Exploring strategies for multimodal BCIs in an enriched environment

Tristan Venot, Arthur Desbois, Marie-Constance Corsi, Laurent Hugueville, Ludovic Saint-Bauzel, Fabrizio de Vico Fallani

Exploring strategies for multimodal BCIs in an enriched environment

1st Tristan Venot
Sorbonne Université, Paris Brain Institute - ICM,
CNRS, Inria, Inserm, AP-HP, Hôpital de la Pitié Salpêtrière,
Paris, France
tristan.venot@inria.fr

2nd Arthur Desbois
Sorbonne Université, Paris Brain Institute - ICM,
CNRS, Inria, Inserm, AP-HP, Hôpital de la Pitié Salpêtrière,
Paris, France
arthur.desbois@inria.fr

3rd Marie Constance Corsi
Sorbonne Université, Paris Brain Institute - ICM,
CNRS, Inria, Inserm, AP-HP, Hôpital de la Pitié Salpêtrière,
Paris, France
ORCID: 0000-0002-6262-5036

4th Laurent Hugueville
Sorbonne Université, Paris Brain Institute - ICM,
CNRS, Inserm, AP-HP, Hôpital de la Pitié Salpêtrière,
Paris, France
laurent.hugueville@sorbonne-universite.fr

5th Ludovic Saint-Bauzel
Institut des Systèmes Intelligents et de Robotiques - ISIR,
Sorbonne Université, CNRS, Inserm,
Paris, France
ludovic.saint-bauzel@sorbonne-universite.fr

6th Fabrizio De Vico Fallani
Sorbonne Université, Paris Brain Institute - ICM,
CNRS, Inria, Inserm, AP-HP, Hôpital de la Pitié Salpêtrière,
Paris, France
fabrizio.de-vico-fallani@inria.fr

***Abstract*—Brain computer interfaces rely on cognitive tasks easy at first sight but that reveal to be complex to perform. In this context, providing engaging feedback and subject's embodiment is one of the keys for the overall system performance. However, noninvasive brain activity alone has been demonstrated to be often insufficient to precisely control all the degrees of freedom of complex external devices such as a robotic arm. Here, we developed a hybrid BCI that also integrates eye-tracking technology to improve the overall sense of agency of the subject.**

While this solution has been explored before, the best strategy on how to combine gaze and brain activity to obtain effective results has been poorly studied. To address this gap, we explore two different strategies where the timing to perform motor imagery changes; one strategy could be less intuitive compared to the other and this would result in differences of performance.

I. Introduction

Despite continuous breakthrough in BCI, especially in machine learning [1][2] which is giving promising results of in terms of classification accuracy, some important challenges remain when it comes to controlling devices such as robotic limbs. Indeed, classification is limited to a few number of classes up to four in the best cases. However the mental tasks associated with each class can be sometimes counter intuitive and challenging for the subject. All this put together, flawless control reveals to be difficult. Hence, we use additional sources of control to help in commanding complex systems with a large number of degrees of freedom, seven in the case of a robotic arm.

Creating multi-modal systems with Motor Imagery BCI at its core has been studied in depth in the past [3],[4],[5], so has been the transfer to a robotic control [6],[7],[8],[9]. MI BCIs rely on the mental task of imagining a movement without performing it [10]).In this general context,the timing to perform the MI task is a key question, so far left unaddressed.

To address this question, we create a protocol in which subjects are asked to control a robotic arm both by gaze and MI BCI. In the experimentation, the timing of the MI task differs from one session to another, either before or after the robot's movement. We focus both on neuro-physiological features and classification performance to assess whether a strategy offers better performances. We make the hypothesis that better performances will occur when the motor imagery task is performed during the last movement phase of the robot (i.e. hand closing), in comparison, to a MI task performed before any movement of the robot. We evaluate two strategies that give two different timing to the subject to perform motor imagery task. In this article, we present the method and the protocol used as well as the early results on those two strategies. In the last part, we discuss our results and identify leads of interpretation of our findings.

FDVF and CP acknowledge support from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant Agreement No. 864729). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. This project is also supported by AID DGA.

II. Method and Protocol

A. Material

The system consists of a robotic arm (Pollen Robotic Reachy 7 degrees of freedom) mounted in front of the subject.

The robot, can reach two cans facing the subject. The cans are on an "augmented table", a flat monitor (42") under a Plexiglas screen. The subject is wearing Tobii Pro glasses 3 that record and transmit gaze activity in real time. EEG data are recorded using a Brain Products EEG cap of 64 amplified wet electrodes with amplification and frequency sampled at 500 Hz. Reference and Ground electrodes are respectively placed on TP9 and TP10 positions (at the mastoid level). Impedance level for the electrodes is set to 15 $k\Omega$ with a tolerance of 10 $k\Omega$. The software used for EEG acquisition and BCI control is OpenViBE 3.2.0.

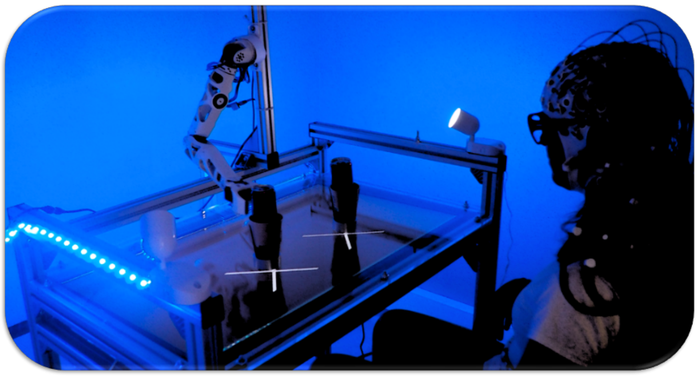


Fig. 1: Experimental setup, composed of the robot Reachy, the Tobii Pro Glasses 3 eyetracker, the 64 Brain products EEG cap and the augmented table

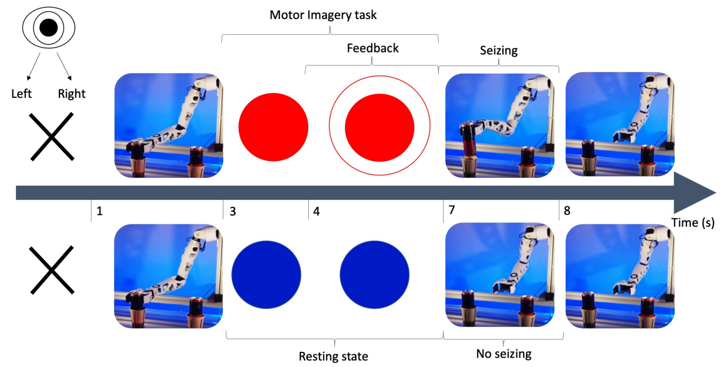


Fig. 2: Strategy 1:the subject selects the target using gaze, then the robot goes to the designated target. The visual stimulus is given, the robot closes its gripper when a red dot appears, otherwise, it does not.

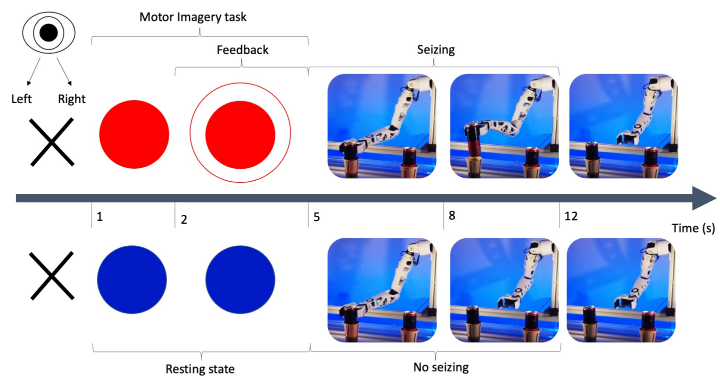


Fig. 3: Strategy 2:the subject selects the target using gaze. The visual stimulus is given. And then the robot goes to the designated target. The robot closes its gripper when a red dot appears, otherwise, it does not.

B. *Experimental Protocol*

The experimentation was performed at the CENIR platform at the Paris Brain Institute in a controlled environment (Faraday cage). The subjects sat on a chair wearing the EEG cap and the eyetracker facing the robot as presented in figure 1. The experimentation is a sequence of both gaze action and mental task related to the control of the robot. First, the targeted object is selected using gaze. Secondly, a robotic action and visual stimulus are shown, depending on the motor imagery task: [seize+lift+drop+red dot] in the case of MI, and [simply going back + blue dot] in the case of resting. The subjects are asked for the motor imagery task to imagine closing their right hand to seize the can. The visual stimuli lasts for 4 seconds during the Training set. During the control set, the visual stimuli lasts for 1 second and a discrete continuous feedback is given (a halo circling the target getting smaller depending on the mental state of the subject - this is directly linked to the classification distance to hyperplane). Those sequences are presented in figure 2 and 3 with the different strategies.

The session is decomposed into 2 phases; first, a calibration phase composed of 3 runs lasting for 7 minutes and 50 seconds, corresponding to 10 trials of MI and 10 of rest. During this phase the robot closes its hand every time it is supposed to be reacting to a motor imagery task. From the data collected in the calibration, we generate an R^2 map and choose what will be the relevant features for the classifier weights both in terms of electrodes and frequencies of interest related to motor imagery patterns, hence in the motor cortex area and in the α and β bands (8 to 35 Hz). We also generate the wilcoxon map in the case where the R^2 is not giving relevant results, we only keep values at $p < 0.05$. This choice is linked to what are the highest significant differences in the map. We then use those selected features of interest to train a 2 class LDA classifier. Secondly, we perform a control phase where the subject is "in control" which means that the incoming samples will be treated as belonging to one class or another. During the MI trial lasting 3 seconds, the classifier attributes a probability to belong to a class to incoming samples. If the majority of incoming samples belongs to the MI class, the hand will close. In the resting state, the robot does not close its hand and goes back to the baseline position. We voluntarily bias the system because we want to be sure that the resting state will always be a *relaxation* state. The control phase is also 3 runs. After each phase, the subject is asked to answer questions regarding their sense of agency

Seven healthy subjects (27.2 $+/-$ 2.1 years old, 1 Male) volunteered for the experimentation. They were all right handed, naive to BCI experimentation, and all signed informed consent. This protocol named *BRACCIO* has been approved by Inria national ethic comity as part of BCI-NET protocol.

C. Method

To compare the different strategies, we use several metrics. For classification, we use sensitivity.

$$Sens = \frac{TP}{TP + FN} \tag{1}$$

The sense of agency assessed via a questionnaire based on Van Acken work [11]. Lastly, given that the classification performance cannot disentangle the subject's and the classifier performance, we also consider neurophysiological features to compare the strategies[12].

In our work, we voluntarily bias the system in order to never do anything during the resting state. Therefore, the accuracy metric (based on the confusion matrix) is not completely adequate to evaluate the performance of the system as it takes into account *True Negatives* and *False Positive* that are associated to the resting state. Hence the need for the sensitivity that focuses more on what the subject perceives, which is only during motor imagery performance. The neurophysiological marker associated to MI is the power spectral density in the α and β bands (8 to 35 Hz) that decreases with regards to the resting state. To estimate the power spectrum, we use Burg auto-regressive (AR) method as it is more relevant to study electro-encephalogram (EEG) data than standard FFT method [13]. The AR model is generated with an order of 20, allowing to establish a certain baseline a certain baseline of comparison and it follows research basis of Bufalari et al [14] as well as Krusienski et al [15]. We evaluate statistical differences between conditions of MI and resting states using R^2 with a focus on electrodes of interest at frequency bands related to MI. The classification algorithm used for the experimentation is a LDA 2 classes where the features (Electrode's power spectrum at certain frequency bands) are selected accordingly to the R^2 and Wilcoxon maps and are specific to each subject. We randomise the order of strategies from one subject to another to limit the learning effect at the group level. However this effect is evaluated via a 2 way ANOVA between strategies and sessions. The pre-processing performed is common average reference (CAR). The window of analysis for the power spectrum is 3 seconds starting 1 second after the visual stimulus to match between Control and Calibration sets.

III. Results

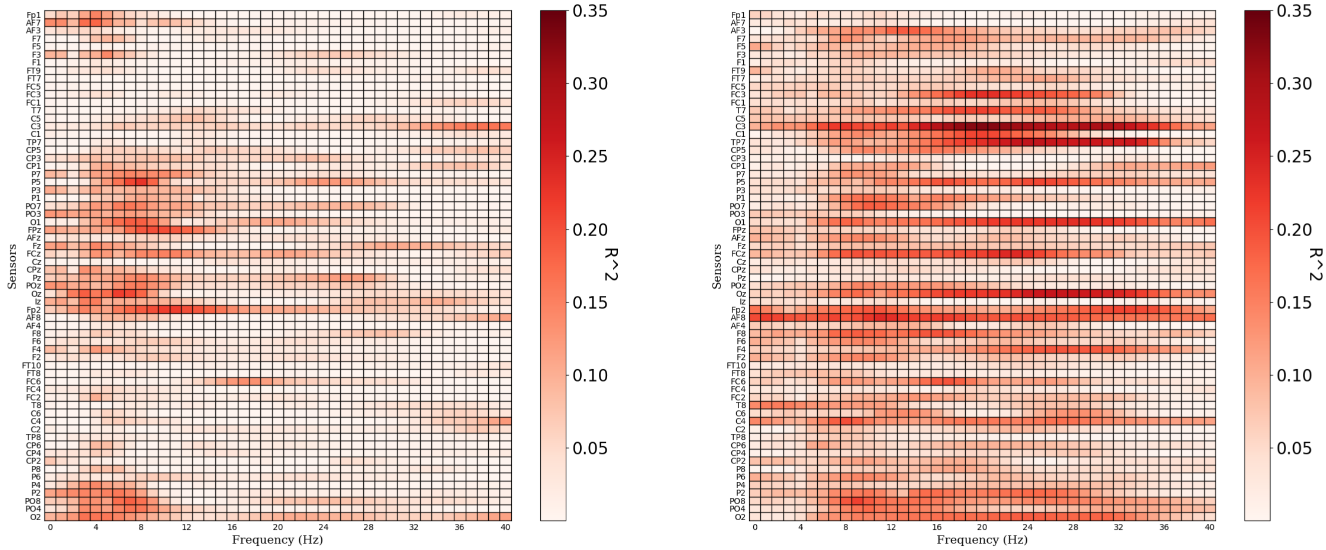
For each subject, we measured the difference of power spectrum between motor imagery and resting, for both training and testing sessions. They all presented significant differences ($p < 0.05$) between the two mental tasks using wilcoxon rank test in the sensorimotor regions (C5 to C6, CP5 to CP6). A subselection of these electrodes is used as classification features, depending on each subject's particular results. In the first figure 6 we show the topography of the testing sessions for the most outperforming subjects at a specific frequency of interest (either in α or β band).

We then evaluate the performances of the subjects offline in terms of classification sensitivity using LDA and support vector machine (SVM) classifier with radial basis function (RBF) kernel. LDA gives a tendency that is confirmed by significant differences with SVM. The performances are based on the control set only in our offline analysis. We use a subset of 80 % for the training and 20 % for the testing. Due to the variability within session and also because the features often change between the calibration and the control phase, it is more relevant to study features when the subjects are actually in control. This is largely due to the fact that in the calibration phase, the subjects are in control of the robotic arm closing, therefore, they may change their internal strategy to complete the task. As it was the case for all subjects, we present here two R^2 maps corresponding to those calibration and control sets in figure 4 from one of the subject to demonstrate this effect. Indeed the figure shows how relevant electrodes appear to be contributing in the mental task in the control set. Moreover, a shift can be observed in the frequencies associated with the mental task. Hence the need to compute the sensitivity based on those new features. This change between calibration and control is known but hardly predictable from one subject to another and between sessions for a same subject. Our evaluation is based on features selected from the calibration and the control sets. We compute the average sensitivity across subjects for each strategy and also for each session to observe potential mechanism of learning.

Results are summarised in the table I for LDA and SVM scores. We evaluate significant differences using permutation ANOVA 2 way test to evaluate both the effect of session and strategy. We did not observe differences in the sensitivity on the Control data set between sessions but we did observe differences between strategies. The tests were applied on $\ln(\frac{x}{1-x})$ where x is the classification score per subject as this allows to highlight deviation in the data without impacting the rank. We present this specific results in figure 5.

IV. Discussion

The presented experimental setup, using multi-modal source of control (BCI and eye tracker) for the control of a robotic arm, shows promising results. The signals obtained for brain patterns associated with motor imagery are excellent in terms of R^2 values between MI and resting state for a large number of subjects. Retrieving those patterns is already a challenge by itself and the concept of the multiple modalities seems relevant as shown by the literature[5]. Studying the strategies regarding the timing to perform a motor imagery task in a BCI context seems to be relevant. Indeed, because we introduce a movement with the robot, we change the way the subject is involved in the overall sequence. Some limitations must be acknowledged, firstly because we only observe significant



(a) Calibration phase

(b) Control phase

Fig. 4: R^2 maps of the two phases in the experimentation, 3 runs of 10 MI vs 10 Rest trials per phase

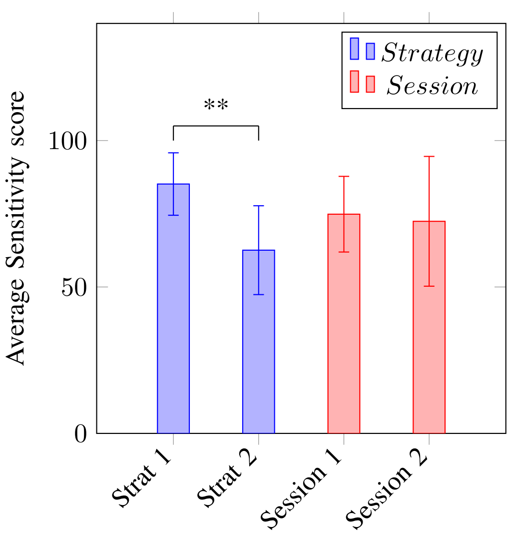


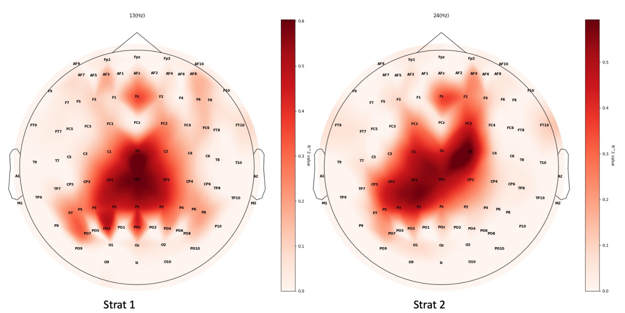
Fig. 5: In blue, average score of the sensitivity for the control set using features of the calibration and control sets between strategies. In red, average score of the sensitivity for the control set using features of the calibration and control sets between sessions. Features are based on both R^2 and Wilcoxon with $p < 0.05$ maps which present differences between MI and resting state, covering all electrodes at each frequency bin from 8 to 40 Hz. Only electrodes from the motor cortex are exploited. Standard deviation for error, (**) $p < 0.01$., SVM classifier

differences for one of the scores and secondly because it varies often from a subject to another as shown in the table I.

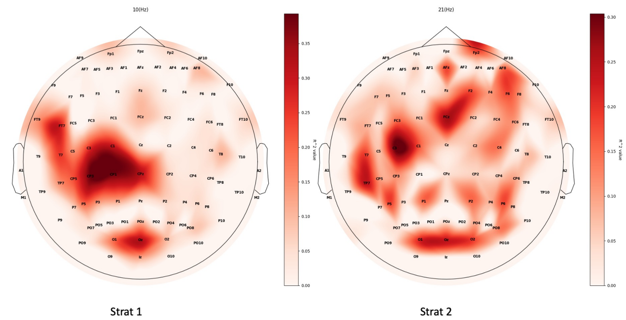
All this taken into consideration, we begin to see a trend between those strategies. In strategy 1, the robot goes to the selected target and then waits for the subject's Motor Imagery task. In this strategy, the robot's hand is at the level of the can, the motivation and the association to the motor imagery task of the closing of the hand are on average higher. In strategy 2, the robot does not move during all the task and goes at the end and grasp the object. This requires self focus for the subject because he is not presented any strong motivation stimuli to focus on. That could explain why we observe in average higher scores for the first strategy than for the second. It is also possible that it requires a higher level of expertise in MI to use strategy 2 as it is solely based on the subject's "imagination". In this specific case, it would be expected to observe less differences between strategies as the "good imaginers" Would always be more focused on themselves that on what is presented. However to discover this effect we would need more subjects to first assess who is or who is not a "good imaginer" as evoked by Allison et al[16].

It is interesting to note that the training effect on sessions does not seem to be relevant as we did not observe significant differences on the sensitivity level. As one can expect, training should take place on a longer span of time, and regular training for a year would be more noticeable than two sessions which could also explain why the statistical difference is only evaluated at $p < 0.05$. Those results need of course more subjects to be substantiated.

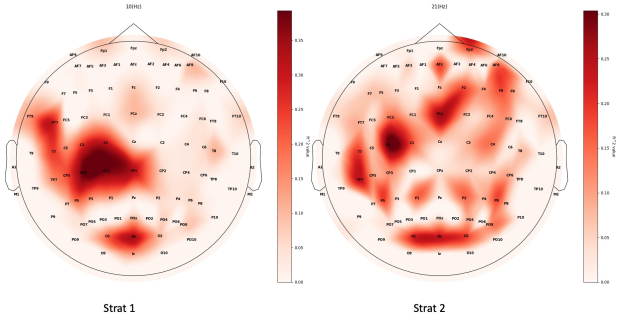
Taken together, our work presents a new framework of multi-modal BCI with the control of an arm in real time and we explore the relevancy of studying the timing to perform motor imagery in this complex system. In this paper, results show that strategy 1 where the robot is placing itself to grasp the can before MI task is giving better result. Those results are however limited by the number of subjects and this obliges to be cautious in terms of conclusions. In further works, we intend to study the ERD maps of the subjects for each strategy to determine whether there are common



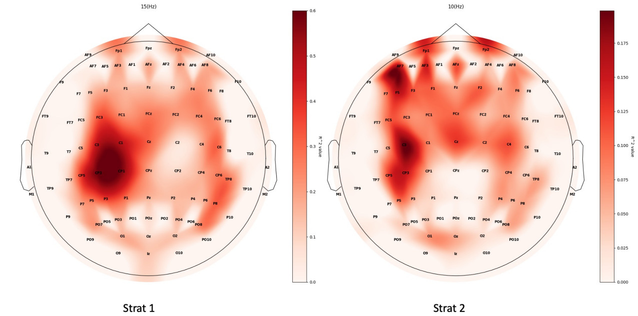
(a) Sub 2



(b) Sub 3



(c) Sub 5



(d) Sub 7

Fig. 6: Topography map for specific frequency bins (in α or β bands), R^2 value between motor imagery and resting state measured in the control phase, for the most performing subjects for the two strategies

patterns to the strategies and also when the highest peak of de-synchronization would be observed. Furthermore we intend to introduce an intermediate strategy to evaluate the effect of the robot's movement itself in the MI.

Having more data will also allow to investigate more deeply the time frequency component. Indeed, the time when motor imagery task is performed from a strategy to another could be easily seen using ERD maps as our early results seem to show.

References

[1] F. Lotte et al. “A review of classification algorithms for EEG-based brain–computer interfaces: a 10 year update”. In: *Journal of Neural Engineering* 15.3 (Feb. 2018). ISSN: 1741-2552. DOI: 10.1088/1741-2552/aab2f2.

[2] F. Yger, M. Berar, and F. Lotte. “Riemannian Approaches in Brain-Computer Interfaces: A Review”. In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 25.10 (Oct. 2017). ISSN: 1534-4320, 1558-0210. DOI: 10.1109/TNSRE.2016.2627016.

[3] H. Wang et al. “Hybrid gaze/EEG brain computer interface for robot arm control on a pick and place task”. In: 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Milan: IEEE, Aug. 2015, pp. 1476–1479. ISBN: 978-1-4244-9271-8. DOI: 10.1109/EMBC.2015.7318649.

[4] J. d. R. Millan et al. “Combining Brain–Computer Interfaces and Assistive Technologies: State-of-the-Art and Challenges”. In: *Frontiers in neuroscience* 4 (Sept. 7, 2010). DOI: 10.3389/fnins.2010.00161.

[5] D. Doherty et al. *Exploring Gaze-Motor Imagery Hybrid Brain-Computer Interface design*. Nov. 2, 2014. DOI: 10.13140/RG.2.1.3444.5287.

[6] J. Zhang and M. Wang. “A survey on robots controlled by motor imagery brain-computer interfaces”. In: *Cognitive Robotics* 1 (Jan. 1, 2021), pp. 12–24. ISSN: 2667-2413. DOI: 10.1016/j.cogr.2021.02.001.

[7] X. Zhang et al. “A survey on deep learning-based non-invasive brain signals: recent advances and new frontiers”. In: *Journal of Neural Engineering* 18.3 (Mar. 5, 2021). ISSN: 1741-2552. DOI: 10.1088/1741-2552/abc902.

[8] Y. Wang et al. “A Human-Robot Interaction System Based on Hybrid Gaze Brain-Machine Interface and Shared Control”. In: *Jiqiren/Robot* 40 (July 2018), pp. 431–439. DOI: 10.13973/j.cnki.robot.180134.

[9] T. D. Sunny et al. “Robotic Arm with Brain – Computer Interfacing”. In: *Procedia Technology*. International Conference on Emerging Trends in Engineering, Science and Technology (ICETEST - 2015) 24 (Jan. 1, 2016), pp. 1089–1096. ISSN: 2212-0173. DOI: 10.1016/j.protcy.2016.05.241.

Strategy \ Subjects	Classifier	1	2	3	4	5	6	7
Strategy 1	LDA	0.98	0.84	0.91	0.71	0.79	0.72	0.93
	SVM	0.64	0.87	0.94	0.9	0.81	0.88	0.94
Strategy 2	LDA	0.77	0.89	0.78	0.67	0.73	0.45	0.66
	SVM	0.82	0.62	0.61	0.62	0.64	0.32	0.7

TABLE I: Subject-specific classification results in terms of sensitivity, for each strategy within Control set, each time 3 runs concatenated to form 30 MI vs 30 Resting state, features selected by hand based on the best component issued from the R^2 and wilcoxon maps in the motor cortex area

[10] A. Guillot and C. Collet. "Contribution from neurophysiological and psychological methods to the study of motor imagery". In: *Brain Research Reviews* 50.2 (Dec. 2005), pp. 387–397. ISSN: 01650173. DOI: 10.1016/j.brainresrev.2005.09.004.

[11] JP. Van Acken. "Tracking the Sense of Agency in BCI Applications". PhD thesis. Aug. 31, 2012.

[12] S. Perdikis and J. d. R. Millan. "Brain-Machine Interfaces: A Tale of Two Learners". In: *IEEE Systems, Man, and Cybernetics Magazine* 6.3 (July 2020). Conference Name: IEEE Systems, Man, and Cybernetics Magazine, pp. 12–19. ISSN: 2333-942X. DOI: 10.1109/MSMC.2019.2958200.

[13] P. Diez et al. "A Comparative Study of the Performance of Different Spectral Estimation Methods for Classification of Mental Tasks". In: *Conference proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society. (EMBC)* 2008 (Feb. 2008), pp. 1155–8. DOI: 10.1109/IEMBS.2008.4649366.

[14] S. Bufalari et al. "Autoregressive spectral analysis in Brain Computer Interface context". In: *Conference proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. (EMBC)* 2006 (2006), pp. 3736–3739. ISSN: 1557-170X. DOI: 10.1109/IEMBS.2006.260238.

[15] D. Krusienski, D. McFarland, and J. Wolpaw. "An Evaluation of Autoregressive Spectral Estimation Model Order for Brain-Computer Interface Applications". In: *Conference proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society. (EMBC)* (Sept. 2006), p. 4.

[16] B. Allison and C. Neuper. "Could Anyone Use a BCI?" In: *Brain-Computer Interfaces: Applying our Minds to Human-Computer Interaction*. Ed. by Desney S. Tan and Anton Nijholt. Human-Computer Interaction Series. London: Springer, 2010, pp. 35–54. ISBN: 978-1-84996-272-8. DOI: 10.1007/978-1-84996-272-8_3.