclusion of globally poor-quality channels are conducted during the offline prefiltering phase, which precedes the online filtering process. Additionally, the ICA matrix required for artifact correction is also computed in this initial stage. It is important to note that, unlike the offline version, the online method does not involve the removal of bad epochs. All subsequent filtering and classification steps proceed in the same manner as in the offline implementation.

Once these preparatory steps are complete, the algorithm is ready to be deployed in an online setting. The first alteration from the original FASTER pipeline involves the treatment of globally bad channels: in the online version, these are no longer recalculated but are instead directly taken from the pretraining phase, thereby eliminating the

need for repeated computation. The most critical modification, however, concerns the ICA component. In the offline FASTER pipeline, ICA decomposition is computationally intensive and thus unsuitable for real-time use. In the online variant, this step is rescheduled: the precomputed ICA transformation matrix is applied directly to incoming EEG data segments, enabling real-time decomposition without the need for re-estimation.

Each resulting independent component is then evaluated based on several artifact-related criteria. These include statistical measures such as the Z-score of its correlation with electrooculographic (EOG) signals, spatial kurtosis, the slope within a defined frequency band, the Hurst exponent, and the median gradient. If any of these parameters exceed predefined thresholds - typically a Z-score greater than 3 - the corresponding component is identified as artifactual and is removed from the signal.

After the ICA step of artifact rejection, the online FASTER pipeline proceeds in a manner consistent with its offline counterpart. Channels marked as bad in the current epoch are removed and subsequently interpolated using data from the surrounding good channels and the globally bad channels identified during pretraining. Finally, a Common Average Reference (CAR) transformation is applied to enhance spatial signal consistency. The output of this process is an artifact-cleaned EEG segment for each incoming epoch, suitable for further real-time analysis or classification.

The implementation of the proposed algorithm can be found in the designated subfolder of the project's GitLab repository:
`https://dev.itk.ppke.hu/adoan/dissertation-codebase/-/tree/main/bionic_apps/artifact_filtering/`.

3.4. The usage of the Online FASTER algorithm

At the outset of each weekly session, a 20-minute guided EEG recording was conducted immediately after headset placement. The first several minutes served to establish a baseline: activity was recorded first with the participant's eyes open and subsequently with eyes closed. These baseline segments provided the data required to estimate the initial parameter set for the Online FASTER artifact rejection filter.

The task phase that followed was cued visually. Over the course of the project, the cueing paradigm evolved considerably. During preparation for Cybathlon 2020, we evaluated two alternative control schemes: (i) a four-class protocol in which each game command corresponded to a distinct mental task, and (ii) a binary protocol comprising one active and one passive task. For Cybathlon 2024, the paradigm was streamlined to

two active mental tasks plus a resting state. In the current implementation, the two active commands are displayed on a screen; at each trial, an arrow highlights one of them, which the pilot must actively imagine. Real-time feedback is provided throughout the task: after online filtering with the FASTER parameters derived from the baseline, the incoming EEG is classified by a pre-trained neural network, and the predicted class is shown to the pilot. Although the new recordings enlarge our dataset, they are used only to refine the session-specific FASTER parameters; the classifier itself is not retrained online. Instead, the best-performing network from prior offline training is fixed for the game session. The flow of control of the BCI game can be observed in Figure 3.2. During gameplay, the continuous EEG stream is segmented into 1-second epochs. Each epoch is passed through the Online FASTER filter and then delivered to the neural network, which assigns the segment to one of the predefined classes. The resulting label is translated immediately into the corresponding in-game command, enabling real-time BCI control. The average processing time for the online FASTER method was 0.123 ± 0.017 s, whereas the EEGNet neural network required an additional 0.061 ± 0.010 s on average. Considering the 1-second EEG segment used for classification, the total latency remains within the optimal time window of 1–2 s, which provides a balance between classification performance and system responsiveness, as suggested by Miladinović et al. [83].

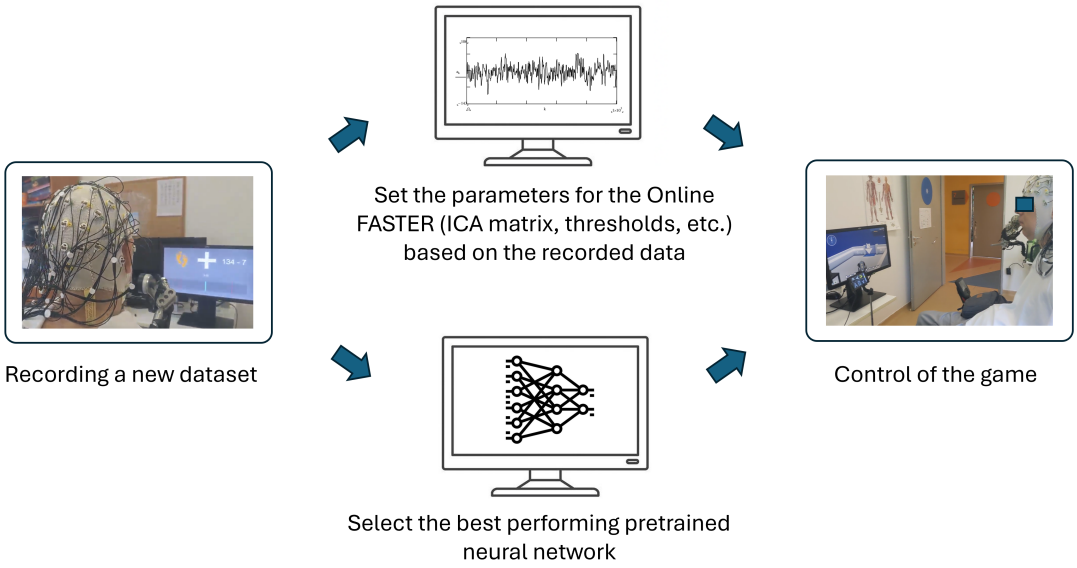

Figure 3.2: Workflow of the online EEG artifact rejection and control system. The recorded data are used to set the parameters of the Online FASTER method and to select the best pretrained neural network for real-time game control.

Chapter 4

The effect of preprocessing steps on the classification accuracy of BCI systems

In this chapter, five neural network architectures are evaluated using data from 105 subjects of the Physionet database. The primary objective is to investigate how various preprocessing techniques - including artifact rejection, transfer learning, cropped training, and frequency filtering - affect classification accuracy across these architectures. The following two subsections first provide a review of the existing literature regarding the impact of preprocessing methods, with particular emphasis on artifact rejection and frequency filtering. Subsequently, the examined neural network models are introduced, and the applied methodologies are described in detail. Finally, the results of this exploratory analysis are presented and summarized.

The complete codebase supporting this research is publicly available at `https://dev.itk.ppke.hu/adoan/dissertation-codebase`. The EEG data used in this study were obtained from the PhysioNet EEG Motor Movement/Imagery dataset, accessible at `https://physionet.org/content/eegmmidb/1.0.0/`.

4.1. Effect of artifact rejection on the classification of EEG signals – literature review

The examined studies concerning the effects of artifact rejection can be found in this section.

Junjie Zhong and Feifei Qi [84] investigated the influence of artifact rejection on the performance of a P300-based BCI in their study. The artifact rejection was achieved through the utilization of ICA, and the classification was performed using Bayesian Linear Discriminant Analysis (BLDA). Their classification results underscored that artifact rejection methodologies can indeed enhance BCI performance by augmenting both accuracy and information transfer rate. However, the study was limited by the use of manual component filtering and a small sample size of only 8 participants.

Minju Kim et al. [85] also examined the effect of various AR methods (such as ICA, adaptive filtering, and Artifact Subspace Reconstruction) on the classification accuracy of the P300 and N200 components. As their results showed, the mean accuracy of ERPs denoised by ICA, which was $62.87\% \pm 8.81\%$, was higher than those by other algorithms or no rejection, using an SVM as a classifier.

M. Mohammadi and M. R. Mosavi [86] compared 2 ICA-based EOG artifact rejection methods with each other and examined the classification accuracy using Linear Discriminant Analysis. In their study, the ICA with wavelet decomposition proved to be the better method, and their comparison yielded better accuracies in the BCI Competition IV 2a [24] dataset than those in the work of Ang et al. [87] and the work of Gouy-Pailler et al. [88]. This work, however, did not aim to compare the raw and artifact-rejected classification accuracies and only examined a small dataset.

Winkler et al. [30] presented an ICA-based AR method, specifically for BCI usage. They used multiple databases to test their system and utilized shrinkage-regularized linear classifiers to evaluate BCI performance. Regarding their result, the AR method had a low effect on BCI classification accuracy in two databases, such as ERP BCI [89] and MI-BCI [90]. They observed the highest value of performance changes in the MI-BCI dataset using lateralized readiness potential as a feature, while the average accuracy did not change significantly; subject-wise, greater disturbances were observable.

Asif Iqbal et al. [91] used the BCI Competition IV 2a. [24] collection of EEG data to test their novel EEGNet and Temporal Convolution Network (TCN)-based classifier and their ocular artifact removal system based on linear regression. The study demonstrated

that their neural network architecture, coupled with an EOG artifact removal system, led to a notable enhancement in BCI classification accuracy, achieving an impressive 80.5% accuracy rate. This represents a 4% improvement compared to the performance of the EEG-TCNet classifier, upon which their network architecture was based. Additionally, the researchers observed a notable reduction in the variance of classification accuracy across test subjects following the removal of EOG artifacts from EEG data.

E. Bou Assi et al. [92] employed their devised method, which integrates ICA with K-means clustering, for the purpose of artifact rejection. In their study, physiologically interpretable features were extracted using methods such as wavelet coherence, wavelet-phase locking value, and band power analysis. These features were then subjected to a statistical test to ascertain significant differences between relaxed and motor imagery states. Features that successfully passed the statistical test were retained and subsequently utilized for classification purposes. They utilized SVM and LDA as classifiers, and they observed that removing artifactual components increases the 2-class classification accuracy from 66% to 88.1% in the Physionet database.

David E. Thompson et al. [93] found that automated artifact rejection methods caused a significant performance decline in the use of P300-based BCIs. They used the simple least squares classification method for the detection of the P3 signal and tested 10 AR algorithms. Their study revealed that methods titled SOBI, JADER, and EFICA induced the least amount of performance decline; however, even with these methods, there was a discernible (up to 10%) decrease in accuracy.

Laura Frølich et al. [94] investigated the effects of different artifact types, such as blinks, lateral eye movements, heartbeat artifacts, muscle artifacts, and mixed artifacts on motor imagery BCI. They used their own developed classifier 'IC_MARC' to label independent components as a type of artifact. The research was conducted on a dataset of 80 subjects with a classical motor-imagery paradigm of 3 classes. They found that if all the 119 EEG channels are used, muscle artifacts adversely affect BCI performance while using only the 48 centrally located channels, artifact effects were reduced.

Mannan et al. [95] conducted a comprehensive investigation to examine the impact of EOG-based artifact rejection on classification accuracy. Their research determined that employing a low-pass filter on the EOG signal with a cutoff frequency of 7 Hz yielded optimal results. Moreover, they demonstrated that the integration of this method with the examined 5 artifact filtering approaches led to notable improvements in classification accuracy when utilizing the CSP classifier on the BCI Competition IV 2a dataset.

Ian Daly et al. [4] presented the Fully Online and Automated Artifact Removal for Brain-Computer Interfacing (FORCe) algorithm for artifact rejection. They made a comparison of the classification accuracy, comparing them on the raw signal and on using the Lagged Auto-Mutual Information Clustering (LAMIC), the FASTER algorithm [6], or their method. They used a small database of 14 subjects with cerebral palsy and revealed significant differences between the original EEG and the FORCe method and between the FORCe method and FASTER.

Pavel Merinov et al. [96] compared 4 AR methods with two benchmark solutions, using the measurements of the BCI Competition IV 2a dataset and CSP as a classifier. They concluded that there was no significant difference between the accuracies of the examined methods. Their research also showed a high degree of subject dependence on classification performance.

Stigt et al. [97] performed their experiments on the TUH Abnormal EEG Corpus dataset, which contains normal and pathological EEG data. First, they designed a CNN to classify artifactual data, and in the second step, they examined whether the rejection of artifacts improves the classification accuracy of the used Shallow ConvNet. They found that artifact rejection did not improve abnormal vs normal EEG classification performance, but it made the training of the network slightly faster.

Yao Chen et al. [98] examined the influence of certain types of artifacts on classification accuracy, such as EMG, EOG, power-line interference, and electrode artifacts. They used ICA to separate the different artifactual components and performed their experiments on a database consisting of 6 subjects' data with 26 scalp electrodes. Their experimental findings indicate that, in terms of artifact types, the order of influence on classification accuracy was as follows: EMG, power-line interference, electrode artifact, and EOG. Specifically, they observed relative increments in accuracy of 5.5%, 4.0%, 3.1%, and 1.7%, respectively, when these artifacts were effectively removed from the EEG signals.

Md Kauf Islam et al. [99] proposed a probability mapping-based method for EEG artifact removal from EEG signals. They used the BCI Competition IV 1, 2a, and 2b datasets to apply their system. They found that removing the artifacts with their method enhances the outputs of the used LDA BCI classifier with windowed means feature, based on receiver operating characteristic curves and an error rate diagram.

Maisha Anjum et al. [100] also researched the difference in classification accuracy by using or neglecting a probability mapping-based artifact detection method. They showed

in their 12-subject database, recorded by the 14-channel Emotiv Pro headset, that AR can increase performance by 16% on average. However, because of the limited number of electrodes used, their original results were not conclusive.

4.2. Effect of Frequency filtering on the classification of EEG signals - literature review

In the pursuit of classifying motor imagery EEG signals, several studies have consistently employed frequency filtering techniques, often focusing on restricting the frequency domain to the mu and beta bands. By narrowing the analysis to these specific frequency ranges, typically associated with motor-related brain activity, researchers aim to extract discriminative features that encode the neural signatures of motor intention and execution. In a comprehensive review conducted by Ali Al-Saegh et al. [101], the authors systematically examined the frequency ranges employed in the reviewed 36 papers. A notable consensus emerged across these studies, with the frequency range from 8 to 25 Hz being consistently included. Remarkably, this frequency range was observed in all analyzed papers, underscoring its widespread adoption and recognition as relevant for motor imagery classification in EEG signals. Conversely, the frequency band from 0 to 5 Hz was notably absent in the majority of the surveyed literature, utilized in only 7 out of 36 studies. This observation highlights a tendency to prioritize higher-frequency bands over lower-frequency ones in the context of motor imagery EEG signal analysis, suggesting a prevailing emphasis on spectral components associated with motor-related brain activity.

In the article of R. Salazar Valas and Roberto A. Vazquez [102], the authors undertook an evaluation of the impact of cutoff frequencies during the pre-processing stage, employing experiments conducted on the BCI Competition IV 2a dataset. Within this framework, three distinct feature extraction methods were employed: a second-order autoregressive model and fractal dimension calculations. Their investigation involved varying bandwidths and lower cutoff frequencies, with subsequent testing of classification accuracy using a linear discriminant classifier. Their work shows that even the frequency range of 0 to 10 Hz could carry significant information, as there were cases where this range also showed accurate classification accuracy.

4.3. Materials and methods

4.3.1 Artifact rejection

In this section I used the offline version of the FASTER algorithm, which is detailed in Chapter 3. While the original article specified a frequency range from 1 to 95 Hz with a notch filter at 50 Hz, my study deviated by utilizing distinct frequency ranges tailored to my specific experimental parameters and objectives. The exact frequencies are detailed in Subsection 4.3.5.2.

4.3.2 Networks for MI signal classification

In this chapter, the two well-known and widely utilized neural networks are described, which I also used for examining the effects of different processing steps. Both of these networks are convolution-based systems and perform well in EEG classification tasks.

4.3.2.1 EEGNet

EEGNet, designed by Lawhern et al. [64], is a compact CNN architecture especially constructed for EEG-based BCIs. It can be applied across different BCI paradigms, it is optimal for training with very limited data, and it has been shown that it generates features that are neurophysiologically interpretable. It consists of three blocks of multiple types of convolution and a final classification layer. In the first block, two convolutional steps are performed in sequence. Initially, a 2D convolution is performed with a kernel size of (1,64), essentially meaning 1D convolution for each channel. The dimension of the kernel is designed to be half of the sampling rate of the data, enabling the system to extract frequency features at 2Hz and above. The number of kernels is F_1, therefore, the output of this layer consists of F_1 feature maps containing the EEG signal in different band-pass frequencies. The second convolutional step is the so-called depthwise convolution, in which the kernel dimension is the number of channels times 1. This operation is for performing spatial filtering, and in this type of convolution, each input channel is convolved with a different kernel. The advantage of depthwise convolution is reducing the number of parameters to fit, as these convolutions are not fully connected to all previous feature maps. In the case of EEG processing, it means that there is a direct way to learn separate spatial filters for each frequency-filtered input, resulting in frequency-specific spatial filtering. This two-stage convolution was inspired by the FBCSP algorithm. Both convolutions are kept linear, and Batch Normalization and ELU nonlinearity are applied

as well. During training, a dropout with a 0.5 probability was applied. In the following block, separable convolution is performed. The structure is depicted in Figure 4.1.

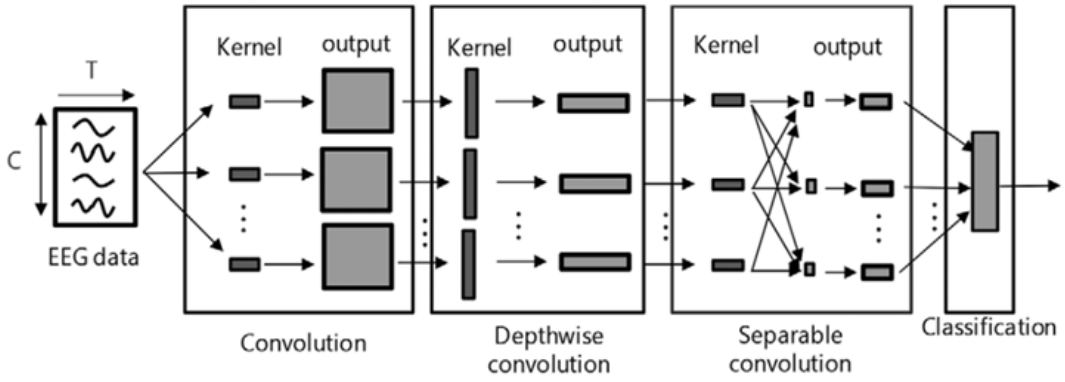

Figure 4.1: The structure of the EEGNet. Raw EEG data undergoes temporal, depthwise, and separable convolution, while a fully connected layer performs the final classification [103].

4.3.2.2 Shallow ConvNet

The next architecture employed in this work is based on the Shallow convolutional neural network (Shallow ConvNet) proposed by Schirrmeister et al. [63], which was specifically designed for the classification of EEG signals. This model comprises two convolutional layers, each containing 40 filters, followed by pooling, flattening, and fully connected layers, culminating in a softmax output layer. The first convolutional layer performs temporal filtering by convolving along the time axis independently for each EEG channel. This operation is analogous to applying a set of linear finite impulse response (FIR) filters to each channel, enabling the network to extract relevant temporal features from the raw signal. Importantly, no spatial mixing occurs at this stage; the operation is channel-wise and captures frequency-specific characteristics of the EEG. The second convolutional layer performs spatial filtering by convolving along the channel dimension at each time step. This layer receives the temporally filtered signals and applies spatial filters that aggregate information across all EEG channels. Due to the use of valid padding and the kernel size being equal to the number of EEG channels, each spatial filter produces a single scalar output at each time point by computing a weighted sum of all channels. This effectively reduces the multi-channel input to a single-channel time series per filter, allowing the network to exploit inter-channel spatial dependencies in a data-driven manner. Following the convolutional stages, the resulting time series are subjected to average pooling, which reduces temporal resolution while retaining the most

salient temporal features. The pooled outputs are then flattened into one-dimensional feature vectors and passed through a fully connected (dense) layer that integrates the learned spatial-temporal features. The final layer of the network is a softmax activation function, whose number of output units corresponds to the number of target classes in the classification task. This architecture, while relatively simple, has been demonstrated to be highly effective for EEG-based classification tasks due to its ability to learn interpretable spatial and temporal features directly from raw EEG data with minimal preprocessing.

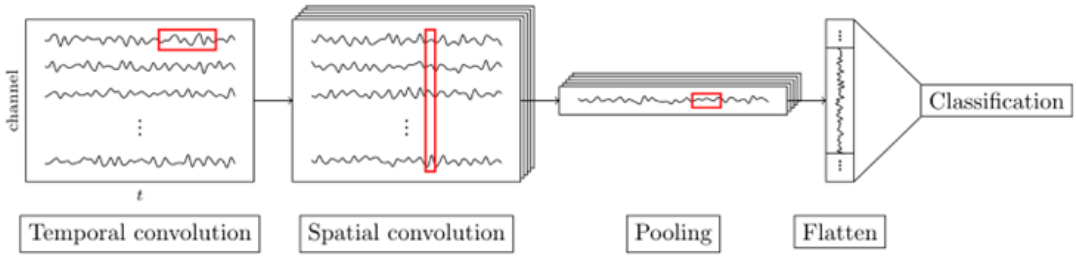

Figure 4.2: The architecture of the Shallow ConvNet classifier. On the raw data, temporal and spatial convolutions are applied, and finally, a fully connected network performs the classification [104].

4.3.3 The 3D representation of the EEG signal

Considering that the spatial configuration of the electrodes can provide informative cues for classification, the EEG signals were represented as a three-dimensional tensor, consisting of two spatial dimensions corresponding to the 2D projection of the electrode layout and one temporal dimension capturing the signal time course, as applied in previous studies [105], [106], [107]. For this transformation, I used a unique arrangement of electrode placements, referred to as dense 3D transformation. In this construction, the Iz electrode is omitted, and the remaining 63 electrodes are rearranged into a 9x7-dimensional rectangle. The arrangement can be seen in Figure 4.3. This arrangement will be the input shape for 2D and 3D convolutional neural networks.

4.3.4 Networks with the 3D representation as the input

4.3.4.1 Conv2D Net

My first proposed network processes the raw EEG tensor described in the previous subsection using two-dimensional convolutional layers whose kernels extend across all time steps. Concretely, each kernel has dimensionality (kx, ky, T), where T equals the number of samples in the temporal dimension. Consequently, the convolution is executed

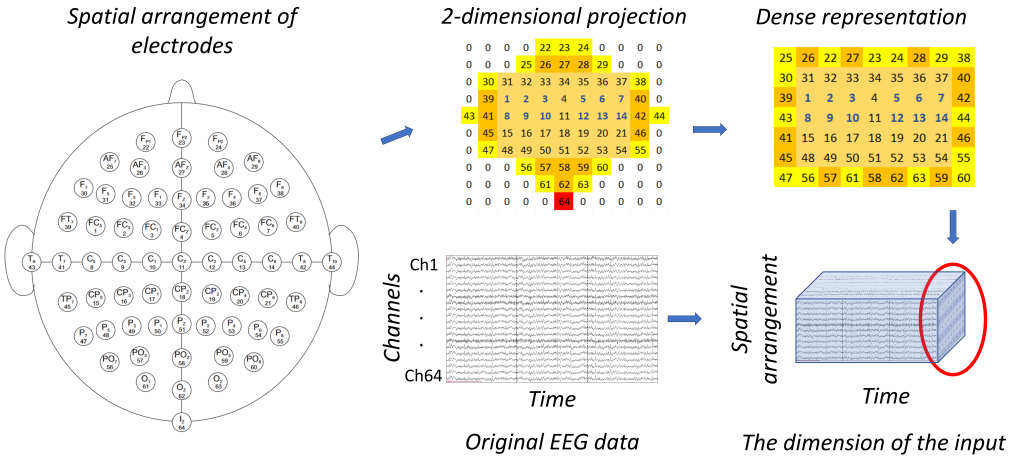

Figure 4.3: 3D representation of the EEG data: the channels x time initial arrangement of the EEG data is rearranged into a 3-dimensional form, where the first two dimensions refer to the spatial arrangement of the electrodes, while the third dimension is time. For the spatial dimension, I used a dense electrode arrangement: while the original distribution contains placeholder zero channels, I rearranged this map to have a dense representation with no such cells [J2].

over the spatial layout of the electrodes, while a weighted summation is applied along the full temporal axis, thereby integrating information from every time point within a single filter response. Prior to convolution, the input undergoes L2-normalisation to ensure scale invariance across channels. The normalised data are then passed through three successive 2D convolutional layers, each followed by a nonlinear activation and, where appropriate, batch normalisation. The output feature maps are subsequently flattened and fed into a two-layer fully connected (dense) sub-network, which serves to synthesize the learned spatial–temporal features. Finally, a softmax output layer produces class membership probabilities, enabling multiclass classification. A schematic representation of the complete architecture is provided in Figure 4.4.

4.3.4.2 Conv3D Net

Neural networks utilizing 3D convolutional layers are commonly applied in video processing tasks [108], and their potential for EEG signal classification has also been demonstrated in several studies [107], [109]. The key distinction between this architecture and conventional 2D convolutional structures lies in the additional convolution operation along the third dimension, which yields a four-dimensional feature representation after

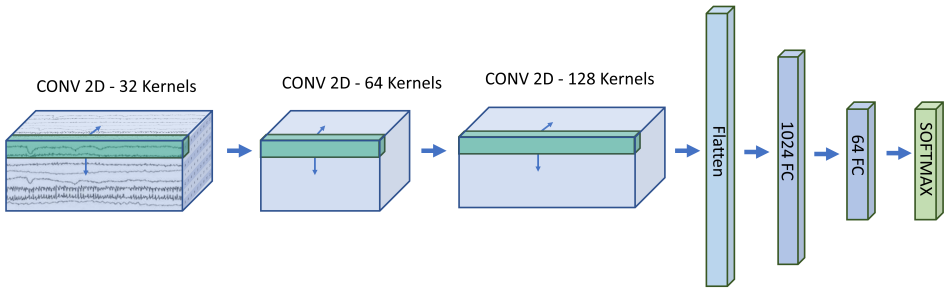

Figure 4.4: The structure of the designed 2D convolutional neural network: The input is the 3D representation of the EEG signal, on which 2D convolution is performed with 32 kernels of the size of [3 x 3 x number of timepoints] in the first layer. In the next two layers of convolution, firstly 64, then 128 kernels of dimension [3 x 3 x number of kernels of the previous layer] are used. Next, the flattened representation is given to two layers of a fully connected network with 1024 and 64 neurons, and finally, a softmax layer is responsible for classification with the output of 4 numbers, as the number of classes [J2].

the second convolutional layer. Although this approach results in increased memory consumption, it offers the advantage of capturing temporal dependencies and features distributed over time.

The architecture I designed begins with L2 normalization, followed by three layers of 3D convolution. The kernel size in the first layer is set to [1, 1, 30], allowing temporal feature extraction without spatial interaction, while the subsequent layers use kernels of size [2, 2, 40] to incorporate both spatial and temporal dimensions. Each convolutional layer employs 32 kernels. Between convolutional operations, batch normalization, and Exponential Linear Unit (ELU) activation functions are applied to improve training stability and non-linearity. Following the convolutional blocks, the feature maps are passed through two fully connected layers, and finally, a softmax layer is used to perform classification. The architecture of this network is illustrated in Figure 4.5.

4.3.4.3 Multi Branch Conv3D Network

The second type of examined 3D CNNs is based on the article of Xinqiao Zhao et al. [105]. It consists of three branches, all of them with two layers of different dimensions of convolutional kernels and three layers of a fully connected network. The last layers are the size of the number of classes; thereafter, the outputs of all three branches are added up, and a softmax operation is performed for the final classification. The detailed structure can be observed in Figure 4.6.

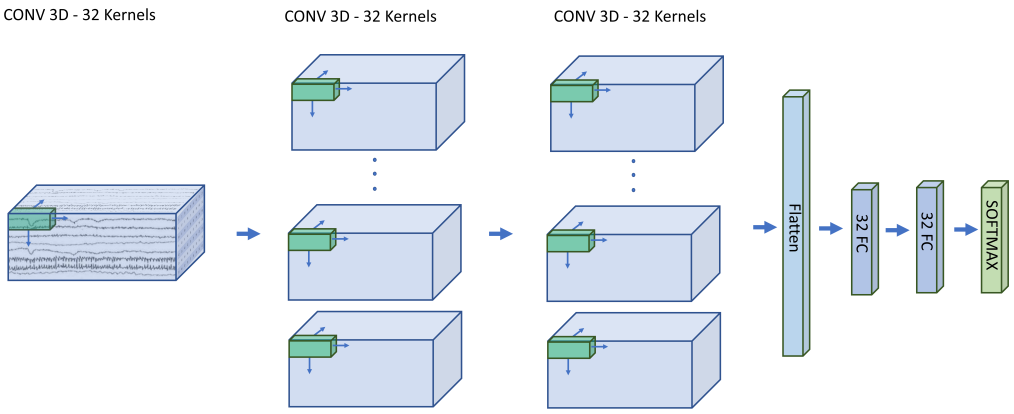

Figure 4.5: The structure of the 3D convolutional network. The input is identical to that of the Conv2D Net; the difference is that in this structure, 3-dimensional convolution is performed. There are 32 kernels in all the layers, the shape is (1 x 1 x 30) in the first, and (2 x 2 x 40) in the second and third layers. The fully connected part consists of two layers with 32 neurons each, and finally a softmax classification layer [J2].

4.3.5 The examined processing techniques

4.3.5.1 Transfer Learning and Finetuning

A deep learning system usually requires a great amount of data to generalize features well and to have reasonable accuracy. Systems for EEG classification only acquire a limited number of samples, as recording and labeling these signals is cumbersome, requires a significant amount of time to process, and requires human intervention. The main idea of transfer learning (TL) application is pre-training a system, or a part of the system, over an independent dataset and transferring these weights as an initial state. These weights are fine-tuned during the actual training phase of the network. In the case of EEG studies, two main types of transfer learning are used [110]. One of the TL approaches is when the feature space generated from the EEG data is similar to one of those tasks, for which a way larger dataset is available. An example is ImageNet as an initial system for classifying EEG samples transformed into images [111]. However, although it represents a promising research direction, this form of transfer learning was not investigated in the present work.

The second approach - the one I actually used in my research - involves using EEG data as the base for pretraining as well, but these data originate from subjects other

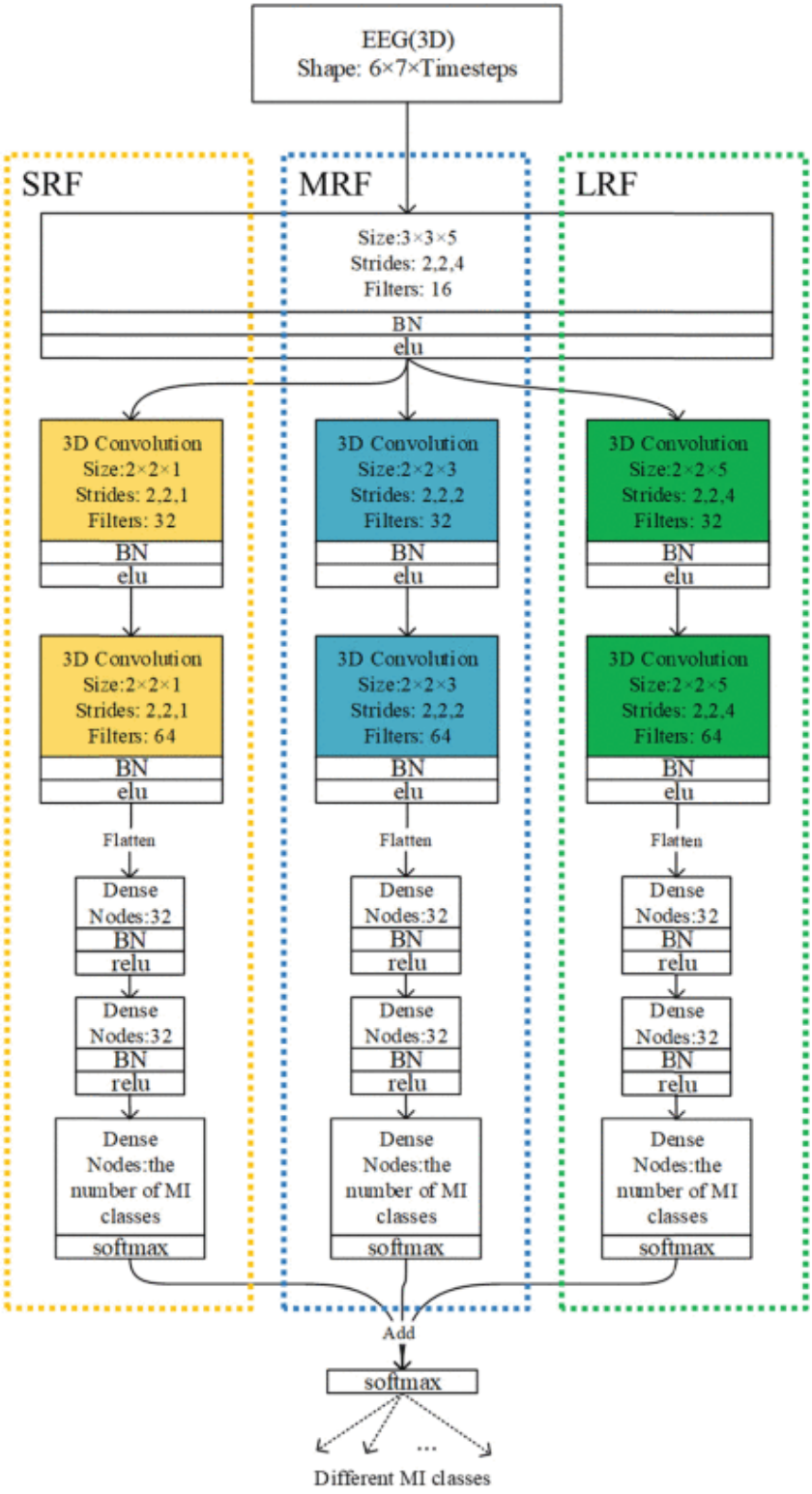

Figure 4.6: The structure of the implemented Multi Branch Conv3D Net, designed by Xinquiao Zhao et al. [105].

than the target individual. In other words, networks pretrained on multiple subjects are then fine-tuned for each individual. Multiple studies used this technique, and it yielded significant improvements in classification accuracy [Au1], [104], [112].

In my research, transfer learning was performed as follows: I iteratively split the 105 subjects in the Physionet database into two groups. The network was pretrained on 90% of the subjects (using 20% of this set for validation), and then fine-tuned and tested on the remaining 10% of the subjects individually. The exact method for training can be observed in Figure 4.7.

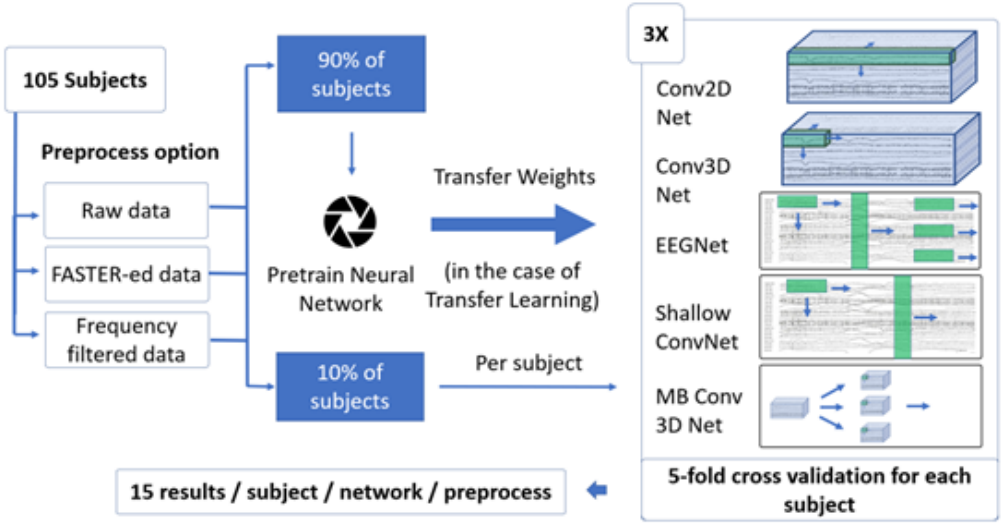

Figure 4.7: Transfer learning and the generation of results. To have statistically significant results, I performed a 5-fold cross-validation for each subject 3 times, resulting in 15 results for each subject, network, and preprocess option. 90% of the subjects' data is used for pretraining the classifier, which is fine-tuned over the individual subjects' training set. The whole process iterates 10 times for every subject to be tested [J2].

4.3.5.2 Frequency filtering

In the later stages of the research, I analyzed the impact of frequency filtering. I examined how classification accuracy is influenced when a band-pass filter is applied to the 0.1–5 Hz frequency range, and similarly evaluated the range between 5 and 75 Hz. Additionally, I investigated the results obtained when frequency filtering is entirely omitted. For the filtering steps, a fifth-order Butterworth filter was used. The Butterworth filter, renowned for its characteristic of being maximally flat, ensures a uniform magnitude response within the pass band. However, a noteworthy drawback of this filter is the substantial width of its transition band [113].

4.3.5.3 Cropped training

To investigate the generalizability of EEG signals with varying onset times, I implemented the cropped training method, which was also employed by Schirrmeister et al. [63]. In this approach, the original data are augmented by systematic temporal shifts, thereby expanding the training and testing datasets. This ensures that neural networks are exposed to a wider range of temporal variations within the EEG signals, facilitating the simulation of different onset times and improving the generalizability of the models.

In the case of my research, I used 1-second-long windows instead of the previous 2-second ones, and performed shifts in 0.1-second steps up to 2 seconds. This resulted in 11 overlapping crops from a single epoch. All samples originating from a given epoch were assigned exclusively to either the training or the test set - no mixing occurred.

The previously described experiment regarding frequency dependence was repeated on this augmented dataset across all networks, both with and without the application of the FASTER method.

4.4. Results

4.4.1 Effect of Artifact Rejection

As can be observed in Figure 4.8 and Table 4.1, without using the TL method, the average of the classification accuracy of Conv 2D, Conv 3D, Shallow ConvNet, and Multi Branch Conv3D Net models are significantly improved due to the FASTER algorithm, while regarding the EEGNet there was no significant improvement. In this scenario, I used 2s-long windows, and during the FASTER algorithm, a frequency filter between 0.1 and 75 Hz was applied. For the raw data, I did not use any frequency filtering.

Upon subject-specific scrutiny, it becomes evident that the influence of artifact rejection is contingent on the particular subject under evaluation. My analysis encompassed a comprehensive approach. Initially, I executed 5-fold cross-validation three times for all four networks, both with and without artifact rejection. To substantiate disparities, I scrutinized the distribution of 15 results obtained for an individual subject with a specific network. If the data followed a normal distribution, I executed a T-test; conversely, if non-normality was detected, a Mann-Whitney U test was performed to ascertain the significance of the observed differences.

I categorized subjects based on the extent of change in the corresponding classification performance across the various networks. Intriguingly, several scenarios arose in which

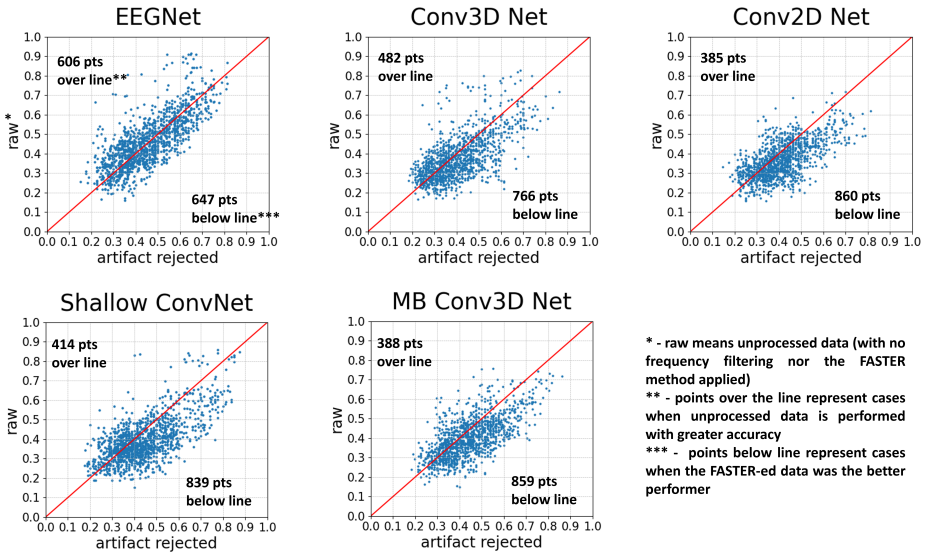

Figure 4.8: The comparison of neural networks with and without using the FASTER method. Each point represents two accuracies for a subject: the one obtained with AR and the other obtained without using the FASTER algorithm. The test was run 12 times to get significant results. The red line indicates the points with no difference between the two options; points over the line run with the raw option as the more accurate, and points below the line where the artifact-rejected version yields better results [J2].

Table 4.1: The accuracy of each classifier with and without artifact rejection, and the P-value of the significance of differences using the Wilcoxon signed-rank test.

Classifier	Original Acc.	AR Acc.	P value
EEGNet	0.460	0.455	0.907
Shallow ConvNet	0.394	0.439	3.46E-17
Conv2D Net	0.367	0.411	4.90E-21
Conv3D Net	0.378	0.405	4.09E-09
Multi Branch Conv3D Net	0.401	0.444	2.05E-19

certain networks led to a notable enhancement, while others yielded a significant decline in the performance for the same subjects. In response to a slight variance in results observed during a second examination, I iteratively conducted the calculations two more times to explore the evolving significance of the observed differences. Finally, I had 4 times 3 accuracy results for each cross-fold iteration and network, meaning four results of significance for each classifier.

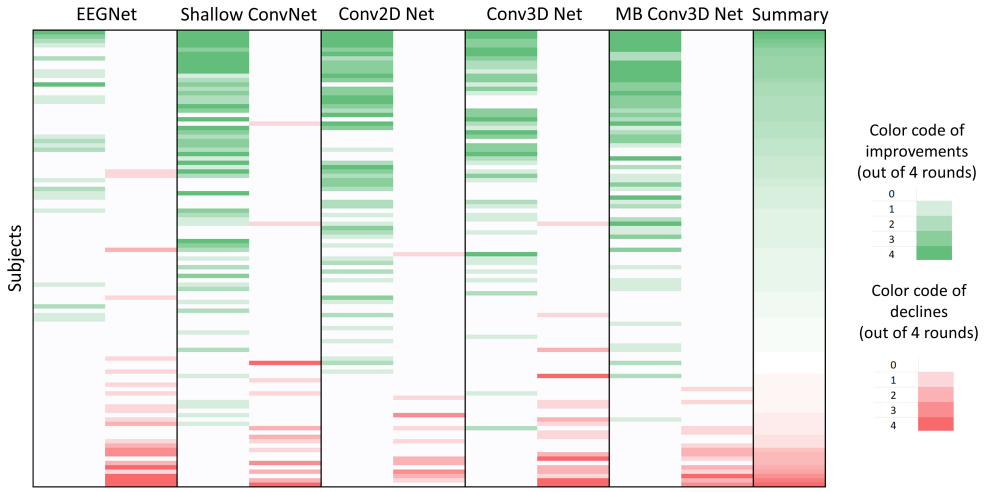

Figure 4.9: Subject dependence of artifact filtering. The shade of color indicates how many runs there were statistically significant improvements (green) or significant declines (red). (The deeper the shade, the more runs were significant.) Significance is based on the T-test in the case of normality and the Mann-Whitney U test in the case of non-normality. Subjects are ordered by summary number, computed as the sum of the significance of improvements minus the sum of the significance of decline [J2].

In my assessment, a numerical score was assigned to each subject, denoted as follows: a value of -1 indicated a significant decline, 0 denoted no significant difference, and +1 represented a discernible increase observed for each computational aspect across all networks. Therefore, the cumulative score per subject ranged from -20 to 20. Twenty-five subjects scored over 8, indicating substantial performance gains from the AR method, with Subject 69 achieving a remarkable 20-point increase. Conversely, some subjects, like Subject 15, experienced significant declines, with only five subjects scoring below -8. Regarding classifiers, EEGNet showed the least improvement, with only three subjects scoring at least 3, while six scored below -3. In contrast, the other four networks had

at least 20 subjects exceeding the three-point mark, with fewer than five scoring -3 or worse. The subject dependence on AR is illustrated in Figure 4.9.

4.4.2 The Effect of Transfer Learning

The classification accuracy obtained by transfer learning was significantly better in every scenario. Based on the signed-rank Wilcoxon test, the learning process performed significantly better in the case of each network, both in the case of unfiltered data (Table 4.2) and artifact-rejected data (Table 4.3). In Figure 4.10, it can be observed that transfer learning does not improve as much in cases of artifact-rejected data as in the case of raw data, resulting in higher classification accuracies in the latter case. (This difference is significant in all the cases except the Multi Branch Conv3D Network.)

Table 4.2: The accuracy of each classifier on unfiltered data, with and without transfer learning, and the P value of significance using the Wilcoxon test.

Classifier	Simple Acc.	TL Acc.	P value
EEGNet	0.461	0.587	1.50E-18
Shallow ConvNet	0.394	0.637	5.83E-19
Conv2D Net	0.366	0.528	7.78E-19
Conv3D Net	0.378	0.56	7.14E-19
Multi Branch Conv3D Net	0.401	0.561	1.30E-18

Table 4.3: The accuracy of each classifier on artifact-rejected data, with and without transfer learning, and the P value of significance using the Wilcoxon test.

Classifier	Simple Acc.	TL Acc.	P value
EEGNet	0.455	0.538	1.67E-17
Shallow ConvNet	0.441	0.559	1.38E-18
Conv2D Net	0.410	0.491	7.20E-17
Conv3D Net	0.405	0.521	7.14E-19
Multi Branch Conv3D Net	0.444	0.557	5.52E-18

4.4.3 Comparison of Neural Networks

In network comparisons, the primary evaluative criterion centers on classification accuracy. As illustrated in Figure 4.11, when transfer learning is not employed, the

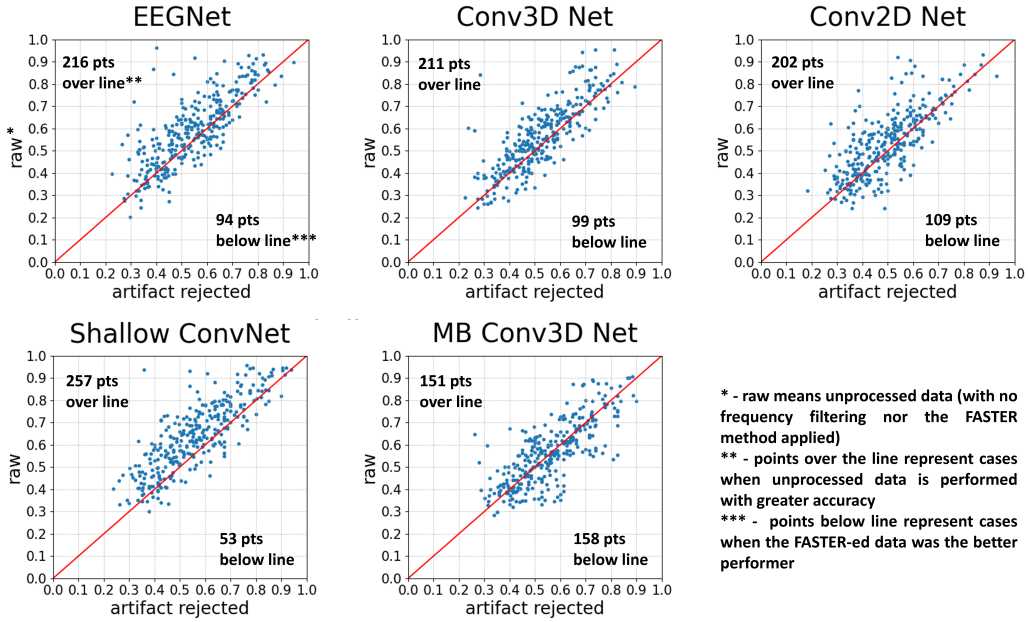

Figure 4.10: The comparison of neural networks with and without using the FASTER method, using transfer learning. Each point represents two accuracies for a subject: the one obtained with AR and the other obtained without using the FASTER algorithm. The test was run 3 times. The red line indicates the points with no difference between the two options, points over the line run with the raw option as the more accurate, and points below the line where the artifact-rejected version yields better results. Generally, by the usage of transfer learning, accuracies without artifact rejection tend to be higher [J2].

highest classification accuracy is achieved by the EEGNet classifier, both in raw and artifact-rejected conditions. However, in the latter case, the differences between the EEGNet, Shallow ConvNet, and MB Conv3D Net were not significant. In the raw data scenario, the Multi Branch Conv3D Net is the second-best performer, followed by the Shallow ConvNet, with my proposed networks trailing.

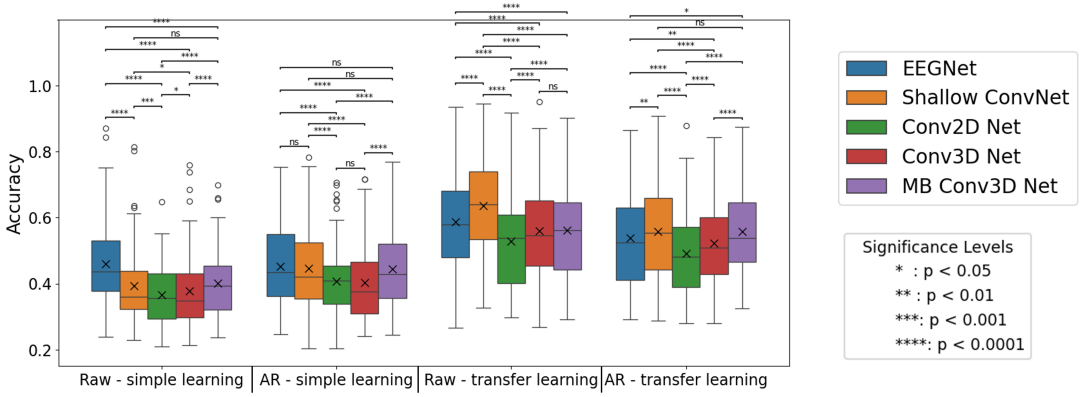

Figure 4.11: The accuracies obtained by neural networks, with and without transfer learning and artifact-rejection. In the box plot, x denotes the mean, while the horizontal line marks the median value. The limits of the box indicate the range of the central 50% of the data. Before applying the TL method, EEG Net was the greatest performer, while after the application, Shallow ConvNet had the highest accuracy. These observations are generally true for the raw and the artifact-rejected data as well. The figure also shows that without transfer learning, networks generally perform better with the FASTER method, while after the application, raw data will get a higher precision [J2].

However, the performance landscape shifts with the introduction of TL. Shallow ConvNet emerges as the top performer, surpassing EEGNet and other classifiers. In the raw data scenario with TL, EEGNet ranks second, followed by the 3D convolutional networks. In the artifact-rejected condition with TL, the MB Conv3D Net outperforms EEGNet, securing the shared first position, with non-significant differences from the Shallow ConvNet. This demonstrates that TL significantly influences the performance hierarchy among different network architectures.

4.4.4 The Effect of Input Representation

Using the dense 3D representation for input did not yield significant improvements in classification performance. This indicates that incorporating spatial information did not lead to higher accuracies compared to networks without this additional spatial data.

The only exception is with the artifact-rejected data using transfer learning, where the MB Conv3D Net slightly outperformed the EEGNet.

4.4.5 The Effect of Frequency Filtering

The outcomes stemming from the frequency filtering analysis have yielded unexpected findings. I compared results using the 0.1 to 5 Hz range and the 5 to 75 Hz range. As observed in Table 4.4, three out of the five neural networks yielded significant results only within the first frequency range (0.1–5 Hz). Shallow ConvNet and EEGNet are the only networks for which the 5 to 75 Hz range also provides results greater than the chance level. However, for EEGNet, these results were still far below the accuracy achieved in the lower frequency range.

Table 4.4: Results of comparing the classification accuracies of neural networks when filters of certain frequency ranges are applied.

Frequency range	EEGNet	Shallow ConvNet	Conv2D Net	Conv3D Net	MB Conv3D Net
0.1-5 Hz – Raw	**0.482**	0.406	0.437	0.410	0.457
5-75 Hz – Raw	0.315	**0.362**	0.262	0.258	0.271
0.1-5 Hz – AR	0.452	0.405	0.421	0.436	**0.462**
5-75 Hz – AR	0.324	**0.381**	0.261	0.261	0.289

4.4.6 The Effect of Simple and Cropped Training

The results derived from simple and cropped learning can be observed in Figure 4.12. Upon comparing results from the simple learning process, we can state that those experiments, when 5 Hz to 75 Hz frequency filtering was performed, had the lowest accuracies. Only the Shallow ConvNet attained moderately higher accuracy, as we have observed in the previous paragraph. The relation between AR data (with 0.1 to 75 Hz frequency filter) and raw data also remains the same, as we have seen in the first part of the results: classification accuracy is enhanced by the application of artifact rejection in the case of four out of five networks, with EEGNet as the only exception. When I apply the mentioned frequency filter without the FASTER algorithm, the case of Shallow ConvNet becomes similar to the EEGNet, in the sense of being the non-artifact-rejected option, the better performer. The partly unexpected finding is the superior performance

of cases where frequency filtering between 0.1 and 5 Hz was applied. These signals include only the lowest part of the frequency spectrum, meaning that they do not contain the mu or beta band.

When applying cropped training, I acquired noteworthy results, depending on the network I used for classification. Contrastingly to the simple learning approach, in the specific context of the Shallow ConvNet and the EEGNet, an inverse trend was observed where the highest performances were achieved in the broader frequency range spanning from 5 to 75 Hz. On the other hand, the worst results were attained by the higher frequency range for the other three networks, where the input feature was the 3D representation. These results show that the frequency ranges of the mu, beta, and gamma bands are the most valuable in the case of EEGNet and Shallow ConvNet, while the other three networks are not able to extract those features effectively. It is important to emphasize that for the two widely used networks to achieve better results at higher frequencies, as reported in the literature, cropped training was necessary.

Within the realm of neural networks, the integration of cropped training demonstrates a parallel effect akin to the observed outcomes with transfer learning. In the absence of cropped training, EEGNet emerges as the preeminent performer. However, upon application, there is a discernible transition where Shallow ConvNet surpasses EEGNet in terms of classification prowess.

4.4.7 Subject Dependence in Simple Learning and Cropped Training

I extended my investigation to explore subject dependence within the context of cropped training. This involved conducting a series of experiments, comprising four sets of three trials each, on both the EEGNet and Shallow ConvNet systems. I selected these two networks because only these two classifiers achieved reasonable accuracy on the frequency-filtered data during cropped training. I meticulously assessed subject-wise performance, focusing on the frequency range of 5 Hz to 45 Hz for the FASTER method. As there were 4-4 significance values for both networks, each subject was assigned a point ranging from -8 to 8, facilitating a comparative analysis between FASTER-processed and completely raw data.

Furthermore, alongside these comparisons, I computed accuracies based on frequency-filtered raw data spanning the 5 Hz to 45 Hz range and juxtaposed these against both raw and FASTER-processed data. These findings are synthesized and presented in Figure 4.13, providing valuable insights into the relative efficacy of different preprocess-

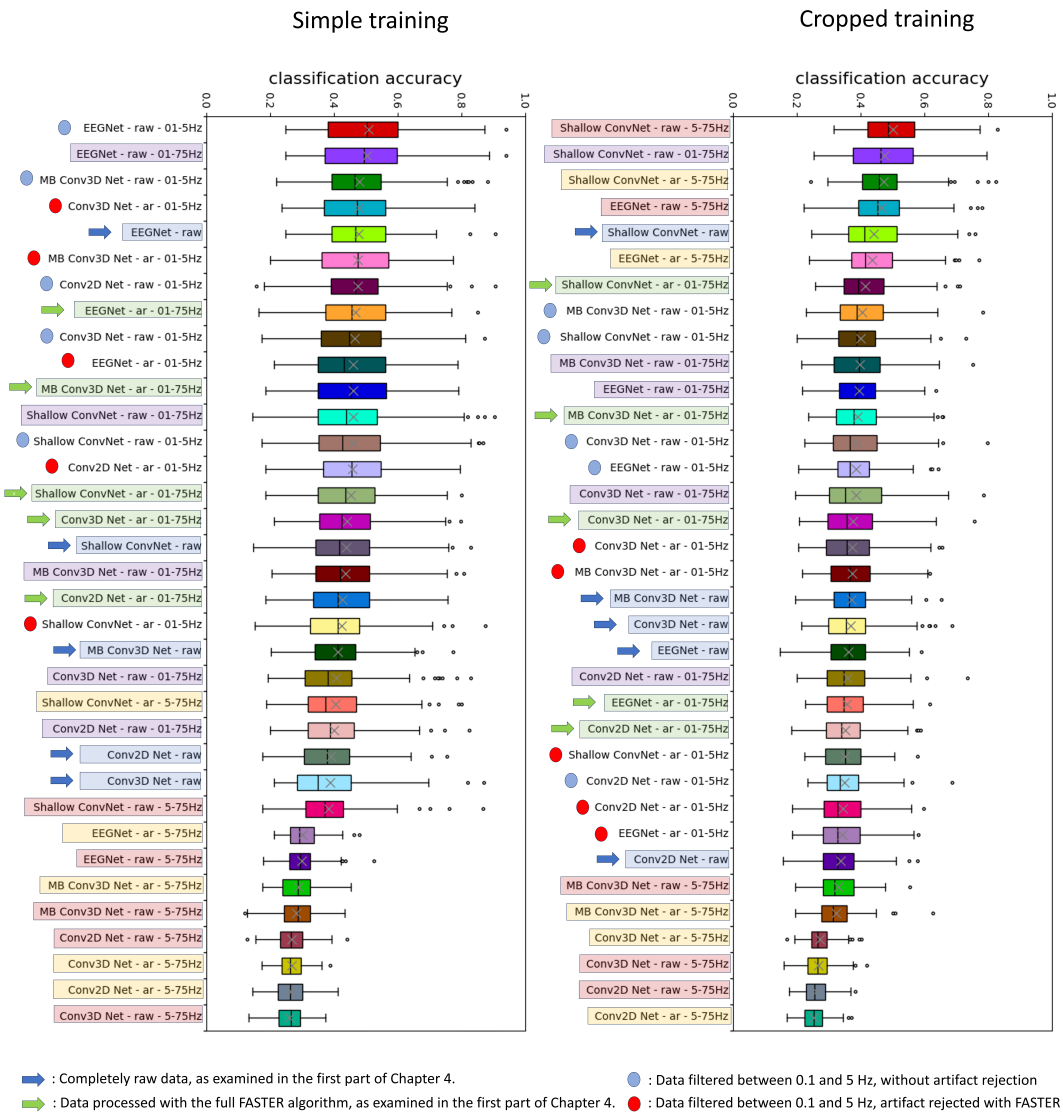

Figure 4.12: The results obtained with and without cropped training, ordered by the classification accuracy. Arrows mark the cases examined in the first part of chapter 4: the differences between the FASTER-ed (green arrow) and the fully raw data (blue arrow). The relation between the two datasets remains the same, as only the EEGNet yields a better raw classification accuracy; however, during cropped training, the Shallow ConvNet also joins this group. Dots marked the results when frequency filtering was presented between 0.1 and 5 Hz. (Red dots for artifact rejected and blue dots for raw data.) Without cropped training, these results tend to be unexpectedly high, indicating that this low-frequency range contains the most relevant information for MI classification on the Physionet database when subject-wise learning is performed without any augmentation of the training dataset. The most noteworthy result is that while during simple learning almost every 5-75 Hz filtered attempt was presented on the lowest part of the list, the cases of EEGNet and Shallow ConvNet jumped to the first places of the cropped training results, corresponding to the extensive literature [J2].

ing techniques across various subjects. As can be observed, the FASTER-applied and frequency-filtered data exhibit similar subject-wise distributions when compared to the raw data, indicating that both methods have comparable effects on subject dependence. However, when comparing these two methods, it is evident that frequency filtering alone yields significantly higher accuracy results. Another important factor to note is that the set of subjects whose classification accuracies improved during cropped training differs substantially from those who showed improvements in the simple training process.

4.5. Discussion

As can be seen from the results, classification accuracy depends on many factors. The slight decrease in classification accuracy observed after applying the FASTER algorithm in the case of EEGNet can be attributed to the possibility that the artifact rejection method may remove not only non-neural components but also neural signal components that are informative for classification. The EEGNet architecture is designed to efficiently extract spatial and temporal EEG features, and it can already achieve higher performance using minimally preprocessed or even raw EEG data. Consequently, it may be more sensitive to the loss of subtle yet discriminative neural information caused by aggressive preprocessing or artifact rejection. For the other four networks, however, when only a limited amount of training data was available, the removal of artifacts significantly boosted performance, suggesting that these architectures are more sensitive to noise and benefit more from the improved signal-to-noise ratio provided by the artifact rejection process.

The subject dependence of the efficacy of artifact rejection is also a noteworthy result. The question that emerges is, what is the reason behind these differences? As each subject has a different scale of artifact contamination, the reasoning can be the amount and quality of contained artifacts. However, it is only partially true because, for instance, subject 15 has the greatest accuracy decline after artifact rejection, but if we examine it closely, multiple artifactual components were rejected there. The real reason can be the number of rejected components on which neural networks can base their classification. If the data from a certain subject contains artifactual yet useful components, the accuracy will decline as a result of the application of the FASTER method.

The application of transfer learning yields a noteworthy improvement for each classification task. This beneficial effect of the process is well-documented in the existing literature [Au1], [110], [111], [112]. What is interesting to note is that transfer learn-

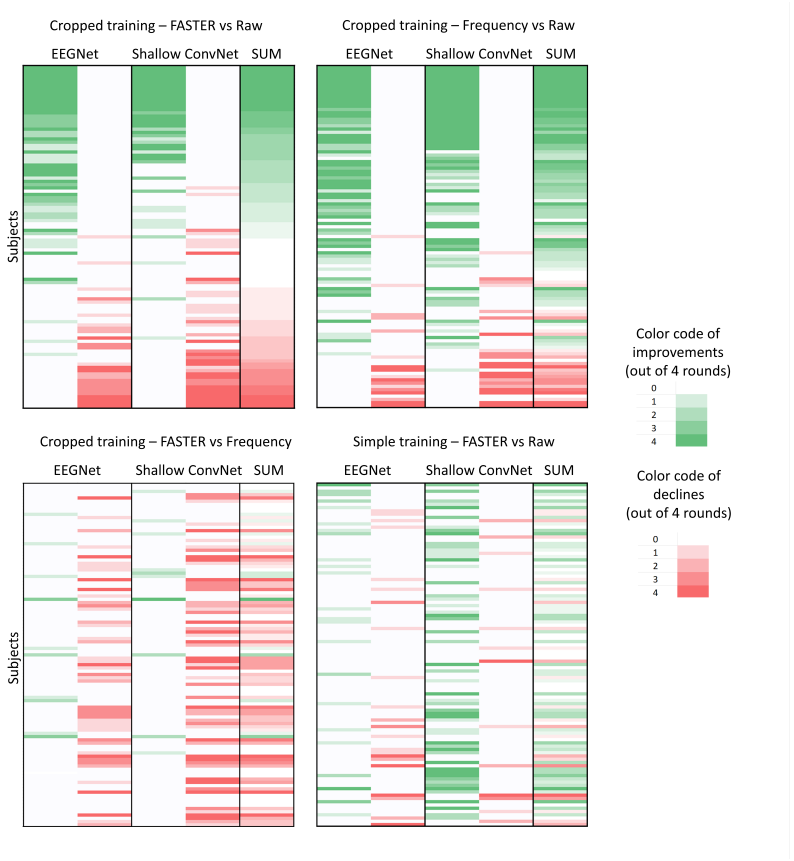

Figure 4.13: Subject-wise effect of the FASTER artifact rejection method and the frequency filtering on the classification performance of the EEGNet and Shallow ConvNet, related to the raw data, during cropped and simple training. The shade of the color means the number of significant increases (green) or declines (red) out of the 4 tests. Subjects are ordered by the score achieved during the AR method compared to the unfiltered data. The order of the subjects is almost the same in the second graph, meaning that frequency filtering has a similar subject-wise effect as the FASTER algorithm, related to the raw data. However, this sequence is merely different from the result from the first part of the article, meaning that during cropped training, it changes which subjects the AR has positive effects on. The sequence is also different on the FASTER compared to the plain frequency filtering graph, where, for most of the subjects, the AR method caused a significant decline. It means that during cropped training, it is more advised to filter only the appropriate frequency ranges and not to use the more complex FASTER method to get better results [J2].

ing improves the raw data classification to the greatest extent, and results obtained this way surpass the accuracy achieved on the artifact-rejected data. A plausible underlying explanation for this phenomenon may be that, with a substantial amount of training data, neural networks can learn to disregard artifacts and focus on neuronal processes. Conversely, the FASTER artifact rejection method may filter out components that are relevant and could have been utilized by neural networks for learning.

The best accuracy is achieved by the Shallow ConvNet with raw data as input and transfer learning, replacing the EEGNet, which was the classifier with the highest accuracy when transfer learning was not applied. This confirms Shallow ConvNet's superior performance in transferring weights from the broader dataset, as was also described earlier [Au1].

As it was presented, adding the spatial dimension using dense 3D representation did not lead to improvements in most cases. This is likely because networks like Shallow ConvNet and EEGNet can inherently learn the spatial relationships between channels through their architectures. Shallow ConvNet achieves this via spatial filtering, while EEGNet uses DepthwiseConv2D layers. These methods allow the networks to autonomously identify and leverage the importance of spatial data and channel interactions, which proves to be more effective for classification performance than explicitly incorporating a dense 3D spatial representation.

Discussing results from the frequency filter section, the conclusion that can be drawn is that when a limited amount of data is given to train the classifier, such as in the case of subject-wise training, even the well-known and widely used Shallow ConvNet and EEGNet tend to concentrate on the lowest frequency ranges, which ranges can contain relevant information. This attention to lower ranges of frequency can be seen in Figure 4.12. An unexpected observation is that networks trained on the 0.1–5 Hz filtered dataset exhibited relatively high accuracies, often surpassing those trained on the 0.1–75 Hz filtered version. In the scenario of lowpass filtering, raw data yielded greater accuracies than the artifact-rejected data, indicating that the previously examined effect of AR does not hold under these circumstances. One possible explanation for the superior classification accuracy of the delta band is its inherently higher amplitude power compared to other frequency bands. This elevated power may provide the classifier with more prominent features, thereby enhancing performance. Additionally, the delta band is closely associated with attentional processes, which may contribute to the presence of underlying features that strongly influence classification accuracy. Furthermore, the limitations of training

data without the use of cropped training or transfer learning might impair the networks' ability to effectively recognize the importance of higher-frequency bands, such as mu and beta. In this work, I demonstrated that the inclusion of cropped training significantly improves the classification of signals containing higher-frequency components, a result that aligns with prior findings in the literature.

It is important to note that when comparing raw data with the artifact-rejected dataset filtered in the 0.1–75 Hz range, we observe the same results: artifact rejection improves performance in four out of five networks, except EEGNet. However, when frequency filtering is applied to the raw data, the Shallow ConvNet without AR also outperforms the version where the FASTER method was applied. For the remaining networks, the performance relationship remains unchanged. Nevertheless, when cropped training was present, the overall picture changed. On one hand, the average classification accuracy of the networks decreases with this method. This decline can be attributed to the test set containing an extended range of data, making it more challenging for the classifiers to produce accurate results. On the other hand, via this method, we can acquire more data to train the classifier, as in the case of transfer learning. This extended version of the dataset is enough for the EEGNet and the Shallow ConvNet to learn from the higher frequency ranges, as expected from the literature. In investigating this aspect, it is shown that EEGNet and Shallow ConvNet are more capable of generalizing data, as they effectively leverage the variability introduced by cropped training. This demonstrates their ability to adapt to different temporal onsets in the EEG signals, a crucial factor in improving classification accuracy and model robustness.

Figure 4.13 demonstrates a notable consistency in the performance trends observed during cropped training across networks trained on frequency-filtered and FASTER-processed datasets, exhibiting a similar propensity for performance enhancement in comparison to the raw dataset. This consistency suggests that, in this case, the performance enhancement caused by the FASTER algorithm can be explained by the effect of frequency filtering. Moreover, if we compare the artifact rejection method to the frequency-filtered performance, the latter gets significantly higher accuracies in the vast majority of subjects. That means that with a simple frequency filter, we can achieve better performance during cropped training than with a complex AR method. Another important observation to note is that the subjects that had higher accuracies with the FASTER method compared to the raw data during cropped training differ from the subjects that were obtained by simple learning. This could be an effect of the phenomenon that dur-

ing cropped training, different parts of the original signals are simultaneously presented, and neural networks can learn about other factors, some of which are filtered out during artifact rejection.

There are only a limited number of papers examining the subjects of the Physionet database regarding the effect of artifact rejection; therefore, it is hard to give a thorough comparison. There are studies where filtering enhances the classification performance in various datasets [92], [99]. Other articles report a decline in accuracy due to artifact rejection [93]. Subject dependence remains a significant factor, as evidenced not only in various articles exploring the BCI Competition IV Dataset 2a [96] but also in this study focusing on the Physionet database. My findings highlight that the efficacy of the FASTER artifact rejection method in terms of classification accuracy is profoundly influenced by the specific subject under consideration.

For the case of frequency dependence, Hauke Dose et al. [104], who also examined the Shallow ConvNet's accuracy on the Physionet database, concluded that this architecture tends to concentrate on the lowest part of the frequency domain. In their experiment, they analyzed the squared frequency responses of the learned temporal filters, and the mean focused on the lowest frequency range (below 10 Hz). These results correspond to my findings. The frequency results obtained through cropped training and the use of Shallow ConvNet and EEGNet align with the broader literature. As imagery movements are described to be mostly classifiable in the mu and beta ranges [19], [114], it is expected that signals filtered between 5 and 75 Hz exhibit higher (or at least similar) accuracies compared to unfiltered signals.

The study reveals how small changes in the preprocessing pipeline can significantly impact classification accuracy, underscoring the need for tailored solutions in EEG-based BCI systems. As discussed by Xu et al. [115], while much progress has been made in neural interface research, translating these advancements into reliable, real-world applications remains challenging. This research contributes to bridging that gap by optimizing processing techniques that can enhance the practicality of BCIs in neurorehabilitation and beyond.

In conclusion, my research indicates that the FASTER method can enhance performance in a subject and network-specific manner. There are subjects where the application of AR comes with an efficiency increase, while in other cases, it comes with a deterioration of the results. Transfer learning proved to be effective in improving the performance of all networks in both raw and artifact-rejected data. However, it was noted that the

accuracy of classification for artifact-rejected data did not improve as significantly as it did for the unfiltered data, resulting in less precision. My findings also revealed an unexpected outcome from frequency filtering, as the tested networks demonstrated strong classification performance based on the low-frequency components during learning. Notably, I observed that higher frequency ranges were more discriminative in the case of EEGNet and Shallow ConvNet when cropped training was applied. In summary, my study underscores the intricate interplay between processing techniques and neural network performance, highlighting the necessity for tailored processing approaches designed for specific subjects and network architectures.

Chapter 5

Summary

5.1. New Scientific Results

Thesis group I

The first group of thesis points focuses on the design, implementation, and real-world deployment of an online artifact rejection system, developed specifically to support our BCI team during the Cybathlon 2020 and 2024 competitions. These international events present a unique environment where Brain-Computer Interface systems must operate under strict real-time conditions, with high demands for technical reliability. The thesis points of this group are based on my second-author article, [J1].

Thesis I.1: *Based on the FASTER (Fully Automated Statistical Thresholding for EEG artifact Rejection) algorithm - originally developed for offline EEG processing, I developed an online artifact removal system which can be used during real-time BCI experiments with an average latency of 0.123 ± 0.017 s.*

Explanation: The additional processing required for EEGNet adds 0.061 ± 0.010 s, and together with the 1-second EEG segment the total delay remains within the recommended 1–2 s window for real-time BCI operation. This confirms that the developed online artifact rejection method satisfies the temporal constraints of practical BCI applications.

Thesis I.2: *I integrated the developed real-time artifact removal system into our team's BCI framework for the 2020 and 2024 Cybathlon events. The system was officially approved for our team in the 2024 competition by the Cybathlon organizing committee, confirming that the solution meets the event's strict technical requirements.*

Explanation: The integration of the online artifact removal system into the Cybathlon BCI framework demonstrates its applicability in a real-time competition environment. The official approval by the Cybathlon organizing committee confirms that the system complied with the event's technical specifications and operated reliably within the constraints of the competition setting.

Thesis group II

The second thesis group investigates the complex and often nonlinear relationships between EEG preprocessing steps and the resulting classification performance in brain-computer interface systems. Preprocessing plays a fundamental role in improving EEG signal quality and reducing noise. However, its effect on classification outcomes is not always intuitive or consistent across different subjects, datasets, and neural network architectures. The theses described here are based on my first-author journal article [J2].

Thesis II.1: *I demonstrated that the impact of the FASTER artifact removal algorithm on classification accuracy strongly depends on the subject under consideration. In a comparative analysis, I found that applying the FASTER algorithm led to improved classification accuracy for 80 out of 105 subjects, with an average improvement of 5.4%. However, this positive effect is not universal: for the remaining 25 subjects, the method resulted in a decrease in accuracy of 4.4% on average.*

Explanation: These findings highlight that the effectiveness of automated artifact removal is not uniform across subjects. The variability suggests that individual differences in EEG signal quality, artifact characteristics, or task-related neural activity can influence whether the FASTER procedure preserves or removes information relevant for classification. Consequently, while the method generally provides notable benefits, its subject-dependent behavior should be taken into account when designing or optimizing BCI pipelines.

Thesis II.2: *I demonstrated that transfer learning consistently improves EEG classification accuracy across multiple pre-training conditions for both raw and artifact-cleaned datasets. The magnitude of these improvements, however, differed substantially: for raw data, transfer learning increased accuracy from 40.0% to 57.5%, whereas for data processed with the FASTER artifact-removal pipeline, the improvement was smaller, from*

43.1% to 53.3%. These values represent averages across all five examined neural network architectures.

Explanation: These results indicate that transfer learning is particularly effective when the input data retains a richer set of signal characteristics, as in the case of raw EEG. The comparatively smaller improvement observed for FASTER-processed data suggests that artifact removal reduces not only noise but also some of the variability and informative structure that transfer learning can exploit. Overall, this pattern highlights that the benefit of transfer learning depends on the preprocessing strategy, with the largest gains obtained when the model can adapt to the full complexity of the unfiltered EEG signals.

Thesis II.3: *I designed two neural networks to assess the role of explicitly modeling spatial dimensions in EEG classification. For this purpose, I introduced a novel dense 3D representation of EEG data that captures the electrode layout more accurately. This representation improved accuracy for certain architectures, yielding a significant 4.43% gain with Conv3D, a non-significant 1.19% gain with Conv2D, and a non-significant 0.15% decrease with the Multi Branch Conv3D model. The results concerning the comparison of manual and learned spatial encodings indicate that explicit spatial encoding does not outperform architectures that implicitly learn spatial relationships: while networks with manual spatial encoding achieved an average accuracy of 53.6%, models relying on implicit spatial feature learning reached 58.0%, with transfer learning applied.*

Explanation: This tendency is most relevant for the configuration combining raw EEG data with transfer learning, which yielded the highest overall performance and therefore carries the greatest significance. In contrast, for data processed with the FASTER algorithm, a slight deviation from this pattern was observed: the Multi Branch Conv3D network slightly outperformed EEGNet, while the Conv2D and Conv3D architectures remained among the lowest-performing models.

Thesis II.4: *I demonstrated that applying frequency filtering yields unexpected patterns in classification performance on the Physionet motor imagery database, with the δ band being a decisive factor. Averaged across all five networks, the classification accuracy reached 46.8% in the 0.1–5 Hz frequency band, whereas it was 30.3% in the 5–75 Hz band. In contrast, when training and testing were performed on cropped EEG segments, the higher-frequency bands - the μ, β, and γ ranges - became more informative for both EEGNet and Shallow ConvNet. Specifically, EEGNet achieved 36.4% accuracy in the*

0.1–5 Hz band and 45.1% in the 5–75 Hz band, while Shallow ConvNet reached 37.7% and 48.8% in the respective bands.

Explanation: As illustrated in Table 5.1, the shift in frequency sensitivity is particularly evident for EEGNet and Shallow ConvNet - the two architectures capable of exploiting the additional variability introduced by cropped training. While simple training favors the low-frequency range due to the limited number of available samples, cropped training provides a much larger set of segments, enabling these models to learn more complex oscillatory patterns associated with the μ and β bands. For the remaining networks, which rely on the dense 3D spatial representation, this shift was not observed in the broader analysis: their performance continued to be dominated by the low-frequency range regardless of training strategy. This underscores that the effect of frequency filtering is strongly architecture-dependent and is modulated by how effectively a model can utilize the increased data quantity generated by cropping.

Table 5.1: Comparison of EEGNet and Shallow ConvNet on different frequency ranges and training strategies.

		0.5–5 Hz	**5–75 Hz**
Simple training	**EEGNet**	48.48%	30.07%
	Shallow ConvNet	44.22%	39.46%
Cropped training	**EEGNet**	36.42%	45.13%
	Shallow ConvNet	37.65%	48.78%

Thesis II.5: *I demonstrated that, during cropped-segment training of EEGNet and Shallow ConvNet, the FASTER method generally improves performance for subjects who also benefit from frequency filtering alone. Moreover, frequency filtering by itself typically produces better results than the full FASTER algorithm, with average accuracies across the two networks of 46.4% for frequency filtering and 43.3% for the full FASTER algorithm. I also showed that the subjects whose accuracy improved with cropped-segment training were largely distinct from those who benefited under full-length segment training.*

Explanation: To assess the consistency of subject-specific preprocessing effects, each subject was assigned a score between –8 and +8, reflecting the number of rounds in which MI classification accuracy significantly increased or decreased following preprocessing. Separate scores were obtained for frequency filtering and for the full FASTER method. The rank correlation between the two score sets was strong (Spearman's $\rho =$

0.824, Kendall's $\tau = 0.689$, $p < 0.001$), indicating that subjects exhibiting consistent improvements under frequency filtering tended to show similar improvement patterns under the full FASTER procedure. This finding suggests that the effectiveness of preprocessing is strongly subject-dependent and that individual signal characteristics influence the benefits observed across different artifact rejection strategies. In line with this observation, Table 5.2 demonstrates that, on average, frequency filtering alone yields higher classification accuracy than the full FASTER algorithm, with both preprocessing methods outperforming the raw data condition.

Table 5.2: Comparison of EEGNet and Shallow ConvNet on different preprocessing strategies.

	Raw data	**Freqency-filtered**	**FASTER filtered**
EEGNet	36.39%	42.98%	41.08%
Shallow ConvNet	43.73%	49.72%	45.47%

5.2. Potential Applications and Benefits

The results of this dissertation offer direct applications in both real-time and offline Brain-Computer Interface systems. The development and deployment of the online artifact rejection system provides a practical solution for ensuring the integrity and reliability of neural signals during live use. This system, officially accepted by the Cybathlon committee, sets a precedent for technically robust artifact handling in competitive environments. The second thesis group provides analytical insights into the non-linear and subject-dependent effects of EEG preprocessing on classification accuracy. These findings can significantly enhance the personalization of BCI pipelines by informing developers when and how to apply specific preprocessing techniques - such as artifact rejection, frequency filtering, or transfer learning - based on subject characteristics and training strategy. The discovery that simpler filtering can sometimes outperform more complex methods (like FASTER) under certain conditions, and that different subjects benefit from different training regimes, promotes the development of adaptive and context-aware BCI frameworks. More broadly, this work contributes to improving the efficiency, generalizability, and interpretability of neural network models applied to EEG data, making BCI systems more accessible, accurate, and responsive to user variability.

Use of AI Assistance

During the preparation of this dissertation, I utilized ChatGPT (developed by OpenAI) to enhance grammar and improve readability. Following the use of this tool, I carefully reviewed and edited the content to ensure accuracy and appropriateness. I take full responsibility for the content and conclusions presented in this work. All results, insights, and critical reasoning are entirely my own. The use of AI tools was limited to supporting clarity and coherence.

Journal publications of the thesis

[J1] C. M. Köllőd, **A. Adolf**, G. Márton, M. Wahdow, W. Fadel, and I. Ulbert, "Closed loop BCI system for Cybathlon 2020", *Brain-Computer Interfaces*, vol. 10, no. 2, pp. 114–128, 2023. DOI: 10.1080/2326263X.2023.2254463.

[J2] **A. Adolf**, C. M. Köllőd, G. Márton, W. Fadel, and I. Ulbert, "The effect of processing techniques on the classification accuracy of brain-computer interface systems", *Brain Sciences*, vol. 14, no. 12, 2024. DOI: 10.3390/brainsci14121272.

Other publications of the author

[Au1] C. M. Köllőd, **A. Adolf**, K. Iván, G. Márton, and I. Ulbert, "Deep Comparisons of Neural Networks from the EEGNet Family", *Electronics*, vol. 12, no. 12, p. 2743, 2023. DOI: 10.3390/electronics12122743.

[Au2] M. Wahdow, M. Alnaanah, W. Fadel, **A. Adolf**, C. M. Köllőd, and I. Ulbert, "Multi frequency band fusion method for EEG signal classification", *Signal, Image and Video Processing*, 2022. DOI: 10.1007/s11760-022-02399-6.

[Au3] C. M. Köllőd, **A. Adolf**, G. Márton, M. Wahdow, W. Fadel, and I. Ulbert, *TTK dataset - 4 class Motor-Imagery EEG*, 2022, https://hdl.handle.net/21.15109/CONCORDA/UOQQVK.

Bibliography

[1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain–computer interfaces for communication and control", *Clinical Neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002. DOI: 10.1016/S1388-2457(02)00057-3.

[2] S. Saha et al., "Progress in Brain Computer Interface: Challenges and Opportunities", *Frontiers in Systems Neuroscience*, vol. 15, p. 578 875, 2021. DOI: 10.3389/fnsys.2021.578875.

[3] X. Jiang, G.-B. Bian, and Z. Tian, "Removal of Artifacts from EEG Signals: A Review", *Sensors*, vol. 19, no. 5, p. 987, 2019. DOI: 10.3390/s19050987.

[4] I. Daly, R. Scherer, M. Billinger, and G. Müller-Putz, "Force: Fully Online and Automated Artifact Removal for Brain-Computer Interfacing", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 23, no. 5, pp. 725–736, 2015. DOI: 10.1109/TNSRE.2014.2346621.

[5] S.-H. Hsu, T. R. Mullen, T.-P. Jung, and G. Cauwenberghs, "Real-Time Adaptive EEG Source Separation Using Online Recursive Independent Component Analysis", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, no. 3, pp. 309–319, 2016. DOI: 10.1109/TNSRE.2015.2508759.

[6] H. Nolan, R. Whelan, and R. B. Reilly, "Faster: Fully Automated Statistical Thresholding for EEG artifact Rejection", *Journal of Neuroscience Methods*, vol. 192, no. 1, pp. 152–162, 2010. DOI: 10.1016/j.jneumeth.2010.07.015.

[7] W. J. Freeman, "Tutorial on neurobiology: From single neurons to brain chaos", *International Journal of Bifurcation and Chaos*, vol. 02, no. 03, pp. 451–482, 1992. DOI: 10.1142/S0218127492000653.

[8] G. Buzsáki, C. A. Anastassiou, and C. Koch, "The origin of extracellular fields and currents — EEG, ECoG, LFP and spikes", *Nature Reviews Neuroscience*, vol. 13, no. 6, pp. 407–420, 2012. DOI: 10.1038/nrn3241.

[9] C. DelGratta, V. Pizzella, F. Tecchio, and G.-L. Romani, "Magnetoencephalography—A noninvasive brain imaging method with 1 ms time resolution", *Reports on Progress in Physics - REP PROGR PHYS*, vol. 64, pp. 1545–1600, 2001. DOI: 10.1088/0034-4885/64/12/204.

[10] V. Scarapicchia, C. Brown, C. Mayo, and J. R. Gawryluk, "Functional Magnetic Resonance Imaging and Functional Near-Infrared Spectroscopy: Insights from Combined Recording Studies", *Frontiers in Human Neuroscience*, vol. 11, p. 419, 2017. DOI: 10.3389/fnhum.2017.00419.

[11] N. K. Logothetis, "The Underpinnings of the BOLD Functional Magnetic Resonance Imaging Signal", *The Journal of Neuroscience*, vol. 23, no. 10, pp. 3963–3971, 2003. DOI: 10.1523/JNEUROSCI.23-10-03963.2003.

[12] N. Naseer and K.-S. Hong, "fNIRS-based brain-computer interfaces: A review", English, *Frontiers in Human Neuroscience*, vol. 9, 2015. DOI: 10.3389/fnhum.2015.00003.

[13] N. Lago and A. Cester, "Flexible and Organic Neural Interfaces: A Review", *Applied Sciences*, vol. 7, no. 12, p. 1292, 2017. DOI: 10.3390/app7121292.

[14] A. F. Jackson and D. J. Bolger, "The neurophysiological bases of EEG and EEG measurement: A review for the rest of us", *Psychophysiology*, vol. 51, no. 11, pp. 1061–1071, 2014. DOI: 10.1111/psyp.12283.

[15] R. Abiri, S. Borhani, E. W. Sellers, Y. Jiang, and X. Zhao, "A comprehensive review of EEG-based brain–computer interface paradigms", *Journal of Neural Engineering*, vol. 16, no. 1, p. 011 001, 2019. DOI: 10.1088/1741-2552/aaf12e.

[16] L. A. Farwell and E. Donchin, "Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials", *Electroencephalography and Clinical Neurophysiology*, vol. 70, no. 6, pp. 510–523, 1988. DOI: 10.1016/0013-4694(88)90149-6.

[17] A. J. Hurst and S. G. Boe, "Imagining the way forward: A review of contemporary motor imagery theory", *Frontiers in Human Neuroscience*, vol. 16, 2022. DOI: 10.3389/fnhum.2022.1033493.

[18] M. Jeannerod and J. Decety, "Mental motor imagery: A window into the representational stages of action", *Current Opinion in Neurobiology*, vol. 5, no. 6, pp. 727–732, 1995. DOI: 10.1016/0959-4388(95)80099-9.

[19] G. Pfurtscheller, C. Brunner, A. Schlögl, and F. H. Lopes da Silva, "Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks", *NeuroImage*, vol. 31, no. 1, pp. 153–159, 2006. DOI: 10.1016/j.neuroimage.2005.12.003.

[20] A. L. Goldberger et al., "Physiobank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals", *Circulation*, vol. 101, no. 23, e215–e220, 2000. DOI: 10.1161/01.CIR.101.23.e215.

[21] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, "Bci2000: A general-purpose brain-computer interface (BCI) system", *IEEE transactions on bio-medical engineering*, vol. 51, no. 6, pp. 1034–1043, 2004. DOI: 10/b7bs63.

[22] C.-C. Fan, H. Yang, Z.-G. Hou, Z.-L. Ni, S. Chen, and Z. Fang, "Bilinear neural network with 3-D attention for brain decoding of motor imagery movements from the human EEG", *Cognitive Neurodynamics*, 2020. DOI: 10/ghsn4r.

[23] K. Roots, Y. Muhammad, and N. Muhammad, "Fusion Convolutional Neural Network for Cross-Subject EEG Motor Imagery Classification", *Computers*, vol. 9, no. 3, p. 72, 2020. DOI: 10/gk3p86.

[24] M. Tangermann et al., "Review of the BCI Competition IV", *Frontiers in Neuroscience*, vol. 6, 2012.

[25] M. Fatourechi, A. Bashashati, R. K. Ward, and G. E. Birch, "Emg and EOG artifacts in brain computer interface systems: A survey", *Clinical Neurophysiology*, vol. 118, no. 3, pp. 480–494, 2007. DOI: 10.1016/j.clinph.2006.10.019.

[26] M. M. N. Mannan, M. A. Kamran, and M. Y. Jeong, "Identification and Removal of Physiological Artifacts From Electroencephalogram Signals: A Review", *IEEE Access*, vol. 6, pp. 30 630–30 652, 2018. DOI: 10.1109/ACCESS.2018.2842082.

[27] S. Kanoga and Y. Mitsukura, "Review of Artifact Rejection Methods for Electroencephalographic Systems", in 2017.

[28] A. Delorme, S. Makeig, and T. Sejnowski, "Automatic artifact rejection for EEG data using high-order statistics and independent component analysis", *Proceedings of the third international ICA conference*, pp. 9–12, 2001.

[29] T. Radüntz, J. Scouten, O. Hochmuth, and B. Meffert, "Eeg artifact elimination by extraction of ICA-component features using image processing algorithms", *Journal of Neuroscience Methods*, vol. 243, pp. 84–93, 2015. DOI: 10.1016/j.jneumeth.2015.01.030.

[30] I. Winkler, S. Brandl, F. Horn, E. Waldburger, C. Allefeld, and M. Tangermann, "Robust artifactual independent component classification for BCI practitioners", *Journal of Neural Engineering*, vol. 11, no. 3, p. 035 013, 2014. DOI: 10.1088/1741-2560/11/3/035013.

[31] N. Bajaj, J. Requena Carrión, F. Bellotti, R. Berta, and A. De Gloria, "Automatic and tunable algorithm for EEG artifact removal using wavelet decomposition with applications in predictive modeling during auditory tasks", *Biomedical Signal Processing and Control*, vol. 55, p. 101 624, 2020. DOI: 10.1016/j.bspc.2019.101624.

[32] W. De Clercq, A. Vergult, B. Vanrumste, W. Van Paesschen, and S. Van Huffel, "Canonical Correlation Analysis Applied to Remove Muscle Artifacts From the Electroencephalogram", *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 12, pp. 2583–2587, 2006. DOI: 10.1109/TBME.2006.879459.

[33] V. V. Grubov, A. E. Runnova, T. Y. Efremova, and A. E. Hramov, "Artifact removal from EEG data with empirical mode decomposition", vol. 10063, 100631F, 2017. DOI: 10.1117/12.2251371.

[34] L. F. Nicolas-Alonso and J. Gomez-Gil, "Brain Computer Interfaces, a Review", *Sensors*, vol. 12, no. 2, pp. 1211–1279, 2012. DOI: 10.3390/s120201211.

[35] C.-H. Chuang, K.-Y. Chang, C.-S. Huang, and T.-P. Jung, "Ic-u-net: A u-net-based denoising autoencoder using mixtures of independent components for automatic eeg artifact removal", *NeuroImage*, vol. 263, p. 119 586, 2022. DOI: https://doi.org/10.1016/j.neuroimage.2022.119586.

[36] J. F. Hwaidi and T. M. Chen, "Classification of motor imagery eeg signals based on deep autoencoder and convolutional neural network approach", *IEEE Access*, vol. 10, pp. 48 071–48 081, 2022. DOI: 10.1109/ACCESS.2022.3171906.

[37] S. Mirzaei and P. Ghasemi, "Eeg motor imagery classification using dynamic connectivity patterns and convolutional autoencoder", *Biomedical Signal Processing and Control*, vol. 68, p. 102 584, 2021. DOI: https://doi.org/10.1016/j.bspc.2021.102584.

[38] D. J. Krusienski, G. Schalk, D. J. McFarland, and J. R. Wolpaw, "A mu-rhythm matched filter for continuous control of a brain-computer interface", *IEEE transactions on bio-medical engineering*, vol. 54, no. 2, pp. 273–280, 2007. DOI: 10.1109/TBME.2006.886661.

[39] V. J. Samar, A. Bopardikar, R. Rao, and K. Swartz, "Wavelet analysis of neuroelectric waveforms: A conceptual tutorial", *Brain and Language*, vol. 66, no. 1, pp. 7–60, 1999. DOI: 10.1006/brln.1998.2024.

[40] G. Tzanetakis, G. Essl, and P. Cook, "Audio Analysis using the Discrete Wavelet Transform", *Proceedings of the Acoustics and Music Theory Applications*, vol. 66, pp. 318–323, 2001.

[41] G. Othman and D. Q. Zeebaree, "The Applications of Discrete Wavelet Transform in Image Processing: A Review", *Journal of Soft Computing and Data Mining*, vol. 1, no. 2, pp. 31–43, 2020.

[42] H. Ramoser, J. Müller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement", *IEEE transactions on rehabilitation engineering: a publication of the IEEE Engineering in Medicine and Biology Society*, vol. 8, no. 4, pp. 441–446, 2000. DOI: 10.1109/86.895946.

[43] M. Grosse-Wentrup and M. Buss, "Multiclass Common Spatial Patterns and Information Theoretic Feature Extraction", *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 8, pp. 1991–2000, 2008. DOI: 10.1109/TBME.2008.921154.

[44] K. K. Ang, Z. Y. Chin, H. Zhang, and C. Guan, "Filter bank common spatial pattern (fbcsp) in brain-computer interface", in *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, 2008, pp. 2390–2397. DOI: 10.1109/IJCNN.2008.4634130.

[45] S. Aggarwal and N. Chugh, "Review of Machine Learning Techniques for EEG Based Brain Computer Interface", *Archives of Computational Methods in Engineering*, vol. 29, no. 5, pp. 3001–3020, 2022. DOI: 10.1007/s11831-021-09684-6.

[46] A. Schlögl, F. Lee, H. Bischof, and G. Pfurtscheller, "Characterization of four-class motor imagery EEG data for the BCI-competition 2005", *Journal of Neural Engineering*, vol. 2, no. 4, p. L14, 2005. DOI: 10.1088/1741-2560/2/4/L02.

[47] Siuly, Y. Li, and P. (Wen, "Clustering technique-based least square support vector machine for EEG signal classification", *Computer Methods and Programs*

in Biomedicine, vol. 104, no. 3, pp. 358–372, 2011. DOI: 10.1016/j.cmpb.2010.11.014.

[48] W. Hu, X. Geng, M. Yue, L. Wang, and X. Zhang, "Feature Extraction of Motor Imagery EEG Signals Based on PSD CSP Fusion", in *Intelligent Computing Technology and Automation*. IOS Press, 2024, pp. 66–72.

[49] J. Bi, Y. Gao, Z. Peng, and Y. Ma, "Classification of motor imagery using chaotic entropy based on sub-band EEG source localization", *Journal of Neural Engineering*, vol. 21, no. 3, p. 036 016, 2024. DOI: 10.1088/1741-2552/ad4914.

[50] S. Salcedo-Sanz, J. L. Rojo-Álvarez, M. Martínez-Ramón, and G. Camps-Valls, "Support vector machines in engineering: An overview", *WIREs Data Mining and Knowledge Discovery*, vol. 4, no. 3, pp. 234–267, 2014. DOI: 10.1002/widm.1125.

[51] S. S. Haykin, *Neural networks and learning machines*, 3. ed. New York Munich: Prentice-Hall, 2009.

[52] F. Rosenblatt, "The perceptron: A probabilistic model for information storage and organization in the brain.", *Psychological review*, vol. 65, no. 6, p. 386, 1958.

[53] M. A. Nielsen, "Neural Networks and Deep Learning", 2015.

[54] K. O'Shea and R. Nash, "An Introduction to Convolutional Neural Networks", 2015. DOI: 10.48550/arXiv.1511.08458.

[55] K. He, X. Zhang, S. Ren, and J. Sun, *Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification*, 2015. DOI: 10.48550/arXiv.1502.01852.

[56] Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, "A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects", *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 12, pp. 6999–7019, 2022. DOI: 10.1109/TNNLS.2021.3084827.

[57] H. Gholamalinezhad and H. Khosravi, "Pooling Methods in Deep Neural Networks, a Review", 2020. DOI: 10.48550/arXiv.2009.07485.

[58] A. Stergiou and R. Poppe, "AdaPool: Exponential Adaptive Pooling for Information-Retaining Downsampling", *IEEE Transactions on Image Processing*, vol. 32, pp. 251–266, 2023. DOI: 10.1109/TIP.2022.3227503.

[59] J. Long, E. Shelhamer, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation", 2015, pp. 3431–3440.

[60] Y. Hu et al., "A Cross-Space CNN With Customized Characteristics for Motor Imagery EEG Classification", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, pp. 1554–1565, 2023. DOI: 10.1109/TNSRE.2023.3249831.

[61] A. Echtioui, A. Mlaouah, W. Zouch, M. Ghorbel, C. Mhiri, and H. Hamam, "A Novel Convolutional Neural Network Classification Approach of Motor-Imagery EEG Recording Based on Deep Learning", *Applied Sciences*, vol. 11, no. 21, p. 9948, 2021. DOI: 10.3390/app11219948.

[62] M. Riyad, M. Khalil, and A. Adib, "Mi-EEGNET: A novel convolutional neural network for motor imagery classification", *Journal of Neuroscience Methods*, vol. 353, p. 109 037, 2021. DOI: 10.1016/j.jneumeth.2020.109037.

[63] R. Schirrmeister et al., "Deep learning with convolutional neural networks for EEG decoding and visualization: Convolutional Neural Networks in EEG Analysis", *Human Brain Mapping*, vol. 38, 2017. DOI: 10.1002/hbm.23730.

[64] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, "Eegnet: A Compact Convolutional Network for EEG-based Brain-Computer Interfaces", *Journal of Neural Engineering*, vol. 15, no. 5, p. 056 013, 2018. DOI: 10.1088/1741-2552/aace8c.

[65] P. Gong, P. Wang, Y. Zhou, and D. Zhang, "A Spiking Neural Network With Adaptive Graph Convolution and LSTM for EEG-Based Brain-Computer Interfaces", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, pp. 1440–1450, 2023. DOI: 10.1109/TNSRE.2023.3246989.

[66] H. Li, M. Ding, R. Zhang, and C. Xiu, "Motor imagery EEG classification algorithm based on CNN-LSTM feature fusion network", *Biomedical Signal Processing and Control*, vol. 72, p. 103 342, 2022. DOI: 10.1016/j.bspc.2021.103342.

[67] P. Wang, A. Jiang, X. Liu, J. Shang, and L. Zhang, "Lstm-Based EEG Classification in Motor Imagery Tasks", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 11, pp. 2086–2095, 2018. DOI: 10.1109/TNSRE.2018.2876129.

[68] F. Xu et al., "A framework for motor imagery with LSTM neural network", *Computer Methods and Programs in Biomedicine*, vol. 218, p. 106 692, 2022. DOI: 10.1016/j.cmpb.2022.106692.

[69] D. Zhang, L. Yao, X. Zhang, S. Wang, W. Chen, and R. Boots, "Eeg-based Intention Recognition from Spatio-Temporal Representations via Cascade and Parallel Convolutional Recurrent Neural Networks", 2017. DOI: 10.48550/arXiv.1708.06578.

[70] A. Vaswani et al., "Attention Is All You Need", 2023. DOI: 10.48550/arXiv.1706.03762.

[71] T. Lin, Y. Wang, X. Liu, and X. Qiu, "A survey of transformers", *AI Open*, vol. 3, pp. 111–132, 2022. DOI: https://doi.org/10.1016/j.aiopen.2022.10.001.

[72] H.-J. Ahn, D.-H. Lee, J.-H. Jeong, and S.-W. Lee, "Multiscale Convolutional Transformer for EEG Classification of Mental Imagery in Different Modalities", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, pp. 646–656, 2023. DOI: 10.1109/TNSRE.2022.3229330.

[73] A. Hameed et al., "Temporal–spatial transformer based motor imagery classification for BCI using independent component analysis", *Biomedical Signal Processing and Control*, vol. 87, p. 105 359, 2024. DOI: 10.1016/j.bspc.2023.105359.

[74] L. Hu, W. Hong, and L. Liu, "Msatnet: Multi-scale adaptive transformer network for motor imagery classification", *Frontiers in Neuroscience*, vol. 17, 2023. DOI: 10.3389/fnins.2023.1173778.

[75] J. Luo et al., "A shallow mirror transformer for subject-independent motor imagery BCI", *Computers in Biology and Medicine*, vol. 164, p. 107 254, 2023. DOI: 10.1016/j.compbiomed.2023.107254.

[76] Y. Ma, Y. Song, and F. Gao, "A novel hybrid CNN-Transformer model for EEG Motor Imagery classification", in *2022 International Joint Conference on Neural Networks (IJCNN)*, 2022, pp. 1–8. DOI: 10.1109/IJCNN55064.2022.9892821.

[77] M. A. Pfeffer, S. S. H. Ling, and J. K. W. Wong, "Exploring the frontier: Transformer-based models in EEG signal analysis for brain-computer interfaces", *Computers in Biology and Medicine*, vol. 178, p. 108 705, 2024. DOI: 10.1016/j.compbiomed.2024.108705.

[78] Y. Song, Q. Zheng, B. Liu, and X. Gao, "Eeg conformer: Convolutional transformer for eeg decoding and visualization", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, pp. 710–719, 2023. DOI: 10.1109/TNSRE.2022.3230250.

[79] W. Zhao, X. Jiang, B. Zhang, S. Xiao, and S. Weng, "CTNet: A convolutional transformer network for EEG-based motor imagery classification", en, *Scientific Reports*, vol. 14, no. 1, p. 20 237, 2024. DOI: 10.1038/s41598-024-71118-7.

[80] J. Xie et al., "A Transformer-Based Approach Combining Deep Learning Network and Spatial-Temporal Information for Raw EEG Classification", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 30, pp. 2126–2136, 2022. DOI: 10.1109/TNSRE.2022.3194600.

[81] P. Novello, G. Poëtte, D. Lugato, and P. M. Congedo, "Goal-Oriented Sensitivity Analysis of Hyperparameters in Deep Learning", *Journal of Scientific Computing*, vol. 94, no. 3, p. 45, 2023. DOI: 10.1007/s10915-022-02083-4.

[82] L. Liao, H. Li, W. Shang, and L. Ma, "An Empirical Study of the Impact of Hyperparameter Tuning and Model Optimization on the Performance Properties of Deep Neural Networks", *ACM Trans. Softw. Eng. Methodol.*, vol. 31, no. 3, 53:1–53:40, 2022. DOI: 10.1145/3506695.

[83] A. Miladinović, A. Accardo, J. Jarmolowska, U. Marusic, and M. Ajčević, "Optimizing real-time mi-bci performance in post-stroke patients: Impact of time window duration on classification accuracy and responsiveness", *Sensors*, vol. 24, no. 18, 2024. DOI: 10.3390/s24186125.

[84] J. Zhong and F. Qi, "Study on the effect of artefact rejection on BCI performance", in *2018 IEEE 23rd International Conference on Digital Signal Processing (DSP)*, 2018, pp. 1–5. DOI: 10.1109/ICDSP.2018.8631586.

[85] M. Kim and S.-P. Kim, "A comparsion of artifact rejection methods for a BCI using event related potentials", in *2018 6th International Conference on Brain-Computer Interface (BCI)*, 2018, pp. 1–4. DOI: 10.1109/IWW-BCI.2018.8311530.

[86] M. Mohammadi and M. R. Mosavi, "Comparison of two methods of removing EOG artifacts for use in a motor imagery-based brain computer interface", *Evolving Systems*, vol. 12, no. 2, pp. 527–540, 2021. DOI: 10.1007/s12530-019-09311-7.

[87] K. K. Ang, Z. Y. Chin, H. Zhang, and C. Guan, "Filter Bank Common Spatial Pattern (FBCSP) in Brain-Computer Interface", in *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, 2008, pp. 2390–2397. DOI: 10.1109/IJCNN.2008.4634130.

[88] C. Gouy-Pailler, M. Congedo, C. Brunner, C. Jutten, and G. Pfurtscheller, "Non-stationary brain source separation for multiclass motor imagery", *IEEE transactions on bio-medical engineering*, vol. 57, no. 2, pp. 469–478, 2010. DOI: 10.1109/TBME.2009.2032162.

[89] M. Schreuder, T. Rost, and M. Tangermann, "Listen, You are Writing! Speeding up Online Spelling with a Dynamic Auditory BCI", *Frontiers in Neuroscience*, vol. 5, 2011. DOI: 10.3389/fnins.2011.00112.

[90] B. Blankertz et al., "Neurophysiological predictor of SMR-based BCI performance", *NeuroImage*, vol. 51, no. 4, pp. 1303–1309, 2010. DOI: 10.1016/j.neuroimage.2010.03.022.

[91] M. Iqbal, M. M. Rahman, and S. Shubha, *Effect of EOG Artifact Removal on EEG Motor-Imagery Classification*. IEEE, 2022, vol. 25th International Conference on Computer and Information Technology (ICCIT).

[92] E. B. Assi, S. Rihana, and M. Sawan, "33% Classification Accuracy Improvement in a Motor Imagery Brain Computer Interface", *Journal of Biomedical Science and Engineering*, vol. 10, no. 6, pp. 326–341, 2017. DOI: 10.4236/jbise.2017.106025.

[93] D. E. Thompson, M. R. Mowla, K. J. Dhuyvetter, J. W. Tillman, and J. E. Huggins, "Automated Artifact Rejection Algorithms Harm P3 Speller Brain-Computer Interface Performance", *Brain Computer Interfaces (Abingdon, England)*, vol. 6, no. 4, pp. 141–148, 2019. DOI: 10.1080/2326263X.2020.1734401.

[94] L. Frølich, I. Winkler, K.-R. Müller, and W. Samek, "Investigating effects of different artefact types on motor imagery BCI", in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2015, pp. 1942–1945. DOI: 10.1109/EMBC.2015.7318764.

[95] M. M. N. Mannan, M. A. Kamran, S. Kang, and M. Y. Jeong, "Effect of EOG Signal Filtering on the Removal of Ocular Artifacts and EEG-Based Brain-Computer Interface: A Comprehensive Study", *Complexity*, vol. 2018, e4853741, 2018. DOI: 10.1155/2018/4853741.

[96] P. Merinov, M. Belyaev, and E. Krivov, "The comparison of automatic artifact removal methods with robust classification strategies in terms of EEG classification accuracy", in *2015 International Conference on Biomedical Engineering and Computational Technologies (SIBIRCON)*, 2015, pp. 221–224. DOI: 10.1109/SIBIRCON.2015.7361887.

[97] M. N. van Stigt, C. R. Camps, J. M. Coutinho, H. A. Marquering, B. S. Doelkahar, and W. V. Potters, "The effect of artifact rejection on the performance of a convolutional neural network based algorithm for binary EEG data classification", *Biomedical Signal Processing and Control*, vol. 85, p. 105 032, 2023. DOI: 10.1016/j.bspc.2023.105032.

[98] Y. Chen, C. Zhang, and X. Wu, "To Assess the Influence of Artifacts on Motor Imagery Based BCI", in *2019 IEEE 4th International Conference on Signal and Image Processing (ICSIP)*, 2019, pp. 925–929. DOI: 10.1109/SIPROCESS.2019.8868896.

[99] M. K. Islam, P. Ghorbanzadeh, and A. Rastegarnia, "Probability mapping based artifact detection and removal from single-channel EEG signals for brain–computer interface applications", *Journal of Neuroscience Methods*, vol. 360, p. 109 249, 2021. DOI: 10.1016/j.jneumeth.2021.109249.

[100] M. Anjum, N. Sakib, and M. K. Islam, "Effect of artifact removal on EEG based motor imagery BCI applications", SPIE, 2024, p. 9. DOI: 10.1117/12.3020879.

[101] A. Al-Saegh, S. A. Dawwd, and J. M. Abdul-Jabbar, "Deep learning for motor imagery EEG-based classification: A review", *Biomedical Signal Processing and Control*, vol. 63, p. 102 172, 2021. DOI: 10.1016/j.bspc.2020.102172.

[102] R. Salazar-Varas and R. A. Vazquez, "Evaluating the effect of the cutoff frequencies during the pre-processing stage of motor imagery EEG signals classification", *Biomedical Signal Processing and Control*, vol. 54, p. 101 592, 2019. DOI: 10.1016/j.bspc.2019.101592.

[103] M. Benomar, S. Cao, M. Vishwanath, K. Vo, and H. Cao, "Investigation of EEG-Based Biometric Identification Using State-of-the-Art Neural Architectures on a Real-Time Raspberry Pi-Based System", *Sensors*, vol. 22, no. 23, p. 9547, 2022. DOI: 10.3390/s22239547.

[104] H. Dose, J. S. Møller, H. K. Iversen, and S. Puthusserypady, "An end-to-end deep learning approach to MI-EEG signal classification for BCIs", *Expert Systems with Applications*, vol. 114, pp. 532–542, 2018. DOI: 10.1016/j.eswa.2018.08.031.

[105] X. Zhao, H. Zhang, G. Zhu, F. You, S. Kuang, and L. Sun, "A multi-branch 3d convolutional neural network for EEG-based motor imagery classification", *IEEE transactions on neural systems and rehabilitation engineering*, vol. 27, no. 10, pp. 2164–2177, 2019.

[106] T. Liu and D. Yang, "A densely connected multi-branch 3d convolutional neural network for motor imagery EEG decoding", *Brain Sciences*, vol. 11, no. 2, p. 197, 2021.

[107] L. Yang, Y. Song, X. Jia, K. Ma, and L. Xie, "Two-branch 3d convolutional neural network for motor imagery EEG decoding", *Journal of Neural Engineering*, vol. 18, no. 4, p. 0460c7, 2021. DOI: 10.1088/1741-2552/ac17d6.

[108] L. Hedegaard and A. Iosifidis, "Continual 3d Convolutional Neural Networks for Real-time Processing of Videos", in *Computer Vision – ECCV 2022*, S. Avidan, G. Brostow, M. Cissé, G. M. Farinella, and T. Hassner, Eds., Springer Nature Switzerland, Cham, 2022, pp. 369–385. DOI: 10.1007/978-3-031-19772-7_22.

[109] E. S. Salama, R. A. El-Khoribi, M. E. Shoman, and M. A. W. Shalaby, "Eeg-based emotion recognition using 3d convolutional neural networks", *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 8, 2018.

[110] Z. Wan, R. Yang, M. Huang, N. Zeng, and X. Liu, "A review on transfer learning in EEG signal analysis", *Neurocomputing*, vol. 421, pp. 1–14, 2021. DOI: 10.1016/j.neucom.2020.09.017.

[111] C. Tan, F. Sun, and W. Zhang, "Deep Transfer Learning for EEG-Based Brain Computer Interface", in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2018, pp. 916–920. DOI: 10.1109/ICASSP.2018.8462115.

[112] F. Mattioli, C. Porcaro, and G. Baldassarre, "A 1d CNN for high accuracy classification and transfer learning in motor imagery EEG-based brain-computer interface", *Journal of Neural Engineering*, vol. 18, no. 6, p. 066 053, 2022. DOI: 10.1088/1741-2552/ac4430.

[113] J. Walker, "Filters", *IEE Colloquium on How to Design RF Circuits (Ref. No.00/027)*, Conferences, pp. 5/1–5/16, DOI: 10.1049/ic:20000144.

[114] D. J. McFarland, L. A. Miner, T. M. Vaughan, and J. R. Wolpaw, "Mu and Beta Rhythm Topographies During Motor Imagery and Actual Movements", *Brain Topography*, vol. 12, no. 3, pp. 177–186, 2000. DOI: 10.1023/A:1023437823106.

[115] S. Xu, Y. Liu, H. Lee, and W. Li, "Neural interfaces: Bridging the brain to the world beyond healthcare", *Exploration*, vol. 4, no. 5, p. 20 230 146, 2024. DOI: 10.1002/EXP.20230146.