A Sliding Window Common Spatial Pattern for Enhancing Motor Imagery Classification in EEG-BCI

Pramod Gaur, Harsh Gupta, Anirban Chowdhury, Karl McCreadie, Ram Bilas Pachori, *Senior Member, IEEE*, and Hui Wang, *Senior Member, IEEE*

***Abstract*— Accurate binary classification of electroencephalography (EEG) signals is a challenging task for the development of motor imagery (MI) brain–computer interface (BCI) systems. In this study, two sliding window techniques are proposed to enhance the binary classification of MI. The first one calculates the longest consecutive repetition (LCR) of the sequence of prediction of all the sliding windows and is named SW-LCR. The second calculates the mode of the sequence of prediction of all the sliding windows and is named SW-Mode. Common spatial pattern (CSP) is used for extracting features with linear discriminant analysis (LDA) used for classification of each time window. Both the SW-LCR and SW-Mode are applied on publicly available BCI Competition IV-2a data set of healthy individuals and on a stroke patients' data set. Compared with the existing state of the art, the SW-LCR performed better in the case of healthy individuals and SW-Mode performed better on stroke patients' data set for left- versus right-hand MI with lower standard deviation. For both the data sets, the classification accuracy (CA) was approximately 80% and kappa (κ) was 0.6. The results show that the sliding window-based prediction of MI using SW-LCR and SW-Mode is robust against intertrial and intersession inconsistencies in the time of activation within a trial and thus can lead to a reliable performance in a neurorehabilitative BCI setting.**

***Index Terms*— Brain–computer interface (BCI), common spatial patterns (CSPs), electroencephalography (EEG), linear discriminant analysis (LDA), motor imagery (MI), neurorehabilitation.**

I. Introduction

A BRAIN–COMPUTER interface (BCI) provides a way for people to control external assistive devices and provides a communication pathway for severely motor-impaired people having damaged neuromuscular pathways [1], [2] by translating neurophysiological signals into commands used to control external devices [3]. Motor imagery (MI) is a neuronal activity that occurs when a subject voluntarily imagines making a movement without actually performing it, for example, imagination of the movement of the right hand.

BCI is also being used for motor recovery in patients who have suffered from a stroke or other injury that often leaves them with the inability to move. This is done by measuring changes in neuronal activity over the motor cortex during an MI task while providing them with a reward in the form of a virtual avatar performing the same movement and/or with a robotic device [4]–[7].

It is clear that the topographies and responses corresponding to limb movements obtained from the beta (13–25 Hz) rhythm are distinct in comparison to the mu (8–13 Hz) rhythm. During limb movements, an increase in the oscillatory power is seen in the beta rhythm in the ipsilateral sensorimotor cortex. At the same time, a decrease in oscillatory power is seen in the contralateral sensorimotor cortex of the mu rhythm [8]–[10]. The BCI system aims to translate the changes observed in the mu and beta rhythms into meaningful commands. A major issue in BCI systems reported in the literature is the nonstationarity present in the recorded neurophysiological signals, such as electroencephalography (EEG). This problem occurs when EEG signals originate from multiple sources. Moreover, the recorded EEG signals have a low signal-to-noise ratio (SNR) [11]. There are several other types of artifacts such as electrical power line, electromyogram (EMG), and electrooculogram (EOG) interference that may contribute to low SNR. Therefore, to ensure a high performing system, a high SNR must be achieved from the recorded EEG. This is done by removing artifacts in the preprocessing stage, which helps to increase the feature set separability of multiple MI tasks [2].

Maddirala and Shaik [12] were able to separate useful EEG signal sources from unwanted signals, such as EOG and EMG. A BCI monitoring system with integrated augmented reality (AR) glasses was proposed by Angrisani *et al.* [13], which handles various applications of maintenance and industrial inspection. They built a low-cost wearable and noninvasive monitoring system for single-channel steady-state visual evoked potential (SSVEP)-based BCI [13]. Another group

Manuscript received October 27, 2020; revised December 30, 2020; accepted January 1, 2021. Date of publication January 18, 2021; date of current version February 2, 2021. The Associate Editor coordinating the review process was Mohamad Forouzanfar. *(Corresponding author: Pramod Gaur.)*

Pramod Gaur is with the Department of Computer Science, BITS Pilani Dubai Campus, 345055 Dubai, United Arab Emirates (e-mail: pgaur@dubai.bits-pilani.ac.in).

Harsh Gupta is with the Department of Computer Science and Engineering, LNMIIT, Jaipur 302031, India (e-mail: 18ucs188@lnmiit.ac.in).

Anirban Chowdhury is with the School of Computer Science and Electronic Engineering, University of Essex, Colchester CO4 3SQ, U.K. (e-mail: a.chowdhury@essex.ac.uk).

Karl McCreadie and Hui Wang are with the School of Computing, Engineering and Intelligent Systems, Ulster University, Northern Ireland BT37 0QB, U.K. (e-mail: k.mccreadie@ulster.ac.uk; h.wang@ulster.ac.uk).

Ram Bilas Pachori is with the Department of Electrical Engineering, IIT Indore, Indore 453552, India (e-mail: pachori@iiti.ac.in).

Digital Object Identifier 10.1109/TIM.2021.3051996

developed a wearable EEG-based device used to assess stress conditions in real-time EEG. The amplitude variations in the EEG signals recorded between the left and right regions of prefrontal areas were analyzed using a single EEG differential channel [14]. Another research group also proposed a design using off-the-shelf components with SSVEP-based instrument to improve the performance of a real-time application in BCI [15].

Moreover, EEG signal tends to have poor time–frequency localization, which involves a fixed set of basis functions such as short-time Fourier transform (STFT) and wavelet transform (WT). It also exhibits highly nonstationary characteristics and suffers from high intersubject variability, thus resulting in often poor task detection accuracy and high error rates in BCI systems [16], [17]. Another study [18] studied BCI Competition IV data set 2a using common spatial patterns (CSPs) for spatial filtering and linear discriminant analysis (LDA) for classification, taking a fixed time segment of 2 s from 2.5 to 4.5 s for training and testing [19]. The study examined multiple time segments to see whether it was possible to extract more discriminable features. Recently, a multichannel/single-channel empirical mode decomposition (EMD)-based filtering approach was developed for binary/multiple-class classification problems [20]–[22]. Similarly, a combination of multivariate EMD and canonical correlation analysis (CCA) was explored [23] to denoise and remove artifacts, which helped them to enhance the quality of the EEG signal by enhancing the SNR.

CSP algorithm variants have been studied in the BCI literature [24], [25] to obtain highly separable features in terms of spatial patterns. Another research group computed the feature set as the spatial information present in the sample covariance matrix of the recorded EEG signal [26]. Recently, power spectral covariance matrices were used to detect the sleep stages exploiting the frequency information [27]. There are different areas where these features have helped to obtain better classification, such as radar image processing [28] and image processing [29]. In a recent study, the feasibility of using spiking neural networks (SNNs) has also been demonstrated in classifying MI from EEG signals [30]. Some other studies explored the utility of wavelet packet transformation [31] and feature priority analysis [32].

One of the major drawbacks in commonly found methods as described above is the handling of intersession transfer, which reduces the CA significantly due to inherent nonstationarity in the EEG data. A common approach to solving this problem is to find the most generalizable features. However, the caveat is that the selection of time point within the trial period is generally chosen heuristically. However, such choices are always suboptimal and suffer heavily due to session-to-session and even trial-to-trial variabilities. In order to minimize the effect of such inconsistencies, in this work, two novel sliding window-based CSP techniques have been introduced, which try to articulate the longest consecutive repetition (SW-LCR) and mode (SW-Mode) of the predicted class labels of each window to decode MI tasks performed by the participants. Thus, the LCR and mode-based approaches are independent of the choice of a particular time point, which gives the generalized feature extraction techniques such as CSP an added advantage to augment its performance. The results were obtained for both healthy individuals (using BCI Competition IV-2a data set) and stroke patients (data set reported in [33]) to validate the efficiency of the proposed approaches. The major contributions of this article are highlighted as follows.

1) The proposed LCR-based approach achieved superior performance for left- versus right-hand MI classification compared to the previous methods with reduced intersubject variability in the case of healthy individuals.
2) The proposed mode-based approach significantly ($p < 0.05$) outperformed the existing benchmark in the case of stroke patients data set.
3) Both the LCR and mode-based approaches achieved comparable performance between the healthy individuals' and stroke patients' data sets with higher accuracy (close to 80%), which makes it a suitable method for handling nonstationarity that could also be applicable for neurorehabilitative BCI designs.

This article is organized as follows. Section II describes the data set and methods used in this study. Section III discusses the CSP method. Section IV discusses the LDA classifier. Section V discusses the proposed method and algorithm. Section VI presents the results and discussion, which includes a comparison of the results reported in the literature using the same data sets [18], [33]–[35], while Section VII concludes the study.

II. Data Set

A. Healthy Individuals' Data Set

The BCI Competition IV Data set 2a was used for validating the results on healthy individuals [34]. This data set contains EEG data from nine healthy subjects seated in a comfortable armchair in front of a computer screen. The cue-based BCI paradigm consists of four different MI tasks, namely the imagination of movement of the left hand (class 1), right hand (class 2), both feet (class 3), and tongue (class 4). Each subject has two sessions, namely training and test sessions held on different days. Each session comprises 72 trials per class making a total of 288 trials. At the start of the trial (t = 0 s), a fixation cross is shown on a black screen accompanied by a short acoustic warning tone. At t = 2 s, an arrow is shown for 1.25 s and its direction: left, right, down, or up which corresponds to one of the four classes: 1, 2, 3, or 4. The subjects were asked to perform the MI task until t = 6 s after which a short break was given. The paradigm is shown in Fig. 1(a). EEG signals were recorded monopolarly with the left mastoid serving as a reference and the right mastoid as ground using 22 Ag/AgCl electrodes with interelectrode distances of 3.5 cm. All signals were then bandpass filtered from 0.5 to 100 Hz (with a 50-Hz notch filter enabled). The recorded EEG signals were sampled at a frequency of 250 Hz. For this study, the classification of all six possible combinations of the four classes was used to show the effectiveness of the proposed methods.

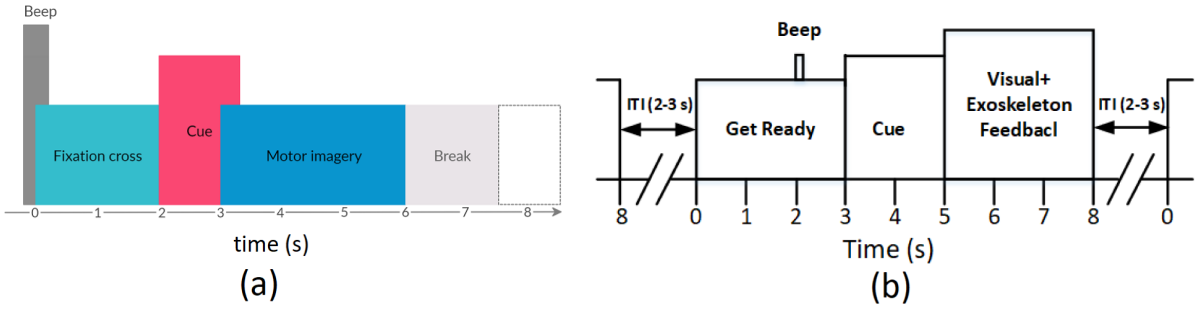

Fig. 1. Timing diagram of the data sets. (a) BCI Competition IV-2a [34]. (b) Stroke patients data set [33].

B. Stroke Patients' Data Set

In order to validate the proposed methodology for neurorehabilitative applications, another data set reported in [33] was also used. The data set consists of left- versus right-hand MI data from ten hemiparetic stroke patients who received multimodal neurofeedback in terms of visual feedback on a computer screen and proprioceptive feedback using a hand exoskeleton, which were contingent on their brain activation. There were two training sessions and an online feedback session each consisting of 40 trials randomly distributed equally into left- and right-hand tasks, i.e., the chance level of prediction was 50%. A trial starts with a "Get Ready" cue, which lasts for 3 s, followed by a cue for indicating the task (left-/right-hand tasks) to perform which lasts for another 5 s during the training session, while during the online feedback session, this period was divided into two parts: the first 2 s was for displaying the cue and the next 3 s was for displaying the visual feedback and providing the exoskeleton-based proprioceptive neurofeedback. There was an intertrial interval (ITI) between 2 and 3 s between two consecutive trials. The timing diagram of the trial is shown in Fig. 1(b). The data acquisition was done at a 500-Hz sampling rate and initially bandpass filtered between 0.1 and 100 Hz with a notch filter at 50 Hz to cancel the power line noise. There were 12 EEG channels, namely, F3, FC3, C3, CP3, P3, FCz, CPz, F4, FC4, C4, CP4, and P4. Only the training data from the first two sessions were used to train the classifier before testing on the data from the online feedback session. A detailed description of the data set can be found in [33].

III. CSP

CSP aims to learn spatial filters that minimize the variance of a class while maximizing the variance of another. It is often helpful to bandpass filter the multichannel EEG signals [19], [36]. The band power in any given frequency band gives the variance of the filtered EEG signals in the selected band. The CSP method obtains optimal discrimination for MI-based BCI tasks based on band-power features [36]. The CSP method uses the spatial filters x by optimizing the function

$$Q(x) = \frac{x^T P_1^T P_1 x}{x^T P_2^T P_2 x} = \frac{x^T C_1 x}{x^T C_2 x} \tag{1}$$

where T signifies the transpose of the matrix and P_i gives the training data matrix with sample points as rows and channels as columns. The spatial covariance matrix for a particular class i is C_i.

There are many ways to solve this optimization problem, but the technique used in this study works by initially visualizing that the function $Q(x)$ is unchanged, if the filter x is rescaled. In fact, $Q(kx) = Q(x)$, where k gives a real constant indicating that the rescaling of filter x is arbitrary. Therefore, minimizing $Q(x)$ is comparable to minimizing $x^T C_1 x$ subject to the constraint $x^T C_2 x = 1$ as there is always a possible way to find a rescaling factor of x such that $x^T C_2 x = 1$. This constrained optimization problem amounts to minimizing the following function using the Lagrange multiplier method:

$$L(\beta, x) = x^T C_1 x - \beta(x^T C_2 x - 1). \tag{2}$$

The derivative of L with regard to x is 0, where β is the Lagrange multiplier, and the filters x minimizing L are such that

$$\frac{\partial L}{\partial x} = 2x^T C_1 - 2\beta x^T C_2 = 0$$
$$\iff C_1 x = \beta C_2 x \iff C_2^{-1} C_1 x = \beta x. \tag{3}$$

Now, this is a standard eigenvalue problem. Hence, the eigenvectors of $Z = C_2^{-1} C_1$ are used to obtain the spatial filters minimizing (1) corresponding to both the largest and the lowest eigenvalues. The features are extracted as the logarithm of EEG signal variance in the selected band after the projection of filters x using the CSP matrix [18].

IV. LDA Classifier

The LDA-based classifier tries to reduce the dimensionality and at the same time protect most of the class discriminatory information [4], [5]. In this article, the proposed method uses an LDA classifier, which is commonly used in EEG-based BCI applications to find the optimum combination of features that provide a better discrimination. For a set of two classes represented by x_1 and x_2, the classification of the n-dimensional sample points $x = \{x_l, x_2, x_3, \ldots, x_n\}$, n_1 samples to the class w_1 and n_2 samples to the class w_2. A line $y = w^t x$ maximizing the discrimination between the two considered classes is tried to archive from a set of all possible lines. In the study, to obtain a good projection vector, the measure of separation between the two classes is required. The mean vector of each class in

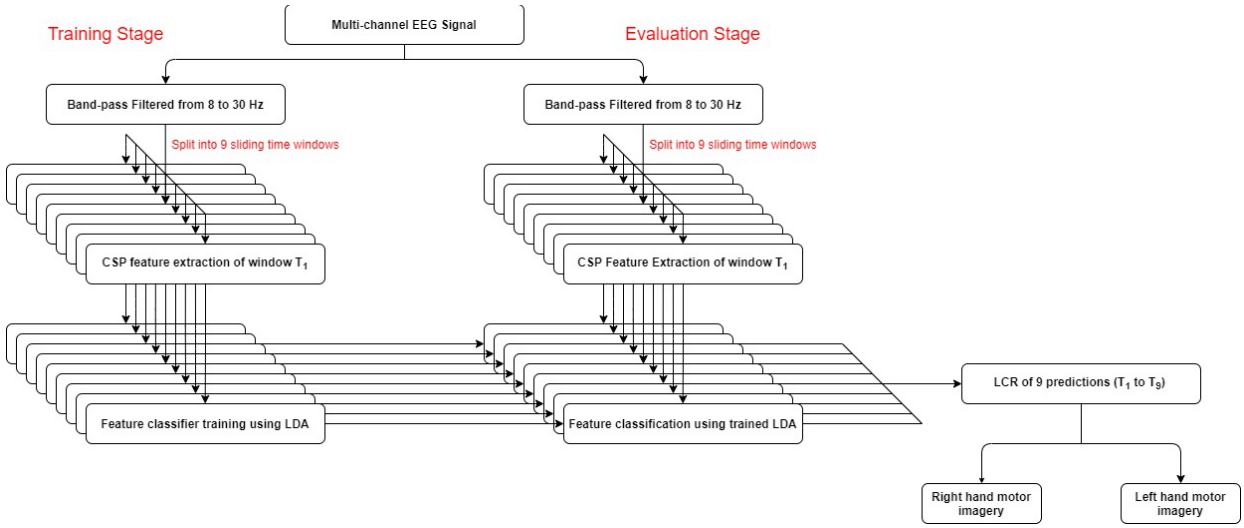

Fig. 2. Block diagram for the proposed method.

x-space and y-space is given by the following equations:

$$\mu_i = \frac{1}{N_i} \sum_{x \in w_i} x \tag{4}$$

and

$$v_i = \frac{1}{N_i} \sum_{y \in w_i} y = \frac{1}{N_i} \sum_{y \in w_i} w^t x = w^t \mu_i. \tag{5}$$

The distance between two projected means is known as the objective function, which is given by the following equation:

$$J(w) = |v_1 - v_2| = |w^t(\mu_1 - \mu_2)|. \tag{6}$$

The standard deviation between classes has not been considered. Therefore, the distance measured between projected means may not always be a good measure. To overcome this limitation, Fisher's LDA classifier has been proposed. This method enhances the LDA classifier by determining a decision boundary or a hyperplane in the feature space to classify the features into distinct classes. It determines the separation boundary between two given distributions in terms of the ratio of two group variances as given by the following equation:

$$J(w) = \frac{\sigma^2_{\text{between}}}{\sigma^2_{\text{within}}} = \frac{w^t(\mu_1 - \mu_2)^2}{w^t S_1 w + w^t S_2 w} \tag{7}$$

where S_1 and S_2 are the variances of the feature distribution between two classes w_1 and w_2, respectively, and μ_1 and μ_2 are the mean of classes. The maximum separation between two classes can be shown by

$$w^* = (S_1 + S_2)^{-1}(\mu_1 - \mu_2). \tag{8}$$

w^* is the weight vector that provides the optimum direction for the projection of the data. The following equation is used by the decision boundary in Fisher's LDA to classify the feature vector $d(m)$ as:

$$p(m) = d(m)w^t + b \tag{9}$$

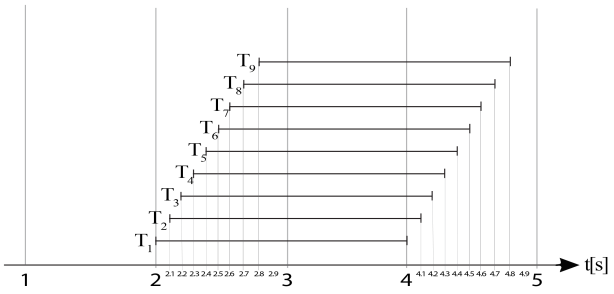

Fig. 3. Sliding time windows T_1–T_9.

where b is the threshold or bias. The features are assigned to one of the classes based on the sign of the $p(m)$ [10].

V. Methodology

A. SW-LCR

In this study, the SW-LCR is used to assign the final prediction. LCR is an element in a sequence, which has the longest subsequence containing just itself, i.e., in "121 111 222," "1111" is the longest subsequence containing only 1. For example, in sequence "112 221 121," 2 is the LCR. If there is more than one candidate for LCR, then the earliest one is chosen. For example, in sequence "121 222 111," 2 as LCR is chosen. The block diagram for the proposed methodology is shown in Fig. 2.

A total of nine overlapping time windows of 2 s, each represented as T_1–T_9 (Fig. 3), were extracted while performing MI tasks. There was a difference of 0.1 s between each successive sliding window, in which the first window (T_1) started at 2.0 s after the start of the trial. Furthermore, the second window (T_2) was from 2.1 to 4.1 s; similarly, the ninth window (T_9) was from 2.8 to 4.8 s. EEG data were bandpass filtered between 8 and 30 Hz. Then, CSP features were extracted by taking three pairs of spatial filters from 100% of the training data. LDA was used to classify CSP features in each sliding

time window. A sequence of nine predicted classes were obtained corresponding to each sliding time window, and the LCR of the sequence was then calculated to assign the final prediction.

B. SW-Mode

Along with LCR, another variant was used called SW-Mode. In this approach, out of the nine overlapping windows as described for SW-LCR, the number of occurrence of classes 1 and 2 is counted. The class containing the maximum count is attributed to that particular trial. This mode-based approach is especially important when the person is unable to concentrate on the task for a long time such as in the case of stroke patients although the overall focus is on the cued task despite fluctuations in concentration. Such a scenario could generate sequences for example, "122 121 211," where the LCR is "2," but the count is greater for "1." Thus, mode can reduce the misclassifications arising due to such a scenario where LCR would be unable to capture this. This notion supported by the results shown in Section VI, where we can see that the SW-Mode performed significantly better than SW-LCR in the case of stroke patients.

A major advantage of the proposed SW-LCR and SW-Mode-based strategies is that it can improve the CA for single-trial prediction irrespective of any feature extraction or classification algorithm used in the background. This is because the proposed sliding window-based method looks for the features at different time segments within the trial rather than focusing on a single time point, which may vary from trial to trial and session to session. Thus, sliding windows capture the idea that change in the mental state is not instantaneous and spread across the whole trial. This is particularly important when the user has a lesser capacity to engage with the task such as for stroke patients for whom the response time after the cue may vary significantly. Looking at all the small windows of time within a trial and then deciding on the particular class of the trial gives more consistent results. However, the disadvantage is that in real-time BCI, the proposed sliding window techniques can add latency in the issued feedback.

VI. Results and Discussion

There are four MI tasks, and the proposed method is used to classify the CSP features into two classes with six combinations of two MI tasks as follows: left and right (L&R), left and foot (L&F), left and tongue (L&T), right and foot (R&F), right and tongue (R&T), and foot and tongue (F&T). The CA is calculated for all nine subjects for each evaluation session.

Tables I and II report the CA obtained on BCI Competition IV datatset 2a by applying the proposed SW-LCR and SW-Mode methods, respectively. For L&R, L&R, and L&T MI tasks, SW-LCR provided better average classification accuracy (ACA) of 80.02%, 83.64%, and 86.19%, respectively, than in SW-Mode. On the other hand, SW-Mode achieved better ACA of 85.03%, 84.03%, and 73.38%, respectively, in R&F, R&T, and F&T than in SW-LCR.

TABLE I

CA (%) and Standard Deviation (Std) for the Proposed SW-LCR Classification Method Applied to BCI Competition IV Data Set 2a

Subject	Accuracy with proposed SW-LCR method (%)						
	L&R	L&F	L&T	R&F	R&T	F&T	Avg
A01	86.81	97.22	97.22	97.22	100	69.44	91.32
A02	64.58	63.89	65.97	80.56	66.67	73.61	69.21
A03	95.83	93.06	94.44	93.06	94.44	69.44	90.05
A04	67.36	82.64	88.19	89.58	86.81	62.5	79.51
A05	68.06	65.97	74.31	70.83	65.97	68.06	68.87
A06	67.36	70.83	72.22	64.58	71.53	70.14	69.44
A07	80.56	97.92	93.75	93.75	93.75	79.86	89.93
A08	97.22	84.72	92.36	88.19	89.58	78.47	88.43
A09	92.36	96.53	97.22	84.03	82.64	85.42	89.70
Average	80.02	83.64	86.19	84.64	83.49	72.99	
Std	13.45	13.75	12.02	10.95	12.65	7.07	

TABLE II

CA (%) and Standard Deviation (Std) for the Proposed SW-Mode Classification Method Applied to BCI Competition IV Data Set 2a

Subject	Accuracy with proposed SW-Mode method (%)						
	L&R	L&F	L&T	R&F	R&T	F&T	Avg
A01	86.11	96.53	96.53	97.22	100	70.83	91.20
A02	64.58	64.58	65.97	80.56	66.67	75	69.56
A03	95.83	93.06	94.44	93.06	95.14	70.14	90.28
A04	64.58	84.72	88.89	90.28	89.58	63.19	80.21
A05	68.06	65.28	75	72.22	66.67	67.36	69.10
A06	68.75	70.14	70.83	64.58	71.53	70.14	69.33
A07	81.94	97.92	93.06	93.75	93.75	79.86	90.05
A08	97.22	84.03	92.36	88.89	90.28	78.47	88.54
A09	90.97	97.22	97.22	84.72	82.64	85.42	89.70
Average	79.78	83.72	86.03	85.03	84.03	73.38	
Std	13.48	13.82	12.04	10.8	12.77	6.92	

TABLE III

Kappa Value for the Proposed SW-LCR Classification Method Applied to BCI Competition IV Data Set 2a

Subject	Accuracy with proposed SW-LCR method (%)						
	L&R	L&F	L&T	R&F	R&T	F&T	Avg
A01	0.74	0.94	0.94	0.94	1	0.39	0.83
A02	0.29	0.28	0.32	0.61	0.33	0.47	0.38
A03	0.92	0.86	0.89	0.86	0.89	0.39	0.80
A04	0.35	0.65	0.76	0.79	0.74	0.25	0.59
A05	0.36	0.32	0.49	0.42	0.32	0.36	0.38
A06	0.35	0.42	0.44	0.29	0.43	0.4	0.39
A07	0.61	0.96	0.88	0.88	0.88	0.6	0.80
A08	0.94	0.69	0.85	0.76	0.79	0.57	0.77
A09	0.85	0.93	0.94	0.68	0.65	0.71	0.79
Average	0.6	0.67	0.72	0.69	0.67	0.46	

The corresponding kappa values against these CAs have been reported in Tables III and IV for SW-LCR and SW-Mode, respectively.

Table V shows the left and right MI CA of the proposed methods compared to competing comparable methods M1–M4. We have chosen L&R over other binary class combinations for comparison as this is widely used in many BCI applications. Another reason is that the stroke patients' data used in this article are also based on the left- versus right-hand task and it would be easier to compare the results for healthy versus patients in that manner.

TABLE IV

KAPPA VALUE FOR THE SW-MODE CLASSIFICATION METHOD APPLIED TO BCI COMPETITION IV DATA SET 2A

Subject	Accuracy with proposed SW-Mode method (%)						
	L&R	L&F	L&T	R&F	R&T	F&T	Avg
A01	0.72	0.93	0.93	0.94	1	0.42	0.82
A02	0.29	0.29	0.32	0.61	0.33	0.5	0.39
A03	0.92	0.86	0.89	0.86	0.9	0.4	0.81
A04	0.29	0.69	0.78	0.81	0.79	0.26	0.60
A05	0.36	0.31	0.5	0.44	0.33	0.35	0.38
A06	0.38	0.4	0.42	0.29	0.43	0.4	0.39
A07	0.64	0.96	0.86	0.88	0.88	0.6	0.80
A08	0.94	0.68	0.85	0.78	0.81	0.57	0.77
A09	0.82	0.94	0.94	0.69	0.65	0.71	0.79
Average	0.6	0.67	0.72	0.7	0.68	0.47	

TABLE V

CA (%) FOR LEFT- VERSUS RIGHT-HAND MI TASK USING SW-LCR AND SW-MODE COMPARED TO METHODS M1–M4 ON BCI COMPETITION IV DATA SET 2A

Subject	SW-LCR	SW-Mode	M1	M2	M3	M4
A01	86.81	86.11	88.65	**91.49**	90.28	88.89
A02	**64.58**	**64.58**	61.27	60.56	57.64	51.39
A03	95.83	95.83	91.24	94.16	95.14	**96.53**
A04	67.36	64.58	74.14	**76.72**	65.97	70.14
A05	**68.06**	**68.06**	57.04	58.52	61.11	54.86
A06	67.36	68.75	69.44	68.52	65.28	**71.53**
A07	80.56	**81.94**	60	78.57	61.11	81.25
A08	**97.22**	**97.22**	94.03	97.01	91.67	93.75
A09	92.36	90.97	83.85	**93.85**	86.11	93.75
Average	**80.02**	79.78	75.52	79.93	74.92	78.01
Std	**13.45**	13.48	14.39	14.99	15.43	17.01

Method 1 (M1) shows the results based on a tangent space-based transfer learning technique by Gaur *et al.* [21]. The ACA of the SW-LCR ($80.02\% \pm 13.45$) is higher when compared with M1 ($75.52\% \pm 14.39$). Notably, six of the nine subjects have shown improvement, with an improved ACA of 4.5%. Also, the standard deviation (Std) of the accuracy of all subjects has been reduced by roughly 1%. Subject A07 shows an improvement of $>20.5\%$, whereas A05 shows the improvement of $>11\%$, and A09 improves $>8.5\%$. Three subjects (A02, A03, and A08) show improvement of $>3\%$, with subjects A01 and A06 declining in performance by 2.1% and A04 by 7%.

Method 2 (M2) shows the results obtained from the evaluation session from a different study by Gaur *et al.* [20], which uses a subject-specific multivariate empirical mode decomposition-based filtering method (SS-MEMDBF) for preprocessing and implements a Riemannian geometry (RG) framework designed separately for each subject for classification. The ACA of the SW-LCR ($80.02\% \pm 13.45$) is comparable to method M2 ($79.93\% \pm 14.99$). Notably, five out of the nine subjects have shown improvement, with an improved ACA of 0.1%, and also, the Std of the accuracy of all subjects has been reduced by roughly 1.5%. Subject A05 shows an improvement of 9.5%, whereas subject A02 improved by 4%. Three subjects (A03, A07, and A08) show improvement $<2\%$, whereas subjects A06 and A09 declined in performance by $<1.5\%$. A01 has declined by $>4.5\%$ and A04 has declined by $>9\%$. The performance of the SW-LCR is comparable to that of M2 in four out of six combinations although the average accuracy is slightly lower than M2. However, more importantly, it can be observed that the SW-LCR can lead to more uniformity in performance across all the subjects as the standard deviation is lower compared to M2. This is because the LCR-based sliding window approach can enhance the performance of low performing subjects such as A02 and A05 as they suffer from lower SNR and higher nonstationarity in intersession transfer. The use of fixed time points, such as 0.5–2.5 s in the case of M2, is in fact disadvantageous to handle such intersession nonstationarities as the time window is selected based on the training session, which is prone to be shifted during the evaluation session. On the contrary, the SW-LCR does not depend on such an assumption and observes a larger time scale through successive windowing and the label of that particular trial is decided from the outcomes of those windows using LCR. Thus, the small shifts in the time point of activation between training and evaluation sessions are accounted for. M2 also comes with an overhead of an SS-MEMDBF for preprocessing making it difficult to implement online due to higher computational complexity. Another advantage of the LCR-based sliding window approach is that it can be used with any feature extraction or classification algorithms, and therefore, limitations that come with the use of a particular feature type (such as in M2) can be avoided.

Method 3 (M3) implements CSP and uses detection of the covariate shift and adaptive learning [37]. The ACA of the SW-LCR ($80.02\% \pm 13.45$) is better when compared with method M3 ($74.92\% \pm 15.43$). Notably, eight out of the nine subjects have shown improvement, an ACA improvement $>5\%$, and also, the Std of the accuracy of all subjects has been reduced by $<2\%$. Subject A07 shows the improvement of $>19.4\%$, whereas A05 and A02 have shown the improvement of $>6.9\%$. A09 shows the improvement of $>6.2\%$, whereas A08 shows the improvement of $>5.5\%$. A06 shows the improvement of $>2\%$, with two subjects (A03 and A04) showing an improvement of $<1.5\%$ with A01 declining by $<3.5\%$.

Method 4 (M4) uses CSP on bandpass filtered EEG between 8 and 30 Hz before calculating the log variance from three pairs of filters for extraction of features and uses LDA for a binary classification problem [18]. The ACA of the SW-LCR ($80.02\% \pm 13.45$) is higher when compared with method M4 ($78.01\% \pm 17.01$). Notably, three out of the nine subjects have shown improvement, with the ACA improving by $>2\%$, and also, the Std of the accuracy of all subjects has been reduced by $>3.5\%$. A05 and A02 have shown the improvement of $>13.1\%$, whereas A08 shows the improvement of $>3.4\%$. A06 has declined by $<4.2\%$. The remaining five subjects are within 3% of the results.

Apart from M1 to M4, there are a few emerging techniques that were also recently applied on the BCI Competition IV-2a data set. For example, an SNN-based [30] technique yielded an ACA of 75.62% for left- versus right-hand MI task that is lower than the proposed methods of SW-LCR (ACA = 80.02%) and SW-Mode (ACA = 79.78%). Among

TABLE VI

CAs (%) AND KAPPA VALUES FOR THE STROKE PATIENTS' DATA SET USING **SW-LCR**

Subject	5-CV Training Accuracy	Training kappa	Testing Accuracy	Testing kappa
A01	68.75	0.38	85	0.7
A02	87.5	0.75	80	0.6
A03	87.5	0.75	62.5	0.25
A04	81.25	0.63	80	0.6
A05	93.75	0.88	60	0.2
A06	100	1	52.5	0.05
A07	100	1	80	0.6
A08	93.75	0.88	82.5	0.65
A09	81.25	0.63	70	0.4
A10	87.5	0.75	82.5	0.65
Average	88.13	0.77	73.5	0.47
Std	9.52	0.19	11.44	0.23

TABLE VII

CAs (%) AND KAPPA VALUES FOR THE STROKE PATIENTS' DATA SET USING **SW-MODE**

Subject	5-CV Training Accuracy	Training kappa	Testing Accuracy	Testing kappa
A01	81.25	0.61	87.5	0.75
A02	100	1	87.5	0.75
A03	100	1	65	0.3
A04	75	0.5	70	0.4
A05	87.5	0.74	75	0.5
A06	93.75	0.87	65	0.3
A07	87.5	0.75	90	0.8
A08	75	0.49	90	0.8
A09	81.25	0.61	75	0.5
A10	93.75	0.87	95	0.9
Average	87.5	0.74	80	0.6
Std	9.32	0.19	11.24	0.22

the deep learning-based approaches, EEGNet achieved an ACA of 68.98% and NSL-EEGNet [38] achieved 70.6%, which are significantly ($p < 0.05$) lower than the proposed SW-LCR and SW-Mode-based approaches.

The results on the stroke patients' data set are given in Tables VI and VII for the SW-LCR and SW-Mode-based approaches, respectively. The reason we have chosen this data set is that the stroke patients' data have more nonstationarity due to altered neurodynamics and lack of engagement during the task. It can be seen from Table VI that in the case of SW-LCR, the ACA in training (fivefold cross validation) is 88.13% $\pm$ 9.52, whereas in testing, it is 73.5% $\pm$ 11.44. The average kappa value in training (fivefold cross validation) is 0.77 $\pm$ 0.19, whereas for testing, it is 0.47 $\pm$ 0.23. Again, for the SW-Mode-based approach (see Table VII), the ACA in training (fivefold cross validation) is 87.5% $\pm$ 9.32, whereas in testing, it is 80% $\pm$ 11.24. The average kappa value in training (fivefold cross validation) is 0.74 $\pm$ 0.19, whereas for testing, it is 0.6 $\pm$ 0.22. It is important to note that the previous benchmark accuracy on the same data set as reported in [33] was 70.25% with a kappa of 0.41, which is significantly ($p < 0.05$) outperformed by the proposed method in the case of the SW-Mode-based approach (average accuracy 80% with kappa 0.6), and in the case of SW-LCR, also the performance is higher (average accuracy 73.5% with kappa 0.47). A deep learning-based approach was also applied on the same data set

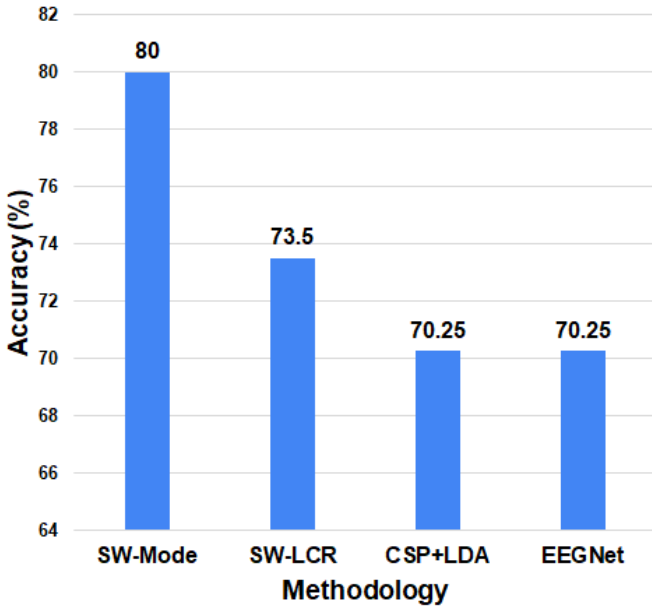

Fig. 4. Comparison of CAs of the proposed methods with previous methods applied on the stroke patients' data set.

previously [35] using the popular EEGNet architecture. They obtained an ACA of 70.25% across all the subjects, which is also outperformed significantly ($p < 0.05$) by the proposed method in the case of the SW-Mode-based approach (average accuracy 80%). These comparisons are shown in Fig. 4 with a bar graph.

It is also interesting to note that the proposed method gave a comparable performance in healthy (BCI Competition IV Data Set 2a) and stroke patients' data sets, which is an indication of the potential robustness of the algorithm. The average accuracy and kappa in the case of healthy individuals are around 80.02% with kappa at 0.6, while in the case of patients, the average accuracy is at 80% (using SW-Mode) with the same kappa value (0.6). The comparison of the distribution of accuracies across different subjects for the proposed methods (SW-LCR and SW-Mode) in the case of healthy and patients is shown in the boxplot given in Fig. 5. However, it can be seen that SW-LCR performed better in the case of healthy individuals than in the case of stroke patients, whereas SW-Mode was better for stroke patients. This might be related to lack of focus during the task often observed in stroke patients as they can get tired quickly and fail to generate LCR of predicted labels needed for good SW-LCR results. On the other hand, as SW-Mode depends on the most frequent predicted label within a trial rather than depending on the longest sequence, it can handle the issue of lack of focus more effectively.

In this study, the proposed method (M1) was compared with three other methods denoted in the manuscript as M2–M4. Among these methods, M1, M3, and M4 are based on CSPs with some variations, whereas M2 is based on multivariate EMD (MEMD)-based filtering and RG. In comparison to the CSP-based methods, the computational time for MEMD is enormous, which impedes its utility for real-time

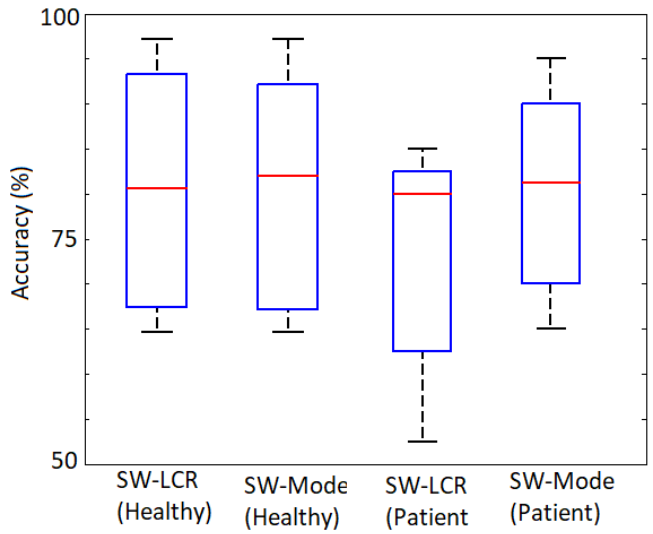

Fig. 5. Comparison of CA distributions of the proposed methods between the healthy and stroke patients' data sets.

application [39]. For example, it takes around 1 h to execute the EMD with 30 000 data points using a standard modern personal computer as reported in [40]. This creates additional dependence on parallel computing architecture such as Cuda or hardware resources such as FPGA [41]. The typical time complexity of MEMD is $O(n \log n)$, where n is the data length [40]. In addition, M2 requires the computation of RG with complexity $O(n^3)$ (n = number of EEG electrodes), which makes the real-time implementation more difficult. The other methods M1, M3, and M4 compute the features using precomputed CSP weight matrix generated during the calibration process and requires only matrix multiplication of the bandpass filtered EEG data with the CSP weight matrix, the processing time of which is in milliseconds. For example, making every single prediction in M1 required around 62.5 ms, which is good enough for issuing intuitive neurofeedback. The hardware used in the experiments was a laptop with Intel® Core i5-8250U as CPU and 8-GB DDR4 RAM.

VII. Conclusion

EEG data are often recorded across session as well as across subject (sometimes from those with brain injuries) and tend to be highly nonstationary with intersubject variability. In this work, two sliding window-based CSP techniques, namely SW-LCR and SW-Mode, were introduced to classify EEG signals into multiple MI tasks. The proposed methods take advantage of examining multiple time segments within a trial rather than depending on a single time point, which is more robust against intertrial and intersession variabilities. With the proposed method, intersubject variability has been reduced to some extent as evidenced by the lower standard deviation. Moreover, the SW-Mode has performed significantly better than the existing state of the art on the stroke patients' data set, whereas SW-LCR performed better on healthy individuals. The performance of the proposed methods is also comparable for both healthy individuals and stroke patients with accuracy around 80% with a higher kappa value of 0.6, which is an important observation since the EEG data from stroke patients suffer from lower SNR and higher nonstationarity due to the presence of altered neurodynamics. Thus, SW-LCR and SW-Mode have the potential to augment the performance of MI-BCI and can pave the way for reliable neurorehabilitative BCI applications.

References

[1] J. R. Wolpaw *et al.*, "Brain-computer interface technology: A review of the first international meeting," *IEEE Trans. Rehabil. Eng.*, vol. 8, no. 2, pp. 164–173, Jun. 2000.

[2] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain–computer interfaces for communication and control," *Clin. Neurophysiol.*, vol. 113, no. 6, pp. 767–791, 2002.

[3] J. Wolpaw and E. W. Wolpaw, *Brain-Computer Interfaces: Principles and Practice*. London, U.K.: OUP, 2012.

[4] F. Lotte, "A tutorial on EEG signal-processing techniques for mental-state recognition in brain–computer interfaces," in *Guide to Brain-Computer Music Interfacing*. London, U.K.: Springer, 2014, pp. 133–161.

[5] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi, "A review of classification algorithms for EEG-based brain–computer interfaces," *J. Neural Eng.*, vol. 4, no. 2, p. R1, 2007.

[6] A. Chowdhury *et al.*, "Active physical practice followed by mental practice using BCI-driven hand exoskeleton: A pilot trial for clinical effectiveness and usability," *IEEE J. Biomed. Health Informat.*, vol. 22, no. 6, pp. 1786–1795, Nov. 2018.

[7] A. Chowdhury, A. Dutta, and G. Prasad, "Corticomuscular co-activation based hybrid brain-computer interface for motor recovery monitoring," *IEEE Access*, vol. 8, pp. 174542–174557, 2020.

[8] P. Herman, G. Prasad, T. M. McGinnity, and D. Coyle, "Comparative analysis of spectral approaches to feature extraction for EEG-based motor imagery classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 16, no. 4, pp. 317–326, Aug. 2008.

[9] V. Gandhi, G. Prasad, D. Coyle, L. Behera, and T. M. McGinnity, "Quantum neural network-based EEG filtering for a brain–computer interface," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 25, no. 2, pp. 278–288, Feb. 2014.

[10] P. Gaur, R. B. Pachori, H. Wang, and G. Prasad, "An empirical mode decomposition based filtering method for classification of motor-imagery EEG signals for enhancing brain-computer interface," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2015, pp. 1–7.

[11] L. F. Nicolas-Alonso and J. Gomez-Gil, "Brain computer interfaces, a review," *Sensors*, vol. 12, no. 2, pp. 1211–1279, 2012.

[12] A. K. Maddirala and R. A. Shaik, "Separation of sources from single-channel EEG signals using independent component analysis," *IEEE Trans. Instrum. Meas.*, vol. 67, no. 2, pp. 382–393, Feb. 2018.

[13] L. Angrisani, P. Arpaia, A. Esposito, and N. Moccaldi, "A wearable brain–computer interface instrument for augmented reality-based inspection in industry 4.0," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 4, pp. 1530–1539, Apr. 2020.

[14] P. Arpaia, N. Moccaldi, R. Prevete, I. Sannino, and A. Tedesco, "A wearable EEG instrument for real-time frontal asymmetry monitoring in worker stress analysis," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 10, pp. 8335–8343, Oct. 2020.

[15] L. Angrisani, P. Arpaia, D. Casinelli, and N. Moccaldi, "A single-channel SSVEP-based instrument with off-the-shelf components for trainingless brain-computer interfaces," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 10, pp. 3616–3625, Oct. 2019.

[16] B. H. Yilmaz, C. M. Yilmaz, and C. Kose, "Diversity in a signal-to-image transformation approach for EEG-based motor imagery task classification," *Med. Biol. Eng. Comput.*, vol. 58, no. 2, pp. 443–459, Feb. 2020.

[17] S. R. Sreeja, Himanshu, and D. Samanta, "Distance-based weighted sparse representation to classify motor imagery EEG signals for BCI," *Multimedia Tools Appl.*, vol. 79, no. 19, pp. 13775–13793, Feb. 2020.

[18] F. Lotte and C. Guan, "Regularizing common spatial patterns to improve BCI designs: Unified theory and new algorithms," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 2, pp. 355–362, Feb. 2011.

[19] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K.-R. Müller, "Optimizing spatial filters for robust EEG single-trial analysis," *IEEE Signal Process. Mag.*, vol. 25, no. 1, pp. 41–56, Jan. 2008.

[20] P. Gaur, R. B. Pachori, H. Wang, and G. Prasad, "A multi-class EEG-based BCI classification using multivariate empirical mode decomposition based filtering and Riemannian geometry," *Expert Syst. Appl.*, vol. 95, pp. 201–211, Apr. 2018.

[21] P. Gaur, K. McCreadie, R. B. Pachori, H. Wang, and G. Prasad, "Tangent space features-based transfer learning classification model for two-class motor imagery brain–computer interface," *Int. J. Neural Syst.*, vol. 29, no. 10, Dec. 2019, Art. no. 1950025.

[22] P. Gaur, R. B. Pachori, H. Wang, and G. Prasad, "An automatic subject specific intrinsic mode function selection for enhancing two-class EEG-based motor imagery-brain computer interface," *IEEE Sensors J.*, vol. 19, no. 16, pp. 6938–6947, Aug. 2019.

[23] X. Chen, X. Xu, A. Liu, M. J. McKeown, and Z. J. Wang, "The use of multivariate EMD and CCA for denoising muscle artifacts from few-channel EEG recordings," *IEEE Trans. Instrum. Meas.*, vol. 67, no. 2, pp. 359–370, Feb. 2018.

[24] K. K. Ang, Z. Y. Chin, C. Wang, C. Guan, and H. Zhang, "Filter bank common spatial pattern algorithm on BCI competition IV datasets 2a and 2b," *Frontiers Neurosci.*, vol. 6, p. 39, Mar. 2012.

[25] H. Zhang, H. Yang, and C. Guan, "Bayesian learning for spatial filtering in an EEG-based brain–computer interface," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 24, no. 7, pp. 1049–1060, Jul. 2013.

[26] A. Barachant, S. Bonnet, M. Congedo, and C. Jutten, "Multiclass brain–computer interface classification by Riemannian geometry," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 4, pp. 920–928, Apr. 2012.

[27] Y. Li, K. M. Wong, and H. DeBruin, "EEG signal classification based on a Riemannian distance measure," in *Proc. IEEE Toronto Int. Conf. Sci. Technol. Humanity (TIC-STH)*, Sep. 2009, pp. 268–273.

[28] F. Barbaresco, "Innovative tools for radar signal processing based on Cartan's geometry of SPD matrices & information geometry," in *Proc. IEEE Radar Conf.*, May 2008, pp. 1–6.

[29] O. Tuzel, F. Porikli, and P. Meer, "Pedestrian detection via classification on Riemannian manifolds," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 10, pp. 1713–1727, Oct. 2008.

[30] C. D. G. Virgilio, J. H. A. Sossa, J. M. Antelis, and L. E. Falcón, "Spiking neural networks applied to the classification of motor tasks in EEG signals," *Neural Netw.*, vol. 122, pp. 130–143, Feb. 2020.

[31] 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, p. 7, 2020.

[32] S. Li and H. Feng, "EEG signal classification method based on feature priority analysis and CNN," in *Proc. Int. Conf. Commun., Inf. Syst. Comput. Eng. (CISCE)*, Jul. 2019, pp. 403–406.

[33] A. Chowdhury, H. Raza, Y. K. Meena, A. Dutta, and G. Prasad, "Online covariate shift detection-based adaptive brain–computer interface to trigger hand exoskeleton feedback for neuro-rehabilitation," *IEEE Trans. Cognit. Develop. Syst.*, vol. 10, no. 4, pp. 1070–1080, Dec. 2018.

[34] 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., Graz, Austria, Tech. Rep., 2008, pp. 136–142. [Online]. Available: http://www.bbci.de/competition/iv/desc_2a.pdf

[35] H. Raza, A. Chowdhury, and S. Bhattacharyya, "Deep learning based prediction of EEG motor imagery of stroke patients' for neuro-rehabilitation application," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2020, pp. 1–8.

[36] H. Ramoser, J. Muller-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.

[37] H. Raza, H. Cecotti, Y. Li, and G. Prasad, "Adaptive learning with covariate shift-detection for motor imagery-based brain–computer interface," *Soft Computing*, vol. 20, no. 8, pp. 1–12, 2015.

[38] H. Raza, A. Chowdhury, S. Bhattacharyya, and S. Samothrakis, "Single-trial EEG classification with EEGNet and neural structured learning for improving BCI performance," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2020, pp. 1–8.

[39] Y.-H. Wang, C.-H. Yeh, H.-W.-V. Young, K. Hu, and M.-T. Lo, "On the computational complexity of the empirical mode decomposition algorithm," *Phys. A, Stat. Mech. Appl.*, vol. 400, pp. 159–167, Apr. 2014.

[40] L.-C. Wu *et al.*, "A novel preprocessing method using Hilbert Huang transform for MALDI-TOF and SELDI-TOF mass spectrometry data," *PLoS ONE*, vol. 5, no. 8, pp. 1–15, Aug. 2010.

[41] M.-H. Lee, K.-K. Shyu, P.-L. Lee, C.-M. Huang, and Y.-J. Chiu, "Hardware implementation of EMD using DSP and FPGA for online signal processing," *IEEE Trans. Ind. Electron.*, vol. 58, no. 6, pp. 2473–2481, Jun. 2011.