

## Article

# Motor Imagery Classification Using Effective Channel Selection of Multichannel EEG

Abdullah Al Shiam <sup>1</sup>, Kazi Mahmudul Hassan <sup>2</sup>, Md. Rabiul Islam <sup>3</sup>, Ahmed M. M. Almassri <sup>4</sup>,  
Hiroaki Wagatsuma <sup>5</sup> and Md. Khademul Islam Molla <sup>6,\*</sup>

<sup>1</sup> Department of Computer Science and Engineering, Sheikh Hasina University, Netrokona 2400, Bangladesh; shiam.cse@shu.edu.bd

<sup>2</sup> Department of Computer Science and Engineering, Jatiya Kabi Kazi Nazrul Islam University, Trishal, Mymensingh 2224, Bangladesh; kmhassan@jkkniu.edu.bd

<sup>3</sup> Department of Medicine, University of Texas Health Science Center, San Antonio, TX 78229, USA; rabiul@sip.tuat.ac.jp

<sup>4</sup> Department of Intelligent Robotics, Faculty of Engineering, Toyama Prefectural University, Toyama 939-0398, Japan; almassri@pu-toyama.ac.jp

<sup>5</sup> Department of Human Intelligence Systems, Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology, Fukuoka 808-0196, Japan; waga@brain.kyutech.ac.jp

<sup>6</sup> Department of Computer Science and Engineering, The University of Rajshahi, Rajshahi 6205, Bangladesh

\* Correspondence: khademul.cse@ru.ac.bd

**Abstract:** Electroencephalography (EEG) is effectively employed to describe cognitive patterns corresponding to different tasks of motor functions for brain–computer interface (BCI) implementation. Explicit information processing is necessary to reduce the computational complexity of practical BCI systems. This paper presents an entropy-based approach to select effective EEG channels for motor imagery (MI) classification in brain–computer interface (BCI) systems. The method identifies channels with higher entropy scores, which is an indication of greater information content. It discards redundant or noisy channels leading to reduced computational complexity and improved classification accuracy. High entropy means a more disordered pattern, whereas low entropy means a less disordered pattern with less information. The entropy of each channel for individual trials is calculated. The weight of each channel is represented by the mean entropy of the channel over all the trials. A set of channels with higher mean entropy are selected as effective channels for MI classification. A limited number of sub-band signals are created by decomposing the selected channels. To extract the spatial features, the common spatial pattern (CSP) is applied to each sub-band space of EEG signals. The CSP-based features are used to classify the right-hand and right-foot MI tasks using a support vector machine (SVM). The effectiveness of the proposed approach is validated using two publicly available EEG datasets, known as BCI competition III–IV(A) and BCI competition IV–I. The experimental results demonstrate that the proposed approach surpasses cutting-edge techniques.

**Keywords:** brain–computer interface; channel selection; electroencephalography; entropy-based information; motor imagery



**Citation:** Shiam, A.A.; Hassan, K.M.; Islam, M.R.; Almassri, A.M.M.; Wagatsuma, H.; Molla, M.K.I. Motor Imagery Classification Using Effective Channel Selection of Multichannel EEG. *Brain Sci.* **2024**, *14*, 462. <https://doi.org/10.3390/brainsci14050462>

Academic Editor: Claudio Lucchiari

Received: 5 April 2024

Revised: 26 April 2024

Accepted: 29 April 2024

Published: 3 May 2024



**Copyright:** © 2024 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

A brain–computer interface (BCI) is a system that enables humans to communicate with the environment without depending on a peripheral limb or muscle [1,2]. BCI can assist people in carrying out regular tasks without moving their bodies. BCI applications extend from communication and rehabilitation to entertainment [3]. Noninvasive methods such as functional magnetic resonance imaging (fMRI), magnetoencephalography (MEG), and electroencephalography (EEG) can be used to monitor brain activity. The EEG-based BCI application has become the most popular because it is very easy to implement and is cost-effective [4–8]. A common form of mental task employing EEG-based BCI is motor imagery (MI), which demands users to articulate particular limb or muscle movements

without performing actual action. Event-related synchronization (ERS) and event-related desynchronization (ERD), which MI-BCI systems may detect, are oscillatory activity in sensorimotor rhythms that are produced by these imaginations [9–11]. The MI classification using multichannel EEG is investigated in this study to identify MI-related brain activity patterns. The aim of this work is to improve the performance of BCI systems. Although there is a noticeable advancement in the study of MI classification using EEG, its performance is affected by the high dimensionality of EEG, intersubject variability, and noise. Moreover, it is still not clear how well the conventional design of channels captures brain signals that are effective for MI categorization [12]. This study focuses on the selection of optimal channels for MI classification using multichannel EEG to overcome the traditional limitation. The channels with redundant and noisy information are discarded such that the MI classification is improved. Rigorous experiments are conducted to demonstrate the efficacy of the proposed method.

An entropy-based channel selection mechanism is introduced to implement the MI classification system in this paper. The EEG signals with a reduced number of channels are decomposed into a finite number of sub-bands. The common spatial pattern-based features are calculated from the sub-bands and used for the classification.

The rest of this paper is organized as follows: Section 2 discusses the related works that already have been conducted on MI classification; Section 3 describes the data used in this study; Section 4 explains the technique; Section 5 illustrates the experimental results; and Sections 6 and 7 include the discussion and some concluding remarks, respectively.

## 2. Related Works

For a variety of applications, many researchers have been striving to create EEG-based BCI systems [12–14]. For better performance, most BCI systems gather EEG inputs from several scalp electrodes. However, it takes a lot of time and effort to set up a BCI system with the sensors. Additionally, it is a significant obstacle to integrating BCI into commonplace applications [13,15,16]. A survey includes new frontiers in applying deep learning for brain signal analysis [17]. Furthermore, the presence of large EEG channels introduces issues such as distortion, excessive computational complexity, and data redundancy [18,19]. Therefore, while designing BCI systems, the selection of a reduced number of effective channels becomes demanding. Several approaches have been introduced to select the effective number of EEG channels. Some techniques rank the channels under independent evaluation standards without employing any classifier. In contrast, other methods evaluate the significance of channels by considering the effectiveness of particular classifiers [20].

Graph Sequence Neural Network (GSNN) for accurately decoding patterns of motor imagery from EEGs is proposed in [21]. The model utilizes a node domain attention selection network. It focuses on training sample datasets specific to each subject at different distraction themes, adapting to subjects' influences, and reducing noise samples [21]. Specifically, it illustrates the difficulty in deciding between a global pooled architecture and a hierarchical pooled architecture for graph classification in MI (Motor Imagery). The errors of linear discriminant analysis (LDA) classifier test have been employed in [22] as criteria to select potential EEG channels. These approaches have not achieved the required level of classification accuracy. A correlation-based channel selection with a regularized common spatial pattern (CCS-RCSP) is implemented in [23] for MI classification. In order to correctly identify the MI tasks, a support vector machine (SVM) classifier with the radial basis function (RBF) kernel is used. In [24], a set of the relevant channels are selected using the binary gravitational search algorithm (BGSA). A bandpass filter was used there to reduce the noise. Additionally, artifacts like electromyography (EMG) and electrooculography (EOG) are eliminated using the blind source separation (BSS) technique. Signals are then processed in time and wavelet domains to extract features. The efficiency of their channel selection strategy is then evaluated using the SVM classifier.

In [25], a unique channel selection method based on the correlation coefficient (CSCC) is used. For each channel, a collection of uncorrelated channels is chosen. To pick one of

these groups, the Fisher score is computed after extracting the features of each group using filter-bank CSP (FBCSP). A group of channels is then chosen which has the highest Fisher score. A channel selection approach is suggested in [19] using the spatiotemporal information in EEG data. The selection process has been conceptualized as an optimization problem. It is solved via a computationally efficient approach by incorporating sparsity restrictions.

The common spatial pattern (CSP)-based method is a known technique for EEG channel selection but it is greatly affected by the frequency spectrum and temporal window of the EEG segments. In [26], the time-constrained sparse group spatial pattern (TSGSP) method is introduced to improve CSP. The effects of frequency and temporal events are taken into account concurrently. A cross-correlation-based discriminant criterion (XCDC) is proposed in [27] to measure the significance of a channel for differentiating the mental states of various motor imagery (MI) tasks. The chi-square statistics is used to select a subset of channels in [28]. The sparse common spatial pattern (SCSP) approach is used to choose the MI-relevant channels that correlate to high sparse CSP filter coefficients [18]. Using the least absolute shrinkage and selection operator (LASSO) algorithm, the CSP-rank for multiple frequency bands (CSP-R-MF)-based approaches suggest in [29] to pick the MI-relevant channels for each frequency band.

The effective EEG channels (DCRCC) are identified in [30] using a dynamic channel relevance (DCR) score. The lowest redundancy maximum relevance paradigm is introduced for choosing the appropriate channels. The Savitzky–Golay filter is used to preprocess the EEG signals, and the sliding window method is employed to divide the resulting segments into predetermined lengths. Three previously identified channels are also considered as possible options for discovering additional related channels. The variance ratio dispersion score (VRDS) and interclass feature distance (ICFD) of small EEG channel groups are two metrics used in the local region CSP (LRFCS) method proposed in [31] to choose channels. The robust feature selection (RFS) algorithm used in the feature compression and channel ranking (FCCR) approach selects the MI-relevant channels by clustering the feature dimension [32]. These EEG channel selection methods rely on temporal, spectral, or spatial characteristics. It is used to classify MI tasks in relation to regionally different ERD/ERS patterns. However, when categorizing delicate and mixed ERD/ERS patterns, these traits may have limitations.

The existing methods for channel selection exhibit relatively poor classification performance because they often fail to accurately identify and eliminate all deceptive channels. In several studies, the number of optimized features is still large, which is less informative and error-prone. Techniques based on DCR score lack candidate channels, which causes performance degradation. In this paper, an entropy score used for effective EEG channel selection is proposed to improve MI classification performance. It mitigates the mentioned limitations of the existing methods. The entropy quantifies the pattern's unpredictability or uncertainty, which is roughly comparable to the signal's information content. High entropy means a more disordered pattern, whereas low entropy means a less disordered pattern with less information. The Shannon entropy is computed for each of the channels to obtain entropy scores. Then, a group of high-entropy channels for further processing are dynamically selected. Later, filtering techniques are applied to selected channels to decompose into different frequency bands. Using CSP, the spatial characteristics are extracted from each sub-band and fused to create the feature vector. The EEG features derived from the effective channels are fed to the support vector machine (SVM) classifier. The effectiveness of the suggested approach is assessed using an EEG dataset that is freely accessible. To demonstrate the superiority of the EEG channel selection methodology used in this study, the experimental findings are contrasted with the MI classification methods that are currently in use.

### 3. Data Description

#### 3.1. Dataset 1

A publicly available dataset, BCI Competition III (IVA), is used in this study to evaluate the proposed approach. The dataset was received from the Berlin BCI group. It includes the Campus Benjamin Franklin of the Charité, Department of Neurology, Neurophysics Group (Gabriel Curio), and Fraunhofer FIRST, Intelligent Data Analysis Group (Klaus-Robert Müller, Benjamin Blankertz). A detailed description of the dataset is available in [33]. Five healthy subjects (aa, al, aw, av, and ay) aged 24 to 25 completed MI tasks on each trial to record the dataset as EEG signals. They were instructed not to move their hands or arms during the recordings and not to move their eyes. The dataset consists of four sessions, each with 70 runs, for a total of 280 runs per subject. Out of 280 runs, 168, 224, 84, 56, and 28 runs for subjects aa, al, aw, av, and ay, respectively, made up the training set, whereas the remaining runs made up the test set. During the 3.5 s that the visual cues suggested, respondents were instructed to carry out one of two motor imagery tasks: right hand or right foot movement. One of the aforementioned activities was given to the users to consider, with goal cue intervals of 1.75 to 2.25 s, during which the subject might unwind. The 128-channel Ag/AgCl electrode cap was used to record the EEG signals. Positions were measured for 118 EEG channels using the worldwide 10/20 system. The bandpass filter was used to filter the measured MI-based EEG signals between 0.05 and 200 Hz, and the signals were subsequently digitized at 1000 Hz with a 16-bit precision. The dataset is downsampled at 100 Hz for further processing. Table 1 represents the summary of Dataset 1.

**Table 1.** Summary of the Dataset 1.

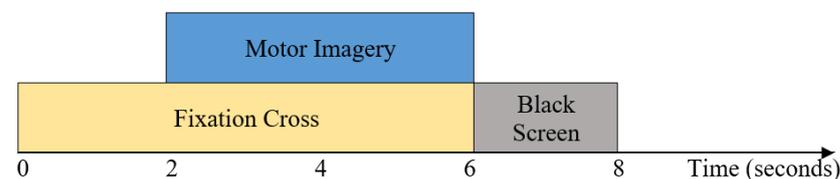
Dataset Name	BCI Competition III IVA
Subjects	aa, al, aw, av, ay
Number of Subjects	5
Channels	118
Sample Frequency	100 Hz
Classes	Left Hand, Right Hand, Foot
Tasks	Each subject completed any two of the classes
Visual Cues Duration	3.5 s
Total Trials	280 per subject (140 trials for each of the two classes)
Samples per Trial	350 (3.5 s × 100 Hz)

#### 3.2. Dataset 2

The second dataset used in this study is obtained from BCI Competition IV Dataset I [34], which is also publicly available. The summary of Dataset 2 is presented in Table 2. There are seven healthy subjects, namely a, b, c, d, e, f, and g. The EEG signals were recorded using 59 channels. The sampling frequency of Dataset 2 is 100 Hz. The left hand, right hand, and foot are the three classes of motor imagery tasks that are involved. Each subject completed any two of the classes. Four seconds of visual cues in the shape of arrows were displayed on the screen, and during that time, the subject completed a specific mental task. The signals were interspersed with two seconds of a blank screen and two seconds of a fixation cross. The fixation cross was superimposed on the cues and shown for six seconds. Thus, each trial lasts for 4 s, and then there is a 4 s rest interval. There are 200 trials in total for each subject and 100 trials for each of the two classes that are chosen by the particular subject. A trial lasts 4 s, and there are 400 samples in a trial (4 s × 100 Hz). Different parameters of Dataset 2 are illustrated in Figure 1.

**Table 2.** Summary of the Dataset 2.

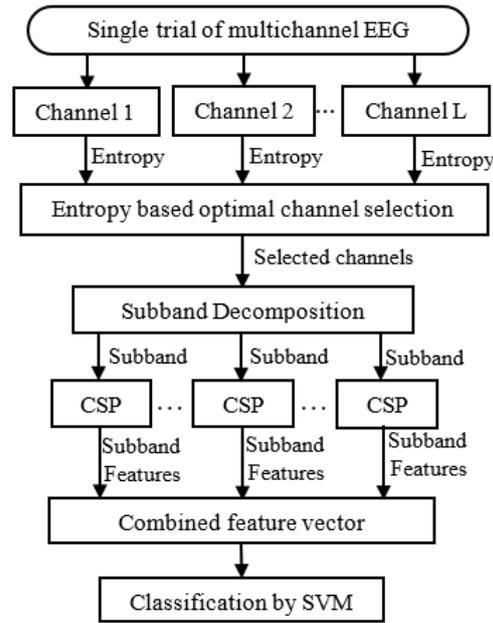
Dataset Name	BCI Competition IV Dataset I
Subjects	a, b, c, d, e, f, g
Number of Subjects	7
Channels	59
Sample Frequency	100 Hz
Classes	Left Hand, Right Hand, Foot
Tasks	Each subject completed any two of the classes
Visual Cues Duration	4 s
Total Trials	200 per subject (100 trials for each of the two classes)
Samples per Trial	400 (4 s × 100 Hz)

**Figure 1.** The timing sequence of BCI competition IV dataset I (Dataset 2).

#### 4. Methodology

The number of channels is increased day by day to enhance the spatial resolution of EEG-based BCI implementation. All the channels used in the recording of EEG are not effectively usable in different applications. The issue for selecting application-dependent effective channels is the demand to improve EEG classification performance. Such a method is implemented in this study. The methodology starts with the calculation of Shannon entropy for each channel in a trial. The entropy measure is used to identify channels having significant information. Based on the entropy scores, an optimal number of effective channels are selected for further analysis. Subsequently, the EEG signals with the selected channels are decomposed into distinct sub-bands. The narrowband approach enables the features of frequency-specific components of the EEG data. The sub-band trials are generated by aggregating similar sub-bands obtained from all selected channels. Common spatial pattern (CSP)-based features are then extracted from the sub-band trials. A feature vector is constructed by combining the features obtained from different sub-bands. The vector encapsulates the relevant information extracted from the EEG data. Finally, classification is performed using a support vector machine (SVM). The block diagram of the proposed channel-selection-based approach is illustrated in Figure 2. The following steps summarize the proposed method:

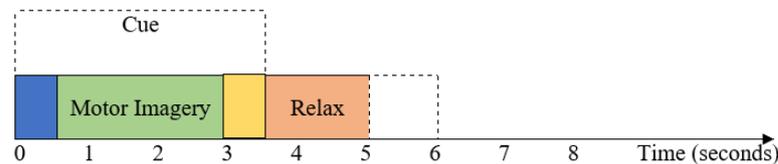
- The entropy of each channel of a trial is calculated. The Shannon entropy is used here.
- The obtained entropy score is used to select the adequate number of channels.
- The EEG signals of the selected channels are decomposed into a finite set of sub-bands using the Butterworth bandpass filter.
- Sub-band trial is generated by an accumulated similar sub-band obtained from all of the selected channels. Common spatial pattern (CSP)-based features are extracted from the sub-band trial.
- A feature vector is formulated by combining the features extracted from each sub-band trial.
- The classification is performed by implementing a support vector machine (SVM) along with the feature vector.



**Figure 2.** Block diagram of the proposed MI classification system based on channel selection.

#### 4.1. Data Preprocessing

For Dataset 1, the visual cues are illustrated for 3.5 s, as shown in Figure 3. The EEG signals from 0.5 s to 3 s after the cue are used in this study. Pre- and postimagination are thought to last for the first 0.5 s (0–0.5 s) and the last 0.5 s (3–3.5 s), respectively. For Dataset 2, the visual cues are illustrated for 4 s, as shown in Figure 1. The EEG signals from 2 s to 6 s after the cue are used in this study for Dataset 2. Then, the EEG signal of both datasets is passed through a fifth-order Butterworth bandpass filter with a frequency range from 8 Hz to 35 Hz.



**Figure 3.** The timing sequence of the BCI competition III dataset IVA (Dataset 1).

#### 4.2. Effective Channel Selection

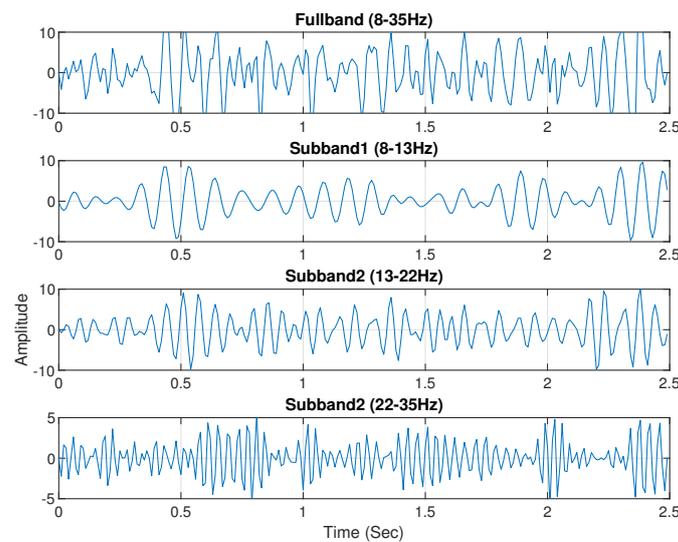
All the EEG channels are not equally effective for motor imagery (MI) classification. Efficient channel selection considerably improves the classification accuracy of MI tasks. An entropy-based information-theoretic approach is implemented here to select the effective EEG channels. The multichannel EEG is denoted as  $X_l^m = [x_l^m(1)x_l^m(2) \dots x_l^m(N)]^T$ , where  $m$  is the trial index ( $m = 1, 2, \dots, M$ ),  $l$  is the channel index ( $l = 1, 2, \dots, L$ ), and  $n$  is the index of sample ( $n = 1, 2, \dots, N$ ). The entropy measures the information of the pattern [35–39]. The Shannon entropy score is used in this study to identify the effective EEG channels. For a specific single subject, the average entropy of the  $l$ -th channel over  $M$  trials is defined as

$$h_l = -\frac{1}{M} \sum_{m=1}^M \sum_{n=1}^N p(x_l^m(n)) \log p(x_l^m(n)), \quad (1)$$

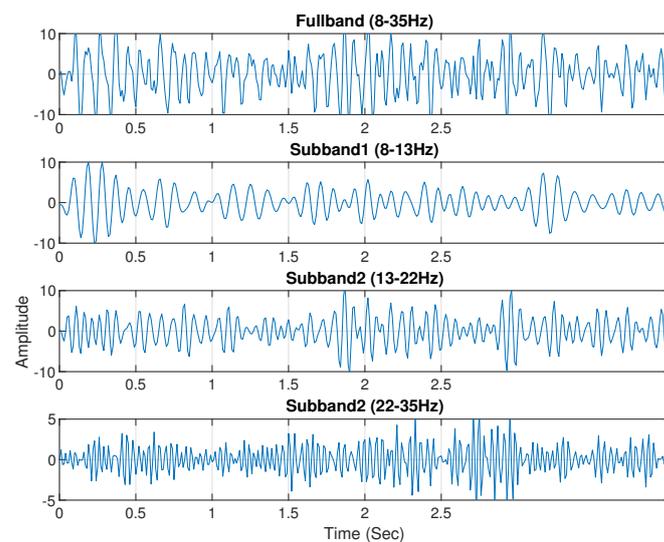
where  $p(x_l^m(n))$  is the probability of the occurrence of the  $n$ -th sample on the  $l$ -th channel of the  $m$ -th trial. The entropy vector  $H = [h_1, h_2, \dots, h_l, \dots, h_L]$  is sorted in descending order, and the channels with higher entropy scores are chosen as the effective channel to improve MI classification performance. The EEG data of reduced channels consisting of the selected channels only are denoted by  $X'$ .

### 4.3. Sub-Band Decomposition

Electrophysiological noises frequently contaminate the multichannel raw EEG. Furthermore, some narrowband components of the EEG signal respond more strongly to the specific MI task. Therefore, the use of sub-band signals of selected EEG channels would be more accurate in classifying MI activities. According to a previous study, the majority of brain activities connected to MI tasks take place at the frequency range of 7–30 Hz [40,41]. Then, the EEG signal is split into four sub-bands using a zero-phase Butterworth bandpass filter. The extracted four sub-bands are Mu (8–13 Hz), low beta (13–22 Hz), high beta (22–35 Hz), and full-band (8–35 Hz). Figure 4 and 5 shows the sub-bands of channel C3 for subject ‘a’ of the BCI Competition III (IVA) dataset and BCI Competition IV (I) dataset.



**Figure 4.** Channel C3 of Subject ‘a’ from the BCI Competition III IVA dataset yielded four unique sub-bands: 8–35 Hz, 8–13 Hz, 13–22 Hz, and 22–35 Hz.



**Figure 5.** The four sub-bands: 8–35 Hz, 8–13 Hz, 13–22 Hz, and 22–35 Hz of channel C3 selected from Subject ‘a’ of BCI Competition IV dataset I.

### 4.4. Feature Extraction

The potential feature extraction is one of the most important steps in BCI implementation. The features have a direct influence on the BCI system performance. Recent studies

have looked at modifying or developing feature extraction algorithms [42]. The most popular and well-known feature extraction technique in multichannel EEG-based BCI systems is the CSP [41,43,44]. Multichannel EEG data are projected onto a low-dimensional spatial subspace using the CSP technique. The CSP determines spatial filters that maximize the variance ratio between two classes of data. The filter serves to extract effective features for classification. The most discriminative spatial information between the MI tasks to execute the effective classification of EEG signals is applicable in BCI implementation. Let  $X'_{c_1}$  and  $X'_{c_2}$  denote the training trials of EEG signals with selected effective channels from class  $c_1$  and  $c_2$ , respectively. The CSP method derives the features by simultaneously diagonalizing both classes' covariance matrices. The EEG data are transformed into a projection matrix using a spatial filter  $w \in R^l$  that maximizes the variances between two classes [45].

$$w = \arg \max \frac{w^T B_1 w}{w^T B_2 w} \quad \text{s.t. } \|w\|_2 = 1. \quad (2)$$

By resolving an extended eigenvalue problem, the optimal solution to Equation (2) is achieved. The largest and lowest eigenvalues  $V$  generated by the eigenvectors are included in the spatial filter  $w = [w_1, w_2, \dots, w_{2V}] \in R^{l \times 2V}$ . Thus, the feature vector of any trial of sub-band  $s$  is constructed as  $f^s = [f_1^{(s)}, f_2^{(s)}, \dots, f_{2V}^{(s)}]$ . The log-transformation is used to normalize the elements of  $f_v^{(s)}$

$$f_v^{(s)} = \log\left(\text{var}\left(w_v^T x\right)\right), \quad v = 1, 2, \dots, 2V, \quad (3)$$

where  $\text{var}(\cdot)$  represents the variance. The CSP is used to extract a subset of features from each sub-band. The features obtained from each sub-band are concatenated as  $F = [f^1, f^2, \dots, f^s]$  to create the feature vector for the particular trial, where  $f^s$  represents the features of  $s$  sub-band and  $s = 1, 2, \dots, S$  ( $S$  represents all sub-bands together). The feature vector  $F$  is used to classify the MI tasks using SVM.

#### 4.5. Classification

The proposed channel selection approach to MI Classification is evaluated using a support vector machine (SVM) classifier with a radial basis function (RBF) kernel. The effectiveness of the SVM is also compared with a linear discriminant analysis (LDA) classifier and an artificial neural network (ANN). The ANN is implemented with one hidden layer of size 10, sigmoid activation function, and one output layer of size 2 with linear function. The detailed description of SVM, LDA, and ANN is available here [46], [47], and [48], respectively.

### 5. Experimental Results

The performance of the proposed channel-selection-based method is assessed using the BCI competition III (IVA) dataset and the BCI competition IV(I) dataset. Right hand and right foot movement are two motor imagery tasks that are considered for classification. The signals of each channel of EEG data are decomposed into four sub-bands described in the Sub-band Decomposition Section. The CSP is applied to each sub-band of EEG signals to extract spatial information. From each of the four sub-bands, two sets of spatial filter pairs are extracted, producing four features. Thus 16 (4 feature  $\times$  4 sub-band)-dimensional feature vector is derived to feed to the classifier. The classification accuracy of the individual subject is evaluated using the  $k$ -fold ( $k = 5$ ) cross-validation approach. The dataset is randomly divided into  $k$  equal groups. One group is designated for testing, while the rest is for training. The procedure is carried out over  $k$  times. The classification accuracy specified in Equation (4) is obtained by averaging the outcomes of the  $k$  iterations. With the use of the BCI Competition III (IVA) and IV (I) datasets, several experiments

are conducted to illustrate the performance of the proposed approach method. Several performance metrics, Accuracy, Precision, Recall, and F1 score, are computed as follows:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \quad (4)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (5)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (6)$$

$$\text{F1 score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (7)$$

where TP, TN, FP, and FN stand for true positive, true negative, false positive, and false negative, respectively. The classification performance of Dataset 1 and 2 with various classifiers is summarized in Table 3 and Table 4, respectively.

**Table 3.** Performance of the proposed method using SVM, ANN, and LDA classifiers in terms of classification accuracy (Acc.), precision (Prec.), recall (Rec.), and F1 score (F1) of dataset 1. Boldface denotes the best result.

Subjects	No. of Channels	Classifier											
		SVM				ANN				LDA			
		Acc.	Prec.	Rec.	F1	Acc.	Prec.	Rec.	F1	Acc.	Prec.	Rec.	F1
aa	30	<b>86.43</b>	<b>84.96</b>	<b>87.14</b>	<b>86.04</b>	83.75	84.38	82.16	83.26	81.43	80.73	82.29	81.50
al	62	97.86	97.24	<b>98.56</b>	97.90	98.14	<b>97.56</b>	97.67	97.61	<b>98.21</b>	97.29	99.29	<b>98.28</b>
av	67	<b>78.93</b>	<b>78.64</b>	<b>77.96</b>	<b>78.30</b>	71.21	73.12	65.04	68.84	68.57	68.12	72.57	70.27
aw	63	<b>97.86</b>	<b>96.67</b>	<b>97.88</b>	<b>97.27</b>	96.43	95.33	97.19	96.25	96.07	96.55	<b>97.86</b>	97.20
ay	59	95.72	93.98	95.36	94.67	94.07	94.16	<b>95.56</b>	94.85	<b>96.07</b>	<b>95.60</b>	95.00	<b>95.30</b>
<b>Mean</b>		<b>91.36</b>	<b>90.30</b>	<b>91.38</b>	<b>90.83</b>	91.08	88.91	87.52	88.16	88.07	87.66	89.40	88.51

**Table 4.** Performance of the proposed method using SVM, ANN, and LDA classifiers in terms of classification accuracy (Acc.), precision (Prec.), recall (Rec.), and F1 score (F1) of dataset 2. Boldface denotes the best result.

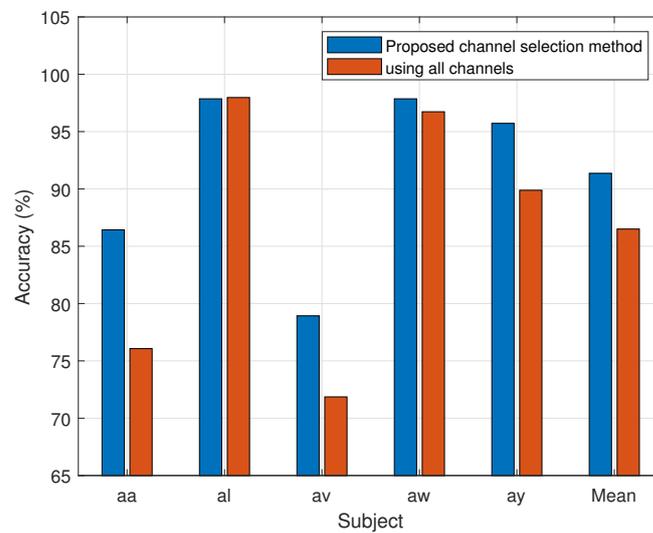
Subjects	No. of Channels	Classifier											
		SVM				ANN				LDA			
		Acc.	Prec.	Rec.	F1	Acc.	Prec.	Rec.	F1	Acc.	Prec.	Rec.	F1
a	23	<b>80.00</b>	78.30	78.78	78.54	79.00	73.08	<b>78.89</b>	75.87	78.50	<b>79.41</b>	76.43	77.89
b	31	<b>67.00</b>	<b>67.00</b>	<b>67.00</b>	<b>67.00</b>	61.80	62.44	61.00	61.71	59.50	58.88	62.00	60.40
c	41	<b>80.00</b>	<b>79.45</b>	<b>81.00</b>	<b>80.22</b>	73.90	74.44	74.20	74.32	72.00	71.43	74.00	72.69
d	46	<b>91.00</b>	<b>90.41</b>	91.00	<b>90.7</b>	88.20	87.79	89.10	88.44	89.50	87.38	<b>93.00</b>	90.10
e	39	90.00	90.68	90.00	90.34	<b>91.00</b>	<b>91.88</b>	90.00	90.93	<b>91.00</b>	90.04	<b>92.60</b>	<b>91.29</b>
f	31	<b>91.00</b>	<b>90.15</b>	<b>90.00</b>	<b>90.07</b>	85.95	88.69	83.00	85.75	86.50	84.75	89.00	86.82
g	17	<b>80.50</b>	79.25	<b>82.00</b>	<b>80.6</b>	79.50	79.39	80.40	79.89	79.50	<b>83.27</b>	77.00	80.01
<b>Mean</b>		<b>82.79</b>	<b>82.18</b>	<b>82.83</b>	<b>82.50</b>	79.91	79.67	79.51	79.56	79.50	79.31	80.58	79.89

Table 3 illustrates that the SVM achieves the highest classification accuracy for individual subjects except 'al' and 'ay'. The LDA performs better than that of SVM with these two subjects. The average classification accuracy of the SVM classifier outperforms ANN and LDA by 2.64% and 3.29%, respectively. Table 3 also demonstrates that for most of the cases, the F1 score, recall, and precision of the SVM classifier achieve higher performance. The mean value of the SVM classifier's F1 score, recall, and precision is higher than the ANN and LDA classifiers. For subjects 'aa', 'av', and 'aw', SVM achieves the highest accuracy, F1 score, recall, and precision performance than ANN and LDA. The ANN achieves the highest recall for the subject 'ay', and the LDA obtains the highest accuracy,

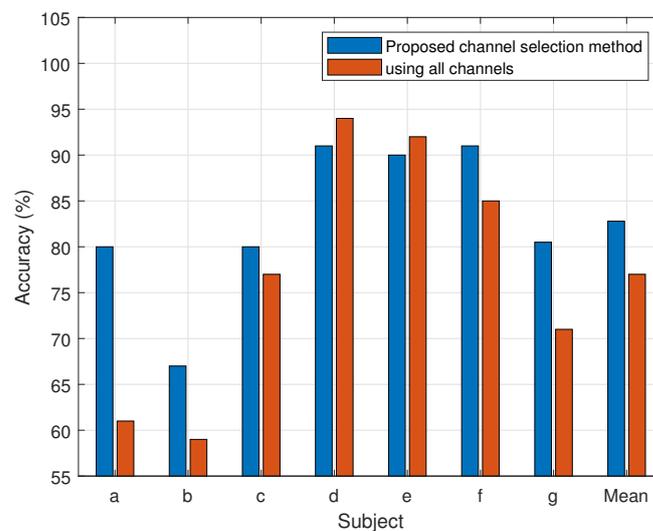
precision, and F1 score for the subject 'ay'. Table 4 illustrates that SVM achieves the highest classification accuracy across all the subjects except 'e'. It is noted that the best accuracy of subject 'e' is achieved by both ANN and LDA. The mean accuracy of SVM is 2.88% and 3.29% higher than that of ANN and LDA. The SVM classifier demonstrates superior performance compared with both the ANN and LDA classifiers in terms of mean accuracy, precision, recall, and F1 score across all the subjects. The implementation of the method is available at <https://github.com/Shiam-Cse-Ru/EEG> (accessed on 17 April 2024).

## 6. Discussion

Figures 6 and 7 show the results of dataset 1 and 2, respectively, for the individual subjects in terms of classification accuracy for MI-BCI utilizing all channels, as well as the proposed channel selection method. All the channels are used to implement the conventional CSP-SVM technique. Performance with all channels is consistently worse than that of the proposed method. It has been noted that the channel selection strategy performs better than the average accuracy across all subjects and for each subject individually.



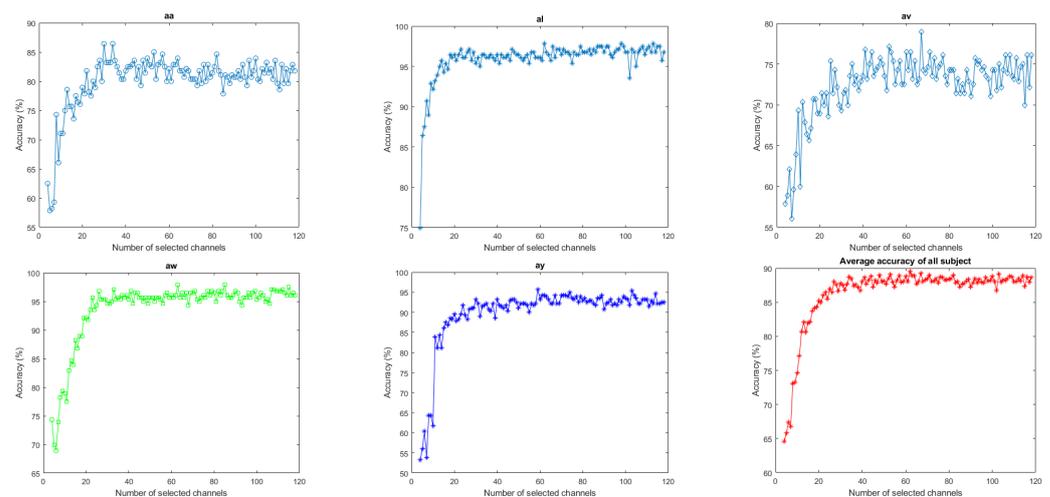
**Figure 6.** Comparison of the proposed channel selection method's performance (accuracy) vs. all channels of Dataset 1.



**Figure 7.** Comparison of the proposed channel selection method's performance (accuracy) vs. all channels of Dataset 2.

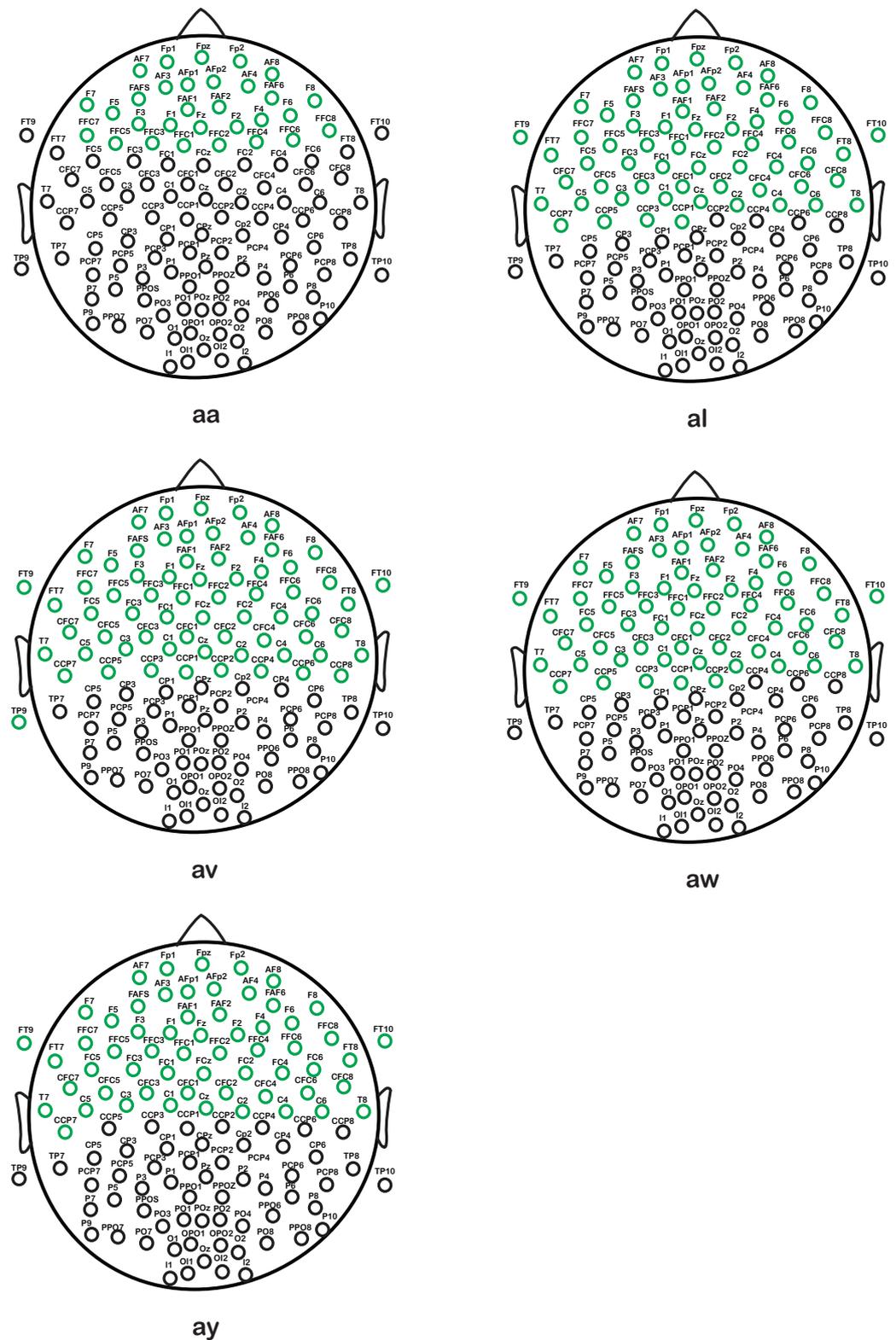
The proposed approach selects the least possible channels for each subject with maximum performance. The results of the experiment show that higher accuracy is achieved for a subject using fewer effective channels (see Figures 6 and 7). The average classification accuracy of the proposed channel selection method with SVM is 4.86% greater than that of the use of all channels. In [49,50], the 30 channels are manually selected to classify the MI tasks. The information represented by EEG is very much subject-dependent. A uniform method of manual channel selection does not guarantee effectiveness in an intersubject scenario. The EEG data and subject-dependent channel selection approach are supposed to work better for MI task classification. Considering the underlying neurological assumptions, an entropy-based approach is implemented in this study to select an effective channel for BCI implementation. It is based on subject-specific EEG data.

The number of channels used for the MI classification is one of the crucial variables affecting accuracy. Individual subject performance is shown in Figure 8 as a function of the number of effective channels selected by the proposed entropy-based approach. Although the classification accuracy of the majority of subjects initially improved with the number of channels, whereas the end performance is lower than the best accuracy reached with an optimal number of selected channels. It occurs because the use of fewer channels initially includes fewer features, leading to lower accuracy. The inclusion of a significantly higher number of channels of redundant information reduces the accuracy. With an effective number of selected channels, the highest accuracy is attained for each subject. In this experiment, the number of channels with the highest average subject-level accuracy is taken into account.



**Figure 8.** Individual subject performance (classification accuracy) and their average as a function of a number of selected channels (using an SVM classifier) of Dataset 1.

Figure 9 shows the selected effective channels by the proposed channel selection approach. It demonstrates how different combinations of channels are used to recognize the MI tasks. The distribution of the selected channels in the cerebral cortex's motor regions is evident from Figure 9. The selected channels are 30, 62, 67, 63, and 59 for the subjects "aa", "al", "av", "aw", and "ay", respectively, as shown in Figure 9. With the aforementioned carefully selected channels, the proposed method achieves the maximum classification accuracy.



**Figure 9.** Distribution of the selected effective channels (green circles) of individual subjects of Dataset 1.

The performance of the proposed method is compared with recent channel selection techniques, as illustrated in Tables 5 and 6. The number of selected channels is indicated within the parentheses. It is noted that individual methods have selected different numbers of channels for different subjects. The challenging task is to fix up for the classification of

motor imagery tasks with optimal performance. In comparison with previous techniques, the proposed method performs better in terms of accuracy across all the subjects. It is compared with the recently developed channel selection algorithm, such as dynamic channel relevance [30], channel selection using a correlation coefficient [25], CSP-R-MF [29], LRFCSF [31], CCS-RCSP [23], and FCCR [32] to understand its superiority. The dynamic channel relevance (DCR) score was utilized to find the effective EEG channels in [30] with a mean classification accuracy of 87.58%. The proposed method achieves the highest mean classification accuracy (91.36%) compared with other recently developed methods.

**Table 5.** Performance comparison in terms of accuracy (%) with Dataset 1 using different cutting-edge channel-selection-based methods. Bold font is used to indicate the best outcome.

Methods	Subjects					Mean $\pm$ STD
	Accuracy (Number of Selected Channels)					
	aa	al	av	aw	ay	
stdWC [51]	84.50 (16)	98.10 (19)	72.80 (12)	95.10 (13)	92.50 (20)	88.80 $\pm$ 9.10
DCRCC [30]	<b>93.60 (24)</b>	79.2 (33)	<b>94.6 (11)</b>	85.54 (26)	84.94 (31)	87.58 $\pm$ 6.46
CSCC [25]	89.29 (10)	<b>98.21 (7)</b>	73.47 (12)	92.86 (9)	89.29 (7)	88.62 $\pm$ 9.22
CSP-R-MF [29]	82.14 (12)	96.42 (12)	72.14 (12)	84.38 (12)	94.28 (12)	85.87 $\pm$ 9.83
LRFCSF [31]	83.93 (22)	96.42 (7)	74.49 (6)	88.84 (7)	89.29 (11)	86.59 $\pm$ 8.10
CCS-RCSP [23]	83.03 (42)	96.42 (33)	70.91 (52)	92.41 (14)	92.46 (67)	87.05 $\pm$ 10.28
FCCR [32]	78.57 (10)	<b>98.21 (10)</b>	72.45 (5)	87.05 (15)	93.25 (9)	85.90 $\pm$ 10.51
<b>Proposed Method</b>	86.36 (60)	97.51 (60)	76.20 (60)	<b>95.81 (60)</b>	<b>95.38 (60)</b>	<b>90.25 <math>\pm</math> 8.98</b>

The correlation coefficient is used in [25] for selecting distinctive channels and achieved 88.62 % mean classification accuracy. Using the efficient channel selection strategy, the proposed method improves noticeable accuracy. It has been noted that the CCS-RCSP technique obtains a mean accuracy of 87.05% using correlation coefficients to pick MI-related channels. The technique described in [29], CSP-rank for multiple frequency bands (CSP-R-MF), selects the MI-relevant channels for each frequency band. They extracted features from those channels using the least absolute shrinkage and selection operator (LASSO) technique and obtained a mean accuracy of 85.87%. The proposed method outperforms the methods described in [25,30,51], CSP-R-MF [29], LRFCSF [31], CCS-RCSP [23], and FCCR [32], by 2.56%, 3.78%, 2.74%, 5.49%, 4.77%, 4.31%, and 5.46%, respectively, in terms of classification accuracy using Dataset 1. Table 6 shows the comparative performance of proposed method with recently developed methods using Dataset 2. It shows that the methods FCCR and CSP-R-MF consider only 'a', 'b', 'f', and 'g' subjects, while others consider all the subjects. The mean accuracy of the proposed method outperforms all the recently developed methods, as shown in Table 6. Considering the overall performance of MI task classification, the proposed information-theoretic EEG channel selection method is effective for BCI implementation. The traditional machine learning approaches SVM, LDA, and ANN are employed in this study rather than the deep learning methods. The proposed approach exhibits several advantages. The mentioned traditional methods are suitable for limited training data and less prone to overfitting. Less training data is more preferable for BCI implementation. The deep learning approach sometimes results in better performance, but it requires higher computational power and a high volume of training data. Such requirements make deep learning less preferable in the BCI paradigm. Additionally, this study focuses on effective EEG channel selection to enhance the MI classification performance. Thus, the proposed method provides a pragmatic and effective approach to implementing an effective BCI system. The experimental findings exhibit the effectiveness of the proposed approach to EEG classification. The results can lead to a broader acceptance and incorporation of BCIs into daily life applications like communication, rehabilitation, and other areas. The main limitation of the proposed channel-selection-based method is

that it is subject-dependent and applicable to classify only two MI tasks. It is not tested with the data collected in a real real-world application environment.

**Table 6.** Performance comparison in terms of accuracy (%) with Dataset 2 using different cutting-edge channel-selection-based methods. Bold font is used to indicate the best outcome.

Methods	Subjects							Mean $\pm$ STD
	Accuracy (Number of Selected Channels)							
	a	b	c	d	e	f	g	
Improved SFFS [52]	69.00 (6)	63.00 (15)	87.00 (26)	94.00 (29)	<b>96.00 (19)</b>	65.00 (8)	72.00 (22)	78.00 $\pm$ 14.00
FCCR [32]	77.00 (14)	71.00 (9)				76.50 (9)	77.00 (18)	75.36 $\pm$ 2.53
GSFS [53]	75.00 (6)	<b>72.00 (13)</b>	<b>89.00 (11)</b>	78.00 (9)	84.00 (14)	78.00 (15)	83.00 (12)	79.80 $\pm$ 5.38
CSP-R-MF [41]	<b>81.50 (24)</b>	63.00 (24)				79.00 (24)	<b>87.50 (24)</b>	77.75 $\pm$ 9.06
CSRI [54]	72.80 (8)	66.18 (12)	83.72 (26)	<b>96.10 (13)</b>	83.50 (15)	76.33 (14)	87.33 (17)	80.85 $\pm$ 9.18
Proposed Method	80.00 (23)	67.00 (31)	80.00 (41)	91.00 (46)	90.00 (39)	<b>91.00 (31)</b>	80.00 (17)	<b>82.79 <math>\pm</math> 8.73</b>

## 7. Conclusions

In this study, an effective EEG channel selection approach is introduced to enhance motor imagery classification performance. The aim of the proposed method is to reduce redundant information by selecting effective channels. The use of a higher number of irrelevant EEG channels can increase the redundant information, which may cause the degraded performance of MI task classification. Effective channel selection is an efficient way to overcome the mentioned problems, leading to improved EEG classification accuracy. After the selection of potential channels, the EEG signals are decomposed into four subbands, and CSP-based features are extracted from individual sub-bands, yielding multiband features. The feature vector is derived by combining the features obtained from the sub-band signals. An SVM is employed to classify the MI tasks using the obtained feature vector. The proposed system improves the binary MI classification performance with two publicly available datasets compared with the state-of-the-art methods. The future extension of this work is to develop a method for subject-independent EEG channel selection and implementation for the multiclass MI classification paradigm.

**Author Contributions:** Conceptualization, A.A.S. and K.M.H.; methodology, A.A.S. and K.M.H.; experiment; A.A.S. and K.M.H.; validation, A.A.S., K.M.H. and M.R.I.; formal analysis, A.A.S., K.M.H., A.M.M.A. and M.K.I.M.; investigation, A.A.S. and K.M.H.; data curation, A.A.S., K.M.H. and A.M.M.A.; writing—original draft preparation, A.A.S. and K.M.H.; writing—review and editing, A.A.S., H.W. and M.K.I.M.; visualization, A.A.S., K.M.H. and M.R.I.; supervision, M.K.I.M. and H.W.; funding acquisition, H.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by JSPS KAKENHI Grant Numbers JP17H06383, JP24K07387.

**Institutional Review Board Statement:** Ethical review and approval were waived for this study, due to the use of publicly available datasets.

**Informed Consent Statement:** This study used publicly available dataset where informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The proposed model is evaluated with a benchmark dataset, namely the BCI Competition III and IV (IVA and I) datasets, which are publicly available. The dataset link is provided here [www.bbc.de/competition/iii/](http://www.bbc.de/competition/iii/) (accessed on 17 April 2024) and [www.bbc.de/competition/iv/](http://www.bbc.de/competition/iv/) (accessed on 17 April 2024).

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

## Abbreviations

The following abbreviations are used in this manuscript:

fMRI	Functional Magnetic Response Imaging
MEG	Magnetoencephalogram
EOG	Electrooculography
ERD	Event-Related Desynchronization
ERS	Event-related synchronization
CSP	Common Spatial Pattern
SVM	Support Vector Machine
ANN	Artificial Neural Network
LDA	Linear Discriminant Analysis
BSS	Blind Source Separation
RCSP	Regularized Common Spatial Pattern
RBF	Radial Basis Function
FBCSP	Filter-Bank Common Spatial Pattern
TSGSP	Time Constrained Sparse Group Spatial Pattern
SCSP	Sparse Common Spatial Pattern
DCR	Dynamic Channel Relevance

## References

1. Wolpaw, J.R.; Birbaumer, N.; Heetderks, W.J.; McFarland, D.J.; Peckham, P.H.; Schalk, G.; Donchin, E.; Quatrano, L.A.; Robinson, C.J.; Vaughan, T.M.; et al. Brain-computer interface technology: A review of the first international meeting. *IEEE Trans. Rehabil. Eng.* **2000**, *8*, 164–173. [[CrossRef](#)] [[PubMed](#)]
2. Miah, A.S.M.; Islam, M.R.; Molla, M.K.I. Motor imagery classification using subband tangent space mapping. In Proceedings of the 2017 20th International Conference of Computer and Information Technology (ICIT), Dhaka, Bangladesh, 22–24 December 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 1–5.
3. Wolpaw, J.R.; McFarland, D.J.; Neat, G.W.; Forneris, C.A. An EEG-based brain-computer interface for cursor control. *Electroencephalogr. Clin. Neurophysiol.* **1991**, *78*, 252–259. [[CrossRef](#)] [[PubMed](#)]
4. Gaur, P.; Pachori, R.B.; Wang, H.; Prasad, G. A multi-class EEG-based BCI classification using multivariate empirical mode decomposition based filtering and Riemannian geometry. *Expert Syst. Appl.* **2018**, *95*, 201–211. [[CrossRef](#)]
5. Li, D.; Wang, J.; Xu, J.; Fang, X. Densely feature fusion based on convolutional neural networks for motor imagery EEG classification. *IEEE Access* **2019**, *7*, 132720–132730. [[CrossRef](#)]
6. Miah, A.S.M.; Islam, M.R.; Molla, M.K.I. EEG classification for MI-BCI using CSP with averaging covariance matrices: An experimental study. In Proceedings of the 2019 International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2), Rajshahi, Bangladesh, 11–12 July 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 1–5.
7. Miah, A.S.M.; Ahmed, S.R.A.; Ahmed, M.R.; Bayat, O.; Duru, A.D.; Molla, M.K.I. Motor-Imagery BCI task classification using riemannian geometry and averaging with mean absolute deviation. In Proceedings of the 2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT), Istanbul, Turkey, 24–26 April 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 1–7.
8. Miah, A.S.M.; Rahim, M.A.; Shin, J. Motor-imagery classification using riemannian geometry with median absolute deviation. *Electronics* **2020**, *9*, 1584. [[CrossRef](#)]
9. Pfurtscheller, G.; Da Silva, F.L. Event-related EEG/MEG synchronization and desynchronization: Basic principles. *Clin. Neurophysiol.* **1999**, *110*, 1842–1857. [[CrossRef](#)] [[PubMed](#)]
10. Pfurtscheller, G.; Neuper, C. Motor imagery and direct brain-computer communication. *Proc. IEEE* **2001**, *89*, 1123–1134. [[CrossRef](#)]
11. Miah, A.S.M.; Mouly, M.A.; Debnath, C.; Shin, J.; Sadakatul Bari, S. Event-Related Potential Classification based on EEG data using xDWT with MDM and KNN. In Proceedings of the Computing Science, Communication and Security: Second International Conference, COMS2 2021, Gujarat, India, 6–7 February 2021; Revised Selected Papers; Springer: Berlin/Heidelberg, Germany, 2021; pp. 112–126.
12. Joy, M.M.H.; Hasan, M.; Miah, A.S.M.; Ahmed, A.; Tohfa, S.A.; Bhuaiyan, M.F.I.; Zannat, A.; Rashid, M.M. Multiclass mi-task classification using logistic regression and filter bank common spatial patterns. In Proceedings of the Computing Science, Communication and Security: First International Conference, COMS2 2020, Gujarat, India, 26–27 March 2020; Revised Selected Papers; Springer: Berlin/Heidelberg, Germany, 2020; pp. 160–170.
13. Miah, A.S.M.; Shin, J.; Hasan, M.A.M.; Molla, M.K.I.; Okuyama, Y.; Tomioka, Y. Movie Oriented Positive Negative Emotion Classification from EEG Signal using Wavelet transformation and Machine learning Approaches. In Proceedings of the 2022 IEEE 15th International Symposium on Embedded Multicore/Many-core Systems-on-Chip (MCSoc), Penang, Malaysia, 19–22 December 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 26–31.
14. Kabir, M.H.; Mahmood, S.; Al Shiam, A.; Musa Miah, A.S.; Shin, J.; Molla, M.K.I. Investigating Feature Selection Techniques to Enhance the Performance of EEG-Based Motor Imagery Tasks Classification. *Mathematics* **2023**, *11*, 1921. [[CrossRef](#)]

15. Miah, A.S.M.; Shin, J.; Islam, M.M.; Molla, M.K.I. Natural Human Emotion Recognition Based on Various Mixed Reality (MR) Games and Electroencephalography (EEG) Signals. In Proceedings of the 2022 IEEE 5th Eurasian Conference on Educational Innovation (ECEI), Taipei, Taiwan, 10–12 February 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 408–411.
16. Zobaed, T.; Ahmed, S.R.A.; Miah, A.S.M.; Binta, S.M.; Ahmed, M.R.A.; Rashid, M. Real time sleep onset detection from single channel EEG signal using block sample entropy. In Proceedings of the IOP Conference Series: Materials Science and Engineering, Thi-Qar, Iraq, 15–16 July 2020; IOP Publishing: Bristol, UK, 2020; Volume 928, p. 032021.
17. Zhang, X.; Yao, L.; Wang, X.; Monaghan, J.; Mcalpine, D.; Zhang, Y. A survey on deep learning-based non-invasive brain signals: recent advances and new frontiers. *J. Neural Eng.* **2021**, *18*, 031002. [[CrossRef](#)]
18. Arvaneh, M.; Guan, C.; Ang, K.K.; Quek, C. Optimizing the channel selection and classification accuracy in EEG-based BCI. *IEEE Trans. Biomed. Eng.* **2011**, *58*, 1865–1873. [[CrossRef](#)]
19. Qi, F.; Wu, W.; Yu, Z.L.; Gu, Z.; Wen, Z.; Yu, T.; Li, Y. Spatiotemporal-filtering-based channel selection for single-trial EEG classification. *IEEE Trans. Cybern.* **2020**, *51*, 558–567. [[CrossRef](#)] [[PubMed](#)]
20. Alotaiby, T.; El-Samie, F.E.A.; Alshebeili, S.A.; Ahmad, I. A review of channel selection algorithms for EEG signal processing. *EURASIP J. Adv. Signal Process.* **2015**, *2015*, 66. [[CrossRef](#)]
21. Cai, S.; Li, H.; Wu, Q.; Liu, J.; Zhang, Y. Motor imagery decoding in the presence of distraction using graph sequence neural networks. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2022**, *30*, 1716–1726. [[CrossRef](#)] [[PubMed](#)]
22. Sannelli, C.; Dickhaus, T.; Halder, S.; Hammer, E.M.; Müller, K.R.; Blankertz, B. On optimal channel configurations for SMR-based brain–computer interfaces. *Brain Topogr.* **2010**, *23*, 186–193. [[CrossRef](#)]
23. Jin, J.; Miao, Y.; Daly, I.; Zuo, C.; Hu, D.; Cichocki, A. Correlation-based channel selection and regularized feature optimization for MI-based BCI. *Neural Netw.* **2019**, *118*, 262–270. [[CrossRef](#)] [[PubMed](#)]
24. Ghaemi, A.; Rashedi, E.; Pourrahimi, A.M.; Kamandar, M.; Rahdari, F. Automatic channel selection in EEG signals for classification of left or right hand movement in Brain Computer Interfaces using improved binary gravitation search algorithm. *Biomed. Signal Process. Control* **2017**, *33*, 109–118. [[CrossRef](#)]
25. Park, Y.; Chung, W. Optimal channel selection using correlation coefficient for CSP based EEG classification. *IEEE Access* **2020**, *8*, 111514–111521. [[CrossRef](#)]
26. Zhang, Y.; Nam, C.S.; Zhou, G.; Jin, J.; Wang, X.; Cichocki, A. Temporally constrained sparse group spatial patterns for motor imagery BCI. *IEEE Trans. Cybern.* **2018**, *49*, 3322–3332. [[CrossRef](#)] [[PubMed](#)]
27. Yu, J.; Yu, Z.L. Cross-correlation based discriminant criterion for channel selection in motor imagery BCI systems. *J. Neural Eng.* **2021**, *18*, 046083. [[CrossRef](#)]
28. Baig, M.Z.; Aslam, N.; Shum, H.P. Filtering techniques for channel selection in motor imagery EEG applications: A survey. *Artif. Intell. Rev.* **2020**, *53*, 1207–1232. [[CrossRef](#)]
29. Feng, J.K.; Jin, J.; Daly, I.; Zhou, J.; Niu, Y.; Wang, X.; Cichocki, A. An optimized channel selection method based on multifrequency CSP-rank for motor imagery-based BCI system. *Comput. Intell. Neurosci.* **2019**, *2019*, 8068357. [[CrossRef](#)]
30. Tiwari, A.; Chaturvedi, A. A novel channel selection method for BCI classification using dynamic channel relevance. *IEEE Access* **2021**, *9*, 126698–126716. [[CrossRef](#)]
31. Park, Y.; Chung, W. Frequency-optimized local region common spatial pattern approach for motor imagery classification. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2019**, *27*, 1378–1388. [[CrossRef](#)] [[PubMed](#)]
32. Han, J.; Zhao, Y.; Sun, H.; Chen, J.; Ke, A.; Xu, G.; Zhang, H.; Zhou, J.; Wang, C. A fast, open EEG classification framework based on feature compression and channel ranking. *Front. Neurosci.* **2018**, *12*, 217. [[CrossRef](#)] [[PubMed](#)]
33. Dornhege, G.; Blankertz, B.; Curio, G.; Müller, K.R. Boosting bit rates in noninvasive EEG single-trial classifications by feature combination and multiclass paradigms. *IEEE Trans. Biomed. Eng.* **2004**, *51*, 993–1002. [[CrossRef](#)] [[PubMed](#)]
34. BBCI Competition IV—Motor Imagery Data Set. Available online: <https://www.bbci.de/competition/iv/> (accessed on 17 April 2024).
35. Zarei, A.; Asl, B.M. Automatic seizure detection using orthogonal matching pursuit, discrete wavelet transform, and entropy based features of EEG signals. *Comput. Biol. Med.* **2021**, *131*, 104250. [[CrossRef](#)] [[PubMed](#)]
36. Roy, G.; Bhoi, A.; Bhaumik, S. A comparative approach for MI-based EEG signals classification using energy, power and entropy. *IRBM* **2022**, *43*, 434–446. [[CrossRef](#)]
37. Catherine Joy, R.; Thomas George, S.; Albert Rajan, A.; Subathra, M. Detection of ADHD from EEG signals using different entropy measures and ANN. *Clin. EEG Neurosci.* **2022**, *53*, 12–23. [[CrossRef](#)] [[PubMed](#)]
38. Ra, J.S.; Li, T.; Li, Y. A novel permutation entropy-based EEG channel selection for improving epileptic seizure prediction. *Sensors* **2021**, *21*, 7972. [[CrossRef](#)]
39. Acharya, U.R.; Fujita, H.; Sudarshan, V.K.; Bhat, S.; Koh, J.E. Application of entropies for automated diagnosis of epilepsy using EEG signals: A review. *Knowl.-Based Syst.* **2015**, *88*, 85–96. [[CrossRef](#)]
40. 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](#)] [[PubMed](#)]
41. Molla, M.K.I.; Al Shiam, A.; Islam, M.R.; Tanaka, T. Discriminative feature selection-based motor imagery classification using EEG signal. *IEEE Access* **2020**, *8*, 98255–98265. [[CrossRef](#)]
42. Molla, M.K.I.; Morikawa, N.; Islam, M.R.; Tanaka, T. Data-adaptive spatiotemporal ERP cleaning for single-trial BCI implementation. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2018**, *26*, 1334–1344. [[CrossRef](#)] [[PubMed](#)]

43. 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*, 1–9. [[CrossRef](#)]
44. Gu, J.; Wei, M.; Guo, Y.; Wang, H. Common spatial pattern with L21-norm. *Neural Process. Lett.* **2021**, *53*, 3619–3638. [[CrossRef](#)]
45. Al Shiam, A.; Islam, M.R.; Tanaka, T.; Molla, M.K.I. Electroencephalography based motor imagery classification using unsupervised feature selection. In Proceedings of the 2019 International Conference on Cyberworlds (CW), Kyoto, Japan, 2–4 October 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 239–246.
46. Vapnik, V. *The Nature of Statistical Learning Theory*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2013.
47. Ye, J.; Li, Q. A two-stage linear discriminant analysis via QR-decomposition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2005**, *27*, 929–941. [[PubMed](#)]
48. Haykin, S. *Neural Networks: A Comprehensive Foundation*; Prentice Hall PTR: Hoboken, NJ, USA, 1998.
49. Tam, W.K.; Tong, K.y.; Meng, F.; Gao, S. A minimal set of electrodes for motor imagery BCI to control an assistive device in chronic stroke subjects: A multi-session study. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2011**, *19*, 617–627. [[CrossRef](#)] [[PubMed](#)]
50. Sreeja, S.; Rabha, J.; Samanta, D.; Mitra, P.; Sarma, M. Classification of motor imagery based EEG signals using sparsity approach. In Proceedings of the Intelligent Human Computer Interaction: 9th International Conference, IHCI 2017, Evry, France, 11–13 December 2017; Proceedings 9; Springer International Publishing: Berlin/Heidelberg, Germany, 2017; pp. 47–59.
51. Mahamune, R.; Laskar, S.H. An automatic channel selection method based on the standard deviation of wavelet coefficients for motor imagery based brain–computer interfacing. *Int. J. Imaging Syst. Technol.* **2023**, *33*, 714–728. [[CrossRef](#)]
52. Qiu, Z.; Jin, J.; Lam, H.K.; Zhang, Y.; Wang, X.; Cichocki, A. Improved SFFS method for channel selection in motor imagery based BCI. *Neurocomputing* **2016**, *207*, 519–527. [[CrossRef](#)]
53. Radman, M.; Chaibakhsh, A.; Nariman-zadeh, N.; He, H. Generalized sequential forward selection method for channel selection in EEG signals for classification of left or right hand movement in BCI. In Proceedings of the 2019 9th International Conference on Computer and Knowledge Engineering (ICCKE), Mashhad, Iran, 24–25 October 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 137–142.
54. Tiwari, A.; Chaturvedi, A. Automatic EEG channel selection for multiclass brain-computer interface classification using multiobjective improved firefly algorithm. *Multimed. Tools Appl.* **2023**, *82*, 5405–5433. [[CrossRef](#)]

**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.