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

ADOLF András

Theses of the PhD Dissertation

Supervisor: Dr. ULBERT István, DSc

Pázmány Péter Catholic University

Roska Tamás Doctoral School of Sciences and Technology

Budapest, 2025

1 Introduction

BCI systems are emerging as transformative tools in assistive technologies, offering individuals with severe motor disabilities a new channel for communication and interaction [1]. 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 [2].

While online artifact rejection methods such as FORCe (Fully Online and Automated Artifact Removal for Brain-Computer Interfacing) [3] and ORICA (Online Recursive Independent Component Analysis) [4] 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 [5] 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.

2 Methods and Tools

2.1 The Online FASTER algorithm

I designed and implemented an artifact removal procedure capable of quasi-real-time filtering, which I named the Online FASTER algorithm. The procedure is based on the widely used and validated FASTER algorithm [5], which I adapted for online operation. The functioning of the algorithm consists of an offline preparatory phase and further online filtering steps.

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.

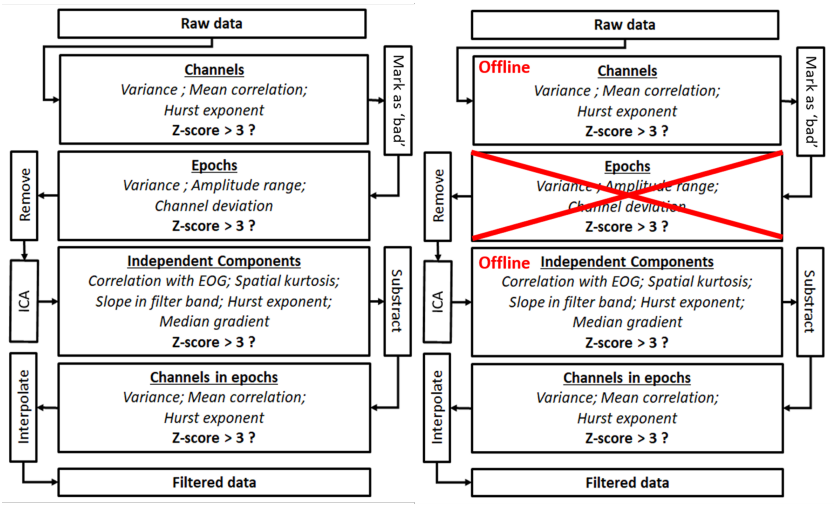

Figure 1: The algorithm of the Offline FASTER method (left), with the designed Online FASTER method (right). In the proposed online algorithm, the identification and exclusion of globally poor-quality channels are conducted during the offline prefiltering phase, which precedes the online filtering process. 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 computation of ICA components. 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.

Following 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/`.

2.2 The Physionet database

My research was performed on the EEG motor movement/imagery dataset recorded by Schalk et al. as part of the Physionet Database [6]. Data was recorded with a 64-channel 10-10 EEG system with 160 Hz sampling frequency and by using the BCI2000 framework, without hardware filters [7]. 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. The dataset is accessible at `https://physionet.org/content/eegmmidb/1.0.0/`.

2.3 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 [8], [9]. For this transformation, I used a unique arrangement of electrode placements, referred to as a 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 2.

2.4 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 over

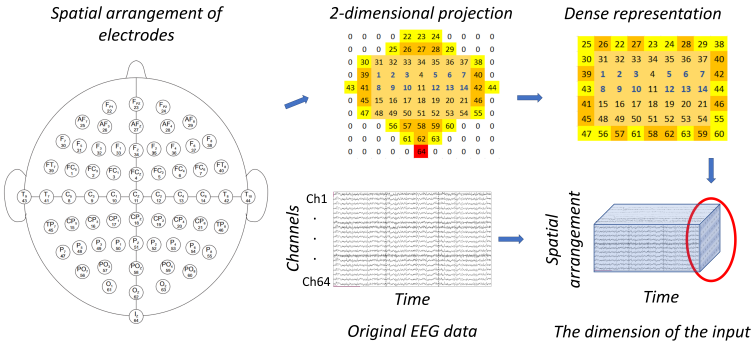

Figure 2: 3D representation of the EEG data: the channels x time initial arrangement of the EEG data is rearranged into a 3-dimensional form, where the first two dimensions refer to the spatial arrangement of the electrodes, while the third dimension is time. For the spatial dimension, I used a dense electrode arrangement: while the original distribution contains placeholder zero channels, I rearranged this map to have a dense representation with no such cells [J2].

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 soft max output layer produces class membership probabilities, enabling multiclass classification. A schematic representation of the complete architecture is provided in Figure 3.

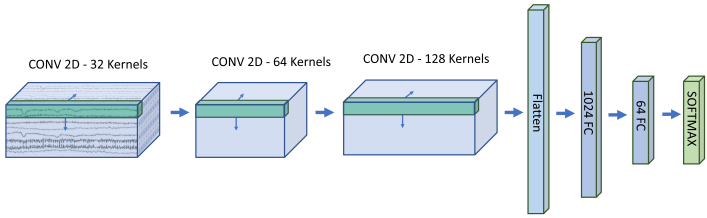

Figure 3: The structure of the designed 2D convolutional neural network: The input is the 3D representation of the EEG signal, on which 2D convolution is performed with 32 kernels of the size of [3 x 3 x number of timepoints] in the first layer. In the next two layers of convolution, firstly 64, then 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].

2.5 Conv3D Net

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 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.

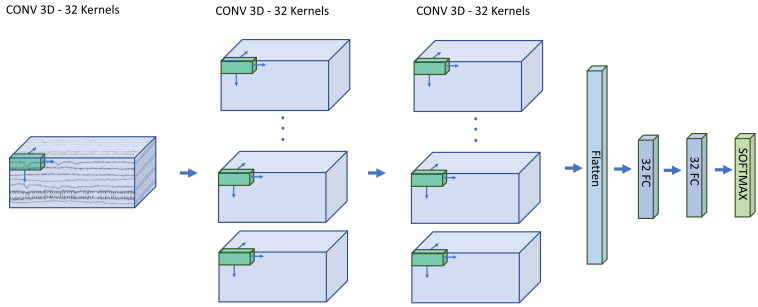

Figure 4: The structure of the 3D convolutional network. The input is identical to that of the 2D CNN, t;e difference is that in this structure, 3-dimensional convolution is performed. There are 32 kernels in all the layers, the shape is (1 x 1 x 30) in the first, and (2 x 2 x 40) in the second and third layers. The fully connected part consists of two layers with 32 neurons each, and finally a softmax classification layer [J2].

2.6 Additional Implemented Neural Networks

In addition to the networks I designed, I also implemented other machine learning architectures, among which EEGNet and Shallow ConvNet are widely used structures.

2.7 EEGNet

The first examined network is one of the most well-known systems used for classifying motor imagery (MI) EEG signals, EEGNet, which was introduced by V. J. Lawhern et al. [10]. The length of the convolutional kernel in the first layer was set to half of the sampling frequency, and no further modifications were applied. The main layers of the network are the following: a two-dimensional convolutional layer, a depthwise two-dimensional convolutional layer, a separable two-dimensional convolutional layer, and finally, a fully connected classification layer.

2.8 Shallow ConvNet

I also implemented the Shallow ConvNet architecture published by Schirrmeister et al. [11], which has been successfully applied in several studies for the classification of imagined movements. The network consists of a temporal convolutional layer, a spatial convolutional layer, a mean pooling layer, and a linear classification fully connected layer.

2.9 Multi Branch Conv3D Net

A network using 3D convolution, similar to the one I designed, was implemented based on the article by Xinqiao Zhao et al. [8]. The architecture consists of three parallel branches, each containing two layers with convolutional kernels of different sizes, followed by three fully connected layers. The size of the final layers equals the number of classes; thereafter, the outputs of the three branches are summed, and the final classification is performed using a softmax operation.

2.10 Transfer learning and finetuning

A deep learning system generally requires a vast amount of training data in order to generalize features effectively and achieve adequate accuracy. In the case of EEG data, such a large dataset is usually not available, as data collection is extremely time- and labor-intensive. The core idea of Transfer Learning (TL) is to pretrain the system, or parts of it, on data independent of the target dataset, and then fine-tune the pre-initialized weights using the actual dataset.

There are two widespread approaches to transfer learning. In the first approach, the EEG signal is transformed into a representation for which large datasets or pretrained networks already exist. A typical example is ImageNet, where we generate 3-channel images from EEG data as input to the network. 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 than the target individual. In other words, networks pretrained on multiple subjects are then fine-tuned for each individual.

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 5.

2.11 Investigation of the Effect of 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

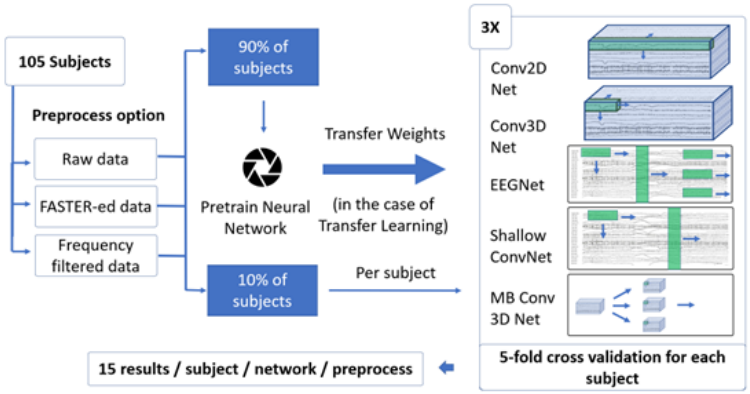

Figure 5: Transfer learning and the generation of results. To obtain statistically significant results, I applied 5-fold cross-validation for each subject. This resulted in a total of 15 measurements per subject, per network, and per preprocessing step. In the case of applying transfer learning, the network was pretrained using data from 90% of the subjects, and then fine-tuned on the training data of the remaining 10% of subjects. The entire process was repeated ten times so that each subject was used as a test subject at least once [J2].

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.

2.12 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. [11]. 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.

2.13 Code availability

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.

The complete codebase supporting this research - including the designed and tested networks, transfer learning, cropped training, artifact rejection, and all the further mentioned methods - is publicly available at `https://dev.itk.ppke.hu/adoan/dissertation-codebase`.

3 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.***

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. A detailed description of the algorithm can be found in the first chapter of the *Methods and Methodologies* section.

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.***

Cybathlon is an international competition held every four years, where, in the brain-computer interface section, pilots with physical disabilities - delegated by different countries - are required to control a computer game using only their thoughts. The use of non-brain-derived signals for control is strictly prohibited, making a proper, real-time artifact removal algorithm especially crucial. 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.***

I showed that the impact of artifact rejection varies depending on the individual subject being analyzed. My analysis encompassed a compre-

hensive 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 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.

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 6.

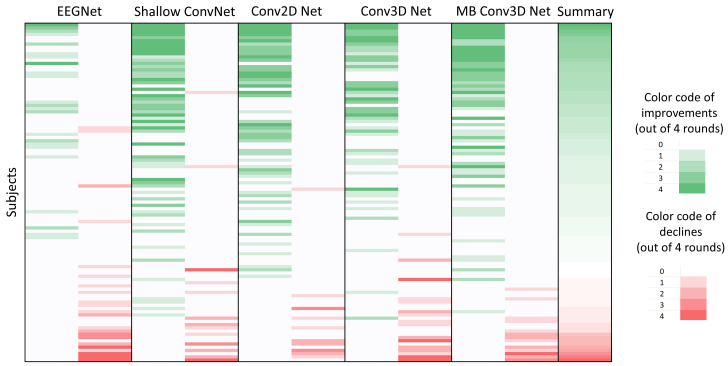

Figure 6: Subject-dependence of artifact filtering. The color shades indicate how many runs resulted in statistically significant improvement (green) or deterioration (red). (The darker the shade, the more runs showed a significant difference.) Statistical significance was determined using a t-test for normally distributed data, and the Mann–Whitney U test otherwise. Subjects were ranked based on a composite score calculated as the difference between the number of significantly improved runs and the number of significantly deteriorated runs [J2].

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.***

I demonstrated that the classification accuracy obtained by transfer learning is significantly better in every scenario. Based on the signed-rank Wilcoxon test, the learning process performs significantly better in the case of each network, both in the case of unfiltered data (Table 1) and artifact-rejected data (Table 2). In Figure 7, 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.) 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.

Table 1: 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 2: 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

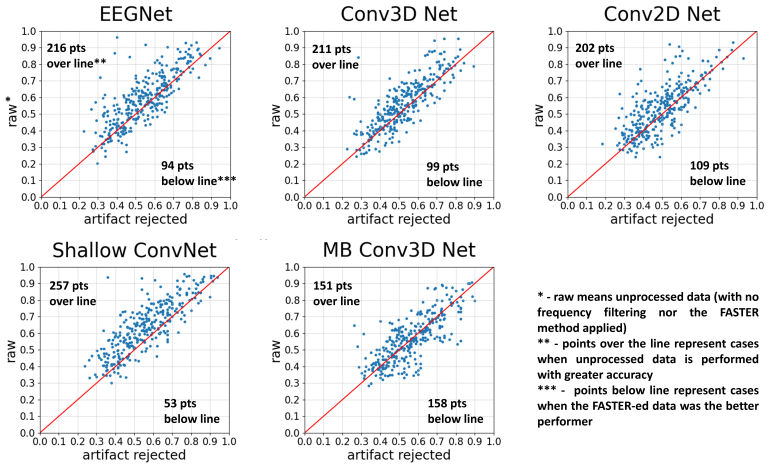

Figure 7: The comparison of neural networks with and without using the FASTER method, using transfer learning. Each point represents two accuracies for a subject: the one obtained with AR and the other obtained without using the FASTER algorithm. The test was run 3 times. The red line indicates the points 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].

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.***

I compared the five examined neural networks using classification accuracy as the primary evaluation criterion. As illustrated in Figure 8, when transfer learning is not employed, the 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.

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.

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. 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.

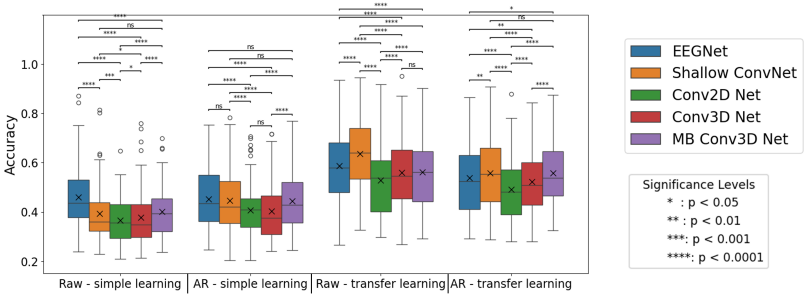

Figure 8: Classification accuracy of different neural networks with and without transfer learning, and with and without artifact rejection. In the boxplot, "x" denotes the sample mean, and the horizontal line denotes the median. The box boundaries represent the interquartile range (middle 50% of the data). Before applying transfer learning, EEGNet achieved the highest performance, while after transfer learning, Shallow ConvNet proved most effective. The figure shows that networks that adaptively utilize the relative spatial arrangement of channels (EEGNet and Shallow ConvNet) outperform their counterparts that rely on explicit spatial modeling (Conv2D, Conv3D, and MB Conv3D) [J2].

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.***

I showed that the outcomes stemming from the frequency filtering analysis yield unexpected findings. I compared results using the 0.1 to 5 Hz range and the 5 to 75 Hz range. As observed in Table 3, 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 3: 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

As illustrated in Table 4, 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 4: 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.***

I showed that, as seen in the first two subplots of Figure 9, the order of subjects is nearly identical across the two experimental results, which suggests that frequency filtering has a similar subject-specific effect as the FASTER algorithm compared to the raw data. However, this ranking differs from the results presented in the first thesis point, meaning that FASTER impacts different subjects depending on whether cropped or full-segment training is applied.

I observed that when comparing FASTER to simple frequency filtering, the subject ranking changes again: for most participants, FASTER led to significant performance degradation. This suggests that, under cropped training conditions, applying only band-specific filtering is preferable to the more complex FASTER procedure for achieving better results.

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

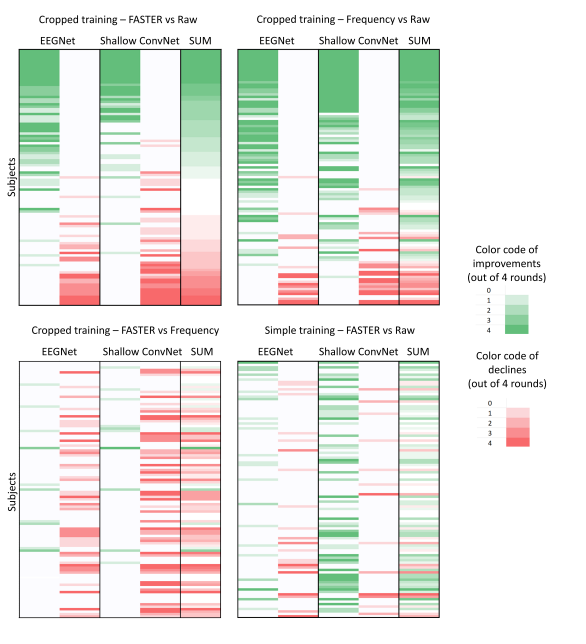

Figure 9: Subject-wise effect of the FASTER artifact rejection method and the frequency filtering on the classification performance of the EEGNet and Shallow ConvNet, related to the raw data, during cropped and simple training. The shade of the color means the number of significant increases (green) or declines (red) out of the 4 tests. Subjects are ordered by the score achieved during the AR method compared to the unfiltered data [J2].

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

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

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 in the writing process.

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.

Publications of the author

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.

References

[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] 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.

[3] 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.

[4] 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.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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.