Narrow Window Feature Extraction for EEG-Motor Imagery Classification using k-NN and Voting Scheme

Adi Wijaya
Electrical Engineering and IT Department
Universitas Gadjah Mada
Yogyakarta, Indonesia
adiwijaya@mail.ugm.ac.id

Teguh Bharata Adji
Electrical Engineering and IT Department
Universitas Gadjah Mada
Yogyakarta, Indonesia
adji@ugm.ac.id

Noor Akhmad Setiawan
Electrical Engineering and IT Department
Universitas Gadjah Mada
Yogyakarta, Indonesia
noorwewe@mail.ugm.ac.id

***Abstract*—Achieving consistent accuracy still big challenge in EEG based Motor Imagery classification since the nature of EEG signal is non-stationary, intra-subject and inter-subject dependent. To address this problems, we propose the feature extraction scheme employing statistical measurements in narrow window with channel instantiation approach. In this study, k-Nearest Neighbor is used and a voting scheme as final decision where the most detection in certain class will be a winner. In this channel instantiation scheme, where EEG channel become instance or record, seventeen EEG channels with motor related activity is used to reduce from 118 channels. We investigate five narrow windows combination in the proposed methods, i.e.: one, two, three, four and five windows. BCI competition III Dataset IVa is used to evaluate our proposed methods. Experimental results show that one window with all channel and a combination of five windows with reduced channel outperform all prior research with highest accuracy and lowest standard deviation. This results indicate that our proposed methods achieve consistent accuracy and promising for reliable BCI systems.**

Keywords*—*EEG, motor imagery, narrow window, channel instantiation, voting scheme

I. Introduction

The term Brain-Computer Interface (BCI) is a system which translate the human brain signals and a communication technique to control a device [1]–[3]. Because of its great potential in areas such as rehabilitation and entertainment, BCI has been studied and developed actively by researchers [4] over the past two decades [5] that is usually recorded using electroencephalogram (EEG) [6]. Motor Imagery (MI) task is one of the most studied types of EEG signals in BCI systems [7]. MI is a mental activity about special motor movement without actual execution [8] or motor output [9]. Many studies have proved that MI plays a crucial role in motor skill learning, prosthesis control and rehabilitation of motor abilities [10].

In BCI, EEG signal is one of the most popular techniques due to low cost and non-invasive nature of EEG [3]. However, the nature of EEG signal has too much noise, [11] non-stationary [12], [13] and subject-dependent [12], [14], [15] that affect the classification results [11]. Therefore, processing of EEG signals, which directly affects the classification accuracy, still represents an important and challenging issue [6]. EEG signal recognition is the key technology of MI-BCI that includes feature extraction and classification [16], [17].

Many studies have been conducted to address those such problems using various signal processing technique such as: time domain, frequency domain, time-frequency domain and spatial domain. Among these approach, time-frequency and spatial domain are commonly used by many studies. These signal processing technique are used as feature extraction and combined with classifier both single classifier or ensemble technique. Several researches also used feature selection since EEG signal recorded with multi-channel.

In spatial domain, Common Spatial Pattern (CSP) is one of most popular technique in EEG based MI classification. Therefore, many studies have been used and improve CSP such as: CSP with sparse time-frequency segment common spatial pattern [18], CSP combined with Differential Evolution as feature selection [19], CSP combined with symmetrical feature [20] and Filter Bank Regularized Common Spatial Pattern [21]. Another approach based on time-frequency signal processing technique by employing three popular wavelet technique has been conducted by [6].

Based on prior research results, CSP still need improvement since CSP very sensitive to noise in nature and often over-fit with small training sets [22]. Existing studies that employed CSP yield excellent accuracy; however, the result still not consistent where some subsets gained low accuracy. Therefore, how to improve the detection performance of EEG based MI is still a challenging issue for the development of BCI systems [22].

In this study, we propose narrow window feature extraction with seven statistical features to tackle non-stationary nature of EEG signal and employing instance-based learning with k-Nearest Neighbor (k-NN) as classifier. K-NN proven as promising classifier in EEG based MI classification [6]. Since instance-based learning need more data, we employ data transformation approach called channel instantiation where EEG multi-channel data transformed into instance which introduced by [22].

The motivation of this work was to analyse the effectiveness of the narrow window feature extraction in channel instantiation approach using k-NN and voting scheme for making accurate and consistent detection of EEG based MI. Another motivation was to analyse a feature reduction with minimum channel by select only motor related activity channels that reduced computation time meanwhile maintain excellent and consistent accuracy. Thus, the proposed method expectedly suitable for reliable BCI systems for further implementation.

978-1-5386-8402-3/18/$31.00 ©2018 IEEE

II. Related Works

EEG based MI classification require two main tasks i.e. feature extraction and classification. In feature extraction, many approaches have been used combined with statistical feature. Since in EEG signal recorded in multi-channel, channel selection also considered by many researchers to reduce computation time yet maintain high accuracy. In this study, the most relevant prior researches are selected as related works based on dataset used (BCI competition III dataset IVa) and method validation with 10-fold cross validation.

In [19], they utilized a differential evolution (DE) based technique to select most relevant features in EEG based MI and CSP as feature extractor (CSP+DE-FS). They achieved excellent accuracy and small standard deviation with 96.02% and 3.77% respectively. However, CSP+DE-FS method is slow compared to the typical feature selection algorithms and the even more slower because of employing DE as a wrapper technique.

Another CSP based feature extraction is used by [21] with improved CSP so called Filter Bank Regularized CSP (FBRCSP). FBRCSP use eighteen selected channel based on the Homunculus Theory and achieve 86.23% and 9.55% for accuracy and standard deviation respectively. FBRCSP not only feature extractor but also classifier based on distance measurement by calculate the distance between each feature vector and each class mean vector. However, FBRCSP need suitable selected parameter set that caused its varying performance.

Another promising study introduced by [18]. They improve CSP with sparse time-frequency segment common spatial pattern (STFSCSP) combined with Discernibility of Feature Sets (DFS) criteria that dedicated for spatial filter optimization and Weighted Naïve Bayesian Classifier (WNBC). They gained excellent result with accuracy about 92.66% and standard deviation about 7.78%. However, the cost of more computation at classification task both in training and testing stages.

A novel feature extraction so called symmetrical feature that built upon CSP (CSP+SF) was introduced by [20]. They achieve 82.06% and 13.4% for accuracy and standard deviation. Although CSP+SF has a robust characteristic of invariant EEG data compared with the previous CSP power band; however, the average results showed that the new feature type has lower performance in terms of power than the original CSP.

Besides CSP, another approach based on signal processing technique by employing three popular technique, i.e.: Empirical Mode Decomposition, Discrete Wavelet Transform, and Wavelet Packet Decomposition has been conducted by [6]. They utilized higher order statistic (HOS) as main feature extractor, k-Nearest Neighbour (k-NN) as classifier and choose only three channels (C3, Cz, C4). Their method achieved excellent result with 92.8% and 2.93% for accuracy and standard deviation respectively. However, their proposed method does not offer the best classification accuracy for all subjects with subject "av" gained lowest accuracy that below 90%.

III. Material and Methods

A. Dataset

Dataset IVa from BCI competition III [23] is used in this study. This data set was recorded from five healthy subjects ("aa", "al", "av", "aw", "ay"). EEG signal recorded for each subject with comfortable chair with their arms resting on the armrests. In this task, all subjects performed three types of motor imageries i.e.: right foot, left hand and right hand. However, for the competition, only right hand and right foot were provided. The recording signals were measured based on extended international 10/20-system with 118 EEG channels. Although signals digitized at 1000 Hz with 16 bit, the data down-sampled at 100 Hz (by picking each 10th sample) also available for analysis. In this study, this 100 Hz down-sampled data is used for EEG based MI classification task. Each subject performed 280 trials in total with the composition for training and testing sets. In this study, although every subject has separate training and testing sets, they were combined into one dataset due to the low number of trials and imbalance between training and testing sets.

B. Narrow Window Feature Extraction

In this study, narrow window feature extraction approach is used as shown in Fig. 1. This strategy is selected due to tackling the non-stationary nature of EEG signal, where smaller or narrower windows, will exhibit signal stationary [24]. As initial step, original dataset which contains EEG signal is filtered using 4^{th} order Butterworth band-pass filter as commonly used in EEG signal processing [14], [25], [26]. EEG signal were filtered in frequency range 8-30 Hz as same range with [19], [27]–[30]. After filtered, a time slot or window between 0.5 – 2 seconds is chosen for further process. This time slot is narrower window than some prior researches, such as: 0.5 – 3.5 seconds [14], 0 – 3.5 seconds [6], [19] and 0.5 – 2.5 seconds [21]. The selected time slot, which is range about 1.5 seconds, consists of 150 data points because 100 Hz down-sampled data is used in this study. These 150 data points then divided into 5 windows for further feature extraction process. So each windows consist 30 data points, which much enough as sample size for statistical calculation.

The next step is conduct feature extraction by using seven statistical features for each window. The following seven statistical features are chosen for EEG based MI classification:

- Mean Absolute Value (mav):

$$\mathrm{mav} = \frac{1}{n}\sum_{i=1}^{N} |x_n| \quad (1)$$

- Root Means Square (rms):

$$\mathrm{rms} = \sqrt{\frac{1}{n}\sum_{1}^{N} x_n^2} \quad (2)$$

- Standard Deviation (σ):

$$\sigma = \sqrt{\frac{1}{n}\sum_{1}^{N}(x_n - \mu)^2} \quad (3)$$

- Skewness:

$$\text{skewness} = \sqrt{\frac{1}{n}\sum_{1}^{N}\frac{(x_n-\mu)^3}{\sigma^3}} \quad (4)$$

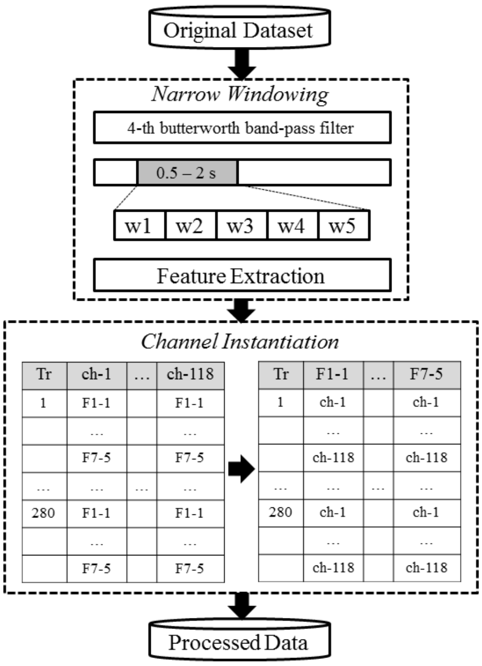

Fig. 1. Narrow window feature extraction and channel instantiation blog diagram

- Kurtosis:

$$\text{kurtosis} = \sqrt{\frac{1}{n}\sum_{1}^{N}\frac{(x_n-\mu)^4}{\sigma^4}} \quad (5)$$

- Variance to Means Ratio (vmr):

$$\text{vmr} = \sigma^2/\mu \quad (6)$$

- Coefficient of Variation (cv):

$$\text{cv} = \sigma/\mu \quad (7).$$

After feature extraction calculated, the next step is channel instantiation step. Channel instantiation means a transformation from column into row, from feature into instance or record. In original dataset, EEG signal were recorded in multi-channel, so the recorded data actually belongs to channel. In this study, each channel has 35 features, comes from 7 statistical features multiple with 5 windows, for each trial. The channel then transformed into row or record and 35 statistical features become column or feature. Since each subject in BCI competition III Dataset IVa has 280 trials, so each subject has 280 x 118 records. In this study, we also use a 17-selected channels (FC3, FC1, FCz, FC2, FC4, C5, C3, C1, Cz, C2, C4, C6, CP3, CP1, CPz, CP2, CP4) since we were only concerned with motor related activity as used by [31] and according to the homunculus theory that represents motor activity area [32].

This dataset transformation based on channel instantiation that proposed by [22]. However, their proposed method only transform in class level not trial level and then perform classification. In our proposed method, we perform classification in trial level not class level which is a mandatory task that should be performed in BCI competition. This feature extraction and channel instantiation will produce a processed data that will be used further for the classification task.

C. k-NN and Voting Scheme

Processed data from feature selection stage than feed to 10-fold cross validation (CV) scheme as shown in Fig.2. In 10-fold CV, data will split into 10 subsets mutually exclusive with equal size where nine subsets are used for training and one subset is used for testing and this process repeats 10 times where the testing data is different for each process. As shown in Fig. 2, channel selection is conducted on processed data with 17 channels that related to motor activity.

In 10-fold CV, k-NN is used as classifier since k-NN support incremental learning, able to model complex decision spaces having hyper-polygonal shapes that may not be as easily describable by other learning algorithms such as decision trees and widely used in the area of pattern recognition [33, p. 423]. k-NN searches the pattern space for the k training instance that are closest to the testing instance. These k training instance are the k “nearest neighbors” of the testing instance. Closeness is defined in terms of a distance metric, such as Euclidean distance, Canberra distance, Manhattan distance, etc. In this study, Canberra distance is used as numerical measurement in k-NN parameter. After 10-fold CV conducted, the detected data then saved to MySQL database and then we calculate the accuracy based on voting scheme.

In this voting scheme, the most certain class (class 1 or class 2) detected as a final decision for each trial. Since in this study, one single trial has 118 detected records for all channel scheme and 17 detected records for selected channel scheme. The decision whether class 1 or class 2 based on which channel is most detected to certain class. Since there are 118 channels, we create simple majority voting threshold where detection more than half of 118 for all channel scheme and 17 for selected channel scheme belongs to certain class, so the decision is belongs to its class.

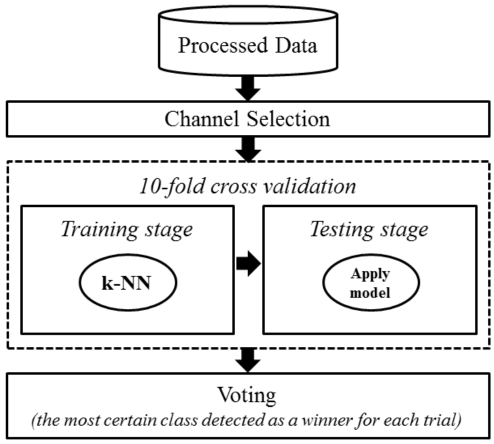

Fig. 2. k-NN detection and voting scheme block diagram

D. Evaluation

In this study, overall accuracy is used as main evaluation since many prior research used accuracy as main evaluation. This simply calculate based on number of true detection divided to total number of trial. In BCI competition III dataset IV-a, all subjects (sub-dataset) has 280 trials. Although every subject consisted different train and test sets, they were combined into one dataset due to the low number of trials as used by some prior research [6], [18], [19]. This evaluation calculated after voting scheme is conducted.

IV. Results and Discussion

The experiments are conducted using a computing platform based on Intel i7 CPU and 8 GB RAM with Microsoft Windows 8.1 64-bit operating system. In development tools, Rapidminer 8.1 Educational License and MySQL are used. Matlab R2014b is employed as signal filtering tool and HeidiSQL version 9.4 as MySQL editor for calculating accuracy based on stored data from Rapidminer process with comply to voting scheme.

In this study, we propose two groups based on number of channel used, i.e.: all channels and selected channel. In all channel, only single window of feature extraction is used. Meanwhile in selected channel, a combination of windows are used. Table I and II show the performance results of our proposed methods.

A. All Channel Classification Results

In Table I, all single window gain excellent accuracy with 5-th window (*w5*) gain the best accuracy and lowest standard deviation with 99.71% and 0.16% respectively. All single window gain perfect accuracy for subject “ay”. The *w5* gain most consistent result compared to other single window. This window then used in combination windows scheme which utilize selected channel only.

TABLE I. Single Window with All Channel

w	Accuracy (%)					*Average*
	aa	*al*	*av*	*aw*	*ay*	
w1	97.86	100	99.64	100	100	99.5 ± 0.9
w2	98.21	98.93	100	100	100	99.43 ± 0.82
w3	98.21	99.64	100	100	100	99.57 ± 0.78
w4	98.21	99.64	98.93	100	100	99.36 ± 0.78
w5	99.64	99.64	99.64	99.64	100	**99.71 ± 0.16**

As shown in Table I, all single window empirically gain excellent and consistent performance with low standard deviation (all below 1%). This findings prove that the narrow window is effective in tackling EEG signal non-stationary [24]. However, this scheme utilize all channels (118 channels). In online BCI system, the using of many channels is not preferable since it will need more resource in EEG signal recording, time and computation [19], etc. Thus, reducing channel is one of solution to tackle those such issues.

B. Selected Channel Classification results

Table II presents five our other proposed methods and its results. As shown in Table II, single windows comes from the best windows, *w5* and the rest comes from combination of two until five windows. In these proposed methods, only seventeen channels included in computation. Single window (*w5*) gain lowest performance both in accuracy and standard deviation. Meanwhile, the performance increase with the number of windows increase. A five windows combination, consists of 1-st, 2-nd, 3-rd, 4-th and 5-th windows (*w12345*), gain the best performance in matter of highest accuracy and lowest standard deviation with 99.41% and 0.46% respectively. The *w12345* with selected channel classification performance almost similar with the best single window (*w5*) with all channels (99.21% ± 0.45% vs. 99.71% ± 0.16%). These findings prove that combination of several narrow window even with minimum channel still gain excellent performance in matter of excellent accuracy and consistent performance in matter of low in standard deviation.

TABLE II. Comibined Windows with Selected Channel

w	Accuracy (%)					Average (%)
	aa	*al*	*av*	*aw*	*ay*	
w[a]	86.79	80.00	91.79	87.14	92.50	87.64 ± 5
w[b]	93.21	95.00	95.71	96.79	97.14	95.57 ± 1.57
w[c]	96.79	97.50	96.79	99.29	98.93	97.86 ± 1.18
w[d]	98.57	99.29	98.57	98.93	99.64	99 ± 0.47
w[e]	98.57	99.29	99.64	98.93	99.64	**99.21 ± 0.46**

[a.] w5, [b.] w12, [c.] w123, [d.] w1234, [e.] w12345

In this study, one trial of EEG based MI task has 17 instances to be detected whether class 1 (right hand) or class 2 (right foot). Since one trial has 17 detection results, at least 9 true detections as true detection for corresponding trial. In other word, false detection assigned to corresponding trial when number of misdetection higher or equal to 9. Table III presents averaged number of misdetection distribution among 280 trials for all subjects.

TABLE III. Misdetection Distribution of Averaged Results from All Subjects for Combined Windows with Selected Channel

range	#misdetection distribution				
	w5	*w12*	*w123*	*w1234*	*w12345*
0 – 2	41	76.2	105.4	**129**	**150.8**
3 – 5	**107.2**	**125**	**123.8**	115.4	102.2
6 – 8	97.2	66.4	44.8	32.8	24.8
>=9	34.6	12.4	6	2.8	2.2
Total	280	280	280	280	280
#True Det.	245.4	267.6	274	277.2	277.8
Accuracy (%)	87.64	95.57	97.86	99	**99.21**

As shown in Table III, *w5* has highest number of misdetection that caused false detection and decrease its accuracy (range >=9); meanwhile *w12345* has lowest number of misdetection. More windows combination increase number of misdetection in lower range (0 – 2 until 6 – 8) that cause increase true detection and finally increase its accuracy (#True Det. divided by Total). In this study, it needs at least 2 narrow windows in order to achieve excellent accuracy (more than 90%). This findings proof that narrow window is still effective although with channel reduction.

In order to analyze the effectiveness of the proposed methods, we compare our proposed methods to the most relevant prior research. Five prior researches, as describe in Section 2, are selected as benchmark and comparison. Table IV presents its comparison.

TABLE IV. PERFORMANCE COMPARISON TO PRIOR RESEARCH

Authors	Method	Accuracy (%)*					
		aa	*al*	*av*	*aw*	*ay*	*average*
Baig *et al.* [19]	CSP+DE-FS	95.8	98.8	89.8	99.2	96.5	96.02 ± 3.77
Kevric and Subasi [6]	MSPCA+WPD+HOS	96	92.3	88.9	95.4	91.4	92.8 ± 2.93
Park S. H. *et al.* [21]	FBRCSP	91.07	94.64	75	76.78	93.65	86.23 ± 9.55
Miao *et al.* [18]	STFSCSP	95.2	98.58	79.41	97.78	95.02	92.66 ± 7.78
Park S. M. *et al.* [20]	CSP+SF	72.62	95.92	63.54	89.85	88.38	82.06 ± 13.4
Proposed Method #1	*w5+all-ch+k-NN+VS*	**99.64**	**99.64**	**99.64**	**99.64**	**100**	**99.71 ± 0.16**
Proposed Method #2	*w5+17-ch+k-NN+VS*	86.79	80.00	**91.79**	87.14	92.50	87.64 ± 5
Proposed Method #3	*w12+17-ch+k-NN+VS*	93.21	95.00	**95.71**	96.79	**97.14**	95.57 ± 1.57
Proposed Method #4	*w123+17-ch+k-NN+VS*	**96.79**	97.50	**96.79**	**99.29**	**98.93**	**97.86 ± 1.18**
Proposed Method #5	*w1234+17-ch+k-NN+VS*	**98.57**	**99.29**	**98.57**	**98.93**	**99.64**	**99 ± 0.47**
Proposed Method #6	*w12345+17-ch+k-NN+VS*	**98.57**	**99.29**	**99.64**	**98.93**	**99.64**	**99.21 ± 0.46**

*bold number means the best results compared to prior researches, underlined number means the best results among prior researches

As shown in Table IV, all prior researches yielded low accuracy for subject "av" (below 90%) and three of them gained wide standard deviation (more than 5%). The best performance form the prior research is hold by Baig *et al.* [19] with 96.02% and 3.77% for accuracy and standard deviation respectively. Compared to our proposed method, especially the proposed methods with selected channel, 4 of 5 our proposed methods outperform all averaged results of the prior researches both in accuracy and standard deviation. Among these results, 2 of 5 our proposed methods outperform not only in averaged result but also outperform in every subjects.

The effectiveness of the proposed methods is a combination strategy from feature extraction scheme until classification scheme. First, the narrow window and its combination in tackling non-stationary EEG signal [24]. Second, the using of higher order statistic (skewness and kurtosis) as statistical feature extraction method [6], mean average value and root mean square [22], [34]. Finally, the using of channel instantiation approach [22] that create more instance that effective for k-NN as instance-based classifier in EEG based MI classification [6].

V. CONCLUSION

In this study, a combination of narrow window feature selection and seventeen motor related activity EEG channels were used for classification task of MI based EEG signal. A multi-channel with 2-class dataset taken from BCI competition III Dataset IVa is used for this purpose. We conducted and evaluated the proposed methods by using promising approach so called channel instantiation as data transformation, k-NN algorithm and voting scheme for final decision. From the experimental results, all proposed methods with all channel yield excellent accuracy and consistent for all subjects. The proposed methods with selected seventeen channels also effective and gain consistent excellent accuracy with at least three narrow windows combination. These results show that the proposed method has the potential to obtain a reliable EEG based MI classification and can be used practically in online BCI systems such as: controlling a wheelchair, rehabilitation therapies for the stroke rehabilitation or improve motor rehabilitation outcomes. However, several issue still open for future works such as: a need for automatic feature selection to reduce the computational time and apply model in multi-class EEG based MI classification to test the robustness of the proposed method.

REFERENCES

[1] D. H. Krishna, I. A. Pasha, and T. S. Savithri, "Classification of EEG Motor Imagery Multi Class Signals Based on Cross Correlation," *Procedia Comput. Sci.*, vol. 85, pp. 490–495, 2016.

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

[3] H. Mirvaziri and Z. S. Mobarakeh, "Improvement of EEG-based motor imagery classification using ring topology-based particle swarm optimization," *Biomed. Signal Process. Control*, vol. 32, pp. 69–75, 2017.

[4] S. Liang, K.-S. Choi, J. Qin, W.-M. Pang, Q. Wang, and P.-A. Heng, "Improving the discrimination of hand motor imagery via virtual reality based visual guidance," *Comput. Methods Programs Biomed.*, vol. 132, pp. 63–74, 2016.

[5] Y. Yu *et al.*, "Toward brain-actuated car applications: Self-paced control with a motor imagery-based brain-computer interface," *Comput. Biol. Med.*, 2016.

[6] J. Kevric and A. Subasi, "Comparison of signal decomposition methods in classification of EEG signals for motor-imagery BCI system," *Biomed. Signal Process. Control*, vol. 31, pp. 398–406, 2017.

[7] P. J. García-Laencina, G. Rodríguez-Bermudez, and J. Roca-Dorda, "Exploring dimensionality reduction of EEG features in motor imagery task classification," *Expert Syst. Appl.*, vol. 41, no. 11, pp. 5285–5295, 2014.

[8] K. McInnes, C. Friesen, and S. Boe, "Specific brain lesions impair explicit motor imagery ability: A systematic review of the evidence," *Arch. Phys. Med. Rehabil.*, vol. 97, no. 3, pp. 478–489, 2016.

[9] T. Zhang *et al.*, "Structural and functional correlates of motor

imagery BCI performance: Insights from the patterns of Fronto-Parietal Attention Network," *Neuroimage*, 2016.

[10] H. Burianová *et al.*, "Multimodal functional imaging of motor imagery using a novel paradigm," *Neuroimage*, vol. 71, pp. 50–58, 2013.

[11] L. He, B. Liu, D. Hu, Y. Wen, M. Wan, and J. Long, "Motor imagery EEG signals analysis based on Bayesian network with Gaussian distribution," *Neurocomputing*, vol. 188, pp. 217–224, 2016.

[12] T. Kayikcioglu and O. Aydemir, "A polynomial fitting and k-NN based approach for improving classification of motor imagery BCI data," *Pattern Recognit. Lett.*, vol. 31, no. 11, pp. 1207–1215, Aug. 2010.

[13] S. Chatterjee, N. Ray Choudhury, and R. Bose, "Detection of epileptic seizure and seizure-free EEG signals employing generalised S -transform," *IET Sci. Meas. Technol.*, vol. 11, no. 7, pp. 847–855, 2017.

[14] J. Meng, L. Yao, X. Sheng, D. Zhang, and X. Zhu, "Simultaneously Optimizing Spatial Spectral Features Based on Mutual Information for EEG Classification," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 1, pp. 227–240, 2015.

[15] D. Li, H. Zhang, M. S. Khan, and F. Mi, "A self-adaptive frequency selection common spatial pattern and least squares twin support vector machine for motor imagery electroencephalography recognition," *Biomed. Signal Process. Control*, vol. 41, pp. 222–232, 2018.

[16] L. Duan, M. Bao, J. Miao, Y. Xu, and J. Chen, "Classification Based on Multilayer Extreme Learning Machine for Motor Imagery Task from EEG Signals," *Procedia Comput. Sci.*, vol. 88, pp. 176–184, 2016.

[17] J. S. Suri, A. Kumar, G. K. Singh, and M. K. Ahirwal, "Sub-band classification of decomposed single event-related potential co-variants for multi-class brain–computer interface: a qualitative and quantitative approach," *IET Sci. Meas. Technol.*, vol. 10, no. 4, pp. 355–363, 2016.

[18] M. Miao, H. Zeng, A. Wang, C. Zhao, and F. Liu, "Discriminative spatial-frequency-temporal feature extraction and classification of motor imagery EEG: An sparse regression and Weighted Naïve Bayesian Classifier-based approach," *J. Neurosci. Methods*, vol. 278, pp. 13–24, 2017.

[19] M. Z. Baig, N. Aslam, H. P. H. Shum, and L. Zhang, "Differential Evolution Algorithm as a Tool for Optimal Feature Subset Selection in Motor Imagery EEG," *Expert Syst. Appl.*, vol. 90, pp. 184–195, 2017.

[20] S. M. Park, X. Yu, P. Chum, W. Y. Lee, and K. B. Sim, "Symmetrical feature for interpreting motor imagery EEG signals in the brain???computer interface," *Optik (Stuttg).*, vol. 129, pp. 163–171, 2017.

[21] S.-H. Park, D. Lee, and S.-G. Lee, "Filter Bank Regularized Common Spatial Pattern Ensemble for Small Sample Motor Imagery Classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 4320, no. c, pp. 1–1, 2017.

[22] Siuly, H. Wang, and Y. Zhang, "Detection of motor imagery EEG signals employing Naïve Bayes based learning process," *Meas. J. Int. Meas. Confed.*, vol. 86, pp. 148–158, 2016.

[23] B. Blankertz *et al.*, "The BCI competition III: Validating alternative approaches to actual BCI problems," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 14, no. 2, pp. 153–159, 2006.

[24] Siuly and Y. Li, "A novel statistical algorithm for multiclass EEG signal classification," *Eng. Appl. Artif. Intell.*, vol. 34, pp. 154–167, 2014.

[25] L. Sun, Z. Feng, B. Chen, and N. Lu, "A contralateral channel guided model for EEG based motor imagery classification," *Biomed. Signal Process. Control*, vol. 41, pp. 1–9, 2018.

[26] T. Uehara, M. Sartori, T. Tanaka, and S. Fiori, "Robust Averaging of Covariances for EEG Recordings Classification in Motor Imagery Brain-Computer Interfaces," *Neural Comput.*, 2017.

[27] L. Duan, Z. Hongxin, M. S. Khan, and M. Fang, "Recognition of motor imagery tasks for BCI using CSP and chaotic PSO twin SVM," *J. China Univ. Posts Telecommun.*, vol. 24, no. 3, pp. 83–90, 2017.

[28] C. Park, C. C. Took, and D. P. Mandic, "Augmented Complex Common Spatial Patterns for Classification of Noncircular EEG From Motor Imagery Tasks," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 1, pp. 1–10, Jan. 2014.

[29] W. Y. Hsu, "EEG-based motor imagery classification using neuro-fuzzy prediction and wavelet fractal features," *J. Neurosci. Methods*, vol. 189, no. 2, pp. 295–302, 2010.

[30] X. Tang, N. Zhang, J. Zhou, and Q. Liu, "Hidden-layer visible deep stacking network optimized by PSO for motor imagery EEG recognition," *Neurocomputing*, vol. 234, no. December 2016, pp. 1–10, 2017.

[31] B. J. Edelman, B. Baxter, and B. He, "EEG Source Imaging Enhances the Decoding of Complex Right-Hand Motor Imagery Tasks," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 1, pp. 4–14, 2016.

[32] C. C. J. M. De Klerk, M. H. Johnson, and V. Southgate, "An EEG study on the somatotopic organisation of sensorimotor cortex activation during action execution and observation in infancy," *Dev. Cogn. Neurosci.*, vol. 15, pp. 1–10, 2015.

[33] J. Han, M. Kamber, and J. Pei, *Data Mining*, 11th ed. Morgan Kaufman, 2011.

[34] P. K. Pattnaik and J. Sarraf, "Brain Computer Interface issues on hand movement," *J. King Saud Univ. - Comput. Inf. Sci.*, 2016.