Deep Learning prediction of BOLD-fMRI signals from simultaneous EEG during motor imagery

1st Pietro Stabile[1,2] pietro1.stabile@mail.polimi.it
2nd Neil Mehta[1,3] neilashm@andrew.cmu.edu
3rd Inês Gonçalves[1] ines.p.goncalves@tecnico.ulisboa.pt
4th Andrea Farabbi[2] andrea.farabbi@polimi.it
5th Athanasios Vourvopoulos[1] athanasios.vourvopoulos@tecnico.ulisboa.pt
6th Pulkit Grover[3] pgrover@andrew.cmu.edu
7th Patrícia Figueiredo[1] patricia.figueiredo@tecnico.ulisboa.pt

[1]Institute for Systems and Robotics (ISR) - Lisboa and Department of Bioengineering, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal.
[2]Department of Electrical, Information and Bioengineering (DEIB), Politecnico di Milano, Milan, Italy.
[3]Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, USA.

***Abstract*—Simultaneous EEG-fMRI acquisitions leverage the complementary strengths of the two functional neuroimaging modalities, with promising applications in the development of neurofeedback (NF) brain-computer interfaces (BCIs). While fMRI provides superior mapping of brain activity, EEG is more accessible for NF-BCI interventions. Previous work has attempted to identify the EEG features that best predict fMRI activity patterns, but the performance achieved is still poor. In this work, we leverage a well-established deep learning network for the classification of EEG signals, EEGNet, and propose an extension to the regression task of predicting the fMRI signal at a specific time point from concurrent EEG data (R-EEGNet). We target the activity of the somatomotor network (SMN) during the execution of two motor imagery (MI) tasks used in NF-BCIs for motor rehabilitation in stroke patients. For this purpose, we use a simultaneous EEG-fMRI dataset collected from 15 healthy subjects while executing the tasks in two separate sessions. The fMRI data are analyzed to extract a time series of MI activity for each subject and task, and the R-EEGNet model is trained to predict each fMRI time sample from a 15-seconds segment of multi-channel EEG data. We evaluated the proposed R-EEGNet model performance in comparison with a conventional machine learning model (Group Lasso) trained on EEG spectral features, as well as with a Naïve model based on the EEG somatomotor rhythm. We found that R-EEGNet achieved a similar performance to Group Lasso, both being significantly superior to the Naïve model. Our results provide the first demonstration of the ability of a subject-specific deep learning model to predict fMRI motor signals based directly on the EEG signal, without the need to extract spectral features. Future work should improve model performance through further hyperparameter optimization and the exploitation of data augmentation to cope with the typically small size of EEG-fMRI datasets.**

***Index Terms*—EEG-fMRI, Deep Learning, EEGNet, Neurofeedback, Brain computer interfaces, Motor Imagery.**

I. Introduction

The simultaneous acquisition of electroencephalography (EEG) and functional magnetic resonance imaging (fMRI) is motivated by the strong complementarity between the two modalities. While EEG measures brain activity through the electric potentials reaching the scalp with high temporal resolution but poor spatial resolution, fMRI captures the relatively slow changes in blood flow coupled to brain activity with excellent spatial localization [1] [2]. Simultaneous EEG-fMRI was first explored in epilepsy research and since then has been applied in several areas of cognitive neuroscience. More recently, it has gained popularity in the context of neurofeedback (NF) brain-computer interfaces (BCIs) [3] [4] [5] [6]. Indeed, the two techniques also greatly differ in their accessibility, with EEG presenting a significant advantage due to its portability and low cost. This is particularly relevant in NF-BCI applications, which typically require participants to undergo multiple, relatively long sessions. Therefore, it would be desirable to exploit simultaneous EEG-fMRI data to find EEG features reflecting the activity of brain networks finely mapped by fMRI, so that they could then be used on their own in EEG NF-BCIs.

Previous research has explored this problem of EEG to fMRI prediction using conventional machine learning (ML). Notably, Meir-Hasson et al. [7] introduced the EEG fingerprint method to predict the fMRI activity of the amygdala during a NF paradigm, which has been used in a series of subsequent studies. They employed linear regression with Ridge regularization based on the time-frequency representation of EEG data from a single electrode. In a similar approach, Cury et al. [8] proposed linear regression with a modified penalty to predict NF-fMRI scores from multi-channel EEG spectral power during motor imagery (MI) tasks. Simoes et al. [9] instead used nonlinear random forest regression with different kinds of EEG features for fMRI signal prediction during facial expression processing.

To the best of our knowledge, there is a lack of studies

utilizing deep learning (DL) for predicting fMRI from EEG. In response, we propose adapting EEGNet [10], a well-established and widely used DL model for EEG classification, to predict fMRI activity. EEGNet is designed for EEG-based BCIs and performs EEG classification on amplitude time series of EEG signals. The model offers several advantages, including applicability across various BCI paradigms (such as evoked potentials, error potentials, and sensory motor rhythms classification), the ability to train with limited data, and the generation of interpretable features [10].

Our approach is to retain EEGNet's feature extraction layers, and hence its ability to capture relevant spatiotemporal patterns in EEG signals, while adapting its final output to suit the regression-based nature of fMRI signal prediction. We thus introduced a novel architecture which we referred to as R-EEGNet. We aim to create subject-specific models capable of predicting fMRI activity of the somatomotor network (SMN) across MI sessions spaced over time. Moreover, we aim to predict fMRI signals reflecting trial-evoked (TE) brain activity, as in previous works, as well as spontaneous trial-by-trial (TBT) fluctuations in brain activity. We also provide insights into model explainability and compare the results obtained with the proposed R-EEGNet with a state-of-the-art ML model based on Group Lasso as well as a literature-based Naïve model [11].

II. Materials and Methods

A. Data

1) Participants and protocol: Data were acquired at Hospital da Luz Imaging Center in Lisbon, from 15 healthy subjects (6 females, 9 males; mean age 24.4 ± 2.8 years), as part of the NeurAugVR Project, with approval from the hospital's Ethics Committee. The protocol consisted of two acquisition sessions conducted approximately two weeks apart. In each session, participants were studied while performing two MI tasks: abstract MI, based on the Graz MI training paradigm, and MI combined with motor observation (MIMO), based on embodied stimulus presented through a virtual reality (VR) as part of the NeuRow training paradigm [12] [13].

Graz MI task (referred to as Graz from now on) required subjects to imagine left or right arm movements in response to a visual cue on a screen. In contrast, NeuRow MIMO task (referred as NeuRow) involved immersing participants in a VR environment, where they observed an avatar of their own arms rowing a boat to match the visual cue. Each session consisted of three runs, with 18 MI trials per run, each lasting around 10 seconds, interspersed with rest periods. This resulted in approximately 17 minutes of EEG-fMRI data per Graz or NeuRow session.

2) EEG-fMRI acquisition: BOLD-fMRI data were acquired on a 3T MRI Siemens Vida system with a 64-channel head RF coil. Whole-brain functional images were acquired using a T2*-weighted 2D multi-slice EPI (voxel size = 2.2 x 2.2 x 2.2 mm^3, TR = 1260 msec, echo time TE = 30msec, flip angle = 70°, SMS-3, GRAPPA-2, 60 axial slices). Whole-brain structural images were acquired on the same scanner using a MPRAGE sequence (voxel size = 1 x 1 x 1 mm^3, TR = 2300 msec, TE = 2.98 msec).

EEG data were collected concurrently using an MR-compatible 32-channels EEG system (Brain Products GmbH, Gilching, Germany) placed according to the 10-10 standard system. The reference electrode was positioned close to the FCz location, and one electrode placed on the back for electrocardiogram (ECG) recording. EEG, ECG, and fMRI data were acquired simultaneously in a continuous way and synchronized by the means of a Syncbox (Brainproduct) device. The sampling rate of EEG and ECG was 10kHz and no filters were applied during the recording phase.

3) EEG-fMRI pre-processing: fMRI data pre-processing was performed using FMRIB's software library (FSL) tools, including: EPI distortion correction, motion correction, high-pass temporal filtering >0.01 Hz; low-pass spatial filtering with a Gaussian kernel with full width at half maximum of 3.3 mm; and registration to MNI standard space.

EEG data pre-processing was performed using Matlab tools, including correction of fMRI-related artifacts, resampling to 250 Hz, bad channel removal, re-referencing, ICA decomposition, and removal of components related to eye and muscle artifacts. Remaining artifacts such as muscle activity bursts were mitigated using burst interpolation, followed by outlier correction and signal standardization.

4) fMRI mapping and signal extraction: The preprocessed fMRI data were submitted to a first-level voxelwise general linear model (GLM) analysis using regressors of interest (stimulus functions and their derivatives) and nuisance regressors (motion outliers and parameters). Group-level analysis of task-related fMRI changes identified a map of significantly activated brain regions across subjects, sessions, and runs. Brain activation maps were defined for each run, session and subject, as the intersection of the group-level activation map with the corresponding run-level map.

For each subject, session and run, TE fMRI signals were computed by averaging voxel time series within the respective maps. Additionally, TBT fMRI signals were obtained by first removing the contribution of task-related regressors of interest from voxel time series and then averaging the residuals within the same maps. Finally, both TE and TBT fMRI signals were normalized (zero mean, unit variance) and concatenated across runs, to yield the target signals for each subject and session.

B. EEG-fMRI modeling

1) EEG-fMRI dataset: The EEG-fMRI dataset to be used for model estimation was created by pairing each fMRI sample with a 15-seconds segment of preceding EEG data. This approach is based on the known delay of the haemodynamic response measured by the fMRI signal, allowing for this to vary between 0 and 15-seconds. Indeed, although a canonical delay of around 6 seconds is often used, a more flexible model is important to take into account for potential variations across different brain regions and individuals [14]. We filtered EEG data applying a bandpass filter (8–30 Hz) using FIR filters (Hann window, order 32) to capture MI-related activity

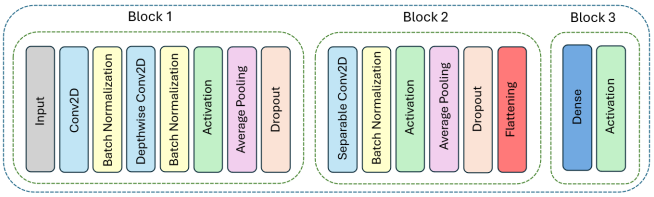

Fig. 1. *R-EEGNet model architecture.* The model consists of 3 blocks, including an initial temporal convolution layer followed by spatial depthwise convolution, and the separable convolution to aggregate temporal and spatial features from EEG data. Batch normalization and dropout layers enhance generalization, while average pooling allows data size reduction.

[15], and then further down-sampling to 80 Hz to reduce computational load and improve generalization.

We generated 60 subject- and condition-specific EEG-fMRI datasets from 15 subjects and 4 conditions, corresponding to the 2 tasks, each with TE and TBT activities conditions: Graz - TE, Graz - TBT, NeuRow - TE, and NeuRow - TBT. For each subject and condition, the three runs of each session were concatenated while the two session were kept separate. The input data were structured as a $[S \times N \times C \times T]$ tensor, where S is the number of sessions ($S = 2$), N is the number of EEG-fMRI time samples over each session ($N = 792$), C is the number of EEG channels ($C = 31$), and T is the number of EEG time points in each 15-seconds segment ($T = 3750$). The outputs were $[S \times N]$ tensors.

2) R-EEGNet model architecture: The architecture of our proposed R-EEGNet model is presented in Fig. 1. The model retains the original Block 1 and Block 2 of EEGNet. Block 1 performs temporal and spatial filtering using two sequential convolutional steps. It first applies 1D convolutions with F1 temporal filters designed to decompose EEG signals into different frequency bands. This is followed by a depthwise convolutional layer that applies D spatial filters to each frequency-specific input map. Batch normalization and an ELU activation layer follow. Block 2 combines extracted features through separable convolutions. It first applies a depthwise convolution (groups = 1) to summarize individual feature maps in time, followed by F2 pointwise convolutions to merge these summaries across feature maps. This approach differs from traditional convolutions by decoupling relationships both within and across feature maps, thus reducing the parameter count and improving feature extraction efficiency. Finally, we modified Block 3 of the original EEGNet architecture, by replacing the original multiple neurons with sigmoid activations by a single neuron with a linear activation function, omitting the bias term. This change resulted in a regressor version of EEGNet, which we call R-EEGNet.

3) Model training and cross-validation: A nested M-K-fold cross-validation (CV) approach was employed to optimize hyperparameters and evaluate model performance. The inner loop partitioned the data in each session into the three runs ($K = 3$) for hyperparameter optimization. The outer loop partitioned the data into the two sessions ($M = 2$) to evaluate the performance of the model selected in each session on the other session. Models were trained using Adam optimizer, with the mean squared error (MSE) as the loss function.

We identified the following key hyperparameters to be optimized: $F1$ (number of temporal filters), D (number of spatial filter groups), and $F2$ (number of pointwise filters). A grid search was conducted to identify the filter settings that yielded the lowest MSE averaged across inner cross-validation loops. Candidate values were selected to maintain the total number of model parameters close to the original EEGNet architecture, preserving its compactness. Given that the standard EEGNet model has $N = 1646$ parameters, we tested filter combinations that resulted in a model with parameters n within the range:

$$\frac{2}{3}N \leq n \leq \frac{4}{3}N \tag{1}$$

This range was selected to avoid reducing the model's performance by making it too small, while also preventing overfitting by making it excessively large. An example of the R-EEGNet model with specific hyperparameters is presented in Fig. 2. Model hyperparameters learning rate and batch size were manually tuned through empirical trials. Batch size was set to 32 during hyperparameter optimization.

4) Model evaluation: To assess model performance and compare across different models, we computed the Pearson correlation between the estimated and the true fMRI signals and took the average across the two session in the CV-outer loop. The proposed DL model R-EEGNet was compared with two reference models: a state-of-the-art ML model, Group Lasso; and a literature-based model, Naïve model.

For Group Lasso, lagged linear regression was performed based on EEG spectral power using Group Lasso regularization [11]. The EEG spectral power were previously extracted by performing time-frequency decomposition of the EEG signal using Morlet Wavelets in 1 Hz frequency bins. The EEG spectral power time series in each channel was lagged between 0 and 15-seconds (to match the 15-seconds segment of EEG data used in R-EEGNet). The resulting multi-channel, multi-frequency, and multi-lag EEG time series were entered as regressors. The training and cross-validation scheme was identical to the one used for R-EEGNet.

For the Naïve model, we computed the total EEG band power in the alpha and beta bands (8-30 Hz), took the average between channels C3 and C4, and shifted the resulting time series by the canonical hemodynamic delay of 6 seconds.

III. Results

A. fMRI mapping and signal extraction

The group activation maps obtained for both MI tasks, Graz and NeuRow, across all runs, sessions and subjects, are presented in Fig. 3. As expected, the SMN is activated as well as parietal and occipital brain areas associated with visuomotor coordination, which are typically engaged in MI tasks [16]. Illustrative examples of the fMRI signals representing TE activity of these networks in one representative subject, session and run, are also presented in Fig. 3

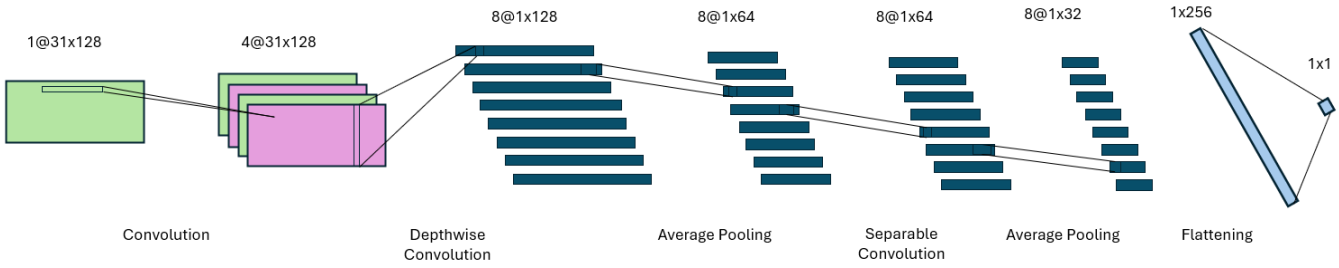

Fig. 2. *R-EEGNet model hyperparameters.* Here EEG data are structured as $[31x128]$ for a given example. The model consists of an initial temporal convolution layer ($F1 = 4$), followed by spatial depthwise convolution ($D = 2$) and separable convolution ($F2 = 8$), and includes average pooling layers with kernels of size $(1, 2)$ and $(1, 2)$.

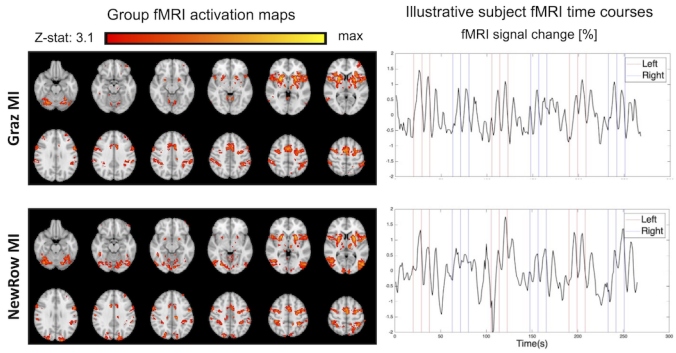

Fig. 3. *fMRI mapping and signal extraction for the Graz and NeuRow MI tasks.* (Left) Group-level fMRI activation (Z-stat) maps across runs, sessions and subjects; and (Right) Illustrative examples of the fMRI signals (percent signal change) representing TE activity of these networks in one representative subject, session and run.

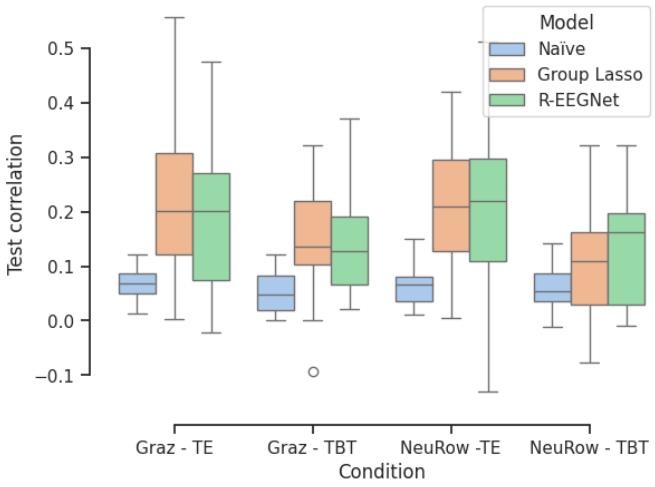

Fig. 4. *Model test performance.* Distributions across subjects of the Pearson correlation values between the predicted and true fMRI signals, obtained for the R-EEGNet model (green), in comparison with the Group Lasso model (red) and the Naive model (blue).

B. R-EEGNet model performance

The performance of the proposed R-EEGNet model across all subjects, for each of the four tested conditions, is presented in Fig. 4. We found that R-EEGNet and Group Lasso models performed similarly (median Pearson correlation values across subjects around 0.20 for the TE conditions and 0.15 for the TBT conditions), and better than the Naïve model (less that 0.10 for all conditions).

C. R-EEGNet model explainability

We investigated model explainability by analyzing the frequency response of the temporal kernels, and the weights of the spatial kernels, as shown in Fig. 5 for an illustrative subject and condition. The temporal kernels reveal which frequencies are enhanced or suppressed by the first convolutional layer, highlighting the EEG frequencies most relevant for prediction. Meanwhile, the weights of the spatial kernels determine how these extracted frequencies are further emphasized or attenuated across different EEG channels.

IV. Discussion and Conclusion

We developed and validated a DL approach, R-EEGNet, to predict fMRI responses during MI tasks directly from simultaneous EEG data, by leveraging the well-established DL EEGNet architecture originally used for EEG classification and adapting it to the problem of fMRI signal prediction in a regression task. Using simultaneous EEG-fMRI acquisitions from 15 subjects in two separate sessions, we trained subject- and condition-specific models capable of predicting fMRI activity across sessions spaced over time. We showed that R-EEGNet model was able to predict both TE brain activity as well as TBT brain activity fluctuations in two different MI tasks. The models clearly outperformed a Naïve model based on the EEG somatomotor rhythm, and were comparable to state-of-the-art linear regression based on EEG spectral features using Group Lasso regularization [11] inspired by [8], highlighting the potential of DL approaches to predict fMRI signals directly from the EEG signal amplitude during MI training sessions.

The use of models for TBT fMRI prediction demonstrated lower performance in both Graz and NeuRow cases when compared to predicting the fMRI signal with task-specific contributions. This was expected given that TBT signals reflect ongoing, spontaneous fluctuations of brain activity, unlike the

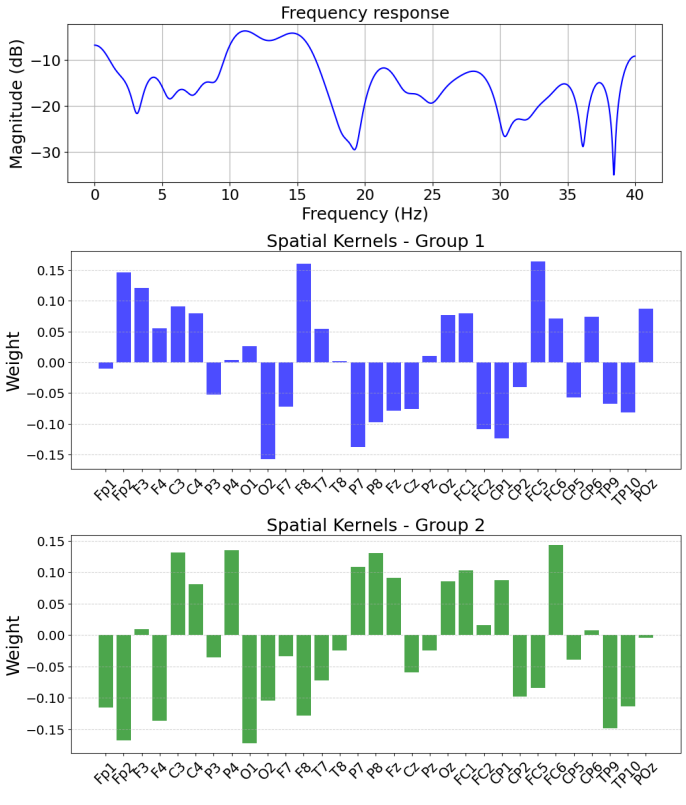

Fig. 5. *Model explainability for representative subject 07, NeuRow Task, Session 1 as training set.* (Top) Frequency response of temporal kernel: Plot of kernel number 5 out of 8 (F1 = 8). We can observe that frequencies in range *10 - 16 Hz* are enhanced by this kernel. (Bottom) Spatial kernels: Plot of group-1 and group-2 (D = 2) kernel acting on EEG channels processed by the temporal kernel number 5 (from the top).

task signals, which correspond to brain activation driven by the periodic paradigm timings. The latter not only reflect larger changes in brain activity, but they are also more likely to be correlated between modalities given the common external driver (i.e., the task paradigm). Nonetheless, it is remarkable that some predictive power can still be achieved for spontaneous fluctuations in brain activity measured by fMRI from those measured by EEG.

In sum, our results provide the first demonstration of the applicability of a DL approach to the problem of predicting both TE activity and TBT fluctuations during MI, with potential for applications in NF-BCI interventions. Future work should aim to improve model performance through further hyperparameter optimization and the exploitation of data augmentation to cope with the typically small size of EEG-fMRI datasets.

Funding

This work was supported by LARSyS FCT funding [DOI: https://doi. org/ 10.54499/ LA/P/ 0083/ 2020, https:// doi. org/ 10. 54499/ UIDP/ 50009/ 2020, and https:// doi. org/ 10. 54499/ UIDB/ 50009/ 2020], NOISyS project [DOI: 10.54499/2022.02283.PTDC], PRR project Center for Responsible AI [grant C645008882-00000055], and FCT grant PTDC/CCI-COM/31485/2017.

References

[1] R. Abreu, A. Leal, and P. Figueiredo, "EEG-informed fMRI: A review of data analysis methods," Frontiers in Human Neuroscience, vol. 12, p. 29, 2018. DOI: 10.3389/fnhum.2018.00029.

[2] Jorge J, van der Zwaag W, Figueiredo P. EEG-fMRI integration for the study of human brain function. Neuroimage. 2014 Nov 15;102 Pt 1:24-34. doi: 10.1016/j.neuroimage.2013.05.114. Epub 2013 May 31. PMID: 23732883

[3] M. Fleury, P. Figueiredo, A. Vourvopoulos, and A. Lécuyer, "Two is better? combining EEG and fMRI for BCI and neurofeedback: a systematic review.," J. Neural Eng., vol. 20, no. 5, Nov. 2023, doi: 10.1088/1741-2552/ad06e1.

[4] H.-J. Hwang, K. Kwon, and C.-H. Im, "Neurofeedback-based motor imagery training for brain-computer interface (BCI).," J. Neurosci. Methods, vol. 179, no. 1, pp. 150–156, Apr. 2009, doi: 10.1016/j.jneumeth.2009.01.015.

[5] Caetano G, Esteves I, Vourvopoulos A, Fleury M, Figueiredo P. NeuXus open-source tool for real-time artifact reduction in simultaneous EEG-fMRI. Neuroimage. 2023 Oct 15;280:120353. doi: 10.1016/j.neuroimage.2023.120353. Epub 2023 Aug 29. PMID: 37652114.

[6] Lioi, G., Cury, C., Perronnet, L. et al. Simultaneous EEG-fMRI during a neurofeedback task, a brain imaging dataset for multimodal data integration. Sci Data 7, 173 (2020). https://doi.org/10.1038/s41597-020-0498-3

[7] Y. Meir-Hasson, S. Kinreich, I. Podlipsky, T. Hendler, and N. Intrator, "An EEG Finger-Print of fMRI deep regional activation," Neuroimage, vol. 102, pp. 128–141, Nov. 2014, doi: 10.1016/J.NEUROIMAGE.2013.11.004.

[8] C. Cury, P. Maurel, R. Gribonval, and C. Barillot, "A Sparse EEG-Informed fMRI Model for Hybrid EEG-fMRI Neurofeedback Prediction," Front. Neurosci., vol. 13, p. 1451, Jan. 2020, doi: 10.3389/FNINS.2019.01451/BIBTEX

[9] M. Simões et al., "How much of the BOLD-fMRI signal can be approximated from simultaneous EEG data: relevance for the transfer and dissemination of neurofeedback interventions.," J. Neural Eng., vol. 17, no. 4, p. 46007, Jul. 2020, doi: 10.1088/1741-2552/ab9a98

[10] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, "EEGNet: a compact convolutional neural network for EEG-based brain-computer interfaces.," J. Neural Eng., vol. 15, no. 5, p. 56013, Oct. 2018, doi: 10.1088/1741- 2552/aace8c.

[11] Mehta N, Gonçalves I, Montagna A, Fleury M, Caetano G, Esteves U, Vourvopoulos A, Grover P, Figueiredo P, "Time-varying EEG spectral power predicts evoked and spontaneous BOLD-fMRI motor brain activity", MAGMA, Book of Abstracts ESMRMB 2024 Online 40th Annual Scientific Meeting 2–5 October 2024.

[12] Vourvopoulos A, Ferreira A, I Badia SB. NeuRow: an immersive VR environment for motor-imagery training with the use of brain-computer interfaces and vibrotactile feedback. In: International Conference on Physiological Computing Systems. SciTePress; 2016. Vol. 2. p. 43–53.

[13] Vourvopoulos A, Jorge C, Abreu R, Figueiredo P, Fernandes JC, Bermudez i Badia S, "Efficacy and brain imaging correlates of an immersive motor imagery BCI-driven VR system for upper limb motor rehabilitation: A clinical case report", Frontiers in human neuroscience. 2019;13:244.

[14] D. A. Handwerker, J. M. Ollinger, and M. D'Esposito, "Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses.," Neuroimage, vol. 21, no. 4, pp. 1639–1651, Apr. 2004, doi: 10.1016/j.neuroimage.2003.11.029.

[15] Pfurtscheller G, Lopes da Silva FH. Event-related EEG/MEG synchronization and desynchronization: basic principles. Clin Neurophysiol. 1999 Nov;110(11):1842-57. doi: 10.1016/s1388-2457(99)00141-8. PMID: 10576479.

[16] Nunes JD, Vourvopoulos A, Blanco-Mora DA, Jorge C, Fernandes J-C, Bermudez i Badia S, et al. (2023) Brain activation by a VR-based motor imagery and observation task: An fMRI study. PLoS ONE 18(9): e0291528. https://doi.org/10.1371/journal.pone.0291528