

ASTERI: Image-based representation of EEG signals for motor imagery classification

Juliana Carneiro Gomes

Universidade de Pernambuco

Marcelo Cairrão de Araújo Rodrigues

Universidade Federal de Pernambuco

Wellington Pinheiro dos Santos (✉ wellington.santos@ufpe.br)

Universidade Federal de Pernambuco <https://orcid.org/0000-0003-2558-6602>

Research Article

Keywords: Motor imagery, BCI, Representation of signals as images, Transfer Learning

Posted Date: April 7th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1041386/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Research on Biomedical Engineering on April 19th, 2022. See the published version at <https://doi.org/10.1007/s42600-022-00215-1>.

ASTERI: Image-based representation of EEG signals for motor imagery classification

Juliana Carneiro Gomes ·
Marcelo Cairrão Araújo Rodrigues ·
Wellington Pinheiro dos Santos

the date of receipt and acceptance should be inserted later

Abstract Purpose: Electroencephalography (EEG) signals are valuable in the monitoring and investigation of neurological diseases and in the control of brain-machine interfaces (BCI). However, these signals are noisy and are non-linear and non-stationary in nature. Signal analysis is an expensive task and can lead to misdiagnosis. Deep learning can be used to overcome these challenges. The most used deep architectures are based on convolutional neural networks (CNNs). Representing EEG signals as images can be useful to use deep architectures based on CNNs in solutions based on intelligent EEG-based systems. **Methods:** In this work, we propose the ASTERI method, for representing EEG signals in two dimensions using the backprojection reconstruction method. To validate the proposal, experiments were performed with motor imagery EEG signals from the BCI Competition IV 2b motor imagery database. To extract attributes from the EEG signal windows, we used the ASTERI representation and investigated the pre-trained networks VGG16 and LeNet. To classify the final classification of imagined movements to the right and to the left, we used Random Forests. **Results:** For the original database, the validation with 100 trees achieved an average accuracy and kappa index of 88.97% and 0.78, respectively. In the case of augmented databases with synthetic instances, the average accuracies were 89.10% and 89.13% for increases of 50% and 150% for each class, respectively. **Conclusion:** The experimental results showed that the ASTERI method, together with hybrid deep architectures based on CNN and Random Forests, can generate competitive results in solving motor imagery problems. The representation of signals as images can also contribute to the exploration of new solutions of applied neuroscience problems based on deep architectures.

Keywords Motor imagery · BCI · Representation of signals as images · Transfer Learning

Corresponding author: Wellington Pinheiro dos Santos, E-mail: wellington.santos@ufpe.br, ORCID
00000003-2558-6602.

Juliana Carneiro Gomes

Polytechnique School of the University of Pernambuco, Recife, Brazil E-mail:
jcg@ecomp.poli.br

Marcelo Cairrão Araújo Rodrigues

Department of Physiology and Pharmacology, Federal University of Pernambuco, Recife, Brazil E-mail:
marcelo.carodrigues@ufpe.br

Wellington Pinheiro dos Santos

Department of Biomedical Engineering, Federal University of Pernambuco, Recife, Brazil
E-mail: wellington.santos@ufpe.br

1 Introduction

1.1 Context and motivation

Electroencephalography (EEG) is a technique that detects brain activity over a period of time. This means that it detects voltage fluctuations due to the flow of ionic current between neurons in the brain. Currently, EEG signals are valuable in the monitoring and investigation of neurological diseases. Abnormalities such as excessive and uncontrolled brain activity, slow rhythms or altered synchronization may be indicative of neuronal dysfunction or brain degeneration, for instance (Abiri et al., 2019, Craik et al., 2019).

In clinical practice, electroencephalography has been used in several contexts: diagnosis of epilepsy (Chen et al., 2020, da Silva Junior et al., 2019, Ullah et al., 2018), Alzheimer's disease (AD) (Cassani et al., 2018, Ieracitano et al., 2019), Parkinson's disease (de Oliveira et al., 2020, Oh et al., 2020), autism (Ibrahim et al., 2018), in people with brain injury resulting from tumor or stroke (Kawano et al., 2017), and in the evaluation of sleep disorders (Islam et al., 2018).

Electroencephalography is a promising technique due to its low cost when compared to other diagnostic support techniques. It is also portable and can be easily applied to the bedside of inpatients, clinics and home care cases. Electroencephalography equipment is easy to handle and the exam is non-invasive and painless Ieracitano et al. (2019).

Despite these many benefits, EEG signals are highly complex, non-linear and nonstationary Ieracitano et al. (2019). The signals generally occupy a range between 0 and 100 Hz and an amplitude between 2 μV and 200 μV Niedermeyer et al. (1985). It is important to highlight that electroencephalographic signals can also be quite noisy, due to artifacts of muscle, eye, tongue and heart movement. Non-physiological artifacts are also present, such as 60 Hz artifacts resulting from interference from the electrical network, artifacts due to electrode malposition, and electrical and electromagnetic interference from other medical and hospital equipment Rampal et al. (2018).

Analyzing EEG signals is a difficult task that requires visual inspection by a well-trained medical specialist. Therefore, this process becomes susceptible to the subjectivity of the specialist, in addition to demanding considerable time Thomas et al. (2020), TjepkemaCloostermans et al. (2018). As a consequence, diagnostic errors are common, which can lead to poor prognosis Benbadis (2010). Amin and Benbadis (2019) point out that approximately 25% to 30% of patients diagnosed with epilepsy, through EEG signs, do not have epilepsy. Some patients had been misdiagnosed for more than 10 years Amin and Benbadis (2019).

Although there are other ways for noise reduction or signal recovery such as curve fitting, attenuation filter for noisy signal (Luccas et al., 2021), several researchers have

sought to overcome these challenges through the use of machine learning (ML) (AlAbboudi et al., 2020, Gayathri et al., 2018, Gemein et al., 2020, Ieracitano et al., 2021, Isa et al., 2019, Mardini et al., 2020, Rashid et al., 2018). Among the various areas within machine learning, deep learning (DL) algorithms have stood out in recent years. Deep learning is based on learning machines that seek to imitate the functioning of the human brain through hierarchical architectures. Unlike classical ML methods, based on the explicit extraction of attributes, DL algorithms may be able to handle data in its original format or raw data, for example, images. Thus, these methods have multiple levels of representation, obtained through the composition of simple and non-linear modules, where each level represents the data in an increasingly abstract way. In this way, these deeper layers can amplify aspects of the input data that are important in the classification process (LeCun et al., 2015, Polat and Özerdem, 2020, Silva et al., 2019). In this way, DL methods implicitly extract attributes. Among deep architectures, convolutional neural networks (CNN) are quite popular, being able to capture an input image, assign learnable weights to various aspects and objects of the image, and thus classify the image (Mao et al., 2020).

EEG recordings are commonly used as one-dimensional signals in machine learning problems. Thus, in this approach, it is necessary to choose a set of representative attributes to represent the signal windows in time through physical and statistical attributes in the time and frequency domains. However, representing these signals in two dimensions can be useful in the context of deep learning. Several researchers have proposed ways to represent and classify EEG signals as (Fu et al., 2014, Ieracitano et al., 2019, Lee and Choi, 2018, Polat and Özerdem, 2020, Srinivasan et al., 2012) images.

The main objective of this work is to propose the ASTERI method of representing EEG signals in images, in order to facilitate the use of deep learning machines based on CNNs to solve EEG motor imagery problems. The ASTERI method uses the classic Backprojection image reconstruction method, originally applied in the reconstruction of computed tomography images (Gordon et al., 1975, Hsieh, 2003). Backprojection is a simple and fast summing method capable of obtaining images from sinograms. In this study, we will use biosignal windows as pseudo-sinograms (Gomes et al., 2020a, 2021a). The method was named ASTERI, the Greek word for "star", because of the pattern that is generated in the representation of the signal in time as an image. The pattern similar to the design of a stylized star is the result of applying the Backprojection algorithm to the signal in time.

To validate the method of representing signals in images proposed in this work, we used as a problem of interest the technique of classification of motor images (MI) based on EEG to control brain-machine interfaces (BCI). MI is based on intentional mental rehearsals of motor behavior without an associated movement or external stimulus. Thus, only with imagination, the kinesthetic memory of a previously performed movement is activated, giving the impression that it is being performed again (Andrade et al., 2020, Batula et al., 2017, Padfield et al., 2019). In this work, we use the 2b database of the BCI IV Competition (2008). This signal base is composed of brain signals from nine volunteers who imagined two types of general movements: right hand and left hand movements.

After generating the image banks referring to the nine subjects' motor image signals, we tested two pre-trained deep architectures: VGG16 and LeNet. These deep neural networks were used to extract features from ASTERI images that represent EEG signal windows. Then, we applied swarm intelligence methods for attribute selection and Random Forests for classification. Finally, the results were evaluated according to the accuracy and kappa index, being compared with the results of the state of the art.

The structure of the subsequent sections is organized as follows: in Section 1.2, we comment on works related to the representation of EEG signals as images and related to motor imagery classification; in Section 2.1, we present the theoretical concepts necessary for a good understanding of this work, and also our proposed method; in Section 3 we show the results and discussion. Finally, our conclusions and future works are described in Section 5.

1.2 Related works

1.2.1 Representation of signals as images for use of Convolutional Neural Networks

Several works have proposed methods of representing signals in images aiming at the use of Convolutional Neural Networks. Kwon et al. (2020) tested an approach to representing speech signals as spectrograms. The general objective of the work was to recognize emotions through these signals. The work methodology consisted of applying a pre-processing to speech signals using an adaptive threshold. Then, the signals were converted into 128×128 images using the Short Term Fourier Transform (STFT) and the Fast Transform of (FFT). Finally, the authors designed a new CNN architecture, called Deep Stride CNN (DSCNN). The network uses steps of 2×2 directly in the convolution layers as downsampling. Furthermore, the network is formed by convolutional layers of 7 and fully connected layers of 2 layers that feed the Softmax classifier. This model was evaluated in two databases: IEMOCAP and RAVDESS. The first base counts with speech signals from 10 actors, who represented seven emotions: anger, fear, happiness, disgust, sadness, excitement, and neutral. The RAVDESS dataset is composed of 24 actors, who enact eight emotions (anger, calm, happiness, sadness, surprise, disgust, fear, and neutral). The proposed method obtained an average accuracy of 81.75% for the IEMOCAP base, and 79.5% for the RAVNESS base.

Similarly, Khare and Bajaj (2020) sought to differentiate emotions with CNNs. However, in this case, the authors used electroencephalographic signals made available in a public database. The database consists of recordings of 20 students using 24 channels. During brain monitoring, volunteers watched videos, intended to arouse one of four emotions: fear, happiness, sadness, and relaxation. First, the work proposed a pre-processing of signals with Butterworth bandpass filters. To represent the signals as images, they tested the Smoothed pseudo-Wigner-Ville (SPWVD) distribution. The work points out that the SPWVD transformation provides a better spatio-temporal resolution, providing a direct localization of the signal energy. The generated images served as input for three pre-trained CNNs (AlexNet, Resnet50 and Vgg16) and a fourth CNN configured with four convolutional layers and 2 fully connected layers. Average classification accuracy was 90.98%, 91.91%, 92.71%, and 93.01%, for tests performed with AlexNet, Resnet50, Vgg16, and CNN configurable, respectively.

Yildirim et al. (2019) proposed a method for automatic detection of diabetes mellitus (DM) using heart beats from electrocardiogram (ECG) signals. Yildirim et al. (2019) also tested the following pre-trained networks: AlexNet, VggNet, ResNet, and DenseNet, in order to overcome the challenge of using a small database. To transform the one-dimensional signals into two-dimensional, the authors generated grayscale spectrograms of size 437×501 . To test the method, they used a database of 30 volunteers: 15 healthy and 15 diagnosed with DM. Finally, the pre-trained DenseNet model achieved the best performance with an average accuracy of 97.62 ± 2.30 .

Taghizadegan et al. (2021) sought to predict obstructive sleep apnea (OSA) events through polysomnographic signals (EEG, ECG, respiratory signals, oximetry, and eye movements). The authors designed a semi-automatic method using CNNs and recurrence plots. The recurrence plot (RP) is a tool to visually represent the dynamic recurrence of the system. The work methodology consisted of representing the signals in RP and thus feeding pre-trained CNNs. In this case, each type of signal fed a different network. Finally, the output of these CNNs was merged using the weighted majority voting (WMV) method based on accuracy. This proposal was tested in two databases: “Dublin Sleep Apnea database” and “MIT-BIH polysomnographic signal”. The first database consists of EEG and ECG signals from 15 volunteers. The second database is composed of EEG, ECG and respiratory signs from 16 men. Considering four scenarios (30, 60, 90, and 120 seconds before the OSA events), the authors observed that the best prediction occurs closer to the event, that is, at 30-second intervals. For the “Dublin Sleep Apnea” database, the ShuffletNet network achieved an accuracy of $90.45\% \pm 1.23$. For the MIT-BIH database, the average accuracies were $90.72\% \pm 1.50$ and $90.23\% \pm 1.35$, for the ShuffleNet and Resnet networks, respectively.

Another way of representing signals as images was addressed by Ieracitano et al. (2019). The researchers sought to differentiate between people with Alzheimer’s disease (AD), people with moderate cognitive impairment (MCI) and healthy individuals. They proposed to represent EEG signals as spectral density images (PSD) of size $19 \times F$. The image dimensions represent the 19 EEG channels, while F is the number of discrete frequency values in the range of 0.5 – 32 Hz. The images then served as inputs to a custom CNN that classifies the images into two or three classes. To test the proposed method, the authors recruited 189 volunteers from the Centro Neurolesi Bonino-Pulejo de Messina, IRCCS, including 63 subjects with AD, 63 subjects with MCI, and 63 volunteers for the control group. The signals were pre-processed and the CNN was trained and tested, achieving an accuracy of 83.33% in the three-class scenario. Considering a binary classification, the best results were the comparison of AD with the control group (92.95%) and of MCI with the control group (91.88%). The authors also tested the classification of PSD images using usual classifiers such as SVM, MLP and LDA, but the results with custom CNN still remained the best.

1.2.2 Classification of EEG-based Motor Imagery for Control of Brain-Machine Interfaces

Zhang et al. (2021) tested ways to adapt CNNs to improve motor image classification. To overcome the small amount of EEG signals per person, the study proposed a subject-independent classification. The idea is to train the network using data from several patients, with the exception of the target subject. Then, the pre-trained model is adapted and adjusted with the data of the chosen subject. For this, the researchers proposed different strategies for adapting the network. First, they tested the classifier’s optimization, leaving the attribute extraction parameters unchanged. Then they tested the adaptation of convolutional layers. In this second scenario, four different proportions were tested, varying the amount of layers retrained with the subject’s data. To evaluate the method, they used a baseline of EEG signals from 54 healthy patients imagining right and left hand movements. The CNN used consisted of spatial and temporal filters with max-pooling, 3 convolutional blocks and a fully connected layer with Softmax. Finally, the study observed that, compared to traditional methods, with training performed with the target subject’s own data, the proposed method presented an accuracy 32.50% higher. Among the scenarios tested, the best results were with adaptation of 3 of the 4 convolutional layers

and the Softmax layer. This indicates that only the first layer was properly trained with data from other patients.

Another innovative approach was tested by Amin et al. (2019). In order to improve motor image classification results, the authors merged several models of CNNs with different depths and filters. Thus, the proposed method is formed by CNN models based on the AlexNet architecture. The attributes extracted by the different models are then concatenated, serving as input to an MLP (MCNN Method) or an autoencoder (CCNN Method). Finally, a Softmax activation function is added. In order to test this proposal, the authors experimented with four types of CNNs, ranging from one to four convolution blocks and max-pooling and filters ranging in size from 10 to 30. The CNN models were pre-trained and evaluated. using the High Gamma Dataset, formed by EEG signals from 20 volunteers. In addition, the models were also tested on the BCI IV-2a (BCID) dataset. The BCID is composed of EEG signals from nine volunteers, who imagined four types of movements: right hand, left hand, foot and tongue. As for training, the study also evaluated two approaches: subject-specific and also using signals from all volunteers for training. Finally, considering the BCID, the tested method presented an average accuracy of 75.7% for the specific training of the discipline. In the case of training with data from all volunteers, the CCNN method outperformed the others, with an average accuracy of 55.34%.

Dai et al. (2020) focused on two approaches to improve motor image classification: customizing the CNN architecture for each subject by adjusting the kernel dimensions, and augmenting the data for each subject. The authors investigated the impact of kernel size between different volunteers and between different imaging sessions for the same person. This analysis revealed 10% differences in accuracy with the variations made. Inspired by these findings, they proposed a CNN with a hybrid convolution scale (HS-CNN). In addition to the architectural adaptations, the work proposed a new method of data augmentation, where signal segments were recombined in the time and frequency domains. To validate the method, the researchers used two BCI Competition IV databases: 2a and 2b. Each base is composed of nine healthy subjects, with 4 and two classes of motor imagery, respectively. The mean classification accuracy was 87.6% for the proposed method, showing an improvement of up to 23.25% and 19.7% for databases 2a and 2b, respectively.

As in the previous work, Zhu et al. (2019) also performed experiments on the BCI Competition IV dataset 2b. However, the authors tested a subject-independent approach. The main idea of the work is to train BCIs with existing data from several individuals and transfer the learning to classify images of a new individual. Thus, the authors pre-processed the signals with bandpass filters and extracted the attributes using the common spatial pattern (CSP) method. In this study, however, the authors abandoned the log normalization operation commonly used in CSP, believing that the operation can remove important temporal information for classification. To this end, they designed a structure capable of reducing the dimensionality of CSP data and with greater discriminative capacity of the information, called separate channel convolutional network (SCCN). To do this, they used a one-dimensional CNN as an encoder to capture information from each channel and reduce redundant information. Then, a recognition block was used, capable of classifying the encoded CSP data. Finally, a softmax layer to complete the classification process. The results obtained were evaluated according to two metrics: accuracy and information transfer rate (ITR). The proposed method presented an average accuracy of 64% and an ITR of 0.83. However, simpler classifiers like LDA had better results (65%).

Zhao et al. (2019) proposed a new method to represent and extract signal attributes from motor images. The authors' strategy was to represent signal windows in three dimensions. For this, the spatial distribution of the electrodes in a 2×2 matrix was maintained, while the positions without electrodes were filled with zero. They then expanded the 2D conformation to 3D using the temporal information. For the extraction of attributes and subsequent classification, a 3D CNN called Multi-branch 3D CNN was designed. Then, three branches formed by different CNNs were used, called Small Receptive Field network (SRF), Medium Receptive Field network (MRF) and Large Receptive Field network (LRF). Each of these branches is made up of three convolutional layers and three dense layers, including the shared layer. The idea with this architecture is that the parameters of the filters chosen for each branch must be different, targeting different receptive fields in the time domain. Finally, these three networks are combined, serving as input to a softmax. This methodology was tested with data from the BCI Competition IV 2a database. To analyze the performance of the proposed CNN, the authors compared the accuracy of each of nine volunteers using CNN 3D Multibranch, SRF only, MRF only, and LRF only. Analyzing the results of networks with a single branch, they realized that it is not possible to choose just one model for all volunteers, as each one stood out in the performance of some subject. These observations point to the need for subject-specific networks. On the other hand, when using the Multi-branch 3D CNN, the results were superior in all cases.

2 Materials and methods

2.1 Proposed method

In this work we present the ASTERI method of representing EEG signal windows in time in the form of an image, to enable the use of convolutional neural networks in applications of intelligent systems based on EEG. To validate the proposal, we applied the proposed form of representation to the solution of the motor imagery problem.

Figure 1 schematically presents a summary of ASTERI. First, EEG signals are organized into classes and windowed into 1 s windows with 0.5 s overlap. Then the signals are filtered to remove artifacts and select frequency bands of interest. In this sense, different filtering approaches are tested in this work.

After pre-processing, the signals are represented in the form of images: each signal window, composed of all the channels used, is interpreted as a pseudo-sinogram. From these pseudo-sinograms, images are reconstructed by the classical method of reconstruction: the rear-projection. Figures 2a and 2b show examples of images obtained with signal windows of subjects 1 and 3.

The third step was to extract and select features. Two ways of extracting attributes were investigated. The first way consists of extracting implicit attributes through pre-trained deep networks with other image datasets. For this study, the architectures chosen were standard ImageNet VGG16 (Hon and Khan, 2017, Tammina, 2019, Theckedath and Sedamkar, 2020) and standard MNIST LeNet (Krishna and Kalluri, 2019, Larabi et al., 2019, Sun et al., 2021, Tan et al., 2018, Zhang et al., 2019). The second way was to extract explicit attributes from each signal window, both in the time domain and in the frequency domain. The chosen attributes are described in detail in Table 1. These 34 features were empirically chosen to combine time and frequency-based analysis using common state-of-the-art

signal features (da Silva Junior et al., 2019, de Freitas et al., 2019, 2022, Espinola et al., 2021a,b, Silva Júnior et al., 2020). We also investigated the fusion of attributes, where the attributes extracted from both forms form a single database. Finally, we applied an attribute selection method based on swarm intelligence, the Evolutionary Search (Rodrigues et al., 2019). Considering the explicit attributes, we could have used a feature selection method to select only the most statistically relevant methods, as used by Faria et al. (2021). However, we preferred to select the most relevant features at the end of the feature extraction process, including also the attributes obtained from the deep neural network.

The fourth step is the investigation of the best classifier configuration. In the proposed deep hybrid architecture, the pre-trained deep neural network is used to extract features from the input images. The feature vector is then passed to the effective classification step, where a Random Forest classifier is present. To define how many trees to use in Random Forest, each configuration is tested with 10-fold cross-validation. Random Forest configurations of 10 to 100 trees were tested, ranging from 10 to 10 trees. The test and training experiments were performed considering 70% of the database instances. The remaining data were used for validation purpose.

Attribute	Mathematical expression	Description	Equation
Mean (μ)	$\mu = \frac{1}{N} \sum_{i=1}^N x_i$	Zero crossings	$ZC = \sum_{i=1}^{N-1} \text{sign}(x_i - x_{i+1}) \cdot x_i - x_{i+1} $ $\text{sign}(x) = \begin{cases} 1, & \text{if } x \geq \text{threshold} \\ 0, & \text{otherwise} \end{cases}$
Variance	$\text{var} = \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu)^2$	Slope Sign Change	$\text{SSC} = \sum_{i=1}^{N-1} \text{if}(x_i - x_{i+1} < (x_i - x_{i+1}))$ $f(x) = \begin{cases} 1, & \text{if } x \geq \text{threshold} \\ 0, & \text{otherwise} \end{cases}$
Standard deviation (σ)	$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \mu)^2}$	Height parameter stability	$H(\text{var})_{\text{stability}} = \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu)^2$
Root mean square	$\text{RMS} = \sqrt{\frac{\sum_{i=1}^N x_i^2}{N}}$	Height parameter mobility	$H(\text{var})_{\text{mobility}} = \frac{\text{var} \left(\frac{d x(t)}{dt} \right)}{\text{var}(x(t))}$
Average Amplitude Change	$\text{AAC} = \frac{1}{N} \sum_{i=1}^N \left \frac{d x(t)}{dt} \right $	Height parameter complexity	$H(\text{var})_{\text{complexity}} = \frac{H(\text{var})_{\text{mobility}} \left(\frac{d x(t)}{dt} \right)}{H(\text{var})_{\text{stability}} \left(x(t) \right)}$
Difference Absolute Deviation	$\text{DASD} = \sqrt{\frac{1}{N} \sum_{i=1}^N \left(\frac{d x(t)}{dt} \right)^2}$	Mean frequency	$MNF = \frac{\sum_{i=1}^M f_i P_i}{\sum_{i=1}^M P_i}$ Where f_i, P_i are the frequency and power of the spectrum, respectively, and M is the length of the frequencies
Integrated Absolute Value	$\text{IAV} = \sum_{i=1}^N x_i$	Median frequency	$MDF = \frac{1}{2} \sum_{i=1}^M P_i$
Logarithm Detector	$\text{LSD} = e^{2 \sum_{i=1}^N \ln(x_i)}$	Mean power	$MNP = \sum_{i=1}^M \frac{P_i}{M}$
Simple Square Integral	$\text{SSI} = \sum_{i=1}^N x_i^2$	Peak frequency	$\text{PKF} = \max(P_i)$
Mean Absolute Value	$\text{MAV} = \frac{1}{N} \sum_{i=1}^N x_i $	Power Spectrum ratio	$\text{PSR} = \frac{\text{PKF}}{\sum_{i=1}^M P_i}$
Mean Logarithm Kernel	$\text{MLOK} = \frac{1}{N} \sum_{i=1}^N \ln(x_i)$	Total Power	$\text{TP} = \sum_{i=1}^M P_i$
Skewness (s)	$s = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^3}{\sigma^3}$	First Spectral Moment	$\text{SM1} = \sum_{i=1}^M f_i P_i$
Kurtosis	$\text{kurt} = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^4}{\sigma^4}$	Second Spectral Moment	$\text{SM2} = \sum_{i=1}^M f_i^2 P_i$
Maximum Amplitude	$\text{MAX} = \max(x_i)$	Third Spectral Moment	$\text{SM3} = \sum_{i=1}^M f_i^3 P_i$
Third Moment	$\text{M3} = \frac{1}{N} \sum_{i=1}^N (x_i)^3$	Variance of Central Frequency	$\text{VCF} = \frac{\text{SM2}}{\text{TP}} - \left(\frac{\text{SM1}}{\text{TP}} \right)^2$
Fourth Moment	$\text{M4} = \frac{1}{N} \sum_{i=1}^N (x_i)^4$	Waveform length	$\text{WL} = \sum_{i=1}^{N-1} x_{i+1} - x_i $
Fifth Moment	$\text{M5} = \frac{1}{N} \sum_{i=1}^N (x_i)^5$	Shannon Entropy	$J = - \sum_{i=1}^M P_i \log(P_i)$

Table 1: List of 34 extracted attributes. These features were empirically chosen to combine time- and frequency-based commonly used features. Source: The authors.

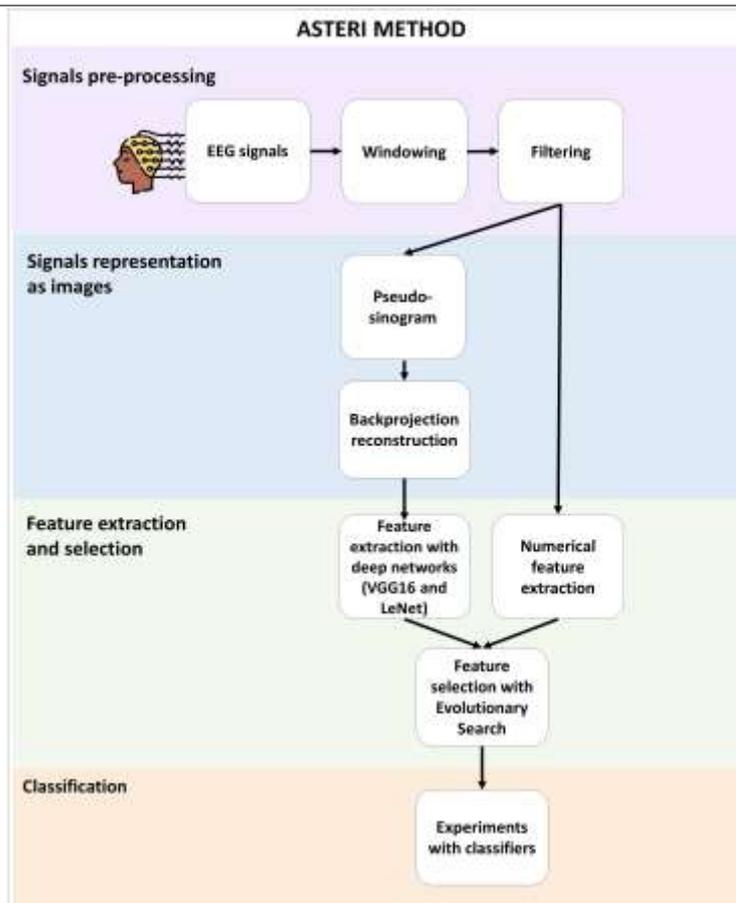


Fig. 1: Proposed method: ASTERI. The method consists of representing EEG signals as images, aiming at the use of deep architectures for attribute extraction. First, the signals are pre-processed. Then, they are represented in two dimensions by applying the backprojection reconstruction method. The third step consists of extracting and selecting attributes by pretrained deep architectures. The last step is classification experiments with test/train subset and validation. Source: The authors.

Experiments were also performed with augmented datasets with synthetic instances. The objective of this step was to analyze scenarios in which more data were collected per patient and analyze the impact of this increase in motor image sections on classification performance. To increase the databases, we applied the SMOTE method (Synthetic

Minority Oversampling Technique) in train/test subset. The validation subset remained only with instances of the original dataset, that is collected EEG signals.

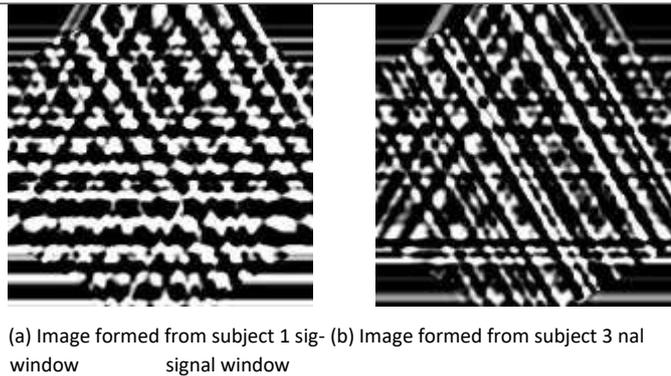


Fig. 2: Images generated with the ASTERI representation method and motor imagery signals

2.2 Database

The proposed method was tested in the BCI Competition IV 2b (Leeb et al., 2008) database. The database is composed of electroencephalographic signals from 9 volunteers. Each of them was evaluated during 5 sessions. At the beginning of each session, some steps were followed: 1) Two minutes with eyes open, looking at a fixed point; 2) One minute with eyes closed; 3) One minute with moving eyes. The purpose of these steps is to estimate the influence of electrooculography (EOG) signals on EEG signals. The electrooculogram was recorded using 3 channels.

After these 4 introductory minutes, the runs of motor images began, monitoring with three EEG channels (positions C3, Cz and C4). Two classes were tested: imagination of the right hand movement and imagination of the left hand movement. In each of the sessions, 6 runs were performed, consisting of 20 trials each. Thus, the base provides 120 image essays for each of the classes for each session. Figure 3 presents a diagram detailing how the acquisition of EEG signals from this base was performed.



Fig. 3: EEG signal acquisition protocol during motor imagery sessions, used in BCI Competition IV database 2b (2008). Source: The author.

The first two imaging sessions were without visual feedback. They consisted of the following steps: first, a fixation cross appeared on the volunteer’s screen; then, an acoustic stimulus was given as an alert; finally, a visual stimulus indicated the class by means of an arrow. Finally, the patient must imagine the indicated movement for a period of 4 seconds. Figure 4 details a run without visual feedback.

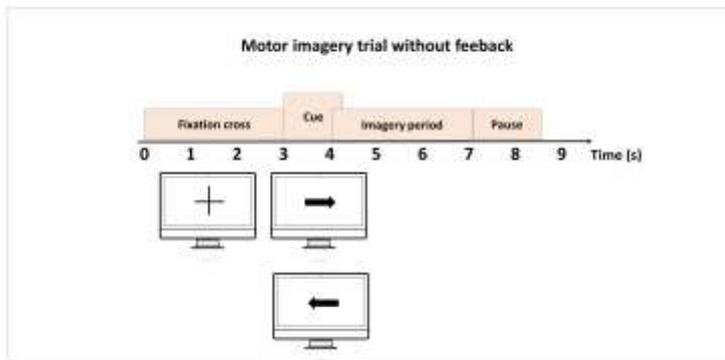


Fig. 4: Example of an EEG acquisition run without application of visual feedback. Source: The author.

In contrast, the final three sessions used visual feedback. At the start of each round, a neutral gray smiley appeared on the screen. Thus, the clue with the class was given in the period of 3.5 to 7.5 seconds. Depending on the cue, the volunteer must move the face to the right or to the left, according to the cue given. During the movement imagination, the smiley turned green or red, indicating whether the imagined movement was in the right or wrong direction, respectively. Figure 5 detail how feedback runs were made.

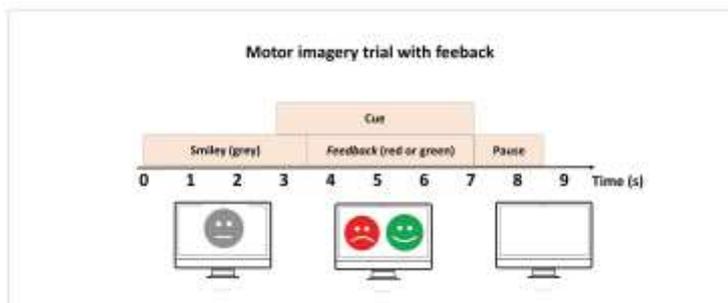


Fig. 5: Example of an EEG acquisition run with application of visual feedback. Source: The author.

2.3 Backprojection

The backprojection reconstruction algorithm was originally proposed in the context of Computed Tomography (CT). In this case, there is a source and a detector of an X-radiation surrounding the body to be irradiated. The reconstruction of this body can be considered

a 2D distribution of some function. This means that we can have a function that represents the linear attenuation coefficients of the objects internal to the body. We then sought to measure the sum or line integrals of the attenuation coefficients along different collected paths or at different angulations. Each result obtained in one direction is called a projection (Gordon et al., 1975, Hsieh, 2003).

The backprojection method is a simple and fast summation method. The projections or Radon Transforms of a sample are line integrals at specific angles. Mathematically, the reconstruction technique can be defined as follows. Assuming a function f of two polar variables r and ϕ , with its line integral defined by $[Rf](l, \theta)$, being at a distance l from the origin and with a angle θ with the vertical axis. Also considering the R operator, corresponding to the Radon transform, we can define for any function p through the Equation 1 (Gordon et al., 1975, Koshev et al., 2016):

$$[R^{-1}p](r, \phi) = \frac{1}{2\pi^2} \int_{-\infty}^{\pi} \int_{-\infty}^{+\infty} \frac{1}{r \cos \theta - \phi - l} p_l(l, \theta) dl d\theta, \quad (1)$$

where $p_l(l, \theta)$ is the partial derivative of p with respect to l . For all points (r, ϕ) , they satisfy the Equation 2 below:

$$[R^{-1}Rf](r, \phi) = f(r, \phi). \quad (2)$$

Therefore, the backprojection method is one of the numerical methods for estimating double integrals in a two-step process. Firstly, you can define:

$$p_c(l, \theta) = \int_{-\infty}^{+\infty} p(l', \theta) q(l - l', \theta) dl'. \quad (3)$$

Secondly, the estimate of f^* from f is obtained by the backprojection given by Equation 4:

$$f^*(r, \phi) = \int_0^{\pi} p_c(r \cos(\theta - \phi) - \theta) d\theta, \quad (4)$$

where this second phase is more computationally costly (Gordon et al., 1975).

There are several ways to represent these projections. Sinograms are the best known and most used form. The name sinogram has its origins in the fact that the single-point Radon Transform is a sinusoid. A sinogram is a collection of flat or single-slice projections arranged in a matrix. It is possible to use the backprojection algorithm to obtain tomographic images from sinograms (Gomes et al., 2020a, Hsieh, 2003).

2.4 Transfer Learning

The training of convolutional neural networks is computationally very expensive, due both to the high demand for processing consumption to adjust the weights of the neurons, and to the use of memory, due to its large number of layers and neurons. The greater the number of layers, the greater the number of parameters to be optimized. Furthermore, in many applications, it is impractical to collect new data and reconstruct models. Therefore, transfer of learning can be an interesting alternative. It is inspired by human capacities to transfer knowledge between multiple activities. In other words, the main idea is that the learning acquired in training a database can be used to solve other problems. Thus, it is possible to use deep learning, without the need to perform numerous (Bouti et al., 2020, Pan and Yang, 2009) calculations.

2.5 Feature extraction: VGG16 and LeNet

Visual geometry group network (VGG16) are 16-layer architectures designed for image classification. They have 13 convolutional layers, where the convolution operation performed with 3×3 kernels, can be described by Equation 5:

$$P_{output} = f(W \cdot P_{input} + b), \quad (5)$$

where P_{input} and P_{output} are the pixels of the input and output images, respectively. The parameters W and b are learned (Rangarajan and Purushothaman, 2020).

Like other deep architectures, the VGG16's convolutional layers act as feature extractors. Initial layers extract simpler attributes, which are combined to form more complex attributes in deeper layers. Therefore, the network has multiple levels of representation, where each level represents data in an increasingly abstract way. In this way, these deeper layers can amplify aspects of the input data that are important in the classification process (LeCun et al., 2015, Polat and Özerdem, 2020).

After each of the convolutional layers, the VGG16 has a non-linear activation layer (Rectified Linear Unit, ReLU) and a maxpooling layer. The network ends up with two fully connected layers (FC). The first FC layer is formed by 4096 neurons. The second layer has the number of neurons corresponding to the number of classes of the problem. Finally, the last layer of the architecture is a softmax, which gives the probability for each class (Özyurt, 2020, Rangarajan and Purushothaman, 2020).

Another deep architecture is LeNet. LeNet is a simpler network than the VGG16, originally developed for handwritten digit recognition. It consists of five layers, two of which are convolutional layers and three are fully connected (Priyadharshini et al., 2019) layers. Likewise, each of the convolutional layers is followed by a clustering layer and an activation function (Kuo, 2016).

2.5.1 Classification: Decision Trees and Random Forests

Decision trees are a type of supervised learning machine that can be used in both classification and regression problems. Decision trees analyze data through a series of attribute-related questions. Each question is contained in a node, which branches into child nodes containing possible answers. The questions then form a hierarchy, modeled as a tree. The starting point of the tree is the root node that has the highest hierarchical level. The endpoint is the leaf, a node with no children. In this way, the algorithm makes decisions following the tree structure. There are several types of decision trees, which are usually distinguished by the way the method traverses the tree (de Freitas Barbosa et al., 2021, Gomes et al., 2020b, Kingsford and Salzberg, 2008). The Random Tree and Random Forest methods are among the most important tree-based models.

According to Geurts et al. (2006), the Random Tree algorithm uses a single tree built by a stochastic process. This method considers only a few randomly selected features at each node of the tree. Random Forests consist of a collection of predictive trees, where each tree depends on a group of random variables (Cutler et al., 2012). In this way, each tree votes for a problem class. Finally, the most voted class is chosen as the classifier's final prediction.

2.6 Data augmentation: SMOTE

SMOTE is a popular database augmentation method. The idea of the algorithm is to perform an interpolation between close instances, both belonging to the same class. Therefore, by introducing new examples, SMOTE aids in the generalizability of the classifier Fernández et al. (2018). Due to its popularity, several works have applied the method, combined with other techniques and proposed modifications (Demidova and Klyueva, 2017, Douzas et al., 2018, Jiang et al., 2016).

The idea behind SMOTE is as follows: the algorithm generates synthetic instances based on data similarities, rather than replicating samples. First, SMOTE selects a sample of the class you want to increase. We call this sample s_1 . It then selects k nearest neighbors from the sample s_1 using the KNN algorithm. In this case, the Euclidean distance is considered. Then, one of the neighboring k samples is randomly selected. We will call this second sample s_2 . Finally, a synthetic sample s_{new} is calculated from the samples s_1 and s_2 . It is generated according to Equation 6. This means that the synthetic sample created is spatially positioned between the samples that generated it (s_1 and s_2). Finally, the sample s_{new} is added to the Chawla et al. (2002) database.

$$s_{new} = (s_1 + s_2) \times n, \quad n \in [0,1]. \quad (6)$$

2.7 Experiments setup

In this work, the codes for pre-processing the motor imagery signals were implemented using the scientific computing environment GNU/Octave¹. The extraction/selection of attributes and classification steps were performed in the open-source Java library Weka².

In the feature extraction step, VGG16 or LeNet deep neural networks were used, pretrained with the IMAGENET and MNIST data sets, respectively. Feature extraction using VGG16 resulted in 1096 attributes, while extraction with LeNet resulted in 500 attributes.

In the best classifier investigation step, we tested different Random Forests. Each classifier configuration was tested under 30 runs using 10-fold cross-validation on the test/train subset (70% of the instances). We also evaluated scenarios with more instances in the test/train subset. For this, we applied the SMOTE algorithm with 2 nearest neighbors, for increases of 50% and 150%. Finally, the best classifier for each train/test subset was tested in the validation subset.

2.8 Metrics

We chose two metrics to evaluate the performance of experiments: accuracy and kappa statistic. Accuracy is the probability that the experiment will provide correct results, i.e. to correct classify the motor imagery as right or left movement. In other words, it is the probability of the true positives and true negatives among all the results. It can be calculated according to the Equation 7 below (Gomes et al., 2021b, Santana et al., 2018):

¹ Open source Software available at: <www.gnu.org/software/octave/index>

² Software available at: <www.cs.waikato.ac.nz/ml/weka/>

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (7)$$

where TP, TN, FP, and FN are the true positives, true negatives, false positives, and false negatives, in this order.

In addition, the Kappa index is a very good measure that can handle very well multiclass problems, as the one proposed here. It is a statistical metric to assess the agreement between the obtained and expected results defined as follows (Gomes et al., 2021b, Santana et al., 2018):

$$\kappa = \frac{\rho_o - \rho_e}{1 - \rho_e}, \quad (8)$$

where ρ_o is observed agreement, or accuracy, and ρ_e is the expected agreement, defined as following (Gomes et al., 2021b, Santana et al., 2018):

$$\rho_e = \frac{(TP + FP)(TP + FN) + (FN + TN)(FP + TN)}{(TP + FP + FN + TN)^2}. \quad (9)$$

The kappa index can range from -1 to 1. Based on its value, the performance of the classifier can be evaluated between very poor and excellent, as shown in Table 2 (McHugh, 2012).

Kappa index	Performance
$\kappa < 0$	Very poor
$0 < \kappa \leq 0.2$	Poor
$0.2 < \kappa \leq 0.4$	Regular
$0.4 < \kappa \leq 0.6$	Good
$0.6 < \kappa \leq 0.8$	Very good
$0.8 < \kappa \leq 1.0$	Excellent

Table 2: Agreement range of the kappa index, as defined by McHugh (2012)

3 Results

Initially, different signal processing and attribute extraction approaches were tested on EEG signals for classification of motor imagery using the ASTERI method. As can be seen in Table 3, regarding the pre-processing of signals, the different approaches included (or not) bandpass filters, wavelet decomposition and statistical thresholds.

In this context, we first used the 5th order Butterworth bandpass filter in setups 1, 3, 6, 7, 8, 9, and 10, aiming to select the alpha and beta frequency bands (8-32 Hz). Many works with 2b dataset include this approach (An et al., 2020, Ang et al., 2012, Hossain et al., 2016, Lee and Choi, 2019), since typically the motor movements are related to these frequency bands (Lee and Choi, 2019).

In addition to filtering, some of these signals have been decomposed into Haar wavelets. Wavelet-based approaches are interesting in the identification of short duration artifacts in signals such as EEG, and may act in the removal of electrooculography artifacts, for instance (Andrade et al., 2020, Bajaj et al., 2020, de Freitas Barbosa et al., 2020).

Thereby, the signals were decomposed into lowpass signals and highpass signals using wavelet function. In the following, lowpass signals were recursively decomposed with the same function until we reach the desired number of levels. In this work, we considered decompositions into 2 or 10 levels. Setups 1, 2, 3, 8, and 9 included this type of analysis.

Among the decomposed signals, a statistical threshold was also applied to the wavelet components of setups 1, 2, 3, and 9. The idea in this case is to disregard the wavelet components that are above or below the established thresholds and then reconstruct the signals back (Bajaj et al., 2020). In these case, two threshold techniques were studied. The first, based on standard deviation, set as threshold the value of $\pm 1.5 \times$ standard deviation, as detailed in the Equation 10 below.

$$\text{Threshold} = 1.5 \times \text{STD}(w), \quad (10)$$

where STD means standard deviation, and w the wavelet coefficients. The second technique consisted of a smoother threshold based on the ATAR (Automatic and Tunable Artefact Removal) algorithm proposed by Bajaj et al. (2020). This algorithm provides different filtering modes and a tunable parameter, which allows for smoother filtering. According to Bajaj et al. (2020), the soft threshold is defined as Equation 11 below:

$$\text{Threshold} = \begin{cases} w, & |w| < \theta_\gamma \\ \frac{1 - e^{\alpha w}}{1 + e^{\alpha w}}, & \text{otherwise} \end{cases} \quad (11)$$

where $\alpha = -\frac{1}{\theta_\gamma} \log \frac{\theta_\alpha - \theta_\gamma}{\theta_\alpha + \theta_\gamma}$ and $\theta_\gamma = 0.8\theta_\alpha$. In addition, θ_α was defined as Equation 12:

$$\theta_\alpha = \begin{cases} f_\beta(r), & f_\beta \geq k_1 \\ k_1, & \text{otherwise} \end{cases} \quad (12)$$

where $f_\beta(r) = k_2 \exp\left(-\beta \frac{100r}{2k_2}\right)$, r is the Interquartile Range of w (Bajaj et al., 2020), k_1 is the 5% quantile of w , k_2 is the 95% quantile of w and $\beta = 0.01$.

As for feature extraction, two pre-trained networks were tested: VGG16 and LeNet. Exceptionally, in setup 5 we have not used deep networks for feature extraction. In this case, explicit time- and frequency-domain attributes were extracted. In setup 10, a fusion of attributes (implicit extracted with VGG16 + explicit) was applied.

3.1 Original dataset

To carry out the experiments with the 10 setups proposed in Table 3, one of the subjects belonging to the database was selected, in order to avoid presenting a huge amount of experimental results. In this case, we selected the subject 8 due to his good results in several studies (Dai et al., 2020, Li et al., 2017). Therefore, it is likely that this subject is the participant who best followed the standards indicated in the motor imagery sessions and, consequently, this subject is a good representative of the full set of subjects. The results for this participant are presented in the tables below. In these initial experiments with subject 8, we considered all signals for training and testing purpose.

Table 4 presents the accuracy values, while Table 5 presents the kappa index values. Observing the mean values of both metrics, it is possible to notice that setting 10

outperformed the other settings, with an average accuracy of 74.53 ± 2.80 and an average kappa index of 0.49 ± 0.06 . If we still consider the standard deviation of these values, we can say that setups 6 and 7 also showed similar results. In contrast, the worst results were obtained by setups 8 and 9, with average accuracies of 63.56 ± 3.94 and 63.49 ± 3.20 , respectively. It is also possible to observe an improvement in performance with the increase in the number of trees used. Random Forests with 100 trees reached the best classification performance.

Setup	Pre-processing			Feature extraction
	Pass-band filter	Wavelet decomposition	Statistical threshold	
Setup 1	Yes	2 levels	Standard Deviation	VGG16
Setup 2	No	10 levels	Soft threshold	VGG16
Setup 3	Yes	10 levels	Standard Deviation	VGG16
Setup 4	No	No	No	VGG16
Setup 5	No	No	No	Explicit feature*
Setup 6	Yes	No	No	VGG16
Setup 7	Yes	No	No	VGG16
Setup 8	Yes	2 levels	No	LeNet
Setup 9	Yes	2 levels	Standard Deviation	LeNet
Setup 10	Yes	No	No	VGG16 + Explicit feature

*The explicit feature extraction do not use the ASTERI method. The extraction is done directly in the signal windows.

Table 3: Different setups used for processing and extracting attributes of EEG signals.

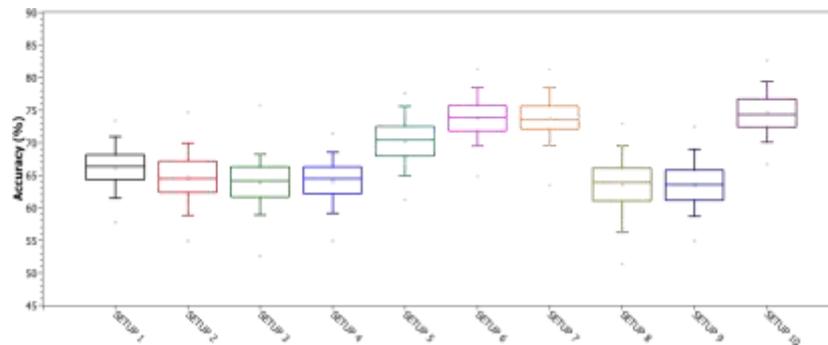
Setup	Accuracy - Subject 8									
	RF 10	RF 20	RF 30	RF 40	RF 50	RF 60	RF 70	RF 80	RF 90	RF 100
Setup 1	63.10 ± 3.14	64.66 ± 3.23	65.33 ± 2.86	65.59 ± 2.89	65.78 ± 2.81	65.94 ± 2.81	66.09 ± 2.80	66.11 ± 2.84	66.05 ± 2.89	66.22 ± 2.81
Setup 2	61.16 ± 3.29	62.92 ± 3.39	61.52 ± 3.16	63.72 ± 3.35	64.11 ± 3.40	64.15 ± 3.36	64.20 ± 3.38	64.39 ± 3.42	64.46 ± 3.40	64.59 ± 3.43
Setup 3	61.81 ± 3.11	63.06 ± 3.05	63.45 ± 3.12	63.52 ± 3.18	63.71 ± 3.17	63.67 ± 3.14	63.73 ± 3.12	63.80 ± 3.14	63.76 ± 3.19	63.91 ± 3.18
Setup 4	62.05 ± 3.14	63.31 ± 3.02	63.46 ± 2.84	63.89 ± 2.96	63.89 ± 2.86	63.95 ± 2.86	63.86 ± 2.81	64.03 ± 2.93	64.05 ± 3.01	64.06 ± 2.86
Setup 5	66.88 ± 2.93	68.54 ± 3.02	69.19 ± 2.97	69.57 ± 2.99	69.90 ± 3.08	69.90 ± 3.03	70.07 ± 3.02	70.16 ± 3.13	70.26 ± 3.16	70.22 ± 3.32
Setup 6	72.41 ± 2.91	72.86 ± 3.10	73.15 ± 2.83	73.34 ± 2.90	73.39 ± 2.86	73.47 ± 2.84	73.47 ± 2.86	73.64 ± 2.75	73.70 ± 2.73	73.81 ± 2.75
Setup 7	72.40 ± 2.85	73.06 ± 2.64	73.15 ± 2.61	73.30 ± 2.61	73.45 ± 2.69	73.60 ± 2.83	73.66 ± 2.79	73.71 ± 2.76	73.79 ± 2.70	73.79 ± 2.73
Setup 8	58.49 ± 3.49	60.40 ± 3.75	61.47 ± 3.82	62.08 ± 3.98	62.52 ± 4.03	62.66 ± 4.00	62.92 ± 3.87	63.14 ± 3.88	63.30 ± 3.87	63.56 ± 3.94
Setup 9	59.50 ± 2.98	61.04 ± 3.12	61.89 ± 3.17	62.34 ± 3.36	62.79 ± 3.20	63.05 ± 3.20	63.23 ± 3.23	63.26 ± 3.13	63.31 ± 3.11	63.49 ± 3.20
Setup 10	73.01 ± 2.89	73.59 ± 3.01	74.03 ± 2.96	74.33 ± 2.83	74.35 ± 2.77	74.50 ± 2.79	74.41 ± 2.77	74.47 ± 2.76	74.56 ± 2.67	74.53 ± 2.80

Table 4: Accuracy values in the classification of motor imagery with EEG signals for subject 8. Different configurations of pre-processing of signals and decision trees were used.

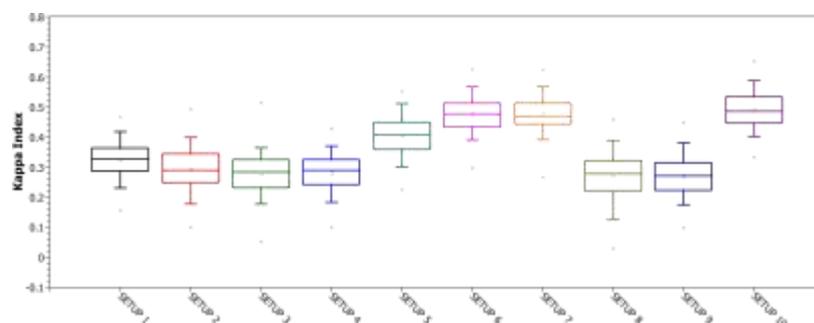
Setup	Kappa index - Subject 8									
	RF 10	RF 20	RF 30	RF 40	RF 50	RF 60	RF 70	RF 80	RF 90	RF 100
Setup 1	0.26 ± 0.06	0.29 ± 0.06	0.31 ± 0.06	0.31 ± 0.06	0.32 ± 0.06	0.32 ± 0.06	0.32 ± 0.06	0.32 ± 0.06	0.32 ± 0.06	0.32 ± 0.06
Setup 2	0.22 ± 0.07	0.26 ± 0.07	0.27 ± 0.07	0.27 ± 0.07	0.28 ± 0.07	0.28 ± 0.07	0.28 ± 0.07	0.29 ± 0.07	0.29 ± 0.07	0.29 ± 0.07
Setup 3	0.24 ± 0.06	0.26 ± 0.06	0.27 ± 0.06	0.27 ± 0.06	0.27 ± 0.06	0.27 ± 0.06	0.27 ± 0.06	0.28 ± 0.06	0.28 ± 0.06	0.28 ± 0.06
Setup 4	0.24 ± 0.06	0.27 ± 0.06	0.27 ± 0.06	0.28 ± 0.06	0.28 ± 0.06	0.28 ± 0.06	0.28 ± 0.06	0.28 ± 0.06	0.28 ± 0.06	0.28 ± 0.06
Setup 5	0.34 ± 0.06	0.37 ± 0.06	0.38 ± 0.06	0.39 ± 0.06	0.40 ± 0.06	0.40 ± 0.06	0.40 ± 0.06	0.40 ± 0.06	0.41 ± 0.06	0.40 ± 0.06
Setup 6	0.45 ± 0.06	0.46 ± 0.06	0.46 ± 0.06	0.47 ± 0.06	0.47 ± 0.06	0.47 ± 0.06	0.47 ± 0.06	0.47 ± 0.05	0.47 ± 0.05	0.48 ± 0.05
Setup 7	0.45 ± 0.06	0.46 ± 0.05	0.46 ± 0.05	0.47 ± 0.05	0.47 ± 0.05	0.47 ± 0.05	0.47 ± 0.06	0.47 ± 0.06	0.48 ± 0.05	0.48 ± 0.05
Setup 8	0.17 ± 0.07	0.21 ± 0.08	0.23 ± 0.08	0.24 ± 0.08	0.25 ± 0.08	0.25 ± 0.08	0.26 ± 0.08	0.26 ± 0.08	0.27 ± 0.08	0.27 ± 0.08
Setup 9	0.19 ± 0.06	0.22 ± 0.06	0.24 ± 0.06	0.25 ± 0.07	0.26 ± 0.06	0.26 ± 0.06	0.26 ± 0.06	0.27 ± 0.06	0.27 ± 0.06	0.27 ± 0.06
Setup 10	0.46 ± 0.06	0.47 ± 0.06	0.48 ± 0.06	0.49 ± 0.06	0.49 ± 0.06	0.49 ± 0.06	0.49 ± 0.06	0.49 ± 0.06	0.49 ± 0.05	0.49 ± 0.06

Table 5: Kappa index in the classification of motor imagery with EEG signals for subject 8. Different configurations of pre-processing of signals and decision trees were used.

The boxplots in Figure 6 show a comparison of the ten configurations. In this case, only the results of the best classifier of each setup were plotted. For example, for setup 1, the boxplot refers to the Random Forest classifier with 100 trees, as it achieved the highest average accuracy. Again, boxplots show better performance for settings 6, 7 and 10.



(a)



(b)

Fig. 6: Subject 8's motor imagery classification performance using EEG signals.

Considering that setup 10 proved to be superior for subject 8, we decided to carry out experiments with the other subjects belonging to the database. This means that the signals from all subjects were subjected to Butterworth bandpass filtering and their attributes were extracted in two ways: explicitly (i.e. time- and frequency-domain features) and implicitly (pre-trained deep neural networks) with the pre-trained VGG16 deep architecture, as pointed out earlier in Table 3.

Tables 6 and 7 show the values of accuracy and kappa index found for each subject, using again the Random Forest classifiers. These training and testing experiments were performed with 70% of the original database instances, while remaining signals were separated for validation. It is possible to observe that, for the nine subjects, the results were better when using 90 or 100 trees in the classification. In addition, both tables show that subjects 4 and 8 presented better classification performance than the other subjects, with average accuracy of 78.20 ± 3.14 and 73.40 ± 3.21 , respectively. On the other hand, subjects 2 (50.54 ± 3.97) and 3 (50.49 ± 3.96) had lower results. Similarly, the kappa index followed this trend.

Finally, Figure 7 presents validation confusion matrices for all nine subjects, considering the original dataset with only collected signals. When observing subject 4, for example, it is possible to notice that among 260 instances referring to imagery of the right hand, 237 were correctly classified, while 23 were classified as imagery of the left hand. Considering the left hand class for this subject, 91.4% of the images were correctly classified. The highest rate of confusion was observed for subject 3. In this case, 16.8% of the left class was classified as right hand class.

Subject	Accuracy									
	RF 10	RF 20	RF 30	RF 40	RF 50	RF 60	RF 70	RF 80	RF 90	RF 100
Subject 1	58.17 ± 3.97	59.89 ± 3.93	60.46 ± 4.08	60.87 ± 4.20	60.96 ± 4.12	61.30 ± 4.02	61.38 ± 4.16	61.29 ± 4.14	61.41 ± 4.08	61.44 ± 3.97
Subject 2	50.29 ± 4.07	50.50 ± 4.08	50.62 ± 4.01	50.50 ± 4.24	50.34 ± 4.21	50.63 ± 4.39	50.63 ± 4.18	50.53 ± 4.33	50.43 ± 4.33	50.54 ± 4.32
Subject 3	49.93 ± 3.95	50.40 ± 3.95	50.50 ± 4.19	50.46 ± 4.11	50.63 ± 4.16	50.54 ± 3.95	50.43 ± 3.87	50.60 ± 3.98	50.44 ± 3.87	50.49 ± 3.96
Subject 4	75.67 ± 3.05	77.20 ± 3.33	77.65 ± 3.23	77.93 ± 3.27	77.89 ± 3.14	78.02 ± 3.16	78.10 ± 3.24	78.16 ± 3.21	78.11 ± 3.17	78.20 ± 3.14
Subject 5	64.18 ± 3.83	66.34 ± 3.82	67.07 ± 3.83	67.60 ± 3.77	67.74 ± 3.85	68.00 ± 3.90	68.26 ± 3.87	68.30 ± 3.92	68.41 ± 3.88	68.31 ± 3.97
Subject 6	56.95 ± 3.88	58.26 ± 4.01	59.70 ± 3.94	59.87 ± 4.12	60.21 ± 4.05	60.34 ± 3.91	60.43 ± 3.98	60.68 ± 3.89	60.68 ± 3.85	60.84 ± 4.04
Subject 7	60.85 ± 3.87	62.06 ± 3.63	62.56 ± 3.76	62.83 ± 3.71	62.95 ± 3.83	62.92 ± 3.90	63.02 ± 3.92	63.01 ± 3.98	62.99 ± 3.86	63.17 ± 3.71
Subject 8	72.01 ± 3.42	72.49 ± 3.08	72.68 ± 3.03	72.86 ± 3.08	72.94 ± 3.22	73.22 ± 3.32	73.26 ± 3.20	73.33 ± 3.34	73.66 ± 3.22	73.80 ± 3.21
Subject 9	65.97 ± 3.96	67.56 ± 3.71	68.34 ± 3.75	68.53 ± 3.69	68.78 ± 3.66	68.97 ± 3.58	69.05 ± 3.54	69.17 ± 3.48	69.21 ± 3.53	69.25 ± 3.59

Table 6: Classification accuracy for the nine subjects. Configuration 10 of pre-processing and extraction of attributes was used. Different decision trees were tested, being trained with subject-specific data.

Subject	Kappa Index									
	RF 10	RF 20	RF 30	RF 40	RF 50	RF 60	RF 70	RF 80	RF 90	RF 100
Subject 1	0.16 ± 0.08	0.20 ± 0.08	0.21 ± 0.08	0.22 ± 0.08	0.22 ± 0.08	0.23 ± 0.08	0.23 ± 0.08	0.23 ± 0.08	0.23 ± 0.08	0.23 ± 0.08
Subject 2	0.01 ± 0.08	0.01 ± 0.08	0.01 ± 0.08	0.01 ± 0.08	0.01 ± 0.08	0.01 ± 0.09	0.01 ± 0.08	0.01 ± 0.09	0.01 ± 0.09	0.01 ± 0.09
Subject 3	0.00 ± 0.08	0.01 ± 0.08	0.01 ± 0.08	0.01 ± 0.08	0.01 ± 0.08	0.01 ± 0.08	0.01 ± 0.08	0.01 ± 0.08	0.01 ± 0.08	0.01 ± 0.08
Subject 4	0.51 ± 0.06	0.54 ± 0.07	0.55 ± 0.06	0.56 ± 0.07	0.56 ± 0.06	0.56 ± 0.06	0.56 ± 0.06	0.56 ± 0.06	0.56 ± 0.06	0.56 ± 0.06
Subject 5	0.28 ± 0.08	0.33 ± 0.08	0.34 ± 0.08	0.35 ± 0.08	0.35 ± 0.08	0.36 ± 0.08	0.37 ± 0.08	0.37 ± 0.08	0.37 ± 0.08	0.37 ± 0.08
Subject 6	0.14 ± 0.08	0.18 ± 0.08	0.19 ± 0.08	0.20 ± 0.08	0.20 ± 0.08	0.21 ± 0.08	0.21 ± 0.08	0.21 ± 0.08	0.21 ± 0.08	0.22 ± 0.08
Subject 7	0.22 ± 0.08	0.24 ± 0.07	0.25 ± 0.08	0.26 ± 0.07	0.26 ± 0.08	0.26 ± 0.08	0.26 ± 0.08	0.26 ± 0.08	0.26 ± 0.08	0.26 ± 0.07
Subject 8	0.44 ± 0.07	0.45 ± 0.06	0.45 ± 0.06	0.46 ± 0.06	0.46 ± 0.06	0.46 ± 0.07	0.47 ± 0.06	0.47 ± 0.07	0.47 ± 0.06	0.47 ± 0.06
Subject 9	0.32 ± 0.08	0.35 ± 0.07	0.37 ± 0.08	0.37 ± 0.07	0.38 ± 0.07	0.38 ± 0.07	0.38 ± 0.07	0.38 ± 0.07	0.38 ± 0.07	0.38 ± 0.07

Table 7: Classification Kappa index for the nine subjects. Configuration 10 of pre-processing and extraction of attributes was used. Different decision trees were tested, being trained with subject-specific data.

3.2 Augmented dataset with SMOTE

In the following, the same pre-processing methodology previously applied in the original database (i.e. using only collected signals) was reproduced for the augmented database, with the creation of synthetic instances using the SMOTE method (increase of 150% considering 2 neighbors). Tables 8 and 9 show the results of the classifiers considering only subject 8.

Unlike the original base, configuration 6 stood out with the enlarged base. In this case, the mean accuracy was 92.47 ± 1.19 and the mean kappa index was 0.85 ± 0.02 . Considering the standard deviation, it is possible to notice that configuration 7 has similar results. However, configuration 4 presented results inferior to the others, with an average accuracy of 86.26 ± 1.39 .

Again, boxplots with the best classifiers for each configuration were plotted, as shown in Figure 8. The boxplots clarify the highlight for settings 6 and 7.

Finally, configuration 6 was chosen for the experiments with the other subjects. In this case, however, the datasets have been expanded using two increase percentages: 50% and 150%. The augmentation was applied to each subject separately and the percentage of increase was applied in each class (right hand and left hand). Furthermore, SMOTE was applied only on train/test subset (70% of the database). Then, both augmented databases were experimented with decision trees using 10-fold cross-validation.

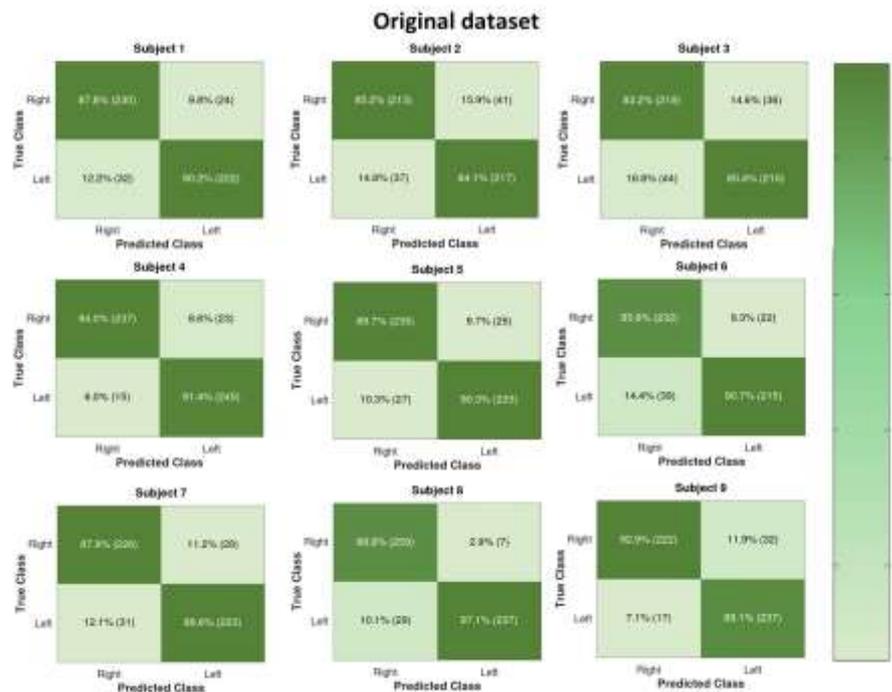


Fig. 7: Confusion matrices for all nine subjects data from Original dataset. Validation was performed with Random Forest classifier with 100 trees.

Setup	Accuracy - Subject 8 (SMOTE)									
	RF 10	RF 20	RF 30	RF 40	RF 50	RF 60	RF 70	RF 80	RF 90	RF 100
Setup 1	84.11 ± 1.45	87.37 ± 1.38	88.58 ± 1.37	89.25 ± 1.28	89.64 ± 1.25	89.93 ± 1.33	90.16 ± 1.27	90.52 ± 1.25	90.45 ± 1.27	90.53 ± 1.25
Setup 2	82.49 ± 1.66	86.19 ± 1.46	87.76 ± 1.32	88.58 ± 1.36	89.13 ± 1.33	89.46 ± 1.28	89.72 ± 1.27	89.96 ± 1.26	90.11 ± 1.24	90.25 ± 1.26
Setup 3	82.84 ± 1.76	86.14 ± 1.57	87.63 ± 1.45	88.41 ± 1.40	88.85 ± 1.47	89.17 ± 1.39	89.39 ± 1.37	89.56 ± 1.38	89.73 ± 1.36	89.83 ± 1.36
Setup 4	82.28 ± 1.64	85.74 ± 1.51	87.10 ± 1.52	87.77 ± 1.47	88.23 ± 1.52	88.51 ± 1.47	88.78 ± 1.46	88.96 ± 1.46	89.14 ± 1.42	89.20 ± 1.43
Setup 5	81.52 ± 1.70	85.97 ± 1.57	84.86 ± 1.52	85.29 ± 1.48	85.61 ± 1.41	85.83 ± 1.48	85.95 ± 1.44	86.04 ± 1.43	86.19 ± 1.41	86.26 ± 1.39
Setup 6	86.35 ± 1.53	89.10 ± 1.38	90.24 ± 1.41	90.98 ± 1.30	91.43 ± 1.23	91.76 ± 1.26	92.03 ± 1.23	92.15 ± 1.23	92.34 ± 1.24	92.47 ± 1.19
Setup 7	86.30 ± 1.54	89.05 ± 1.37	90.30 ± 1.33	91.04 ± 1.26	91.44 ± 1.19	91.82 ± 1.18	92.03 ± 1.21	92.19 ± 1.20	92.39 ± 1.16	92.44 ± 1.17
Setup 8	81.94 ± 2.09	86.34 ± 1.99	88.17 ± 1.85	88.97 ± 1.78	89.58 ± 1.72	89.97 ± 1.71	90.26 ± 1.69	90.49 ± 1.71	90.66 ± 1.67	90.82 ± 1.75
Setup 9	82.03 ± 1.69	86.25 ± 1.56	87.87 ± 1.49	88.70 ± 1.48	89.26 ± 1.39	89.64 ± 1.39	89.90 ± 1.37	90.12 ± 1.35	90.30 ± 1.33	90.46 ± 1.36
Setup 10	83.17 ± 1.66	85.79 ± 1.49	86.97 ± 1.40	87.82 ± 1.40	88.32 ± 1.30	88.73 ± 1.36	89.04 ± 1.30	89.20 ± 1.32	89.42 ± 1.28	89.56 ± 1.28

Table 8: Classification accuracy for subject 8. Different signal pre-processing and decision trees configurations were used for the expanded database by 150%.

For the database with an increase of 150%, participants 2,3 and 8 stood out with accuracies around 93% and kappa index around 0.86, according to Tables 10 and 11. For the database with an increase of 50%, the results were intermediate. In other words,

higher than the results with the original database and lower than the base obtained with a higher magnification. In this case, subjects 4 and 8 showed better performance, similar to the original dataset. For subject 4, the mean accuracy obtained was 85.08 ± 2.36 (Table 12) and mean kappa index was 0.70 ± 0.05 (Table 13). In contrast, subject 1 performed worse, with mean accuracy of 84.35 ± 2.38 and mean kappa index of 0.69 ± 0.05 .

Setup	Kappa index - Subject 8									
	RF 10	RF 20	RF 30	RF 40	RF 50	RF 60	RF 70	RF 80	RF 90	RF 100
Setup 1	0.26 ± 0.06	0.29 ± 0.06	0.31 ± 0.06	0.31 ± 0.06	0.32 ± 0.06	0.32 ± 0.06	0.32 ± 0.06	0.32 ± 0.06	0.32 ± 0.06	0.32 ± 0.06
Setup 2	0.22 ± 0.07	0.26 ± 0.07	0.27 ± 0.07	0.27 ± 0.07	0.28 ± 0.07	0.28 ± 0.07	0.28 ± 0.07	0.29 ± 0.07	0.29 ± 0.07	0.29 ± 0.07
Setup 3	0.24 ± 0.06	0.26 ± 0.06	0.27 ± 0.06	0.27 ± 0.06	0.27 ± 0.06	0.27 ± 0.06	0.27 ± 0.06	0.28 ± 0.06	0.28 ± 0.06	0.28 ± 0.06
Setup 4	0.24 ± 0.06	0.27 ± 0.06	0.27 ± 0.06	0.28 ± 0.06	0.28 ± 0.06	0.28 ± 0.06	0.28 ± 0.06	0.28 ± 0.06	0.28 ± 0.06	0.28 ± 0.06
Setup 5	0.34 ± 0.06	0.37 ± 0.06	0.38 ± 0.06	0.39 ± 0.06	0.40 ± 0.06	0.40 ± 0.06	0.40 ± 0.06	0.40 ± 0.06	0.41 ± 0.06	0.40 ± 0.06
Setup 6	0.45 ± 0.06	0.46 ± 0.06	0.46 ± 0.06	0.47 ± 0.06	0.47 ± 0.06	0.47 ± 0.06	0.47 ± 0.06	0.47 ± 0.05	0.47 ± 0.05	0.48 ± 0.05
Setup 7	0.45 ± 0.06	0.46 ± 0.05	0.46 ± 0.05	0.47 ± 0.05	0.47 ± 0.05	0.47 ± 0.06	0.47 ± 0.06	0.47 ± 0.06	0.48 ± 0.05	0.48 ± 0.05
Setup 8	0.17 ± 0.07	0.21 ± 0.08	0.23 ± 0.08	0.24 ± 0.08	0.25 ± 0.08	0.26 ± 0.08	0.26 ± 0.08	0.27 ± 0.08	0.27 ± 0.08	0.27 ± 0.08
Setup 9	0.19 ± 0.06	0.22 ± 0.06	0.24 ± 0.06	0.25 ± 0.07	0.26 ± 0.06	0.26 ± 0.06	0.26 ± 0.06	0.27 ± 0.06	0.27 ± 0.06	0.27 ± 0.06
Setup 10	0.46 ± 0.06	0.47 ± 0.06	0.48 ± 0.06	0.49 ± 0.06	0.49 ± 0.06	0.49 ± 0.06	0.49 ± 0.06	0.49 ± 0.06	0.49 ± 0.05	0.49 ± 0.06

Table 9: Classification Kappa index for subject 8. Different signal pre-processing and decision trees configurations were used for the expanded database by 150%.

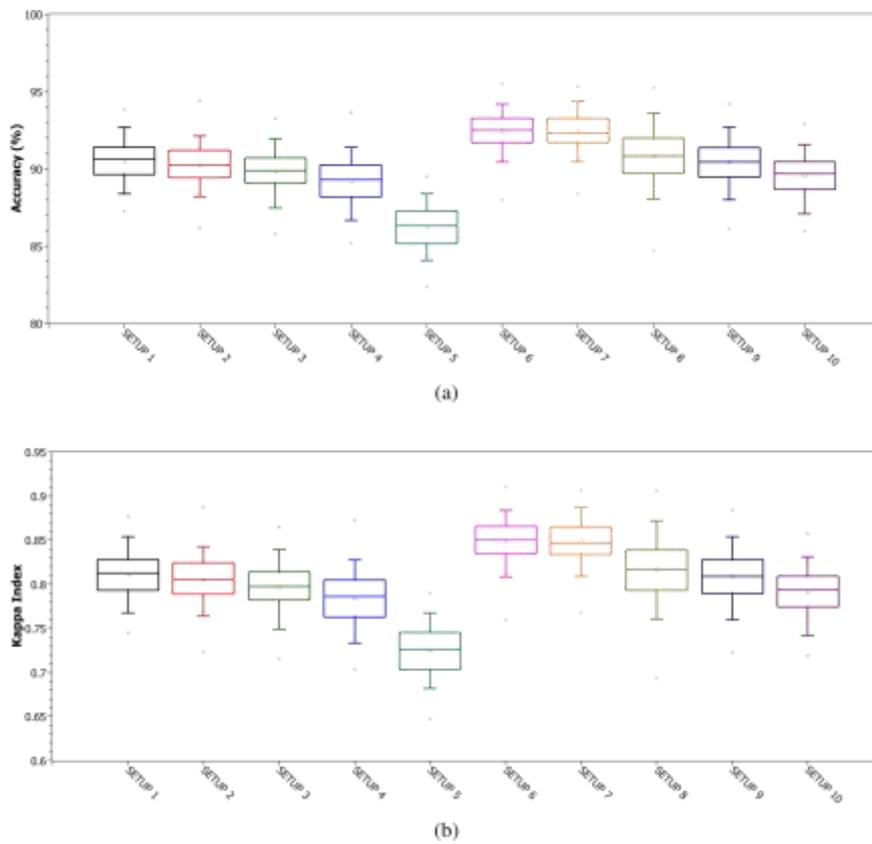


Fig. 8: Subject 8's motor imagery classification performance using the augmented database with the SMOTE method.

Furthermore, Figures 10 and 9 present validation confusion matrices for all nine subjects, considering both datasets (expanded by 50% and 100%, respectively). In these cases, Random Forest with 100 trees was applied, considering its superior performance in most train/test cases.

By observing the confusion matrix from Figure 9, we can see that all subjects had a good performance, with a hit rate of the "Right" class between 83.6% (Subject 3) and 94.4% (Subject 4). In the case of the "Left" class, the hit rate varied between 84.8% (Subject 2) and 95.9% (Subject 8).

In addition, the dataset augmented by 50% (Figure 10) also present great results. In this case, subject 9 had superior results in the classification of "Right" class (94.4%), while subject 2 had the worse result in the same classification (84.2%). Considering the "Left" class, the results achieved by all subjects were in the range of 85.9–96.0%. Finally, Table 14 presents a summary of all validation experiments, considering accuracy and kappa index. As detailed before, the best classifier configuration for each subject (obtained in the training/test step) was applied in the validation subset.

Subject	Accuracy (SMOTE 150%)									
	RF 10	RF 20	RF 30	RF 40	RF 50	RF 60	RF 70	RF 80	RF 90	RF 100
Subject 1	81.89 ± 1.98	86.05 ± 1.75	87.54 ± 1.75	88.36 ± 1.71	88.92 ± 1.62	89.33 ± 1.67	89.53 ± 1.57	89.78 ± 1.56	89.93 ± 1.53	90.09 ± 1.45
Subject 2	79.87 ± 2.07	85.81 ± 1.67	88.50 ± 1.61	89.95 ± 1.63	91.03 ± 1.56	91.64 ± 1.51	92.19 ± 1.46	92.56 ± 1.42	92.93 ± 1.35	93.25 ± 1.27
Subject 3	78.05 ± 2.19	84.98 ± 1.86	87.94 ± 1.78	89.73 ± 1.61	90.79 ± 1.58	91.56 ± 1.42	92.04 ± 1.48	92.50 ± 1.44	93.00 ± 1.47	93.27 ± 1.33
Subject 4	88.05 ± 1.67	89.54 ± 1.57	90.06 ± 1.53	90.34 ± 1.51	90.48 ± 1.50	90.59 ± 1.47	90.62 ± 1.47	90.67 ± 1.46	90.71 ± 1.40	90.71 ± 1.44
Subject 5	83.82 ± 1.98	86.78 ± 1.92	88.00 ± 1.90	88.66 ± 1.82	88.95 ± 1.81	89.22 ± 1.78	89.29 ± 1.76	89.40 ± 1.82	89.54 ± 1.78	89.66 ± 1.74
Subject 6	80.89 ± 1.92	85.42 ± 1.83	87.24 ± 1.70	88.27 ± 1.78	89.62 ± 1.76	89.48 ± 1.66	89.83 ± 1.72	90.11 ± 1.68	90.36 ± 1.70	90.53 ± 1.64
Subject 7	82.70 ± 1.99	86.19 ± 1.79	87.54 ± 1.79	88.19 ± 1.79	88.66 ± 1.73	89.12 ± 1.72	89.29 ± 1.73	89.48 ± 1.71	89.58 ± 1.67	89.63 ± 1.63
Subject 8	86.45 ± 1.73	89.38 ± 1.61	90.71 ± 1.54	91.47 ± 1.43	91.92 ± 1.39	92.21 ± 1.42	92.48 ± 1.42	92.64 ± 1.35	92.81 ± 1.34	92.91 ± 1.33
Subject 9	83.91 ± 1.94	86.37 ± 1.87	87.31 ± 1.72	87.81 ± 1.65	88.09 ± 1.63	88.77 ± 1.70	88.80 ± 1.63	88.51 ± 1.61	88.63 ± 1.54	88.70 ± 1.59

Table 10: Classification accuracy of the nine subjects, using the augmented database by 150%. Configuration 6 of pre-processing and attribute extraction was used. Different decision trees were trained with subject-specific data.

Subject	Kappa Index (SMOTE 150%)									
	RF 10	RF 20	RF 30	RF 40	RF 50	RF 60	RF 70	RF 80	RF 90	RF 100
Subject 1	0.64 ± 0.04	0.72 ± 0.03	0.75 ± 0.03	0.77 ± 0.03	0.78 ± 0.03	0.79 ± 0.03	0.79 ± 0.03	0.80 ± 0.03	0.80 ± 0.03	0.80 ± 0.03
Subject 2	0.60 ± 0.04	0.72 ± 0.03	0.77 ± 0.03	0.80 ± 0.03	0.82 ± 0.03	0.83 ± 0.03	0.84 ± 0.03	0.85 ± 0.03	0.86 ± 0.03	0.86 ± 0.03
Subject 3	0.57 ± 0.04	0.70 ± 0.04	0.76 ± 0.04	0.79 ± 0.03	0.82 ± 0.03	0.83 ± 0.03	0.84 ± 0.03	0.85 ± 0.03	0.86 ± 0.03	0.87 ± 0.03
Subject 4	0.76 ± 0.03	0.79 ± 0.03	0.80 ± 0.03	0.81 ± 0.03	0.81 ± 0.03	0.81 ± 0.03	0.81 ± 0.03	0.81 ± 0.03	0.81 ± 0.03	0.81 ± 0.03
Subject 5	0.68 ± 0.04	0.74 ± 0.04	0.76 ± 0.04	0.77 ± 0.04	0.78 ± 0.04	0.78 ± 0.04	0.79 ± 0.04	0.79 ± 0.04	0.79 ± 0.04	0.79 ± 0.04
Subject 6	0.62 ± 0.04	0.71 ± 0.04	0.74 ± 0.03	0.77 ± 0.04	0.78 ± 0.04	0.79 ± 0.03	0.80 ± 0.03	0.80 ± 0.03	0.81 ± 0.03	0.81 ± 0.03
Subject 7	0.65 ± 0.04	0.72 ± 0.04	0.75 ± 0.04	0.76 ± 0.04	0.77 ± 0.03	0.78 ± 0.03	0.79 ± 0.03	0.79 ± 0.03	0.79 ± 0.03	0.79 ± 0.03
Subject 8	0.73 ± 0.03	0.79 ± 0.03	0.81 ± 0.03	0.83 ± 0.03	0.84 ± 0.03	0.84 ± 0.03	0.85 ± 0.03	0.85 ± 0.03	0.86 ± 0.03	0.86 ± 0.03
Subject 9	0.68 ± 0.04	0.73 ± 0.04	0.75 ± 0.03	0.76 ± 0.03	0.76 ± 0.03	0.77 ± 0.03	0.77 ± 0.03	0.77 ± 0.03	0.77 ± 0.03	0.77 ± 0.03

Table 11: Kappa index in the classification of the nine subjects, using the augmented database by 150%. Configuration 6 of pre-processing and attribute extraction was used. Different decision trees were trained with subject-specific data.

4 Discussion

According to the experimental results, we can notice a highlight for configurations 10 and 6 for the original and extended databases with SMOTE, respectively. According to Table

3, both configurations make use of signal pre-processing with Butterworth bandpass filters, indicating success in removing artifacts from this technique. On the other hand, both configurations do not use wavelet decomposition techniques combined with statistical thresholds. This indicates that these latter approaches likely removed important information from the signals. Despite this, future experiments can apply a greater number of levels in the wavelet decomposition, since only 2 and 10 levels were tested. It is possible that this change will bring interesting results.

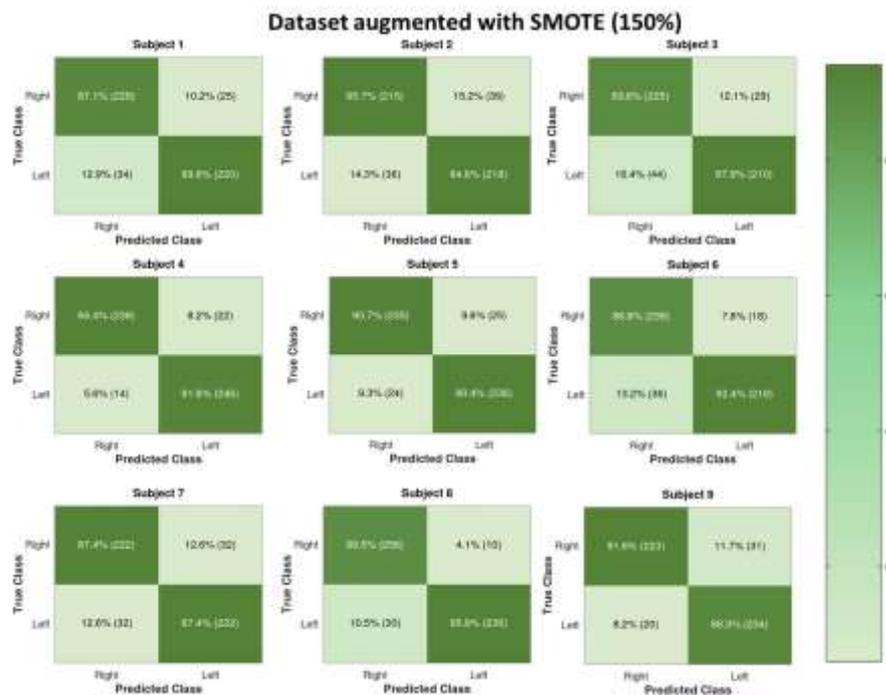


Fig. 9: Confusion matrices for all nine subjects data from expanded dataset by 150%. Validation was performed with Random Forest classifier with 100 trees.

Subject	Accuracy (SMOTE 50%)										
	RF 10	RF 20	RF 30	RF 40	RF 50	RF 60	RF 70	RF 80	RF 90	RF 100	
Subject 1	71.01 ± 3.00	74.09 ± 2.79	75.57 ± 2.86	76.40 ± 2.72	76.81 ± 2.72	77.37 ± 2.71	77.47 ± 2.70	77.71 ± 2.76	77.93 ± 2.84	77.99 ± 2.73	
Subject 2	66.53 ± 3.33	70.80 ± 3.18	72.94 ± 3.05	74.45 ± 3.06	75.59 ± 2.98	76.25 ± 3.07	76.74 ± 3.03	77.12 ± 3.02	77.59 ± 3.07	77.82 ± 3.07	
Subject 3	65.98 ± 3.06	70.57 ± 3.03	72.78 ± 3.09	74.46 ± 2.89	75.48 ± 2.79	76.49 ± 2.92	77.06 ± 2.94	77.56 ± 2.79	78.02 ± 2.83	78.14 ± 2.84	
Subject 4	82.55 ± 2.56	83.93 ± 2.37	84.38 ± 2.39	84.53 ± 2.40	84.71 ± 2.39	84.90 ± 2.44	84.96 ± 2.37	85.06 ± 2.33	85.05 ± 2.36	85.08 ± 2.35	
Subject 5	75.48 ± 2.93	77.87 ± 2.69	78.95 ± 2.49	79.55 ± 2.58	79.89 ± 2.59	79.97 ± 2.65	80.06 ± 2.58	80.12 ± 2.58	80.26 ± 2.53	80.35 ± 2.48	
Subject 6	69.88 ± 3.09	73.61 ± 3.16	74.92 ± 3.05	75.70 ± 2.99	76.34 ± 2.90	76.83 ± 2.88	77.29 ± 2.97	77.52 ± 3.00	77.83 ± 3.02	77.90 ± 3.02	
Subject 7	72.60 ± 3.07	75.64 ± 2.92	76.92 ± 3.19	77.64 ± 3.11	77.94 ± 2.99	78.37 ± 2.94	78.42 ± 2.89	78.60 ± 3.00	78.70 ± 2.98	78.80 ± 2.87	
Subject 8	79.30 ± 2.62	81.26 ± 2.58	82.35 ± 2.59	82.89 ± 2.54	83.35 ± 2.42	83.60 ± 2.43	83.87 ± 2.30	84.10 ± 2.34	84.31 ± 2.35	84.35 ± 2.38	
Subject 9	75.69 ± 2.92	77.60 ± 2.91	78.42 ± 3.04	78.83 ± 2.92	79.17 ± 2.88	79.23 ± 2.89	79.36 ± 2.79	79.43 ± 2.75	79.45 ± 2.74	79.58 ± 2.84	

Table 12: Accuracy in the classification of the nine subjects, using the expanded database by 50%. Configuration 6 of pre-processing and attribute extraction was used. Different decision trees were trained with subject-specific data.

Subject	Kappa Index (SMOTE 50%)									
	RF 10	RF 20	RF 30	RF 40	RF 50	RF 60	RF 70	RF 80	RF 90	RF 100
Subject 1	0.42 ± 0.06	0.48 ± 0.06	0.51 ± 0.06	0.53 ± 0.05	0.54 ± 0.05	0.55 ± 0.05	0.55 ± 0.05	0.55 ± 0.06	0.56 ± 0.06	0.56 ± 0.05
Subject 2	0.33 ± 0.07	0.42 ± 0.06	0.46 ± 0.06	0.49 ± 0.06	0.51 ± 0.06	0.52 ± 0.06	0.53 ± 0.06	0.54 ± 0.06	0.55 ± 0.06	0.56 ± 0.06
Subject 3	0.32 ± 0.06	0.41 ± 0.06	0.46 ± 0.06	0.49 ± 0.06	0.51 ± 0.06	0.53 ± 0.06	0.54 ± 0.06	0.55 ± 0.06	0.56 ± 0.06	0.56 ± 0.06
Subject 4	0.65 ± 0.05	0.68 ± 0.05	0.69 ± 0.05	0.69 ± 0.05	0.69 ± 0.05	0.70 ± 0.05	0.70 ± 0.05	0.70 ± 0.05	0.70 ± 0.05	0.70 ± 0.05
Subject 5	0.51 ± 0.06	0.56 ± 0.05	0.58 ± 0.05	0.59 ± 0.05	0.60 ± 0.05	0.60 ± 0.05	0.60 ± 0.05	0.60 ± 0.05	0.61 ± 0.05	0.61 ± 0.05
Subject 6	0.40 ± 0.06	0.47 ± 0.06	0.50 ± 0.06	0.51 ± 0.06	0.53 ± 0.06	0.54 ± 0.06	0.55 ± 0.06	0.55 ± 0.06	0.56 ± 0.06	0.56 ± 0.06
Subject 7	0.45 ± 0.06	0.51 ± 0.06	0.54 ± 0.06	0.55 ± 0.06	0.56 ± 0.06	0.56 ± 0.06	0.57 ± 0.06	0.57 ± 0.06	0.57 ± 0.06	0.58 ± 0.06
Subject 8	0.59 ± 0.05	0.63 ± 0.05	0.65 ± 0.05	0.66 ± 0.05	0.67 ± 0.05	0.67 ± 0.05	0.68 ± 0.05	0.68 ± 0.05	0.69 ± 0.05	0.69 ± 0.05
Subject 9	0.51 ± 0.06	0.55 ± 0.06	0.57 ± 0.06	0.58 ± 0.06	0.58 ± 0.06	0.58 ± 0.06	0.59 ± 0.06	0.59 ± 0.06	0.59 ± 0.05	0.59 ± 0.06

Table 13: Kappa index in the classification of the nine subjects, using the expanded database by 50%. Configuration 6 of pre-processing and attribute extraction was used. Different decision trees were trained with subject-specific data.

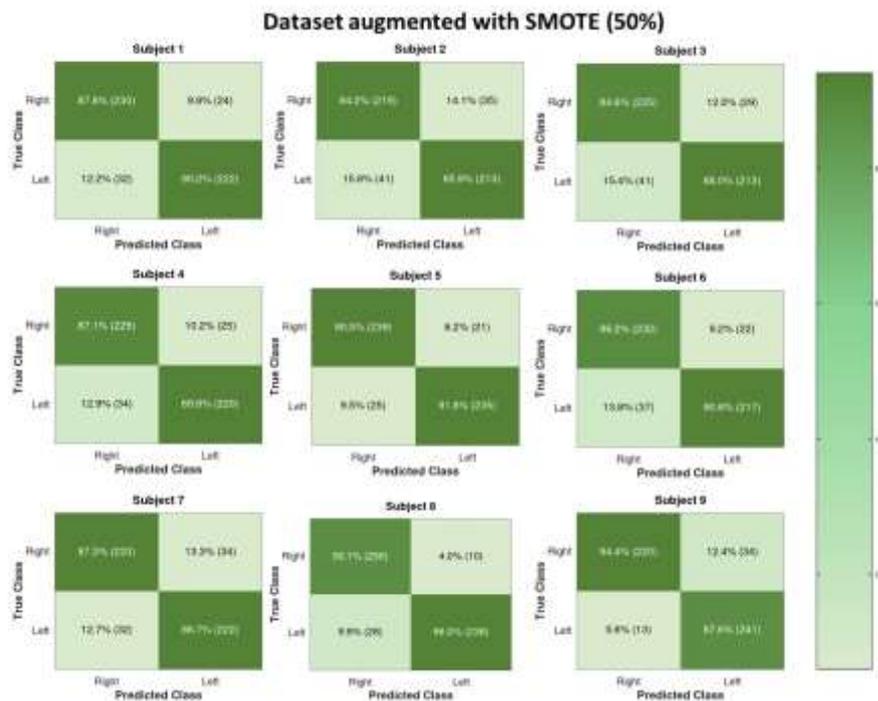


Fig. 10: Confusion matrices for all nine subjects data from expanded dataset by 50%. Validation was performed with Random Forest classifier with 100 trees.

Also, using the SMOTE method to extend the database, more experiments were designed. However, when we look at the validation results, we obtained similar results with all three approaches (with and without expansion), as pointed in the Table 14. For subject 8, for instance, it achieved classification accuracy of 93.23%, 92.86% and 92.48% for original, expanded by 50% and 150% datasets, respectively. With respect to the kappa index, it achieved 0.86 for both original and with 50% expansion, and 0.85 for the third dataset. These results indicate that the collection of more data per patient is not necessary.

That is, the data collected by the database 2b from BCI Competition IV is sufficient for training and generalization by using the ASTERI method, preferably using the fusion of implicit and explicit attributes to improve results.

Subject	Original		SMOTE 50%		SMOTE 150%	
	Accuracy (%)	Kappa index	Accuracy (%)	Kappa index	Accuracy (%)	Kappa index
Subject 1	89.17	0.78	88.98	0.78	88.39	0.77
Subject 2	84.65	0.69	85.04	0.70	85.24	0.70
Subject 3	84.25	0.69	86.22	0.72	85.63	0.71
Subject 4	92.69	0.85	91.54	0.83	93.08	0.86
Subject 5	90.00	0.80	91.15	0.82	90.58	0.81
Subject 6	87.99	0.76	88.39	0.77	89.37	0.79
Subject 7	88.39	0.77	87.01	0.74	87.40	0.75
Subject 8	93.23	0.86	92.86	0.86	92.48	0.85
Subject 9	90.35	0.81	90.75	0.82	89.96	0.80

Table 14: Validation results of the nine subjects, considering the databases with and without the addition of synthetic instances.

These results are interesting, especially in assistive technology rehabilitation scenarios, where acquiring small amounts of signals is less tiring and tedious. In addition, by increasing the collection time of the sessions, there is a risk of the volunteer becoming more dispersed and less efficient in the imagery of the indicated class. Consequently, the signal quality may also deteriorate.

Finally, the Table 15 below provides a comparison between the results presented in this work and other state-of-the-art works, using the same database (BCI IV - 2b). Thereby, it is clear that ASTERI method achieved results superior to many state-of-art works in the last 5 years. This result is valuable, considering the hybrid nature of the ASTERI method, which combines deep networks for attribute extraction and decision trees in classification. Thus, EEG signal classification problems can be solved in a less complex way and with satisfactory results.

Furthermore, the work of Dai et al. (2020) also expanded the data, tripling their number. That is, the increase applied was even greater than the two cases proposed in this work. The method used in this case was not SMOTE. The idea applied was the recombination of signals belonging to the same class and the same subject. Despite the relevant results, the method may have added high frequency noise, the peaks and valleys of different signal segments may have been put together.

We also observed that many of these studies considered accuracy only as an evaluation metric (Dai et al., 2020, Li et al., 2017, Lu et al., 2016). However, the BCI competition uses the kappa index to compare results, as it is a more rigorous metric.

Subjects	Accuracy				
	RF 100	G Dai et al (2020)	Na Lu et al (2016)	Tabar et al (2017)	Li et al (2017)
Subject 1	89.17	80.5	81	76	84.6
Subject 2	84.65	70.6	65	65.8	66.2
Subject 3	84.25	85.6	66	75.3	62.9
Subject 4	92.69	94.6	98	95.3	95.8
Subject 5	90.00	98.3	93	83	89.2
Subject 6	87.99	86.6	88	79.5	97.9
Subject 7	88.39	89.6	82	74.5	82.1
Subject 8	93.23	95.6	94	75.3	86.2
Subject 9	90.35	87.4	91	73.3	97.1
Mean	88.97	87.6	84	77.6	84.7

Table 15: Comparison of the proposed method with others of the state of the art.

5 Conclusion

In this work, we propose a new method to represent EEG signals in images: the ASTERI method. ASTERI generates pseudo-sinograms, i.e. pre-processed signal windows on which the rear-projection reconstruction method is applied, generating images. The main objective of this type of representation is to facilitate the use of deep architectures such as CNNs in applications with EEG signals.

To validate this approach, we applied the method to motor imagery signals from the BCI Competition IV 2b database. The signals were windowed and then we tested different preprocessing approaches. Among the 10 approaches tested, the best results were found when using a bandpass filter (8-32 Hz) and extracting attributes with the pre-trained VGG16 network.

Thus, the average accuracy obtained among the nine volunteers was 88.97% and the average kappa index was 0.78, using Random Forest with 100 trees in validation step. In addition to experiments with the collected data, we also expanded the database with synthetic instances generated by the SMOTE method, for each subject. For a 50% increase in the instances of each class, the accuracy and average kappa index were 89.10% and 0.78, respectively. For an increase of 150%, the average accuracy obtained was 89.13%. Thereby, we can see that the results with and without database expansion were very close with ASTERI method. This shows that the signals collected in the motor imagery sections of dataset 2b are sufficient for training, generalization and subsequent control of brain-machine interfaces. This result is valuable in the sense that increasing the acquisition time can be tiring and affect the participants' concentration, deteriorating the quality of the signals.

Finally, these results indicate that ASTERI is effective in classifying motor imagery and generates results that are competitive with the state-of-art. In addition, ASTERI is also interesting for its low complexity, especially for the use of pre-trained deep networks combined with Random Forests and unique image reconstruction method for all types of EEG signals, regardless of the subject. In the future, the ASTERI method should be tested in other motor imagery databases and in other applications based on EEG signals.

Acknowledgements

The authors are grateful to the Brazilian research agencies Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco, FACEPE, grant IBPG-1557-1.03/18, and Conselho Nacional de Desenvolvimento Científico e Tecnológico, CNPq, grant 314896/2018-0, for the partial financial support of this research.

Conflict of Interest

All authors declare they have no conflicts of interest.

Compliance with Ethical Standards

This study was funded by the Brazilian research agencies FACEPE and CNPq.

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

References

- R. Abiri, S. Borhani, E. W. Sellers, Y. Jiang, and X. Zhao. A comprehensive review of eegbased brain–computer interface paradigms. *Journal of neural engineering*, 16(1):011001, 2019.
- M. AlAbboudi, M. Majed, F. Hassan, and A. B. Nassif. Eeg wheelchair for people of determination. In *2020 Advances in Science and Engineering Technology International Conferences (ASET)*, pages 1–5. IEEE, 2020.
- S. U. Amin, M. Alsulaiman, G. Muhammad, M. A. Mekhtiche, and M. S. Hossain. Deep learning for eeg motor imagery classification based on multi-layer cnns feature fusion. *Future Generation computer systems*, 101:542–554, 2019.
- U. Amin and S. R. Benbadis. The role of eeg in the erroneous diagnosis of epilepsy. *Journal of Clinical Neurophysiology*, 36(4):294–297, 2019.
- S. An, S. Kim, P. Chikontwe, and S. H. Park. Few-shot relation learning with attention for eeg-based motor imagery classification. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 10933–10938. IEEE, 2020.
- M. K. Andrade, M. A. d. Santana, G. Moreno, I. Oliveira, J. Santos, M. C. A. Rodrigues, and W. P. d. Santos. An eeg brain-computer interface to classify motor imagery signals. In *Biomedical signal processing*, pages 83–98. Springer, 2020.
- K. K. Ang, Z. Y. Chin, C. Wang, C. Guan, and H. Zhang. Filter bank common spatial pattern algorithm on bci competition iv datasets 2a and 2b. *Frontiers in neuroscience*, 6:39, 2012.
- N. Bajaj, J. R. Carrión, F. Bellotti, R. Berta, and A. De Gloria. Automatic and tunable algorithm for eeg artifact removal using wavelet decomposition with applications in predictive modeling during auditory tasks. *Biomedical Signal Processing and Control*, 55:101624, 2020.
- A. M. Batula, J. A. Mark, Y. E. Kim, and H. Ayaz. Comparison of brain activation during motor imagery and motor movement using fnirs. *Computational intelligence and neuroscience*, 2017, 2017.
- S. R. Benbadis. The tragedy of over-read eegs and wrong diagnoses of epilepsy. *Expert Review of Neurotherapeutics*, 10(3):343–346, 2010. doi: 10.1586/ern.09.157. URL <https://doi.org/10.1586/ern.09.157>.

- A. Bouti, M. A. Mahraz, J. Riffi, and H. Tairi. A robust system for road sign detection and classification using lenet architecture based on convolutional neural network. *Soft Computing*, 24(9):6721–6733, 2020.
- R. Cassani, M. Estrellas, R. San-Martin, F. J. Fraga, and T. H. Falk. Systematic review on resting-state eeg for alzheimer’s disease diagnosis and progression assessment. *Disease markers*, 2018, 2018.
- N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer. Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16:321–357, 2002.
- Z. Chen, G. Lu, Z. Xie, and W. Shang. A unified framework and method for eeg-based early epileptic seizure detection and epilepsy diagnosis. *IEEE Access*, 8:20080–20092, 2020.
- A. Craik, Y. He, and J. L. Contreras-Vidal. Deep learning for electroencephalogram (eeg) classification tasks: a review. *Journal of neural engineering*, 16(3):031001, 2019.
- A. Cutler, D. R. Cutler, and J. R. Stevens. Random forests. In *Ensemble machine learning*, pages 157–175. Springer, 2012.
- M. da Silva Junior, R. C. de Freitas, W. P. dos Santos, W. W. A. da Silva, M. C. A. Rodrigues, and E. F. Q. Conde. Exploratory study of the effect of binaural beat stimulation on the eeg activity pattern in resting state using artificial neural networks. *Cognitive Systems Research*, 54:1–20, 2019.
- G. Dai, J. Zhou, J. Huang, and N. Wang. Hs-cnn: a cnn with hybrid convolution scale for eeg motor imagery classification. *Journal of neural engineering*, 17(1):016025, 2020.
- R. C. de Freitas, R. Alves, A. G. da Silva Filho, R. E. de Souza, B. L. Bezerra, and W. P. dos Santos. Electromyography-controlled car: A proof of concept based on surface electromyography, extreme learning machines and low-cost open hardware. *Computers & Electrical Engineering*, 73:167–179, 2019.
- R. C. de Freitas, G. R. Naik, M. J. S. Valença, B. L. D. Bezerra, R. E. de Souza, and W. P. dos Santos. Surface electromyography classification using extreme learning machines and echo state networks. *Research on Biomedical Engineering*, 2022:1–22, 2022.
- V. A. de Freitas Barbosa, M. A. de Santana, M. K. S. Andrade, R. d. C. F. de Lima, and W. P. dos Santos. Deep-wavelet neural networks for breast cancer early diagnosis using mammary termographies. In *Deep Learning for Data Analytics*, pages 99–124. Elsevier, 2020.
- V. A. de Freitas Barbosa, J. C. Gomes, M. A. de Santana, E. d. A. Jeniffer, R. G. de Souza, R. E. de Souza, and W. P. dos Santos. Heg. ia: An intelligent system to support diagnosis of covid-19 based on blood tests. *Research on Biomedical Engineering*, pages 1–18, 2021.
- A. P. S. de Oliveira, M. A. de Santana, M. K. S. Andrade, J. C. Gomes, M. C. Rodrigues, and W. P. dos Santos. Early diagnosis of parkinson’s disease using eeg, machine learning and partial directed coherence. *Research on Biomedical Engineering*, 36(3):311–331, 2020.
- L. Demidova and I. Klyueva. Svm classification: Optimization with the smote algorithm for the class imbalance problem. In *2017 6th Mediterranean Conference on Embedded Computing (MECO)*, pages 1–4. IEEE, 2017.
- G. Douzas, F. Bacao, and F. Last. Improving imbalanced learning through a heuristic oversampling method based on k-means and smote. *Information Sciences*, 465:1–20, 2018.
- C. W. Espinola, J. C. Gomes, J. M. S. Pereira, and W. P. dos Santos. Detection of major depressive disorder using vocal acoustic analysis and machine learning—an exploratory study. *Research on Biomedical Engineering*, 37(1):53–64, 2021a.
- C. W. Espinola, J. C. Gomes, J. M. S. Pereira, and W. P. dos Santos. Vocal acoustic analysis and machine learning for the identification of schizophrenia. *Research on Biomedical Engineering*, 37(1):33–46, 2021b.

- V. d. A. Faria, M. Azimbagirad, G. Viani Arruda, J. Fernandes Pavoni, J. Cezar Felipe, E. M. C. M. F. dos Santos, L. O. Murta Junior, et al. Prediction of radiation-related dental caries through pyradiomics features and artificial neural network on panoramic radiography. *Journal of Digital Imaging*, 34(5):1237–1248, 2021.
- A. Fernández, S. Garcia, F. Herrera, and N. V. Chawla. Smote for learning from imbalanced data: progress and challenges, marking the 15-year anniversary. *Journal of artificial intelligence research*, 61:863–905, 2018.
- K. Fu, J. Qu, Y. Chai, and Y. Dong. Classification of seizure based on the time-frequency image of eeg signals using hht and svm. *Biomedical Signal Processing and Control*, 13: 15–22, 2014.
- G. Gayathri, G. Udupa, G. Nair, and S. Poorna. Eeg-controlled prosthetic arm for micromechanical tasks. In *Proceedings of the Second International Conference on Computational Intelligence and Informatics*, pages 281–291. Springer, 2018.
- L. A. Gemein, R. T. Schirrmeyer, P. Chraba_{szcz}, D. Wilson, J. Boedecker, A. SchulzeBonhage, F. Hutter, and T. Ball. Machine-learning-based diagnostics of eeg pathology. *NeuroImage*, 220:117021, 2020.
- P. Geurts, D. Ernst, and L. Wehenkel. Extremely randomized trees. *Machine Learning*, 63 (1):3–42, 2006.
- J. C. Gomes, V. A. Barbosa, D. E. Ribeiro, R. E. de Souza, and W. P. dos Santos. Electrical impedance tomography image reconstruction based on backprojection and extreme learning machines. *Research on Biomedical Engineering*, 36(4):399–410, 2020a.
- J. C. Gomes, V. A. d. F. Barbosa, M. A. Santana, J. Bandeira, M. J. S. Valença, R. E. de Souza, A. M. Ismael, and W. P. dos Santos. Ikonos: An intelligent tool to support diagnosis of covid-19 by texture analysis of x-ray images. *Research on Biomedical Engineering*, pages 1–14, 2020b.
- J. C. Gomes, M. A. de Santana, C. L. de Lima, R. E. de Souza, and W. P. dos Santos. Using extreme learning machines and the backprojection algorithm as an alternative to reconstruct electrical impedance tomography images. In *Biomedical Computing for Breast Cancer Detection and Diagnosis*, pages 16–27. IGI Global, 2021a.
- J. C. Gomes, A. I. Masood, L. H. d. S. Silva, J. R. B. da Cruz Ferreira, A. A. F. Júnior, A. L. dos Santos Rocha, L. C. P. de Oliveira, N. R. C. da Silva, B. J. T. Fernandes, and W. P. Dos Santos. Covid-19 diagnosis by combining rt-pcr and pseudo-convolutional machines to characterize virus sequences. *Scientific Reports*, 11(1):1–28, 2021b.
- R. Gordon, G. T. Herman, and S. A. Johnson. Image reconstruction from projections. *Scientific American*, 233(4):56–71, 1975.
- M. Hon and N. M. Khan. Towards Alzheimer’s disease classification through transfer learning. In *2017 IEEE International conference on bioinformatics and biomedicine (BIBM)*, pages 1166–1169. IEEE, 2017.
- I. Hossain, A. Khosravi, and S. Nahavandhi. Active transfer learning and selective instancetransfer with active learning for motor imagery based bci. In *2016 International Joint Conference on Neural Networks (IJCNN)*, pages 4048–4055. IEEE, 2016.
- J. Hsieh. *Computed tomography: principles, design, artifacts, and recent advances*, volume 114. SPIE press, 2003.
- S. Ibrahim, R. Djemal, and A. Alsuwailem. Electroencephalography (eeg) signal processing for epilepsy and autism spectrum disorder diagnosis. *Biocybernetics and Biomedical Engineering*, 38(1):16–26, 2018.

- C. Ieracitano, N. Mammone, A. Bramanti, A. Hussain, and F. C. Morabito. A convolutional neural network approach for classification of dementia stages based on 2d-spectral representation of eeg recordings. *Neurocomputing*, 323:96–107, 2019.
- C. Ieracitano, N. Mammone, A. Hussain, and F. C. Morabito. A novel explainable machine learning approach for eeg-based brain-computer interface systems. *Neural Computing and Applications*, pages 1–14, 2021.
- N. M. Isa, A. Amir, M. Ilyas, and M. Razalli. Motor imagery classification in brain computer interface (bci) based on eeg signal by using machine learning technique. *Bulletin of Electrical Engineering and Informatics*, 8(1):269–275, 2019.
- M. R. Islam, M. A. Rahim, H. Akter, R. Kabir, and J. Shin. Optimal imf selection of emd for sleep disorder diagnosis using eeg signals. In *Proceedings of the 3rd International Conference on Applications in Information Technology*, pages 96–101, 2018.
- K. Jiang, J. Lu, and K. Xia. A novel algorithm for imbalance data classification based on genetic algorithm improved smote. *Arabian journal for science and engineering*, 41(8): 3255–3266, 2016.
- T. Kawano, N. Hattori, Y. Uno, K. Kitajo, M. Hatakenaka, H. Yagura, H. Fujimoto, T. Yoshioka, M. Nagasako, H. Otomune, et al. Large-scale phase synchrony reflects clinical status after stroke: An eeg study. *Neurorehabilitation and Neural Repair*, 31(6):561–570, 2017.
- S. K. Khare and V. Bajaj. Time-frequency representation and convolutional neural networkbased emotion recognition. *IEEE transactions on neural networks and learning systems*, 2020.
- C. Kingsford and S. L. Salzberg. What are decision trees? *Nature biotechnology*, 26(9): 1011–1013, 2008.
- N. Koshev, E. S. Helou, and E. X. Miqueles. Fast backprojection techniques for high resolution tomography. *arXiv preprint arXiv:1608.03589*, 2016.
- S. T. Krishna and H. K. Kalluri. Deep learning and transfer learning approaches for image classification. *International Journal of Recent Technology and Engineering (IJRTE)*, 7 (5S4):427–432, 2019.
- C.-C. J. Kuo. Understanding convolutional neural networks with a mathematical model. *Journal of Visual Communication and Image Representation*, 41:406–413, 2016.
- S. Kwon et al. A cnn-assisted enhanced audio signal processing for speech emotion recognition. *Sensors*, 20(1):183, 2020.
- M. E. A. Larabi, S. Chaib, K. Bakhti, and M. S. Karoui. Transfer learning for changes detection in optical remote sensing imagery. In *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium*, pages 1582–1585. IEEE, 2019.
- Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. *nature*, 521(7553):436–444, 2015.
- H. K. Lee and Y.-S. Choi. A convolution neural networks scheme for classification of motor imagery eeg based on wavelet time-frequency image. In *2018 International Conference on Information Networking (ICOIN)*, pages 906–909. IEEE, 2018.
- H. K. Lee and Y.-S. Choi. Application of continuous wavelet transform and convolutional neural network in decoding motor imagery brain-computer interface. *Entropy*, 21(12): 1199, 2019.
- R. Leeb, C. Brunner, G. Müller-Putz, A. Schlögl, and G. Pfurtscheller. Bci competition 2008–graz data set b. *Graz University of Technology, Austria*, pages 1–6, 2008.
- M.-a. Li, W. Zhu, H.-n. Liu, and J.-f. Yang. Adaptive feature extraction of motor imagery eeg with optimal wavelet packets and se-isomap. *Applied Sciences*, 7(4):390, 2017.
- N. Lu, T. Li, X. Ren, and H. Miao. A deep learning scheme for motor imagery classification based on restricted boltzmann machines. *IEEE transactions on neural systems and rehabilitation engineering*, 25(6):566–576, 2016.

- V. A. Luccas, M. Azimbagirad, L. C. Cabrelli, F. W. Grillo, T. Z. Pavan, L. Otavio Murta, J. Elias, and A. A. Carneiro. 3d automated breast ultrasound imaging: equipment and image reconstruction. *Research on Biomedical Engineering*, 37(3):519–533, 2021.
- W. Mao, H. Fathurrahman, Y. Lee, and T. Chang. Eeg dataset classification using cnn method. *Journal of physics: conference series*, 1456(1):012017, 2020.
- W. Mardini, M. M. B. Yassein, R. Al-Rawashdeh, S. Aljawarneh, Y. Khamayseh, and O. Meqdadi. Enhanced detection of epileptic seizure using eeg signals in combination with machine learning classifiers. *IEEE Access*, 8:24046–24055, 2020.
- M. L. McHugh. Interrater reliability: the kappa statistic. *Biochemia medica*, 22(3):276–282, 2012.
- E. Niedermeyer, W. Froescher, and R. Fisher. Epileptic seizure disorders. *Journal of neurology*, 232(1):1–12, 1985.
- S. L. Oh, Y. Hagiwara, U. Raghavendra, R. Yuvaraj, N. Arunkumar, M. Murugappan, and U. R. Acharya. A deep learning approach for parkinson’s disease diagnosis from eeg signals. *Neural Computing and Applications*, 32(15):10927–10933, 2020.
- F. Özyurt. Efficient deep feature selection for remote sensing image recognition with fused deep learning architectures. *The Journal of Supercomputing*, 76(11):8413–8431, 2020.
- N. Padfield, J. Zabalza, H. Zhao, V. Masero, and J. Ren. Eeg-based brain-computer interfaces using motor-imagery: Techniques and challenges. *Sensors*, 19(6):1423, 2019.
- S. J. Pan and Q. Yang. A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, 22(10):1345–1359, 2009.
- H. Polat and M. S. Özerdem. Automatic detection of cursor movements from the eeg signals via deep learning approach. In *2020 5th International Conference on Computer Science and Engineering (UBMK)*, pages 327–332. IEEE, 2020.
- R. A. Priyadarshini, S. Arivazhagan, M. Arun, and A. Mirnalini. Maize leaf disease classification using deep convolutional neural networks. *Neural Computing and Applications*, 31(12):8887–8895, 2019.
- N. Rampal, C. B. Maciel, and L. J. Hirsch. Electroencephalography and artifact in the intensive care unit. *Atlas of Artifacts in Clinical Neurophysiology*, page 59, 2018.
- A. K. Rangarajan and R. Purushothaman. Disease classification in eggplant using pretrained vgg16 and msvm. *Scientific reports*, 10(1):1–11, 2020.
- M. Rashid, N. Sulaiman, M. Mustafa, S. Khatun, and B. S. Bari. The classification of eeg signal using different machine learning techniques for bci application. In *International Conference on Robot Intelligence Technology and Applications*, pages 207–221. Springer, 2018.
- A. L. Rodrigues, M. A. de Santana, W. W. Azevedo, R. S. Bezerra, V. A. Barbosa, R. C. de Lima, and W. P. dos Santos. Identification of mammary lesions in thermographic images: feature selection study using genetic algorithms and particle swarm optimization. *Research on Biomedical Engineering*, 35(3):213–222, 2019.
- M. A. d. Santana, J. M. S. Pereira, F. L. d. Silva, N. M. d. Lima, F. N. d. Sousa, G. M. S. d. Arruda, R. d. C. F. d. Lima, W. W. A. d. Silva, and W. P. d. Santos. Breast cancer diagnosis based on mammary thermography and extreme learning machines. *Research on Biomedical Engineering*, 34:45–53, 2018.
- I. R. Silva, G. S. Silva, R. G. de Souza, W. P. dos Santos, and A. d. A. Roberta. Model based on deep feature extraction for diagnosis of alzheimer’s disease. In *2019 International Joint Conference on Neural Networks (IJCNN)*, pages 1–7. IEEE, 2019.
- M. d. Silva Júnior, R. C. d. Freitas, W. W. A. d. Silva, M. C. A. Rodrigues, E. F. Q. Conde, and W. P. d. Santos. Using artificial neural networks on multi-channel eeg data to detect the

- effect of binaural stimuli in resting state. In *Biomedical Signal Processing*, pages 99–136. Springer, 2020.
- K. Srinivasan, J. Dauwels, and M. R. Reddy. Multichannel eeg compression: Wavelet-based image and volumetric coding approach. *IEEE journal of biomedical and health informatics*, 17(1):113–120, 2012.
- Y. Sun, S. Liu, T. Zhao, Z. Zou, B. Shen, Y. Yu, S. Zhang, and H. Zhang. A new hydrogen sensor fault diagnosis method based on transfer learning with LeNet-5. *Frontiers in Neurorobotics*, 15:57, 2021.
- Y. Taghizadegan, N. J. Dabanloo, K. Maghooli, and A. Sheikhani. Obstructive sleep apnea event prediction using recurrence plots and convolutional neural networks (rp-cnns) from polysomnographic signals. *Biomedical Signal Processing and Control*, 69:102928, 2021.
- S. Tammina. Transfer learning using VGG-16 with deep convolutional neural network for classifying images. *International Journal of Scientific and Research Publications (IJSRP)*, 9(10):143–150, 2019.
- C. Tan, F. Sun, T. Kong, W. Zhang, C. Yang, and C. Liu. A survey on deep transfer learning. In *International conference on artificial neural networks*, pages 270–279. Springer, 2018.
- D. Theckedath and R. Sedamkar. Detecting affect states using VGG16, ResNet50 and SE-ResNet50 networks. *SN Computer Science*, 1(2):1–7, 2020.
- J. Thomas, J. Jin, P. Thangavel, E. Bagheri, R. Yuvaraj, J. Dauwels, R. Rathakrishnan, J. J. Halford, S. S. Cash, and B. Westover. Automated detection of interictal epileptiform discharges from scalp electroencephalograms by convolutional neural networks. *International Journal of Neural Systems*, 30(11):2050030, 2020.
- M. C. Tjepkema-Cloostermans, R. C. de Carvalho, and M. J. van Putten. Deep learning for detection of focal epileptiform discharges from scalp eeg recordings. *Clinical neurophysiology*, 129(10):2191–2196, 2018.
- I. Ullah, M. Hussain, H. Aboalsamh, et al. An automated system for epilepsy detection using eeg brain signals based on deep learning approach. *Expert Systems with Applications*, 107:61–71, 2018.
- O. Yildirim, M. Talo, B. Ay, U. B. Baloglu, G. Aydin, and U. R. Acharya. Automated detection of diabetic subject using pre-trained 2d-cnn models with frequency spectrum images extracted from heart rate signals. *Computers in biology and medicine*, 113:103387, 2019.
- K. Zhang, N. Robinson, S.-W. Lee, and C. Guan. Adaptive transfer learning for eeg motor imagery classification with deep convolutional neural network. *Neural Networks*, 136: 1–10, 2021.
- S. Zhang, F. Sun, N. Wang, C. Zhang, Q. Yu, M. Zhang, P. Babyn, and H. Zhong. Computeraided diagnosis (CAD) of pulmonary nodule of thoracic CT image using transfer learning. *Journal of digital imaging*, 32(6):995–1007, 2019.
- X. Zhao, H. Zhang, G. Zhu, F. You, S. Kuang, and L. Sun. A multi-branch 3d convolutional neural network for eeg-based motor imagery classification. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 27(10):2164–2177, 2019.
- X. Zhu, P. Li, C. Li, D. Yao, R. Zhang, and P. Xu. Separated channel convolutional neural network to realize the training free motor imagery bci systems. *Biomedical Signal Processing and Control*, 49:396–403, 2019.