Deep Riemannian Neural Architectures for Domain Adaptation in Burst cVEP-based Brain Computer Interface

Sébastien Velut, Sylvain Chevallier, Marie-Constance Corsi, Frédéric Dehais

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

Deep Riemannian Neural Architectures for Domain Adaptation in Burst cVEP-based Brain Computer Interface

Velut Sébastien [1,2], Chevallier Sylvain[2], Corsi Marie-Constance[3], Dehais Frédéric [1]

1- Fédération ENAC ISAE-SUPAERO ONERA,
Université de Toulouse, 10 avenue Edouard Belin, 31400, Toulouse - France
2- A&O - LISN - Université Paris-Saclay
1 rue René Thom, 91190 Gif-sur-Yvette - France
3- Sorbonne Université, Institut du Cerveau - Paris Brain Institute -ICM,
CNRS, Inria, Inserm, AP-HP, Hopital de la Pitié Salpetriere, F-75013, Paris, France

Abstract. Code modulated Visually Evoked Potentials (cVEP) is an emerging paradigm for Brain-Computer Interfaces (BCIs) that offers reduced calibration times. However, cVEP-based BCIs still encounter challenges related to cross-session/subject variabilities. As Riemannian approaches have demonstrated good robustness to these variabilities, we propose the first study of deep Riemannian neural architectures, namely SPDNets, on cVEP-based BCIs. To evaluate their performance with respect to subject variabilities, we conduct classification tasks in a domain adaptation framework using a burst cVEP open dataset. This study demonstrates that SPDNet yields the best accuracy with single-subject calibration and promising results in domain adaptation.

1 Introduction

Code modulated Visually Evoked Potentials (cVEP) have gained popularity in the Brain-Computer Interface (BCI) community [1]. This approach employs pseudo-random visual flickers, providing advantages such as shorter calibration times, as only one code needs to be learned. Alternative decoding methods, like bitwise-decoding [2], have enabled self-paced BCI with flexible decoding period. Despite these advancements, cVEP-based BCIs remain primarily studied in lab settings due to the persistent need for recalibration before each use. This limitation is related to cross-session and cross-subject variabilities common to all BCI paradigms. These sources of variability in BCI are diverse [3], encompassing anatomical differences such as variations in grey matter quantity, human factors like differences in education level and lifestyle habits, or physiological factors like fatigue, concentration levels, and stress levels. Additionally, neurophysiological disparities, such as variations in modulations of spectral power across specific frequencies, also contribute to these variabilities. To address these sources of variability, extensive research has been conducted [4, 3] to propose new approaches. There are two main settings for evaluating transfer learning approaches, depending on the quantity of information available for a target subject. In the most independent setting, refer to as *Domain Generalization*, no information from the target subject is at hand thus the model is trained on data

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

from n-1 subjects and tested on the n-th (target) subject. If some information could be obtained from the target subject, the so-called *Domain Adaptation* setting uses a small portion of the n-th subject's data for training [5]. In the context of Deep Learning (DL) algorithms, the model could be initially trained on n-1 subjects, followed by freezing one or more layers and retraining the model on the data from the last subject [6]. To achieve reasonable performances in DG/DA settings, an important preprocessing step is data alignment, which aims to align the data from all subjects to ensure a similar feature space across subjects [5]. While these settings have shown promise in addressing inter-subject variability, there is still room for improvement in terms of accuracy and robustness. In this context, Riemannian geometry has emerged as a powerful tool for enhancing the performance of BCIs. Riemannian techniques have the ability to capture the intrinsic structure of the data on a curved manifold, enabling more effective and efficient classification algorithms [7, 8]. Moreover, recentering the data before classification has been proven to enhance accuracy when using transfer learning techniques [9].

In this study, we propose a novel approach that adapts Riemannian deep learning model, namely SPDNet, for cVEP-based BCIs to address cross-subject variability. We aim to compare our method with state-of-the-art CNN model for cVEP in terms of accuracy and robustness across subjects. For this study, we adapted SPDNet [10] and its batch normalization version [11] (SPDBNNet) for a cVEP dataset [2] with 12 participants. We evaluated these models on two different transfer learning settings: Domain Generalization (DG) and Domain Adaptation (DA). They are compared with Single Subject (SS) training baseline (no transfer learning), where one model per subject is trained with longer calibration phase than DA. In the following sections, we will first provide in Sect. 2 some preliminaries to redefine key definitions of Riemannian manifolds and context related to the cVEP. Next, we will present in Sect. 3 the methodology used in our study, followed in Sect. 4 by the results obtained and the implications of our findings. Finally, we conclude the paper with Sect. 5.

2 Preliminaries

2.1 Riemannian manifold

Let $\mathcal{P}_n$ be the set of the $n \times n$ symmetric definite-positive (SPD) matrices : $\mathcal{P}_n : \{P \in R^{n \times n}, P^\top = P, u^\top P u = 0,\ \forall u \in R^n\}$ endowed with the affine-invariant Riemannian (AIR) distance. While other distance could be considered, the properties of the AIR and the Frechet mean (center of mass of a SPD matrices) are interesting for processing EEG and their definition can be found in [8],[7]

2.2 cVEP and Burst cVEP

Code modulated VEP relies on pseudo-random code of 0s and 1s, often referred to as an m-sequence or Gold codes, and is generated using a linear feedback

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

shift register. Once the initial sequence is produced, the subsequent sequences are then generated by phase-shifting this initial sequence. Indeed, it is crucial for these sequences to be as uncorrelated as possible to ensure a maximal discriminability between the sequences. The burst cVEP is a variation of the cVEP. Burst cVEP employs short bursts of irregular visual flashes at a slower rate [2]. This method evokes a distinct brain response, known as P100 (i.e., a positive deflection that occurs 100 ms after the brief burst of flash) with double the amplitude of classical cVEP m-sequences, as shown in Figure 1. Those clearer responses to onsets facilitate the detection and classification of the code. Additionally, those code could be more comfortable for the user, as it is possible to reduce the flash amplitude, generating thus less visual fatigue in response to those rapidly blinking stimuli.

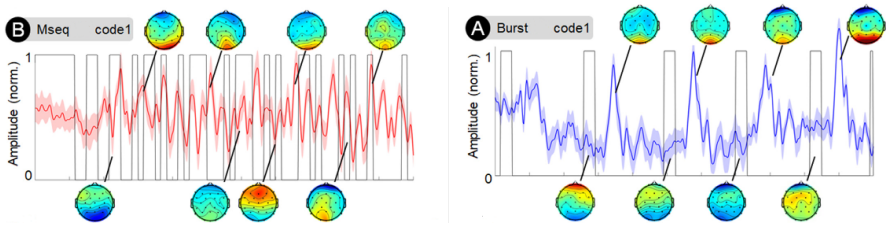

Fig. 1: cVEP relying on m-sequence or burst. Left: The black line illustrates a prototypical m-sequence cVEP, animating the presentation of the flash with alternating plateaus of '1' (flash on) and '0' (flash off). The red line represents the averaged and normalized cerebral response to this alternating on/off visual stimulation. Right: Example of Burst cVEP graphs, consisting of brief flash presentations. The blue line depicts the averaged and normalized cerebral response to this alternating on/off visual stimulation. Adapted from [2]

3 Material and Method

3.1 Datasets

The data used for this study come from the dataset created with the study of Cabrera-Castillos et al.[1] [2]. In this experiment, 12 participants were instructed to concentrate on four targets cued sequentially for 2.2 seconds in random order for each of 15 blocks (60 trials in total). To retrieve the dataset easily and to facilitate the reproducibility of this study, we have implemented this dataset in MOABB [12].

3.2 Method

The emergence and advantages of Riemannian methods in recent years have led to significant improvements in the performance of BCIs, particularly in the

[1]available at `https://zenodo.org/record/8255618`

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

context of motor imagery tasks [8]. However, there have been relatively few studies exploring the potential of Riemannian classifiers in the domain of cVEP-based BCIs. Motivated by this gap in the literature, we sought to investigate the performance of Riemannian classifiers (3 layers of BiMap-ReEig, followed by a LogEig layer, a flatten and a Linear layer; for the SPDBNNet, the batch normalisation layers are not domain specific) in comparison to a state-of-the-art deep learning algorithm, an optimized CNN which has demonstrated excellent results for Burst cVEP [2].

We evaluated different settings for transfer learning, namely domain generalization and domain adaptation. Let k be the target participant index (evaluated in testing set) and $I = [1, ...12]$ the participants' index. For domain generalization, let be $\Omega_S^{\mathrm{DG}} = \{X_i, \forall i \in I \backslash k\}$ the training set with X_i the data of the i-th participant. The model is trained on Ω_S and tested on the testing set $\Omega_T^{\mathrm{DG}} = \{X_k\}$. In Domain Adaptation setting, a subset $X_k^{\mathrm{DA}} \in X_k$ of trials from subject k are available in $\Omega_S^{\mathrm{DA}} = \Omega_S^{\mathrm{DG}} \cup X_k^{\mathrm{DA}}$ and, indeed, excluded from testing set $\Omega_T^{\mathrm{DA}} = \Omega_T^{\mathrm{DG}} \backslash X_k^{\mathrm{DA}}$. In this study, X_k^{DA} are the first 16 (out of 60) trials of X_k, representing roughly 32 sec of calibration for a subject. The baseline is a single subject setting, using X_k^{SS} 32 (out of 60) trials for training, that represents a circa 110s calibration to train a subject/session-specific model. For the training we used the following parameters : lr=0.001,optimiser=Adam or RiemannianAdam,loss=CrossEntropy,epochs=20,batchsize=64.

All channels of the 32-electrode EEG cap are used, EEG data is bandpass filtered between 1-25 Hz, and the signals are re-referenced to the average. Epochs were created by extracting 0.25s windows following each frame and labeling them with the corresponding bit of the code. For the Riemannian algorithms, a spatial XDAWN filter, on class 1, is estimated and applied before computing covariance matrices with Ledoit-Wolf estimator. A recentering step is applied to whiten the covariance matrices. For the CNN, the data were normalized as commonly done with neural nets and a recentering step is applied too. Due to the unbalanced nature of Burst cVEP, the dataset has circa 4.85 times more 0 than 1. To address this imbalance, we randomly removed 0-labeled epochs to match 1-labeled epochs in training set. We did not balance the testing set, to reflect situation occurring in a real world online BCI application. To avoid bias, we repeated 10 times the process of each pipelines and then average the different scores.

4 Results and discussion

The statistical differences shown in this section are obtained via the Stouffer's method that combines p-values resulting from the Wilcoxon signed-rank test. Then a Bonferroni correction was performed.

Figure 2 illustrates the performance comparison of different DL models in different settings. The DA setting exhibits significantly higher accuracy than DG setting($p \leqslant 0.05$) but, no difference is found between DG and SS setting or between DA and SS setting. This result is something in contrary to what we expected compared to other studies. There are much more training samples

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

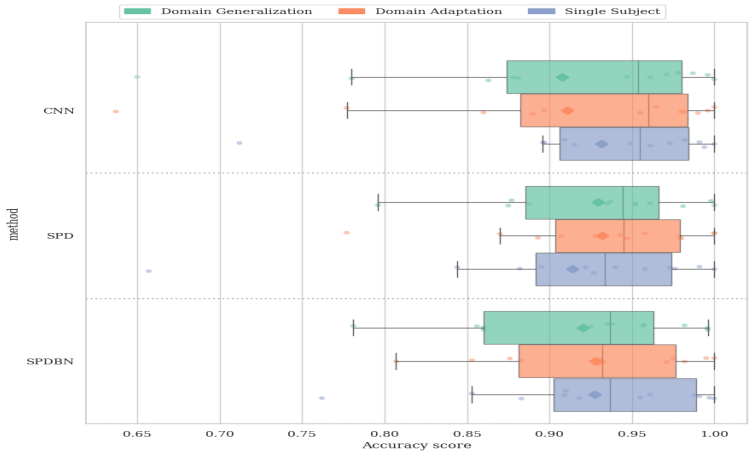

Fig. 2: Average accuracy scores for all subjects (10 repetitions). The box range is between the first and third quartile. The black line in the box is the median and the dark color square point is the mean

in DG and DA settings than in SS setting. It can explained that SS setting is not significantly higher. It indicates that riemannian algorithms and the CNN performs equally well between the settings.

When comparing the same setting (DG/DA/SS) across different models, all three models show comparable performance without significant differences, with one exception: in the SS setting, the CNN and the SPDBNNet achieves better accuracy than the SPDNet. Notably, we successfully matched the accuracy of an optimized CNN using Riemannian DL models, achieving very good performance (circa 93% accuracy). It should be noted the hyperparameters of the CNN are well optimized as reported in [2], while we did not conduct any hyperparameter search for Riemannian DL models, it is thus likely that the SPDNets could match the CNN performances when carefully tuned. Indeed, we will conduct such optimization and propose an ablation study in a near future.

However, although the CNN appears more consistent across different settings, it exhibits more outliers compared to the SPDNet and SPDBNNet. This indicates a preferable behavior for the Riemannian DL models in a context of transfer learning. These results were obtained by predicting a code in a fixed time window of less than 2 seconds. For instance, the mean prediction time for SPDBNNet in the SS settings was 1.438 seconds, while for CNN in the SS settings, it was 1.458 seconds. However, it is important to note that the training time of SPDBNNet is longer than that of SPDNet, which was 1.442 seconds, which in turn is longer than the training time of the CNN.

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

5 Conclusion

This article presented two Riemannian deep learning (DL) models that were on par with the state-of-the-art model in terms of accuracy and decision time for cVEP-based BCIs. These models were more robust to outliers and had better accuracy for the lowest-performing subjects. This first study, which employed a vanilla SPDNet architecture without optimization, showed promising results. We plan to work on more complex approaches for transfer learning to tackle cross-subject variabilities. In this study, we restricted our analysis to cross-subject evaluation, but we are confident that cross-session evaluations for a given subject might greatly benefit from these findings.

References

[1] V. Martinez-Cagigal, J. Thielen, E. Santamaria-Vazquez, S. Perez-Velasco, P. Desain, and R. Hornero. Brain-computer interfaces based on code-modulated visual evoked potentials (c-VEP): a literature review. *Journal of Neural Engineering*, page 22, 2021.

[2] K. Cabrera Castillos, S. Ladouce, L. Darmet, and F. Dehais. Burst c-VEP based BCI: Optimizing stimulus design for enhanced classification with minimal calibration data and improved user experience. *NeuroImage*, page 11, 2023.

[3] S. Saha and M. Baumert. Intra- and inter-subject variability in EEG-based sensorimotor brain computer interface: A review. *Frontiers in Human Neuroscience*, page 8, 2020.

[4] S. Khazem, S. Chevallier, Q. Barthélemy, K. Haroun, and C. Noûs. Minimizing subject-dependent calibration for bci with riemannian transfer learning. In *IEEE/EMBS Conference on Neural Engineering (NER)*, pages 523–526, 2021.

[5] F. Fahimi, Z. Zhang, W. Boon Goh, T. Lee, K. Keng Ang, and C. Guan. Inter-subject transfer learning with end-to-end deep convolutional neural network for EEG-based BCI. *Journal of Neural Engineering*, page 13, 2018.

[6] B. Aristimunha, R. Y. de Camargo, W. H. Lopez Pinaya, S. Chevallier, A. Gramfort, and C. Rommel. Evaluating the structure of cognitive task with transfer learning. In *NeurIPS workshop AI4Science*, page 19, 2023.

[7] A. Barachant, S. Bonnet, M. Congedo, and C. Jutten. Classification of covariance matrices using a Riemannian-based kernel for BCI applications. *Neurocomputing*, 112:172–178, 2013.

[8] F. Yger, M. Berar, and F. Lotte. Riemannian approaches in brain-computer interfaces: A review. *IEEE TNSRE*, 25(10):1753–1762, 2017.

[9] P. Zanini, M. Congedo, C. Jutten, S. Said, and Y. Berthoumieu. Transfer learning: a riemannian geometry framework with applications to brain-computer interfaces. *IEEE TBME*, page 12, 2018.

[10] Z. Huang and L. Van Gool. A riemannian network for SPD matrix learning. In *AAAI*, page 9, 2017.

[11] R. Kobler, J. Hirayama, Q. Zhao, and M. Kawanabe. SPD domain-specific batch normalization to crack interpretable unsupervised domain adaptation in EEG. In *NeurIPS*, 2022.

[12] S. Chevallier, I. Carrara, B. Aristimunha, P. Guetschel, S. Sedlar, B. Lopes, S. Velut, S. Khazem, and T. Moreau. The largest eeg-based bci reproducibility study for open science: the moabb benchmark. *arXiv preprint arXiv:2404.15319*, 2024.