Heading for motor imagery brain-computer interfaces (MI-BCIs) usable out-of-the-lab: Impact of dry electrode setup on classification accuracy

Maria-Isabel Casso, Camille Jeunet, Raphaëlle N. Roy

Heading for motor imagery brain-computer interfaces (MI-BCIs) usable out-of-the-lab: Impact of dry electrode setup on classification accuracy

Maria-Isabel Casso[1], Camille Jeunet[2], and Raphaëlle N. Roy[1]

***Abstract*— A primary challenge to make motor-imagery Brain-Computer Interfaces (MI-BCIs) technologies usable and actually used out-of-the-lab consists of providing EEG systems that are efficient in terms of classification accuracy- and easy to install, e.g., using a minimal number of dry electrodes. We hypothesize that the optimal signal processing method might depend on the number of (dry) electrodes that are used. Therefore, we compared for the first time the classification accuracy associated with different dry electrode setups, i.e., 7 configurations from 8 to 32 channels, and various signal processing methods, namely (1) regularized Common Spatial Pattern (rCSP) + Linear Discriminant Analysis, (2) rCSP + Support Vector Machine (SVM), (3) Minimum Distance to Riemannian Mean and (4) SVM in the Riemannian Tangent Space. This offline comparison was performed on the data of 10 participants (one session each). Our results suggest that whatever the method, MI-BCI performances drop significantly for 8 and 12 channels ($p < 0.01$). Moreover, method 3 was associated with the lowest performances ($p < 0.05$). Finally, post-doc analyses suggest that methods 1 and 2 perform best with the highest numbers of electrodes 28 and 32. For method 4 the best performances are obtained using 20 and 24 channels, which seems to be the optimal combination ($p < 0.05$). These results show the importance of selecting the signal processing pipeline as a function of the number of electrodes used.**

I. INTRODUCTION

Brain-Computer Interfaces (BCIs) are closed-loop systems that operate with brain signals obtained by recording methods such as electroencephalography (EEG) or functional near infra-red spectroscopy (fNIRS). These systems provide a user with the possibility of controlling other systems by translating their brain activity into commands [1]. A typical BCI paradigm, or task, is motor imagery (MI), a process in which the user imagines a limb movement, which generates specific EEG power amplitude variations [2]. The signal processing and machine learning steps necessary to achieve a high performance in BCIs, and more particularly in EEG-based MI-BCIs, have been designed and benchmarked by the growing BCI community [3]. An important step towards out-of-the-lab MI-BCI use is to evaluate i) how dry-EEG systems perform when using a reduced number of electrodes, and ii) which processing pipeline enable the achievement of high classification accuracy with setups including a low number of dry electrodes. Indeed, gel electrode systems present some challenges such as the time needed for subject preparation, signal stabilization, and equipment cleaning for every EEG recording [4], as well as the user's comfort that is critical for acceptability [5]. In a survey from 2015 on BCI users with spinal cord injury [6], it was concluded that the simplicity of a BCI setup was as crucial as the BCI functions. Among the surveyed patients, 89% demonstrated an acceptance for dry electrodes compared to 62% who advocated for gel electrodes. Moreover, in recent years the need for out-of-the-lab and wearable EEG equipment has become a major need in BCI research. A survey revealed that about 90% of neurologists agreed that wireless EEG systems were useful and very much needed and concluded that more work investigating the efficiency of various algorithms in this context should be done [7]. Several studies have been performed over the years to compare both types of electrodes, usually by using a similar electrode setup for both dry and gel systems [8], [9], [10], [11], including the performance analysis of new dry electrode wearable and mobile patents [12], [13], [14], [15]. In [8], an 8 dry-electrode helmet was compared to a 22-gel electrode setup with 2 subjects in 2 MI sessions. Classification accuracies of 67.5% and 71.9% were obtained with the dry electrode system using random forest and Support Vector Machine (SVM) algorithms respectively, with a preliminary spatial filtering step using Common Spatial Patterns (CSP), while 80% was reached with the gel system. Also, in [9], similar performances were obtained with dry and gel electrode EEG systems. The data quality of 2 dry electrode EEG systems was compared in [16], with the conclusion that both were of satisfying quality and much more usable than EEG systems using gel. Nonetheless, to our knowledge no study has yet directly compared the impact of the electrode setup, and more specifically the electrode number, using a dry EEG system, for several classification pipelines. In order to head towards out-of-the-lab MI-BCI use, this preliminary study evaluates how 7 dry-electrode setups -decreasing from 32 to 8 electrodes- impact the classification accuracy of left- vs. right-hand MI with four state-of-the art classification pipelines.

II. METHODS

A. Graz-BCI paradigm

The paradigm of this study is based on the Graz-BCI experimental paradigm [17] implemented in OpenVIbe [18]. During the MI-BCI session, participants initially have to wait for 40s to stabilize the signal and begin the trials. At each trial, they have to focus on a centered green cross displayed on screen and 2s later, a red arrow pointing either left or right

[1] Maria-Isabel Casso and Raphaëlle N. Roy are with ISAE-SUPAERO, Université de Toulouse, France. maria.casso-echalar@isae-supaero.fr raphaelle.roy@isae-supaero.fr

[2] Camille Jeunet is with the CLLE Lab (University Toulouse Jean Jaurès and CNRS) camille.jeunet@univ-tlse2.fr

cues for left or right-hand MI, respectively. Participants have to perform the instructed task until the cross disappears, after which they can relax for 3s before the new trial starts (figure 1). Participants took part in one session including 20 trials per task, pseudorandomly distributed.

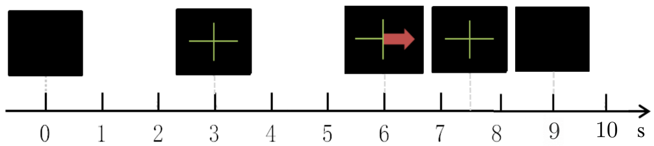


Fig. 1: Experimental paradigm

B. Participants and Data Acquisition

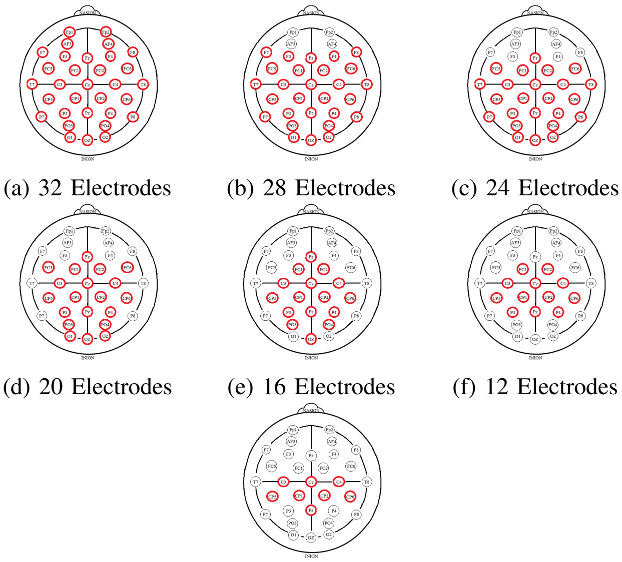
Ten volunteer students, 8 of them BCI-naive, 6 females, with an average age of 25.5 years old (± 1.13 years), participated in this study. The later took place between Sept. and Oct. 2020, following the sanitary protocol in force at the time to prevent COVID-19 contagion. The recording sessions were carried in a computer room, with common distractions. Each subject participated in two recording sessions, one for training and the second for online feedback. In this paper only the data from the training session is considered, as it allows for the offline investigation of different electrode setups and classification pipelines. All participants were seated in a chair facing a 19in. monitor at an approximate distance of 1m. Signals were recorded with an Enobio32 wireless EEG cap using the following 32 comb-shaped dry electrodes positioned according to the 10-20 system: FP1, FP2, AF3, AF4, F7, F3, FZ, F4, F8, FC5, FC1, FC2, FC6, C3, CZ, C4, T7, T8, CP5, CP1, CP2, CP6, P7, P3, PZ, P4, P8, PO3, PO4, O1, OZ, O2.The Correlated-Multiple-Sampling (CMS) and Driven Right Leg (DRL) channels where placed on the left and right mastoid respectively. The participants were instructed to stand still, and that the MI should involve as much finger movement as possible.

C. Electrode setups

Starting from the 32-electrode setup that covers the whole scalp, we defined 7 electrode placement configurations by removing 4 electrodes at each step, until reaching the minimum 8-electrode setup centered on C3 and C4 above the motor areas (figure 2).

D. Processing Pipelines

Four standard and state-of-the-art classification pipelines -here called "methods"- were used to evaluate the impact of the dry-electrode setup on MI classification accuracy. OpenVIbe was used for offline data processing. The first method (rCSP+LDA) was composed of a regularized CSP filtering step and a Linear Discriminant Analysis classification step (LDA); the second (rCSP+SVM) also comprised a rCSP step, but with an SVM classification. The last two methods were based on Riemannian Geometry: Minimum Distance to Mean (MDRM) and SVM in the Riemannian Tangent Space (RTS+SVM). All methods were tested using a 10-fold cross-validation procedure.



(a) 32 Electrodes (b) 28 Electrodes (c) 24 Electrodes

(d) 20 Electrodes (e) 16 Electrodes (f) 12 Electrodes

(g) 8 Electrodes

Fig. 2: Electrode setups: from 32 whole scalp coverage to 8 central electrodes above motor areas.

1) Preprocessing

The data was bandpass-filtered between 8 to 30 Hz with a Butterworth filter. Epoching was performed using 1s-long sliding windows with1/16s overlap [19]. The classification pipeline was applied to each window. Then, the outputs were averaged in order to obtain a mean classification accuracy for each trial.

2) Method 1: rCSP with LDA classifier (rCSP+LDA)

The main idea of the CSP algorithm is to find a linear transform to enhance the discriminability of data and thereby ease the classification process. This method aims to find spatial filters that maximize the variance of the signal for one class and minimize it for the others, and so on, for all classes [3]. Due to its susceptibility to noise, CSP regularization (rCSP) has been proven to be a more robust option [20]. This pipeline follows the Tikhonov method, which regularizes the CSP algorithm's objective function by adding a penalty function and a regularization parameter defined manually, which in our case was 0.5.We obtained 6 filters per condition with a chunk average covariance update. Next, the filtered data is used for training an LDA classifier that finds a linear combination to separate data into two classes (namely, left- and right-hand MI)[3]. Although the estimates might not be optimal, linear classifiers have proven to be efficient with single-trial classification [3].

3) Method 2: rCSP with SVM classifier (rCSP+SVM)

Here, the same rCSP spatial filtering step as in method 1 was used before classification, but combined with an SVM. The SVM classifier projects the data to a higher-dimensional space with kernel functions and finds an hyperplane that will separate the classes with the highest margin possible. The optimization problem and classifier's performance depends not only on the training data but also on the right choice of a penalty and kernel parameters [3]. For this study, this

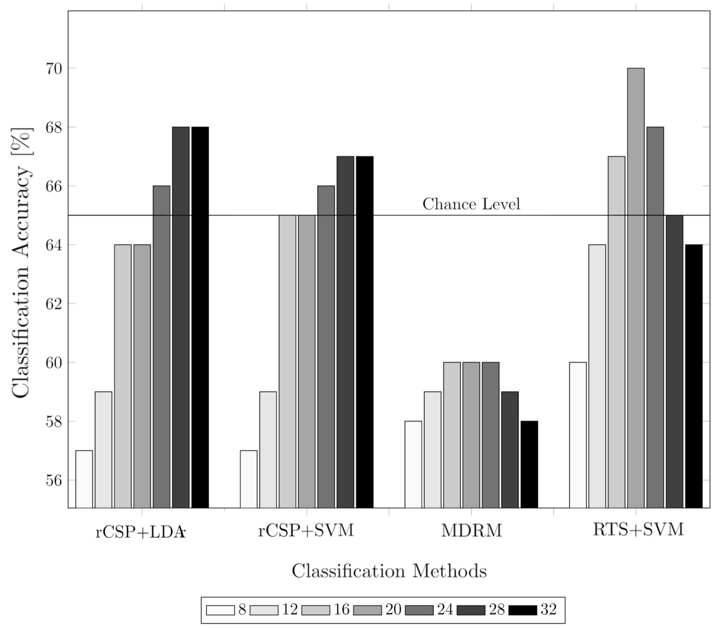


Fig. 3: Average classification accuracy for all electrode setups and classification pipelines

pipeline's penalty term was 200 and we used the Radial Basis Function (RBF) kernel, based on the parameter selection scheme presented in [21].

4) Method 3: Minimum Distance to Riemannian Mean (MDRM)

A significant advantage of methods based on Riemann geometry is that the algorithms directly process the data covariance matrices that have embedded spatial information necessary for classification, thus saving computation time. For the MDRM, it is first necessary to estimate the covariance matrix of training data, denominated as the Sample Covariance Matrix (SCM), in our case, done with an Oracle Approximating Shrinkage (OAS) estimator. The SCM matrix is considered a positive definite matrix, thus belonging to the Riemannian manifold. The optimization problem becomes finding the minimum mean Riemannian geodesic distance between a set of training data SCM matrices and every intra-class covariance matrix [22][23].

5) Method 4: SVM in the Riemannian Tangent Space (RTS+SVM)

As with an SVM classifier, classification with the Riemannian Tangent Space is based on hyperplane projection. This main idea of this method is to map the SCM matrices into the Riemannian tangent space, found in the data's geometric mean [23]. This projection allows finding spatial vectors, treated as feature vectors for classification of dimension $n(n+1)/2$, where n is the number of variables [22], [24].The classifier used in this method is SVM with the same penalty term and kernel configurations as method 2.

III. RESULTS

The classification accuracy for each electrode setup was obtained using the four methods described in section II. Figure 3 presents the average classification accuracy for each processing pipelines and electrode setup. As a comparison for the obtained accuracies, the adjusted chance level limit for a 95% confidence ($\alpha = 5\,\%$), for two classes and 40 trials has been computed [25].

A statistical analysis (2-way ANOVA for repeated measures and Tukey's post-hoc tests) also revealed that the electrode setup/number had a significant impact on classification accuracy across all methods ($F(6, 54) = 8.23, p < 0.01$) with a significant drop in classification accuracy for the 12 and 8 electrode setups compared to the other ones ($p < 0.01$). There was also a significant interaction between the electrode setup and the method on the classification accuracy ($F(18, 162) = 2.58, p < 0.01$). Indeed, there was a significant drop in accuracy for the 8 and 12 electrode setups compared to the other setups for both the rCSP + LDA and rCSP + SVM pipelines, with the highest accuracy reached for the 32 and 28 electrode setups ($p < 0.05$). However, for the MDRM pipeline, there was no significant difference in classification accuracy between electrode setups, the classification accuracy remaining stable, although the significantly lowest of all pipelines ($p < 0.05$). Lastly, regarding the RTS + SVM pipeline, there was a significant drop in classification accuracy for the 8 electrode setup compared to the highest accuracy reached with this method, i.e., with the 20 and 24 electrode setups ($p < 0.05$).

IV. DISCUSSION

The BCI research community's need for usable BCI systems has been repeatedly reported [4], [6], [7]. A first step to head towards out-of-the-lab BCI use would be to provide users with a portable and dry-EEG system. We hypothesise that the optimal signal processing method might depend on the number of (dry) electrodes that are used to record the brain activity. Therefore, we compared for the first time the classification accuracy associated with different dry electrode setups (7 configurations from 8 to 32 channels) and various state-of-the-art signal processing methods (rCSP+LDA, rCSP+SVM, MDRM and RTS+SVM) through the offline analysis of the data of 10 participants who took part in one session each. Our results show the importance of selecting the signal processing pipeline as a function of the number of electrodes and suggest that a minimum of 16 electrodes is required to maintain the classification accuracy above the adjusted chance level for the rCSP+LDA and rCSP+SVM pipelines and a minimum of 12 for the RTS+SVM pipeline. All results from the MDRM pipeline remained below the adjusted chance level. The highest classification accuracy was reached using the RTS+SVM pipeline with an average of 70% for the 20-electrode setup. Interestingly, this pipeline's classification accuracy increased when reducing the setups from 32 to 20 electrodes; the setups with 32 and 28 electrodes attained an average accuracy below the adjusted chance level. This increase in accuracy suggests that the signal from the frontal electrodes (kept only in the 32 and 28 electrode setups) might harm the MI classification process when no spatial filtering step is applied, or that there is some overfitting phenomenon involved with this pipeline. Hence, a setup of 12 or 16 dry electrodes centered on C3 and C4 is the best option to maintain the classification accuracy to an acceptable level.

These preliminary results have to be taken with caution as the data used were only gathered from 10 participants -eight of them BCI-naive- during a single training session. The lack of BCI experience among the subjects [26] and the sanitary restrictions to prevent COVID-19 contagion that limited the number of participants and recording sessions in this study might have also influenced the classification performance. Nevertheless, this study is a first step to evaluating portable MI-BCI for out-of-the-lab BCI to advise research headed for practical, easily operated options. In future studies, further data should be collected in order to assess the relevance of this promising result on long-term performance, but also to evaluate the influence that the selected method may have on the participants' ability to learn how to control an MI-BCI efficiently.

ACKNOWLEDGMENT

This work was funded by an ANITI (Artificial and Natural Intelligence Toulouse Institute) Research Chair allocated to Pr Frédéric Dehais.

REFERENCES

[1] J. Wolpaw and E. W. Wolpaw, *Brain-computer interfaces: principles and practice*. OUP USA, 2012.

[2] M. Clerc, L. Bougrain, and F. Lotte, *Brain-Computer Interfaces 1*. Wiley-ISTE, 2016.

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

[4] C. Guger, G. Krausz, B. Z. Allison, and G. Edlinger, "Comparison of dry and gel based electrodes for p300 brain–computer interfaces," *Front Neurosci*, vol. 6, p. 60, 2012.

[5] W. Ko and J. Hynecek, "Dry electrodes and electrode amplifiers," *Biomedical electrode technology*, pp. 169–181, 1974.

[6] J. E. Huggins, P. A. Wren, and K. L. Gruis, "What would brain-computer interface users want? opinions and priorities of potential users with amyotrophic lateral sclerosis," *Amyotrophic Lateral Sclerosis*, vol. 12, no. 5, pp. 318–324, 2011.

[7] A. J. Casson, S. Smith, J. S. Duncan, and E. Rodriguez-Villegas, "Wearable eeg: what is it, why is it needed and what does it entail?," in *Int Conf IEEE Eng Med Biol Soc*, pp. 5867–5870, 2008.

[8] L. Angrisani, P. Arpaia, F. Donnarumma, A. Esposito, M. Frosolone, G. Improta, N. Moccaldi, A. Natalizio, and M. Parvis, "Instrumentation for motor imagery-based brain computer interfaces relying on dry electrodes: a functional analysis," in *IEEE Int Instrum Meas Techn Conf*, pp. 1–6, IEEE, 2020.

[9] C. Guger, G. Krausz, and G. Edlinger, *Brain-computer interface control with dry EEG electrodes*. Citeseer, 2011.

[10] J. W. Kam, S. Griffin, A. Shen, S. Patel, H. Hinrichs, H.-J. Heinze, L. Y. Deouell, and R. T. Knight, "Systematic comparison between a wireless eeg system with dry electrodes and a wired eeg system with wet electrodes," *NeuroImage*, vol. 184, pp. 119–129, 2019.

[11] K. E. Mathewson, T. J. Harrison, and S. A. Kizuk, "High and dry? comparing active dry eeg electrodes to active and passive wet electrodes," *Psychophysiology*, vol. 54, no. 1, pp. 74–82, 2017.

[12] J. Xu, S. Mitra, C. Van Hoof, R. F. Yazicioglu, and K. A. Makinwa, "Active electrodes for wearable eeg acquisition: Review and electronics design methodology," *IEEE Rev Biomed Eng*, vol. 10, pp. 187–198, 2017.

[13] T. O. Zander, M. Lehne, K. Ihme, S. Jatzev, J. Correia, C. Kothe, B. Picht, and F. Nijboer, "A dry eeg-system for scientific research and brain–computer interfaces," *Front Neurosci*, vol. 5, p. 53, 2011.

[14] A. J. Casson, "Wearable eeg and beyond," *Biomed Eng Letters*, vol. 9, no. 1, pp. 53–71, 2019.

[15] Y. Chi, Y.-T. Wang, Y. Wang, C. Maier, T.-P. Jung, and G. Cauwenberghs, "Dry and noncontact EEG sensors for mobile brain–computer interfaces," *IEEE Trans Neur Sys Rehab Eng*, vol. 20, pp. 228–235, 2011.

[16] H. Rieiro, C. Diaz-Piedra, J. M. Morales, A. Catena, S. Romero, J. Roca-Gonzalez, L. J. Fuentes, and L. L. Di Stasi, "Validation of electroencephalographic recordings obtained with a consumer-grade, single dry electrode, low-cost device: A comparative study," *Sensors*, vol. 19, no. 12, p. 2808, 2019.

[17] G. Pfurtscheller and C. Neuper, "Motor imagery and direct brain-computer communication," *Proc IEEE*, vol. 89, pp. 1123–1134, 2001.

[18] Y. Renard, F. Lotte, G. Gibert, M. Congedo, E. Maby, V. Delannoy, O. Bertrand, and A. Lécuyer, "Openvibe: An open-source software platform to design, test, and use brain–computer interfaces in real and virtual environments," *Presence: teleoperators and virtual environments*, vol. 19, no. 1, pp. 35–53, 2010.

[19] C. Jeunet, E. Jahanpour, and F. Lotte, "Why standard brain-computer interface (bci) training protocols should be changed: an experimental study," *J Neur Eng*, vol. 13, no. 3, p. 036024, 2016.

[20] F. Lotte and C. Guan, "Regularizing common spatial patterns to improve bci designs: unified theory and new algorithms," *IEEE Trans Biomed Eng*, vol. 58, no. 2, pp. 355–362, 2010.

[21] A. Rojas and A. Nandi, "Practical scheme for fast detection and classification of rolling-element bearing faults using support vector machines," *Mech Syst Sign Process*, vol. 20, pp. 1523–1536, 2006.

[22] M. Congedo, A. Barachant, and R. Bhatia, "Riemannian geometry for eeg-based brain-computer interfaces; a primer and a review," *Brain-Computer Interfaces*, vol. 4, no. 3, pp. 155–174, 2017.

[23] A. Barachant, S. Bonnet, M. Congedo, and C. Jutten, "Multiclass brain–computer interface classification by riemannian geometry," *IEEE Trans Biomed Eng*, vol. 59, no. 4, pp. 920–928, 2011.

[24] S. Kumar, K. Mamun, and A. Sharma, "Csp-tsm: Optimizing the performance of riemannian tangent space mapping using common spatial pattern for MI-BCI," *Comput Biol Med*, vol. 91, pp. 231–242, 2017.

[25] G. Müller-Putz, R. Scherer, C. Brunner, R. Leeb, and G. Pfurtscheller, "Better than random: a closer look on bci results," *International Journal of Bioelectromagnetism*, vol. 10, pp. 52–55, 2008.

[26] B. Blankertz and C. Vidaurre, "Towards a cure for bci illiteracy: machine learning based co-adaptive learning," *BMC Neuroscience*, vol. 10, no. 1, pp. 1–2, 2009.