

RESEARCH

Open Access

SPECTRA: a tool for enhanced brain wave signal recognition



Shiu Kumar^{1*} , Tatsuhiko Tsunoda^{2,3,4†} and Alok Sharma^{2,3,5,6†}

From 19th International Conference on Bioinformatics 2020 (InCoB2020)
Virtual. 25-29 November 2020

*Correspondence:
shiu.kumar@fnu.ac.fj
†Tatsuhiko Tsunoda and Alok
Sharma authors have shared
last authorship
¹ School of Electrical
and Electronics Engineering,
Fiji National University, Suva,
Fiji
Full list of author information
is available at the end of the
article

Abstract

Background: Brain wave signal recognition has gained increased attention in neuro-rehabilitation applications. This has driven the development of brain-computer interface (BCI) systems. Brain wave signals are acquired using electroencephalography (EEG) sensors, processed and decoded to identify the category to which the signal belongs. Once the signal category is determined, it can be used to control external devices. However, the success of such a system essentially relies on significant feature extraction and classification algorithms. One of the commonly used feature extraction technique for BCI systems is common spatial pattern (CSP).

Results: The performance of the proposed spatial-frequency-temporal feature extraction (SPECTRA) predictor is analysed using three public benchmark datasets. Our proposed predictor outperformed other competing methods achieving lowest average error rates of 8.55%, 17.90% and 20.26%, and highest average kappa coefficient values of 0.829, 0.643 and 0.595 for BCI Competition III dataset IVa, BCI Competition IV dataset I and BCI Competition IV dataset IIb, respectively.

Conclusions: Our proposed SPECTRA predictor effectively finds features that are more separable and shows improvement in brain wave signal recognition that can be instrumental in developing improved real-time BCI systems that are computationally efficient.

Keywords: Brain computer interface (BCI), Common spatial pattern (CSP), Common spatio-spectral pattern (CSP), Electroencephalography (EEG), Motor imagery (MI), Tangent space mapping (TSM)

Introduction

Brain-computer interface (BCI) is one of the emerging technologies for neuro-rehabilitation that offers paralyzed people a non-muscular channel of control and communication to the external world [1, 2]. Complex and unique brain wave patterns are generated for each different brain activity and it is very difficult to manually decode and identify the different categories of these brain wave signals. Therefore, many research work [3–9]



are being carried out to automatically identify the different categories of the brain wave signals with high accuracy as it is useful in many applications such as seizure detection [10, 11], sleep stage recognition [12], emotion or stress recognition [13, 14], neuro-rehabilitation [15, 16], and gaming [17, 18].

Non-invasive electroencephalography (EEG) sensors are usually placed around the scalp to capture the brain waves generated by the brain activities. The EEG signals can then be mapped to various commands for controlling external devices after a chain of refined signal processing and machine learning procedures such as filtering, feature extraction and classification of the EEG signal. Thus, deliberately generating different brain wave patterns will enable individuals to control external devices.

P300 [1, 19–21] and motor imagery (MI) [18, 22–25] are the two methods for obtaining the EEG signal for EEG-based BCI systems. In P300, usually the parietal and occipital areas are used to obtain the distinctive EEG signals 300 ms after the visual stimulus. This study focuses on the MI-based BCI that uses the sensorimotor rhythms. The Mu and Beta rhythms are recorded from the sensorimotor cortex region of the scalp producing different patterns for different MI tasks, which can be processed and used for BCI control.

In MI-based BCI systems there are several problems that need to be taken care of such as the pre-processing algorithm for noise removal/reduction, selection of the frequency band(s), feature extraction and classification. Filtering is usually applied as the pre-processing algorithm. With a vast range of filtering methods available, common average filtering [26, 27], Laplacian filtering [28, 29] and FIR bandpass filtering [30] are the most commonly used filtering methods. A number of feature extraction methods [31–43] and classification algorithms [23, 32, 44, 45] have been proposed due to the fact that the reliability and feasibility of MI-based BCI systems largely depend on robust and effective feature extraction and classification of EEG signals. Common spatial pattern (CSP) [9, 23, 36, 37, 39, 41–43, 46–55] has been widely used for feature extraction of EEG signals for MI-based BCIs. The selection of frequency bands plays a major role in extracting significant CSP features from MI EEG signals. The optimal frequency bands are generally subject-dependent and manually tuning the frequency band is a challenging and tedious task. To tackle this problem, various frequency band selection approaches have been proposed [9, 38, 50, 56–61]. Novi et al. [59] proposed sub-band common spatial pattern (SBCSP) approach. In their approach, EEG signals are decomposed into multiple non-overlapping sub-bands and the CSP features extracted from each sub-band are fused together and used for classification. Filter bank CSP (FBCSP) was proposed by Ang et al. [56], which uses multiple overlapping sub-bands for decomposing the EEG signals into a number of sub-bands. The CSP features obtained from these sub-bands are then fused together and feature selection is employed for selecting important features. To improve the FBCSP approach, a discriminative FBCSP (DFBCSP) [61] approach was proposed. It utilizes Fishers ratio for choosing the significant subject-dependent sub-bands. Wei and Wei [38] proposed a binary particle swarm optimization method for selecting significant sub-bands from a set of pre-determined sub-bands.

A sparse filter bank CSP (SFBCSP) [62] approach, which utilized multiple sub-bands for optimizing the sparse patterns has also been proposed. Sparse Bayesian learning is increasingly gaining widespread attention and used for various purposes such as

feature selection [42] and classification [63]. Zhang et al. [42] proposed a sparse Bayesian learning of filter bank (SBLFB) approach in which sparse Bayesian learning is used for automatically selecting the significant features. A spatial-frequency-temporal optimized feature sparse representation based classification (SFTOFSRC) [36] has been proposed with a focus of optimizing CSP features in subject adapted space-frequency-time patterns.

In this work, we mainly focus on the feature extraction process. Feature extraction is one of the essential steps in machine learning and signal processing having a vast impact on the performance of the algorithms in these fields. The extraction of significant features is essential as selecting redundant or insignificant features will degrade the performance of the system. This work extends our previous work on CSP-TSM (tangent space mapping) [64] approach. In the CSP-TSM approach, a single window is used to extract the CSP and TSM features followed by feature selection using least absolute shrinkage and selection operator (Lasso). In this paper, we propose to use multiple temporal delayed windows to extract features that are more separable. Using multiple windows give rise to problems such as the window size and number of windows to use. Therefore, these problems are also addressed in this work. Furthermore, we take advantage of the common spatio-spectral pattern (CSSP) approach. In CSSP, a temporal time delayed signal is inserted to the raw signal. The value of time delay τ used also influences the performance of the CSSP algorithm. The problem of selecting the appropriate τ value is also addressed in this work. Thus, this work combines the CSP-TSM and CSSP approaches to take advantage of both approaches which boosts the performance of the overall system.

The TSM approach uses Riemannian distance to Riemannian mean which provides superior information about the class membership compared to the CSP approach that utilises Euclidean distance to its mean. On the other hand, the CSSP approach improves the spatial resolution of the signal. Therefore, taking advantage of these approaches i.e. appropriately combining CSP-TSM and CSSP approaches should yield features that are more effective and significant in classifying the MI EEG signals. To validate and compare our approach with other competing methods, public benchmark datasets: BCI Competition III dataset IVa, BCI Competition IV dataset I and BCI Competition IV dataset IIb have been used. The proposed scheme successfully extracts more significant features which accounts for the reduced error rates achieved (refer to results section) for all three datasets. Promising results are obtained, thus the proposed scheme can play a key role in developing improved MI-based BCI systems.

The main contributions of this work are as follows:

- We have combined CSSP with the CSP-TSM approach resulting in CSSP-TSM. TSM is retained as it gives superior information about the class membership while the use of CSSP improves the spatial resolution of the signal and thus further boosts the overall performance of the system.
- Use of CSSP involves inserting a temporal delayed window to the trial signal. Therefore, we have proposed the use of multiple overlapping temporal windows for extracting more significant features. We have addressed the problem of the number of multiple windows to use with the proposed scheme and how the multiple windows obtained can be combined to result in CSP-TSM and CSSP-TSM approaches

for improved performance. Also, the time delay τ used influences the performance of the system and varies among different subjects. Therefore, a cross-validation approach has been proposed for the selection of time delay τ in order to obtain optimal performance for each of the subjects.

- Several feature selection methods have been evaluated to determine which feature selection method is best for selecting significant features. Use of F-score showed superior performance in selecting the significant features over other feature selection methods (Lasso—used in original CSP-TSM approach, mutual information, and sparse Bayesian learning). Thus, F-score is recommended for feature selection over other methods that have been evaluated and have been used in this work.

Results

The processing in this work has been carried out using Matlab. Moreover, all training and test have been performed using each subject's data, i.e. the data from other subjects are not used. In this study, the MI EEG data between 0.5 and 2.5 s (i.e. 200 sample points for dataset 1 and 2, and 500 sample points for dataset 3) after the visual cue have been extracted and used for further processing to obtain the results of all the competing methods. Common average referencing has been used as the pre-processing step for each individual EEG trial. A Butterworth bandpass filter has been used for filtering and classification is done using the SVM classifier (which is trained using training data) for all the methods. A 7–30 Hz wide band have been used for the conventional CSP approach. To make a fair comparison six spatial filters have been used for all the methods while keeping all other parameter settings the same as proposed by the reported works. The performance of all the experiments conducted have been evaluated using $10 \times$ tenfold cross validation. The values after the \pm sign in Tables 1, 2, 3 represent the standard deviation.

The error rates of the proposed scheme compared to other competing methods for dataset 1, dataset 2 and dataset 3 are given in Tables 1, 2 and 3, respectively. The results from Tables 1, 2, and 3 shows that the proposed scheme yields the best results obtaining

Table 1 Error rates (%) of proposed SPECTRA predictor and competing methods for dataset 1

Subject	CSP	CSSP	FBCSP	DFBCSP	SFBCSP	SBLFB	CSP-TSM	SPECTRA (Proposed)
<i>aa</i>	21.00 ± 5.31	17.00 ± 7.34	17.14 ± 8.19	9.65 ± 5.01	18.43 ± 7.45	16.79 ± 8.93	16.79 ± 6.29	10.36 ± 6.10
<i>al</i>	3.86 ± 3.63	3.07 ± 3.03	1.29 ± 1.18	1.00 ± 1.91	1.64 ± 1.36	1.36 ± 1.23	2.14 ± 3.53	1.07 ± 2.51
<i>av</i>	28.29 ± 7.46	28.86 ± 7.10	30.36 ± 8.23	31.21 ± 8.92	29.93 ± 6.44	28.07 ± 8.45	24.90 ± 9.10	21.67 ± 5.93
<i>aw</i>	10.36 ± 5.10	8.43 ± 5.09	6.50 ± 4.55	4.64 ± 4.75	9.29 ± 5.85	5.57 ± 4.90	4.54 ± 2.80	3.93 ± 3.29
<i>ay</i>	3.86 ± 4.11	4.29 ± 3.75	5.07 ± 4.68	8.21 ± 5.06	12.79 ± 5.96	11.00 ± 6.03	3.21 ± 2.53	5.71 ± 3.94
Average	13.47 ± 5.18	12.33 ± 5.30	12.07 ± 5.51	10.94 ± 5.13	14.14 ± 5.57	12.56 ± 6.07	10.31 ± 4.85	8.55 ± 4.35

The lowest error rates for each of the subjects are indicated in bold

Table 2 Error rate (%) of proposed SPECTRA predictor and competing methods for dataset 2

Subject	CSP	CSSP	FBCSP	DFBCSP	SFBCSP	SBLFB	CSP-TSM	SPECTRA (Proposed)
<i>a</i>	13.20 ± 8.07	13.65 ± 8.19	19.10 ± 9.35	16.80 ± 7.81	17.40 ± 5.93	19.10 ± 9.73	11.98 ± 6.64	13.00 ± 8.05
<i>b</i>	42.80 ± 12.25	42.70 ± 11.38	44.70 ± 11.27	42.90 ± 9.75	45.30 ± 6.59	41.50 ± 11.12	40.95 ± 10.32	42.50 ± 8.28
<i>c</i>	43.70 ± 11.24	39.95 ± 10.21	35.70 ± 9.58	35.20 ± 8.51	43.00 ± 11.62	33.20 ± 12.53	32.16 ± 8.68	27.90 ± 11.53
<i>d</i>	22.40 ± 8.82	14.60 ± 8.75	22.20 ± 8.99	23.50 ± 8.41	29.50 ± 10.13	11.50 ± 7.91	15.74 ± 7.89	13.53 ± 6.12
<i>e</i>	18.00 ± 9.74	18.05 ± 9.18	14.00 ± 9.15	18.30 ± 8.84	24.70 ± 10.34	11.60 ± 6.88	9.85 ± 5.77	9.50 ± 8.13
<i>f</i>	22.50 ± 10.84	18.55 ± 8.39	19.60 ± 8.56	14.30 ± 8.57	20.90 ± 6.45	21.20 ± 11.98	14.09 ± 7.13	13.20 ± 6.94
<i>g</i>	7.10 ± 5.06	6.35 ± 4.92	6.90 ± 6.62	9.00 ± 5.05	9.70 ± 4.97	5.90 ± 5.41	7.86 ± 5.67	5.67 ± 4.69
Average	24.24 ± 9.43	21.98 ± 8.72	23.17 ± 9.07	22.86 ± 8.13	27.21 ± 8.00	20.57 ± 9.36	18.94 ± 7.44	17.90 ± 7.68

The lowest error rates for each of the subjects are indicated in bold

Table 3 Error rate (%) of proposed SPECTRA predictor and competing methods for dataset 3

Subject	CSP	CSSP	FBCSP	DFBCSP	SFBCSP	SBLFB	CSP-TSM	SPECTRA (Proposed)
<i>B0103T</i>	23.19 ± 10.14	25.31 ± 9.99	19.00 ± 8.47	23.25 ± 11.23	26.50 ± 9.24	25.25 ± 10.33	22.50 ± 9.44	22.23 ± 12.03
<i>B0203T</i>	41.94 ± 11.04	42.94 ± 11.74	45.63 ± 11.93	40.76 ± 12.45	42.75 ± 12.84	40.75 ± 11.99	39.50 ± 10.43	41.02 ± 11.69
<i>B0303T</i>	46.69 ± 9.38	48.44 ± 10.82	49.13 ± 13.54	50.50 ± 12.87	44.97 ± 11.65	50.68 ± 13.34	49.06 ± 11.09	47.36 ± 11.97
<i>B0403T</i>	0.75 ± 2.04	0.63 ± 0.60	1.75 ± 1.61	0.75 ± 0.69	0.38 ± 0.35	0.88 ± 0.73	0.75 ± 2.23	1.46 ± 2.69
<i>B0503T</i>	17.85 ± 8.70	42.25 ± 16.33	28.50 ± 8.85	25.00 ± 10.71	25.02 ± 7.38	20.21 ± 10.26	17.18 ± 10.25	13.78 ± 10.51
<i>B0603T</i>	35.19 ± 11.05	23.81 ± 10.94	24.38 ± 9.80	20.88 ± 10.38	20.06 ± 10.70	25.12 ± 12.32	23.01 ± 9.35	19.31 ± 9.44
<i>B0703T</i>	14.50 ± 8.56	13.81 ± 8.11	15.50 ± 6.83	12.13 ± 9.05	12.25 ± 7.47	11.88 ± 9.39	13.81 ± 7.89	11.67 ± 6.71
<i>B0803T</i>	13.06 ± 8.43	14.50 ± 8.56	18.88 ± 11.68	11.13 ± 6.95	12.38 ± 7.63	11.13 ± 8.95	11.44 ± 7.95	9.79 ± 6.29
<i>B0903T</i>	19.13 ± 9.96	17.25 ± 8.66	20.88 ± 10.07	22.25 ± 10.80	25.00 ± 9.62	19.38 ± 10.58	19.75 ± 9.47	15.63 ± 8.65
Average	23.59 ± 8.81	25.44 ± 9.67	24.85 ± 9.39	22.96 ± 9.61	23.26 ± 8.67	22.81 ± 9.97	21.89 ± 8.67	20.26 ± 8.89

The lowest error rates for each of the subjects are indicated in bold

the lowest average error rates on all three datasets. The proposed scheme shows an improvement in the average error rates (improvement of 1.76% for datasets 1, 1.04% for dataset 2 and 1.63% for dataset 3) compared to the previously best performing CSP-TSM algorithm. An improvement in error rates of 4.92% for datasets 1, 6.34% for dataset 2 and 3.33% for dataset 3 are shown compared to the conventional CSP approach. Considering the performance of the individual subjects, 2 out of 5 subjects for dataset 1, 4 out of 7 subjects for dataset 2 and 5 out of 9 subjects for dataset 3 achieved the lowest error rates

using the proposed SPECTRA predictor. Overall, 15 out of 21 subjects showed improvement in performance compared to the CSP-TSM approach with subject “*aa*” of dataset 1 showing highest decrease in error rate (6.43%). Out of these 15 subjects, 13 subjects showed greater than 1% reduction in the error rate. This indicates the advantage of our proposed SPECTRA predictor in comparison to the CSP-TSM approach. It should also be noted that for 4 subjects out of all the subjects used in the evaluation, the error rates increased using the proposed SPECTRA predictor when compared to the CSP-TSM approach with highest increase being 2.50% for the subject “*ay*” of dataset 1. This may be improved or overcome by incorporating automatic selection of the parameter n that is subject dependent and will be explored in future works. Our proposed predictor also performed well compared to the TFPO-CSP [51] approach that was evaluated using dataset 1 (achieved error rate of 10.19%) and dataset 2 (achieved error rate of 20.63%).

To add on, the authors of the SBLFB approach have used linear discriminant analysis (LDA) as the classifier. We have used SVM classifier to make a fair comparison between the different methods. It should be noted that the SBLFB approach achieved a slightly better error rate of 11.89% on dataset 1 when the LDA classifier was employed. Furthermore, the authors in [65] proposed an iterative spatio-spectral patterns learning (ISSPL) approach. They evaluated their method using dataset 1 obtaining an average error rate of 5.79%. However, they used a window size of 3.5 s for extracting the trials and thus it cannot be compared with our method as more data is being used by ISSPL approach. Similarly, a cross-correlation based logistic regression (CC-LR) [66] method achieved an average error rate of 6.09% on dataset 1. However, they only used the training data from the competition and evaluated their method using threefold cross-validation. Thus, we cannot compare this method with SPECTRA. In [67], the authors proposed using multiscale principal component analysis for de-noising the EEG signal and extracted higher-order statistics features from wavelet packet decomposition sub-bands. The method was also evaluated using dataset 1, achieving an average error rate of 7.2%. However, it also used 3.5 s window for extracting the trials and hence cannot be directly compared with our approach. In future, we will explore the effect of using multiscale principal component analysis for de-noising the EEG signal with our proposed approach. Also, we will explore the effect of other feature extraction approaches [68–70] and deep learning methods [71] with our current work.

Furthermore, to validate the reliability of the results that has been achieved, Cohen’s kappa coefficient κ is used. Tables 4, 5 and 6 shows the κ values obtained

Table 4 Cohen’s kappa coefficient values of proposed and competing methods for dataset 1

Subject	CSP	CSSP	FBCSP	DFBCSP	SFBCSP	SBLFB	CSP-TSM	SPECTRA
<i>aa</i>	0.613	0.659	0.601	0.816	0.394	0.664	0.636	0.793
<i>al</i>	0.927	0.940	0.970	0.976	0.917	0.973	0.950	0.979
<i>av</i>	0.426	0.423	0.384	0.329	0.389	0.439	0.543	0.567
<i>aw</i>	0.800	0.837	0.837	0.906	0.743	0.889	0.896	0.921
<i>ay</i>	0.903	0.926	0.881	0.847	0.763	0.780	0.950	0.886
Average	0.734	0.757	0.735	0.775	0.641	0.749	0.795	0.829

The best values for each of the subjects are highlighted in bold

Table 5 Cohen's kappa coefficient values of proposed and competing methods for dataset 2

Subject	CSP	CSSP	FBCSP	DFBCSP	SFBCSP	SBLFB	CSP-TSM	SPECTRA
<i>a</i>	0.736	0.727	0.618	0.664	0.652	0.618	0.757	0.730
<i>b</i>	0.144	0.146	0.106	0.142	0.094	0.170	0.206	0.150
<i>c</i>	0.126	0.201	0.286	0.290	0.140	0.336	0.359	0.418
<i>d</i>	0.552	0.708	0.556	0.530	0.410	0.770	0.696	0.733
<i>e</i>	0.640	0.639	0.720	0.634	0.506	0.768	0.826	0.807
<i>f</i>	0.550	0.629	0.608	0.714	0.582	0.576	0.717	0.777
<i>g</i>	0.858	0.873	0.862	0.820	0.806	0.882	0.851	0.887
Average	0.515	0.560	0.537	0.542	0.456	0.589	0.630	0.643

The best values for each of the subjects are highlighted in bold

Table 6 Cohen's kappa coefficient values of proposed and competing methods for dataset 3

Subject	CSP	CSSP	FBCSP	DFBCSP	SFBCSP	SBLFB	CSP-TSM	SPECTRA
<i>B0103T</i>	0.536	0.494	0.620	0.535	0.470	0.495	0.550	0.554
<i>B0203T</i>	0.161	0.141	0.088	0.185	0.145	0.185	0.210	0.179
<i>B0303T</i>	0.106	0.031	0.018	0.010	0.100	0.014	0.010	0.054
<i>B0403T</i>	0.983	0.988	0.965	0.985	0.993	0.983	0.985	0.971
<i>B0503T</i>	0.650	0.115	0.430	0.500	0.499	0.595	0.655	0.725
<i>B0603T</i>	0.296	0.524	0.513	0.583	0.598	0.457	0.530	0.614
<i>B0703T</i>	0.710	0.724	0.690	0.758	0.755	0.763	0.724	0.767
<i>B0803T</i>	0.739	0.710	0.623	0.778	0.753	0.776	0.771	0.804
<i>B0903T</i>	0.618	0.655	0.583	0.555	0.500	0.613	0.605	0.688
Average	0.534	0.487	0.503	0.543	0.535	0.542	0.560	0.595

The best values for each of the subjects are highlighted in bold

using each of the methods for dataset 1, dataset 2 and dataset 3, respectively. It can be seen from Table 8 (in methods section) that a higher value of κ means a greater strength of agreement. A higher strength of agreement means that the results are more reliable. Our proposed scheme attained the best average κ values for all the 3 datasets. This shows that the results of the proposed scheme are more reliable when compared with other competing methods. Considering the average κ values, a very good strength of the agreement of classes for dataset 1 and good prediction of classes for dataset 2 and dataset 3 have been achieved. It can be noted that the κ values for some subjects (such as subject "*av*" of dataset 1, subjects "*b*" and "*c*" of dataset 2 and subjects "*B0203T*" and "*B0303T*" of dataset 3) are very low. These results are consistent with the results of other methods and may be mainly due to low quality signals being recorded, which are contaminated by noise. Considering individual subjects 4 out of 5 for dataset 1, 5 out of 7 for dataset 2 and 6 out of 9 subjects for dataset 3 achieved good or very good strength of agreement using the proposed scheme. While on the other hand, 3 out of 5 subjects for dataset 1, 6 out of 7 subjects for dataset 2 and 5 out of 9 subjects for dataset 3 attained the best κ values using the proposed scheme.

In this work, we have used a single wide band to keep the computational complexity of the proposed method low as using multiple sub-bands would result in an

Table 7 Test time required by different algorithms for single-trial MI EEG signal classification

Method	CSP	CSSP	FBCSP	DFBCSP	SFBCSP	SBLFB	CSP-TSM	SPECTRA
Time (ms)	2.30	4.30	14.22	10.80	19.86	13.10	2.60	6.10

increased computational complexity. However, using multiple sub-bands may further improve the performance of the system and will be studied in future. Table 7 shows the time taken to process and classify a MI EEG signal for different methods (Matlab running on a personal computer at 2.4 GHz (Intel(R) Core(TM) i5) has been used for all processing). Our proposed SPECTRA predictor takes 6.10 ms to process and classify a trial of EEG signal. Thus, the proposed scheme is suitable for real-time applications and is computationally efficient for portable devices. Our proposed approach also takes less time to process and classify a trial compared to other competing methods such as DFBCSP, SFBCSP and SBLFB. SPECTRA takes more time to process and classify a trial compared to CSP, CSSP and CSP-TSM as it employs these approaches.

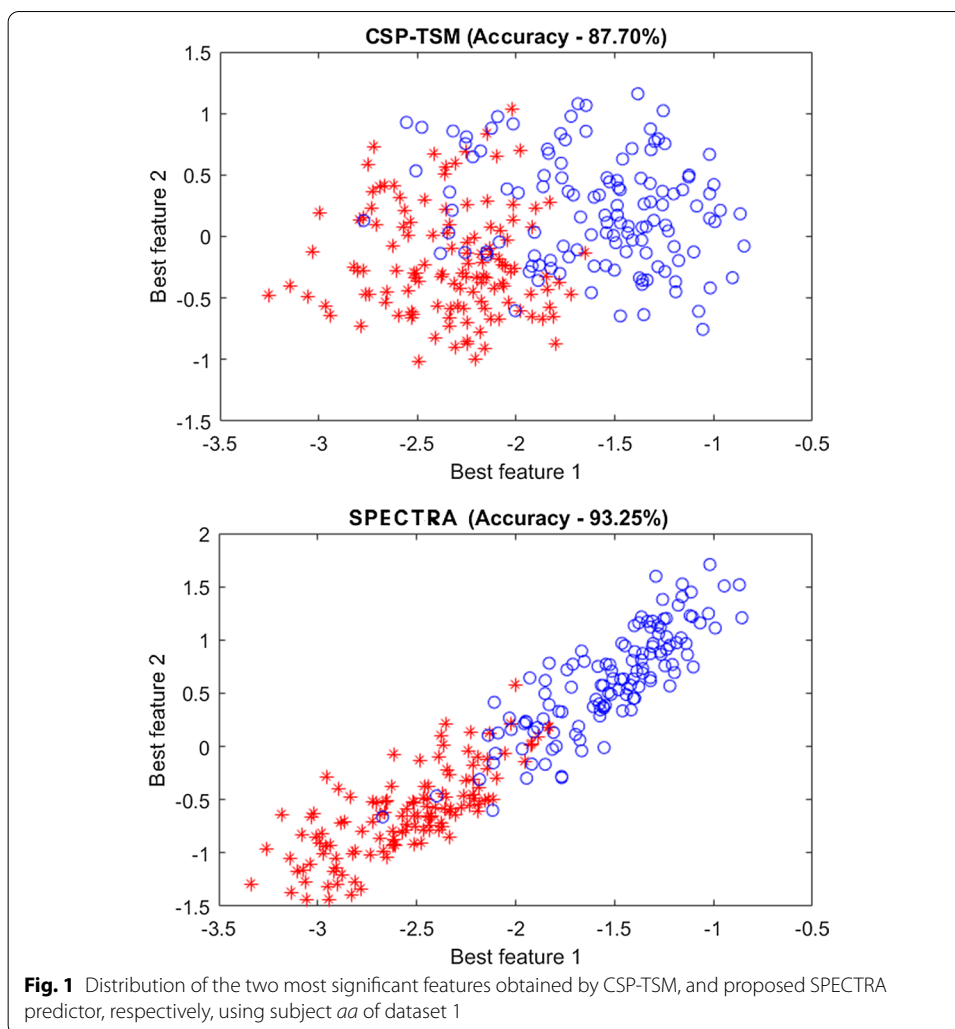
Discussion

In this study, we have performed feature selection using F-score in order to remove redundant features so that only significant features are used. Top $r=10$ features has been selected [64]. Figure 1 shows the feature distribution of the top two features for CSP-TSM and the proposed scheme. It can be seen that using the proposed scheme effectively finds more separable features that accounts for the improved performance and usefulness of the proposed system.

Furthermore, as mentioned before, we have only used dataset 2 for selecting the parameter n . This has been done so that we do not have to tune the parameters for each new dataset that is used and so that the parameters selected can perform well on all datasets. This will reduce the training time by not having to perform parameter selection on other datasets. It is seen that the parameters selected for the proposed method in this work performed well as promising results have been obtained for all the three datasets.

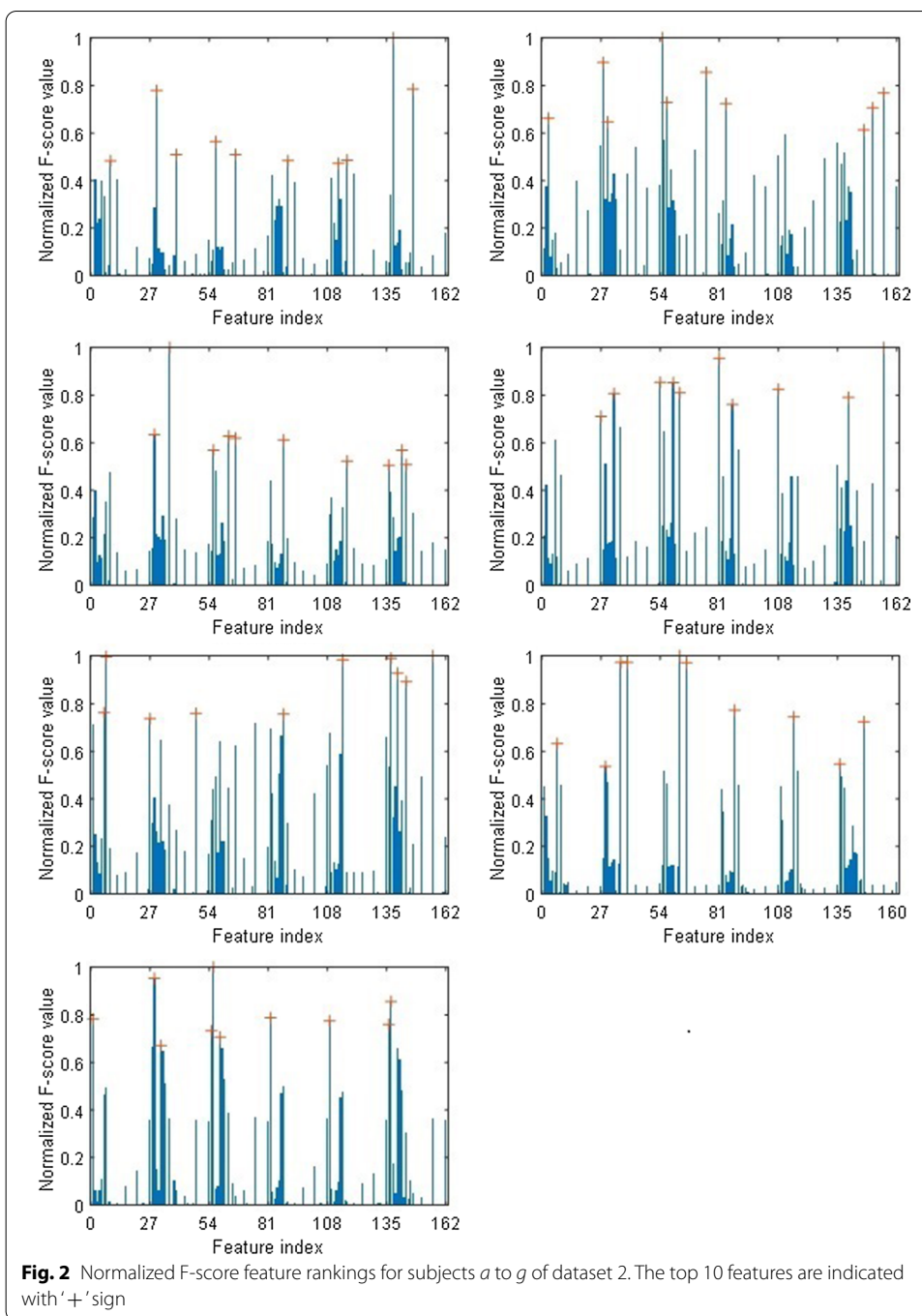
To show the significance of the proposed method, we have performed paired t-test with 1% significance level. The average individual error rate results of the proposed scheme have been compared with that of the 2nd best method (CSP-TSM). The p-value obtained was 0.0036, which shows that significant improvements are achieved by the proposed scheme.

Moreover, there are various ways of combining the temporal windows for CSSP-TSM approach as this can be done by simply using only two temporal windows. Figure 2 shows the normalized F-score ranking of the features for the subjects of dataset 2. The number of features obtained by each CSP-TSM or CSSP-TSM process is 27 (6 CSP features and 21 TSM features), therefore a total of 162 (6×27) features are obtained. CSP-TSM and CSSP-TSM processes refer to the blocks performing CSP-TSM and CSSP-TSM, as shown in Fig. 1, having 3 CSP-TSM and 3 CSSP-TSM processes. The output of each of these processes is a combination of CSP and TSM features. It can be seen from Fig. 2 that all the CSP-TSM and CSSP-TSM processes give more separable features. Hence the framework given in Fig. 1 has been adopted. In this work we are performing

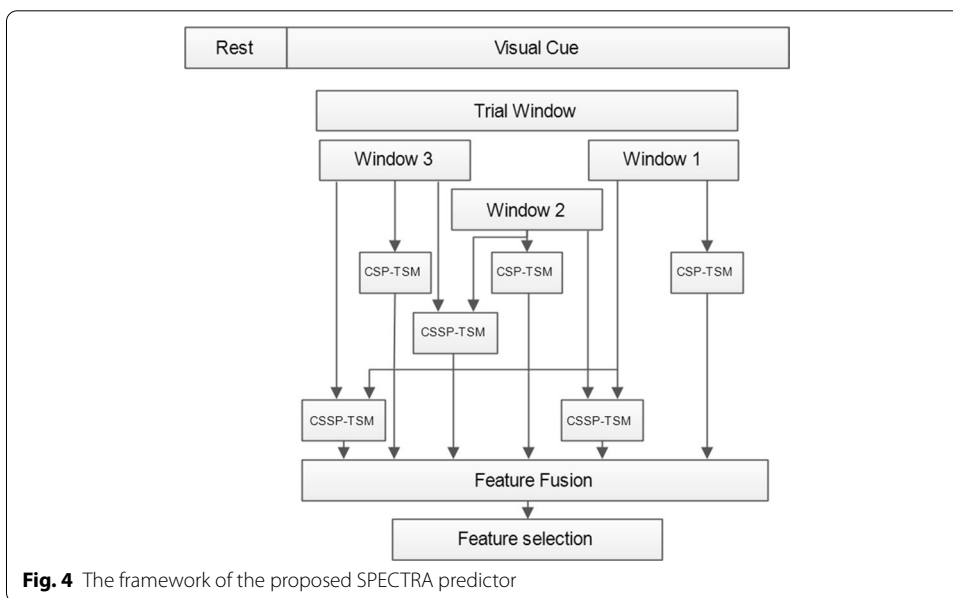
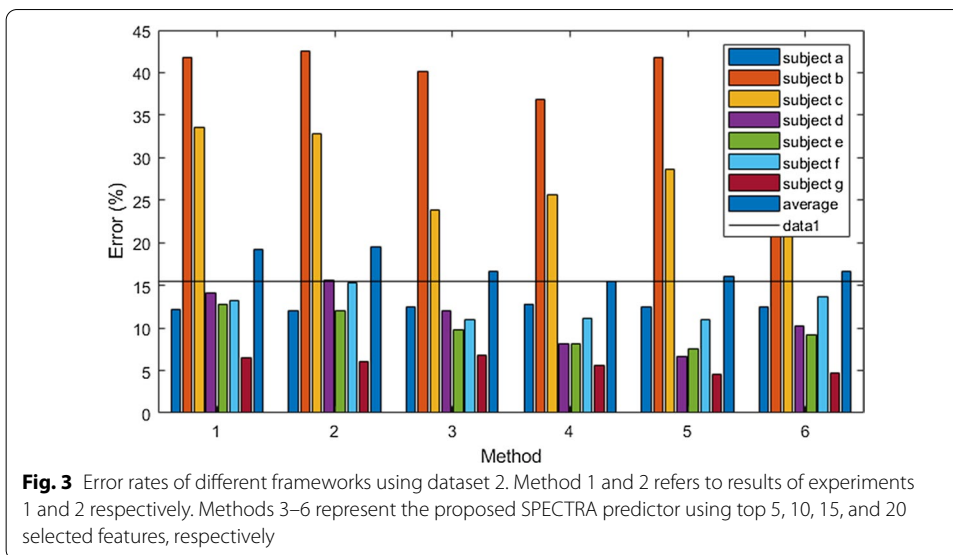


feature selection rather than selecting several CSP-TSM and CSSP-TSM processes only. This has been proposed after evaluating different frameworks. We have evaluated selecting only features of several top k CSP-TSM or CSSP-TSM processes from the 6 processes shown in Fig. 4 (refer to methods section). To select these top k CSP-TSM or CSSP-TSM processes, we again used the F-score. Two experiments were conducted for this. Experiment 1 used individual F-score feature ranking to select top k CSP-TSM or CSP-TSSM processes i.e. the top k CSP-TSM or CSSP-TSM processes with highest individual feature rankings were selected. In experiment 2, the average of the F-score feature rankings of all features of each CSP-TSM and CSSP-TSM processes were used to select the top k CSP-TSM or CSSP-TSM processes. We have used $k=4$ (similar to the band selection procedure in [61]) for experiments 1 and 2. It is evident from Fig. 3 that our proposed scheme with top 10 features selected gives the best result.

To add on, the results of BCI Competition were obtained using specific test data only (which was specifically for BCI competition only). Cross-validation using all the data is a more effective way to test a model's performance and has been mostly



utilized to compare the different methods proposed for BCI applications. This is the reason why the results of BCI Competition have not been compared with our work as done by other researchers. Moreover, as mentioned earlier, the value of parameter n selected did not produce the optimal results for all individual subjects and will be investigated in future works. We will also consider other feature extraction, feature selection and classifiers [72] for future works.



Convolutional neural network (CNN) has gained a lot of attention over the recent years. Therefore, in future, we will evaluate the use of CNN for MI-EEG signal recognition by developing hybrid models utilizing CNN with SPECTRA. Furthermore, good performance is noted by CNN on image data, therefore, DeepInsight [71] will be used to transform the EEG signal to image before being fed as input the CNN model. Long short-term memory network has also performed well for MI-EEG signal recognition [73] and we will also consider using LSTM network to further improve the performance of the proposed SPECTRA predictor.

Conclusions

In this work, we have utilised the CSP-TSM approach with multiple temporal delayed windows for extracting more separable features, using CSP and CSSP methods. Parameters such as the temporal delay and number of windows have been optimized. F-score for feature selection is proposed over Lasso that is used by the CSP-TSM approach due to its reliability and enhanced ability in selecting significant features. Our proposed scheme out-performed other competing approaches and achieved the lowest average error rates and highest average Cohen's kappa coefficient values. A fixed wide band has been used for all evaluation. Developing sophisticated algorithms which will automatically learn filter bands that will give optimal performance for each subject may further improve the performance of the proposed system. Our proposed scheme can be potentially used for the development of improved and computationally efficient BCI systems.

Methods

Public benchmark datasets

We have evaluated the performance of the proposed scheme using 3 datasets that are publicly available: BCI Competition III dataset IVa [74], BCI Competition IV dataset I [75] and, BCI Competition IV dataset IIb [75] referred to as dataset 1, dataset 2 and dataset 3 from here onwards, respectively.

All the three datasets contain two class MI tasks. The EEG signals of right hand and left foot MI tasks recorded from five subjects using 118 channels of EEG signals is contained in dataset 1. The signals sampled at 100 Hz are used with each subject having 140 trials for each task. Dataset 2 contains MI EEG signals of seven subjects recorded using 59 channels at 1000 Hz. The down sampled data at 100 Hz is used and it contains 200 trials for each subject containing almost equal number of trials for each MI tasks. Dataset 3 contains EEG signals of nine subjects. It contains 3 channels right hand and left hand MI tasks sampled at 250 Hz. As used in [62], we have only used data from session three for evaluation. Each subject contains 80 trials of each MI task. For a complete explanation of the datasets, refer to <http://www.bbci.de/competition/>.

CSP feature extraction

CSP has become one of the most popular and widely used techniques for feature extraction of MI EEG signals. Spatial filters W_{csp} are learned by the CSP algorithm, which maximizes the variance of one class while minimizing the variance of the other class. This offers an effective method to approximate the discerning information of the MI tasks. Given an EEG signal $X_i \in R^{C \times T}$ where i denotes the i -th trial, c denotes the number of channels data contained by the EEG signal and τ is the number of sample points. The learned spatial filters are used to transform the EEG signal to a new time series using (1).

$$Z_i = W_{CSP}^T X_i \quad (1)$$

The variance based CSP features are then extracted from the spatially transformed signal Z_i using (2), where f_i^k is the k -th feature of the i -th trial and $\text{var}(Z_i^j)$ denotes the variance of j -th row of Z_i . Refer to [76] for a detailed description of the CSP algorithm.

$$f_i^k = \log \left(\frac{\text{var}(Z_i^k)}{\sum_{j=1}^{2m} \text{var}(Z_i^j)} \right) \quad (2)$$

CSP-TSM feature extraction

The CSP-TSM approach has been proposed for extracting significant tangent space features while keeping the computational complexity low [52]. The concept of Riemannian geometry is utilized by the CSP-TSM approach. The normalized covariance matrix Σ_i of each of the spatially filtered trial Z_i is calculated. The Riemannian distance δ_R is then computed using (3), where Σ is the Riemannian mean of all the trial covariance matrices Σ_i (from the training set) and is calculated using (4), the logarithmic mapping $\text{Log}_\Sigma(\Sigma_i)$ is given by (5) and s_i represents the normalized tangent space vector (also referred to as tangent space features). The upper(\cdot) in (3) means vectorizing the upper triangular portion of the symmetric matrix and multiplying the out-of-diagonal elements [77] by $\sqrt{2}$.

$$\delta_R(\Sigma, \Sigma_i) = \|\text{Log}_\Sigma(\Sigma_i)\|_\Sigma = \|\text{upper}(\Sigma^{-1/2} \text{Log}_\Sigma(\Sigma_i) \Sigma^{-1/2})\|_2 = \|s_i\|_2 \quad (3)$$

$$\Sigma = R(\Sigma_i) = \underset{\Sigma \in \Sigma(n)}{\text{argmin}} \sum_{i=1}^N \delta_R^2(\Sigma, \Sigma_i) \quad (4)$$

$$\text{Log}_\Sigma(\Sigma_i) = \Sigma^{1/2} \log(\Sigma^{-1/2} \Sigma_i \Sigma^{-1/2}) \Sigma^{1/2} \quad (5)$$

The above process maps all the trial covariance matrices Σ_i into the tangent space. Thus, the features obtained from tangent space mapping are fused together with the CSP features and significant features are selected. The selected features are then used for classification. A complete description of the CSP-TSM approach can be obtained from our preceding work [64].

Proposed approach

In this study, we propose an effective subject-dependent method of feature extraction by utilizing the CSP-TSM approach. The general conceptual framework of the proposed methodology for obtaining significant features is shown in Fig. 4. Usually, only a single window of 2.0–3.0 s is used for MI-based BCI applications. Here, we propose to use n multiple temporal delayed windows in two different ways. Firstly, the variance based CSP features and TSM features are computed for each of the $n=3$ windows (the choice of n used is explained in the following sub-section). Secondly, the CSSP approach is utilized for extracting further information. CSSP method involves inserting a temporal delayed window to the trial signal and performing CSP on this modified trial signal that is obtained. The CSSP approach was proposed for improving the performance of CSP. The time delay value τ influences the performance of the system and needs to be chosen carefully. In this work, the time delay (τ sample points) has been selected using the cross-validation technique. All combinations of the n windows are used for obtaining new CSSP trial windows given by (6), where W_i is the i -th window of the original signal (refer to Fig. 4), $W_{CSSP}^{i,i+j}$ is the signal obtained by inserting the W_{i+j} temporal delayed

window to window W_i and $i = 1 : n - 1$. CSP variance-based features and TSM features are attained from the windows obtained from (6).

$$W_{CSSP}^{i,i+j} = \begin{bmatrix} W_i \\ W_{i+j} \end{bmatrix}; j = 1 : n - i \quad (6)$$

All the features obtained are fused together to form the feature vector. Using (7) the F-score ranking of the features is then computed, where \bar{F}_i is average value of the i -th feature, \bar{F}_i^+ and \bar{F}_i^- are the average values of the i -th feature for the positive and negative samples respectively, N^+ and N^- refers to the total number of positive and negative samples respectively and $F_{k,i}$ refers to the k -th sample of the i -th feature. The positive samples for all the three datasets were right-hand MI task samples while the negative samples were left-foot MI task samples for dataset 1 and 2 and left hand MI task samples for dataset 3. The F-score values obtained are arranged in descending order and the top r features are selected, which are classified using support vector machine (SVM) classifier.

$$F_{score}(i) = \frac{(\bar{F}_i^+ - \bar{F}_i)^2 + (\bar{F}_i^- - \bar{F}_i)^2}{\frac{1}{N^+ - 1} \sum_{k=1}^{N^+} (F_{k,i}^+ - \bar{F}_i^+)^2 + \frac{1}{N^- - 1} \sum_{k=1}^{N^-} (F_{k,i}^- - \bar{F}_i^-)^2} \quad (7)$$

SVM is a supervised learning technique and has been effectively used for both regression and classification problems. A hyperplane that maximizes the separation of the support vectors is determined by the SVM algorithm. In this study we employed an SVM classifier having radial basis kernel function. The use of kernel function allows non-linear data to be mapped to a linearly separable higher dimensional plane.

Parameter selection

Multiple temporal delayed windows have been utilized in this study. Two factors are of importance in this process: window size and temporal time delay τ between windows. Different subjects have different response rate to the onset cue. Therefore, determining the exact location of the MI task in the EEG signals needs to be investigated and clustering methods [78–80] can be utilised for this purpose. We have fixed the window size to 2.0 s in our work, the same as used by most of the researchers [34, 48, 58, 62]. To determine the best τ value that would yield the optimal performance, we have conducted the following experiments. Firstly, the τ value was varied from 10 to 100% of the sampling frequency for each of the datasets and the results are shown in Fig. 5. In selecting the τ value, it is very important to consider real time BCI applications. Considering that real time BCI applications will also be portable, the computational complexity should be kept to a minimum. Therefore, it is desirable to select the smallest τ value that will produce near to optimal results. From Fig. 5, it can be seen that using 10% of sampling frequency as the τ value gives near optimal performance for all three datasets. To further refine the τ parameter (since now it is clear that using larger τ values would not improve the performance), τ values from 1 to 10% of sampling frequency were evaluated (results shown in Fig. 6). It can be noted from Fig. 6 that for dataset 1 and dataset 2 only 10 (10% of 100) sample points are shown whereas for dataset 3, 25 (10% of 250) sample

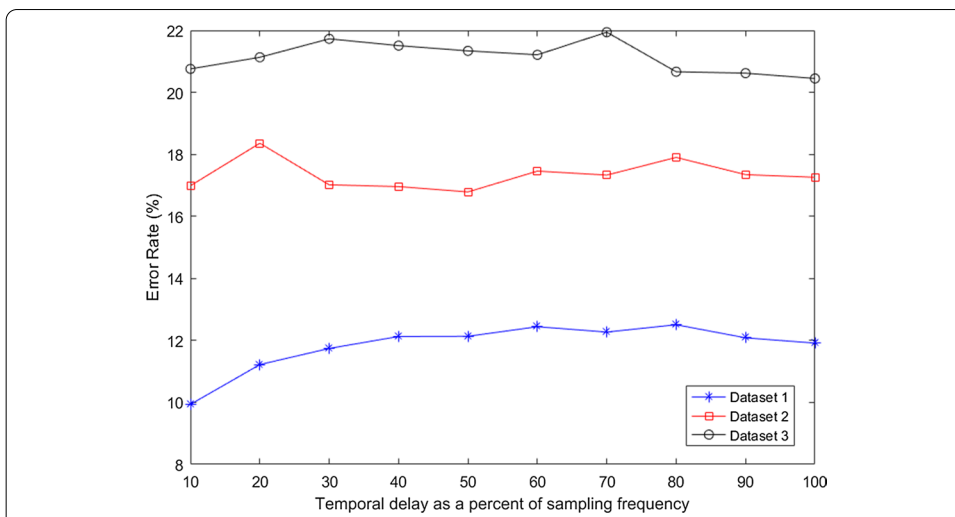


Fig. 5 Average error rates for different values of temporal delay as a percentage of sampling frequency

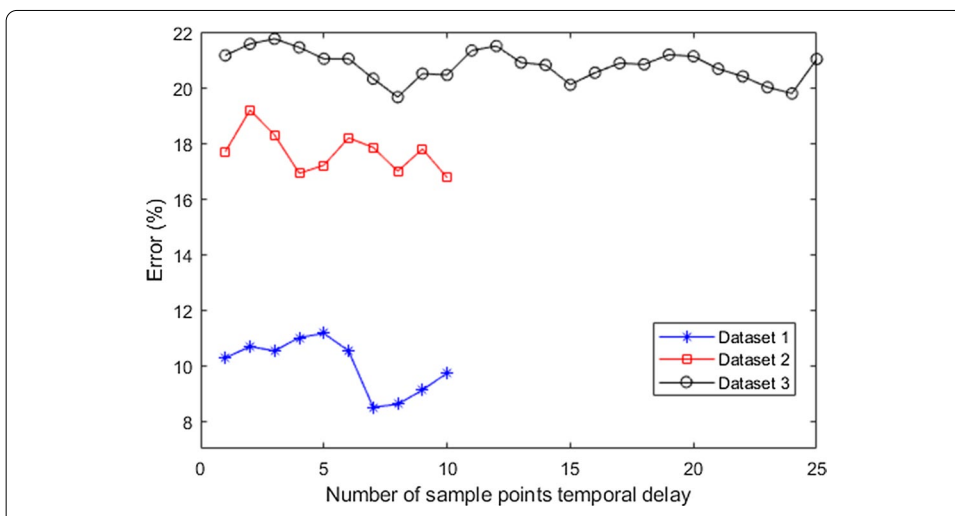
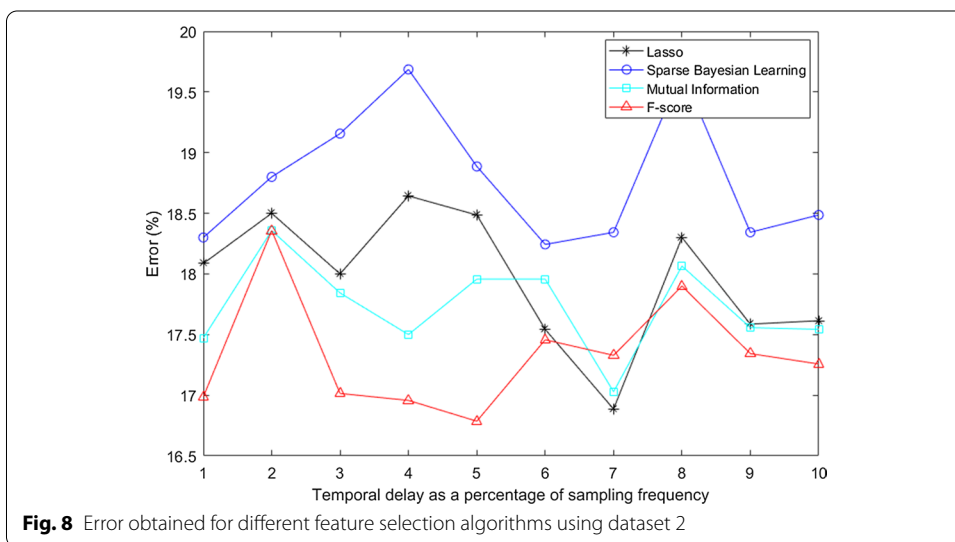
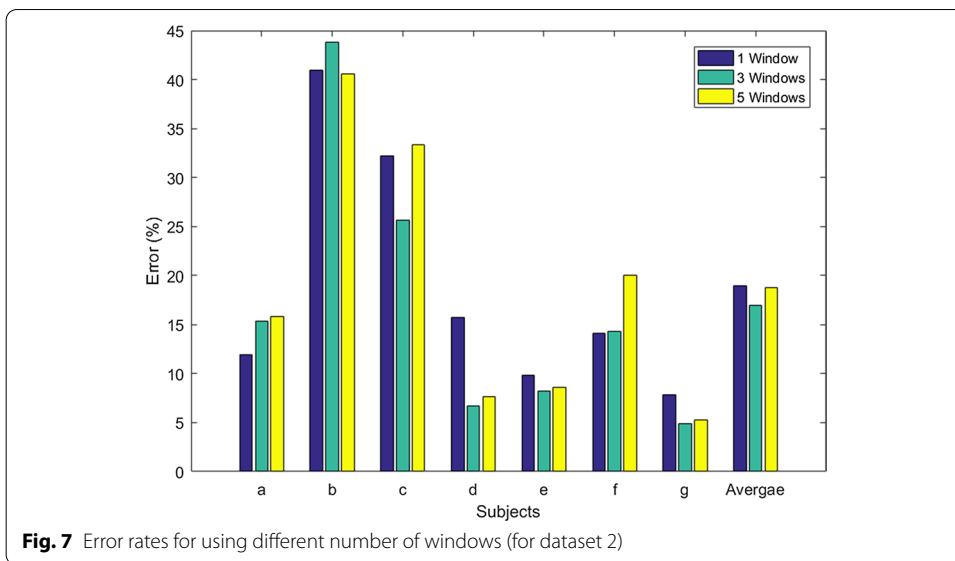


Fig. 6 Average error rates for different values of temporal delay

points are shown due to the signals being sampled at different frequencies. It can also be noted from Fig. 6 that optimal performance is obtained for different subjects at different τ values. Thus, another tenfold cross-validation has been performed on the training data (which is obtained from the initial tenfold cross-validation) for the selection of subject-dependent τ values that will give optimal performance. In this way, the test samples are not used during parameter tuning.

The other parameter that needed to be selected was n , the number of windows. We have evaluated $n = [1, 3, 5]$ and the results are shown in Fig. 7. We have randomly selected dataset 2 for selecting the parameter n . Using only 1 window will result in the CSP-TSM approach. It is evident from Fig. 7 that using a high number of windows did not enhance the system performance and would increase the computational complexity of the system. All subjects except subjects a and b performed well using 3 windows



compared to using 1 or 5 windows. Therefore, to retain a low computation complexity of the proposed scheme while also producing optimal performance, we have chosen $n = 3$.

We have also evaluated four different feature selection algorithms (Lasso [52, 81], sparse Bayesian learning [42], mutual information [9] and F-score based feature selection algorithms) in order to choose the best performing algorithm. Figure 8 shows the error rates obtained for different feature selection algorithms using dataset 2. It can be noted that using F-score yields the minimum error rates for almost all temporal delay values showing that it is a robust and reliable feature selection method. This is the reason why we have used F-score in this work for feature selection instead of the Lasso method as used in CSP-TSM approach.

Table 8 Strength of agreement for different Cohen's kappa coefficient values

κ	< 0.20	0.21–0.40	0.41–0.60	0.61–0.80	0.81–1.0
Strength	Poor	Fair	Moderate	Good	Very Good

Performance measures

To appropriately rank and compare our proposed scheme with competing methods, two performance measures, error rate and Cohen's kappa coefficient (κ) have been used. Error rate is a commonly used measure for evaluating the performance of BCI systems, which shows the percentage of trials that are classified incorrectly. κ is utilised for validating the reliability of the results which statistically accesses the consistency of agreement among two classes. κ is calculated using (8), where p_e is the chance of agreement (in percentage) that is expected and p_a is the actual agreement (in percentage). Table 8 shows the strength of agreement for different κ values [82].

$$\kappa = \frac{p_a - p_e}{1 - p_e} \quad (8)$$

Abbreviations

BCI: Brain computer interface; CSP: Common spatial pattern; CSSP: Common spatio-spectral pattern; DFBCSP: Discriminant filter bank common spatial pattern; EEG: Electroencephalography; FBCSP: Filter bank CSP; LDA: Linear discriminant analysis; MI: Motor imagery; SBCSP: Sub-band common spatial pattern; SBLFB: Sparse Bayesian learning of filter banks; SFBCSP: Sparse filter bank CSP; SFTOFSRC: Spatial-frequency-temporal optimized feature sparse representation-based classification; SPECTRA: Spatial-frequency-temporal feature extraction; SVM: Support vector machine; TFPO: Temporal filter parameter optimization.

Acknowledgements

We would like to thank the Berlin BCI group for publicly providing the motor imagery EEG datasets (BCI Competition III dataset IVa, BCI Competition IV dataset I and BCI Competition IV dataset IIb).

About this supplement

This article has been published as part of BMC Bioinformatics Volume 22 Supplement 6, 2021: 19th International Conference on Bioinformatics 2020 (InCoB2020). The full contents of the supplement are available online at <https://bmcbioinformatics.biomedcentral.com/articles/supplements/volume-22-supplement-6>.

Authors' contributions

SK and AS conceived the project. SK performed the analysis and wrote the manuscript under the guidance AS. TT provided computational resources. All authors read and approved the final manuscript.

Funding

This research work and the publication charge for this article is funded by JST CREST (Grant Number: JPMJCR1412), Japan, RIKEN, Center for Integrative Medical Sciences, Japan and the College Research Committee (CRC) of Fiji National University, Fiji. The funding bodies did not have any role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Availability of data and materials

The datasets used in this study are publicly available at <http://www.bbc.de/competition>.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹ School of Electrical and Electronics Engineering, Fiji National University, Suva, Fiji. ² Laboratory for Medical Science Mathematics, RIKEN Center for Integrative Medical Sciences, Yokohama 230-0045, Japan. ³ Department of Medical Science Mathematics, Medical Research Institute, Tokyo Medical and Dental University (TMDU), Tokyo 113-8510, Japan. ⁴ Laboratory for Medical Science Mathematics, Department of Biological Sciences, Graduate School of Science, University of Tokyo, Tokyo 113-0033, Japan. ⁵ School of Engineering and Physics, The University of the South Pacific, Suva, Fiji. ⁶ Institute for Integrated and Intelligent Systems, Griffith University, Nathan, Brisbane, QLD, Australia.

Received: 15 March 2021 Accepted: 21 March 2021

Published: 2 June 2021

References

- Bhattacharyya S, Konar A, Tibarewala DN. Motor imagery, P300 and error-related EEG-based robot arm movement control for rehabilitation purpose. *Med Biol Eng Comput.* 2014;52(12):1007–17.
- Ramos-Murguialday A, Broetz D, Rea M, Läer L, Yilmaz Ö, Brasil FL, Liberati G, Curado MR, Garcia-Cossio E, Vyziotis A, et al. Brain-machine interface in chronic stroke rehabilitation: a controlled study. *Ann Neurol.* 2013;74(1):100–8.
- Luo T-J, Zhou C-L, Chao F. Exploring spatial-frequency-sequential relationships for motor imagery classification with recurrent neural network. *BMC Bioinform.* 2018;19(1):344.
- Frölich L, Andersen TS, Mørup M. Rigorous optimisation of multilinear discriminant analysis with Tucker and PARAFAC structures. *BMC Bioinformatics.* 2018;19(1):197.
- Richhariya B, Tanveer M. EEG signal classification using universum support vector machine. *Expert Syst Appl.* 2018;106:169–82.
- Rahman MA, Khanam F, Ahmad M, Uddin MS. Multiclass EEG signal classification utilizing Rényi min-entropy-based feature selection from wavelet packet transformation. *Brain Inform.* 2020;7:7.
- Bajaj V, Taran S, Khare SK, Sengur A. Feature extraction method for classification of alertness and drowsiness states EEG signals. *Appl Acoust.* 2020;163:107224.
- Sharma R, Chopra K. EEG signal analysis and detection of stress using classification techniques. *J Inf Optim Sci.* 2020;41(1):229–38.
- Kumar S, Sharma A, Tsunoda T. An improved discriminative filter bank selection approach for motor imagery EEG signal classification using mutual information. *BMC Bioinform.* 2017;18(16):545.
- Gao Y, Gao B, Chen Q, Liu J, Zhang Y. Deep convolutional neural network-based epileptic electroencephalogram (EEG) signal classification. *Front Neurol.* 2020;11:375.
- Zhou D, Li X. Epilepsy EEG signal classification algorithm based on improved RBF. *Front Neurosci.* 2020;14:606.
- Yuan Y, Jia K, Ma F, Xun G, Wang Y, Su L, Zhang A. A hybrid self-attention deep learning framework for multivariate sleep stage classification. *BMC Bioinform.* 2019;20(16):586.
- Yin Z, Liu L, Chen J, Zhao B, Wang Y. Locally robust EEG feature selection for individual-independent emotion recognition. *Expert Syst Appl.* 2020;162:113768.
- Liu J, Wu G, Luo Y, Qiu S, Yang S, Li W, Bi Y. EEG-based emotion classification using a deep neural network and sparse autoencoder. *Front Syst Neurosci.* 2020;14:43.
- Naseer N, Ayaz H, Dehais F. Portable and wearable brain technologies for neuroenhancement and neurorehabilitation. *Biomed Res Int.* 2018;2018:2.
- Chowdhury A, Raza H, Meena YK, Dutta A, Prasad G. An EEG-EMG correlation-based brain-computer interface for hand orthosis supported neuro-rehabilitation. *J Neurosci Methods.* 2019;312:1–11.
- Asensio-Cubero J, Gan JQ, Palaniappan R. Multiresolution analysis over graphs for a motor imagery based online BCI game. *Comput Biol Med.* 2016;68(Supplement C):21–6.
- Bordoloi S, Sharmah U, Hazarika SM. Motor imagery based BCI for a maze game. In: *4th International Conference on Intelligent Human Computer Interaction (IHCI); Kharagpur.* 2012: 1–6.
- Akram F, Han H-S, Kim T-S. A P300-based word typing brain computer interface system using a smart dictionary and random forest classifier. In: *The Eighth International Multi-Conference on Computing in the Global Information Technology: 2013.* 106–109.
- Akram F, Metwally MK, Hee-Sok H, Hyun-Jae J, Tae-Seong K. A novel P300-based BCI system for words typing. In: *International Winter Workshop on Brain-Computer Interface (BCI): 18–20 February 2013.* 24–25.
- Kleih SC, Kuafmann T, Zickler C, Halder S, Leotta F, Cincotti F, Aloise F, Riccio A, Herbert C, Mattia D, et al. Out of the frying pan into the fire—the P300-based BCI faces real-world challenges. *Prog Brain Res.* 2011;194:27–46.
- Alonso-Valerdi LM, Salido-Ruiz RA, Ramirez-Mendoza RA. Motor imagery based brain-computer interfaces: an emerging technology to rehabilitate motor deficits. *Neuropsychologia.* 2015;79(Part B):354–63.
- Kumar S, Sharma A, Mamun K, Tsunoda T. A deep learning approach for motor imagery EEG signal classification. In: *3rd Asia-Pacific World Congress on Computer Science and Engineering: 4th–6th December; Denarau Island, Fiji.* 2016.
- Liu Y, Li M, Zhang H, Wang H, Li J, Jia J, Wu Y, Zhang L. A tensor-based scheme for stroke patients' motor imagery EEG analysis in BCI-FES rehabilitation training. *J Neurosci Methods.* 2014;222:238–49.
- Pfurtscheller G, Neuper C. Motor imagery and direct brain-computer communication. *Proc IEEE.* 2001;89(7):1123–34.
- McFarland DJ, McCane LM, David SV, Wolpaw JR. Spatial filter selection for EEG-based communication. *Electroencephalogr Clin Neurophysiol.* 1997;103(3):386–94.
- Kawala-Sterniuk A, Podpora M, Pelc M, Blaszczyzyn M, Gorzelanczyk EJ, Martinek R, Ozana S. Comparison of smoothing filters in analysis of eeg data for the medical diagnostics purposes. *Sensors (Basel, Switzerland).* 2020;20(3):807.

28. McFarland DJ. The advantages of the surface Laplacian in brain–computer interface research. *Int J Psychophysiol.* 2015;97(3):271–6.
29. Bradshaw LA, Wikswo JP. Spatial filter approach for evaluation of the surface Laplacian of the electroencephalogram and magnetoencephalogram. *Ann Biomed Eng.* 2001;29(3):202–13.
30. Ghani U, Wasim M, Khan US, Mubasher Saleem M, Hassan A, Rashid N, Islam Tiwana M, Hamza A, Kashif A. Efficient FIR filter implementations for multichannel BCIs using Xilinx system generator. *Biomed Res Int.* 2018;2018:9861350.
31. Aghaei AS, Mahanta MS, Plataniotis KN. Separable common spatio-spectral patterns for motor imagery BCI systems. *IEEE Trans Biomed Eng.* 2016;63(1):15–29.
32. Dong E, Li C, Li L, Du S, Belkacem AN, Chen C. Classification of multi-class motor imagery with a novel hierarchical SVM algorithm for brain–computer interfaces. *Med Biol Eng Comput.* 2017;55(10):1809–18.
33. El Bahy MM, Hosny M, Mohamed WA, Ibrahim S. EEG signal classification using neural network and support vector machine in brain computer interface. In: *Proceedings of the International Conference on Advanced Intelligent Systems and Informatics*. Edited by Hassanién AE, Shaalan K, Gaber T, Azar AT, Tolba MF. Cham: Springer International Publishing; 2017: 246–256.
34. Gaur P, Pachori RB, Wang H, Prasad G. A multi-class EEG-based BCI classification using multivariate empirical mode decomposition based filtering and Riemannian geometry. *Expert Syst Appl.* 2018;95(Supplement C):201–11.
35. Luo J, Feng Z, Zhang J, Lu N. Dynamic frequency feature selection based approach for classification of motor imageries. *Comput Biol Med.* 2016;75:45–53.
36. Miao M, Wang A, Liu F. A spatial-frequency-temporal optimized feature sparse representation-based classification method for motor imagery EEG pattern recognition. *Med Biol Eng Comput.* 2017;55(9):1589–603.
37. Mingai L, Shuoda G, Jinfu Y, Yanjun S. A novel EEG feature extraction method based on OEMD and CSP algorithm. *J Intell Fuzzy Syst* 2016:1–13.
38. Wei Q, Wei Z. Binary particle swarm optimization for frequency band selection in motor imagery based brain–computer interfaces. *Bio-Med Mater Eng.* 2015;26(s1):S1523–32.
39. Yang B, Li H, Wang Q, Zhang Y. Subject-based feature extraction by using fisher WPD-CSP in brain–computer interfaces. *Comput Methods Programs Biomed.* 2016;129:21–8.
40. Yuksel A, Olmez T. A neural network-based optimal spatial filter design method for motor imagery classification. *PLoS ONE.* 2015;10(5):e0125039.
41. Zhang S, Zheng Y, Wang D, Wang L, Ma J, Zhang J, Xu W, Li D, Zhang D. Application of a common spatial pattern-based algorithm for an fNIRS-based motor imagery brain-computer interface. *Neurosci Lett.* 2017;655(Supplement C):35–40.
42. Zhang Y, Wang Y, Jin J, Wang X. Sparse Bayesian learning for obtaining sparsity of EEG frequency bands based feature vectors in motor imagery classification. *Int J Neural Syst.* 2017;27(02):1650032.
43. Kumar S, Sharma A, Tsunoda T. Brain wave classification using long short-term memory network based OPTICAL predictor. *Sci Rep.* 2019;9(1):9153.
44. Hamzah N, Norhazman H, Zaini N, Sani M. Classification of EEG signals based on different motor movement using multi-layer perceptron artificial neural network. *J Biol Sci.* 2016;16(7):265–71.
45. Ma Y, Ding X, She Q, Luo Z, Potter T, Zhang Y. Classification of motor imagery EEG signals with support vector machines and particle swarm optimization. *Comput Math Methods Med.* 2016;2016:8.
46. Hooda N, Kumar N. Cognitive imagery classification of EEG signals using csp-based feature selection method. *IETE Tech Rev* 2019:1–12.
47. Wang J, Feng Z, Lu N, Sun L, Luo J. An information fusion scheme based common spatial pattern method for classification of motor imagery tasks. *Biomed Signal Process Control.* 2018;46:10–7.
48. Nguyen T, Hettiarachchi I, Khatami A, Gordon-Brown L, Lim CP, Nahavandi S. Classification of multi-class BCI data by common spatial pattern and fuzzy system. *IEEE Access.* 2018;6:27873–84.
49. Alotaiby TN, Alshebeili SA, Alotaibi FM, Alrshoud SR. Epileptic seizure prediction using CSP and LDA for scalp EEG signals. *Comput Intell Neurosci.* 2017;2017:1240323–1240323.
50. Kumar S, Sharma A, Tsunoda T. Subject-specific-frequency-band for motor imagery EEG signal recognition based on common spatial spectral pattern. *Lecture Notes in Artificial Intelligence: Sub-series of Lecture Notes in Computer Science* 2019, 11671.
51. Kumar S, Sharma A. A new parameter tuning approach for enhanced motor imagery EEG signal classification. *Med Biol Eng Comput.* 2018;56(10):1861–74.
52. Kumar S, Mamun K, Sharma A. CSP-TSM: optimizing the performance of Riemannian tangent space mapping using common spatial pattern for MI-BCI. *Comput Biol Med.* 2017;91(Supplement C):231–42.
53. Sharma R, Kumar S, Tsunoda T, Patil A, Sharma A. Predicting MoRFs in protein sequences using HMM profiles. *BMC Bioinform.* 2016;17(Suppl 19):251–8.
54. Kumar S, Sharma R, Sharma A, Tsunoda T. Decimation filter with common spatial pattern and fishers discriminant analysis for motor imagery classification. In: *2016 International Joint Conference on Neural Networks (IJCNN): 24–29 July 2016; Vancouver, Canada.* 2090–2095.
55. Kumar S, Sharma A, Mamun K, Tsunoda T. Application of cepstrum analysis and linear predictive coding for motor imaginary task classification. In: *2nd Asia-Pacific World congress on computer science & engineering: 2–4 December 2015; Shangri-La Fijian Resort, Fiji.*
56. Ang KK, Chin ZY, Zhang H, Guan C. Filter bank common spatial pattern (FBCSP) in brain–computer interface. In: *IEEE international joint conference on neural networks (IEEE World Congress on Computational Intelligence): 1–8 June 2008; Hong Kong.* 2390–2397.
57. Arvaneh M, Umilta A, Robertson IH. Filter bank common spatial patterns in mental workload estimation. In: *37th annual international conference of the IEEE engineering in medicine and biology society (EMBC): 25–29 August 2015.* 4749–4752.
58. Das AK, Suresh S, Sundararajan N. A discriminative subject-specific spatio-spectral filter selection approach for EEG based motor-imagery task classification. *Expert Syst Appl.* 2016;64:375–84.

59. Novi Q, Cuntai G, Dat TH, Ping X. Sub-band common spatial pattern (SBCSP) for brain–computer interface. In: *3rd International IEEE/EMBS conference on neural engineering: 2–5 May 2007 2007*. 204–207.
60. Raza H, Cecotti H, Prasad G. Optimising frequency band selection with forward-addition and backward-elimination algorithms in EEG-based brain–computer interfaces. In: *2015 International Joint Conference on Neural Networks (IJCNN): 12–17 July 2015 2015*. 1–7.
61. Thomas KP, Cuntai G, Lau CT, Vinod AP, Keng KA. A new discriminative common spatial pattern method for motor imagery brain computer interfaces. *IEEE Trans Biomed Eng.* 2009;56(11):2730–3.
62. Zhang Y, Zhou G, Jin J, Wang X, Cichocki A. Optimizing spatial patterns with sparse filter bands for motor-imagery based brain–computer interface. *J Neurosci Methods.* 2015;255:85–91.
63. Younghak S, Seungchan L, Junho L, Heung-No L. Sparse representation-based classification scheme for motor imagery-based brain–computer interface systems. *J Neural Eng.* 2012;9(5):056002.
64. Kumar S, Mamun K, Sharma A. CSP-TSM: optimizing the performance of Riemannian tangent space mapping using common spatial pattern for MI-BCI. *Comput Biol Med.* 2017;91:231–42.
65. Wu W, Gao X, Hong B, Gao S. Classifying single-trial EEG during motor imagery by iterative spatio-spectral patterns learning (ISSPL). *IEEE Trans Biomed Eng.* 2008;55(6):1733–43.
66. Li Y, Wen P. Modified CC-LR algorithm with three diverse feature sets for motor imagery tasks classification in EEG based brain–computer interface. *Comput Methods Programs Biomed.* 2014;113(3):767–80.
67. Kevric J, Subasi A. Comparison of signal decomposition methods in classification of EEG signals for motor-imagery BCI system. *Biomed Signal Process Control.* 2017;31:398–406.
68. Zabalza J, Ren J, Zheng J, Zhao H, Qing C, Yang Z, Du P, Marshall S. Novel segmented stacked autoencoder for effective dimensionality reduction and feature extraction in hyperspectral imaging. *Neurocomputing.* 2016;185:1–10.
69. Zabalza J, Ren J, Yang M, Zhang Y, Wang J, Marshall S, Han J. Novel Folded-PCA for improved feature extraction and data reduction with hyperspectral imaging and SAR in remote sensing. *ISPRS J Photogramm Remote Sens.* 2014;93:112–22.
70. Sharma A, Paliwal KK, Imoto S, Miyano S. A feature selection method using improved regularized linear discriminant analysis. *Mach Vis Appl.* 2014;25(3):775–86.
71. Sharma A, Vans E, Shigemizu D, Boroevich KA, Tsunoda T. DeepInsight: a methodology to transform a non-image data to an image for convolution neural network architecture. *Sci Rep.* 2019;9(1):11399.
72. Padfield N, Zabalza J, Zhao H, Masero V, Ren J. EEG-based brain–computer interfaces using motor-imagery: techniques and challenges. *Sensors.* 2019;19(6):1423.
73. Kumar S, Sharma R, Sharma A. OPTICAL+: a frequency-based deep learning scheme for recognizing brain wave signals. *PeerJ Comput Scis.* 2021;7:e375.
74. Dornhege G, Blankertz B, Curio G, Müller K. Boosting bit rates in noninvasive EEG single-trial classifications by feature combination and multiclass paradigms. *IEEE Trans Biomed Eng.* 2004;51(6):993–1002.
75. Blankertz B, Dornhege G, Krauledat M, Müller K-R, Curio G. The non-invasive Berlin brain–computer interface: fast acquisition of effective performance in untrained subjects. *Neuroimage.* 2007;37(2):539–50.
76. Kumar S, Sharma R, Sharma A, Tsunoda T. Decimation filter with common spatial pattern and fishers discriminant analysis for motor imagery classification In: *IEEE World congress on computational intelligence: 24–29th July; Vancouver, Canada.* 2016.
77. Tuzel O, Porikli F, Meer P. Pedestrian detection via classification on Riemannian manifolds. *IEEE Trans Pattern Anal Mach Intell.* 2008;30(10):1713–27.
78. Sharma A, Kamola PJ, Tsunoda T. 2D–EM clustering approach for high-dimensional data through folding feature vectors. *BMC Bioinform.* 2017;18(16):547.
79. Sharma A, Boroevich K, Shigemizu D, Kamatani Y, Kubo M, Tsunoda T. Hierarchical maximum likelihood clustering approach. *IEEE Trans Biomed Eng.* 2017;64(1):112–22.
80. Sharma A, Shigemizu D, Boroevich KA, López Y, Kamatani Y, Kubo M, Tsunoda T. Stepwise iterative maximum likelihood clustering approach. *BMC Bioinform.* 2016;17(319):1–14.
81. Tibshirani R. Regression shrinkage and selection via the Lasso. *J R Stat Soc B.* 1996;58(1):267–88.
82. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33(1):159–74.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

