

## Article

# Posthoc Interpretability of Neural Responses by Grouping Subject Motor Imagery Skills Using CNN-Based Connectivity

Diego Fabian Collazos-Huertas <sup>1,\*</sup>, Andrés Marino Álvarez-Meza <sup>1,†</sup>, David Augusto Cárdenas-Peña <sup>2,†</sup>, Germán Albeiro Castaño-Duque <sup>3,†</sup> and César Germán Castellanos-Domínguez <sup>1,†</sup>

<sup>1</sup> Signal Processing and Recognition Group, Universidad Nacional de Colombia, Manizales 170003, Colombia

<sup>2</sup> Automatics Research Group, Universidad Tecnológica de Pereria, Pereira 660003, Colombia

<sup>3</sup> Cultura de la Calidad en la Educación Research Group, Universidad Nacional de Colombia, Manizales 170003, Colombia

\* Correspondence: dfcollazosh@unal.edu.co

† These authors contributed equally to this work.

**Abstract:** Motor Imagery (MI) refers to imagining the mental representation of motor movements without overt motor activity, enhancing physical action execution and neural plasticity with potential applications in medical and professional fields like rehabilitation and education. Currently, the most promising approach for implementing the MI paradigm is the Brain-Computer Interface (BCI), which uses Electroencephalogram (EEG) sensors to detect brain activity. However, MI-BCI control depends on a synergy between user skills and EEG signal analysis. Thus, decoding brain neural responses recorded by scalp electrodes poses still challenging due to substantial limitations, such as non-stationarity and poor spatial resolution. Also, an estimated third of people need more skills to accurately perform MI tasks, leading to underperforming MI-BCI systems. As a strategy to deal with BCI-inefficiency, this study identifies subjects with poor motor performance at the early stages of BCI training by assessing and interpreting the neural responses elicited by MI across the evaluated subject set. Using connectivity features extracted from class activation maps, we propose a Convolutional Neural Network-based framework for learning relevant information from high-dimensional dynamical data to distinguish between MI tasks while preserving the post-hoc interpretability of neural responses. Two approaches deal with inter/intra-subject variability of MI EEG data: (a) Extracting functional connectivity from spatiotemporal class activation maps through a novel kernel-based cross-spectral distribution estimator, (b) Clustering the subjects according to their achieved classifier accuracy, aiming to find common and discriminative patterns of motor skills. According to the validation results obtained on a bi-class database, an average accuracy enhancement of 10% is achieved compared to the baseline EEGNet approach, reducing the number of “poor skill” subjects from 40% to 20%. Overall, the proposed method can be used to help explain brain neural responses even in subjects with deficient MI skills, who have neural responses with high variability and poor EEG-BCI performance.

**Keywords:** EEG Network; motor imagery; motor skills; functional connectivity; class activation maps



**Citation:** Collazos-Huertas, D.F.; Álvarez-Meza, A.M.; Cárdenas-Peña, D.A.; Castaño-Duque, G.A.; Castellanos-Domínguez, C.G. Posthoc Interpretability of Neural Responses by Grouping Subject Motor Imagery Skills Using CNN-Based Connectivity. *Sensors* **2023**, *23*, 2750. <https://doi.org/10.3390/s23052750>

Academic Editor: Vassilis S. Kodogiannis

Received: 30 January 2023

Revised: 19 February 2023

Accepted: 27 February 2023

Published: 2 March 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

As an exercise in dynamic simulation, Motor Imagery (MI) involves practicing the representation of a given motor movement in working memory without any overt motor activity. Simulation of MI mechanisms supervised and supported by cognitive systems activate brain neural systems that improve motor learning ability and neural plasticity, having a high potential for numerous medical and professional applications [1–5]. It has been identified as a promising method for enhancing motor proficiency, with significant educational implications. Furthermore, it has also demonstrated potential for rehabilitating children with neurological disorders [3,6]. Due to the fact that the Media and Information

Literacy methodology proposed by UNESCO covers several cognitive competencies, this aspect becomes essential, as highlighted in [7]. Nevertheless, accurate decoding of brain neural responses elicited by MI tasks poses a challenge since spectral/spatial/temporal features reflect the ease/difficulty subjects (MI ability) face when adhering to the mental imagery paradigm. In this scenario, individual motor skills seriously impact the MI system's execution.

In practice, MI systems have become the most common application in Brain-Computer Interface (BCI) research to enable communication and control over computer applications and external devices directly from brain activity [8]. BCI systems are typically designed to use Electroencephalogram (EEG) sensors due to their non-invasiveness, high time resolution, and relatively low cost [9]. However, despite these advantages, EEG recordings from electrodes have several limitations, including non-stationarity, a low signal-to-noise ratio (SNR), and poor spatial resolution, among other reported challenges [10]. Thus, effective control of MI-BCI depends on a careful balance between user skills and EEG-montage constraints [11]. Unfortunately, only a small number of individuals possess the necessary skills to control their BCIs effectively. Moreover, an estimated 15–30% of people lack the skills needed to complete MI tasks accurately, resulting in poor performance (known as BCI-inefficiency) [12]. As a result, the application of the imagery paradigm is often limited to laboratory settings, as the EEG-based signal analysis and implementation of BCIs with acceptable reliability have become increasingly complex procedures.

A growing body of research has focused on developing strategies to address BCI-inefficiency, with several promising approaches emerging. One such approach is to adapt the MI BCI system to specific frameworks of application, as demonstrated by recent studies [13–16]. In addition, these studies have explored various ways to optimize the performance of MI, including by customizing the system's parameters to suit the user's needs and abilities. Another promising strategy involves developing personalized systems that account for individual cognitive variability [17,18]. Considering factors such as attentional capacity, memory, and learning styles, these personalized systems can help improve the effectiveness of MI-BCI control, even for those with limited motor skills. Likewise, identifying subjects with a poor motor performance at the early stages of BCI training is another essential strategy for improving MI efficiency [19]. By identifying those who may struggle with BCI control early on, interventions can be implemented to help these individuals develop the necessary skills and optimize their performance.

In turn, determining the lack of motor skills at an early stage of EEG BCI system training implies assessing and interpreting the neural responses elicited by MI across the considered subject set. Yet, MI responses may depend on factors like brain structure variability, diversity of cognitive strategy, and earned individual expertise [12]. A widely-used approach to deal with EEG variability is developing BCI processing pipelines with optimized spatial pattern transformations that may account for conventional approaches [20,21]. Deep learning (DL) methods, especially Convolutional Neural Networks (CNN), have been investigated to optimize signal analysis since they automatically extract dynamic data streams from neural responses to address the complexity of EEG data associated with MI tasks and the variability of users [22–24]. Nevertheless, DL models generate high-level abstract features processed by black boxes with meaningless neuro-physiological explainability, resulting in a severe disadvantage in ensuring adequate reliability compared to their high performance [25].

Compared to other deep learning models, CNN-based classifiers offer several key advantages, including the ability to learn end-to-end from raw data and requiring fewer tuning parameters. Besides, end-to-end learning allows for a more streamlined and automated process, reducing the need for manual feature engineering and enabling the model to learn features directly from the data [26–29]. Moreover, visualization of feature maps from convolutional layers or class-discriminative features computed by the Class Activation Maps (CAM) is among the most used technique for a post-hoc explanation of well-trained CNN learning models [30–32]. Generally, different approaches to visualizing

the learned CNN filters fall into two classes: extracting trained weights at a selected layer and from reduced low-dimensional representations weighted across the layers, like developed for Motor Imagery in [33,34]. To better encode the diversity of stochastic behaviors, the reduced feature representations are clustered, separating the individuals into groups according to selected metrics of MI BCI performance, like developed in [35]. Despite this, the widespread use of DL models is prevented in functional neuroimaging because of the difficulty of training on high-dimensional low-sample-size datasets and the unclear relevance of the resulting predictive markers [36]. Therefore, CNN-based EEG decoding should be improved by verifying the learned weights' discriminating capability and relationship to neurophysiological features, as stated in [37].

This work presents a CNN framework to learn MI patterns from high-dimensional EEG dynamical data. The latter is achieved while ensuring that the interpretability of the neural responses is maintained through the use of connectivity features extracted from class activation maps. Also, to address the issue of inter/intra-subject variability in MI-EEG data, our work adopts two different approaches. The first approach involves the extraction of functional connectivity from spatiotemporal CAMs. Notably, we coupled the well-known EEGnet architecture and a novel kernel-based cross-spectral distribution estimator founded on a Gaussian Functional Connectivity (GFC) measure. This approach helps to overcome the limitations of traditional connectivity measures by capturing both the linear and non-linear relationships between the different brain regions. The second strategy involves clustering the subjects based on their achieved classifier accuracy and relevant channel-based connectivities. This clustering aims to identify common and discriminative patterns of motor skills. By doing so, the proposed approach can better account for the inter-subject variability often encountered in MI-EEG data. The validation results from a bi-class database demonstrate that the proposed approach effectively enhances the physiological explanation of brain neural responses. This is particularly true for subjects with poor MI skills and those with high variability and deficient EEG BCI performance. Overall, the proposed method represents a significant contribution to the field of neural response interpretation, and it has the potential to facilitate the development of more effective and accurate BCI systems. Of note, the present work continues our framework in [38] by eliminating the characterization stage to make the developed deep model take advantage of big data. A more complex database with more subjects is then used to evaluate the resulting end-to-end model. Furthermore, we introduce a CNN framework to learn from high-dimensional dynamical data while maintaining post-hoc interpretability using connectivity features extracted from class activation maps.

The agenda is as follows: Section 2 defines the used EEG-Net-based Classification framework of MI tasks, the Score-Weighted Visual Class Activation Maps estimated from EEG-Net, the GFC fundamentals, and the two-sample Kolmogorov-Smirnov test used to ensure the significance of extracted functional connectivity measures. Section 3 describes the validated EEG data set and the parameter setting of the CNN training framework. Section 4 explains the classification results of CAM-based EEGnet Mask, clustering of Motor Imagery Neural Responses using Individual GFC measures, and the Enhanced Interpretability from GFC patterns according to the clusterized motor skills. Lastly, Section 5 gives critical insights into their supplied performance and addresses some limitations and possibilities of the presented approach.

## 2. Materials and Methods

We present the fundamentals of EEGnet-based discrimination and the approach of CAM using a GFC-based EEG representation to improve the posthoc DL interpretability of elicited neural responses. In addition, we present a two-sample Kolmogorov-Smirnov test used to prune the GFC-based channel relationships matrix that is performed to enhance individuals' grouping according to their MI skills.

### 2.1. EEGnet-Based Classification of Motor Imagery Tasks

This convolutional neural network architecture includes a stack of spatial and temporal layers,  $\{\mathbf{W}_l \in \mathbb{R}^{P_{l-1} \times P_l}, \mathbf{b}_l \in \mathbb{R}^{P_l}: l \in L\}$ , fed by the input-output set,  $\{\mathbf{X}_n \in \mathbb{R}^{C \times T}, \boldsymbol{\lambda}_n \in \{0, 1\}^K: n \in N\}$ . Here,  $P_l \in \mathbb{N}$  denotes the  $l$ -th layer unit set,  $L \in \mathbb{N}$  is the network depth,  $C \in \mathbb{N}$  is the number of EEG channels,  $T \in \mathbb{N}$  is the recording length,  $N \in \mathbb{N}$  is the number of trials captured for each of  $K \in \mathbb{N}$  MI classes of MI tasks.

The EEGnet model, noted as  $\mathcal{M}(\mathbf{X}) : \mathbb{R}^{C \times T} \rightarrow [0, 1]^K$ , performs prediction  $\hat{\boldsymbol{\lambda}} \in [0, 1]^K$  of each one-hot class membership,  $\boldsymbol{\lambda}$ , as below:

$$\mathcal{M}(\mathbf{X}) = (\tilde{\mathbf{X}}_L \circ \dots \circ \tilde{\mathbf{X}}_1)(\mathbf{X}) \rightarrow \hat{\boldsymbol{\lambda}} \quad (1)$$

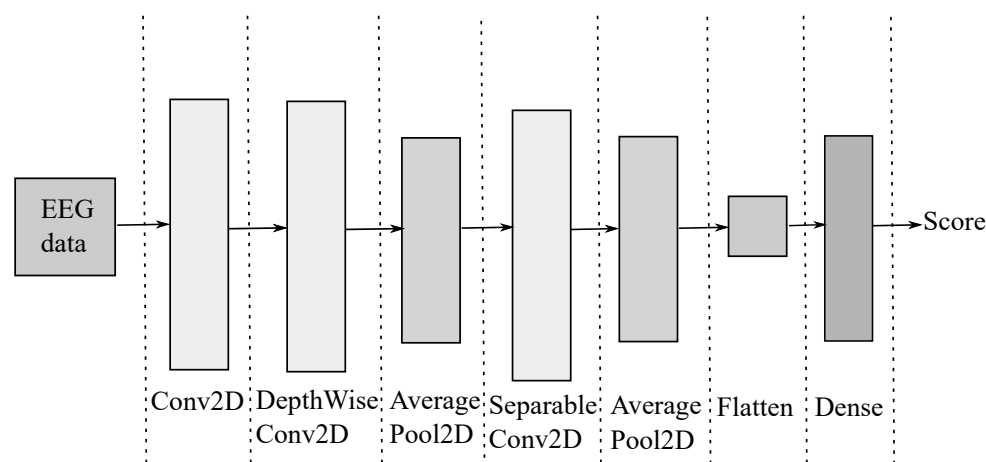
where  $\tilde{\mathbf{X}}_l = \phi_l(\tilde{\mathbf{X}}_{l-1} \otimes \mathbf{W}_l + \mathbf{b}_l) \in \mathbb{R}^{P_l}$  is a feature map holding  $P_l \in \mathbb{N}$  elements at  $l$ -th layer,  $\zeta_l: \mathbb{R}^{P_{l-1}} \rightarrow \mathbb{R}^{P_l}$  is a representation learning function,  $\mathbf{b}_l \in \mathbb{R}^{P_l}$  is a bias term, and  $\phi_l(\cdot)$  is a non-linear activation that is fixed as a sigmoid and softmax function for the bi-class and multiclass scenarios, respectively. Notations  $\circ$  and  $\otimes$  stand for function composition and proper tensor operation (i.e., 1D/2D convolution or fully connected-based product), respectively. Of note, we set  $\tilde{\mathbf{X}}_0 = \mathbf{X}$ ,  $\tilde{\mathbf{X}}_L = \hat{\boldsymbol{\lambda}}$  in the iterative computation of feature maps.

The solution in (1) implies the parameter set computation of  $\Theta = \{\mathbf{W}_l, \mathbf{b}_l: l \in L\}$ , performed within the following minimizing framework:

$$\Theta^* = \arg \min_{\Theta} \mathbb{E}\{\mathcal{L}(\boldsymbol{\lambda}_n, \hat{\boldsymbol{\lambda}}_n | \Theta) + \gamma \Omega(\Theta) : \forall n \in N\}, \quad (2)$$

where  $\mathcal{L} : \{0, 1\}^K \times [0, 1]^K \rightarrow \mathbb{R}$  is a given loss function,  $\Omega(\cdot)$  is a regularization function,  $\gamma \in \mathbb{R}^+$  is a trade-off, while  $\mathbb{E}\{\cdot\}$  stands for the expectation operator.

In order to support EEG classification, the compact EEGNet convolutional network minimizes (2), providing convolutional kernel connectivity between inputs and output feature maps that can apply to different MI paradigms. As seen in Figure 1, the EEGNet pipeline begins with a temporal convolution to learn frequency filters, followed by a depth-wise convolution connected to each feature map to learn frequency-specific spatial filters. Further, the separable convolution combines a depthwise convolution, which learns a temporal summary for each feature map individually, followed by a point-wise convolution, which learns how to mix the feature maps optimally for class-membership prediction.



**Figure 1.** EEGnet main sketch. First column: input EEG. Second column: temporal convolution (filter bank). Third and fourth columns: spatial filtering convolution. Fifth and sixth columns: temporal summary. Last column: output label prediction.

### 2.2. Score-Weighted Visual Class Activation Maps from EEG-Net

To highlight the most relevant input features, we conduct a CAM-based relevance analysis over the EEG-Net pipeline to favor the post-hoc interpretability of those relevant

spatio-temporal brain patterns that contribute the most to discriminating the elicited MI neural responses. To this end, we employ the Score-Weighted Visual Explanation approach, termed Score-CAM, that eliminates the dependence on gradient-based interpretability methods through the EEGNet network forward passing score on target class and a further linear combination of relevance weights. Thus, we extract the upsampled class-specific CAM matrix,  $A_l(\hat{\lambda}_k) \in \mathbb{R}^{C \times T}$ , that regards  $k$ -th estimated label,  $\hat{\lambda}_k \in [0, 1]$ , as follows [39]:

$$A_l(\hat{\lambda}_k) = \bar{\mu} \left( \sum_{\forall d \in D_l} \alpha_l^d(\hat{\lambda}_k) \tilde{X}_l^d \right), \quad \forall k \in K \quad (3)$$

where  $\bar{\mu}(\cdot)$  is the resulting function composition between the upsampling function  $\mu(\cdot)$  and the activation function in the form  $\text{ReLU}\{\eta\} = \max(0, \eta) \forall \eta \in \mathbb{R}$ , the matrix  $\tilde{X}_l^d \in \mathbb{R}^{P_l}$  computed on the trial basis, sizing  $P_l = C_l \times T_l$ , holds  $d$ -th feature map at  $l$  layer,  $D_l \in \mathbb{N}$  is the number of filters implementing the 1D/2D convolutional operations by the  $\mathcal{M}(\cdot)$  architecture, and  $\alpha_l^d(\hat{\lambda}_k) \in [0, 1]$  is a combination weight that is computed as follows:

$$\alpha_l^d(\hat{\lambda}_k) = \frac{\tilde{\alpha}_l^d(\hat{\lambda}_k)}{\sum_{d=1}^{D_l} \tilde{\alpha}_l^d(\hat{\lambda}_k)}, \quad (4)$$

where the estimate  $\tilde{\alpha}_l^d(\hat{\lambda}_k) \in \mathbb{R}^+$  is computed in a post-hoc way (namely, after the EEG-Net training) by  $k$ -th class prediction  $\mathcal{M}(\cdot|k)$  as  $\tilde{\alpha}_l^d(\hat{\lambda}_k) = \mathcal{M}(\mathbf{X} \odot \mu(\tilde{X}_l^d)|k) - \mathcal{M}(\mathbf{X}|k)$ , where notation  $\odot$  stands for the Hadamard product. Note that the Score-CAM representation is constrained by ReLU-based thresholding in (3) to assemble positive definite values of relevance into  $A_l(\hat{\lambda}_k)$ , implying discriminant EEG inputs that increase the output neuron's activation rather than suppressing behaviors [40].

Consequently, the weights estimated in Equation (4) support the CAM-based approach to boost Spatio-temporal EEG patterns, represented by the element-wise feature maps (or masks), which are salient in terms of discriminating between MI labels.

### 2.3. Pruned Gaussian Functional Connectivity from Score-CAM

Consider two Score-CAM-based EEG records  $\mathbf{a}, \mathbf{a}' \in A$  for a given EEGNet layer  $l$  and predicted class  $\hat{\lambda}_k$ . Their correlation can be expressed as a generalized, stationary kernel,  $\kappa: \mathbb{R}^T \times \mathbb{R}^T \rightarrow \mathbb{R}$ , if their spectral representations meet the following assumption [41]:

$$\kappa(\mathbf{a} - \mathbf{a}') = \int_{\tilde{\mathbf{f}} \in \Xi} \exp(j2\pi(\mathbf{a} - \mathbf{a}')^\top \tilde{\mathbf{f}}) \tilde{S}_{aa'}(\tilde{\mathbf{f}}) d\tilde{\mathbf{f}}, \quad (5)$$

where  $\tilde{\mathbf{f}} \in \Xi$  is a vector in the spectral domain and  $\tilde{S}_{aa'}(\tilde{\mathbf{f}}) \in \mathbb{C}$  is the cross-spectral density function with  $\tilde{S}_{aa'}(\tilde{\mathbf{f}}) = dP_{aa'}(\tilde{\mathbf{f}})/d\tilde{\mathbf{f}}$ , and  $P_{aa'}(\tilde{\mathbf{f}}) \in [0, 1]$  is the cross-spectral distribution. The cross-spectral distribution between CAM-based records in Equation (5) can be computed as follows [42]:

$$P_{aa'}(\Xi) = 2 \int_{\tilde{\mathbf{f}} \in \Xi} \mathcal{F}\{\kappa(\mathbf{a} - \mathbf{a}')\} d\tilde{\mathbf{f}}. \quad (6)$$

$\mathcal{F}\{\cdot\}$  stands for the Fourier transform. Then, nonlinear interactions between CAM-based boosted EEG channels are coded for a more accurate depiction of neural activity.

Of note, a stationary kernel is employed to preserve temporal dynamics of EEG signals, essential for MI classification accuracy. Consequently, the Gaussian kernel is favored in pattern analysis and machine learning for its versatility in approximating functions and its mathematical tractability [43]. These properties make it a prime candidate for calculating the Kernel-based Cross-Spectral Functional Connectivity in Equation (6), fixing the pairwise relationships between  $\mathbf{a}$  and  $\mathbf{a}'$  as:



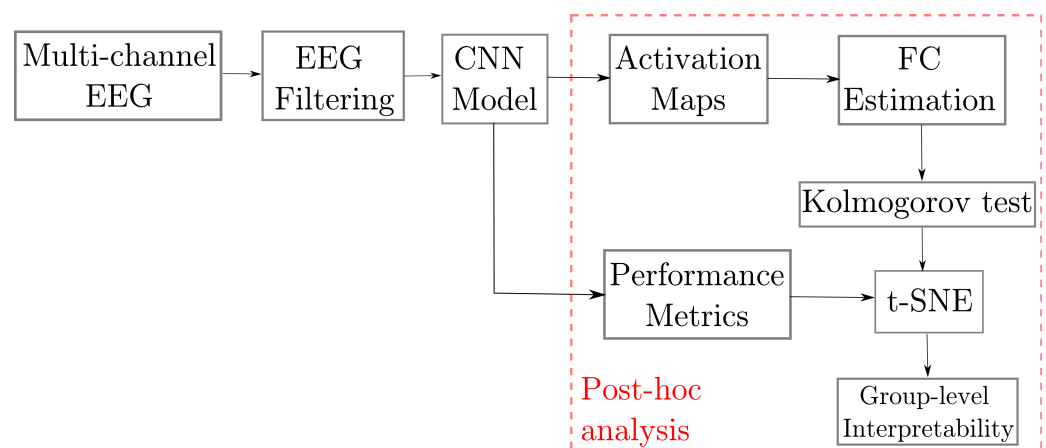
$$\kappa_G(\mathbf{a} - \mathbf{a}'; \sigma) = \exp\left(\frac{-\|\mathbf{a} - \mathbf{a}'\|_2^2}{2\sigma^2}\right), \quad (7)$$

where  $\|\cdot\|_2$  is the L2 distance and  $\sigma \in \mathbb{R}^+$  is a given bandwidth hyperparameter, commonly fixed based on the median of the input distances.

Finally, to prune the Gaussian Functional Connectivity (GFC), as in Equations (6) and (7), we need to determine which connections are crucial for class separation. A high correlation in the functional connectivity matrix does not necessarily lead to higher class separation, so we use the two-sample Kolmogorov–Smirnov (2KS) test, as in [44], to address this. The null hypothesis is that both samples come from the same distribution. In turn, we group trials for each CAM GFC connection for a subject by label to form samples representing class membership. Each pair is then tested with the 2KS test and connections with  $p$ -values  $\leq 0.05$  are kept, implying the samples are from different distributions and the classes are separable. Finally, a  $p$ -value matrix indicating relevant connections is generated.

#### 2.4. Post-Hoc Grouping of Subject Motor Imagery Skills

We evaluate the proposed approach for improving the post-hoc interpretability of the Deep Learning-based model using the Class Activation Maps and the pruned Gaussian Functional Connectivity to characterize the inter/intra-subject variability of the brain signals. This purpose is addressed within the evaluation pipeline shown in Figure 2, including the following stages: (i) Preprocessing of raw EEG data to be fed the DL model; (ii) Bi-class discrimination of motor-evoked tasks within a shallow ConvNets framework (EEGNet); (iii) In the post-hoc analysis, we compute the CAMs over the Conv2D layer, followed by a clustering analysis performed on a reduced set of extracted features, using the well-known t-distributed Stochastic Neighbor Embedding (t-SNE), to enhance the explainability of produced neural responses regarding motor skills among subjects. Also, pruned GFC is used for visual inspection results and clustering enhancement.

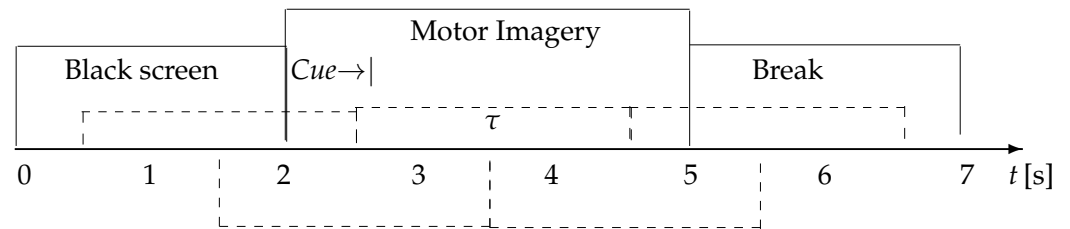


**Figure 2.** Guideline of the proposed framework for enhanced post-hoc interpretability of MI neural responses using connectivity measures extracted from EEGNet CAMs and clustering subjects, according to the EEG MI performance.

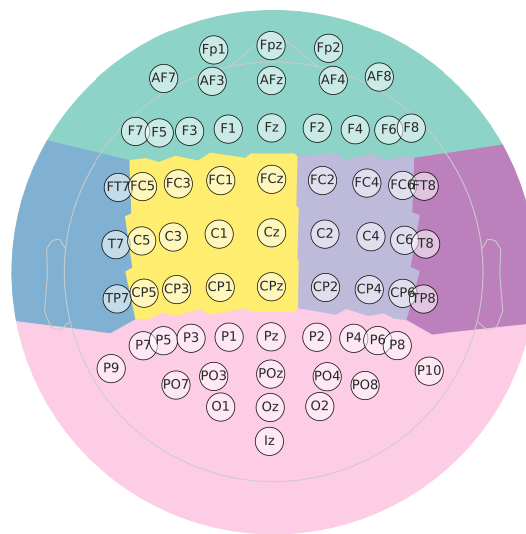
#### 2.5. GigaScience Database

In this collection, available at <http://gigadb.org/dataset/100295>, accessed on 1 October 2022, EEG data from fifty (50) subjects have been assembled according to the experimental paradigm for MI, as shown in Figure 3. The MI task execution starts with a fixation cross appearing on a black screen within 2 s. Afterward, for 3 s, a cue instruction appeared on the screen, requesting that each subject imagine the finger movements (initially, with the forefinger and then the small finger touching their thumb). Then, a blank screen appeared at the start of the break, lasting randomly for a short period between 4.1 to 4.8 s. These procedures were repeated 20 times for each MI class in a single testing run, collecting

100 trials per individual (each one lasting  $T = 7$  s) for either labeled task (left hand or right hand). Data were acquired at 512 Hz sampling rate by a 10-10 placement using an C-electrode montage ( $C = 64$ ), see Figure 4.



**Figure 3.** GigaScience database timeline of the evaluated motor imagery paradigm.



**Figure 4.** Topographic map for EEG representation. Besides, it highlights in color the main parts of the brain (Frontal, Central right, Posterior right, Posterior, Posterior left, Central left).

### 3. Experimental Set-Up

#### 3.1. Parameter Setting of Trained CNN Framework

In the preprocessing stage, each raw channel is passed through a bandpass filter within [4–40] Hz using a five-order Butterworth filter, as suggested in [45]. To enhance the physiological interpretation of implemented experimental paradigm, we analyze the dynamic at the representative time segment from 2.5 s to 5 s (i.e., motor imagery interval). Further, the EEG signal set is standardized across the channels with the z-score approach.

Afterward, the resulting preprocessed EEG multi-channel data feeds the EEGNet framework with a shallow ConvNet architecture. The CNN model is implemented by downsampling all cropped EEG recordings to 128 for feeding the EEGNet model, fixing the parameter set to the values as displayed in Table 1. The CNN learner includes a categorical cross-entropy loss function and the Adam optimizer.

Table 1 brings a detailed description of the employed network architecture for which the following values are fixed: The number of band-pass feature sets extracted from EEG data by a set of 2D convolutional filters  $F_1 = 16$ ; 2D CNN filter length is 128 Hz (i.e., half of the sampling rate); and the number of temporal filters along with their associated spatial filters is adjusted to  $F_2 = DF_1$  in SeparableConv2D. Besides, validation is performed by running 100 training iterations (epochs for performing validation stopping), storing the model weights that reach the lowest loss. The model validation is conducted on a GPU device at a Google Collaboratory session, using TensorFlow and the Keras API.

**Table 1.** EEGNet architecture. Input shape is  $(1, C, T)$ , where  $C$  is number of channels,  $T$ —number of time points,  $F_1$ —number of temporal filters,  $D$ —depth multiplier (number of spatial filters),  $F_2$ —number of point-wise filters, and  $N$ —number of classes, respectively. For the Dropout layer, we use  $p = 0.5$  for within-subject classification. Notations \* stand for  $16 * D * F_1 + F_2 * (D * F_1)$  and (\*\*) for  $N * (F_2 * T // 32)$ .

Layer Name	Conv2D	Depthwise Conv2D	Separable Conv2D	Flatten	Dense
# filters	$F_1$	$D * F_1$	$F_2$		(**)
Size	$(1, 64)$	$(C, 1)$	$(1, 16)$		
# params	$64 * F_1$	$C * D * F_1$	(*)		
Output	$(F_1, C, T)$	$(D * F_1, 1, T)$	$(F_2, 1, T // 4)$	$(F_2 * (T // 32))$	N
Options	Activation = Linear Mode = same	Activation = Linear Mode = same Depth = $D$ max_norm = 1	Activation = Linear Mode = same		
	BatchNorm = True	BatchNorm = True Activation = ELU AvgPool2D = $(1, 4)$ Dropout* - $p = 0.25$ or $p = 0.5$	BatchNorm = True Activation = ELU AvgPool2D = $(1, 8)$ Dropout* - $p = 0.25$ or $p = 0.5$		

The final step was to perform a group analysis of individuals using the K-means approach for clustering the performance metrics and Kolmogorov test values extracted from connectivity measures. Note that the feature set is z-scored to feed the partition algorithm. Besides, we apply the well-known t-distributed Stochastic Neighbor Embedding (t-SNE) for dimensionality reduction [46,47], selecting the first two components.

### 3.2. Quality Assessment

The proposed methodology is evaluated in terms of the EEGNet binary classification estimated through the cross-validation strategy, which consists of prearranging two sets of points: 80% for training and the remaining 20% for hold-out. The following metrics assess the classifier performance [47]: accuracy, kappa value, f1-score, precision, and recall. Besides, a Nemenyi post-hoc statistical test is performed at a significance level of 0.05 [48]. The latter aims to compare subject-dependent EEGNet baseline performance vs. our CAM-based enhancement.

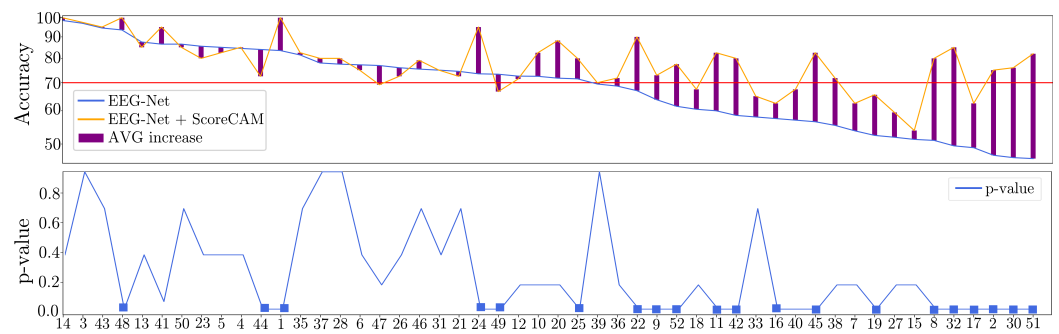
## 4. Results and Discussion

### 4.1. Classification Results of CAM-Based EEGnet Masks

Figure 5 displays the classification accuracy results obtained by the EEGnet model (colored with a blue line) with parameters tuned as described above. In order to better understand the values, we rank them in decreasing order of mean accuracy to reveal that more than half of them fall below the 70% level (horizontal red line) and may need more MI skills. Otherwise, a system with so many poor-performing individuals makes MI training ineffective.

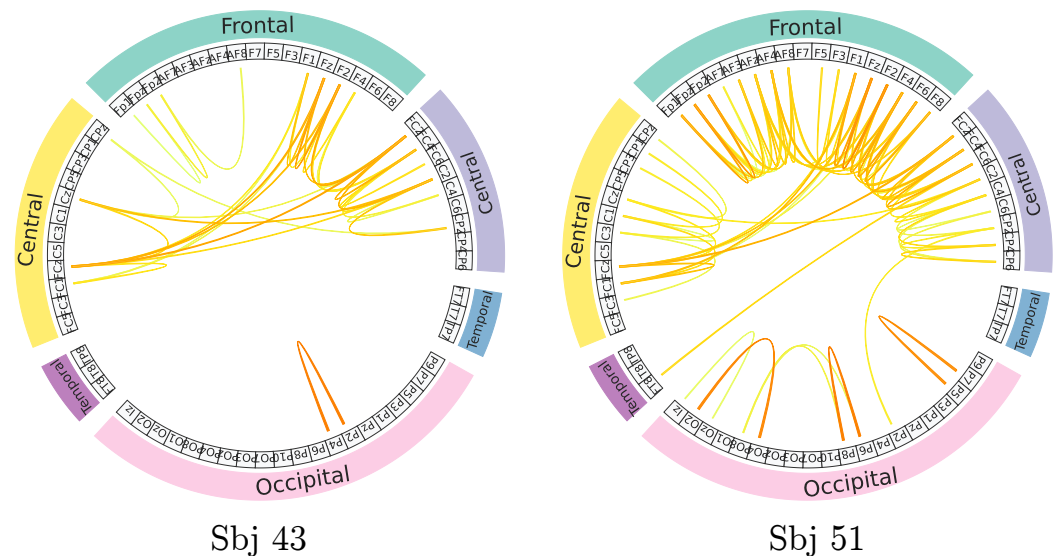
Results for individual masks are also displayed (colored with an orange line) and suggest the improved performance of the classifier. The EEGNet approach enhanced by CAM-based representations is thus more effective at dealing with hardly discernible neural responses, reducing the number of subjects labeled as inadequate in MI from 21 (~40%) to 11 (~20%). Furthermore, concerning the Nemenyi post-hoc statistical test, the results demonstrate a significant difference between our CAM-based enhancement and the benchmark EEGNet discrimination approach. Specifically, this difference was most pronounced in subjects with lower performance levels. These findings highlight the potential of our proposed approach to improve the accuracy and reliability of EEG-based analyses, particularly for individuals with poorer MI skills.





**Figure 5.** Subject-dependent MI discrimination results. (**Up**): Classification accuracy achieved for Bi-class MI tasks (left-hand and right-hand). Note: the red line at 70% level shows frequently used for fixing the poor MI coordination skill threshold under the subjects are considered as worse-performing. (**Bottom**): Obtained  $p$ -value per subject using the Nemenyi post-hoc test (square marker represents a  $p$ -value  $< 0.05$ ).

A mapping of GFC connectivity based on CAM-based masks is also shown to understand individual neural responses better. For illustration purposes, the topoplot representations extracted from GFC connectivity maps are contrasted for the most successful and poorest performing subjects (#43 and #51, respectively). The displayed topoplot of each subject in Figure 6 shows high response amplitudes distributed evenly over the Frontal and central areas, as expected for the neural responses elicited by MI paradigms [49]. There are, however, substantial differences between both subjects in the assessed relevant GFC links.



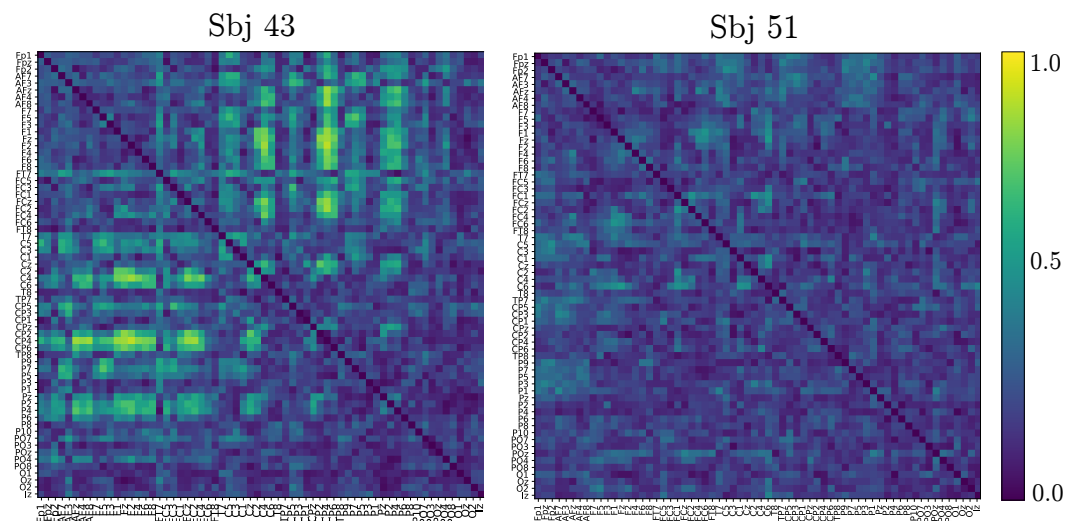
**Figure 6.** Connectivity maps are estimated for the best-performing (left side) and poorest-performing (right side) subjects. Stronger correlations between nodes are represented by darker edges linking two EEG channels. The right plot shows the brain regions (i.e., Temporal, Frontal, Occipital, Parietal, and Central) colored differently to improve spatial interpretation.

A notable finding is that the neural responses in the best-performing individual produced much lower discriminant links, primarily concentrated in the Sensory-Motor region. At the same time, the discriminative links of #43 are more related to the contralateral hemisphere. There may be an explanation for this effect if the skilled subject is more focused on performing tasks based on the MI paradigm.

By contrast, neural connections throughout the brain in the subject with poor performance are twice as high as in the skilled individual. Furthermore, the relevant links of #51 include associations between occipital and frontal areas in the cerebral cortex with channels not primarily related to the imagination of motor activities, suggesting a spu-

rious electrophysiological mechanism during MI performance. Because of these factors, poor-performing subjects tend to be less accurate in distinguishing between MI tasks [50].

Based on the above information, we calculate the connectivity matrix after applying a two-sample Kolmogorov-Smirnov test, as carried out in [51] for the goodness of fit. As can be observed in Figure 7, the best-performing subject produces most of the statistically significant connections all above the sensory motor region following the MI paradigm. In contrast, the poorest-performing subject holds low statistically significant links, making the activity over the motor cortex barely distinguishable from the remaining brain areas.

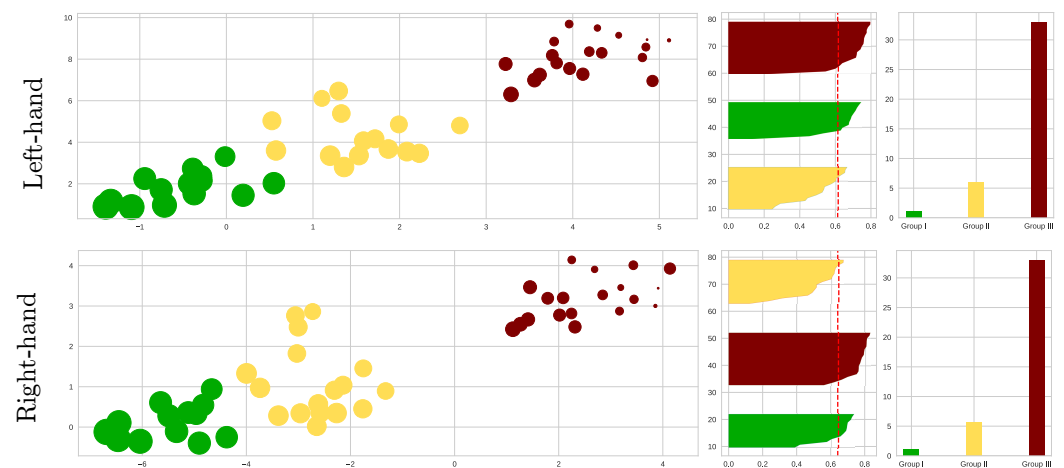


**Figure 7.** Connectivity matrix after two-sample Kolmogorov-Smirnov test obtained by the best (left side) and poorest-performing (right side) subjects. The pictured GFC matrices include both MI tasks and are computed at 90-percentile of normalized relevance weights.

#### 4.2. Clustering of Motor Imagery Neural Responses Using Individual GFC Measures

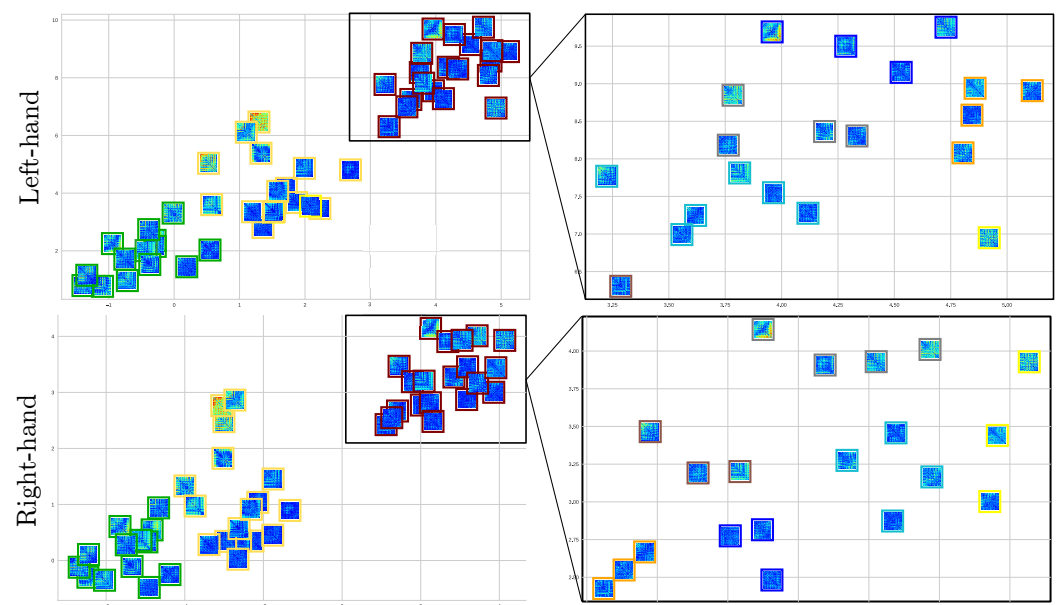
As inputs to the partition algorithm, we include several classifier performance metrics since we assume that the more distinguishable the elicited neural responses between each labeled task, the better the skills of a subject in performing the MI paradigm. In addition to the EEG-Net classifier metrics, the Kolmogorov test values are also evaluated and assessed for the Gaussian FC measures of the individual ScoreCAMs. For encoding the stochastic behavior of metrics, the clustering algorithm is fed by their reduced space representation using the t-SNE projection, as described in [42].

As shown in the left scatter plots of Figure 8, we performed the  $k$ -means algorithm on the reduced feature set for each MI task, separating the individuals into three groups according to their motor skills: The best-performing subjects (termed Group I and colored in green); The partition with intermediate MI abilities (Group II in yellow); Group III (red color) with the worst-performing individuals. We separately analyze both Motor Imagery Tasks due to Hand Dominance influence, as discussed in [52]. The grouping quality is assessed by the Silhouette coefficient displayed in the center plots, showing differences in the partition density: the more inadequate the motor skills, the more spread the partition. A further aspect of evaluation is the effect of CAM masks on enhanced classifier performance, which increases as the individual's MI skills decrease (see right-sidebar plots).



**Figure 8.** K-means clustering visualization of the EEG-Net classifier performance metrics and the Kolmogorov test values extracted from connectivity information. Left-side scatter plots of individuals (data points) estimated for the three considered groups of motor skills. Note that the point size represents the subject's accuracy. The central plot shows the clustering metrics of the Silhouette coefficient, while the right-side plot displays the classifier performance, averaged across all subjects of each partition.

By coloring the squares into three clusters, Figure 9 enhances the explanation of the extracted connectivity measures with statistical significance and shows that the partition density of best-performing subjects (green square) by both MI tasks is alike and well clustered, reaching high silhouette values. For GII, three subjects are outliers, making this fact more evident for the data point labeled as #20, especially in the right-hand task. This result may be explained since this subject holds very few trials.



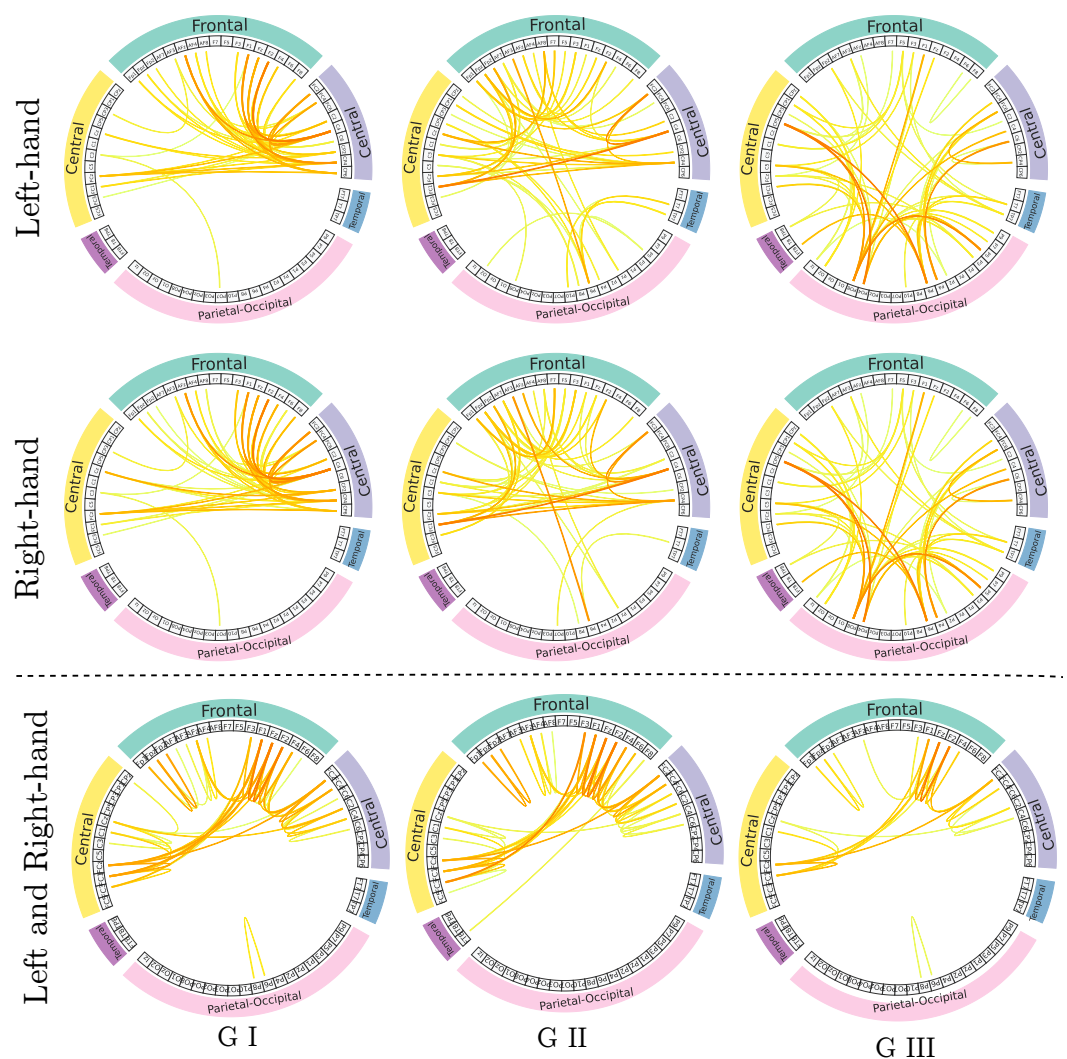
**Figure 9.** Clustering of motor skills obtained by the CNN-based approach with connectivity features for each MI task. (**Left side**): Subject partitions are colored for each subject group: GI (green squares), GII (yellow), and GIII (red). (**Right side**): Detailed analysis of the worst-performing GIII by splitting all the participants into subgroups, fixing  $k = 6$ .

Lastly, there is a tendency for the GIII cluster to fragment into smaller partitions. This issue is clarified by recalculating the clustering algorithm and exploring the number of subgroups within  $k = \{3, 4, 5, 6\}$ . According to the right-side column, fixing  $k = 3$  provides the highest silhouette value, being more evident in the scatter plot estimated for

the right-hand task. Apparently, the MI paradigm is performed quite distinctly by each GIII individual, thus worsening the classifier accuracy.

#### 4.3. Enhanced Interpretability from GFC Patterns According to Clusterized Motor Skills

To bring a physiological foundation to clustered motor skills, we analyzed the brain neural responses elicited by MI paradigms through the GFC patterns extracted from the Score-based CAMs and averaged across each subject partition. As explained above, we compute the pairwise functional connectivity graph that contributes the most to discriminating between the labeled MI-related tasks, selecting electrode links that fall within the 90-percentile of normalized relevance weight calculations, as shown in Figure 10. However, two analysis scenarios are considered: The statistically significant connections computed by the Komolgorov-Smirnov test (see Figure 7) and the connectivity set without passing under the hypothesis test.



**Figure 10.** Graphical representation of the brain Gaussian Functional Connectivity estimated across each cluster of individual motor skills. Connectograms of each MI task (**top and middle rows**) are computed without the Komolgorov-Smirnov test and conducting the hypothesis test (**bottom row**). The GFC set is calculated at the 90 percentile of normalized relevance weights.

The top and middle rows in the former scenario display the connectograms computed for each MI task, respectively. As can be seen, each Group of individual skills triggers different sets of connections in terms of nodes and relevant links, which may lead to specific network integration and segregation changes, affecting the system's accuracy. Accordingly, the connectivity graphs estimated for the Group of the best-skilled subjects hold high evoked response amplitudes densely spread over the Frontal, Central, and Sensory-Motor areas involved in motor imagery. At the same time, GII yields a connectivity set similar to GI, with most links spreading over the primary motor cortex. However, several EEG electrodes generate discriminating links over the temporal and parietal-occipital areas, which are assumed uncorrelated with MI responses. This situation worsens in the Group of poor-performing individuals with a link set covering the whole scalp surface, regardless of the MI task. Moreover, the most robust links are elicited over the parietal-occipital cortical areas, which are devoted mainly to visuospatial processes. This finding related to relevant connectivities outside the primary motor cortex may account for the strength of DL models for improving accuracy at the expense of searching for any discriminant patterns of neural activity. Patterns may be found even among spurious artifacts in the data acquisition procedure or related to irregularly elicited activations performed in the MI paradigm.

As shown in the bottom row, the connectivity maps for the second scenario are derived with Kolmogorov-Smirnov boosting masks, resulting in sets of discriminant links generated by electrodes just over the primary motor cortex, regardless of the partition of motor skills. It is worth noting that the connectograms achieved by GI are similar for either considered analysis scenario. The fact that the best-performing subjects can achieve very high accuracy even with simple classifiers is widely known. Nonetheless, the more the accuracy of subject partitions, the more robust links elicited by the MI responses.

## 5. Concluding Remarks

We introduce an approach to improving the post-hoc interpretability of neural responses using connectivity features extracted from class activation maps. The Spatiotemporal CAMs are computed using a CNN model and employed to deal with the inter/intra-subject variability for distinguishing between bi-class motor imagery tasks. Specifically, we extract the functional connectivity between EEG channels through a novel kernel-based cross-spectral distribution estimator to cluster subjects according to their achieved classifier accuracy, aiming to find common and discriminative patterns of motor skills. Based on validation results on a tested bi-class database, the proposed approach proves its usefulness for enhancing the physiological explanation of MI responses, even in cases of inferior performance. Although the approach may facilitate the increased use of neural response classifiers based on deep learning, the following aspects are to be highlighted for its implementation:

**Deep-Learning based Classifier Performance.** The deep model is implemented through EEGNet, a compact CNN specially developed for EEG-based brain-computer interfaces, which allows for reaching a bi-class accuracy of 69.7%. The additional use of proposed connectivity masks extracted from the electrode contribution increases the accuracy to 78.2% (EEGNet+ScoreCam) on average over the whole subject set. Table 2 delivers the accuracy of a few state-of-the-art approaches with similar discriminative deep learning models reported recently, indicating that the proposed spatiotemporal CAM approach is competitive. The works in [53,54] are higher accurate, though no interpretable machine-learning models are offered. By contrast, the CNN-based approach with a close value to our accuracy is suggested in [55], which performs data augmentation and employs Cropping methods for improving machine learning models, allowing plotting topographic heads for spatial interpretation.



**Table 2.** Comparison of bi-class accuracy achieved by state-of-the-art approaches in GigaScience collection. \* The baseline approach (FBCSP+LDA) using conventional pattern recognition methods is also reported for comparison. The proposed CNN-based approach introducing spatiotemporal ScoreCAMs is competitive (marked in bold).

Approach	Accuracy	Interpretability
FBCSP+LDA *	67.7 ± 13.6	–
LSTM+Optical [56]	68.2 ± 9.0	–
EKGnetv4+EA [57]	73.4	✓
DC1JNN [58]	76.50	✓
MINE+EEGnet [59]	76.6 ± 12.48	✓
Cropping+CNN [55]	86.5	✓
Pretrained ShuffleNet [53]	87.7	–
CSP+ST+CNN-PIM [54]	97.67 ± 2.0	–
EEGNet	69.7 ± 14.5	✓
EEGNet+ScoreCam (ours)	78.2 ± 10.9	✓

Therefore, a broader class of training procedures should be considered to improve the explainability of CNN-based models. To this extent, several strategies have been suggested recently for raising the classifier performance of CNN in discriminating MI tasks, like enhancing their convergence in [60,61], or extracting data at different time-resolutions [62]; this latter strategy may be a better choice for interpretation.

Enhanced interpretability using CNN-based Connectivity Features. This proposed approach uses functional connectivity with kernel-based cross-spectral distribution estimators to improve the computation of joint and discriminative patterns of EEG channel relationships. These correlations are essential in interpreting patterns of the primary motor area that become active and are elicited by motor imagery tasks. In this regard, Table 2 illustrates the existing balance between desirable high classifier performance and meaningful explainability of results to be achieved by deep learning models.

CAMs have previously been reported for CNN classifiers to discover shared EEG features across subject sets. However, we suggest extracting Functional Connectivity measures instead of using the trained weights (like in [33,63,64]) to identify brain regions contributing most to the classification of MI tasks. A convenient trade-off between interpretability and accuracy is thus achieved.

Clustering of subjects according to their motor skills. This model illustrates a potential approach to enhancing CNN classifiers to discover shared features related to EEG MI tasks across subject sets. The use of CNN models is also increasing when screening subjects with poor skills. In [24], for example, authors aim to detect users who are inefficient in using BCI systems. However, in addition to identifying poor performers, explaining why they cannot produce the desired sensorimotor patterns is essential. On the other hand, the strength of DL models lies in their ability to improve accuracy at the expense of finding any discriminant patterns of neural activity. As a result, DL models may estimate patterns unrelated to the primary motor cortex, i.e., they are responses with no physiological interpretation according to the MI paradigm.

Instead, the proposed method allows for determining the actual spatial locations of neural responses. Thus, we propose clustering motor skills into three partitions: Group I with high accuracy (good performing of MI paradigms and not subject to artifacts), GII with fair accuracy (adequate MI performing but subject to artifacts), and GIII with low accuracy (very variable MI performance of tasks and also subject to artifacts).

As future work, the authors plan to address the issue of high intra-class variability and overfitting, specifically in the worst-performing subjects, by incorporating more powerful functional connectivity estimators, for instance, using regularization techniques based on Rényi's entropy [65]. Besides, we plan to test more powerful DL frameworks, including attention mechanisms for decoding stochastic dynamics, as suggested in [66]. Further during-training analysis of the inter-class variability should also be conducted, specifically



focusing on how each subject performs within different experiment runs. Our goal is to determine if subjects learn MI tasks as they progress through the runs and whether this leads to better performance in subsequent runs. Validation of other MI EEG data will be carried out.

**Author Contributions:** Conceptualization, D.F.C.-H., A.M.Á.-M. and C.G.C.-D.; methodology, D.F.C.-H., D.A.C.-P. and C.G.C.-D.; software, D.F.C.-H. and A.M.Á.-M.; validation, D.F.C.-H., A.M.Á.-M. and D.A.C.-P.; formal analysis, A.M.Á.-M., G.A.C.-D. and C.G.C.-D.; investigation, D.F.C.-H., D.A.C.-P., A.M.Á.-M. and C.G.C.-D.; data curation, D.F.C.-H.; writing original draft preparation, D.F.C.-H., C.G.C.-D.; writing—review and editing, A.M.Á.-M., G.A.C.-D. and C.G.C.-D.; visualization, D.F.C.-H.; supervision, A.M.Á.-M. and C.G.C.-D. All authors have read and agreed to the published version of the manuscript.

**Funding:** Authors want to thank to the project “PROGRAMA DE INVESTIGACIÓN RECONSTRUCCIÓN DEL TEJIDO SOCIAL EN ZONAS DE POSCONFLICTO EN COLOMBIA Código SIGP: 57579” con el proyecto de investigación “Fortalecimiento docente desde la alfabetización mediática Informativa y la CTel, como estrategia didáctico-pedagógica y soporte para la recuperación de la confianza del tejido social afectado por el conflicto. Código SIGP 58950. Financiado en el marco de la convocatoria Colombia Científica, Contrato No. FP44842-213-2018”.

**Institutional Review Board Statement:** The experiment was approved by the Institutional Review Board of Gwangju Institute of Science and Technology.

**Informed Consent Statement:** All subjects gave written informed consent to collect information on brain signals.

**Data Availability Statement:** Publicly available datasets were analyzed in this study. This data can be found here: <http://gigadb.org/dataset/100295>, accessed on 1 October 2022.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Moran, A.P.; O’shea, H. Motor Imagery Practice and Cognitive Processes. *Front. Psychol.* **2020**, *11*, 394. [[CrossRef](#)] [[PubMed](#)]
2. Bahmani, M.; Babak, M.; Land, W.; Howard, J.; Diekfuss, J.; Abdollahipour, R. Children’s motor imagery modality dominance modulates the role of attentional focus in motor skill learning. *Hum. Mov. Sci.* **2021**, *75*, 102742. [[CrossRef](#)] [[PubMed](#)]
3. Behrendt, F.; Zumbrennen, V.; Brem, L.; Suica, Z.; Gäumann, S.; Ziller, C.; Gerth, U.; Schuster-Amft, C. Effect of motor imagery training on motor learning in children and adolescents: A systematic review and meta-analysis. *Int. J. Environ. Res. Public Health* **2021**, *18*, 9467. [[CrossRef](#)] [[PubMed](#)]
4. Said, R.R.; Heyat, M.B.B.; Song, K.; Tian, C.; Wu, Z. A Systematic Review of Virtual Reality and Robot Therapy as Recent Rehabilitation Technologies Using EEG-Brain-Computer Interface Based on Movement-Related Cortical Potentials. *Biosensors* **2022**, *12*, 1134. [[CrossRef](#)] [[PubMed](#)]
5. Alharbi, H. Identifying Thematics in a Brain-Computer Interface Research. *Comput. Intell. Neurosci.* **2023**, *2023*, 1–15. [[CrossRef](#)]
6. Souto, D.; Cruz, T.; Fontes, P.; Batista, R.; Haase, V. Motor Imagery Development in Children: Changes in Speed and Accuracy with Increasing Age. *Front. Pediatr.* **2020**, *8*, 100. [[CrossRef](#)]
7. Frau-Meigs, D. *Media Education. A Kit for Teachers, Students, Parents and Professionals*; Unesco: Paris, France, 2007.
8. Lyu, X.; Ding, P.; Li, S.; Dong, Y.; Su, L.; Zhao, L.; Gong, A.; Fu, Y. Human factors engineering of BCI: An evaluation for satisfaction of BCI based on motor imagery. *Cogn. Neurodynamics* **2022**, *17*, 1–14. [[CrossRef](#)]
9. Jeong, J.; Cho, J.; Lee, Y.; Lee, S.; Shin, G.; Kweon, Y.; Millán, J.; Müller, K.; Lee, S. 2020 International brain-computer interface competition: A review. *Front. Hum. Neurosci.* **2022**, *16*, 898300. [[CrossRef](#)]
10. Värbu, K.; Muhammad, N.; Muhammad, Y. Past, present, and future of EEG-based BCI applications. *Sensors* **2022**, *22*, 3331. [[CrossRef](#)]
11. Thompson, M.C. Critiquing the Concept of BCI Illiteracy. *Sci. Eng. Ethics* **2019**, *25*, 1217–1233. [[CrossRef](#)]
12. Becker, S.; Dhindsa, K.; Mousapour, L.; Al Dabagh, Y. BCI Illiteracy: It’s Us, Not Them. Optimizing BCIs for Individual Brains. In Proceedings of the 2022 10th International Winter Conference on Brain-Computer Interface (BCI), IEEE, Gangwon-do, Republic of Korea, 21–23 February 2022; pp. 1–3.

13. Vavoulis, A.; Figueiredo, P.; Vourvopoulos, A. A Review of Online Classification Performance in Motor Imagery-Based Brain–Computer Interfaces for Stroke Neurorehabilitation. *Signals* **2023**, *4*, 73–86. [[CrossRef](#)]
14. Shi, T.W.; Chang, G.M.; Qiang, J.F.; Ren, L.; Cui, W.H. Brain computer interface system based on monocular vision and motor imagery for UAV indoor space target searching. *Biomed. Signal Process. Control* **2023**, *79*, 104114. [[CrossRef](#)]
15. Naser, M.; Bhattacharya, S. Towards Practical BCI-Driven Wheelchairs: A Systematic Review Study. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2023**, *31*, 1030–1044. [[CrossRef](#)]
16. Tao, T.; Jia, Y.; Xu, G.; Liang, R.; Zhang, Q.; Chen, L.; Gao, Y.; Chen, R.; Zheng, X.; Yu, Y. Enhancement of motor imagery training efficiency by an online adaptive training paradigm integrated with error related potential. *J. Neural Eng.* **2023**, *20*, 016029. [[CrossRef](#)]
17. Tao, L.; Cao, T.; Wang, Q.; Liu, D.; Sun, J. Distribution Adaptation and Classification Framework Based on Multiple Kernel Learning for Motor Imagery BCI Illiteracy. *Sensors* **2022**, *22*, 6572. [[CrossRef](#)]
18. Ma, Y.; Gong, A.; Nan, W.; Ding, P.; Wang, F.; Fu, Y. Personalized Brain–Computer Interface and Its Applications. *J. Pers. Med.* **2023**, *13*, 46. [[CrossRef](#)]
19. Subasi, A. Artificial Intelligence in Brain Computer Interface. In Proceedings of the 2022 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), Ankara, Turkey, 9–11 June 2022; pp. 1–7.
20. Jiao, Y.; Zhou, T.; Yao, L.; Zhou, G.; Wang, X.; Zhang, Y. Multi-View Multi-Scale Optimization of Feature Representation for EEG Classification Improvement. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2020**, *28*, 2589–2597. [[CrossRef](#)]
21. Wang, T.; Du, S.; Dong, E. A novel method to reduce the motor imagery BCI illiteracy. *Med. Biol. Eng. Comput.* **2021**, *59*, 1–13. [[CrossRef](#)]
22. Mahmud, M.; Kaiser, M.; Hussain, A.; Vassanelli, S. Applications of Deep Learning and Reinforcement Learning to Biological Data. *IEEE Trans. Neural Netw. Learn. Syst.* **2018**, *29*, 2063–2079. [[CrossRef](#)]
23. Guragai, B.; AlShorman, O.; Masadeh, M.; Heyat, M.B.B. A survey on deep learning classification algorithms for motor imagery. In Proceedings of the 2020 32nd international conference on microelectronics (ICM), IEEE, Aqaba, Jordan, 14–17 December 2020; pp. 1–4.
24. Tibrewal, N.; Leeuwis, N.; Alimardani, M. Classification of motor imagery EEG using deep learning increases performance in inefficient BCI users. *PLoS ONE* **2022**, *17*, e0268880. [[CrossRef](#)]
25. Bang, J.; Lee, S. Interpretable Convolutional Neural Networks for Subject-Independent Motor Imagery Classification. In Proceedings of the 2022 10th International Winter Conference on Brain-Computer Interface (BCI), Gangwon-do, Republic of Korea, 21–23 February 2022; pp. 1–5.
26. Aggarwal, S.; Chugh, N. Signal processing techniques for motor imagery brain computer interface: A review. *Array* **2019**, *1*, 100003. [[CrossRef](#)]
27. Collazos-Huertas, D.; Álvarez-Meza, A.; Acosta-Medina, C.; Castaño-Duque, G.; Castellanos-Dominguez, G. CNN-based framework using spatial dropping for enhanced interpretation of neural activity in motor imagery classification. *Brain Inform.* **2020**, *7*, 8. [[CrossRef](#)] [[PubMed](#)]
28. Kim, S.; Lee, D.; Lee, S. Rethinking CNN Architecture for Enhancing Decoding Performance of Motor Imagery-Based EEG Signals. *IEEE Access* **2022**, *10*, 96984–96996. [[CrossRef](#)]
29. Garg, D.; Verma, G.; Singh, A. A review of Deep Learning based methods for Affect Analysis using Physiological Signals. *Multimed. Tools Appl.* **2023**, 1–46. [[CrossRef](#)]
30. Xiaoguang, L.; Shicheng, X.; Xiaodong, W.; Tie, L.; Hongrui, W.; Xiuling, L. A compact multi-branch 1D convolutional neural network for EEG-based motor imagery classification. *Biomed. Signal Process. Control* **2023**, *81*, 104456.
31. Collazos-Huertas, D.; Álvarez-Meza, A.; Castellanos-Dominguez, G. Image-Based Learning Using Gradient Class Activation Maps for Enhanced Physiological Interpretability of Motor Imagery Skills. *Appl. Sci.* **2022**, *12*, 1695. [[CrossRef](#)]
32. Fujiwara, Y.; Ushiba, J. Deep Residual Convolutional Neural Networks for Brain–Computer Interface to Visualize Neural Processing of Hand Movements in the Human Brain. *Front. Comput. Neurosci.* **2022**, *16*, 882290. [[CrossRef](#)]
33. Izzuddin, T.; Safri, N.; Othman, M. Compact convolutional neural network (CNN) based on SincNet for end-to-end motor imagery decoding and analysis. *Biocybern. Biomed. Eng.* **2021**, *41*, 1629–1645. [[CrossRef](#)]
34. Kumar, N.; Michmizos, K. A neurophysiologically interpretable deep neural network predicts complex movement components from brain activity. *Sci. Rep.* **2022**, *12*, 1101. [[CrossRef](#)]
35. Tobón-Henao, M.; Álvarez-Meza, A.; Castellanos-Dominguez, G. Subject-Dependent Artifact Removal for Enhancing Motor Imagery Classifier Performance under Poor Skills. *Sensors* **2022**, *22*, 5771. [[CrossRef](#)]
36. Rahman, M.; Mahmood, U.; Lewis, N.; Gazula, H.; Fedorov, A.; Fu, Z.; Calhoun, V.; Plis, S. Interpreting models interpreting brain dynamics. *Sci. Rep.* **2022**, *12*, 12023. [[CrossRef](#)] [[PubMed](#)]
37. Borra, D.; Fantozzi, S.; Magosso, E. Interpretable and lightweight convolutional neural network for EEG decoding: Application to movement execution and imagination. *Neural Netw.* **2020**, *129*, 55–74. [[CrossRef](#)] [[PubMed](#)]
38. Caicedo-Acosta, J.; Castaño, G.A.; Acosta-Medina, C.; Alvarez-Meza, A.; Castellanos-Dominguez, G. Deep neural regression prediction of motor imagery skills using EEG functional connectivity indicators. *Sensors* **2021**, *21*, 1932. [[CrossRef](#)] [[PubMed](#)]

39. Wang, H.; Wang, Z.; Du, M.; Yang, F.; Zhang, Z.; Ding, S.; Mardziel, P.; Hu, X. Score-CAM: Score-weighted visual explanations for convolutional neural networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Seattle, WA, USA, 14–19 June 2020; pp. 24–25.
40. Zeiler, M.; Fergus, R. Visualizing and understanding convolutional networks. In Proceedings of the Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, 6–12 September 2014; Proceedings, Part I 13; Springer: Berlin/Heidelberg, Germany, 2014; pp. 818–833.
41. Wackernagel, H. *Multivariate Geostatistics: An Introduction with Applications*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2003.
42. García-Murillo, D.G.; Álvarez-Meza, A.M.; Castellanos-Domínguez, G. Single-Trial Kernel-Based Functional Connectivity for Enhanced Feature Extraction in Motor-Related Tasks. *Sensors* **2021**, *21*, 2750. [[CrossRef](#)] [[PubMed](#)]
43. Álvarez-Meza, A.M.; Cárdenas-Pena, D.; Castellanos-Domínguez, G. Unsupervised kernel function building using maximization of information potential variability. In *Iberoamerican Congress on Pattern Recognition*; Springer: Berlin/Heidelberg, Germany, 2014; pp. 335–342.
44. Gu, L.; Yu, Z.; Ma, T.; Wang, H.; Li, Z.; Fan, H. Random matrix theory for analysing the brain functional network in lower limb motor imagery. In Proceedings of the 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), IEEE, Montreal, QC, Canada, 20–24 July 2020; pp. 506–509.
45. Li, C.; Qin, C.; Fang, J. Motor-imagery classification model for brain-computer interface: A sparse group filter bank representation model. *arXiv* **2021**, arXiv:2108.12295.
46. Van der Maaten, L.; Hinton, G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **2008**, *9*, 2579–2605.
47. Géron, A. *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow*; O’Reilly Media, Inc.: Sebastopol, CA, USA, 2022.
48. Peterson, V.; Merk, T.; Bush, A.; Nikulin, V.; Kühn, A.A.; Neumann, W.J.; Richardson, R.M. Movement decoding using spatio-spectral features of cortical and subcortical local field potentials. *Exp. Neurol.* **2023**, *359*, 114261. [[CrossRef](#)]
49. Liu, Y.; Wang, Z.; Huang, S.; Wang, W.; Ming, D. EEG characteristic investigation of the sixth-finger motor imagery and optimal channel selection for classification. *J. Neural Eng.* **2022**, *19*, 016001. [[CrossRef](#)]
50. Putzolu, M.; Samogin, J.; Cosentino, C.; Mezzarobba, S.; Bonassi, G.; Lagravinese, G.; Vato, A.; Mantini, D.; Avanzino, L.; Pelosin, E. Neural oscillations during motor imagery of complex gait: An HdEEG study. *Sci. Rep.* **2022**, *12*, 4314. [[CrossRef](#)]
51. Strypsteen, T.; BertrandSenior Member, A. Bandwidth-efficient distributed neural network architectures with application to neuro-sensor networks. *IEEE J. Biomed. Health Inform.* **2022**, *27*, 1–12. [[CrossRef](#)]
52. Nergård, K.; Endestad, T.; Torresen, J. Effect of Hand Dominance When Decoding Motor Imagery Grasping Tasks. In Proceedings of the Computational Neuroscience: Third Latin American Workshop, LAWCN 2021, São Luís, Brazil, 8–10 December 2021; Revised Selected Papers; Springer: Berlin/Heidelberg, Germany, 2022; pp. 233–249.
53. Sadiq, M.T.; Aziz, M.Z.; Almogren, A.; Yousaf, A.; Siuly, S.; Rehman, A.U. Exploiting pretrained CNN models for the development of an EEG-based robust BCI framework. *Comput. Biol. Med.* **2022**, *143*, 105242. [[CrossRef](#)] [[PubMed](#)]
54. Chacon-Murguia, M.; Rivas-Posada, E. A CNN-based modular classification scheme for motor imagery using a novel EEG sampling protocol suitable for IoT healthcare systems. *Neural Comput. Appl.* **2022**, 1–22. [[CrossRef](#)]
55. George, O.; Smith, R.; Madiraju, P.; Yahyasoltani, N.; Ahamed, S.I. Data augmentation strategies for EEG-based motor imagery decoding. *Heliyon* **2022**, *8*, e10240. [[CrossRef](#)]
56. Kumar, S.; Sharma, A.; Tsunoda, T. Brain wave classification using long short-term memory network based OPTICAL predictor. *Sci. Rep.* **2019**, *9*, 9153. [[CrossRef](#)] [[PubMed](#)]
57. Xu, L.; Xu, M.; Ma, Z.; Wang, K.; Jung, T.; Ming, D. Enhancing transfer performance across datasets for brain-computer interfaces using a combination of alignment strategies and adaptive batch normalization. *J. Neural Eng.* **2021**, *18*, 0460e5. [[CrossRef](#)]
58. Zhao, X.; Zhao, J.; Liu, C.; Cai, W. Deep neural network with joint distribution matching for cross-subject motor imagery brain-computer interfaces. *BioMed Res. Int.* **2020**, *2020*, 7285057. [[CrossRef](#)]
59. Jeon, E.; Ko, W.; Yoon, J.; Suk, H. Mutual Information-Driven Subject-Invariant and Class-Relevant Deep Representation Learning in BCI. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *34*, 1–11. [[CrossRef](#)]
60. Deng, X.; Zhang, B.; Yu, N.n.; Liu, K.; Sun, K. Advanced TSGL-EEGNet for Motor Imagery EEG-Based Brain-Computer Interfaces. *IEEE Access* **2021**, *9*, 25118–25130. [[CrossRef](#)]
61. Li, D.; Yang, B.; Ma, J.; Qiu, W. Three-Class Motor Imagery Classification Based on SELU-EEGNet. In Proceedings of the 8th International Conference on Computing and Artificial Intelligence ICCAI ’22, Tianjin, China, 18–21 March 2022; Association for Computing Machinery: New York, NY, USA, 2022; pp. 522–527.
62. Riyad, M.; Khalil, M.; Adib, A. MR-EEGNet: An Efficient ConvNets for Motor Imagery Classification. In *Advanced Intelligent Systems for Sustainable Development (AI2SD2020) Volume 1*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 722–729.
63. Cui, J.; Lan, Z.; Liu, Y.; Li, R.; Li, F.; Sourina, O.; Müller-Wittig, W. A compact and interpretable convolutional neural network for cross-subject driver drowsiness detection from single-channel EEG. *Methods* **2022**, *202*, 173–184. [[CrossRef](#)]
64. Niu, X.; Lu, N.; Kang, J.; Cui, Z. Knowledge-driven feature component interpretable network for motor imagery classification. *J. Neural Eng.* **2022**, *19*, 016032. [[CrossRef](#)]

65. Suhail, T.; Indiradevi, K.; Suhara, E.; Poovathinal, S.; Ayyappan, A. Distinguishing cognitive states using electroencephalography local activation and functional connectivity patterns. *Biomed. Signal Process. Control* **2022**, *77*, 103742. [[CrossRef](#)]
66. Hossain, K.M.; Islam, M.A.; Hossain, S.; Nijholt, A.; Ahad, M.A.R. Status of deep learning for EEG-based brain-computer interface applications. *Front. Comput. Neurosci.* **2023**, *16*, 1006763. [[CrossRef](#)] [[PubMed](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.