RESEARCH ARTICLE

Multi-Task EEG Signal Classification Using Correlation-Based IMF Selection and Multi-Class CSP

N. ALIZADEH, S. AFRAKHTEH, AND M. R. MOSAVI [ORCID]

Department of Electrical Engineering, Iran University of Science and Technology, Tehran 16846-13114, Iran

Corresponding author: M. R. Mosavi (m_mosavi@iust.ac.ir)

ABSTRACT In the context of motor imagery (MI)-based brain-computer interface (BCI) systems, a great amount of research has been studied for attaining higher classification performance by extracting discriminative features from MI-based electroencephalogram (EEG) signals. In this study, we propose an innovative approach for classifying multi-class MI-EEG signals, which consists of a signal processing technique based on empirical mode decomposition (EMD) and multi-class common spatial patterns (MCCSP). Specifically, after applying the EMD, we propose selecting the best intrinsic mode functions (IMF) as the substitution to the original EEG signal for the next stage of processing. The metric we used for the selection is based on the cross-correlation of each decomposed IMF with the original signal. Next, we extend the CSP algorithm to the MCCSP to be utilized as the feature extractor. We applied our technique to the BCI competition IV (2a). Results revealed that the proposed technique improved classification accuracy significantly compared to the original case when applying MCCSP directly to the original EEG channel data. Moreover, the K-nearest neighbor (KNN) achieved the highest mean classification accuracy rate of 91.28%. Our findings suggest that a promising elevated classification accuracy of 96.71% can be achieved by raising the feature dimension through MCCSP. Compared to state-of-the-art algorithms, the performance of the proposed method is highly convincing and motivating for future studies.

INDEX TERMS BCI, cross-correlation, classification, EEG, EMD, MCCSP.

I. INTRODUCTION

Brain-computer interfaces (BCIs) provide an effective link between human brain activity and the world of technology and computers, which can prepare a new non-neuromuscular output to the central nervous system (CNS) that is non-hormonal too [1]. BCI systems not only enable intelligent daily functioning for persons with disabilities, but they can also regenerate or substitute for an activity or natural output lost due to damages or diseases [2], [3], [4], [5], [6], [7].

Electroencephalography (EEG) is a non-invasive method generally used to record signals arising from the electric field produced by the brain [8]. The motor imagery (MI) task uses

The associate editor coordinating the review of this manuscript and approving it for publication was Md. Kafiul Islam [ORCID].

EEG data to simulate the movement of a body part without actually moving it [9], [10]. To successfully deploy a BCI system, MI-EEG processing consists of two key stages: feature extraction and classification. Feature extraction is carried out in order to isolate the features associated with a specific MI class and their differences from other classes. Most of the features used in BCI systems are based on spatial, temporal, or spectral analyses. Ramoser et al. developed common spatial patterns (CSP) as a fundamental method for extracting discriminative features from multi-channel EEG data during left- and right-hand MI movements [11]. The effectiveness of these patterns in categorizing various MI-EEG tasks has since led to many studies utilizing a variety of CSP versions to achieve high classification accuracy [12], [13], [14], [15], [16], [17], [18], [19]. Furthermore, significant works have been done to develop effective feature extraction using

discrete wavelet transforms (DWT) [20], wavelet packet decomposition (WPD) [21], and fast Fourier transformation (FFT) [22], [23].

The empirical mode decomposition (EMD) algorithm, as one of the data-driven processing methods of non-linear and non-stationary EEG signals, was used to decompose a signal into its intrinsic mode functions (IMF) and a residual using an iterative approach. EMD decomposition-based methods are popular because they require no predefined functions like Fourier or classical time-frequency transformations and are less susceptible to a priori assumptions on input data. Thus, researchers have done a great deal of research on feature extraction and classification of EEG signals using various EMD approaches. In this regard, generating features based on EMD and lifting wavelet transform for compression of EEG signals and classification using ANNs was presented in [24]. In [25], the classification of MI movements was performed using multivariate empirical mode decomposition (MEMD) and short-time Fourier transform (STFT). After selecting the third IMF, STFT was applied to divide signals into 8 shorter time frames. Finally, the mean accuracy rate of 86.43% and 87.14% was achieved by evaluating ANFIS and PNN classifiers [25]. The instantaneous phase-based features in [26] were obtained by applying the Hilbert transform to four types of IMF signals of the same type in each of the 22 EEG channels. Finally, a long short-term memory (LSTM) network has been operated to classify the four-class MI features and a maximum mean accuracy rate of 89.89% has been accomplished [26]. Ji et al. in [27] first decomposed EEG signal by DWT and then an appropriate sub-band signal is applied to EMD. Then, using the reconstructed signal spectral analysis, entropy values were obtained and classified by the SVM classifier. Their approach led to a classification accuracy of 85.71% on BCI-IV2a. A combination of EMD and principal component analysis (PCA) has been studied in [28]. PCA was performed for dimension reduction of feature vectors consisting of entropy values of the first IMF, and average accuracy of 78.65% was achieved by carrying out the LDA classifier. Table 1 summarizes the reviewed works in MI-EEG signal classification.

In MI-EEG signal processing, all signal components are not constructive and do not include useful information. Therefore, it's desired to find the most informative signal component for detecting different MI movements. Additionally, blindly selecting the first or third IMF, without considering all data channels is not optimal. Therefore, to tackle this issue, we propose a novel method for accurate MI-EEG classification, which performs as a new feature extraction algorithm based on EMD combined with spatial patterns. First, a stage of pre-processing of the raw EEG data is carried out to extract frequency bands carrying MI information (μ and β rhythms). The next step contains applying EMD to multi-channel EEG signals acquired from the MI movements of the right hand, left hand, both feet and tongue. Then, to enhance the performance of EMD in terms of selecting the most effective and practical IMF, an IMF selection criterion

called cross-correlation is applied to the first four IMFs. In the next stage, we use the selected IMF of each EEG channel as their alternations. Additionally, since the study's primary focus is on four-class data, we extend the CSP to the multi-class CSP (MCCSP) to extract features. With the MCCSP, not only we can develop spatial patterns for four MI classes, but also for dimension reduction. In other words, the main contributions of the presented study are as follows:

- Suggesting the new feature extraction technique exploiting the combination of the IMF selection and the MCCSP approach (i.e., IMF selection + MCCSP).
- Proposing the cross-correlation to find the most relevant IMF: As the first part of our proposal (i.e., IMF selection), we suggest employing cross-correlation to find the most informative IMF signal in terms of its similarity to the original signal while being exposed to artifact reduction by applying signal decomposition.
- Utilization of MCCSP in the feature extraction stage, which is an extension of the CSP approach: the rationale behind using this strategy is its capability to increase discriminability in the feature extraction stage by integrating the abilities of both techniques.
- Applicability of the proposed technique as the input for different classifiers such as K-nearest neighbor (KNN), MLPNN, and decision tree (DT).

The rest of the paper is categorized as follows: Section II introduces details about the dataset used in this paper and the idea of EMD, MCCSP, and our proposed method is described. The experimental results of applying the proposed method are explained in Section III. Then, discussions are presented in Section IV, and finally, the conclusions of this paper are provided in Section V.

II. MATERIALS AND METHODS
Here, we first provide a description of the framework we have proposed. The dataset used in this study and various parts of the suggested method is then thoroughly explicated.

A. PROPOSED FRAMEWORK
As previously discussed, decomposition-based approaches for processing EEG signals in various frequency sub-bands can be useful for extracting significant informatics data components. The EMD method has been widely employed for MI-EEG classification due to its ability to overcome the restrictions of linear time-frequency (T-F) methods. Therefore, in our proposed framework, the Butterworth-filtered MI-EEG signals have been decomposed into four IMF modes by applying the EMD. In the next stage, we investigate employing the IMF's content method to find a new signal that represents the original signal for the further phases of processing. Therefore, if there is a specific criterion for selecting the most significant IMF, the original signal information can only be shown with one IMF component. To select the most effective mode, we use the cross-correlation coefficient,

TABLE 1. An overview of the works proposed in the field of MI studies.

| Author (s)	Year	Technique	Classifier	Class number	Accuracy (%)
H. Ramoser et al. [11]	2000	CSP	Linear classifier	2	90.80
Cheng et al. [13]	2016	RCSPTL	LDA	2	74.68
Park et al. [14]	2019	FBCSP	SVM	2	89.13
Wang et al. [16]	2012	CSP-L1	FDA	4	92.88
Afrakhteh et al. [17]	2020	CSP	FNNPSOGSA	2	97.82
Mosavi et al. [18]	2021	CSP	ANFIS-CPSO	2	96.43
Miao et al. [19]	2021	CTFSP	SVM	3	85.00
Pattnaik et al. [20]	2016	DWT	ANN	2	80.71
Zhang et al. [21]	2016	AR model and WPD	SVM	5	96.84
K. Keerthi Krishnan et al. [23]	2021	VMD	EEGNet	4	94.41
Syed Khairul Bashar et al. [25]	2016	MEMD + STFT	ANFIS	2	86.43
Tosun et al. [26]	2021	HHT + eigenvalues	LSTM	4	89.89
Ji et al. [27]	2019	DWT and EMD combined with approximate entropy	SVM	4	85.71
Md. Toky Foysal Talukdar et al. [28]	2014	EMD + PCA	LDA	2	78.65
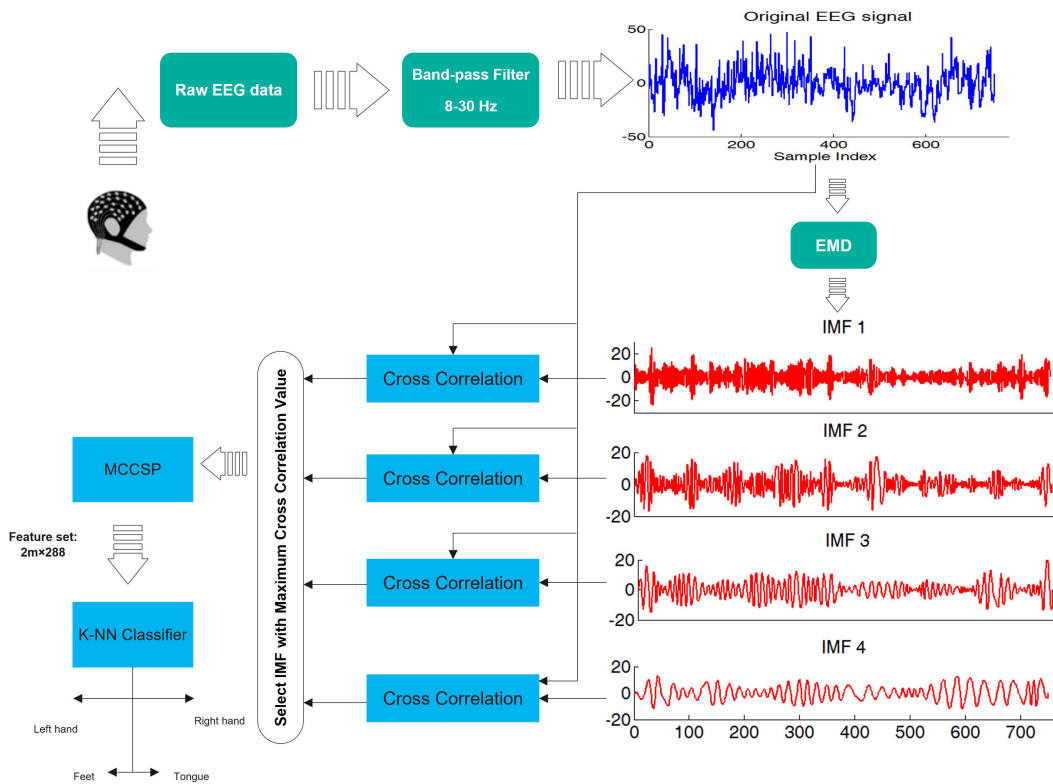
FIGURE 1. Block diagram of the proposed technique for EEG signal classification.

which can be expressed as Eq. (1):

$$c_{cor} = \frac{\sum_{l=1}^{S} \left(C_j(l) - \bar{C}_j \right) (x(l) - \bar{x})}{\sqrt{\sum_{l=1}^{S} \left(C_j(l) - \bar{C}_j \right)^2 \sum_{l=1}^{S} (x(l) - \bar{x})^2}} \quad (1)$$

where c_{cor} computes the correlation coefficient between the original EEG signal and an IMF. S denotes the sampling frequency of the signal. C_j is the $j-th$ IMF and $\bar{C}_j = \frac{1}{S}\sum_{l=1}^{S} C_j(l)$. x is the original signal and $\bar{x} = \frac{1}{S}\sum_{l=1}^{s} x(l)$.

A high correlation coefficient between the original signal and each of its IMFs, which is denoted by the cross-correlation, signifies the corresponding IMF better represents the original EEG signal and contains more useful information. This means that the most effective IMF is

selected based on the highest correlation coefficient value. The advantages of correlation-based IMF selection are as follows:

- The original signal contains noise, and by this algorithm, a kind of noise removal operation is performed.
- By selecting the most effective IMF from the four IMFs obtained, the impractical information of the signal is released, and instead of processing the four components of the main signal, in the next steps, only the meaningful part of the signal is processed, which declines the computational and processing cost. Furthermore, it allows for the reduction of the feature dimension by selecting just one IMF from a possible four IMFs.
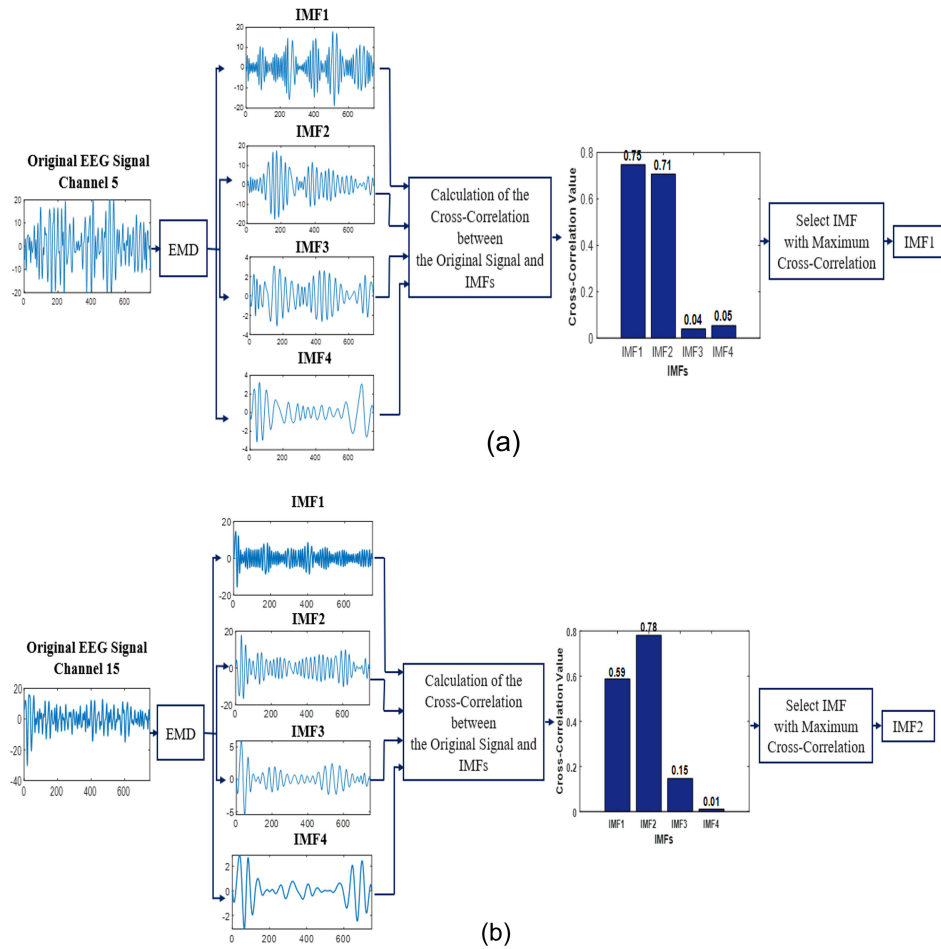
FIGURE 2. The band-pass filtered MI-EEG signals and their 4 IMFs from (a) channel 5, and (b) channel 15. The most significant IMF has been selected based on the cross-correlation coefficient between the original signal and each of its IMFs.

Therefore, considering this factor, the best IMF is selected from the four IMFs for the feature extraction process. Fig. 1 shows the structure of our proposed technique. An illustration of the filtered original signal and its four IMFs in relation to two different EEG-data channels is presented in Fig. 2. The c_{cor} coefficients of each IMFs are presented to select the most effective one for further analysis. For instance, in Fig.2(b), the EEG signal and its second selected IMF have the highest coefficient of 0.78. It should be noted that we utilized the first four IMFs in our proposed strategy because the first IMFs are the most informative component of the signal. Additionally, oscillations become insignificant as the number of IMFs is increased beyond four, which reduces the correlation between the original signal and the IMFs that correspond to them. This demonstrates that there is a relatively low likelihood of choosing IMFs with higher indexes in the proposed strategy. Furthermore, the consideration of additional IMFs presents an increment in computational complexity, as well as necessitating the computational task of computing the cross-correlation of the original EEG signal with each respective IMF.

In the next step, these selected IMFs are used as the input of MCCSP for the feature extraction. An advantage of MCCSP in addition to its high discriminability is the dimension reduction, which leads to reducing the computational complexity. The number of $2m$ feature vectors is derived through MCCSP, which can be controlled by setting the parameter m. Then, these extracted features of various MI-EEG classes are used to evaluate classification accuracies by the KNN classifier. The dataset used in this paper, as well as the details of the various processing tools of our proposed method, are discussed in the next sections.

B. EEG DATA PRE-PROCESSING

Our proposed framework has been evaluated utilizing dataset 2a from BCI Competition IV [29]. The dataset includes EEG signals from nine healthy individuals visualizing four MI movements: left hand, right hand, feet, and tongue. Each subject must spend a considerable amount of time and effort gathering sufficient numbers of labeled samples for a
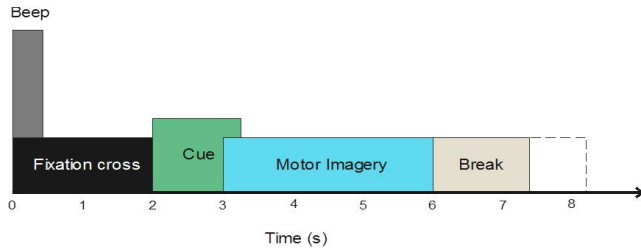
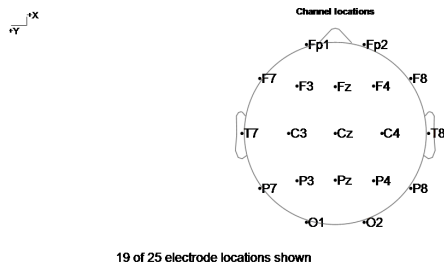**FIGURE 3.** The timing diagram of the cue-based paradigm.

FIGURE 4. Channel location of the EEG electrode of BCI competition IV dataset 2a extracted with EEG-LAB in MATLAB.

subject-specific model due to the substantial inter-session and subject variability. Several cross-subject and cross-session approaches have been considered to address this issue [30], [31]. Fig. 3 demonstrates the timing scheme for cue-based subject guidance. 22 EEG recording Ag/AgCl electrodes and 3 EOG channels are used to record data, although the EOG channels should not be used for classification. Therefore, the placement of the 22 electrodes on the scalp follows the international 10-20 system which is shown in Fig. 4 [32]. The EEG signals were sampled at 250 Hz, and ambient noise filtering was performed by the 0.5-100 Hz band-pass filter. Line noise is removed via a 50-Hz notch filter.

After data acquisition, a preprocessing stage is carried out to remove noise and artifacts from the raw EEG signals in order to reduce their impact on the efficiency of the feature extraction and classification phases. EOG and EMG are common artifacts that interfere with EEG signals owing to eye and muscle movements [33], [34]. In this paper, EEG signals were filtered using a third-order Butterworth filter with an 8-30 Hz frequency passband. By setting this frequency band, the $\mu(8-14\,Hz)$ and $\beta(14-30\,Hz)$ rhythms are placed in this frequency range, which contains MI information [35]. Furthermore, this filtering eliminates EOG and EMG noise from the EEG signal. Other representations of the EEG signals used in this paper are shown in Fig. 5 and Fig. 6. Specifically, power spectral density (PSD) of the selected frequency band (8-30Hz) has been shown in Fig. 5. On the other hand, the power of the EEG signal recorded from channel 14 is plotted in Fig. 6. All of this information is highly important for choosing the right frequency band and channel to record the relevant EEG data for the proposed BCI application. Finally, the EMD algorithm is applied to the Butterworth-filtered EEG signals; details are provided in the next section.

C. EMPIRICAL MODE DECOMPOSITION

Data analysis is an integral part of the world of research and machine learning. In 1998, Huang et al. [36] presented a new signal decomposition-based method called EMD which is a data-dependent multi-resolution method for analyzing non-linear and non-stationary signals like EEG. By decomposing a signal using the EMD process, a series of IMFs are obtained that must have the following two principal conditions: First, the number of zero-crossing points must be the same as the number of extrema points in the entire dataset or at most differ by one. Additionally, the mean value of the upper and lower envelopes constructed from the local extrema is zero.

Using EMD, the input EEG signal $x(t)$ can be decomposed into some IMFs in the following steps:

1) Selecting all the local maxima and minima.
2) After determining the extrema points, the upper and lower envelopes are obtained by connecting all the maxima and minima points separately using a cubic spline curve.
3) The average of the two upper and lower envelopes is defined as Eq. (2):

$$m_1(t) = \frac{e_{\text{lower}}(t) + e_{\text{upper}}(t)}{2} \qquad (2)$$

where $e_{\text{upper}}(t)$ and $e_{\text{lower}}(t)$ are the upper and lower envelope curves, respectively.

4) Then, the difference between the original signal $x(t)$ and $m_1(t)$ is calculated by Eq. (3):

$$x(t) - m_1(t) = h_1(t) \qquad (3)$$

where $h_1(t)$ is the first proto-IMF and ideally, satisfies the IMF definition. t denotes the time index. The above procedure is designated as the sifting process and is repeated several times until the first IMF is obtained. The next sifting process is performed again on $h_1(t)$ using Eq. (4):

$$h_1(t) - m_{11}(t) = h_{11}(t) \qquad (4)$$

where $h_1(t)$ is considered as the data. If the sifting process is repeated q times, we will have:

$$h_{1(q-1)}(t) - m_{1q}(t) = h_{1q}(t) \qquad (5)$$

where $h_{1q}(t)$ is the first IMF:

$$h_{1q} = C_1 \qquad (6)$$

where C_1 is assigned as the first IMF component which satisfies the definition of IMF. The criterion for stopping the sifting process is the standard deviation (SD) which is obtained through the results of two consecutive sifting consummations as follows:

$$SD = \sum_{t=0}^{T} \left[\frac{\left| \left(h_{1(q-1)}(t) - h_{1q}(t) \right)^2 \right|}{h_{1(q-1)}^2(t)} \right] \qquad (7)$$
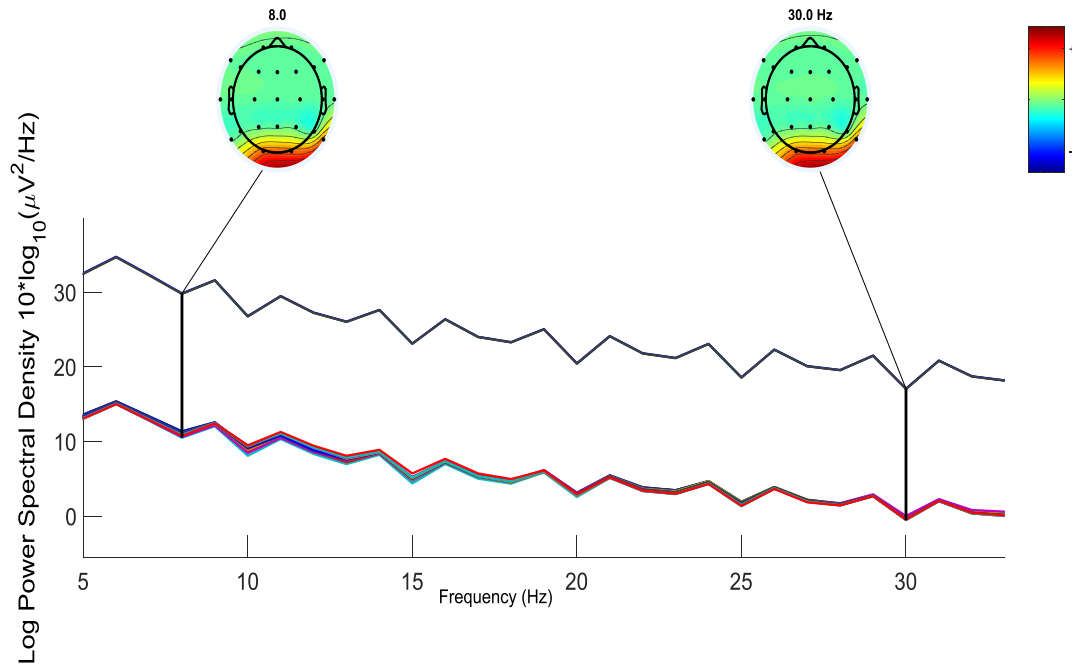
FIGURE 5. PSD of the frequency band in the range of 8Hz to 30Hz.
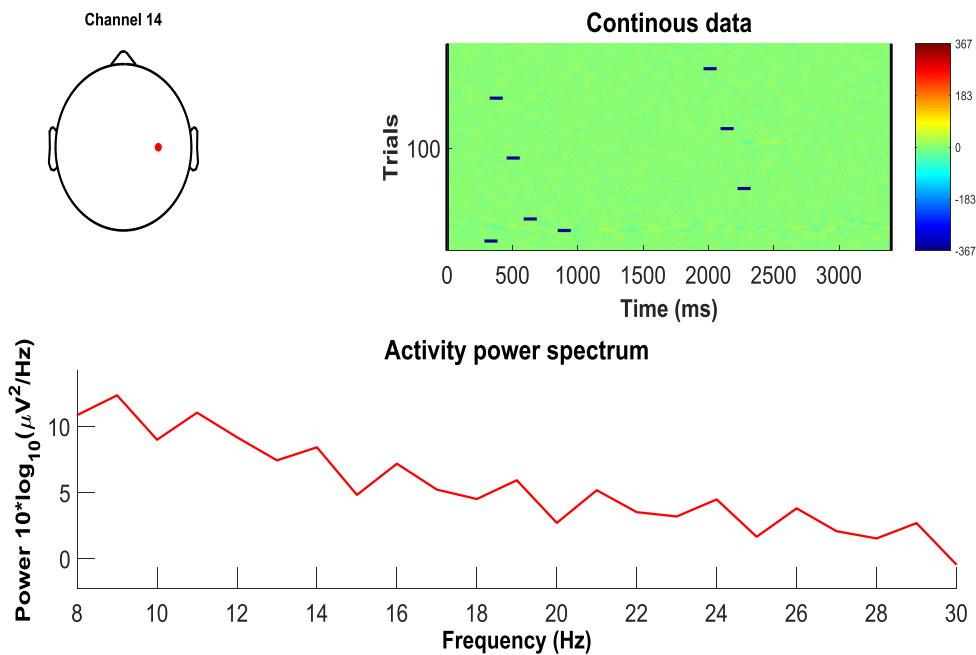
FIGURE 6. Power representation of EEG signal for channel 14 in the frequency range of 8-30Hz.

5) After obtaining the first IMF (i.e., C_1), it is separated from the rest of the data using Eq. (8):

$$r_1(t) = x(t) - C_1 \qquad (8)$$

where $x(t)$ is the original signal and C_1 is the first IMF. r_1 is the residue that still contains critical information about the original data on which the EMD decomposition method is performed again to extract the other IMFs.

6) Finally, the original signal is decomposed into a number of IMFs that can be reconstructed as Eq. (9):

$$x'(t) = \sum_{i=1}^{p} C_p + res_p \qquad (9)$$

where $x'(t)$ is the original reconstructed signal. $C_1(t), C_2(t), \ldots, C_p(t)$ are IMF components as a result

of the EMD and p Specifies the number of IMFs. res $_p$ is the last residue signal.

In the proposed framework, we perform an EMD algorithm on the bandpass-filtered EEG signals of all 22 channels and retain the first four IMFs for the next analysis. Finally, the selected IMFs are applied to the MCCSP to extract the input data of the classifier. In the next section, MCCSP as a feature extraction algorithm is explained in detail.

D. MULTI-CLASS COMMON SPATIAL PATTERNS

In the proposed study, feature vectors for the classifier are obtained from the CSP approach. The MCCSP algorithm considered in this study generalizes spatial patterns for multi-class data. CSP has been developed as one of the effective and widely used data-dependent spatial filters in EEG-BCI systems to separate the distinctive patterns of binary MI. The principal idea of this algorithm is to maximize the difference between two classes of EEG data using a projection matrix that maps the data into a low-dimensional spatial space [37].

Given that our goal is to classify the four classes of MI tasks, we first subtract raw EEG data of each class from its mean and obtain $[X_{C_i}]_{N \times T}$ for $i = 1, 2, 3, 4$ where i is the class number index. N represents the number of EEG channels and T denotes the number of time samples of each EEG channel. This process is called common average referencing and is usually done to remove noise. The covariance matrix for each of the classes is then computed as Eq. (10):

$$R_{ci} = X_{ci}{}^T X_{ci}, \quad i = 1, 2, 3, 4 \tag{10}$$

where i is the index of class number and since the problem is considered for the four classes, the covariance matrix is calculated for each of the four classes. X^T indicates the transpose of a matrix.

Then, we form the combined covariance matrix which is the sum of the covariance matrices of different classes:

$$R = \sum_{i=1}^{N_c} R_{ci} \tag{11}$$

where N_c indicates the total number of classes. Decomposition of the resulting combined covariance matrix into its eigenvalues and eigenvectors is defined as Eq. (12):

$$R = U_0 \Lambda U_0{}^T \tag{12}$$

where U_0 represents a $N \times N$ unitary matrix that contains the principal components and $[\Lambda]_{N \times N}$ is a diagonal matrix of eigenvalues. Therefore, the whitening transformation matrix is constructed as Eq. (13):

$$W = \Lambda^{-1/2} U_0^T \tag{13}$$

This operation specifies components that have non-zero eigenvalues. Then, the conversion of the covariance matrix R_{ci} to S_{ci} is done with the following mapping:

$$S_{ci} = W R_{ci} W^T \tag{14}$$

where i represents the class number. The extraction of spatial values and spatial vectors of the covariance matrix in the new space S for class i will be as follows. Then, Eigen decomposition of the covariance matrix in the new space S_{ci} for $i - th$ class can be noted as Eq. (15):

$$S_{ci} = U_i \Lambda_{ci} U_i{}^T, \quad i = 1, 2, 3, 4 \tag{15}$$

where U_i is the common principal components matrix for class i. Then, m principal components are selected from S_{ci} corresponding to the maximum eigenvalues, and m principal components having the minimum eigenvalues are selected and denoted by U_i^s. Finally, spatial filter for the $i - th$ class will be as Eq. (16):

$$SF_{ci} = (U_i^s)^T W \quad i = 1, 2, 3, 4 \tag{16}$$

where SF_{ci} is the spatial filter for class i. Then, with the obtained spatial filter, the decomposition of X_{ci} can be written as Eq. (17):

$$X_{ci} = SP_{ci} Z_{ci} \quad i = 1, 2, 3, 4 \tag{17}$$

where SP_{ci} is the pseudoinverse of SF_{ci} and can be considered as the spatial patterns matrix for i-th class. Z_{ci} is a new time series as the projection of X_{ci} into the space of CSP and can be written as follows:

$$Z_{ci} = SF_{ci} X_{ci} \tag{18}$$

where X_{ci} is EEG data of $i - th$ class and SF_{ci} is its corresponding spatial filter obtained using Eq. (16).

The process is repeated for all MI classes to obtain spatial filters for all classes. Next, by applying these spatial patterns, feature vectors are obtained for all classes as the output of MCCSP. The main idea of MCCSP for more than two conditions is that it computes spatial patterns for each class against all other conditions. Similar to CSP, MCCSP minimizes the variance of all other classes by maximizing the variance of the considered class. As a result, MCCSP can also lead to distinguishable information.

Finally, after extracting the features using MCCSP, we use the KNN approach in the classification stage. The pseudo-code in Algorithm 1 describes all stages of our suggested framework. The performance evaluation of the proposed method on the dataset introduced in section II-B is provided in the next section.

III. RESULTS

The main goal of our study is to accurately classify four MI (right hand, left hand, feet, and tongue) EEG-based mental tasks by extracting effective features. The results have been evaluated by applying the proposed framework on dataset 2 a from the BCI Competition IV. As described, two sessions of data acquisition were performed for each subject. Each session consisted of 6 runs, 72 trials for each class, and finally a total of 288 trials for each session. First, MI-EEG signals associated with various runs have been extracted in the pre-processing stage. Then, the data of four MI classes

---

Algorithm 1 Proposed Algorithm for Accurate Classification of MI-EEGs

---

Input: $[\mathbf{X_c}]_{N \times T}$, m, K, N_e

                                                             ▷ $[\mathbf{X_c}]_{N \times T}$: EEG Channel Data,
                                                          ▷ m: Parameter in MCCSP,
                                                                  ▷ K: KNN Parameter,
                                      ▷ N_e: Number of IMFs at the output of EMD

Output: *Predicted_Class*

1: **Artifact reduction of $[\mathbf{X_c}]_{N \times} \times \mathbf{T}$**   ▷ Using band-pass filter considering cut-off frequencies: 8, 30 Hz.
2: **for** $i = 1; i = N; i + +$ **do**
3:     *Original* $= \mathbf{X_c}$(i, 1 : *end*)   ▷ Extracting i-th channel's signal
4:     **IMF**$(1 : N_e)$ = EMD (*Original*, N_e)   ▷ Extracting IMFs
5:     **for** $j = 1; j = N_e; j + +$ **do**
6:        C_{cor} (i, j) $\leftarrow$ *Cross_correlation*(**IMFj**, *Original*)   ▷ Using Eq. (1)
7:     **end for**
8:     *Index*(i) = find(**C(i, j)** == Max(**C(i, j)**))   ▷ Find the IMF with maximum cross-correlation
9:     *New_signal*(i) = *IMF*(*Index*(i))   ▷ Extracting new signal for on behalf or original
10: **end for**
11: *Feature_set* = MCCSP(*New_signal*, *m*)   ▷ Feature Extraction using MCCSP
12: *Predicted_Class* = KNN(Feature_set, *K*)   ▷ Applying KNN classifier to the extracted features

---

and their labels are separated. Additionally, to retain the μ and β frequency bands, the raw EEG signals are bandpass filtered by utilizing a 3^{rt} order Butterworth filter in the frequency range of 8 Hz $-$ 30 Hz.

Each data channel has a dimension of $4 \times 750 \times 22$ after applying EMD, where 4 indicates the number of IMFs. The input of MCCSP as feature extraction approach (supervised algorithm), is a 3D dataset containing 72 trials per class. Each trial consists of data obtained from 22 channels. Each channel contains 750-time samples representing the most appropriate IMF selected during the correlation-based IMF selection process. As a result of projection in the MCCSP approach, the obtained spatial patterns are multiplied by the input data related to each class to generate the feature vectors for various MI-EEG classes. Employing MCCSP results in the extraction of $2m$ feature vectors. Therefore, considering the parameter m, the dimension of the feature data obtained is $288 \times 2m$, where 288 indicates the number of trials. Finally, KNN was used as the classifier. The main reason behind choosing KNN as classifier is its capability to adapt to the multi-class classification problem as well as its ability to control the performance by simply changing its parameter (i.e., K). The performance evaluation has been done using 5×5 fold cross-validation technique. Fig. 7 displays the graphic depiction of the train-test split for each fold.

The classification accuracy results can be affected by various parameters. By setting the parameter m in MCCSP and K in the KNN classifier, the test accuracy rates are obtained for all nine subjects in terms of m and K, which is listed in Table 2. As shown in this table, we set the parameters $m = 3, 5, 7$ and $K = 3, 4, 5$.

To provide the impact of EMD, the classification accuracies for the case where just MCCSP is employed are also presented (without EMD). Finally, we reported the mean and
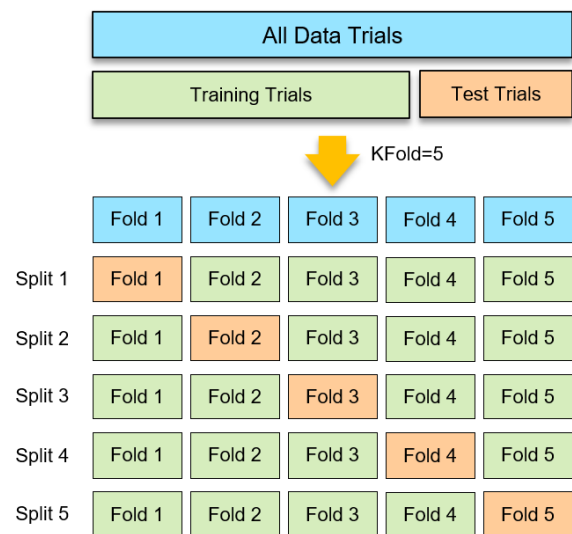
FIGURE 7. 5 × 5 fold cross-validation visualization. The number kf equals the number of folds.

standard deviation (STD) across all folds. The results demonstrate the impact of the proposed architecture in improving the classification accuracy for all subjects. It is abundantly clear that the proposed framework can offer superior classification accuracy performance.

To analyze the impact of both m and K parameters on classification performance (%), two-dimensional (2-D) heat maps are provided in Fig. 8 for different subjects. It is clear from the heatmaps that as m rises, the recognition accuracy of the subjects significantly increases. A greater classification performance has also been achieved with smaller K values in the KNN classifier.

FIGURE 8. 2-D heat maps created for the classification accuracy of several subjects in terms of *m* and *K*.

TABLE 2. Classification test accuracies (%) for the proposed method in terms of m and K. The better results have been shown in boldface.
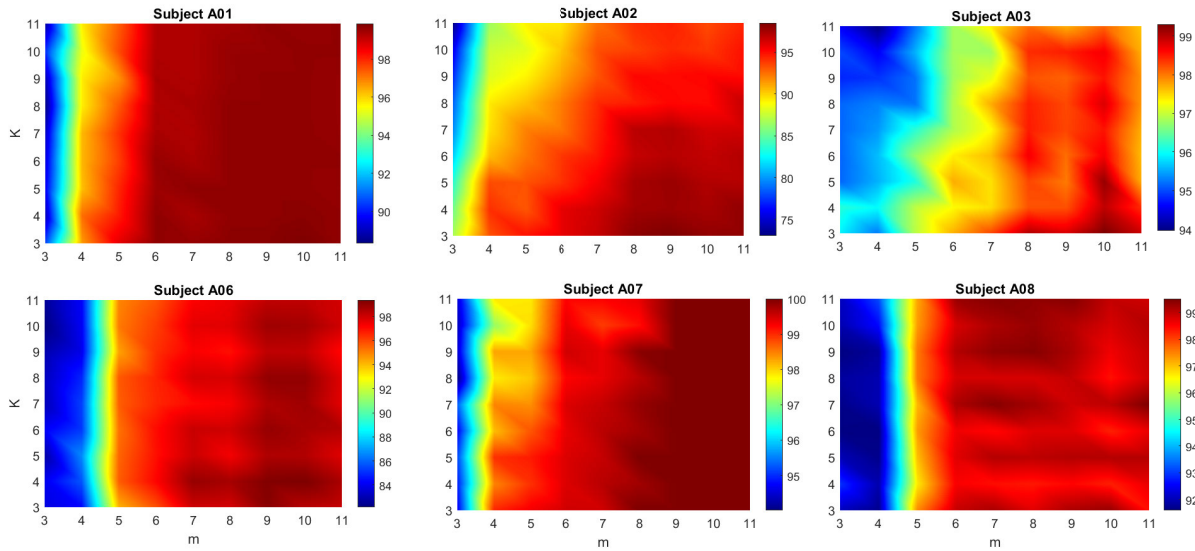
Subject	Method	K=3, m=3	K=3, m=5	K=3, m=7	K=4, m=3	K=4, m=5	K=4, m=7	K=5, m=3	K=5, m=5	K=5, m=7
A01	MCCSP	87.99±4.24	94.30±2.99	94.93±3.41	87.92±3.12	93.95±2.97	93.68±3.12	88.76±3.29	92.64±3.08	93.88±2.31
	Proposed	**90.05±4.39**	**98.33±2.22**	**99.65±0.71**	**88.48±3.39**	**98.11±1.32**	**99.58±0.76**	**89.87±3.87**	**97.71±1.78**	**99.51±0.79**
A02	MCCSP	72.62±5.40	77.36±3.90	79.99±5.04	71.18±5.67	74.38±4.85	77.07±3.10	71.14±6.25	75.42±4.87	76.87±5.60
	Proposed	**87.47±3.93**	**94.31±2.38**	**96.73±2.29**	**84.82±4.63**	**94.51±2.68**	**96.38±2.38**	**85.42±4.28**	**93.68±3.61**	**95.69±3.56**
A03	MCCSP	91.87±2.87	94.17±2.91	95.28±2.28	91.38±3.5	93.66±1.89	93.67±3.31	92.22±3.89	94.39±3.01	95.08±2.41
	Proposed	**96.04±2.21**	**97.22±1.80**	**98.11±1.75**	**95.83±3.09**	**97.01±1.96**	**97.70±1.74**	**95.35±2.28**	**96.86±1.89**	**97.77±1.54**
A04	MCCSP	67.24±6.06	78.97±4.82	84.93±5.9	64.95±5.82	77.80±6.91	85.69±4.22	65.68±6.47	75.89±5.44	83.26±5.06
	Proposed	**87.30±4.22**	**92.23±2.91**	**91.67±2.15**	**85.21±4.30**	**91.33±3.86**	**92.00±2.94**	**85.20±3.89**	**91.46±3.09**	**91.86±3.02**
A05	MCCSP	74.50±4.30	78.62±3.96	72.84±4.80	73.20±4.40	76.31±4.01	70.77±3.82	72.41±4.61	74.79±4.02	67.72±6.21
	Proposed	**94.21±2.26**	**95.55±3.02**	**93.19±2.47**	**93.38±3.77**	**96.10±2.46**	**92.76±3.85**	**93.18±3.04**	**93.54±2.61**	**93.44±2.27**
A06	MCCSP	64.74±4.76	63.40±5.22	62.59±4.61	63.32±6.17	63.53±4.8	60.08±5.77	62.25±6.7	58.91±6.67	59.84±5.7
	Proposed	**85.14±2.95**	**93.61±2.58**	**98.12±1.39**	**84.16±3.96**	**95.27±2.92**	**99.17±1.12**	**83.87±3.7**	**94.56±2.67**	**97.85±2.12**
A07	MCCSP	94.37±2.69	93.70±4.01	94.78±3.56	93.46±2.64	94.31±3.58	93.77±2.81	93.12±3.30	92.50±4.01	93.82±2.85
	Proposed	**95.21±2.78**	**99.03±0.89**	**99.58±0.76**	**94.64±2.83**	**98.89±1.21**	**99.37±0.70**	**95.13±2.62**	**98.82±0.99**	**99.64±0.69**
A08	MCCSP	82.99±3.75	98.75±1.76	98.33±1.83	81.12±4.29	98.74±1.29	98.33±1.53	81.94±3.47	98.54±1.46	98.13±2.04
	Proposed	**92.02±3.23**	**98.89±1.08**	**99.16±0.65**	**92.51±2.73**	**96.88±2.15**	**98.95±1.12**	**92.09±3.40**	**96.94±2.24**	**99.51±0.49**
A09	MCCSP	84.09±5.37	91.60±3.91	92.31±3.80	83.52±3.75	91.79±3.01	91.81±3.43	84.87±3.10	91.12±3.38	92.18±3.33
	Proposed	**94.11±3.30**	**94.37±2.81**	**94.17±3.12**	**92.20±3.99**	**93.39±3.03**	**94.26±2.54**	**92.48±2.89**	**93.54±3.21**	**94.16±2.28**
Mean	MCCSP	80.05±10.71	85.65±11.6	86.22±12.21	78.89±11.21	84.94±12.13	84.99±12.93	79.15±11.59	83.80±13.09	84.53±13.61
	Proposed	**91.28±3.95**	**95.95±2.50**	**96.71±2.98**	**90.14±4.53**	**95.72±2.38**	**96.69±2.98**	**90.28±4.41**	**95.23±2.43**	**96.60±2.92**

In addition to accuracy, we provided sensitivity and specificity rates to evaluate the classifier performance. Sensitivity and specificity results of both proposed and MCCSP methods for all 9 subjects, assuming $m = 3$ and $K = 3$ are given in Table 3. Cohen's kappa values of our proposed framework are also reported in this table. Cohen's kappa is a measure used to assess the performance of a multi-class classification problem that describes how much better the classifier performs over the performance of a classifier that simply predicts at random according to each class [38]. Cohen's kappa is computed with Eq. (19):

$$kappa = \frac{acc - r_e}{1 - r_e} \qquad (19)$$

where acc indicates the overall accuracy of the classification and r_e denotes the random classification result.

Table 3 indicates that subject A03 has achieved the highest kappa rate of 0.909 when using the proposed framework

and subject A06 had the lowest kappa value of 0.593. Additionally, by utilizing the only MCCSP structure, the highest kappa value can be reported for subject A07 (0.859), while the lowest kappa value can be noted for subject A04 (0.111). In comparison to the MCCSP method, our suggested approach generally raises kappa rates for all 9 subjects. Also, we provided the confusion matrix for different parameter settings in Fig. 9.

Table 4 expresses the performances of popular classifiers consisting of DT, MLPNN, and KNN for both proposed and MCCSP methods by assuming values of $m = 5$ and $K = 3$. Thus, the feature set obtained from these approaches is applied to the input of DT [39] and MLPNN. The mean accuracy across all subjects obtained by the DT is equal to 85.37% for our proposed method, while it is 69.39% for the MCCSP method. An MLPNN consisting of three layers including an input layer, a hidden layer, and an output layer has been used to evaluate the classification performance of
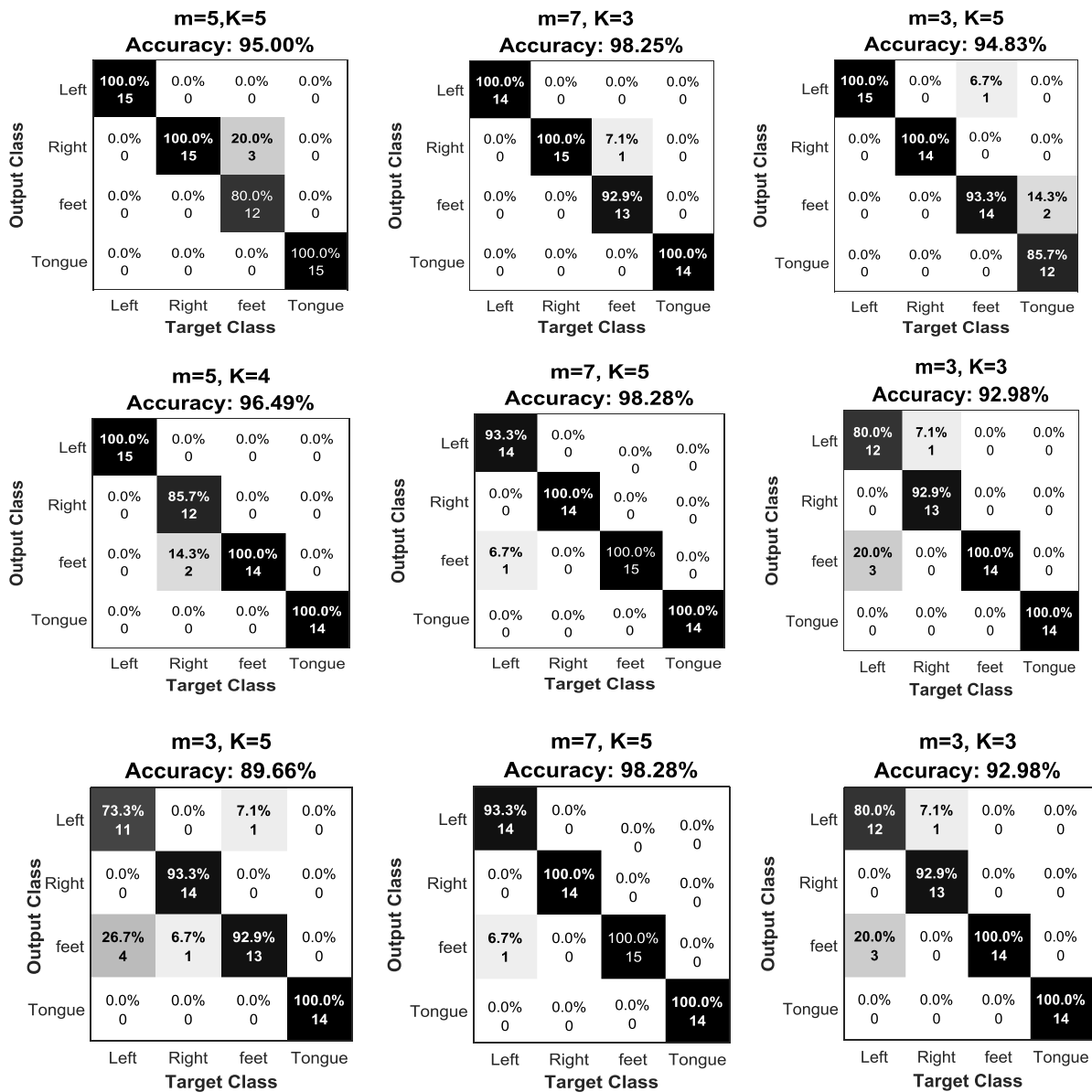
FIGURE 9. Confusion matrix of applying proposed method in terms of *m* and *K* parameters for 9 different subjects.

the proposed study. The number of hidden layer neurons is 15 and the output of each neuron is obtained using ' tanh ' as activation function [40]. The average accuracy of 81.09% has been obtained utilizing the proposed framework and described MLPNN classifier.

IV. DISCUSSION

We applied our proposed technique to the 4-class EEG data of BCI IV-2a. The classification performance of the proposed architecture is affected by the *m* parameter of MCCSP that modifies the dimension of the feature vectors. As depicted in Table 2 based classification of the features derived by the proposed framework has achieved the highest mean accuracy rate of 96.71% with an STD of ±2.98 for the 9 subjects (with m = 7). While *m* = 3 and *m* = 5 have yielded

the mean accuracy and STD values of 91.28 ± 3.95% and 95.95 ± 2.50%, respectively. It can be concluded from both Table 2 and generated 2-D heat maps in Fig. 8 that classification accuracy rates have improved for all subjects as the parameter *m* increases. Of course, by selecting smaller *m*, the performances are still high, which shows the strength of the MCCSP-based in lower dimensions. In other words, selecting a smaller m, results in a reduction in the dimension of the MCCSP-based features, which in turn reduces the computational complexity. It is, however, necessary to strike a compromise between complexity and classification accuracy. So that in order to get high accuracy rates, we must raise *m*, which results in more complexity, and in order to achieve reduced complexity, we may decrease *m*. It is therefore preferable to select appropriate *m* values that offer

TABLE 3. Sensitivity, specificity, and kappa values of proposed and MCCSP methods for BCI competition IV dataset 2a. The better results have been shown in boldface.

	Method	A01	A02	A03	A04	A05	A06	A07	A08	A09
Sensitivity	MCCSP	0.860	0.759	0.915	0.663	0.769	0.690	**0.950**	0.850	0.839
	Proposed	**0.913**	**0.875**	**0.965**	**0.882**	**0.948**	**0.850**	0.947	**0.916**	**0.932**
Specificity	MCCSP	0.954	0.919	0.971	0.888	0.922	0.896	0.982	0.948	0.946
	Proposed	**0.971**	**0.959**	**0.988**	**0.960**	**0.982**	**0.949**	**0.987**	**0.971**	**0.977**
Kappa	MCCSP	0.632	0.350	0.774	0.111	0.381	0.172	0.859	0.593	0.571
	Proposed	**0.774**	**0.672**	**0.909**	**0.683**	**0.859**	**0.593**	**0.862**	**0.774**	**0.816**

us both manageable complexity and higher accuracy results. Therefore, the major benefit of MCCSP is that a high rate of accuracy can be achieved with a smaller feature dimension.

KNN classifier accuracy can be raised or decreased by adjusting K parameter. The mean accuracy rates in Table 2 show that considering the constant value of m equal to 3, with increasing the value of K, the classification accuracy decreases. Because the resulting average accuracies for $K = 3$, $K = 4$ and $K = 5$ are achieved $91.28 \pm 3.95\%$, $90.14 \pm 4.53\%$ and $90.28 \pm 4.41\%$, respectively. Additionally, 2-D heat maps prove more sensibly that for a constant m, the classification accuracy has been increased by selecting a smaller K parameter. Consequently, although expanding the number of K results in a reduction in classification test accuracies, more neighbors vote to assign the test sample to one of four MI classes, which raises the confidence in the classification and makes the decision more robust.

Differentially, we presented the accuracies in two categories so that we were able to investigate the impact of EMD-based signal decomposition on classification performance (See Table 2). The highest average accuracy rate of 86.22% with STD of ±12.21 has been obtained for the MCCSP method by setting the parameters $m = 7$ and $K = 3$. On the other hand, the lowest accuracy rate for the MCCSP method has been achieved as 78.89% with STD of ±11.21. However, the lowest mean accuracy rate obtained from the proposed method is $90.14 \pm 4.53\%$. The results of the proposed and MCCSP method in Table 2 show that for all m and K values, our proposed approach is much more accurate than the MCCSP method. According to results in Table 2, for $m = 7$ and $K = 5$, average accuracy values and STDs are $96.60 \pm 2.92\%$ and $84.53 \pm 13.61\%$ for the proposed and MCCSP methods, respectively. Thus, our proposed approach generates 14.2% improvement over the MCCSP method. Compared to the MCCSP method, our suggested approach has the lowest inter-subject STD, which offers a more consistent classification with higher accuracy. The reason for the better classification performance in terms of test accuracy results is that the EMD algorithm can be used for noise filtering and redundant information removal in EEG signals. According to the results, we see that by extracting MCCSP-based MI-EEG features from the only cross-correlation-based selected IMF components, high classification rates can be achieved. According to the kappa values presented in Table 3, the difference in kappa value between the best subject (A02) and worst subject (A03) is 0.12 for

TABLE 4. BCI competition IV dataset 2a classification accuracy results (%) of applying the proposed method for KNN, DT, and MLPNN classifiers. The better results have been shown in boldface.

Sub	Method	KNN	DT [39]	MLPNN [40]
A01	MCCSP	94.30	71.96	75.00
	Proposed	**98.33**	**90.29**	**91.37**
A02	MCCSP	77.36	78.94	75.00
	Proposed	**94.31**	**92.98**	**88.82**
A03	MCCSP	94.17	74.67	61.83
	Proposed	**97.22**	**84.41**	**70.15**
A04	MCCSP	78.97	62.50	46.94
	Proposed	**92.23**	**83.69**	**75.86**
A05	MCCSP	78.62	54.48	72.41
	Proposed	**95.55**	**89.23**	**87.93**
A06	MCCSP	63.40	45.81	31.57
	Proposed	**93.61**	**80.91**	**61.00**
A07	MCCSP	93.70	74.73	55.83
	Proposed	**99.03**	**80.18**	**80.55**
A08	MCCSP	98.75	87.54	57.02
	Proposed	**98.89**	**84.01**	**84.48**
A09	MCCSP	91.60	73.92	77.58
	Proposed	**94.37**	**82.62**	**89.65**
Average	MCCSP	85.65	69.39	61.46
	Proposed	**95.95**	**85.37**	**81.09**

the proposed framework, whereas this value is reported as 0.73 for the MCCSP method.

A comparison of the classification performance of the suggested method using other popular classifiers is given in Table 4. Among the three classifiers, the KNN classifier achieves the highest classification accuracy of 95.95% for $m = 5$, considering the 4 nearest neighbors to participate in the unknown sample label decision. Thus, the KNN classifier improves the classification performance by 12.3% compared to the DT classifier. Moreover, it should be noted that MLPNN requires many parameter adjustments and KNN is a simpler classifier. However, in general, these results confirm the power of the proposed feature extraction algorithm in extracting the optimal features of MI-based EEG signals. In addition, the performance of the proposed algorithm is compared with the MCCSP method when using different classifiers. The results also show that our proposed approach's improvement over the MCCSP case is not significantly affected by the type of classifier. As a result, it demonstrates the feature set's strength.

Table 5 and Table 6 are presented so that we were able to compare our proposed study with other works in this area. All the works introduced to use the same BCI competition dataset IV-2a. Results revealed that the proposed technique

TABLE 5. Classification accuracy of the proposed method compared to other works in this area.

Feature extraction	Classifier	Acc (%)									
		A01	A02	A03	A04	A05	A06	A07	A08	A09	AVG
HHT [26]	LSTM	90.55	90.50	90.24	90.85	88.41	87.5	90.80	91.46	88.72	89.89
WPT [41]	MLPNN	80.00	82.00	84.00	74.00	78.00	85.00	85.00	79.00	74.00	80.11
PCA [42]	SVM	71.43	67.50	64.29	57.50	87.86	58.93	85.71	79.29	73.93	71.83
TSSM [43]	SVM	80.0	58.70	86.30	68.20	60.30	59.20	84.40	84.00	89.60	74.50
WT [44]	1D-CNN	82.14	75.05	90.10	94.61	92.30	86.75	76.60	80.50	79.55	84.18
CWT [45]	LAFFN	93.25	88.43	92.48	94.40	90.55	92.09	92.48	89.01	91.51	91.58
FAWT [46]	RF	90.32	58.06	82.25	93.5	74.19	91.93	83.87	96.77	95.16	85.11
Time-frequency [47]	CNN-GRU	99.21	99.37	99.39	99.68	99.84	98.66	99.37	99.59	99.49	99.40
(OvO)-divCSP [48]	(OvR)-SVM	87.29	63.82	88.40	74.24	61.60	60.76	92.64	83.13	79.51	76.82
Proposed method (m=5, K=5)	KNN	**97.71**	**93.68**	**96.86**	**91.46**	**93.54**	**94.56**	**98.82**	**96.94**	**93.54**	**95.23**
Proposed method (m=7, K=5)	KNN	**99.51**	**95.69**	**97.77**	**91.86**	**93.44**	**97.85**	**99.64**	**99.51**	**94.16**	**96.60**
Proposed method (m=9, K=5)	KNN	**100**	**98.31**	**98.28**	**94.83**	**96.61**	**100**	**100**	**100**	**96.55**	**98.29**

TABLE 6. Kappa comparison of the proposed method and the state-of-art approaches.

Method	Best Kappa values									
	A01	A02	A03	A04	A05	A06	A07	A08	A09	Mean
TSSM + SVM [43]	0.700	0.320	0.750	0.540	0.320	0.340	0.70	0.690	0.770	0.571
WT + 1D-CNN [44]	0.696	0.587	1.000	1.000	0.830	0.800	0.716	0.626	0.640	0.766
LAFFN [45]	0.907	0.837	0.894	0.921	0.866	0.889	0.894	0.845	0.881	0.881
CNN-GRU [47]	0.985	0.996	0.992	0.993	0.992	0.986	0.988	0.992	0.993	0.991
(OvO)-divCSP + (OvR)-SVM [48]	0.830	0.520	0.850	0.660	0.490	0.480	0.900	0.780	0.730	0.690
Ego-CNNs [49]	0.970	0.890	0.770	0.990	0.860	0.650	0.880	0.940	0.940	0.880
Proposed method (m=5, K=5)	**0.953**	**0.859**	**0.909**	**0.770**	**0.812**	**0.859**	**0.954**	**0.911**	**0.906**	**0.881**
Proposed method (m=7, K=5)	**0.954**	**0.906**	**0.952**	**0.819**	**0.862**	**0.953**	**1.00**	**1.00**	**0.908**	**0.928**
Proposed method (m=9, K=5)	**1.00**	**0.954**	**0.954**	**0.862**	**0.909**	**1.00**	**1.00**	**1.00**	**0.908**	**0.954**

is superior to the others in terms of classification accuracy and Kappa values. Furthermore, results indicate that higher accuracy and kappa rates can be achieved by setting different m values. Besides, although deep networks in [47] can achieve better results in this field, they are more complex, requiring various parameters to be adjusted. However, our proposed method can perform high-accuracy classification performance using only a simple KNN classifier. This is possible if effective and discriminative features are acquired from multi-class EEG signals.

As the main findings, the feature extraction method presented in this work utilized two powerful real-world signal processing techniques, based on EMD and CSP. By selecting an effective IMF based on the cross-correlation coefficient between the original band-pass filtered EEG signal and each of its corresponding IMFs, the traditional EMD algorithm has been enhanced. As a result, non-linear and non-stationary EEG signals have less noise and redundant information. In Table 2, the impact of the EMD-based methodology on the performance assessment of the proposed framework has been examined. Our findings demonstrate that the features derived from the proposed structure are capable of making higher discriminability among various MI-EEG tasks compared to using the original signal. On the other hand, CSP is a reliable technique for maximizing the discriminability between the variance characteristics from two classes of MI tasks. As a result, the MCCSP technique has been modified for the multi-class data mode, and the test classification accuracy rates demonstrate this method's high capacity to

extract useful features from various MI patterns. The suggested MCCSP method is also able to achieve a promising classification accuracy for the testing set, even with a small number of feature vectors in the training phase, which is measured as 91.28%, 95.95%, and 96.71% for $m = 3, 5, 7$.

All the analyses were performed in MATLAB environment on an Intel 3.5GHz core i7 PC with 12 GB of RAM.

V. CONCLUSION

This paper has presented a new feature extraction approach for multiclass MI-EEG classification utilizing EMD and MCCSP. An intelligent selection of the most effective IMF based on the correlation coefficient has been completed. Then, for the feature extraction of the selected IMFs, the MCCSP algorithm, which is based on spatial patterns, was applied. In multiclass cases, the considered MCCSP technique computes spatial patterns of each MI class against all other classes. In the next stage, multiple classifiers have been operated to estimate the quality of the proposed method in deriving discriminative information between MI-related tasks of BCI competition IV. The results of the classification performance using the KNN classifier proves the high power of the obtained feature set in the EEG-based BCIs. Selecting only one IMF from the resulting set of IMFs where the original signal information is stored not only reduces the computational load and dimensions of the processed vectors, but also removes the redundant information of EEG signals, and higher quality features are extracted in the MCCSP. The proposed method significantly improved the classification

accuracy of multi-class MI signals by extracting high-quality feature vectors that are obtained by applying IMF selection and the MCCSP. On the other hand, further control on feature dimension reduction is supplied by adjusting the parameter m in MCCSP. As the classification performance evaluated using various classifiers, like KNN and DT, is too persuasive, this high-quality feature extraction algorithm will confidently provide a new pathway for forthcoming works in the applications of MI-based BCIs. MCCSP can be adjusted in the frequency domain to obtain spatial patterns particular to distinct frequency bands, or a number of optimization algorithms can be employed to improve MLPNN classifier performance, etc.

REFERENCES

[1] M.-J. Schneider, J. J. Fins, and J. R. Wolpaw, "Scientific and engineering advances often bring with them," in *Brain–Computer Interfaces: Principles and Practice*. Oxford Univ. Press, 2012, p. 373.

[2] Y. Yu, Z. Zhou, E. Yin, J. Jiang, J. Tang, Y. Liu, and D. Hu, "Toward brain-actuated car applications: Self-paced control with a motor imagery-based brain-computer interface," *Comput. Biol. Med.*, vol. 77, pp. 148–155, Oct. 2016.

[3] R. Zhang, Y. Li, Y. Yan, H. Zhang, S. Wu, T. Yu, and Z. Gu, "Control of a wheelchair in an indoor environment based on a brain–computer interface and automated navigation," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 24, no. 1, pp. 128–139, Jan. 2016.

[4] R. Chai, G. R. Naik, T. N. Nguyen, S. H. Ling, Y. Tran, A. Craig, and H. T. Nguyen, "Driver fatigue classification with independent component by entropy rate bound minimization analysis in an EEG-based system," *IEEE J. Biomed. Health Informat.*, vol. 21, no. 3, pp. 715–724, May 2017.

[5] L.-D. Liao, C.-Y. Chen, I.-J. Wang, S.-F. Chen, S.-Y. Li, B.-W. Chen, J.-Y. Chang, and C.-T. Lin, "Gaming control using a wearable and wireless EEG-based brain–computer interface device with novel dry foam-based sensors," *J. NeuroEngineering Rehabil.*, vol. 9, no. 1, pp. 1–12, Dec. 2012.

[6] R. Bousseta, I. El Ouakouak, M. Gharbi, and F. Regragui, "EEG based brain computer interface for controlling a robot arm movement through thought," *IRBM*, vol. 39, no. 2, pp. 129–135, Apr. 2018.

[7] N. Kosmyna, F. Tarpin-Bernard, N. Bonnefond, and B. Rivet, "Feasibility of BCI control in a realistic smart home environment," *Frontiers Human Neurosci.*, vol. 10, p. 416, Aug. 2016.

[8] N. Jamil, A. N. Belkacem, S. Ouhbi, and A. Lakas, "Noninvasive electroencephalography equipment for assistive, adaptive, and rehabilitative brain–computer interfaces: A systematic literature review," *Sensors*, vol. 21, no. 14, p. 4754, Jul. 2021.

[9] J. Luo, X. Gao, X. Zhu, B. Wang, N. Lu, and J. Wang, "Motor imagery EEG classification based on ensemble support vector learning," *Comput. Methods Programs Biomed.*, vol. 193, Sep. 2020, Art. no. 105464.

[10] A. Vučković and F. Sepulveda, "A two-stage four-class BCI based on imaginary movements of the left and the right wrist," *Med. Eng. Phys.*, vol. 34, no. 7, pp. 964–971, Sep. 2012.

[11] H. Ramoser, J. Müller-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.

[12] M. Yan, Z. Lv, W. Sun, and N. Bi, "An improved common spatial pattern combined with channel-selection strategy for electroencephalography-based emotion recognition," *Med. Eng. Phys.*, vol. 83, pp. 130–141, Sep. 2020.

[13] M. Cheng, Z. Lu, and H. Wang, "Regularized common spatial patterns with subject-to-subject transfer of EEG signals," *Cognit. Neurodynamics*, vol. 11, no. 2, pp. 173–181, Apr. 2017.

[14] Y. Park and W. Chung, "Selective feature generation method based on time domain parameters and correlation coefficients for filter-bank-CSP BCI systems," *Sensors*, vol. 19, no. 17, p. 3769, Aug. 2019.

[15] K. Keng Ang, Z. Yang Chin, H. Zhang, and C. Guan, "Filter bank common spatial pattern (FBCSP) in brain–computer interface," in *Proc. IEEE Int. Joint Conf. Neural Netw. (IEEE World Congr. Comput. Intelligence)*, Jun. 2008, pp. 2390–2397.

[16] H. Wang, Q. Tang, and W. Zheng, "L1-norm-based common spatial patterns," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 3, pp. 653–662, Mar. 2012.

[17] S. Afrakhteh, M.-R. Mosavi, M. Khishe, and A. Ayatollahi, "Accurate classification of EEG signals using neural networks trained by hybrid population-physic-based algorithm," *Int. J. Autom. Comput.*, vol. 17, no. 1, pp. 108–122, Feb. 2020.

[18] M. R. Mosavi, A. Ayatollahi, and S. Afrakhteh, "An efficient method for classifying motor imagery using CPSO-trained ANFIS prediction," *Evolving Syst.*, vol. 12, no. 2, pp. 319–336, Jun. 2021.

[19] Y. Miao, J. Jin, I. Daly, C. Zuo, X. Wang, A. Cichocki, and T. Jung, "Learning common time-frequency-spatial patterns for motor imagery classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 699–707, 2021.

[20] S. Pattnaik, M. Dash, and S. K. Sabut, "DWT-based feature extraction and classification for motor imaginary EEG signals," in *Proc. Int. Conf. Syst. Med. Biol. (ICSMB)*, Jan. 2016, pp. 186–201.

[21] Y. Zhang, B. Liu, X. Ji, and D. Huang, "Classification of EEG signals based on autoregressive model and wavelet packet decomposition," *Neural Process. Lett.*, vol. 45, no. 2, pp. 365–378, Apr. 2017.

[22] H. Hindarto and S. Sumarno, "Feature extraction of electroencephalography signals using fast Fourier transform," *CommIT Communication Inf. Technology J.*, vol. 10, no. 2, pp. 49–52, 2016.

[23] K. Keerthi Krishnan and K. P. Soman, "CNN based classification of motor imaginary using variational mode decomposed EEG-spectrum image," *Biomed. Eng. Lett.*, vol. 11, no. 3, pp. 235–247, Aug. 2021.

[24] J. Sokhal, B. Garg, S. Aggarwal, and R. Jain, "Classification of EEG signals using empirical mode decomposition and lifting wavelet transforms," in *Proc. Int. Conf. Comput., Commun. Autom. (ICCCA)*, May 2017, pp. 1197–1202.

[25] S. K. Bashar and M. I. H. Bhuiyan, "Classification of motor imagery movements using multivariate empirical mode decomposition and short time Fourier transform based hybrid method," *Eng. Sci. Technol., Int. J.*, vol. 19, no. 3, pp. 1457–1464, Sep. 2016.

[26] M. Tosun and O. Çetin, "A new phase-based feature extraction method for four-class motor imagery classification," *Signal, Image Video Process.*, vol. 16, pp. 283–290, Oct. 2022.

[27] N. Ji, L. Ma, H. Dong, and X. Zhang, "EEG signals feature extraction based on DWT and EMD combined with approximate entropy," *Brain Sci.*, vol. 9, no. 8, p. 201, Aug. 2019.

[28] Md. T. F. Talukdar, S. K. Sakib, C. Shahnaz, and S. A. Fattah, "Motor imagery EEG signal classification scheme based on entropy of intrinsic mode function," in *Proc. 8th Int. Conf. Electr. Comput. Eng.*, Dec. 2014, pp. 703–706.

[29] C. Brunner, R. Leeb, G. Müller-Putz, A. Schlögl, and G. Pfurtscheller, "BCI competition 2008–graz data set a," Inst. Knowl. Discovery, Lab. Brain-Comput. Interfaces, Graz Univ. Technol., Vienna, Austria, vol. 16, Tech. Rep., 2008, pp. 1–6.

[30] T. Lotey, P. Keserwani, G. Wasnik, and P. P. Roy, "Cross-session motor imagery EEG classification using self-supervised contrastive learning," in *Proc. 26th Int. Conf. Pattern Recognit. (ICPR)*, Aug. 2022, pp. 975–981.

[31] Y. Xu, X. Huang, and Q. Lan, "Selective cross-subject transfer learning based on Riemannian tangent space for motor imagery brain–computer interface," *Frontiers Neurosci.*, vol. 15, Nov. 2021, Art. no. 779231.

[32] M. Khazi, A. Kumar, and M. Vidya, "Analysis of EEG using 10:20 electrode system," *Int. J. Innov. Res. Sci., Eng. Technol.*, vol. 1, no. 2, pp. 185–191, 2012.

[33] J. Minguillon, M. A. Lopez-Gordo, and F. Pelayo, "Trends in EEG-BCI for daily-life: Requirements for artifact removal," *Biomed. Signal Process. Control*, vol. 31, pp. 407–418, Jan. 2017.

[34] N. Elsayed, Z. Saad, and M. Bayoumi, "Brain computer interface: EEG signal preprocessing issues and solutions," *Int. J. Comput. Appl.*, vol. 169, no. 3, pp. 12–16, Jul. 2017.

[35] D. J. McFarland, L. A. Miner, T. M. Vaughan, and J. R. Wolpaw, "Mu and beta rhythm topographies during motor imagery and actual movements," *Brain Topogr.*, vol. 12, no. 3, pp. 177–186, Feb. 2000.

[36] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. Roy. Soc. London A, Math., Phys. Eng. Sci.*, vol. 454, no. 1971, pp. 903–995, Mar. 1998.

[37] K. P. Thomas, C. Guan, C. T. Lau, A. P. Vinod, and K. K. Ang, "A new discriminative common spatial pattern method for motor imagery brain–computer interfaces," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 11, pp. 2730–2733, Nov. 2009.

[38] S. M. Vieira, U. Kaymak, and J. M. C. Sousa, "Cohen's Kappa coefficient as a performance measure for feature selection," in *Proc. Int. Conf. Fuzzy Syst.*, Jul. 2010, pp. 1–8.

[39] Priyanka and D. Kumar, "Decision tree classifier: A detailed survey," *Int. J. Inf. Decis. Sci.*, vol. 12, pp. 246–269, 2020.

[40] A. Subasi and E. Erçelebi, "Classification of EEG signals using neural network and logistic regression," *Comput. Methods Programs Biomed.*, vol. 78, no. 2, pp. 87–99, May 2005.

[41] M. A. Rahman, F. Khanam, M. Ahmad, and M. S. Uddin, "Multiclass EEG signal classification utilizing Rényi min-entropy-based feature selection from wavelet packet transformation," *Brain Informat.*, vol. 7, no. 1, pp. 1–11, Dec. 2020.

[42] D. Wang, D. Miao, and G. Blohm, "Multi-class motor imagery EEG decoding for brain–computer interfaces," *Frontiers Neurosci.*, vol. 6, p. 151, 2012.

[43] X. Xie, Z. L. Yu, H. Lu, Z. Gu, and Y. Li, "Motor imagery classification based on bilinear sub-manifold learning of symmetric positive-definite matrices," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 6, pp. 504–516, Jun. 2017.

[44] B. Xu, L. Zhang, A. Song, C. Wu, W. Li, D. Zhang, G. Xu, H. Li, and H. Zeng, "Wavelet transform time-frequency image and convolutional network-based motor imagery EEG classification," *IEEE Access*, vol. 7, pp. 6084–6093, 2019.

[45] Z. Yu, W. Chen, and T. Zhang, "Motor imagery EEG classification algorithm based on improved lightweight feature fusion network," *Biomed. Signal Process. Control*, vol. 75, May 2022, Art. no. 103618.

[46] K. Jindal, R. Upadhyay, and H. S. Singh, "A novel EEG channel selection and classification methodology for multi-class motor imagery-based BCI system design," *Int. J. Imag. Syst. Technol.*, vol. 32, no. 4, pp. 1318–1337, Jul. 2022.

[47] J. Liu, F. Ye, and H. Xiong, "Multi-class motor imagery EEG classification method with high accuracy and low individual differences based on hybrid neural network," *J. Neural Eng.*, vol. 18, no. 4, Aug. 2021, Art. no. 0460f1.

[48] G. Liu, L. Tian, and W. Zhou, "Multiscale time-frequency method for multiclass motor imagery brain computer interface," *Comput. Biol. Med.*, vol. 143, Apr. 2022, Art. no. 105299.

[49] J. Jin, H. Sun, I. Daly, S. Li, C. Liu, X. Wang, and A. Cichocki, "A novel classification framework using the graph representations of electroencephalogram for motor imagery based brain–computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 20–29, 2022.

N. ALIZADEH received the B.Sc. degree in electrical engineering from Urmia University, Iran, in 2020. She is currently pursuing the master's degree in digital electronics with the Iran University of Science and Technology (IUST), Tehran, Iran. Her research interests include machine learning, deep learning, artificial neural networks, signal processing, and image processing.

S. AFRAKHTEH received the B.Sc. degree in electrical engineering from the Shiraz University of Technology (SUTech), Shiraz, Iran, in 2015, and the M.Sc. degree in digital electronics and the Ph.D. degree in biomedical engineering from the Iran University of Science and Technology (IUST), Tehran, Iran, in 2017 and 2021, respectively. His research interests include machine learning and medical image and signal processing.

M. R. MOSAVI received the B.S., M.S., and Ph.D. degrees in electronic engineering from the Iran University of Science and Technology (IUST), Tehran, Iran, in 1997, 1998, and 2004, respectively. He is currently a Faculty Member (Full Professor) with the Department of Electrical Engineering, IUST. He is the author of more than 450 scientific publications in journals and international conferences in addition to 12 academic books. His research interest includes circuits and systems design. He is the Editor-in-Chief of *Iranian Journal of Marine Technology* and an Editorial Board Member of *Iranian Journal of Electrical and Electronic Engineering* and *GPS Solutions*.

• • •