Improving Common Spatial Patterns in Brain-computer Interface using Dynamic Time Warping and EEG Normalization

Mohamed A A Mohamed
Automatic control and Systems Engineering Department
University of Sheffield
Sheffield, UK
mmohamed3@sheffield.ac.uk

Salem Mansour
Automatic control and Systems Engineering Department
University of Sheffield
Sheffield, UK
sslmansour1@sheffield.ac.uk

Payam Soulatiantork
Automatic control and Systems Engineering Department
University of Sheffield
Sheffield, UK
p.soulatiantork@sheffield.ac.uk

Kai Keng Ang
Institute for Infocomm Research
*Agency for Science, Technology and Research (A*STAR)*
Singapore
kkang@i2r.a-star.edu.sg

Phua Kok Soon
Institute for Infocomm Research
*Agency for Science, Technology and Research (A*STAR)*
Singapore
ksphua@i2r.a-star.edu.sg

Mahnaz Arvaneh
Automatic control and Systems Engineering Department
University of Sheffield
Sheffield, UK
m.arvaneh@sheffield.ac.uk

***Abstract*—Common spatial patterns (CSP) is widely employed for spatial filtering and feature extraction in electroencephalogram (EEG)-based brain-computer interfaces (BCIs). However, the non-stationary nature of EEG signals can lead to poor estimation of CSP. To address this drawback, this paper proposes a novel algorithm called scaled and warped CSP (SW-CSP). The proposed algorithm enhances the classical CSP by reducing within EEG class non-stationarity across both time and amplitude domains, followed by automatically selecting the most discriminative features. First, the maximum-minimum scaling is applied to align amplitude of each EEG trial to its class average. Next, dynamic time warping (DTW) temporally aligns each scaled EEG trial to its class average. Thereafter, the scaled and warped EEG trials are used to compute the CSP covariance matrices. Finally, the proposed Fisher's score is used to select most discriminative SW-CSP filters. The proposed SW-CSP algorithm is evaluated using two datasets, and compared with the CSP algorithm and the previously proposed DTW-CSP. The results showed that the SW-CSP significantly outperformed both CSP and DTW-CSP ($p<0.003$). Notably, an improvement above 10% was observed in 20% of the participants.**

***Index Terms*—EEG, brain-computer interface, common spatial patterns, non-stationarity, dynamic time warping, Fisher's score.**

I. Introduction

Brain-computer Interface (BCI) could enable people suffering from severe motor disability to communicate their intention and control external devices only using their brain signals [1]. Motor Imagery (MI) is a mental practice in BCI. A MI-based BCI detects motor related brain patterns generated during imagination of movement of different limbs, e.g. right and left hand, and converts them to commands for controlling an external device [1].

Electroencephalogram (EEG) is commonly used in BCI. Using non-invasive electrodes attached on the user's scalp, EEG records the electrical activities resulting from interaction of neurons [2]. EEG signals are used to detect different mental states and diagnose different neurological conditions such as epilepsy [3]. However, EEG is known to be non-linear, noisy, non-stationary and prone to artifacts [2]. These issues make the single trial EEG classification in MI-based BCI very challenging. Moreover, EEG has a relatively poor spatial resolution due to the volume conduction effect [4]. The volume conduction term refers to the fact that the electrical potentials generated by a brain source can be picked up by EEG electrodes at different locations. Henceforth, several advanced techniques are used to improve the performance of the single trial EEG classification in BCI applications, such as independent component analysis (ICA) [5], auto-regressive modeling [6], and spectral filtering [7].

Common spatial patterns (CSP) is a popular spatial filtering and feature extraction technique in MI-based BCIs [4]. CSP can effectively enhance the spatial resolution of EEG signals and increase the discrimination between the variances of two classes by weighting each electrode according to its importance. However, CSP is highly sensitive to artifacts and non-stationarity present in EEG signals. CSP is based on the estimation of covariance matrices of EEG. Artifacts and stationarity in EEG signals can negatively affect the estimation of EEG covariance matrices, leading to poor CSP features [8].

Several algorithms have been proposed to improve CSP by regularizing either the CSP objective function or the covariance matrix estimator [9]–[11]. In this paper, we focus on improving the covariance matrix estimate in the CSP algorithm. Although, the previously proposed regularization algorithms outperformed the conventional CSP [10], [11], they mostly required performing a time-consuming cross validation on the training data in order to obtain well suited regularization parameters. Moreover, most of the previous studies did not directly take into account the temporal variations between the EEG trials. A recent study by Azab et al. [8] improved the estimation of the CSP covariance matrix using a dynamic time warping (DTW) approach, such that each EEG trial of the same class was aligned to the average of its class. By reducing temporal variations across the trials from the same class, this study successfully improved the estimation of CSP covariance matrices without the need for regularization. However, the DTW algorithm proposed by Azab et al. [8] aligns the EEG signals only based on their variations in the time domain and ignores the amplitude variations between each EEG trial and its class average, which can lead to imprecise DTW alignment.

This study aims to improve the performance of CSP by reducing within class trial-to-trial variability across both time and amplitude domains. This is achieved by a novel DTW algorithm with amplitude scaling, whereby the scaled and dynamically warped EEG trials are used to calculate the CSP covariance matrices. Our proposed scaled and warped CSP algorithm, called SW-CSP, is followed by a filter selection method based on the Fisher's score. The performance of the proposed SW-CSP algorithm is evaluated using two datasets, the publicly available dataset 2a from BCI Competition IV [12] and a large dataset from [13]. The results are compared with those obtained using the CSP algorithm and the previously proposed DTW-based CSP algorithm [8].

II. METHODOLOGY

The proposed SW-CSP algorithm is applied on the training EEG trials and consists of four steps. First, each EEG trial is linearly scaled to have the same amplitude range as the amplitude range of the average of the trials from the same class. Next, each scaled trial is temporally aligned to the average of the trials from the same class using a DTW-based algorithm. Thereafter, the CSP matrix is calculated using the scaled and temporally warped training trials. Finally, the Fisher's score are used to reject the filters with low performance. The details of the proposed SW-CSP algorithm is further described below.

A. Scaling Amplitude of EEG Trials

Let the data of the l^{th} trial be presented as $\mathbf{d}^l = (\mathbf{X}^l, y^l)$, where $\mathbf{X}^l \subset \mathbb{R}^{s \times ch}$ represents the single-trial EEG matrix with s number of EEG instances and ch number of channels. Subsequently, $y^l \in \{0, 1\}$ denotes the class label of $\mathbf{X}^l$. For each subject the trials are divided into two classes. The average of the trials belonging to the class $y \in \{0, 1\}$ can be computed as

$$\bar{\mathbf{X}}_y = \frac{1}{n_y} \sum_{l=1}^{n_y} \mathbf{X}^l_y, \tag{1}$$

where n_y denotes the number of trials belonging to the class y, and $\mathbf{X}^l_y$ is the l^{th} trial belonging to the class y. EEG signals are usually contaminated with different sources of noise and artifacts. Because of the lack of consistency of this noise across the trials, averaging EEG trials could help reduce the noise.

In the first step of the proposed SW-CSP, a min-max scaling is applied to the original EEG trials $\mathbf{X}^l$ and the average trials of each class $\bar{\mathbf{X}}_y$, such that their amplitudes fall into the range of [-0.5, 0.5]. Thus, the scaled $\mathbf{X}^l$, $\mathbf{X}'^l$, is calculated as

$$\mathbf{X}'^l = -0.5 + \frac{\mathbf{X}^l - \min(\mathbf{X}^l)}{\max(\mathbf{X}^l) - \min(\mathbf{X}^l)}. \tag{2}$$

Similarly, $\bar{\mathbf{X}}_y$ is scaled as follows:

$$\bar{\mathbf{X}}'_y = -0.5 + \frac{\bar{\mathbf{X}}_y - \min(\bar{\mathbf{X}}_y)}{\max(\bar{\mathbf{X}}_y) - \min(\bar{\mathbf{X}}_y)}. \tag{3}$$

Please note the functions min and max respectively find a single minimum and maximum value across all the channels.

B. Temporal Alignment using Dynamic Time Warping

Dynamic Time Warping (DTW) has been well known in speech recognition field e.g. determining the matching of one phrase to another regardless if the phrase is spoken slower or faster than the other [14]. DTW maps two time series that are similar in shape but locally at different phases by minimizing the distance between them.

Scaling the amplitude of EEG trials, reduces the amplitudes variations but ignores the variations in the temporal dimension. Therefore, DTW is employed in order to align the scaled trials $\mathbf{X}'^l$ to the scaled average of the trials from the same class $\bar{\mathbf{X}}'_y$. The algorithm starts by constructing $s-by-s$ distance matrix ,$\mathbf{D}$, which forms all pairwise distances between $\mathbf{X}'^l$ and $\bar{\mathbf{X}}'_y$. $\mathbf{D}(i, j)$ is the Euclidean distance computed by

$$\mathbf{D}(i, j) = \sqrt{\sum_{k=1}^{ch} (\mathbf{X}'^l(i, k) - \bar{\mathbf{X}}'(j, k))^2}, \tag{4}$$

where i and j are time instances of $\mathbf{X}'^l$ and $\bar{\mathbf{X}}'_y$ respectively. After the local cost matrix is built, DTW algorithm formulates an optimal alignment path $\mathbf{P}^*$, that runs through the elements of the local cost matrix. In our study, an alignment path is represented as

$$\mathbf{P} = [p(1), ..., p(k), ..., p(K)], \quad s =< K < 2s - 1 \tag{5}$$

where the k_{th} element of the warping path is represented as: $p(k) = \mathbf{D}(i_k, j_k)$. Generally, the warping path should satisfy the following constrains:

(1) Boundary: This condition ensures that the path starts with $i_1 = j_1 = 1$ and terminates at $i_K = j_K = s$.

(2) Multi-mapping: One element of the first signal can be aligned to more than one element of the second signal.

(3) Monotonicity: The points' time order is preserved (i.e. no returning path). Hence, $i_k \geq i_{k-1}$,$j_k \geq j_{k-1}$.

(4) Continuity: The path transition is limited to the adjacent point, $i_k - i_{k-1} \leq 1$, $j_k - j_{k-1} \leq 1$.

(5) Slop: This condition avoids extreme movement of the path towards one direction.

Given the above-mentioned conditions, an acceptable path transition has one of the following three possibilities: $(i,j) \rightarrow (i+1,j), (i,j) \rightarrow (i,j+1), (i,j) \rightarrow (i+1,j+1)$.

The shortest non-linear alignment between $\mathbf{X'}^{l}$ and $\bar{\mathbf{X}}'_y$, satisfying the above-mentioned conditions, is called optimum warping path $\mathbf{P}^*$, identified as:

$$\mathbf{P}^* = \mathbf{argmin_P}\left(\frac{1}{K}\sum_{k=1}^{K}(p(k))\right). \tag{6}$$

Examining every possible path is a computationally expensive. Hence, the optimal path, $\mathbf{P}^*$, is computed recursively by utilizing a dynamic programming approach where the cumulative distance $\mathbf{d}$ is defined as:

$$\mathbf{d}(i,j) = \mathbf{D}(i,j) + \min[\mathbf{d}(i-1,j), \mathbf{d}(i,j-1), \mathbf{d}(i-1,j-1)]. \tag{7}$$

Finally, given $\mathbf{P}^*$, $\mathbf{X'}^{l}$ is aligned to $\bar{\mathbf{X}}'_y$ as:

$$\mathbf{X'^{l}_{aligned}} = \begin{bmatrix} \mathbf{X}'^{l}(i_1^*,1) & ... & \mathbf{X}'^{l}(i_1^*,ch) \\ ... & ... & ... \\ \mathbf{X}'^{l}(i_K^*,1) & ... & \mathbf{X}'^{l}(i_K^*,ch) \end{bmatrix}, \tag{8}$$

where $[i_1^*, i_2^* .... i_K^*]$ is obtained from $\mathbf{P}^*$, denoting the time indices in which $\mathbf{X'}^{l}$ and $\bar{\mathbf{X}}'_y$ are minimally separated.

C. Scaled Warped Common Spatial Filters (SW-CSP)

SW-CSP filters, $\mathbf{W_{SW}}$, spatially filter the EEG signals, such that the difference between the variances of the two classes is maximized. Thus,

$$\begin{aligned} & maximize \quad \mathbf{W_{SW}^T}\overline{\mathbf{C_1}}\mathbf{W_{SW}} \\ & subject \quad to \quad \mathbf{W_{SW}^T}(\overline{\mathbf{C_1}} + \overline{\mathbf{C_2}})\mathbf{W_{SW}} = \mathbf{I}, \end{aligned} \tag{9}$$

where $\overline{\mathbf{C_y}}$ refers to the average of the covariance matrices of the scaled aligned EEG trials, belonging to the class y. Henceforth, $\overline{\mathbf{C_y}}$ is computed as:

$$\overline{\mathbf{C_y}} = \frac{1}{n_y}\sum_{l=1}^{n_y}\frac{\mathbf{X'^{l}_{aligned}}.(\mathbf{X'^{l}_{aligned}})^{\mathrm{T}}}{\mathrm{trace}(\mathbf{X'^{l}_{aligned}}.(\mathbf{X'^{l}_{aligned}})^{\mathrm{T}})} \tag{10}$$

where T and trace(.) refer to the transpose and trace function respectively. n_y is the number of the trials for class y. Next, the SW-CSP filters are calculated by optimizing the objective function (9) using the joint diagonalization [15]. Thereafter, $\mathbf{X'^{l}_{aligned}}$ is spatially filtered using the SW-CSP filters:

$$\mathbf{Q^l} = \mathbf{W_{SW}}.(\mathbf{X'^{l}_{aligned}})^{\mathrm{T}}, \tag{11}$$

where $\mathbf{Q^l} \subset \mathbb{R}^{ch \times s}$. Finally , the feature vector of the l^{th} trial, $\mathbf{R^l} \subset \mathbb{R}^{ch}$, is obtained as:

$$\mathbf{R^l} = \log(\mathrm{diag}(\mathbf{Q^l}\mathbf{Q^l}^{\mathrm{T}})), \tag{12}$$

where diag(.) returns the diagonal elements of a matrix.

D. Feature Selection based on Fisher's Scores

The Fisher's scores aim to identify the most discriminative features that exhibit small within-class distances and large between-class distances [16]. To calculate these scores, the feature matrix $\mathbf{R_y} \subset \mathbb{R}^{ch \times n_y}$ is constructed for each class y using the training data, where n_y represents the total number of training trials for class y.

The between-class distance, denoted as $\mathbf{S_b}$, is computed as the sum of the squared differences between the average feature vectors of each class and the overall average feature vector, weighted by the number of trials in each class:

$$\mathbf{S_b} = \sum_{y=1}^{2}\frac{n_y}{n}(\bar{\mathbf{R}}_\mathbf{y} - \bar{\mathbf{R}})(\bar{\mathbf{R}}_\mathbf{y} - \bar{\mathbf{R}})^{\mathbf{T}}, \tag{13}$$

Similarly, the within-class distance, denoted as $\mathbf{Sw}$, is calculated as the sum of the squared differences between each trial's feature vector and the average feature vector of its respective class:

$$\mathbf{S_w} = \frac{1}{n}\sum_{y=1}^{2}\sum_{l=1}^{n_y}(\mathbf{R_y^l} - \bar{\mathbf{R}}_\mathbf{y})(\mathbf{R_y^l} - \bar{\mathbf{R}}_\mathbf{y})^{\mathbf{T}}, \tag{14}$$

Here, $\bar{\mathbf{R}}_\mathbf{y}$ and $\bar{\mathbf{R}}$ represent the average feature vectors for class y and both classes, respectively. $\mathbf{R_y^l}$ denotes the features of the l_t trial from class y, and n is the total number of train trials from both classes.

The Fisher's score for the i_t feature, denoted as $\mathbf{F_i}$, is then calculated as the ratio of the between-class distance to the within-class distance for that feature:

$$\mathbf{F_i} = \frac{\mathbf{S_b(i,i)}}{\mathbf{S_w(i,i)}}. \tag{15}$$

E. Classification

On the last stage, we selected the six discriminant features with the highest Fisher's scores and fed them to the classifier. In this study, we employed SVM to evaluate the performance of three different CSP variants: classical CSP, DTW-CSP, and the proposed SW-CSP. SVM is widely recognized as one of the most popular and effective classifiers for MI-based BCIs, particularly in online and real-time applications [17].Lotte et al. in a systematic review study showed that SVM outperforms other state-of-the-art classifiers in this domain. [17].

III. EXPERIMENT

A. Data Description

1) Dataset 2a from BCI Competition IV [12]: The data consists of EEG signals collected from nine different subjects using 22 electrodes. The EEG data from each subject was recorded while performing four different movement imagination tasks (right hand, left hand, both feet and tongue). However, only the imagination of movements of the right and left hands are considered in this study. Moreover, each subject performed two sessions of motor imagery on two different days. Each session consists of 6 runs interrupted by short breaks. One run is comprised of 12 trials per class giving

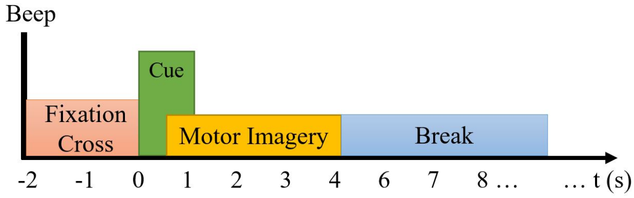

Fig. 1. Guideline for the experiment's timing. Each trial begins with an audible beep, followed by a two-second fixation cross. At time 0, a visual cue appears on the screen, prompting the participant to initiate the motor imagery tasks. Participants will be given four seconds after the cue to perform the motor imagery tasks, followed by a short break.

a total of 288 trials for one session. In this paper, A_i presents the i^{th} subject from this dataset.

2) Dataset from [13]: The dataset contains EEG data from 16 healthy subjects recorded with 27 electrodes. For each subject, two sessions, on two separate days, were recorded while the subjects performing motor imagery for hand movement (either left or right) versus an idle condition. Each session consists of two runs, with each run collecting 80 trials without feedback. In this paper, B_i presents the i^{th} subject from this dataset.

The timing of each trial is described in Fig. 1. Each trial starts with an audible beep and a two-second period of a fixation cross. When time reaches 0, a visual cue appears on the screen, signalling the participant to initiate the motor imagery tasks. After the cue, participants have four seconds to perform the motor imagery tasks. This is followed by a short break before the next trial begins. In this study, the first session for each subject was used for training, while the second session was used for testing.

B. Data Processing

First, offline EEG data analysis was performed using the EEGLAB toolbox and MATLAB 2021a. The EEG signals were bandpass filtered in the frequency range of 1-40 Hz. The filtered EEG data was visually inspected using the EEGLAB tool, and trials with artifacts were eliminated by rejecting those with absolute amplitudes that exceeded 120 microvolts at any electrode [18]. Thereafter, re-referencing using common average referencing was applied to reduce the contribution of common noise sources and enhance the relative differences between electrodes. The signals were recorded at a sampling rate of 250 Hz. Data recorded from 0.5 to 2.5 seconds after the cue were extracted and band-pass filtered from 8 Hz to 35 Hz using a zero-phase elliptic filter. Next, spatial filtering and feature extraction were performed using either SW-CSP, DTW-CSP [8], or CSP. For all methods, the proposed Fisher's scores were used to select the six most discriminative features. The normalized logarithm of variances of the spatially filtered EEG data using these six features was used as classification features. By employing Fisher's scores for feature selection, we were able to select the most discriminating features from each method and make a fair comparison of their performance. Finally, a support vector machine (SVM) was used as the classifier as recommended in [17].

IV. RESULTS

Fig. 2 shows the classification results of the two datsets using the classical CSP, DTW-CSP, and the proposed SW-CSP. As can be seen, the proposed SW-CSP algorithm outperformed CSP by an average of 6% and 2.6% using the first and the second datasets respectively. Interestingly, this improvement is more than 10% for subjects A2, A4, A5, B1 and B5. The results also reveal that the proposed SW-CSP algorithm outperformed DTW-CSP by an average of 4.3% and 2.4% using the first and the second datasets respectively. The non-parametric Wilcoxon signed rank test on all the results showed that the proposed SW-CSP algorithm significantly outperformed both the CSP and DTW-CSP algorithms with the p-values of 0.0025 and 0.0029 respectively.

Fig. 3 illustrates the comparison of an EEG trial before and after DTW alignment, with and without amplitude scaling. The figure demonstrates that aligning the amplitude of the single trial with its class average helps reduce the variability in both temporal and amplitude aspects within subject trials. This approach not only improves the performance of the DTW alignment but also contributes to a better understanding of why the proposed SW-CSP algorithm outperforms the DTW-CSP algorithm, as shown in Fig.3.

Fig.4 demonstrates examples of some selected spatial filters computed using CSP and the proposed SW-CSP algorithm for subject A6. From a neurophysiological perspective, it can be seen that the filters obtained from CSP presented some weights appearing in unexpected locations. Whereas, SW-CSP has shown neurophysiologically more relevant and smoother weights. For example, for the left hand motor imagery, the SW-CSP weights are stronger over the motor cortex area and leans to the right side. However, CSP wrongly identified weights on occipital and parietal parts of the brain for the left hand motor imagery.

V. CONCLUSION

This paper introduces a novel approach to enhance the performance of the CSP algorithm in MI based BCIs. The proposed method, named SW-CSP, focuses on minimizing variability in the time and amplitude domains across EEG trials, thereby reducing non-stationarity within trials. Consequently, it produces more consistent and accurate spatial filters that better capture the underlying neural activity. The experimental results clearly demonstrate the superiority of the SW-CSP algorithm over both the CSP and DTW-CSP algorithms in generating spatial filters that are more relevant to neurophysiology. These findings have significant implications for the practical implementation of EEG-based BCIs by significantly increasing the performance of the MI-based BCI while having a reasonably manageable computational time in an online BCI scenario. Compared to the regularized CSP algorithms proposed in literature, the proposed SW-CSP does nor require any time-consuming cross-validation on the

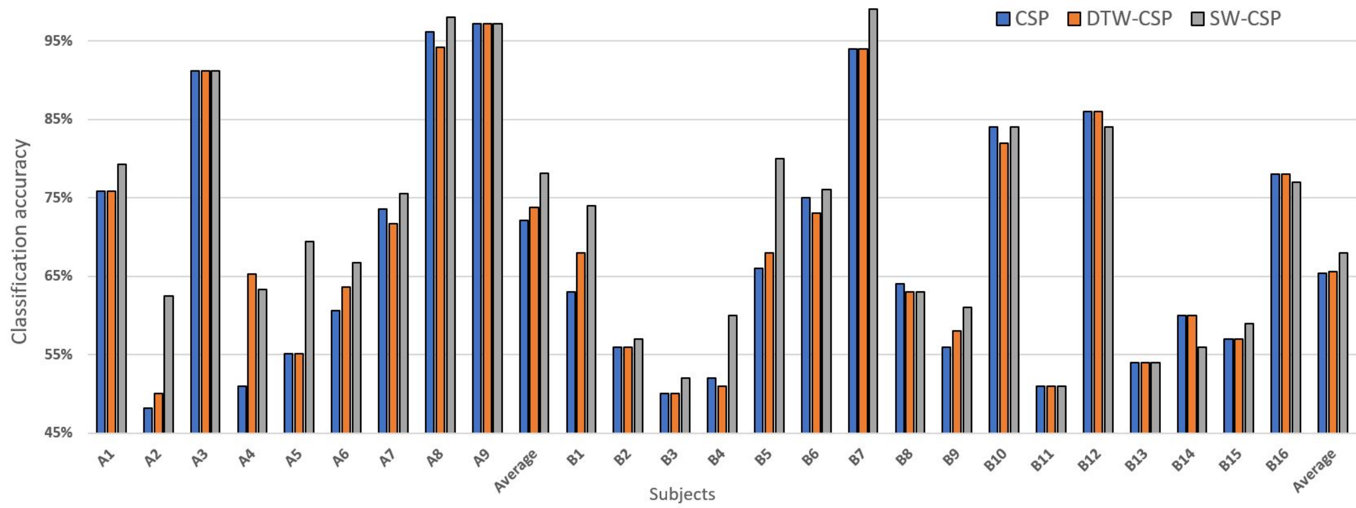

Fig. 2. Comparison of classification results obtained using the classical CSP, DTW-CSP and the proposed SW-CSP. A_i and B_i refer to the subjects i from datasets A and B respectively. The results of the non-parametric Wilcoxon signed rank test demonstrated a significant improvement in the performance of the proposed SW-CSP algorithm compared to classical CSP algorithms, with an average increase of 6% and 2.6% in datasets A_i and B_i, respectively.

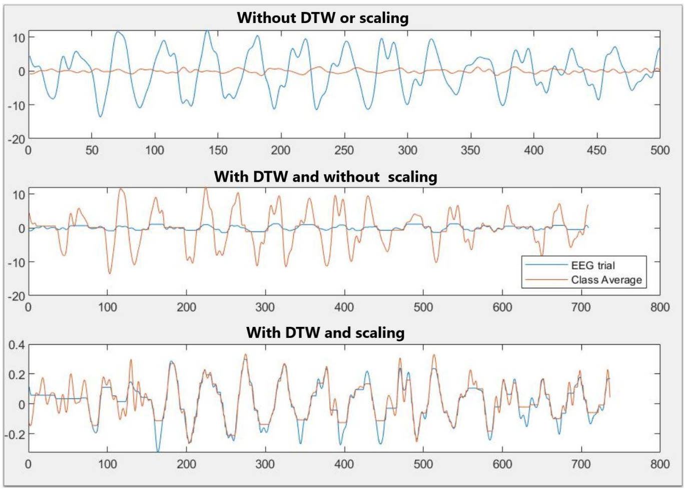

Fig. 3. Comparison of Dynamic Time Warping performance in aligning a single-channel EEG trial to its class average, both before and after amplitude scaling. It can be observed that the scaling of the EEG trials improves the DTW alignment significantly.

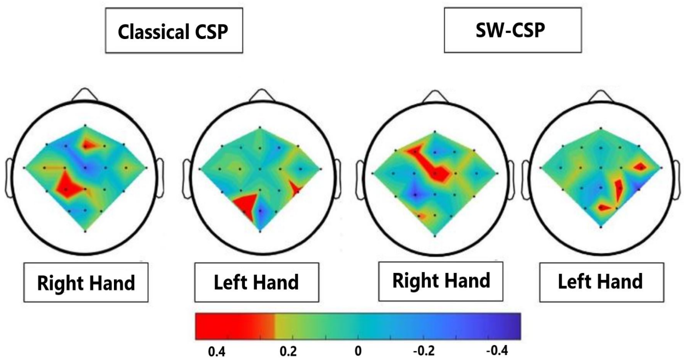

Fig. 4. Visual presentation of the most discriminative spatial filters computed using the classical Common Spatial Patterns (CSP) and the proposed scaled-warped CSP (SW-CSP) algorithm for subject A6. The colors show the weights that filters give to each channel. The filters obtained from CSP appear to have random weights in unexpected locations, while SW-CSP produces smoother and more neurophysiologically relevant weights.

train data in order to find the regularization value. These factors play a crucial role in advancing the development of real-time BCIs. However, it is important to note that further validation on larger datasets is necessary as future work. This validation would offer a more comprehensive understanding of the algorithm's efficacy and further enhance its practical applicability.

References

[1] M. AhmedKhan, R. Das, H. K.Iversen, and S. Puthusserypady, "Review on motor imagery based bci systems for upper limb post-stroke neurorehabilitation: From designing to application," *Computers in Biology and Medicine*, vol. 103843, no. 123, pp. 1–17, 2020.

[2] S. Vaid, P. Singh, and C. Kaur, "Eeg signal analysis for bci interface: A review. in 2015 fifth international conference on advanced computing communication technologies," *IEEE*, pp. (143–147), 2015, February.

[3] S. J. Smith, "Eeg in the diagnosis, classification, and management of patients with epilepsy," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 76, no. suppl 2, pp. ii2–ii7, 2005.

[4] M. Tariq, P. M.Trivailo, and M. Simic, "Classification of left and right knee extension motor imagery using common spatial pattern for bci applications," *Procedia Computer Science*, vol. 159, pp. 2598–2606, 2019.

[5] A. Subasi and M. I. Gursoy, "Eeg signal classification using pca, ica, lda and support vector machines," *Expert Systems with Applications*, vol. 37, no. 12, pp. (8659–8666), 2010.

[6] A. E. Vijayan, D. Sen, and A. Sudheer, "Eeg-based emotion recognition using statistical measures and auto-regressive modeling," *IEEE International Conference on Computational Intelligence Communication Technology*, pp. (587–591), 2015.

[7] A. S. Al-Fahoum and A. A. Al-Fraihat, "Methods of eeg signal features extraction using linear analysis in frequency and time-frequency do-

mains," *International Scholarly Research Notices (ISRN) Neuroscience*, pp. (1–7), 2014.

[8] A. M.Azab, L. Mihaylova, H. Ahmadi, and M. Arvaneh, "Robust common spatial patterns estimation using dynamic time warping to improve bci systems," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. (3897–3901), 2019.

[9] F. Lotte and C. Guan, "Regularizing common spatial patterns to improve bci designs: Theory and algorithms," *IEEE Transactions on biomedical Engineering*, vol. 58, no. 2, pp. (355–362), 2010.

[10] R. Lahiri, A. Rakshit, and A. Konar, "Discriminating motor imagery eeg signals using an improvised regularised csp algorithm," *IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES)*, pp. 1–6, 2016.

[11] H. V. Shenoy, A. P. Vinod, and C. Guan, "Shrinkage estimator based regularization for eeg motor imagery classification," *International Conference on Information, Communications and Signal Processing (ICICS)*, pp. 1–5, 2015.

[12] C.Brunner, R.Leeb, G. Muller-Putz, A.Schlog, and G.Pfurtscheller, "Bci competition 2008- graz data set a," *Institute for Knowledge Discovery (Laboratory of Brain-Computer Interfaces), Graz University of Technology*, vol. 16, pp. 1–6, 2008.

[13] M. Arvaneh, C. Guan, K. K. Ang, T. E. Ward, K. S. G. Chua, C. Wee, K. Kuah, G. J. E. Joseph, K. S. Phua, and C. Wang, "Facilitating motor imagery-based brain–computer interface for stroke patients using passive movement," *Neural Computing and Applications*, vol. 28, no. 11, pp. 3259–3272, 2017.

[14] I.-J. Ding, C.-T. Yen, and Yen-MingHsu, "Developments of machine learning schemes for dynamic time-wrapping-based speech recognition," *Mathematical Problems in Engineering*, pp. (1–10), 2013.

[15] M. Arvaneh, C. Guan, K. K. Ang, and C. Quek, "Spatially sparsed common spatial pattern to improve bci performance," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. (2412–2415), 2011.

[16] M. N. KP and P. Thiyagarajan, "Feature selection using efficient fusion of fisher score and greedy searching for alzheimer's classification," *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 8, pp. 4993–5006, 2022.

[17] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi, "A review of classification algorithms for eeg-based brain–computer interfaces," *Journal of neural engineering*, vol. 4, no. 2, p. R1, 2007.

[18] J. Wang, J. Wang, J. Sun, C. Li, S. Tong, and X. Hong, "The effects of pre-cue posterior alpha on post-cue alpha activity and target processing in visual spatial attention tasks with instructional and probabilistic cues," *Cerebral Cortex*, vol. 33, no. 7, pp. 4056–4069, 2023.