Neurophysiologically meaningful motor imagery EEG simulation with applications to data augmentation

Catalina M. Galván, Rubén D. Spies, Diego H. Milone and Victoria Peterson

***Abstract*—Motor imagery-based Brain-Computer Interfaces (MI-BCIs) have gained a lot of attention due to their potential usability in neurorehabilitation and neuroprosthetics. However, the accurate recognition of MI patterns in electroencephalography signals (EEG) is hindered by several data-related limitations, which restrict the practical utilization of these systems. Moreover, leveraging deep learning (DL) models for MI decoding is challenged by the difficulty of accessing user-specific MI-EEG data on large scales. Simulated MI-EEG signals can be useful to address these issues, providing well-defined data for the validation of decoding models and serving as a data augmentation approach to improve the training of DL models. While substantial efforts have been dedicated to implementing effective data augmentation strategies and model-based EEG signal generation, the simulation of neurophysiologically plausible EEG-like signals has not yet been exploited in the context of data augmentation. Furthermore, none of the existing approaches have integrated user-specific neurophysiological information during the data generation process. Here, we present PySimMIBCI, a framework for generating realistic MI-EEG signals by integrating neurophysiologically meaningful activity into biophysical forward models. By means of PySimMIBCI, different user capabilities to control an MI-BCI can be simulated and fatigue effects can be included in the generated EEG. Results show that our simulated data closely resemble real data. Moreover, a proposed data augmentation strategy based on our simulated user-specific data significantly outperforms other state-of-the-art augmentation approaches, enhancing DL models performance by up to 15%.**

***Index Terms*—Brain-Computer Interfaces, Surface EEG, Motor Imagery, Deep Learning, Data Augmentation, EEG simulation.**

I. Introduction

Decoding methods for the development of Brain-Computer Interfaces (BCIs) are commonly based on non-invasive Electroencephalography (EEG) recordings. Despite the significant progress observed in the last decade, these systems still face several data-related challenges which restrict their use in real-world scenarios [1]. One of the main challenges arises from the inability to determine the actual mental activities of users throughout the experiment based on EEG signals. This is even more evident for endogenous BCIs, which rely on users' intentions regardless of the external stimuli presented. Consequently, the artificial generation of EEG-like signals plays a crucial role in providing well-defined data for the development, validation, and interpretation of decoding models.

Although real experiments and real data will ultimately always be needed for testing and validating experimental hypotheses, collecting large-scale data under specific experimental conditions is expensive, time-consuming, and quite often not possible. Despite the existence of publicly available EEG-BCI databases, they remain scarce, hindering the development of algorithmic solutions for these systems. Motor imagery (MI) BCIs for motor rehabilitation are particularly affected by these data limitations since the MI is an endogenous process hard to master by the BCI user [2]. In addition, the desirable MI-BCIs-based rehabilitation scenario should minimize the time spent in system calibration and maximize the feedback delivery to enhance the rehabilitative impact [3]. Data scarcity in the field of MI-BCIs has become even more critical with the advent of deep learning (DL) models, whose performance strongly depends on the volume of available data for training [4].

To address these challenges, there have been mainly two lines of work: model-based EEG signal generation and data augmentation techniques. Multiple alternatives have been proposed within each of these lines. Regarding data augmentation, one of the simplest approaches consists in generating new EEG trials by combining randomly selected segments of the few available trials [5], [6]. Similarly, channel-level recombination is presented as a data augmentation method in [7]. Another strategy consists of the generation of new training samples by the addition of Gaussian noise to each sample of the original training data, as suggested in [8], [9]. In the same line, in [8], [10] a set of augmentations based on transformations that do not affect the meaningful information of EEG data were used. In the last few years, methods using generative adversarial networks (GANs) have gained much attention among artificial MI-EEG data generation techniques [11]–[15]. For instance, in [11], a GAN model was trained to generate artificial EEG samples that exhibited notable similarity to real data in the time, frequency and time-frequency domains. A cross-subject EEG classification framework using a GAN-based model for augmentation was proposed in [15]. In particular, in order to preserve the spatial features of the EEG signals and enhance the discrimination between different classes, a specific module was introduced in the discriminator. Similarly, in [16], customized neural-based generative models were used to generate steady-state visual evoked potential EEG signals

C. M. Galván, V. Peterson and R. D. Spies are at the Instituto de Matemática Aplicada del Litoral, IMAL, UNL, CONICET, Santa Fe, Argentina (e-mail: cgalvan@santafe-conicet.gov.ar; vpeterson@santafe-conicet.gov.ar; rspies@santafe-conicet.gov.ar)

D. H. Milone is at the Instituto de Investigación en Señales, Sistemas e Inteligencia Computacional, sinc(i), FICH-UNL, CONICET, Santa Fe, Argentina. (e-mail: dmilone@sinc.unl.edu.ar).

for data augmentation purposes. Nevertheless, existing data augmentation methods, whether employing data transformations or neural networks, can generate data that disregard the key neurophysiological information from the EEG, creating signals that might not be observed in real-world scenarios [17].

From a different perspective, multiple works have been dedicated to model-based EEG signal generation through the development of neurophysiologically plausible models. These models aim to capture the underlying mechanisms and dynamics of brain activity, allowing for the creation of synthetic data that closely mimic the patterns observed in real-world recordings. In this line, in [18] the authors generated simulated EEG data by combining biologically plausible activity with user-specific head models based on magnetic resonance imaging (MRI). Similarly, the SimBCI approach [19] stood out by providing a pipeline to simulate EEG data following particular experimental BCI protocols. These advancements in generating meaningful EEG-like data have significantly contributed to the exploration and validation of novel decoding algorithms in the field of EEG analysis and BCIs. However, though relevant in the field, the level of attention, fatigue and user MI brain modulation ability have not yet been included as possible inputs in the state-of-the-art models.

Despite the great efforts invested in implementing effective data augmentation strategies and model-based MI-EEG signal generation, it is noteworthy that neurophysiologically plausible EEG-like signals have not yet been exploited in the context of data augmentation for DL. Additionally, existing neurophysiologically plausible models lack the capability to incorporate user-specific information, which could be valuable when considering data simulation as a strategy for data augmentation.

In this work, improvements are simultaneously proposed in these two specific areas: model-based EEG signal generation and data augmentation for MI-BCI decoding. Specifically, a model-based simulation framework for the generation of realistic MI-EEG-like data built by embedding user-specific neurophysiologically meaningful activity into MRI-based forward models is presented. The main objectives are two, namely: i) account for reliable and well-defined data for building EEG decoding models, and ii) augment training data for DL in a task- and user-specific manner. Our approach, which we named PySimMIBCI, extends the method presented in [19] and can be used to create artificial brain recordings that are electrophysiologically similar to real data for a given BCI user. The proposed framework allows the simulation of different subject modulation capabilities and the introduction of different sources of EEG variability and artifacts, such as fatigue and eye-movement. When user-specific information is included, it becomes an effective tool for data augmentation in the context of DL models for MI-BCIs. For this purpose, a way to construct training batches that keeps a fixed number of real and simulated trials while guaranteeing class balance is proposed. The presented framework is open source and freely available (https://github.com/catalinamagalvan/PySimMIBCI), extending previously existing implementations in the widely accepted MNE-Python Library [20].

II. A FRAMEWORK FOR REALISTIC USER-SPECIFIC MI-EEG DATA GENERATION

A. Generative forward model for BCI-EEG simulations

There are two main typical approaches in the field of EEG modeling: biophysical source modeling and statistical modeling [21]. The former proposes that the EEG signals are obtained by linear combinations of unobservable physiological sources, such as a parcel of cerebral cortex with synchronously firing neurons [22]. In contrast, statistical modeling views sources as statistical entities, in principle unconnected to anatomical concepts [23]. PySimMIBCI is based on the biophysical source modeling approach [21], motivated by the goal to generate artificial signals that are neurologically plausible. Formally, let $\boldsymbol{X} \in \mathbb{R}^{c \times t}$ be the measured EEG observations over time, where c denotes the number of EEG channels and t is the number of samples; let also $\boldsymbol{G} \in \mathbb{R}^{c \times s}$ be the leadfield or gain matrix modeling the effects of tissue conduction from neuronal sources to the scalp, with s being the number of sources, and $\boldsymbol{Z} \in \mathbb{R}^{s \times t}$ comprise the matrix of source activities in the cortex [24], including both signal and noise components. Finally, let $\boldsymbol{N} \in \mathbb{R}^{c \times t}$ be a matrix of surface noise added at the electrode level. The linear superposition model leads to

$$\boldsymbol{X} = \boldsymbol{G}\boldsymbol{Z} + \boldsymbol{N}. \tag{1}$$

This biophysical forward model linearly links the current dipoles located on the vertices of the cortical meshes to the EEG electrodes located on the scalp. In a simulation framework based on a biophysical source model, the generated EEG $\boldsymbol{X}$ is obtained by linear mixing unobserved brain sources $\boldsymbol{Z}$ through the leadfield $\boldsymbol{G}$. Such a leadfield may vary in complexity from a single-sphere model to a physiologically realistic, user-specific head model [19], [21]. To get $\boldsymbol{G}$, we employ an MRI-based model that includes physiological information about head and skull morphology, taking into account the conductivity properties of the tissues [25]. Section II-B describes how $\boldsymbol{Z}$ can be constructed for a left vs. right hand MI scenario. For simplicity, $\boldsymbol{N}$ is not considered in this work.

The framework described above shows how EEG data can be modeled from a biophysical perspective. However, when considering BCIs, it is essential to incorporate specific temporal and spatial information associated with the particular BCI protocol. Therefore, to simulate informative EEG data from the decoding perspective of BCIs, our framework, PySimMIBCI, integrates this information into the biophysical generative model. Each signal component is characterized by three attributes: events, waveform and spatial information, which are equivalent to the *when*, *what* and *where* properties, proposed in [19]. The interpretation of these characteristics is very intuitive: the events information (*when*) defines the specific onsets and durations on a timeline; the waveform information (*what*) specifies the signal content introduced in each event (e.g. synchronizations/desynchronizations in specific EEG bands or particularly evoked responses); and the spatial information (*where*) indicates the cortical location of the source activity. Note that this approach assumes independence among signal components, allowing to analyze the effects of varying these parameters during the generation process.

B. Simulating left vs. right hand MI-EEG data

Different neural signals can be used to design and build a BCI. In particular, BCIs based on the MI paradigm have a clear importance in the context of hand motor rehabilitation since it has been shown that they can provide additional benefits to conventional physiotherapy [26], [27]. As in [19], here we simulate right hand MI vs. left hand MI data. Our artificial MI-EEG-like signals are obtained as a superposition of task-dependent rhythmic activity generated on the motor cortex and task-unrelated activity, which includes background noise and eye-related artifacts. The temporal and spatial properties of each component in the artificial signals are described below.

a) Task-related activity: It is well-known that hand MI causes different event-related synchronization (ERS) and desynchronization (ERD) patterns in multiple bands of the EEG in the sensorimotor cortex [28]. Briefly, an ERD corresponds to a decrease in power, whereas an ERS represents an increase in power relative to a baseline. The most relevant frequency bands involved in MI-EEG modulation have been found to be the α (8–14 Hz) and β (14–30 Hz) bands [29]. The proposed simulation considers the typical ERD in the α band in the contralateral hand motor area [30], [31]. Thus, the generators of this rhythm are assumed to be dipoles in the right and left hand motor areas. For the precise localization of sulcogyral structures in the cortex the *aparc.a2009s* parcellation [32] available in the FreeSurfer[1] package is employed. For the spatial information, circular patches of 30 mm radius, centered on the centroids of the right and left prefrontal area parcels are configured. Data is generated as standard Gaussian-Distributed Signals [33] filtered in order to restrict the frequency content to a narrow subband, with a specific central frequency and bandwidth within the α band. Rhythmic activity in this subband is present in both the left and right hand motor areas throughout the experiment, with a reduction in amplitude (ERD) on the contralateral side during the hand MI tasks. To simulate the acquisition protocol, an artificial timeline with randomly ordered trials of each class is generated.

b) Task-unrelated activity: One type of noise and two types of artifacts are considered in our simulation framework: background noise, blink artifact and eye-movement artifact, as in [34]. Background noise comprises aperiodic neural activity, that is, spectrally pink components related to neural background processes [18]. It is modeled by $1/f^{\lambda}$. This component is consistently present throughout the cortex surface for all frequencies f. In order to generate it, temporally and spatially uncorrelated white Gaussian noise is temporally filtered to have a specific power spectral density (PSD) according to the noise exponent λ. Subsequently, by means of a coloring transformation [35], the resulting components are forced to have a predetermined covariance matrix. Through this spatial filtering, a covariance structure determined by the proximity of sources, that is with closer sources exhibiting a stronger correlation, is obtained. Furthermore, we simulate the effects of eye-related artifacts, specifically blink and eye-movement artifacts. For the blink artifact, the *mne.simulation.add_eog* function of MNE-Python [20] is used. In this implementation, random activation onsets are drawn from an inhomogeneous Poisson process with blink rate oscillating between 4.5 and 17 blinks/minute according to the low and high blink rates [36]. The activation kernel consists of a 250 ms Hanning window. Two activated dipoles are located in the $z = 0$ plane at $\pm 30^{\circ}$ away from the midline (nasion). Eye-movement artifacts are modeled as a Poisson process. In each activation, the eye-movement is generated as random signal with maximum random duration of 500 ms. The two generator dipoles are located as in the blink artifact implementation [20].

Fig. 1 summarizes PySimMIBCI for the task of right hand MI. Firstly, as shown in the left side of the figure, source activity is simulated by combining specific spatial, waveform and events' information. In this example, spatial information comprises right and left hand motor areas and task-related waveform is the characteristic ERD in the α band in the contralateral hand motor area. Events' information involves a typical protocol stimulation timeline for MI: 4 s of MI followed by an inter-trial break of variable duration. All this information is mixed at the source space and gives origin to the source activity $\boldsymbol{Z}$. Through a forward projection, characterized by a leadfield matrix $\boldsymbol{G}$, this information can be accessed at the sensor level, where different types of noises can be added via matrix $\boldsymbol{N}$ resulting in the simulated EEG signal $\boldsymbol{X}$.

C. Simulation of users with different brain modulation capabilities and inclusion of fatigue effects

When developing EEG decoding methods, it is often useful to generate data corresponding to different groups of users with diverse levels of proficiency in controlling an MI-BCI. This enables researchers to test hypotheses and gain insights from the obtained results. To this end, two strategies for generating artificial users with different MI modulation characteristics are proposed here. The first strategy involves manipulating the percentage of α ERD in the contralateral hand motor area described in Section II-B. Higher percentages of ERD indicate stronger modulation and thus would correspond to more proficient users, while less pronounced ERDs would be associated with users with reduced modulation capability. The second strategy simulates scenarios where the user does not perform the required task (also known as null signals in the Evoked Potentials literature [37]). This is done by including a certain proportion of failed MI trials, that is, trials without ERD in the corresponding area. Artificial users with different levels of competence to control an MI-BCI can be then simulated by separately or jointly combining these ideas, as further shown in Section III.

Another important issue is the fact that prolonged engagement in a cognitively demanding task, such as MI, can lead to mental fatigue, degrading mental activities performance [38]. In other words, attentional focus becomes less efficient and more error-prone as the time spent on the task increases. These consequences of mental fatigue are accompanied by changes in the brain activity, particularly in low-frequency bands, resulting in altered EEG signals [39]. In order to generate realistic MI-EEG signals that reflect fatigue-related effects, we propose a way to include them in the simulated data. Several

[1] http://surfer.nmr.mgh.harvard.edu/

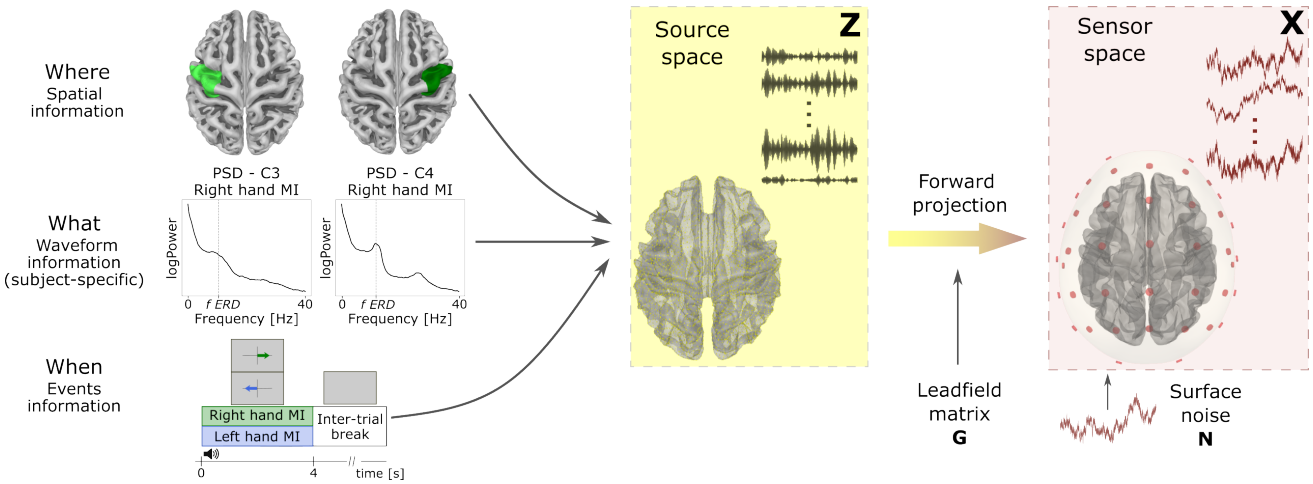

Fig. 1. PySimMIBCI pipeline exemplified by the simulation of right hand MI-EEG data. Source space data is generated by the combination of specific spatial, waveform, and events' information related to the BCI task. A forward operator is then used to get sensor space data and at this level different types of surface noise and artifacts are added.

studies of EEG oscillations have found that both θ (4-8 Hz) and α (8-13 Hz) band powers increase with mental fatigue [40]–[42]. More specifically, in [43], [44] mental fatigue has been associated with progressive increases in both frontal θ and parietal α powers with task duration. Thus, these neural markers serve as a way of simulating mental fatigue. In our simulation framework, θ and α brain generators are assumed to be dipoles in the frontal and parietal areas, respectively. As before, we generate this component by a standard Gaussian-Distributed Signal filtered to get the θ and α band activities. Fatigue can be added from any specific onset time and its level is linearly increased with time, as shown in Section III.

D. User-specific MI-EEG generation for neurophysiologically meaningful data augmentation

Data augmentation techniques aim to enlarge training datasets by artificially increasing its size and diversity. A traditional approach involves creating additional training samples through label-preserving transformations, such as the geometric transformations used in computer vision [46]. In the context of MI-EEG decoding, data consists of a collection of noisy, non-stationary time series from different electrodes and, thus, straightforward geometric transformations are not suitable since they may distort temporal and spatial-domain features. Additionally, while GAN models have been employed to generate MI-EEG data [13] and can generate data through user-specific training, the personalized neuromodulation information associated with MI can potentially be degraded or even lost during the data generation process. This becomes particularly relevant in MI decoding since the models are typically user-specific, i.e. they are trained and tested on data from the same person. To address these limitations, our proposed data simulation framework is conceived for generating user-specific MI-EEG signals for data augmentation.

The first step to generate user-specific data consists in calculating the PSD of real MI-EEG training trials for each channel by the multitaper method [47]. As shown in Fig. 2 a), three different average PSDs of interest are obtained: Average 1), the average over all right hand MI trials for channel C4; Average 2), the average over all left hand MI trials for channel C3, and Average 3), the average over all training trials and over all channels. Next, the fitting oscillations and one-over f (FOOOF) [45] method is employed to get specific parameters from the different average PSDs, as seen in Fig. 2 b). The FOOOF method is a physiologically-informed tool to parameterize neural PSDs in two constituents: aperiodic and periodic components. The aperiodic activity is exponentially decreasing with frequency and it can be described by

$$L(f) = b - \log(f^{\lambda}), \tag{2}$$

where λ is the exponent which characterizes the pattern of aperiodic power across frequencies, b is the offset that specifies the uniform shift in power and f is the frequency [45]. In contrast, periodic components of neural data consist of overlying putative periodic oscillatory peaks, which in FOOOF are individually modeled by a Gaussian-like function

$$P(f) = a \exp\left(\frac{-(f-c)^2}{2\omega^2}\right), \tag{3}$$

where a is the power, c is the central frequency and ω is the width of the peak.

It is reasonable to expect that aperiodic activity will be similarly distributed throughout the cortex and not associated with any specific MI tasks. Hence, for the aperiodic component parameterization, we employed the average PSD computed over all training trials and across all EEG channels (Average 3). On the contrary, taking into account that the two main EEG channels on the hand motor area are C3 and C4 [31], only the information in these channels is used for fitting the periodic components. That is, the average PSDs to extract periodic components are calculated over all training trials belonging to each class for both channels (Average 1 and Average 2).

After the extraction of these parameters from real MI-EEG, user-specific data is simulated as described below. The aperiodic exponent λ is employed to generate the user-specific background activity. Further, the α ERD is generated based

sinc(*i*) Research Institute for Signals, Systems and Computational Intelligence (sinc.unl.edu.ar)
C. Galván, R. Spies, D. H. Milone & V. Peterson; "Neurophysiologically meaningful motor imagery {EEG} simulation with applications to data augmentation"
IEEE Trans. on Neural Systems and Rehab. Eng., 2024.

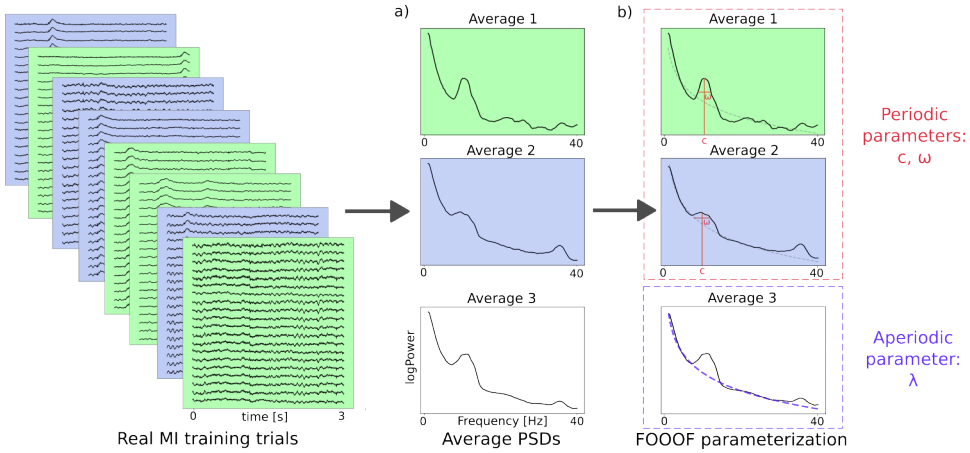

Fig. 2. Extraction of user-specific periodic and aperiodic information. a) First, PSD is calculated for each channel of each training trial. b) After that, periodic and aperiodic parameterization is done over different sets of average PSDs using FOOOF [45].

on the parameters of central frequency and bandwidth of the highest peak in the 7-14 Hz band. If the user has no peak detected in the α band, a generic peak with central frequency $c = 11.5$ Hz and bandwidth $\omega = 3$ Hz is considered, according to the most likely α peak across all real users (Fig. A1 in Supplementary Material).

III. Data and Experimental Setup

In order to evaluate the quality of our simulations, an analysis comparing the artificial MI-EEG data with their real-world counterparts was conducted. Further, to show the impact of our data augmentation approach in DL, we employed real MI-EEG data to train the model and compare its performance before and after the inclusion of simulated data. In both cases, two publicly available datasets were used in this work:

- Dataset-1 [34] comprises two sessions of left vs. right hand MI-EEG data from two users (*b* and *g*). For each user, the first session contains 100 trials per class, while the number of trials for the second session varies among users. Data was recorded at 1000 Hz using 59 electrodes densely distributed over the sensorimotor areas. Since the simulation process requires knowing the exact electrodes' position, only the 41 channels from to the 10-5 international electrode system were used. To train and test DL models data were downsampled to 250 Hz following [48].
- Dataset-2 [49] is a 54-users two-sessions left vs. right hand MI-EEG dataset. Each session comprises 100 trials per class. EEG was recorded at 1000 Hz using 62 electrodes positioned according to the international 10-20 system. For running DL models, 20 channels in the motor region were chosen, as in [49], and, here again following [48], data were downsampled to 250Hz.

As DL model for MI decoding we used the Filter Bank Convolutional Network (FBCNet) proposed in [48], a hybrid model designed for user-specific MI-BCI classification. The main reason for selecting this state-of-the-art model is because it has demonstrated superior performance in several datasets, and its implementation in PyTorch [50] is open and readily accessible. FBCNet comprises four stages: i) spectral filtering of the EEG in multiple narrow bands; ii) extraction of spatial discriminative patterns for each view by a depthwise convolutional layer; iii) compact representation of the temporal information by a variance layer; and iv) classification by a fully connected layer. In all the experiments, the architecture was trained using Adam optimizer with default settings and the log-cross-entropy loss, following the original work [48]. The model was trained using early stopping with a patience of 200 epochs. Once the stopping criterion was reached, the network parameters with the best validation accuracy were restored. The maximum number of training epochs was limited to 1500. The architecture and hyperparameters settings are described in detail in Section 7 of the Supplementary Material.

A. PySimMIBCI for realistic MI-EEG data simulation

As described in Section II-C, users with different modulation characteristics were simulated, and subsequently compared with real MI-EEG data from Dataset-1. User *g* from this dataset was employed as the guide BCI-user since it had obtained the best classification performance in the original study [34]. Neurophysiological information was extracted from this user following the steps described in Section II-D. For each simulated user, 200 artificial trials were generated (100 per class). First, ERD percentage was varied from 50% (ideal user [34]) to 10% (S10) in steps of 10%. Then, we simulated users that do not perform the required task by including an increasing proportion of trials without ERD, from 10% (user SF10) to 40% (user SF40). In the latter cases, the ERD percentage was always set to 50%. A summary of the simulation parameters can be found in Table I.

Next, to illustrate the effects of mental fatigue, we introduced the associated parietal α band and frontal θ band activities, as detailed in Section II-C. As before, user *g* from Dataset-1 was the guide BCI-user and 200 trials were generated. Considering that mental fatigue is most likely to manifest towards the end of a long session, the first half of the trials excluded fatigue effects, while the second half

TABLE I
SIMULATION PARAMETERS FOR ARTIFICIAL USERS WITH DIFFERENT BRAIN MODULATION CAPABILITIES.

User name	ERD [%]	Failed trials [%]
ideal	50	0
S40	40	0
S30	30	0
S20	20	0
S10	10	0
SF10	50	10
SF20	50	20
SF30	50	30
SF40	50	40

included a progressively increasing level of fatigue, simulated by a linearly increasing amplitude with time in both bands.

B. PySimMIBCI as a plausible data augmentation strategy

To leverage PySimMIBCI for data augmentation purposes, 100 trials per class were generated for each real user with the ideal parameters (50% of ERD and 0% bad trials) and the corresponding periodic and aperiodic extracted information, as described in Section II-D. To see the corresponding ablation study, refer to Section 5 of the Supplementary Material.

The proposed data augmentation strategy was compared with state-of-the-art approaches, which were selected to cover the main strategies presented in the literature at which method implementation was clear or publicly available. We briefly describe them below:

- Segmentation and recombination (SR): training trials (2 s) were splitted into 8 segments and new trials were created by concatenating random analogous segments from trials of the same class, as in [51].
- Gaussian noise addition (GNA): copies of each training trial were generated by adding Gaussian noise with zero mean and standard deviation equal to 0.2 [9].
- Time Masking (TM): augmented trials were derived from the original training trials by zeroing out all channels of randomly chosen 100-sample time segments [8].
- Channels Dropout (CD): new training trials were obtained by zeroing out randomly chosen channels from the original training trials. The probability of dropping each channel was 0.2, as in [10].
- Bandstop Filtering (BF): new trials were generated by applying a 5 Hz bandwidth bandstop filter, with random central frequency, between 0 and 40 Hz [8].
- Wasserstein GAN (WGAN): The model in [16] was adapted for generating MI-EEG data. Training was conducted on a user-specific basis.

In all the cases, the size of the obtained augmented dataset was twice the size of the original real MI-EEG dataset.

Baseline and augmented models share the same hyperparameter setting but differ in the training data composition. For all the tested data augmentation strategies, training batches are constructed with a fixed number of real and simulated trials, always guaranteeing class balance. More precisely, each batch comprises N_b trials and has the following structure: N_a augmented trials for each class, randomly sampled from the total

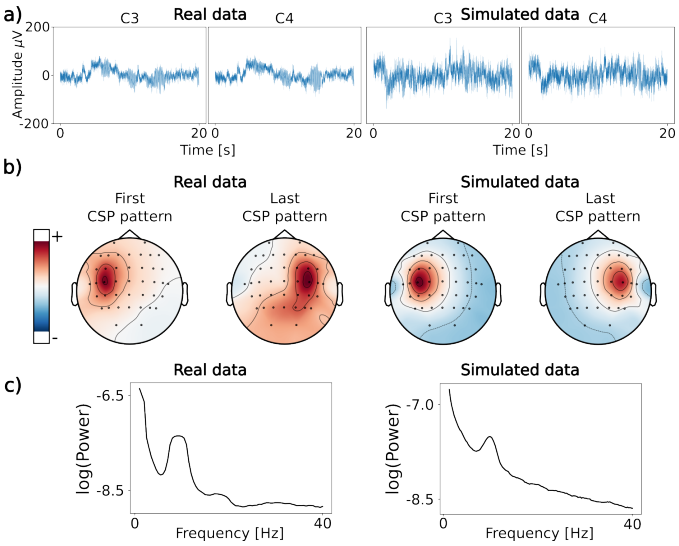

Fig. 3. Temporal, spatial and spectral plots for real and simulated EEG data. a) 20 s of both signals in channels C3 and C4. b) First and last CSP patterns. c) Average PSDs.

of simulated trials, and $N_r = (N_b - 2N_a)/2$ randomly sampled real trials of each class. This sampling approach introduces valuable variability, which has a regularizing effect on model calibration. Simultaneously, it conserves class balance within each batch and fixes real/simulated trials ratio across batches. With these considerations, we ensure that the model benefits from diverse training examples while preserving a consistent data composition across different batches (refer to Fig. A2 in Supplementary Material to see the effect of different sampling methods for batch construction). In this study we used $N_a = 1$ since no differences were found when adding more simulated trials into each training batch (see Fig. A3 in Supplementary Material).

To evaluate the impact of the proposed data augmentation strategy we used the relative change (RC) in accuracy over a baseline performance, defined as follows:

$$RC = \frac{a_a - a_r}{a_r}, \tag{4}$$

where a_a refers to test accuracy with data augmentation, that is, ratio of number of correct predictions over the total number of predictions, and a_r is the reference test accuracy (without data augmentation). If $RC > 0$, data augmentation improves model performance while if $RC < 0$, it deteriorates it.

IV. RESULTS

A. PySimMIBCI for realistic MI-EEG data simulation

To assess the similarity between artificially generated signals and real MI-EEG data, a qualitative comparison based on temporal, spectral and spatial information was first conducted. The simulated data corresponded to the 200 trials from the ideal user (Table I), while the real data comprised 200 trials from session 1 of user g from Dataset-1.

The visual comparison of real and artificial data in the time domain was made at channels C3 and C4. Fig. 3 (a) shows 20 seconds of both simulated and real signals. The spatial

assessment was made by examining the scalp topographies of the first and last CSP patterns (Fig. 3 b)), which represent time-invariant EEG spatial source distributions. These patterns provide insights into the underlying cortical activity of each MI task [52], [53]. Finally, to analyze the spectral contents of both signals, the average PSD across the 200 trials was plotted for C3 and C4 channels (Fig. 3 c)). As it can be observed, spatial, temporal and spectral traces of the simulated MI-EEG very well resemble their real counterparts.

To quantify the similarity between simulated and real data, we computed the cross-correlation between the corresponding channels of different pairs of signals: 1) simulated vs. real signal (simulated-real), 2) two real signals, i.e. user b vs. user *g* from Dataset-1 (real-real), and 3) randomly generated vs. real signal (random-real). It turns out that the differences in the maximum cross-correlation values for the pairs simulated-real and real-real are negligible, and in some cases, the similarity simulated-real exceeds that of the real-real pair. These results can be found in Fig. A4 in the Supplementary Material.

The classification performance of simulated MI-EEG data was also evaluated and compared with real data accuracy in a 10 fold cross-validation (CV) scenario. Eight folds were used for training, one fold was utilized for validation and the remaining fold for testing. Only session 1 (200 trials) of real data was part of this analysis. Regarding simulated data, the 200 trials for each one of the users of Table I were considered.

Fig. 4 a) shows the results of this CV analysis for users *b* and *g* of Dataset-1 and for the nine simulated users described in Table I. Users with decreasing percentages of ERD are represented in an orange palette while a pink palette is used for simulated users with increasing percentages of trials without ERD. As can be seen, simulated data can yield similar classification performance as real MI-EEG data. Further, as expected, lower percentages of ERD and higher proportions of bad trials lead to decreased classification performance.

To show how the effects of fatigue may influence decoding performance, signals with simulated fatigue can be used as training and testing data. In our analysis, the first half of the session described in Section III-A, where no fatigue effects were present, was used as training data. Data from the second half of the session were divided into two test partitions: one with low and one with high fatigue effects. Fig. 4 b) shows the accuracy obtained for the two test sessions. A notable decrease in accuracy is observed when mental fatigue effects are simulated over a prolonged session. Further, Fig. 4 c) illustrates the topographic maps in the θ and α bands. As expected, these plots show that, as simulated mental fatigue level increases throughout the session, powers in θ and α bands in the frontal and parietal areas increase as well.

B. PySimMIBCI as a plausible data augmentation strategy

In this section, we show the impact of using model-based user-specific simulated data as a data augmentation strategy for MI-BCI decoding based on DL. The results correspond to a within-subject cross-session analysis. In this analysis, for each subject, the data from session 1 was used for model training while session 2 data was used for model testing. The training data was further divided into a training and a validation set (25 trials per class). To account for variations in model initialization, this process was repeated 10 times using different random seeds. Fig. 5 shows the RC yielded by each tested data augmentation method in the within-subject cross-session scenario. The mean RC values across 10 random seeds are plotted for a total of 56 users (dark squares: Dataset-1, light circles: Dataset-2). PySimMIBCI achieves the highest average accuracy value ($66.72\% \pm 2.65\%$), followed by GNA ($65.74\% \pm 2.35\%$), STM ($65.64\% \pm 1.71\%$), CD ($65.54\% \pm 2.03\%$), BF ($65.32\% \pm 1.75\%$), SR ($65.21\% \pm 2.28\%$) and WGAN ($64.08\% \pm 2.69\%$). The baseline accuracy (i.e. without data augmentation) was $64.98\% \pm 2.42\%$. Remarkably, PySimMIBCI is the only approach that does not result in a performance decrease of more than 5% for any of the users.

The statistical analysis supports the effectiveness of the proposed method. The Friedman test [54] rejected the null hypothesis (p-value < 0.05) that there is no difference in the central tendency of the methods. Additionally, the post-hoc Nemenyi test [55] revealed that PySimMIBCI, GNA and STM methods are significantly better (p-value < 0.05) than the baseline. It also showed that PySimMIBCI is significantly better than SR, WGAN, BF, CD and STM, but similar to GNA. An evaluation of PySimMIBCI across four additional EEG decoding models with varying degrees of complexity and diverse input types can be found in Supplementary Material, Section 6.

V. Discussions and conclusions

In this work, we introduced PySimMIBCI, a novel approach to generate realistic and user-specific MI-EEG-like signals by embedding neurophysiologically meaningful activity into MRI-based forward models. Different user capabilities to generate distinguishable brain MI patterns were simulated and for the first time the effect of fatigue in EEG signals was considered in artificial data generation. By leveraging these user-specific simulated data for data augmentation we were able to significantly improve the training of DL models with application in MI-BCI decoding.

A complete framework to generate realistic MI-EEG data for right vs. left hand MI scenarios was presented in the first part of this work. Our results convincingly demonstrated the electrophysiological similarity between our simulated data and real MI-EEG data. The temporal traces of the simulated signals exhibited comparable amplitudes and visual appearance to actual EEG data (Fig. 3 a)). Moreover, when analyzing the CSP patterns, we observed a clear correspondence in the spatial properties of the simulated and real data (Fig. 3 b)). These patterns further reinforced the neurophysiological plausibility of the simulations, as it can be verified by the amplitudes in the hand motor areas where ERDs take place in both real and simulated MI-EEG data. Additionally, the spectral characteristics of the simulated signals closely resemble those of real signals, with both exhibiting a typical $1/f^{\lambda}$ shaped spectrum and a noticeable peak in the α band (Fig. 3 c). It is worth clarifying that the α peak is not always as prominent as shown in Fig. 3 c in real MI-EEG data. It

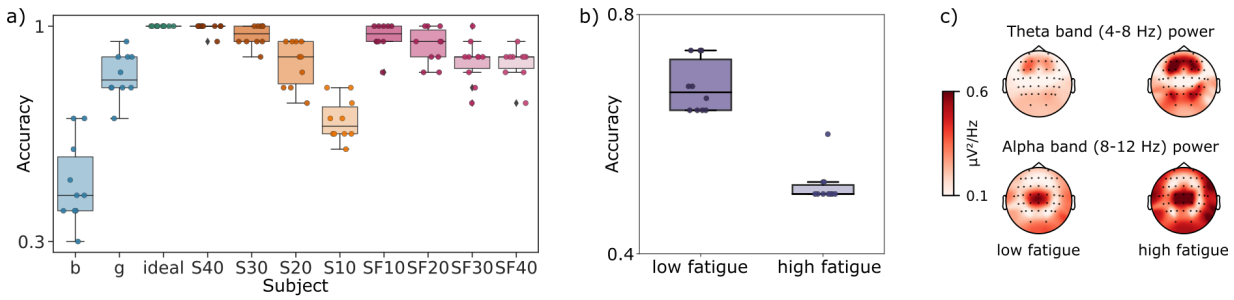

Fig. 4. Results for simulated data corresponding to different brain modulation capabilities and fatigue effects simulation. a) 10-fold CV results for real and simulated data. Each point corresponds to test accuracy for one fold. Light-blue points represent the results obtained for two real MI-EEG users, and green points the results for the ideal simulated user (50% of ERD and 0% of failed trials). Simulated users with decreasing percentages of ERD (50-10%) are in an orange palette while a pink palette was used for simulated users with increasing proportions of failed trials (0-40%). b) Accuracy values obtained for two test partitions, one with low and the other with high mental fatigue. The model was trained on data without fatigue effects. c) Topographic maps in the θ and α bands for the two test partitions with fatigue effects.

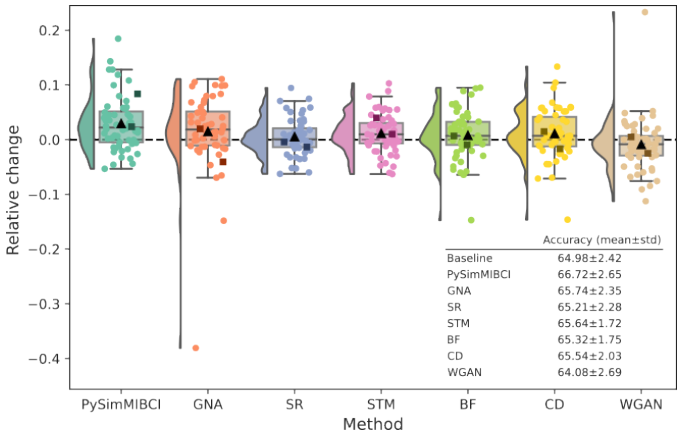

Fig. 5. Data augmentation results in two real MI-EEG datasets. Each point corresponds to the mean RC values across 10 random seeds for one real user (dark squares: Dataset-1, bright circles: Dataset-2). Average RC value across all the users is shown by a black triangle.

should be mentioned also that despite our simulations simply use the typical and well-described ERD in the α band of the contralateral hand motor area, effective results were achieved. In fact, by inspecting the spectrum of real EEG data, a clear task-related α peak was visualized in most of the cases (Fig. A1, Supplementary Material), exhibiting a variety of central frequencies and bandwidths and allowing us to parameterize it. On the contrary, that was not the case for the β band, as only 7 out of the 56 users exhibited a characterizable peak within this frequency range. Naturally, the inclusion of more complex patterns and their impact in the MI-EEG data generation process is undoubtedly an interesting topic to explore in the future.

It is timely to mention here that in our simulations two sets of user-specific parameters are considered: a) the central frequency and bandwidth in the α peak, and b) the exponent of the $1/f^{\lambda}$ aperiodic activity. This approach is in line with recent findings that emphasize the significance of non-oscillatory activity as a predictor of individual motor abilities, complementing the well-studied sensorimotor characteristics [56]. By estimating the periodic parameters we can identify user-specific variation in the task-related ERD activity. Additionally, the user-specific fitting of the aperiodic activity enables to preserve cognitive function capability and retain valuable information about individual visuomotor skills [56].

In the development of an EEG-BCI decoding model, it is essential to assess its performance with varying degrees of data distinguishability and account for typical neurophysiological changes that occur between BCI sessions. To address these issues, users with different modulation capabilities and levels of attention were modeled by employing a range of ERD percentages and introducing different proportions of failed trials, respectively. Through our experiments, we demonstrated that these two strategies effectively simulate users that achieve different classification performances. However, as it becomes evident from the results shown in Fig. 4 a), the two scenarios have different effects from the classification perspective. When the percentage of ERD decreases, the classes become almost completely overlapped, dropping the accuracy to near by-chance levels. In contrast, adding trials without ERD produces data from a new different class, which are randomly assigned to any MI class, causing, as expected, a less pronounced loss in model performance.

When trying to control a BCI, other mental states and cognitive processes might affect the ability to maintain voluntary control of the EEG activity. In particular, during long MI-BCI sessions, the required high concentration and mental effort, often lead to mental fatigue, which results in a decline in system performance [57]. This impact of mental fatigue on EEG patterns has also been observed in patients undergoing MI-BCI stroke rehabilitation, who are even more susceptible to mental fatigue than users without neurological pathologies [58]. In this work, we simulated these effects on the EEG by progressively increasing power in frontal θ and parietal α, as task duration increases. As expected, the inclusion of fatigue resulted in a noticeable decrease in classification performance (Fig. 4 b)), which was consistently accompanied by increases in frontal θ and parietal α powers, as shown in the topography maps in the bands of interest (Fig. 4 c). Although we simulated fatigue effects considering the changes in the α and θ bands, patterns that were observed in previous EEG studies, it is worth noting that other alternative approaches can also be explored. In fact, there is still a lack of consensus

sinc(*i*) Research Institute for Signals, Systems and Computational Intelligence (sinc.unl.edu.ar)
C. Galván, R. Spies, D. H. Milone & V. Peterson; "Neurophysiologically meaningful motor imagery {EEG} simulation with applications to data augmentation"
IEEE Trans. on Neural Systems and Rehab. Eng., 2024.

within the community regarding the EEG effects of mental fatigue [40]. Moreover, further experiments are crucial to assess if the synthetic fatigue effect resembles the genuine mental fatigue impact on the EEG. Importantly, our simulation of fatigue effects in the EEG highlights the possibility of introducing other task-unrelated cognitive states into artificial data generation. This implies that any additional cognitive state can be simulated, provided that the underlying brain dynamics are well-known in the temporal, spectral, and spatial domains. Addressing this aspect is part of our future plans to enhance the realism of the simulated EEG.

Finally, a data augmentation strategy based on our user-specific simulated MI-EEG data was proposed. Experimental results on FBCNet exhibited that the performance of DL models can be effectively improved when these simulated data is introduced within the training phase. It is timely to mention that such improvements were found by adding only 2 simulated samples in each training batch, a quantity three times smaller than the lowest found in the literature [14]. Moreover, our method yields a significant improvement over the baseline and the SR, BF, CD, STM and WGAN methods. This difference could be attributed to the fact that SR, BF, CD and STM generate data by simply corrupting the original trials, and WGAN focuses on generating data with the closest distribution as possible to the original dataset. Given this, neither of these methods guarantees the preservation of relevant neurophysiological information for accurate classification in the generated data. The lack of significant difference between PySimMIBCI and GNA suggests that similar classification performances can be achieved by either of these two strategies. Nevertheless, a major advantage of PySimMIBCI is that it takes into consideration neurophysiological principles and it is based on a model that enables simulation of many realistic scenarios, as opposed to GNA, which merely adds noise to real data. Moreover, an evaluation of PySimMIBCI was conducted across four additional EEG decoding models with varying degrees of complexity and diverse input types (refer to Supplementary Material, Section 6). The results suggest that the proposed method holds promise as a data augmentation strategy, especially in models proficient at extracting band power features in relevant frequency bands. These findings align with the essence of PySimMIBCI, which generates MI-EEG data by incorporating user-specific spectral information.

When moving towards the particular application of MI-BCIs for rehabilitation, effective data augmentation strategies would enable the training of decoding models with just a few trials from the target user. The beginning of the feedback phase could be accelerated and thus the applicability and usability of these systems could be greatly improved. In addition, Pysim-MIBCI for data augmentation could help reduce the levels of BCI inefficiency and alleviate frustration in users unable to generate discriminable patterns necessary for controlling an MI-BCI system [59]. However, analyses with real patient data and under rehabilitation conditions are needed to assess the continued efficacy of our method in this particular scenario.

It is important to mention that the proposed framework can be further improved or expanded in different aspects, some of which are mentioned below. Firstly, only hand MI classes were tested throughout this study. Nevertheless, it is important to highlight that PySimMIBCI can be easily extended to other MI-BCI tasks as long as the spatial, waveform and events' properties are known. Hence, in the future, we will study the generation of other types of MI that are relevant in the field of motor rehabilitation, such as hand MI vs. rest [60] and multi-class scenarios. Moreover, other user-specific information, such as personalized spatial location of generator dipoles in the cortex and the use of personalized head models will be included in the data generation process. Furthermore, it is worth noting that during our data augmentation experiments, the training batches were structured to include a fixed number of real and simulated trials. This might lead to a learning problem caused when training batches are constructed with multi-domain data. That is, training batches have a mix of simulated and real data as in a multi-domain problem. From this perspective, future studies will explore ideas to improve gradients' interference during multi-domain data training such as those of gradient surgery [61].

In conclusion, we presented and validated PySimMIBCI, a framework for generating user-specific and meaningful MI-EEG-like data. These data might be useful not only during the development of new decoding models but also as a plausible data augmentation strategy for DL-based MI-BCIs. Therefore, we strongly believe that this approach holds significant potential for advancing the development of more accurate decoding models for MI-BCI systems. In addition, our framework is open source, freely available, and compatible with implementations in MNE-Python [20], thereby promoting accessibility for a wide community of researchers.

REFERENCES

[1] S. Saha, K. A. Mamun, K. Ahmed, R. Mostafa, G. R. Naik, S. Darvishi, A. H. Khandoker, and M. Baumert, "Progress in brain computer interface: Challenges and opportunities," *Frontiers in Systems Neuroscience*, vol. 15, p. 578875, 2021.

[2] J. Mladenović, J. Frey, S. Pramij, J. Mattout, and F. Lotte, "Towards identifying optimal biased feedback for various user states and traits in motor imagery BCI," *IEEE Transactions on Biomedical Engineering*, vol. 69, no. 3, pp. 1101–1110, 2021.

[3] L. Van Dokkum, T. Ward, and I. Laffont, "Brain computer interfaces for neurorehabilitation-its current status as a rehabilitation strategy post-stroke," *Annals of Physical and Rehabilitation Medicine*, vol. 58, no. 1, pp. 3–8, 2015.

[4] Y. Roy, H. Banville, I. Albuquerque, A. Gramfort, T. H. Falk, and J. Faubert, "Deep learning-based electroencephalography analysis: a systematic review," *Journal of Neural Engineering*, vol. 16, no. 5, p. 051001, 2019.

[5] F. Lotte, "Generating artificial EEG signals to reduce BCI calibration time," in *5th International Brain-Computer Interface Workshop*, 2011, pp. 176–179.

[6] S.-J. Kim, D.-H. Lee, and Y.-W. Choi, "Cropcat: Data augmentation for smoothing the feature distribution of EEG signals," in *2023 11th International Winter Conference on Brain-Computer Interface (BCI)*. IEEE, 2023, pp. 1–4.

[7] Y. Pei, Z. Luo, Y. Yan, H. Yan, J. Jiang, W. Li, L. Xie, and E. Yin, "Data augmentation: Using channel-level recombination to improve classification performance for motor imagery eeg," *Frontiers in Human Neuroscience*, vol. 15, p. 645952, 2021.

[8] M. N. Mohsenvand, M. R. Izadi, and P. Maes, "Contrastive representation learning for electroencephalogram classification," in *Machine Learning for Health*. PMLR, 2020, pp. 238–253.

[9] F. Wang, S.-h. Zhong, J. Peng, J. Jiang, and Y. Liu, "Data augmentation for EEG-based emotion recognition with deep convolutional neural networks," in *International Conference on Multimedia Modeling*. Springer, 2018, pp. 82–93.

[10] A. Saeed, D. Grangier, O. Pietquin, and N. Zeghidour, "Learning from heterogeneous eeg signals with differentiable channel reordering," in *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2021, pp. 1255–1259.

[11] Y. Dong, X. Tang, F. Tan, Q. Li, Y. Wang, H. Zhang, J. Xie, W. Liang, G. Li, and P. Fang, "An approach for eeg data augmentation based on deep convolutional generative adversarial network," in *2022 IEEE International Conference on Cyborg and Bionic Systems (CBS)*. IEEE, 2023, pp. 347–351.

[12] F. Fahimi, S. Dosen, K. K. Ang, N. Mrachacz-Kersting, and C. Guan, "Generative adversarial networks-based data augmentation for brain-computer interface," *IEEE Transactions on Neural Networks and Learning Systems*, 2021.

[13] S. Roy, S. Dora, K. McCreadie, and G. Prasad, "MIEEG-GAN: generating artificial motor imagery electroencephalography signals," in *2020 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2020, pp. 1–8.

[14] Q. Zhang and Y. Liu, "Improving brain computer interface performance by data augmentation with conditional deep convolutional generative adversarial networks," *arXiv preprint arXiv:1806.07108*, 2018.

[15] Y. Song, L. Yang, X. Jia, and L. Xie, "Common spatial generative adversarial networks based EEG data augmentation for cross-subject brain-computer interface," *arXiv preprint arXiv:2102.04456*, 2021.

[16] N. K. N. Aznan, A. Atapour-Abarghouei, S. Bonner, J. D. Connolly, N. Al Moubayed, and T. P. Breckon, "Simulating brain signals: Creating synthetic EEG data via neural-based generative models for improved SSVEP classification," in *2019 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2019, pp. 1–8.

[17] R. Fu, Y. Wang, and C. Jia, "A new data augmentation method for EEG features based on the hybrid model of broad-deep networks," *Expert Systems with Applications*, vol. 202, p. 117386, 2022.

[18] E. Barzegaran, S. Bosse, P. J. Kohler, and A. M. Norcia, "EEGsourcesim: A framework for realistic simulation of EEG scalp data using MRI-based forward models and biologically plausible signals and noise," *Journal of Neuroscience Methods*, vol. 328, p. 108377, 2019.

[19] J. T. Lindgren, A. Merlini, A. Lecuyer, and F. P. Andriulli, "simBCI—a framework for studying BCI methods by simulated EEG," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 11, pp. 2096–2105, 2018.

[20] A. Gramfort, M. Luessi, E. Larson, D. A. Engemann, D. Strohmeier, C. Brodbeck, R. Goj, M. Jas, T. Brooks, L. Parkkonen *et al.*, "MEG and EEG data analysis with MNE-Python," *Frontiers in Neuroscience*, vol. 7, no. 267, pp. 1–13, 2013.

[21] D. Sabbagh, P. Ablin, G. Varoquaux, A. Gramfort, and D. A. Engemann, "Predictive regression modeling with MEG/EEG: from source power to signals and cognitive states," *NeuroImage*, vol. 222, p. 116893, 2020.

[22] L. C. Parra, C. D. Spence, A. D. Gerson, and P. Sajda, "Recipes for the linear analysis of EEG," *Neuroimage*, vol. 28, no. 2, pp. 326–341, 2005.

[23] S. Makeig, A. Bell, T.-P. Jung, and T. J. Sejnowski, "Independent component analysis of electroencephalographic data," *Advances in Neural Information Processing Systems*, vol. 8, 1995.

[24] S. Baillet, J. C. Mosher, and R. M. Leahy, "Electromagnetic brain mapping," *IEEE Signal processing magazine*, vol. 18, no. 6, pp. 14–30, 2001.

[25] C. M. Michel and D. Brunet, "Eeg source imaging: a practical review of the analysis steps," *Frontiers in Neurology*, vol. 10, p. 325, 2019.

[26] A. Zimmermann-Schlatter, C. Schuster, M. A. Puhan, E. Siekierka, and J. Steurer, "Efficacy of motor imagery in post-stroke rehabilitation: a systematic review," *Journal of Neuroengineering and Rehabilitation*, vol. 5, no. 1, pp. 1–10, 2008.

[27] H. C. Dijkerman, M. Ietswaart, M. Johnston, and R. S. MacWalter, "Does motor imagery training improve hand function in chronic stroke patients? a pilot study," *Clinical Rehabilitation*, vol. 18, no. 5, pp. 538–549, 2004.

[28] G. Pfurtscheller and F. L. Da Silva, "Event-related EEG/MEG synchronization and desynchronization: basic principles," *Clinical Neurophysiology*, vol. 110, no. 11, pp. 1842–1857, 1999.

[29] G. Pfurtscheller and C. Neuper, "Motor imagery activates primary sensorimotor area in humans," *Neuroscience letters*, vol. 239, no. 2-3, pp. 65–68, 1997.

[30] Y. Jeon, C. S. Nam, Y.-J. Kim, and M. C. Whang, "Event-related (de) synchronization (ERD/ERS) during motor imagery tasks: Implications for brain–computer interfaces," *International Journal of Industrial Ergonomics*, vol. 41, no. 5, pp. 428–436, 2011.

[31] G. Pfurtscheller, C. Brunner, A. Schlögl, and F. L. Da Silva, "Mu rhythm (de) synchronization and EEG single-trial classification of different motor imagery tasks," *NeuroImage*, vol. 31, no. 1, pp. 153–159, 2006.

[32] C. Destrieux, B. Fischl, A. Dale, and E. Halgren, "Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature," *Neuroimage*, vol. 53, no. 1, pp. 1–15, 2010.

[33] S. J. Orfanidis, *Introduction to signal processing*. Prentice-Hall, Inc., 1995.

[34] M. Tangermann, K.-R. Müller, A. Aertsen, N. Birbaumer, C. Braun, C. Brunner, R. Leeb, C. Mehring, G. Mueller-Putz, G. Nolte *et al.*, "Review of the BCI competition IV," *Frontiers in Neuroscience*, vol. 6, p. 55, 2012.

[35] M. Hossain, "Whitening and coloring transformations for multivariate gaussian data," *A lecture partly based on the ECE662 Spring*, 2014.

[36] A. R. Bentivoglio, S. B. Bressman, E. Cassetta, D. Carretta, P. Tonali, and A. Albanese, "Analysis of blink rate patterns in normal subjects," *Movement disorders*, vol. 12, no. 6, pp. 1028–1034, 1997.

[37] K. B. Walhovd, H. Rosquist, and A. M. Fjell, "P300 amplitude age reductions are not caused by latency jitter," *Psychophysiology*, vol. 45, no. 4, pp. 545–553, 2008.

[38] Y. Tran, A. Craig, R. Craig, R. Chai, and H. Nguyen, "The influence of mental fatigue on brain activity: Evidence from a systematic review with meta-analyses," *Psychophysiology*, vol. 57, no. 5, p. e13554, 2020.

[39] T. Cao, F. Wan, C. M. Wong, J. N. da Cruz, and Y. Hu, "Objective evaluation of fatigue by eeg spectral analysis in steady-state visual evoked potential-based brain-computer interfaces," *Biomedical Engineering online*, vol. 13, no. 1, pp. 1–13, 2014.

[40] T. Jacquet, R. Lepers, B. Poulin-Charronnat, P. Bard, P. Pfister, and B. Pageaux, "Mental fatigue induced by prolonged motor imagery increases perception of effort and the activity of motor areas," *Neuropsychologia*, vol. 150, p. 107701, 2021.

[41] S. K. Lal and A. Craig, "Driver fatigue: electroencephalography and psychological assessment," *Psychophysiology*, vol. 39, no. 3, pp. 313–321, 2002.

[42] J. Phipps-Nelson, J. R. Redman, and S. M. Rajaratnam, "Temporal profile of prolonged, night-time driving performance: breaks from driving temporarily reduce time-on-task fatigue but not sleepiness," *Journal of Sleep Research*, vol. 20, no. 3, pp. 404–415, 2011.

[43] L. J. Trejo, K. Kubitz, R. Rosipal, R. L. Kochavi, and L. D. Montgomery, "EEG-based estimation and classification of mental fatigue," *Psychology*, vol. 6, no. 05, p. 572, 2015.

[44] U. Talukdar and S. M. Hazarika, "Estimation of mental fatigue during EEG based motor imagery," in *Intelligent Human Computer Interaction: 8th International Conference, IHCI 2016, Pilani, India, December 12-13, 2016, Proceedings 8*. Springer, 2017, pp. 122–132.

[45] T. Donoghue, M. Haller, E. J. Peterson, P. Varma, P. Sebastian, R. Gao, T. Noto, A. H. Lara, J. D. Wallis, R. T. Knight *et al.*, "Parameterizing neural power spectra into periodic and aperiodic components," *Nature Neuroscience*, vol. 23, no. 12, pp. 1655–1665, 2020.

[46] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *Journal of big data*, vol. 6, no. 1, pp. 1–48, 2019.

[47] D. B. Percival and A. T. Walden, *Spectral analysis for physical applications*. cambridge University press, 1993.

[48] R. Mane, E. Chew, K. Chua, K. K. Ang, N. Robinson, A. P. Vinod, S.-W. Lee, and C. Guan, "FBCNet: A multi-view convolutional neural network for brain-computer interface," *arXiv preprint arXiv:2104.01233*, 2021.

[49] M.-H. Lee, O.-Y. Kwon, Y.-J. Kim, H.-K. Kim, Y.-E. Lee, J. Williamson, S. Fazli, and S.-W. Lee, "EEG dataset and OpenBMI toolbox for three BCI paradigms: An investigation into BCI illiteracy," *GigaScience*, vol. 8, no. 5, p. giz002, 2019.

[50] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga *et al.*, "Pytorch: An imperative style, high-performance deep learning library," *Advances in Neural Information Processing Systems*, vol. 32, 2019.

[51] F. Lotte, "Signal processing approaches to minimize or suppress calibration time in oscillatory activity-based brain–computer interfaces," *Proceedings of the IEEE*, vol. 103, no. 6, pp. 871–890, 2015.

[52] H. Ramoser, J. Muller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 4, pp. 441–446, 2000.

[53] Y. Wang, S. Gao, and X. Gao, "Common spatial pattern method for channel selection in motor imagery based brain-computer interface," in *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*. IEEE, 2006, pp. 5392–5395.

[54] M. Friedman, "The use of ranks to avoid the assumption of normality implicit in the analysis of variance," *Journal of the American Statistical Association*, vol. 32, no. 200, pp. 675–701, 1937.

[55] P. B. Nemenyi, *Distribution-free multiple comparisons.* Princeton University, 1963.

[56] M. A. Immink, Z. R. Cross, A. Chatburn, J. Baumeister, M. Schlesewsky, and I. Bornkessel-Schlesewsky, "Resting-state aperiodic neural dynamics predict individual differences in visuomotor performance and learning," *Human Movement Science*, vol. 78, p. 102829, 2021.

[57] E. A. Curran and M. J. Stokes, "Learning to control brain activity: A review of the production and control of EEG components for driving brain–computer interface (BCI) systems," *Brain and Cognition*, vol. 51, no. 3, pp. 326–336, 2003.

[58] R. Foong, K. K. Ang, C. Quek, C. Guan, K. S. Phua, C. W. K. Kuah, V. A. Deshmukh, L. H. L. Yam, D. K. Rajeswaran, N. Tang *et al.*, "Assessment of the efficacy of EEG-based MI-BCI with visual feedback and EEG correlates of mental fatigue for upper-limb stroke rehabilitation," *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 3, pp. 786–795, 2019.

[59] G. Dornhege, J. del R. Millán, T. Hinterberger, D. J. McFarland, and K.-R. Müller, *An Introduction to Brain-Computer Interfacing*, 2007, pp. 1–25.

[60] K. K. Ang, C. Guan, K. S. Phua, C. Wang, L. Zhao, W. P. Teo, C. Chen, Y. S. Ng, and E. Chew, "Facilitating effects of transcranial direct current stimulation on motor imagery brain-computer interface with robotic feedback for stroke rehabilitation," *Archives of Physical Medicine and Rehabilitation*, vol. 96, no. 3, pp. S79–S87, 2015.

[61] L. Mansilla, R. Echeveste, D. H. Milone, and E. Ferrante, "Domain generalization via gradient surgery," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 6630–6638.