# Finding the Optimal Time Window for Increased Classification Accuracy during Motor Imagery


D. A. Blanco-Mora[1][a], A. Aldridge[1][b], C. Jorge[1][c], A. Vourvopoulos[3][d], P. Figueiredo[3][e] and S. Bermúdez i Badia[1,2][f]

[1] *Madeira Interactive Techonologies Institute, Universidade da Madeira, Funchal, Portugal*

[2] *Faculdade de Ciências Exatas e da Engenharia, Universidade da Madeira, Funchal, Portugal*

[3] *Institute for Systems and Robotics - Lisboa, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal*



Keywords: Brain-computer Interface, BCI, Motor Imagery, MI, Classification Accuracy, Common Spatial Pattern, CSP, Electroencephalography, EEG, Neurorehabilitation, Stroke.



Abstract: Motor imagery classification using electroencephalography is based on feature extraction over a length of time, and different configurations of settings can alter the performance of a classifier. Nevertheless, there is a lack of standardized settings for motor imagery classification. This work analyzes the effect of age on motor imagery training performance for two common spatial pattern-based classifier pipelines and various configurations of timing parameters, such as epochs, windows, and offsets. Results showed significant ($p \leq 0.01$) inverse correlations between performance and feature quantity, as well as between performance and epoch/window ratio.


## 1 INTRODUCTION

In recent decades, Brain-Computer Interfaces (BCIs) have been used in novel neurorehabilitative techniques, yielding promising results in terms of motor recovery (Cervera et al., 2018). The use of BCIs in neurorehabilitation allows the recruitment and activation of motor regions through Motor Imagery (MI), without the need of active movement. This could potentially result in neuro-plasticity changes in areas considered to be damaged from a stroke (Bai et al., 2020). When combined with serious games and Virtual Reality (VR), BCIs allow for a more intensive neurorehabilitation (Putze, 2019), encouraging patients via immediate feedback (Mubin et al., 2020) and immersing them in engaging, virtual environments (Khan et al., 2020). One important challenge with neurorehabilitation is the timing and efficacy of feedback delivery relating to MI. Feedback delivery plays an important role in BCIs, and efficiently using proprioceptive feedback can improve BCI performance significantly (Ramos-Murguialday et al., 2019). During learning, feedback provided to close the sensorimotor loop should be associated with a MI event to facilitate the recreation of a more realistic and authentic sensation. Similarly, feedback provided during online sessions should be delivered as soon as possible in response to a MI event.

The practice of detecting and classifying MI can be challenging due to the variability and uniqueness of each EEG signal. According to (Ortner et al., 2015), 35% of participants obtained a classification accuracy lower than 70% with only 65% of people controlling MI-based BCI adequately ($\geq$ 70%). The inability to replicate a standard pattern or brain network topology, known as illiteracy, has a negative impact on MI performance classification (Ahn and Jun, 2015). To face this challenge, researchers use strategies for the recruitment of participants to help bypass BCI illiteracy (Ahn et al., 2013), but this is not an option when using BCI as a neurorehabilitative tool for stroke patients, because the impaired neural pathways prevent the use of screening strategies based on BCI illiteracy, and the use of screening strategies will restrict the usability impact to a particular group with a certain type of lesion.


[a] https://orcid.org/0000-0003-2232-0999
[b] https://orcid.org/0000-0003-3733-4736
[c] https://orcid.org/0000-0002-7693-7292
[d] https://orcid.org/0000-0001-9676-8599
[e] https://orcid.org/0000-0002-0743-0869
[f] https://orcid.org/0000-0003-4452-0414





Blanco-Mora, D., Aldridge, A., Jorge, C., Vourvopoulos, A., Figueiredo, P. and Bermúdez i Badia, S.
Finding the Optimal Time Window for Increased Classification Accuracy during Motor Imagery.
DOI: 10.5220/0010316101440151
In *Proceedings of the 14th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2021) - Volume 1: BIODEVICES*, pages 144-151
ISBN: 978-989-758-490-9

Another way researchers are tackling the challenges of MI detection and classification is by using machine learning tools to create methodologies or pipelines that can be generalized across all users. The most common methodologies use frequency features or sensory motor rhythms (SMR), including event-related synchronization/desynchronization (ERS/ERD) (Padfield et al., 2019). Common Spatial Patterns (CSPs) algorithm is a feature extraction method that can learn spatial filters maximizing the discriminability of two classes. CSP filters are widely used to reduce data dimensionality for lighter data processing and maximize the discriminability of classes, for instance, right versus left arm movement or arm versus leg movement. They also express data as a combination of a reduced number of components (Ai et al., 2018). Combined with linear discriminant analysis (LDA), CSP and LDA, are the preferred pipeline and especially useful during online rehabilitation sessions (Lotte et al., 2018).

Even with improved techniques and methodologies, working with MI is challenging because MI classification is difficult to generalize and to improve. For instance, studies have performed different time settings for signal analyses (Chen et al., 2020; Wang et al., 2020), and it is known that ERS/ERD could last for several seconds following the stimulus presentation (Pfurtscheller and Lopes, 1999). Hence, there is a lack of standardized settings for MI due to the variability seen across subjects (Wang et al., 2020). Most BCI research is done with healthy, young participants whereas the stroke population tends to be elderly or aged. Therefore, we designed a study to investigate if our findings generalize to an older population.

In the context of a serious game developed for motor rehabilitation based on MI, we studied the effect of age on MI signals and identified the optimal pipeline and time parameter settings in comparison to a control group of healthy individuals.

# 2 METHODS

To find the optimal configuration and time parameter settings, we have designed a set of different configurations and tested them with NeuRow, a BCI training paradigm in VR, designed for upper-limb motor rehabilitation (Vourvopoulos et al., 2016). The following subsections describe the software, equipment, data acquisition, and processing pipelines we used for this study.

## 2.1 NeuRow

NeuRow uses MI to control avatar movement and haptic feedback in order to increase the sense of embodiment of the user in a closed visual and sensorimotor loop (Vourvopoulos et al., 2019). With the use of a head-mounted or a desk display, the game shows the player in a first-person perspective as an avatar in a kayak floating on a body of water. NeuRow consists of two stages: training and real-time control. The training stage uses signal processing and signal features to train a system to separate EEG signals according to desired events, specifically right and left hand MI. The MI training stage is an adaptation of the Graz paradigm, using directional arrows (Pfurtscheller et al., 2003). Participants perform MI while observing the NeuRow avatar row to the left or right for a total of 20 trials per side, following the training block paradigm shown in Fig. 1.

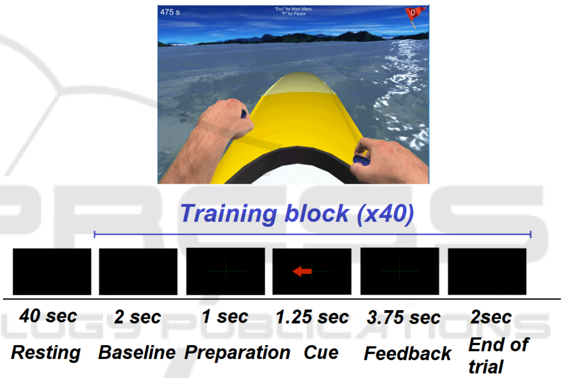


Figure 1: Training Block paradigm.

## 2.2 Setup

The full setup includes a desktop computer, the NeuRow game, an Oculus Rift VR system, the EEG system, and the custom-made haptic feedback system described in (Vourvopoulos et al., 2016), as is shown in Fig. 2. OpenVibe (Renard et al., 2010) was used to translate MI events into game commands for controlling the NeuRow avatar via a Virtual-Reality Peripheral Network. Four stages constitute MI translation: data acquisition, spatial filter training, classifier training, and online game-play.

## 2.3 Participants

All participants are healthy with no known neurological clinical history. The participants have been recruited based on their motivation to participate in the study and divided in two groups: the control group and the age-matched group. To study the applicability

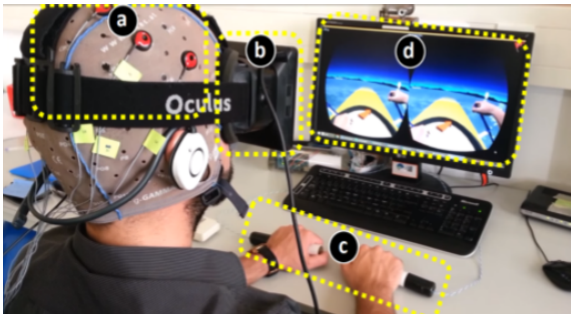


Figure 2: NeuRow setup: A) EEG system, B) The Oculus Rift DK1 Head-Mounted Display, C) Haptic feedback system, D) NeuRow BCI-VR task (Vourvopoulos, 2018).

of our findings for stroke participants and under the approval of the ethical committee at the Hospital of Madeira, Nélio Mendonça (SESARAM), the spouses of stroke patients were recruited as age-matched participants to match the typical age of stroke survivors. The control group consists of six young, healthy participants with an average age of 25 $\pm$5.33 years, 5 males and 1 female. The age-matched group consists of 5 females and 1 male with an average age of 51.33 $\pm$4.97 years. All but one of the participants (one from the control group) are right-handed according to the Edinburgh inventory (Oldfield, 1971).

## 2.4 Data Acquisition

EEG data was acquired using OpenVibe's acquisition server and 32 EEG channels with a sampling frequency of 500Hz and are downsampled to 250Hz. The control participants' datasets are acquired with Liveamp 32 EEG amplifier (Brain Products GmbH, Munich, Germany) at IST in Lisbon, Portugal. The age-matched participants' datasets was acquired with a wireless g.Nautilus from g.tec at the Madeira Interactive Technologies Institute in Funchal, Portugal. Participants went through 20 trials per side doing MI while watching the avatar. Common electrodes for both groups: FC5, F3, Fz, F4, FC6, C3, Cz, C4, CP5, CP1, CP2, CP6, P3, Pz, and P4. Age-matched group included PO3, PO4, C1, and C2. Peripheral electrodes (Fp1, Fp2, O1, O2, among others) are excluded from data processing to reduce the presence of artifacts. The electrode impedances are checked using the g.needAccess software from g.tec (g.tec medical engineering GmbH, Austria).

## 2.5 Processing Pipeline

In the following, the term pipeline, as outlined in this paper, refers to the processing structure used in the classifier training scenario. Two different processing pipelines, both with the same LDA classifier type, are used for comparing classifier performance in OpenVibe. These are referred to as 2-CSP-Boxcar and 4-CSP-Hamming:

- The 2-CSP-Boxcar pipeline, applied in (Suryotrisongko and Samopa, 2015), has a classifier that is based on the power signal of a two-dimensional CSP filter. It extracts power from both output signals of the CSP filter during the sliding window period, according to Equation (1) in which $x$ represents an output signal from the CSP filter.

$$2-CSP-Boxcar_{feature} = log(1+x^2) \tag{1}$$

- The 4-CSP-Hamming pipeline, explained in (Müller-Gerking et al., 1999) and applied in (Irimia et al., 2018), uses the variance contribution of each signal from a four-dimensional CSP filter. It extracts the contribution of each of the four CSP output signals, as shown in Equation (2) (Irimia et al., 2018), with sub-index $p$ representing each output signal.

$$Hamming_{feature} = log\left(\frac{VAR_p}{\sum_{p=1}^{4} VAR_p}\right) \tag{2}$$

EEG data processing for the two pipelines is executed using CSP filter training and classifier training scenarios in OpenVibe. Raw EEG datasets are filtered from 8 to 30 Hz (including the Alpha and Beta bands) in the respective CSP filter training scenarios and then spatially filtered using the previously trained CSP filter for both pipelines. The CSP filter reduces the signal quantity obtained from the number of the EEG channels to a fixed number of representative outputs, and it acquires weights for the EEG channels' signals according to their contribution to each event. Each EEG channel's signal is expressed as a linear combination of the representative outputs. The use of frequency filters limits the signal content to frequencies of interest, helping to denoise the desired signal and eliminate potential constant offset, linear trending, or noise that is caused by the power supply (50/60 Hz) present in the signal.

The 2-CSP-Boxcar pipeline uses the filtered data to train the CSP filter to yield two output signals. The 4-CSP-Hamming pipeline, however, yields 4 output signals. After the CSP filtering, left and right MI trials are windowed either with a Boxcar window or a Hamming window, respectively. Next, EEG feature extraction is applied to obtain the power and variance characteristics to represent the signal. Finally, the extracted features from both pipelines are used to train a LDA classifier to identify right and left MI trials.

### 2.6 Pipeline Settings

To identify the optimal configuration of the pipelines described above, we study a number of settings. As shown in Fig. 3, there are four main settings to consider for the proposed pipelines: epoch length, epoch offset time, sliding window length, and sliding window interval. Here, epoch length is the length of the signal used for feature extraction; epoch offset time determines the start of an epoch; sliding window length is the length of the window that slides across the selected epoch; and sliding window interval (moving steps) is the time between the start of each sliding window. Sliding window displacement does not exist when the sliding window has the same length as the epoch. Relations between the sliding window length and epoch length determine the quantity of data points of the calculated or extracted EEG features as power and variance contribution.

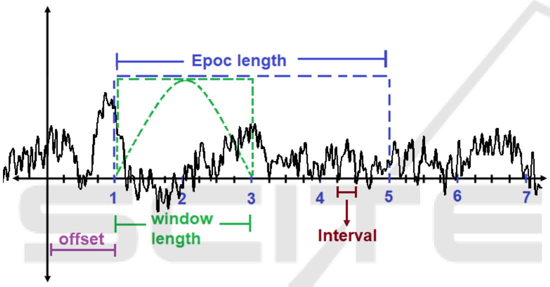


Figure 3: Parameters of a sliding window.

### 2.7 Data Analysis

The acquired EEG datasets are used to train the 2-CSP-boxcar and 4-CSP-Hamming pipelines for various combinations of time parameters. The classifier's accuracy performance for each configuration is computed using 5-fold cross-validation, as the average of the positive predictive value and the negative predictive value from the confusion matrix's classifier performance results. The association between the number of features used and the resulting accuracy performance is tested using Pearson's correlation. The same is done for the relation between accuracy performance and epoch/window ratios. Double-tailed t-tests are used to test for significant results.

The number of features' data points (from now on is shortened as only features) that feed the classifier depends on the epoch lengths, the window lengths, the number of trials, and the number output channels from the CSP filter, as determined by Equation (3).

$$n_{features} = \left\| \frac{l}{p}\left(\frac{w}{l}+1\right)+1 \right\| * trials * CSP_{outputs} \quad (3)$$

In Equation (3), $l$ is epoch length, $w$ stands for sliding window length, $p$ is sliding window interval or moving steps, $\|\ \|$ denotes floor function, * means product, *trials* represents the number of executed trials during MI, and $CSP_{outputs}$ represents the number of CSP signals.

## 3 RESULTS

To study the effects of age on MI and how pipeline settings affect classifier performance, the results section is broken down into: an analysis of features and their corresponding classification performance, pipeline performance analysis.

### 3.1 Features and Classifier Performance

For the analysis of this section, we consider the following variables: pipeline (2-CSP-Boxcar and 4-CSP-Hamming), epoch length (1, 2, 4 s), sliding window length (0.5, 1, 2, 4 s), offset (0.1, 0.5 s), and number of features (Equation (3)), keeping the sliding window time interval as 50 ms. The combination of these parameters totals 36 different configurations to study. Fig. 4 shows the relationship between number of features and the epoch and sliding window lengths. The number of features is constant where the epoch and sliding window lengths are equal. Hereafter, these configurations are mentioned as epoch-lagged because they add a time lag. The number of features increases for longer epoch lengths and shorter sliding window lengths. Configurations that have longer sliding window lengths than epoch lengths are indicated as NaN as it is not possible to have a negative quantity of features.

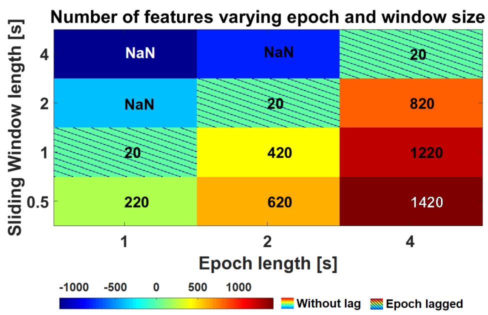


Figure 4: Number of features per CSP output signal ($CSP_{outputs} = 1$) for varying epoch and sliding window lengths (diagonal lines represent epoch-lagged configurations).

The relationship between number of features and classifier performance was tested for correlation, finding Inverse relations between number of features and classifier performance for both pipelines. The analysis revealed correlations for all configurations to be between -0.73 and -0.95 (Table 1). Interestingly, we found a higher correlation between the number of features for the age-matched group than for the control one of about 8%. When comparing the contributions of different pipeline settings (epoch length and sliding window length) to the classification performance, the correlation analysis revealed that the epoch length/sliding window length ratio highly correlates with the quantity of used features (r=0.8712, p=0.002) and performance (-0.77 to -0.98) (Table 2). Consistent with the previous data, correlations are generally stronger for age-matched controls.

Table 1: Correlation values between classifier performance and quantity of features.

| r-value | Offset = 0.1 s | | Offset = 0.5 s | |
|---|---|---|---|---|
| | *Boxcar* | *Hamming* | *Boxcar* | *Hamming* |
| Age-matched | $-0.94^{**}$ | $-0.93^{**}$ | $-0.95^{**}$ | $-0.94^{**}$ |
| Controls | $-0.86^{**}$ | $-0.85^{**}$ | $-0.73^{*}$ | $-0.86^{**}$ |

$*p<\mathbf{0.05}$; $**p<\mathbf{0.01}$

Table 2: Correlation values between classifier performance and epoch/window length ratio.

| r-value | Offset = 0.1 s | | Offset = 0.5 s | |
|---|---|---|---|---|
| | *Boxcar* | *Hamming* | *Boxcar* | *Hamming* |
| Age-matched | $-0.98^{**}$ | $-0.93^{**}$ | $-0.91^{**}$ | $-0.88^{**}$ |
| Controls | $-0.92^{**}$ | $-0.77^{*}$ | $-0.82^{**}$ | $-0.92^{**}$ |

$*p<\mathbf{0.05}$; $**p<\mathbf{0.01}$

## 3.2 Pipeline Performance Analysis

### 3.2.1 Age-matched Participants

The age-matched group's performance results show the highest accuracy percentage of 85% for the 4-CSP-Hamming pipeline (offset of 0.5 s, epoch and window lengths of 1 s) (Fig. 5). The best results for all configurations were achieved with equivalent epoch and sliding window lengths, which as previously explained, are the configurations that add a significant time lag (1 to 4 s), making them unsuitable for applications that rely on real-time feedback. When we consider only the non-epoch-lagged configurations that are suitable for real-time feedback, the best performance is seen for the 4-CSP-Hamming pipeline at 79.09% with an offset of 0.1 seconds, an epoch length of 2 seconds, and a window length of 1 second.

Performances for the 2-CSP-Boxcar pipeline were always lower than for the 4-CSP-Hamming pipeline, both for epoch-lagged configurations (76.67% with an offset of 0.5 s and epoch and window lengths of 4 s)

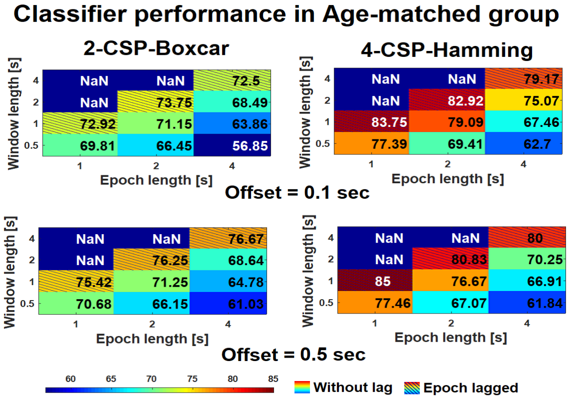


Figure 5: Classifier performance for age-matched group (diagonal lines represent epoch-lagged configurations).

and for non-epoch-lagged ones (71.25% with an offset of 0.5 seconds, epoch length of 2 seconds, and window length of 1 second).

### 3.2.2 Control Participants

Similar to the results of the age-matched group, the control group's best results were for the epoch-lagged configurations. The highest performance was 90% for the 4-CSP-Hamming pipeline when offset equaled 0.1 seconds. The best result for the non-lagged configurations was also obtained with the 4-CSP-Hamming pipeline at 79.68% for an offset of 0.1, an epoch length of 1 second, and a window length of 0.5 seconds (Fig. 6). For the 2-CSP-Boxcar pipeline, the best lagged configuration was recorded at 74.25% when offset equaled 0.5 seconds, and the best non-lagged at 72.41% with 0.1 seconds of offset. Interestingly, there seems to be a tendency for higher accuracies with smaller epoch/sliding window ratios and for equal ratios, a preference for smaller epochs and shorter sliding window lengths.

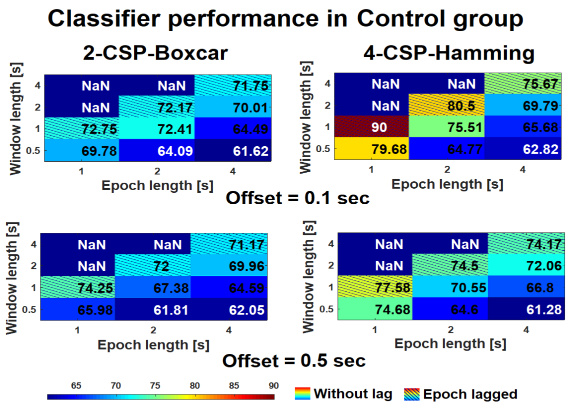


Figure 6: Classifier performance for control group (diagonal lines represent epoch-lagged configurations).

# 4 DISCUSSION

The goal of this study is to shed light on the disparity of pipelines and classification performances reported in MI literature, by performing a systematic comparative study of pipelines and their parameters on two populations. This research is essential for the development of MI-based neurorehabilitation systems as processing pipelines and their specific configurations yield optimal settings that can generalize from age-matched participants to stroke patients. Nevertheless, this study is also relevant for any MI paradigm because it reports on epoch-lagged configurations that are suitable for applications other than neurorehabilitation. Therefore, to understand the behavior and relation of performance with different pipeline configurations and age groups, this section discusses the correlation between feature quantity and classification performance, accuracy performance between pipelines, and the impact of age on performance.

## 4.1 Features and Classifier Performance

The results reveal that classifier performance decreases as epoch and sliding window lengths and their within ratio increase, with significant, negative correlation test results (Table 1, Table 2). These inverse correlations can be related to overfitting (León et al., 2020). The large quantity of features provides the classifier with too much detail and noise, allowing the algorithm to overfit the training data. As part of the training process for the classifier, a portion of the training data is withheld as new data for cross-validation testing. Because of this, we are able to see a negative impact on the classification of unseen data. This might explain why classifier performance decreases when feature quantity increases.

## 4.2 Pipeline Performance

Pipeline performance best results are summarized in Table 3. Epoch-lagged configurations were tested and yielded the highest performances. However, they are inadequate for rehabilitative methods that rely on real-time feedback because they add a time lag equivalent to the epoch and window length. Online, real-time feedback induces more pronounced attention and motor cortical activation; it works as a hybrid control that augments sensory-feedback processing (Ros et al., 2014), improving BCI performance when efficiently delivered (Ramos-Murguialday et al., 2019), and eliminates the need for extensive sessions involving a high number of trials ($> 100$) (Ortner et al., 2015). Therefore, it is desirable to use non-lagged configurations with a window length shorter than the epoch length so that the lag time is as short as the interval length, meaning the classifier output happens at each sliding interval of the sliding window.
The 4-CSP-Hamming pipeline outperformed the 2-CSP-Boxcar pipeline for all of the configurations in both populations. This can be related to the number of CSP output channels with a slightly better performance seen for configurations with 4 CSP channels versus 2 CSP channels (Müller-Gerking et al., 1999). Nevertheless, both pipeline performances are in the adequate control range for MI-based BCIs ($\geq 70\%$) (Ortner et al., 2015).

## 4.3 Impact of Age on Performance

Age matched group obtained similar performance as control group, with no significant differences for all tested configurations. Our results show that using configurations with an epoch/window ratio of 2 obtain similar results to the extended training with high high number of trials (80 per session) and 6 sessions during training, whose maximum performance was 80.7% and minimum was 72.4% (Ortner et al., 2015), and better than 66% for older participants ($72.0 \pm 8.07$ years old) (Chen et al., 2019) with lower SMR lateralization for MI according to aging processes. This is promising for neurorehabilitation adapted for stroke patients because shorter sessions might help alleviate the exhaustiveness of MI training.

## 4.4 Limitations

This work had some limitations such as the low number of participants, equipment differences in both EEG and VR rendering modalities between populations and gender unbalance. Moreover, the existence of pipelines not used in this study, and the applicability of the results for a stroke population. The number of participants in this study, (N=12), might be too low to achieve a high statistical power, but it is greater than the number of participants recruited for BCI performance studies that used datasets from BCI competitions (N=4) and have helped to improve classifier performance. Inconsistencies between the EEG devices and electrode placements might have enhanced the differences between the two populations. Nevertheless, the protocol used to process the data for both groups of participants was the same. Impedance quality was kept on the same level, identical pipelines were used, and non-statistical differences were found between groups. Although it might be interesting to check with several classes of pipelines, doing a larger comparison would be impracticable. The age-

Table 3: Sum-up of results.

| **Configurations** | | **MI-performance** | | |
|---|---|---|---|---|
| | | **Control** | **Age-matched** | **Diff. between-groups** |
| **Epoch-lagged** | 4-CSP-Hamming | ***90*** | *85* | *5* |
| | 2-CSP-Boxcar | *74.25* | *76.67* | *-2.42* |
| | **Diff. between-pipelines** | ***15.75*** | ***8.33*** | ***X*** |
| **Non Epoch-lagged** | 4-CSP-Hamming | ***79.68*** | *79.09* | *0.59* |
| | 2-CSP-Boxcar | *72.41* | *71.25* | *1.16* |
| | **Diff. between-pipelines** | ***17.27*** | *7.84* | ***X*** |
| **Diff. Lagged and non-lagged** | 4-CSP-Hamming | ***10.32*** | *5.91* | *4.41* |
| | 2-CSP-Boxcar | *1.84* | *5.42* | *-3.58* |

matched group is a step closer, but stroke patients may perform differently due to their specific lesion types.

# 5 CONCLUSION

Two populations were recruited for the investigation of age on MI classification, and two CSP-based pipelines were compared to find the optimal pipeline settings for improving classifier performance. Classification performance under multiple pipeline settings for MI in both groups showed no significant effect for age. The results confirmed that the number of features used to train a classifier depends on the epoch-window lengths relation. Classifier performance increased when a smaller epoch/window ratio was applied. However, using an epoch/window ratio equal to one, introduce a lag in time equal to the length of the epoch, which is not desirable for online purposes. By investigating various time parameters that directly affect classification performance, an optimal pipeline configuration was found comprised of the 4-CSP-Hamming pipeline using an epoch/window ratio of 2 and offline of 0.1 seconds for both populations.

## 5.1 Future Work

New datasets will be acquired for the ongoing project, including from stroke participants, increasing the population size and the statistical power of the study. We also plan to compare online classifier performance versus training classifier performance closing the accuracy gap and achieving a more generalizable classifier.

# ACKNOWLEDGEMENTS

This work was supported by by the Portuguese Foundation for Science and Technology through the NeurAugVR FCT project (PTDC/CCI-COM/31485/2017), the NOVA-LINCS-NOVA Laboratory for Computer Science and Informatics (PEest/UID/CEC/04516/2019), the Laboratório de Robótica e Sistemas em Engenharia e Ciência (UIDB/50009/2020) and Scientific Employment Stimulus (CEECIND/01073/2018).

# REFERENCES

Ahn, M., Cho, H., Ahn, S., and Jun, C. (2013). High theta and low alpha powers may be indicative of bci-illiteracy in motor imagery. *PLoS One*, 8(11):1303–1309.

Ahn, M. and Jun, S. C. (2015). Performance variation in motor imagery brain–computer interface: A brief review. *Journal of Neuroscience Methods*, 243:103–110. https://doi.org/10.1016/j.jneumeth.2015.01.033.

Ai, Q., Liu, Q., Meng, W., and Xie, S. Q. (2018). *Advanced Rehabilitative Techonology*, chapter EEG-Based Brain Intention Recognition, pages 135–166. Academic Press. https://doi.org/10.1016/B978-0-12-814597-5.00006-0.

Bai, Z., Fong, K. N. K., Zhang, J. J., Chan, J., and Ting, K. H. (2020). Immediate and long-term effects of bci- based rehabilitation of the upper extremity after stroke: a systematic review and meta- analysis. *Journal of NeuroEngineering and Rehabilitation*, 2:1–20. https://doi.org/10.1186/s12984-020-00686-2.

Cervera, M. A., Soekadar, S. R., Ushiba, J., del R. Millán, J., Liu, M., Birbaumer, N., and Garipelli, G. (2018). Brain-computer interfaces for post-stroke motor rehabilitation: a meta-analysis. *Annals of Clinical and Translational Neurology*, 5(5). https://doi.org/10.1002/acn3.544.

Chen, C., Chen, P., Belkacem, A. N., Lu, L., Xu, R., Tan, W., Li, P., Gao, Q., Shin, D., Wange, C., and Ming, D. (2020). Neural activities classification of left and right finger gestures during motor execution and motor imagery. *Brain-Computer Interfaces*, pages 1–11. https://doi.org/10.1080/2326263X.2020.1782124.

Chen, M. L., Fu, D., Boger, J., and Jiang, N. (2019). Age-related changes in vibro-tactile eeg response and its implications in bci applications: A comparison between older and younger populations. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 27(4). https://doi.org/10.1109/TNSRE.2019.2890968.

Irimia, D. C., Ortner, R., Poboroniuc, M. S., Ignat, B. E., and Guger, C. (2018). High classification accuracy of a motor imagery based brain-computer interface for stroke rehabilitation training. *Frontier in Robotics and AI*, 5(130):9. https://doi.org/10.3389/frobt.2018.00130.

Khan, M. A., Das, R., Iversen, H. K., and Puthusserypady, S. (2020). Review on motor imagery based bci systems for upper limb post-stroke neurorehabilitation: From designing to application. *Computers in Biology and Medicine*, 123. https://doi.org/10.1016/j.compbiomed.2020.103843.

León, J., Escobar, J. J., Ortiz, A., Ortega, J., González, J., Martín-Smith, P., Gan, J. Q., and Damas, M. (2020). Deep learning for eeg-based motor imagery classification: Accuracy-cost trade-off. *Plos One*, 15(6). https://doi.org/10.1371/journal.pone.0234178.

Lotte, F., Bougrain, L., Cichocki, A., Clerc, M., Congedo, M., Rakotomamonjy, A., and Yger, F. (2018). A review of classification algorithms for eeg-based brain–computer interfaces: a 10 year update. *Journal of Neural Engineering*, 15(3):28. https://doi.org/10.1088/1741-2552/aab2f2.

Müller-Gerking, J., Pfurtscheller, G., and Flyvbjerg, H. (1999). Designing optimal spatial filters for single-trial eeg classification in a movement task. *Clincial Neurophysiology*, 5(1):787–798. https://doi.org/10.1016/S1388-2457(98)00038-8.

Mubin, O., Alnajjar, F., Mahmud, A. A., Jishtu, N., and Alsinglawi, B. (2020). Exploring serious games for stroke rehabilitation: a scoping review. *Disability and Rehabilitation: Assistive Technology*, pages 1–7. https://doi.org/10.1080/17483107.2020.1768309.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The edinburgh inventory. *Neuropsychologia*, 9(1):97–113. https://doi.org/10.1016/0028-3932(71)90067-4.

Ortner, R., Scharinger, J., Lechner, A., and Guger, C. (2015). How many people can control a motor imagery based bci using common spatial patterns? In *2015 7th International IEEE/EMBS Conference on Neural Engineering (NER)*, pages 202–205, Montpellier. IEEE. https://doi.org/10.1109/ner.2015.7146595.

Padfield, N., Zabalza, J., Zhao, H., Masero, V., and Ren, J. (2019). Eeg-based brain-computer interfaces using motor-imagery: Techniques and challenges. *Sensors*, 19:1–34. https://doi.org/10.3390/s19061423.

Pfurtscheller, G. and Lopes, F. H. (1999). Event-related eeg/meg synchronization and desynchronization: basic principles. *Clinical Neurophysiology*, 110:1842–1857. https://doi.org/10.1016/S1388-2457(99)00141-8.

Pfurtscheller, G., Neuper, C., Obermaier, B., Krausz, G., Scherer, R., Graimann, B., , and Schrank, C. (2003). Graz-bci: state of the art and clinical applications. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11(2):1–4. https://doi.org/10.1109/tnsre.2003.814454.

Putze, F. (2019). *Brain Art: Brain-Computer Interfaces for Artistic Expression*, chapter Methods and Tools for Using BCI with Augmented and Virtual Reality, pages 433–446. Springer, Cham. https://doi.org/10.1007/978-3-030-14323-7.

Ramos-Murguialday, A., Curado, M. R., Broetz, D., Yilmaz, ., Brasil, F. L., Liberati, and G., … Birbaumer, N. (2019). Brain-machine interface in chronic stroke: Randomized trial long-term follow-up. *American Society of Neurorehabilitation*, 33(3):188–198. https://doi.org/10.1177/1545968319827573.

Renard, Y., Lotte, F., Gilbert, G., Congedo, M., Maby, E., Delannoy, V., Bertrand, O., and Lécuyer, A. (2010). Openvibe: An open-source software platform to design, test and use brain-computer interfaces in real and virtual environments. *Presence Teleoperators and Virtual Environments*, 19(1):35–53. https://doi.org/10.1162/pres.19.1.35.

Ros, T., Baars, B. J., Lanius, R. A., and P.Vuilleumier (2014). Tuning pathological brain oscillations with neurofeedback: a systems neuroscience framework. *Frontiers in Human Neuroscience*, 8:1008. https://doi.org/10.3389/fnhum.2014.01008.

Suryotrisongko, H. and Samopa, F. (2015). Evaluating openbci spiderclaw v1 headwear's electrodes placements for brain-computer interface (bci) motor imagery application. *Procedia Computer Science*, 72:398–405. https://doi.org/10.1016/j.procs.2015.12.155.

Vourvopoulos, A. (2018). *Using Brain-Computer Interaction and Multimodal Virtual-Reality for Augmenting Stroke Neurorehabilitation*. PhD thesis, Universidade da Madeira. https://doi.org/10400.13/1992.

Vourvopoulos, A., Ferreira, A., and Bermudez, S. (2016). Neurow: An immersive vr environment for motor-imagery training with the use of brain-computer interfaces and vibrotactile feedback. In *3rd International Conference on Physiological Computing Systems*, Lisbon, Portugal. SCITEPRESS. https://doi.org/10.5220/0005939400430053.

Vourvopoulos, A., Jorge, C., Abreu, R., Figueiredo, P., Fernandes, J., and Bermudez, S. (2019). Efficacy and brain imaging correlates of an immersive motor imagery bci-driven vr system for upper limb motor rehabilitation: A clinical case report. *Frontiers in Human Neuroscience*, 13(244):567–577. https://doi.org/10.3389/fnhum.2019.00244.

Wang, J., Feng, Z., Ren, X., Lu, N., Luo, J., and Sun, L. (2020). Feature subset and time segment selection for the classification of eeg data based motor imagery. *Biomedical Signal Processing and Control*, 61. https://doi.org/10.1016/j.bspc.2020.102026.