International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Multi-Class Motor Imagery Detection using optimum Channels

[1]Rajesh Bhambare, [2]Manish Jain

Submitted: 26/11/2023 **Revised**: 06/01/2024 **Accepted**: 16/01/2024

Abstract: The Brain-Computer Interface (BCI) finds application in various fields such as robotics and environmental control, particularly benefiting individuals with disabilities. Electroencephalography (EEG) signals serve as a prevalent choice for a typical BCI system due to the non-invasive, cost-effective, and portable nature of electrodes. EEG data is often collected from a large number of channels across the brain, so effective channel selection techniques play a vital role in identifying the best channels for a given application. Channel selection helps to decrease setup time and computational complexity while analyzing EEG signals. By eliminating noisy channels, channel selection can improve system performance. The Filer Bank Common Spatial Pattern (FBCSP) based Convolutional Network (CNN) is used to distinguish between four motor imagery (MI) tasks. A sliding window technique is utilized to generate time-varying data on EEG signals. The results obtained from the experimentation of the proposed method on BCI competition IV dataset 2a demonstrate a noteworthy average accuracy of 92.66% across 22 channels. This performance surpasses that of numerous existing systems. Additionally, when employing the mutual information technique for channel selection, extended experimental results revealed a commendable classification accuracy of 89.1% with five channels and 90.66% with three channels. Notably, the use of three channels exhibited an average kappa value of 0.86. These outcomes underscore the efficacy of our proposed system for real-time BCI development. The robustness of the model is further validated by its ability to achieve an accuracy of 89.3% on BCI competition IV dataset 2b. Thus, our proposed model demonstrates consistent and commendable results across both datasets, affirming its potential for practical and reliable application in brain-computer interface systems.

Keywords*: FBCSP, Mutual Information, Motor Imagery, CNN, Kohen Kappa*

1. Introduction

By circumventing the customary neuronal and muscular processes of the brain, as observed noninvasively through scalp recordings, Electroencephalography (EEG) based Brain-Computer Interfaces (BCI) enable the conversion of neural signals associated with cognitive tasks into actionable commands. These neural signals, harnessed from the scalp, are captured at specific locations referred to as channels. This innovative approach leverages the unique capability of EEG to detect and interpret electrical patterns on the scalp, facilitating the seamless translation of mental activities into tangible commands for controlling external devices or applications. The channels, strategically placed across the scalp, serve as conduits for capturing the intricate interplay of neural signals, paving the way for advancements in neurotechnology and enhancing the potential for direct communication between the human brain and external technologies. Dense EEG electrodes provide deeper insight into the underlying neural activity, but they also produce high-dimensional data and increase noise redundancy. Additionally, a practical BCI system require less preparation time which demands an optimum number of relevant channels [1].

[1]*Department of Electrical and Electronics Engineering Mandsaur University, Mandsaur,(M.P), India*
bhambare.rajesh@gmail.com ORCID ID: 0000-0003-4571-726X
[2]*Department of Electrical and Electronics Engineering Mandsaur University,Mandsaur,(M.P), India*
manish.jain@meu.edu.in ORCID ID: 0009-0003-3335-5342

The human brain is divided into various parts, each of which is linked to a particular task. For example, the motor cortex is linked to movement functions, The occipital cortex plays a crucial role in visual processing within the human brain. Located at the back of the brain, the occipital cortex is primarily responsible for receiving and interpreting visual information from the eyes. It processes visual stimuli such as shapes, colors, and motion, allowing individuals to perceive and make sense of the visual world.

On the other hand, the frontoparietal regions are integral to decision-making processes. These areas, situated towards the front and top of the brain, are involved in higher-order cognitive functions, including decision-making, problem-solving, and cognitive control. The frontoparietal network collaborates with other brain regions to integrate sensory information, assess potential outcomes, and execute appropriate responses.

Together, the specialized functions of the occipital cortex and the frontoparietal regions demonstrate the complexity and specialization of different brain regions, highlighting the intricate network of neural processes that contribute to fundamental cognitive functions such as visual perception and decision-making. Therefore, choosing task-related signal electrodes can speed up setup and increase comfort in non-clinical BCI applications. In this study, motor imagery (MI) multichannel EEG signal processing is seen as a multidimensional classification challenge. The curse of dimensionality associated with multichannel signal information and related time-varying features can be alleviated by a channel selection technique. Even if prior knowledge of the neurological correlate aids in choosing the appropriate channels, this may not be the case for all people [2]. For a certain BCI experiment, there is intersubject variability; therefore, the most prominent channels for one user might not be the same for another individual. So, selecting sensors manually might not be the best approach to boost performance.

The issue of selecting a channel has been studied, focusing on trials including movement and mental tasks. In [3] D. Feess et al. provides a study of channel selection techniques for event-related potentials. It talks about channel transfer inside and across sessions to see how different subjects perform. Similar to this, in [4] Y. Li et al. clarifies the importance of channel characteristics. Several strategies have been put forth in the context of MI, concentrating on the appropriate sensor space. The Common Spatial Pattern (CSP) stands out as one of the most widely adopted feature extraction methods, and its efficacy is exemplified in the work of Y. Wang et al., as presented in [5]. This method serves a dual purpose, not only as a feature extraction technique but also as a channel selection algorithm. Wang and colleagues incorporate a sophisticated approach by considering the coefficients of the projection matrix when ranking the channels.

The essence of the CSP lies in its ability to identify spatial filters that maximize the variance of one class while minimizing the variance of another. By leveraging the coefficients of the projection matrix, Wang et al. enhance the channel selection process, ensuring that the chosen channels contribute optimally to the discrimination between different classes of signals. This meticulous consideration of the projection matrix coefficients adds a layer of refinement to the CSP algorithm, making it a robust and sophisticated tool for extracting discriminative features from multichannel data, particularly in the context of applications such as brain-computer interfaces and neurotechnology. CSP has the ability to differentiate between two classes, artifacts in the EEG signal frequently interfere with it [6]. Other channel selection techniques, including the genetic algorithm (GA) for artificial neural networks [7] and sparse CSP [8], need a lot of processing and are not appropriate for online experimentation. Choosing electrodes may be essential for resolving problems with complex and intrusive equipment. It is difficult because experts dispute on the exact number and placement of EEG electrodes. EEG signals are gathered from various areas of the brain and typically use 32 or 64 electrodes. The analysis of EEG data is more computationally complex when there are many electrodes. Additionally, it increases the possibility of signal overlapping and causes interaction problems [9].

Automated feature extraction and categorization utilizing deep learning techniques has recently been the subject of numerous studies. Such approaches' outcomes have demonstrated that they increase accuracy while reducing the need for laborious preprocessing [10,11]. To improve accuracy for motor imagery categorization, an RBM with a four-layer neural network was used in [10] by Lu, N. et al. In order to capture and learn the spatial and temporal aspects of the MI problem, Zhang et al. presented a hybrid deep neural network approach based on CNN and LSTM [11]. In [12], CNN and short-time Fourier transform (STFT) have been employed by Shovon et al. to classify MI into two categories. A further study by, Wang. P. et al. [13] suggests extracting EEG features using LSTM with the dimension-aggregate approximation (1d-AX) channel weighting method to improve classification accuracy. Using an image-based methodology, a CNN was built in [14] by Yang T. et al. to classify EEG data. The authors from [15] Blankertz B. et al. demonstrated how sliding windows from various time segments of a continuous stream of EEG can extract more distinguishable characteristics. To encourage the learning of effective spatial filters, regularized CSP techniques were suggested in [16] by Lotte F. et al. involving the extraction of features from a set time segment of 2s. The incorporation of time segments, as recommended by various studies [15-18], based on the Filter Bank Common Spatial Pattern (FBCSP) has been a subject of investigation. However, despite these efforts, the observed improvements in accuracy have not been substantial. This suggests that simply adding time segments may not be sufficient to enhance performance significantly.

Addressing the challenge of nonstationary EEG signals, Gaur P. et al. proposed sliding window-based CSP approaches in [19]. By adopting this methodology, the approach takes into consideration both session-to-session and trial-to-trial variability. This adaptive strategy is designed to capture and adapt to the dynamic nature of EEG signals, which can exhibit changes over different sessions and trials.

In addition, recurrent neural networks (RNNs) emerge as a pertinent architecture for training sequential processing models tailored for time-series signal analysis, given the inherently sequential nature of EEG signals. RNNs are well-suited for capturing temporal dependencies in data, allowing them to model the sequential aspects of EEG signals effectively. This utilization of RNNs reflects a broader trend in leveraging advanced neural network architectures to unravel the temporal intricacies inherent in sequential data, with promising applications in EEG signal analysis and other domains involving time-series data. The LSTM network is the most common kind of RNN [20].

We use an overlapping sliding window technique to collect EEG signal time series data in addition to enhancing training data sets. In our research, we introduced a novel channel selection strategy based on mutual information. This approach involves assessing the information content of each channel with respect to the task at hand, allowing us to identify and prioritize the most informative channels. Our proposed strategy stands as a unique contribution to the field, aiming to enhance the efficiency and effectiveness of channel selection in our particular context.

To objectively evaluate the performance of our method, we conducted a thorough quantitative comparison with alternative channel selection approaches. This comparative analysis serves to highlight the strengths and advantages of our proposed strategy over existing methods. By employing rigorous metrics and benchmarks, we sought to demonstrate the superior capabilities of our mutual information-based channel selection in terms of accuracy, robustness, or other relevant performance criteria.

Through this research, we aim to not only present a novel approach to channel selection but also to contribute valuable insights into the comparative landscape of existing methodologies. Our findings provide a foundation for informed decisions in selecting the most appropriate channel selection strategy for specific applications, fostering advancements in the broader field of signal processing and analysis. The subsequent sections of this paper are organized as follows to provide a comprehensive overview of the research methodology, results, and conclusions:

Section II: Literature Review

In this section, we offer a comprehensive review of relevant prior work in the field. This includes an examination of existing techniques, methodologies, and advancements related to four-class Motor Imagery (MI) EEG classification.

Section III: FBCSP-Based Technique Using CNN

This section details the proposed technique, which is grounded in Filter Bank Common Spatial Pattern (FBCSP) and employs Convolutional Neural Networks (CNN) for four-class MI EEG classification. The technique is elaborated, discussing the incorporation of channel selection and the variations in CNN architecture.

Section IV: Comparative Analysis

The results obtained from the proposed method are meticulously compared with those of other state-of-the-art research techniques. This comparative analysis aims to highlight the strengths, weaknesses, and distinctive features of the proposed approach in relation to existing methodologies.

Section V: Experimental Findings and Analysis

This section delves into the experimental outcomes, presenting a thorough analysis of the results. Performance metrics, statistical analyses, and other relevant findings are discussed in detail to provide a comprehensive understanding of the experimental results.

Section VI: Conclusion

The paper concludes in Section VI, summarizing the key findings, implications of the research, and potential avenues for future exploration. This section serves as a culmination of the study, encapsulating the contributions, limitations, and overall significance of the proposed FBCSP-based technique with CNN for four-class MI EEG classification.

2. Related Work

The derivation of features plays a pivotal role in determining the performance of a Brain-Computer Interface (BCI) system that utilizes motor imagery, as highlighted in prior research [21]. The effectiveness of a BCI system is closely tied to the quality and relevance of the features extracted from electroencephalogram (EEG) signals. However, the practicality of manually extracting features for real-time classification is significantly hindered by the non-stationary nature of EEG signals and the sheer volume of data generated, particularly when working with a limited number of electrodes [22].

The non-stationary nature of EEG signals means that their characteristics can change over time, necessitating adaptive and efficient feature extraction methods. Furthermore, handling large volumes of data manually is time-consuming and impractical for real-time applications. In light of these challenges, traditional machine learning-based BCI systems face limitations, prompting the need for a different strategy in the development of EEG-driven BCI systems.

In response to these challenges, there is a growing recognition of the importance of leveraging advanced signal processing techniques, such as deep learning and automated feature extraction, to enhance the efficiency and accuracy of BCI systems. By automating the feature extraction process and harnessing the capabilities of advanced algorithms, researchers aim to overcome the constraints imposed by manual methods and address the dynamic and non-stationary characteristics of EEG signals, paving the way for more effective and practical real-time BCI implementations. Some writers have investigated the use of neural networks to automatically learn features from data in order to overcome these limitations.

However, the classic neural network's weight initialization is a challenging operation because tiny initial weights cause weight diffusion and big initial weights result in subpar local minima [23]. Recent strides in deep learning have ushered in a viable solution for the automatic extraction of features from data through the utilization of dense layers comprising numerous hidden units. This development marks a substantial breakthrough, effectively addressing the limitations inherent in traditional machine learning approaches.

In deep learning architectures, particularly neural networks with multiple layers, the thick layer of hidden units enables the model to automatically learn and extract intricate hierarchical representations from the input data. This automated feature extraction process is characterized by the capacity to discern complex patterns and relationships within the data, overcoming the need for manual feature engineering. As a result, deep learning methods have demonstrated significant prowess in handling intricate tasks, especially in domains like image recognition, natural language processing, and, pertinent to BCI systems, the analysis of EEG signals.

By leveraging deep learning's ability to automatically extract relevant features, BCI systems can more adeptly navigate the challenges posed by the non-stationary nature of EEG signals and the impracticality of manual feature extraction in real-time applications. The use of deep learning architectures represents a paradigm shift, offering a more flexible and adaptive approach to feature extraction, ultimately enhancing the efficiency and performance of BCI systems. Without manually constructing features, deep learning algorithms may identify spatial structure features in any particular dataset. Deep learning is proving to be an effective method for classifying biological data, including EEG signals, which often have a lot of interference because it has been shown to be capable of detecting essential properties despite interference from undesired or external signals [24]. A deep learning model's training process affords the opportunity to select important EEG channels for channel reduction since critical features are given higher weights. Using deep learning, Hurbert Cecotti et al. [25] were able to identify p300 in time-domain EEG recordings. Francesco C.M. et al. [26] employed deep learning to separate prodromal forms of dementia from Alzheimer's disease utilizing raw EEG signals in the healthcare domain. Recurrent convolutional neural networks were utilized by Thodorof et al. In reference to [27], a noteworthy application of deep learning involves the automatic detection of seizures by harnessing spatial, spectral, and temporal information derived from electroencephalogram (EEG) signals. This innovative approach signifies a crucial advancement in the field of medical diagnostics, where deep learning techniques demonstrate their effectiveness in discerning complex patterns indicative of seizure activity within EEG data.

Moreover, the application of deep learning-based EEG analysis extends beyond medical contexts and finds utility in biometrics. The inherent capacity of deep learning models to extract intricate features from EEG signals enables their utilization in the field of biometric authentication. By leveraging the unique patterns present in individual EEG signals, deep learning algorithms can facilitate robust and secure biometric identification methods.

These applications underscore the versatility and efficacy of deep learning in unlocking valuable insights from EEG data, whether for medical diagnoses such as seizure detection or for innovative biometric authentication systems. The adaptability of deep learning models to various domains highlights their potential to revolutionize the analysis and interpretation of complex data, contributing to advancements in both healthcare and security. where Mao et al. [28] were able to correctly identify biometrics from EEG signals.

3. Materials and Method

A. Dataset Description

The Brain-Computer Interface Competition IV (BCI-IV) dataset 2a was used as the source of the study's data [29]. Nine volunteers that performed MI tasks requiring movement of the left hand (class 1), right hand (class II), foot (class III), and tongue (class IV) provided the data for the dataset. The 10-20 system was used to capture the EEG using 22 electrodes at a sampling rate of 250 Hz. Every subject was recorded twice, once for training purposes and once for evaluation purposes, on separate days. There is total 288 trials in each session, or 72 trials in each class.

B. Proposed FBCSP-CNN Method

A Convolutional Neural Network (CNN) based on Filter Bank Common Spatial Pattern (FBCSP) represents an integration of two powerful techniques in the domain of signal processing and machine learning features is proposed for optimum channel selection. A model must be trained using all of the known channels, which slows down learning because it must learn from many different channels, some of which are unrelated to brain motor activity. The design and production of smaller, more portable, and less expensive EEG devices with fewer channels will benefit from the identification of the required channels. We will require the exact positioning of the channels on the scalp in order to use these portable EEG devices with fewer channels [30].

C. Experimental Setup

CSP features are extracted for the current frequency band after scanning over a required list of frequency bands. It passes the target labels, the feature matrix, and the time window of size 20. Calculations are made to create a feature matrix and labels relevant to the current band. It is stored in a dictionary after features for the current band have been extracted. In order to make it simpler to assess or train machine learning models on these band-specific features, CSP features were collected from 8 frequency bands. You can access the properties of each band later for additional processing or modeling because they are linked to the appropriate frequency ranges. The Process of Feature Extraction and Classification using FBCSP is shown in

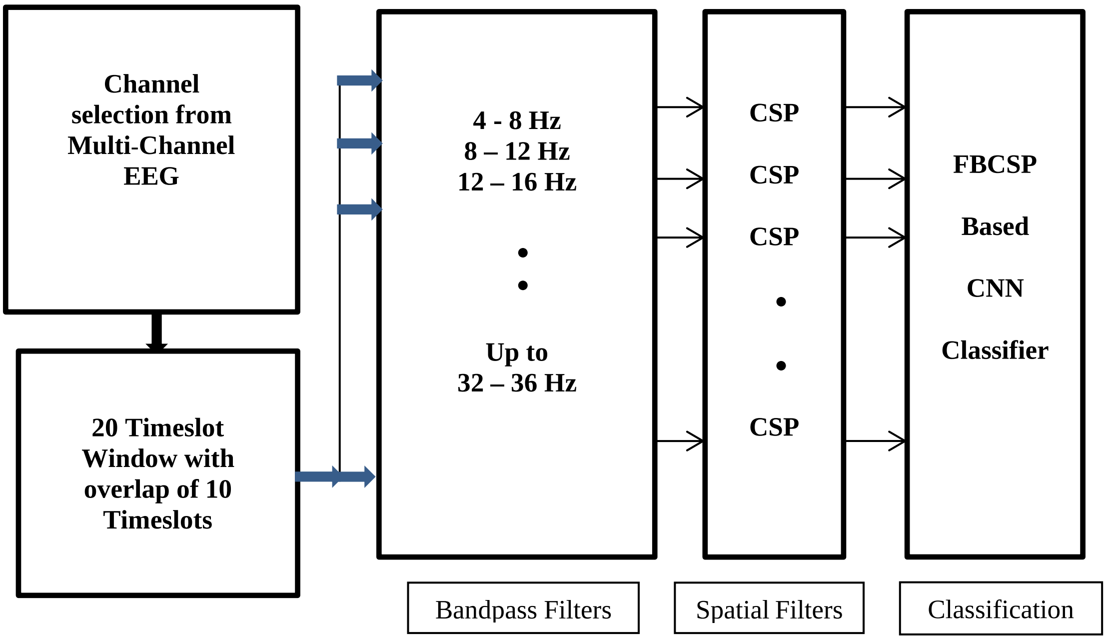


Fig. 1 Process of Feature Extraction and Classification using FBCSP based CNN with Channel Selection

Figure 1. The feature extraction process was carried out after choosing the window for feature extraction. In this investigation, we have proposed the FBCSP based CNN algorithm. Similar to the traditional FBCSP, this algorithm involved four steps. The general framework for the suggested strategy is depicted in Figure 1. In the first phase, a filter bank was used to divide the EEG into 8 frequency passbands, commencing from 4 to 36 Hz with a bandwidth of 4Hz. The mutual information (MI) indicates how much information an attribute under a presumption of independence provides about the class membership. The text describes mutual information as an indicator of association or correlation between variables in both rows and columns. Mutual information provides insights into the relationships between variables X and Y. The formula (1) is provided for calculating mutual information, where p(x, y) is the joint probability distribution function of X and Y, and p(x) and p(y) denote the marginal probability distribution functions of X and Y, respectively.

The interpretation of mutual information is highlighted, noting that a higher mutual information value suggests that the related attribute is more effective in estimating class membership. In other words, variables with higher mutual information values are considered to have a stronger association with the target class, making them more informative for classification tasks.

This concept is particularly relevant in feature selection or attribute ranking, where selecting attributes with higher mutual information can improve the performance of machine learning models by focusing on the most relevant and discriminative features.

It is also possible to calculate the mutual information using equation (2The text introduces additional information related to the calculation of mutual information using entropy terms. The formula for mutual information (I(X;Y)) can be expressed in terms of entropy as follows:

[I(X;Y) = H(Y) - H(Y|X) = H(X) - H(X|Y)]

Where:

- (H(Y)) is the marginal entropy of variable Y.

- (H(X|Y)) is the conditional entropy of variable X given Y.

- (H(X)) is the marginal entropy of variable X.

- (H(Y|X)) is the conditional entropy of variable Y given X.

Additionally, the joint entropy (H(X,Y)) can be related to these terms:

[H(X,Y) = H(X) + H(Y) - I(X;Y)]

These relationships highlight the fundamental connection between mutual information and entropy. Mutual information quantifies the reduction in uncertainty about one variable gained by knowing the other variable, and entropy terms characterize the uncertainty or disorder associated with random variables. In essence, mutual information reflects the information shared between X and Y, considering both marginal and conditional entropies. The feature vector $A_V^i = [a_{B1}^i a_{B2}^i .....a_{Bn}^i]$ is created by concatenating the features received from each band. Here A_V^i represents feature vector of i-th trial, a_{Bj}^i denotes the features received from j-th band of i-th trial, and n gives the total number of bands. The train data's feature vectors for each trial are used to create the feature matrix $A_M[A_V^1; A_V^2; ....A_V^n]$ is

(X,Y)=∑_(y∈Y)∑_(x∈X) 〖p(x,y)log((p(x,y))/(p(x)p(y))〗) (1)

$$I(X,Y) = H(Y) - H(Y \mid X) \tag{2}$$

$$= H(X,Y) - H(X \mid Y) - H(Y \mid X) \tag{3}$$

TABLE I Proposed CNN Model

Layer (type)	**Output Shape**	**Parameters**
conv2d (Conv2D)	(None, 999, 4, 128)	640
max_pooling2d (MaxPooling2D)	(None, 499, 2, 128)	0
dropout (Dropout)	(None, 499, 2, 128)	0
conv2d_1 (Conv2D)	(None, 499, 2, 128)	16512
dropout_1 (Dropout)	(None, 499, 2, 128)	0
conv2d_2 (Conv2D)	(None, 499, 2, 256)	33024
dropout_2 (Dropout)	(None, 499, 2, 256)	0
flatten (Flatten)	(None, 255488)	0
dense (Dense)	(None, 512)	130810368
dropout_3 (Dropout)	(None, 512)	0
dense_1 (Dense)	(None, 256)	131328
dropout_4 (Dropout)	(None, 256)	0
dense_2 (Dense)	(None, 128)	32896
dropout_5 (Dropout)	(None, 128)	0
dense_3 (Dense)	(None, 4)	516
Total Parameters: 131025284 Trainable Parameters: 131025284 Non-trainable Parameters: 0		

The mutual information (MI) is computed by employing equation (3), using the feature matrix. The resulting MI values are denoted as MI = [J1, J2...JL], where JL signifies the mutual information value corresponding to the l-th feature. Each value in the MI array represents the extent to which the associated feature contributes relevant information.

In the context of MI-based Brain-Computer Interface (BCI) systems, three specific EEG electrodes, namely C3, C4, and Cz, are commonly selected as candidate channels. These electrodes play a crucial role in capturing essential characteristics of motor imagery (MI)-specific electrical activity. The choice of C3, C4, and Cz is grounded in the neurophysiological foundation of the human brain.

The rationale behind selecting these three electrodes is rooted in the understanding of the human brain's motor cortex, which serves as the driving force for voluntary muscle movement. The motor cortex is predominantly located in the frontal lobe of the brain and is symmetrically distributed between the left and right hemispheres. By focusing on the EEG signals from electrodes C3, C4, and Cz, MI-based BCI systems strategically target regions associated with motor activity, aiming to capture relevant neural information linked to the intention of movement. This targeted electrode selection aligns with the neuroanatomical considerations of the motor cortex, enhancing the specificity and efficacy of the BCI system for decoding motor-related signals.

4. Evaluation of the Suggested Method in Comparison to Other Methods

Jeong hee Hwang et al. presented a classification framework in [33] based on Long Short-Term Memory (LSTM) networks, aiming to enhance the accuracy of categorizing four-class Motor Imagery (MI) signals. The methodology involves the application of a sliding window technique to generate time-varying EEG signal data. Concurrently, an overlapping-band-based Filter Bank Common Spatial Pattern (FBCSP) is employed to extract spatial features that are unique to each individual.

In terms of performance, the experimental results on BCI competition IV dataset 2a demonstrated an average accuracy of 93.9%. This underscores the effectiveness of the proposed LSTM-based framework in achieving high accuracy in the classification of MI signals. However, it's acknowledged that LSTMs can demand substantial processing power, particularly when dealing with extended or sizable datasets. The use of a sliding window approach, although beneficial for capturing temporal dependencies, further intensifies the computational burden. This heightened demand may pose challenges, especially in real-time applications where timely processing is critical. The trade-off between accuracy and computational efficiency is highlighted here, acknowledging the potential limitations associated with the computational demands of LSTM-based frameworks. Balancing these factors is crucial for determining the suitability of such approaches in various applications, with considerations for real-time processing constraints and the available computing resources. Using the Dynamic Channel Relevance (DCR) score, Anurag Tiwari and Amrita Chaturvedi propose [34] a unique approach to channel selection that finds efficient EEG channels with 85.38% classification accuracy. EEG signals are prone to artifacts and noise. Channel selection quality may suffer if DCR scores unintentionally favor channels that record noise or irrelevant artifacts.

In [35], Q. Ai et al. suggested two other feature extraction algorithms, local characteristic-scale decomposition (LCD) and CSP are combined with the features of functional brain networks to extract discriminative features. The suggested approach achieves an average classification accuracy of 79.67%. A high-dimensional feature space could result from the combination of LCD, CSP, and functional brain network characteristics. Issues with high dimensionality can include higher processing expenses, a higher chance of overfitting, and trouble determining the significance of particular characteristics.

In their work described in [36], Hongtao Wang et al. proposed a classification framework for distinguishing between three distinct motor imagery movements: left-hand, right-hand, and foot movements. The approach recommended by the authors involves the implementation of a cascade structure that combines the one-versus-the-rest filtered-bank common spatial pattern (OVR-FBCSP) method with a multi-kernel relevance vector machine (MK-RVM). The goal of this methodology is to effectively discriminate between the different motor imagery classes.

The OVR-FBCSP approach is commonly used in Brain-Computer Interface (BCI) systems for multiclass classification tasks. It employs a one-versus-the-rest strategy, creating binary classifiers for each class, which are then combined to make multiclass predictions. The filtered-bank common spatial pattern (FBCSP) is utilized to extract discriminative spatial features from EEG signals.

The MK-RVM is employed as the classifier, leveraging the advantages of relevance vector machines in handling high-dimensional data. The multi-kernel aspect suggests the utilization of multiple kernel functions, allowing the model to capture diverse types of information from the input data.

The outcome of this cascade structure, as reported by the authors, is an average classification accuracy of 83.21%. This metric reflects the model's ability to correctly classify motor imagery movements across the three specified classes. The cascade structure, combining feature extraction with FBCSP and classification with MK-RVM, showcases a promising approach for improving accuracy in multiclass motor imagery classification tasks.

It may be difficult to interpret these complex models, making it difficult to comprehend how particular traits or kernels affect the final categorization choice. Reduced generalization performance on untested data could result from overfitting.

The study conducted by Arvaneh M. et al. [37] introduces a novel technique called Sparse Common Spatial Pattern (SCSP) for EEG channel selection. The primary objective of this technique is to sparsify common spatial filters in a way that reduces the number of channels while maintaining a predefined classification accuracy limit. The reported classification accuracy achieved by this method is 78.93%.

The SCSP technique operates on the premise of common spatial patterns, which are commonly used in Brain-Computer Interface (BCI) systems for extracting discriminative features from EEG signals. However, the innovation in this study lies in the sparsification of these spatial filters. By introducing sparsity constraints, the SCSP technique aims to identify and retain only the most relevant channels for maintaining a balance between classification accuracy and channel reduction.

This approach is particularly relevant in scenarios where reducing the number of EEG channels is desirable due to practical constraints, such as limited available electrodes or to minimize computational complexity. The reported classification accuracy of 78.93% suggests that the SCSP technique effectively achieves this channel reduction without compromising classification performance beyond the specified limit.

In summary, the study by Arvaneh M. et al. [37] presents a valuable contribution to the field of EEG-based classification by introducing the SCSP technique, a method that strategically sparsifies common spatial patterns to optimize the balance between classification accuracy and the number of channels used. Even if SCSP is good at choosing a subset of channels, it may be difficult to comprehend the neurophysiological significance of the channels that are chosen. It might not be directly interpretable in terms of the fundamental functions of the brain.

The proposed method gives a cross validation accuracy of 90.66% for 3 channels and 89.10 % for 5 channels which is close to that of 92.66% for 22 channels. The cross-validation accuracy assesses a model's performance across various subjects or datasets. It evaluates robustness, minimizes subject-specific pattern bias, evaluates generalizability, and simulates unknown data scenarios to get a more accurate assessment of real-world performance. It provides a more comprehensive knowledge of a model's performance in a variety of scenarios [38]. Spatial filters over a range of frequency bands are used by FBCSP to effectively extract discriminative characteristics from EEG signals. High-dimensional EEG signals can be processed and handled more effectively thanks to FBCSP, which decreases the dimensionality of EEG data while maintaining pertinent information. A CNN's capacity to automatically learn hierarchical representations is enhanced when FBCSP characteristics are incorporated into the design. This enhances the use and comprehension of the features that are retrieved [39].

The proposed framework is also assessed using dataset 2b from the BCI Competition IV [40]. This dataset comprises nine subjects. The Validation accuracy achieved is 0.893 with Cohen-kappa value of 0.654 using the proposed FBCSP-CNN model. It gives good results on both datasets, BCI IV 2a and BCI IV 2b giving a generalized model.

5. Results and Discussions

TABLE II.

	Classification Report for 5 channels on BCI IV 2a dataset			
Class	***Precision***	***Recall***	***F1-score***	***Support***
0	0.94	0.86	0.90	56
1	0.90	0.90	0.90	42
2	0.88	0.81	0.84	47
3	0.84	0.98	0.91	55

TABLE III

	Classification Report for 5 channels on BCI IV 2a dataset			
Class	***Precision***	***Recall***	***F1-score***	***Support***
0	0.93	0.94	0.93	81
1	0.85	0.91	0.88	70
2	0.90	0.80	0.85	80
3	0.89	0.93	0.91	60

TABLE IV

	Classification Report for 3 channels on BCI IV 2b dataset			
Class	***Precision***	***Recall***	***F1-score***	***Support***
0	0.88	0.77	0.82	155
1	0.79	0.88	0.83	145

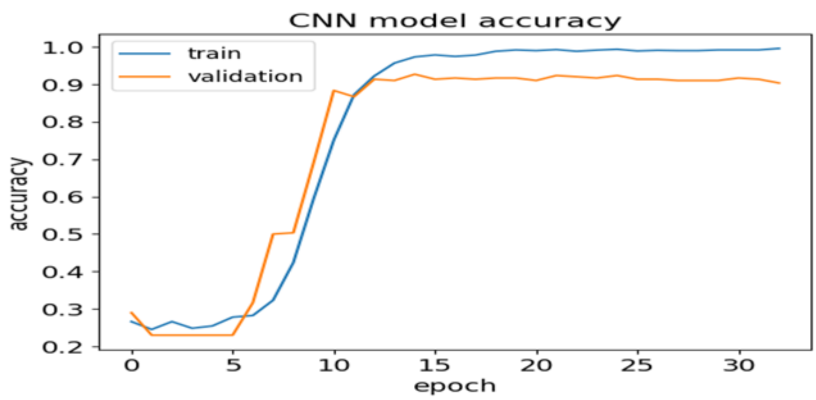
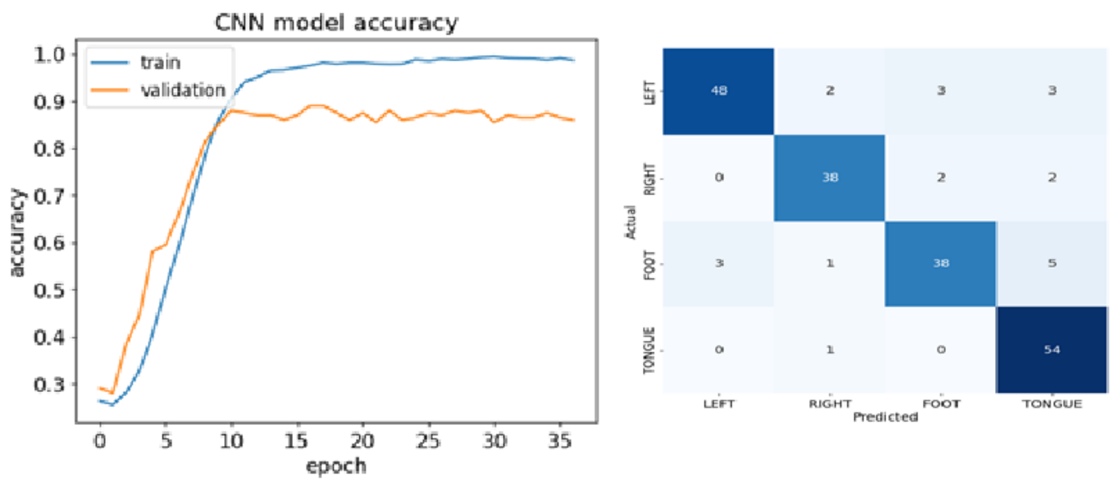
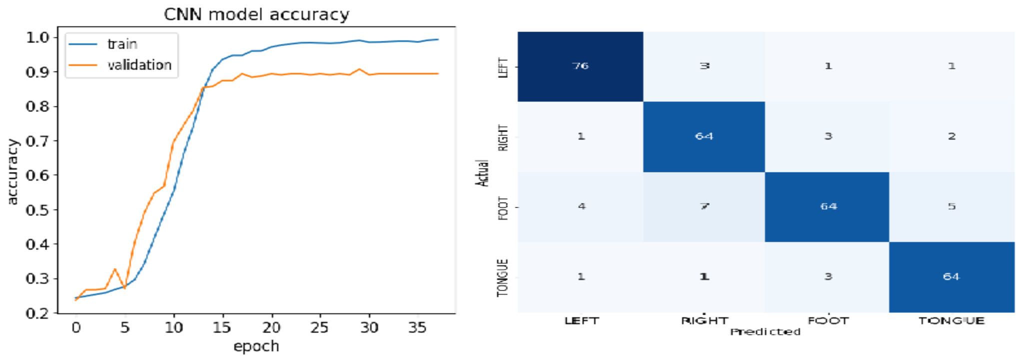


Fig 2. CNN model accuracy curve for 22 channels



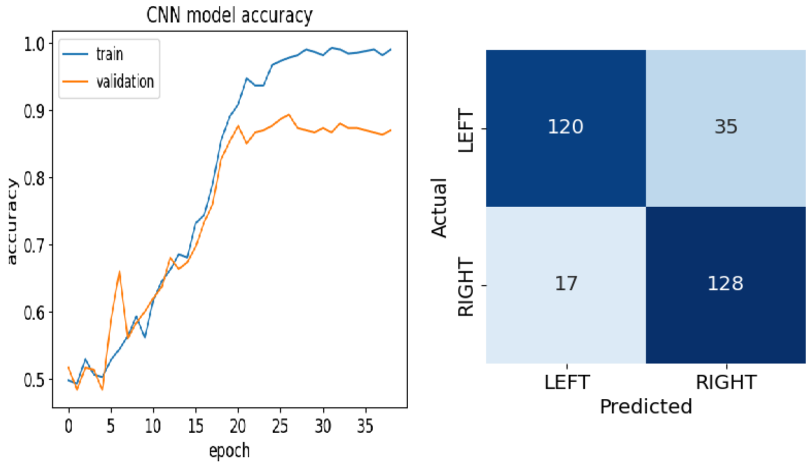
(a) **(b)**

Fig 3. (a) CNN Model Training and Validation Accuracy for 5 channels and (b) Confusion Matrix for the Proposed CNN Model for 5 Channels on BCI IV 2a dataset



(a) **(b)**

Fig 4. (a) CNN Model Training and Validation Accuracy for 3 channels and (b) Confusion Matrix for the Proposed CNN Model for 3 Channels on BCI IV 2a dataset



(a) (b)

Fig 5. (a) CNN Model Training and Validation Accuracy for 3 channels and (b) Confusion Matrix for the Proposed CNN Model for 3 Channels on BCI IV 2b dataset

TABLE 5.

The Subject Independent Classification accuracy of various methods on the MI-Dataset BCI IV 2a			
Classifier	***Feature Extracted***	***No. of Channels***	***Accuracy***
LSTM [33]	FBCSP	7	93.9%
SVM [34]	SCSP	11	85.38%
SRDA (Spectral regression discriminant analysis) [35]	CSP + LCD + Brain Network	3	79.67%
(Multi – Kernel Relevance Vector Machine) [36]	FBCSP + PLV(Phase Locking Value)	9	83.21%
SVM [37]	Sparse CSP	9	78.93%
Proposed method	**FBCSP + CNN** BCI IV 2a dataset	**5**	**89.10 %**
		3	**90.66%**
		22	**92.66%**
	FBCSP + CNN BCI IV 2b dataset	**3**	**89.30%**

6. Conclusion

The experimental results demonstrate the robustness of our proposed strategy in identifying distinguishing brain activity patterns during various MI tasks, leading to an improvement in classification accuracy in four-class MI tasks. Real-time rehabilitation systems can benefit greatly from the high classification accuracy and low processing load. The subject-independent channel selection for BCI applications based on motor imagery was the main emphasis of this paper. In order to achieve this, we looked at the channel reduction strategy that limits the classification accuracy to a reasonable range. By concentrating on channels that reliably transmit task-related information across individuals, channel selection based on MI can improve generalization and increase the model's capacity to generalize to new data. When paired with spatial filtering methods like FBCSP, MI-based channel selection can maximize the selection of pertinent channels for use in later feature extraction, which results in more efficient and condensed feature representations. The proposed method outperforms other methods in comparison like, LSTM-CNN, SVM-SCSP, SRDA, and FBCSP-PLV. The proposed model gives better results on both datasets, BCI IV 2a and BCI IV 2b.

Author contributions

- **Rajesh Bhambare:** Development of ideas, research design, software utilization, data representation, active exploration, writing and reviewing, editing, and on-site study.
- **Dr. Manish Jain:** Curating data, preparing the original draft, developing software, and conducting validation.
- **Conflicts of interest:** The authors affirm that there are no conflicts of interest.

References

[1] X. Zhang, L. Yao, X. Wang,, J. Monaghan, D. Mcalpine, &Y. A. Zhang, " A survey on deep learning based brain computer interface: Recent advances and new frontiers." arXiv:1905.04149, arXiv e-prints, (p. arXiv:1905.04149) , 2019

[2] B. Blankertz, F. Losch, M. Krauledat, G. Dornhege, G. Curio, and K.-R. M¨uller, "The berlin brain–computer Interface: Accurate performance from first-session in BCI-na¨ıve subjects." IEEE Trans. Biomed. Eng., vol. 55, no. 10, pp. 2452–62, Oct. 2008.

[3] D. Feess, M. M. Krell, and J. H. Metzen, "Comparison of sensor selection mechanisms for an ERP-based brain-computer interface," PLoS ONE, vol. 8, no. 7, p. e67543, 2013.

[4] Y. Li, A. Cichocki, and S.-I. Amari, "Blind estimation of channel parameters and source components for EEG signals: A sparse factorization approach," IEEE Trans. Neural Netw., vol. 17, no. 2, pp. 419–31, Mar. 2006.

[5] Y.Wang, S. Gao, and X. Gao, "Common spatial pattern method for channel selection in motor imagery based brain-computer interface." in Proc. Int. Conf. Eng. Med. Biol. Soc., 2005, vol. 5, pp. 5392–5395.

[6] H. Ramoser, J. M¨uller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," IEEE Trans. Rehabil. Eng., vol. 8, no. 4, pp. 441–446, Dec. 2000.

[7] J. Yang et al., "Channel selection and classification of electroencephalogram signals: An artificial neural network and genetic algorithm based approach." Artif. Intell. Med., vol. 55, no. 2, pp. 117–126, Jun. 2012.

[8] M. Arvaneh, C. Guan, K. K. Ang, and C. Quek, "Optimizing the channel selection and classification accuracy in EEG-based BCI." IEEE Trans. Biomed. Eng., vol. 58, no. 6, pp. 1865–1873, Jun. 2011.

[9] Nandini K. Bhandari, and Manish Jain, "EEG Emotion Detection Using Optimum Channels," NeuroQuantology, Volume 20, Issue 8, pp. 1570-1581, July 2022

[10] Lu, N.; Li, T.; Ren, X.; Miao, H. A Deep Learning Scheme for Motor Imagery Classification based on Restricted Boltzmann Machines. IEEE Trans. Neural Syst. Rehabil. Eng. 2016, 25, 566–576.

[11] Zhang, R.; Zong, Q.; Dou, L.; Zhao, X. A novel hybrid deep learning scheme for four-class motor imagery classification. J. Neural Eng. 2019, 16, 066004.

[12] Shovon, T.H.; Al Nazi, Z.; Dash, S.; Hossain, M.F. "Classification of motor imagery EEG signals with multi-input convolutional neural network by augmenting STFT ", In Proceedings of the 2019 5th International Conference on Advances in Electrical Engineering (ICAEE), Dhaka, Bangladesh, 26–28 September 2019; pp. 398–403.

[13] Wang. P., Jiang. A., Liu. X., Shan. J, Zhang, L. LSTM-based EEG classification in motor imagery tasks. IEEE Trans. Neural Syst. Rehabil. Eng. 2018, 26, 2086–2095.S Sivash Sakhavi, Cuntai Guan and Shuicheng Yan, " Learning Temporal Information for Brain-Computer Interface using Convolutional Neural Networks" IEEE Transactions on Neural Networks and Learning Systems, Volume 29, Issue 11, pp. 5619-5629, Nov. 2018

[14] Yang T., Phua, K.S.,Yu, J ,Selvaratnam, T., Toh, V.; Ng, W.H.; So, R.Q. Image-based motor imagery EEG classification using convolutional neural network. In Proceedings of the 2019 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI), Chicago, IL, USA, 19–22 May 2019; pp. 1–4.

[15] Blankertz B.,Tomioka, R.,Lemm, S., Kawanabe, M., Muller, K.-R. Optimizing Spatial filters for Robust EEG

Single-Trial Analysis. IEEE Signal Process. Mag. 2007, 25, 41–56. O.Y. Kwon, M.H. Lee, C. Guan, and S.W. Lee,“Subject-independent brain computer interfaces based on deep convolutional neural networks," IEEE Trans. Neural Netw. Learn. Syst., vol. 31, no. 10, pp. 3839-3852, Oct. 2019.

[16] Lotte F., Guan, C., “Regularizing Common Spatial Patterns to Improve BCI Designs: Unified Theory and New Algorithms”, 58, IEEE Trans. Biomed. Eng. pp. 355–362, 2010.

[17] Ang, K.K.; Chin, Z.Y.; Zhang, H.; Guan, C. Mutual information-based selection of optimal spatial–temporal patterns for single-trial EEG-based BCIs. Pattern Recognit. 2012, 45, 2137–2144.

[18] Yahya-Zoubir, B.; Bentlemsan, M.; Zemouri, E.T.; Ferroudji, K. Adaptive time window for EEG-based motor imagery classification. In Proceedings of the International Conference on Intelligent Information Processing, Security and Advanced Communication, Batna, Algeria, 23–25 November 2015; pp. 1–6.

[19] Gaur P., Gupta, H., Chowdhury, A.,McCreadie, K., Pachori, R.B.,Wang, H. A , “Sliding Window Common Spatial Pattern for Enhancing Motor Imagery Classification in EEG-BCI.”, IEEE Trans. Instrum. Meas. 2021, 70, 4002709.

[20] Liu, Y.; Huang, Y.-X.; Zhang, X.; Qi, W.; Guo, J.; Hu, Y.; Zhang, L.; Su, H.” Deep C-LSTM Neural Network for Epileptic Seizure and Tumor Detection Using High-Dimension EEG Signals”, IEEE Access 2020, 8, 37495–37504.

[21] G.E. Hinton and R.R. Salakhutdinov, “Reducing the dimensionality of data with neural networks,” science, 313(5786), pp.504-507, 2006.

[22] M.P. Hosseini, D. Pompili, K. Elisevich, and H. Soltanian-Zadeh, “Optimized deep learning for EEG big data and seizure prediction BCI via internet of things,” IEEE Transactions on Big Data, 3(4), pp.392-404, 2017.

[23] G.E. Hinton and R.R. Salakhutdinov, “Reducing the dimensionality of data with neural networks,” science, 313(5786), pp.504-507, 2006.

[24] N.S. Kwak, K.R. Müller, and S.W. Lee, “A convolutional neural network for steady state visual evoked potential classification under ambulatory environment,” PloS one, 12(2), p.e0172578, 2017.

[25] H. Cecotti and A. Graser, “Convolutional neural networks for P300 detection with application to brain-computer interfaces,” IEEE Transactions on Pattern Analysis and Machine Intelligence, 33(3), pp.433-445, 2011.

[26] F.C. Morabito et al., “Deep convolutional neural networks for classification of mild cognitive impaired and Alzheimer's disease patients from scalp EEG recordings,” In Research and Technologies for Society and Industry Leveraging a better tomorrow (RTSI), 2016 IEEE 2nd International Forum on (pp. 1-6). IEEE, 2016.

[27] P. Thodoroff, J. Pineau, A. Lim, “Learning robust features using deep learning for automatic seizure detection,” In Machine learning for healthcare conference (pp. 178-190), 2016.

[28] Z. Mao, W.X. Yao, and Y. Huang, “EEG-based biometric identification with deep learning,” In Neural Engineering (NER), 2017, 8th International IEEE/EMBS Conference on (pp. 609-612). IEEE, 2017.

[29] C. Brunner, R. Leeb, G. Mleer-Putz, A. Schlogl, “BCI competition 2008-graz data set A,” Inst. Know. Disc., Lab. Berlin –Comp. Interfaces, Graz University, Austria, 2008 pp. 136-142, vol. 16 [Online]. Available: http://www.bbci.de/competition/iv/desc 2a.pdf

[30] Deogratias Mzurikwao,Woon-Hong Yeo and Xiangxin Li "A Channel Selection Approach Based on Convolutional Neural Network for Multi-channel EEG Motor Imagery Decoding", 2019 IEEE Second International Conference on Artificial Intelligence and Knowledge Engineering (AIKE)

[31] Shiu Kumar, Alok Sharma ,Tatsuhiko Tsunoda ,"An improved discriminative filter bank selection approach for motor imagery EEG signal classification using mutual information”, From 16th International Conference on Bioinformatics (InCoB 2017) Shenzhen, China. September 20–22, 2017

[32] J. Jin, Y. Miao, I. Daly, C. Zuo, D. Hu, and A. Cichocki, ``Correlation based channel selection and regularized feature optimization for MI-based BCI," Neural Netw., vol. 118, pp. 262270, Oct. 2019.

[33] Jeonghee Hwang , Soyoung Park and Jeonghee Chi , “Improving Multi-Class Motor Imagery EEG Classification Using Overlapping Sliding Window and Deep Learning Model”, Electronics,2023, 12, 1186, pp. 1- 16. 202

[34] Anurag Tiwari and Amrita Chaturvedi, “A Novel Channel Selection Method for BCI Classification Using Dynamic Relevance”, IEEE Access, Volume 9, pp. 126698- 126716, September 20, 2021.

[35] Q. Ai, A. Chen, K. Chen, Q. Liu, T. Zhou, S. Xin, and Z. Ji, “Feature extraction of four-class motor imagery EEG signals based on functional brain network," J. Neural Eng., vol. 16, no. 2, Apr. 2019, Art. no. 026032.

[36] Hongtao wang , Tao xu, Cong tang, Hongwei yue, Chuangquan chen, Linfeng x, Zian pei,“ Diverse

Feature Blend Based on Filter-Bank Common Spatial Pattern and Brain Functional Connectivity for Multiple Motor Imagery Detection", IEEE Access, Volume 8, pp. 155590 - 155601, September , 2021.

[37] Arvaneh M, Guan C, Ang KK, Quek C (2011) Optimizing the channel selection and classification accuracy in EEG-based BCI. 58(6) IEEE Trans Biomed Eng: 1865–1873.

[38] Liye Lv , Xueguan Song and Wei Sun, "Modify Leave-One-Out Cross Validation by Moving Validation Samples around Random Normal Distributions: Move-One-Away Cross Validation" MDPI, Appl. Sci. 2020, 10, 2448. pp. 1-13, 2020

[39] Kai KengAng ,ZhengYangChin, ChuanchuWang, CuntaiGuan and HaihongZhang," Filterbank common spatial pattern algorithm on BCI competition IV Datasets 2a and 2b",Frontiers in Neuroscience", Volume6,Article39,pp. 1-9, March 2012.

[40] Brunner, C.; Leeb, R.; Muller-Putz, G.R.; Schlogl, A. BCIcompetition 2008—Graz data set A 6. Graz Univ. Technol. 2008, 16, 1–6.