



## **Logistic Regression based Feature Selection and Two-Stage Detection for EEG based Motor Imagery Classification**

**Adi Wijaya<sup>1,2</sup>**

**Teguh Bharata Adji<sup>1</sup>**

**Noor Ahmad Setiawan<sup>1\*</sup>**

<sup>1</sup>*Department of Electrical and Information Engineering, Universitas Gadjah Mada, Yogyakarta, Indonesia*

<sup>2</sup>*Department of Health Information Management, STIKES Indonesia Maju, Jakarta, Indonesia*

\* Corresponding author's Email: noorwewe@ugm.ac.id

---

**Abstract:** Electroencephalogram (EEG) based motor imagery (MI) classification requires efficient feature extraction and consistent accuracy for reliable brain-computer interface (BCI) systems. Achieving consistent accuracy in EEG-MI classification is still big challenge according to the nature of EEG signal which is subject dependent. To address this problem, we propose a feature selection scheme based on Logistic Regression (LRFS) and two-stage detection (TSD) in channel instantiation approach. In TSD scheme, Linear Discriminant Analysis was utilized in first-stage detection; while Gradient Boosted Tree and k-Nearest Neighbor in second-stage detection. To evaluate the proposed method, two publicly available datasets, BCI competition III-Dataset IVa and BCI competition IV-Dataset 2a, were used. Experimental results show that the proposed method yielded excellent accuracy for both datasets with 95.21% and 94.83%, respectively. These results indicated that the proposed method has consistent accuracy and is promising for reliable BCI systems.

**Keywords:** Motor imagery, Feature selection, Two-stage detection, Gradient boosted tree, Reliable BCI.

---

### **1. Introduction**

Brain-computer interface (BCI) is a communication technique that aims to identify a human brain intents and translate them to control external devices or environment [1–3] that is usually recorded using electroencephalogram (EEG) [4]. BCI techniques have researcher's intention and developed over the past two decades rapidly [5]. One of the most studied types of EEG signals in BCI systems is motor imagery (MI) task [6]. MI is a mental practice of a special motor movement without actual execution [7] or motor output [8]. Many studies have proved that MI plays a crucial role in motor skill learning, prosthesis control and rehabilitation of motor abilities [9].

EEG signal is one of the most popular techniques for monitoring brain activities in BCIs [2] due to low cost and non-invasive nature of EEG. However, the nature of EEG signal has too much noise, [10] non-stationary [11, 12] and subject-dependent [13–15] that affect the classification results [10]. Therefore,

processing of EEG signals, which directly affects the classification accuracy, still represents an important challenge [4, 16]. EEG signal recognition is the key technology of BCI that includes feature extraction and classification [17, 18].

Many studies have been conducted to address those such problems using various signal processing technique combined with machine learning classifier both single classifier or ensemble technique. Several researches also used feature selection since EEG signal recorded with multi-channel [16, 19]. In EEG based MI classification, two publicly available datasets come from BCI competition, namely BCI competition III-Dataset IVa and BCI competition IV-Dataset 2a.

Regarding the first dataset, many studies have been conducted by employing common spatial pattern (CSP) [16, 19–22]. In [20], they proposed the sparse group representation model (SGRM) which aims at finding out the most significant training samples and yielded average accuracy 77.7%. However, their method needs parameter optimization

which takes much time. A dynamic and self-adaptive classification algorithm was proposed by [16] with ability to select EEG channel dynamically. However, their method gained even lower average accuracy with 75.15% and the selected channels differs for each subject. An improved accuracy was produced by [22] study with 82.48%. They proposed CSP-Rank for Multiple Frequency Bands (CSPRMF) and linear discriminant analysis (LDA). Their method also supports channel selection. However, selected channels were different under various frequency bands for each subject and need for further performance improvement. An excellent average accuracy was produced by [19] and [21] with 90.58% and 92.93%, respectively. However, their methods lack of consistent selected channels for each subjects and low consistent performance among all subject with standard deviation more than 5%.

Regarding the second dataset, although BCI competition IV-Dataset 2 recorded in 4 classes, many studies were conducted in 2 classes. Left hand and right hand class were selected in many studies due to corroborate the effectiveness of their method in 2 classes EEG based MI classification. Since this dataset is multi-class in nature, the 2 class classifications of this dataset that have been conducted by previous research have not yet produced excellent performance. Several studies have produced an average accuracy of nearly 90% [16,23–27], while there are still many whose average accuracy is still below 80% [28–30].

Among these studies, many of them use a single classifier, followed by ensemble and deep learning in the form of convolutional neural network (CNN). In single classifier, support vector machine (SVM) were used by [23,29,31], linear discriminant analysis (LDA) [16,30], least square classifier [28] and k-Nearest Neighbor (kNN) [27]. These studies still need further improvement to produce excellent and consistent performance due to either using single classifier or CSP as feature extraction method. Despite its popularity, CSP very sensitive to noise, and often over-fit with small training sets [32]. Therefore, how to improve the recognition performance of EEG based MI is still a vital issue for the development of BCI systems [32] and still ongoing [26].

In this study, we propose a two-stage detection employing Linear Discriminant Analysis (LDA) as first stage detector whereas k-Nearest Neighbor (k-NN) and Gradient Boosted Tree (GBT) as second stage detector. LDA is proven as effective as classifier [33] and k-NN also proven as promising classifier in EEG based MI classification [4]. Meanwhile, GBT as a development and improved

version of Decision Tree (DT) also promising since DT itself is a good classifier. A simple yet effective feature selection also utilized in our proposed method with Logistic Regression (LR) algorithm by utilizing LR ability on feature weighting. LR describes the association between a categorical dependent variable and certain independent variables by using probability scores as the predicted values of the dependent variables [34], which effective used as feature selector. In this study, seven statistical features are used for feature extraction and employed data transformation approach called channel instantiation where EEG multi-channel data transformed into instance which introduced by [32]. It shoul be new paragraph of the proposed method consists of two parts, i.e. (1) a hybrid classifier in form of two-stage detection which is simple yet effective, and (2) simple feature selection based on feature weighting by logistic regression. The combination of these technique will produce not only excellent and consistent performance of the detection but also efficient in time execution. The motivation of this work was to analyse the effectiveness of the two-stage detection in channel instantiation approach for making accurate detection of EEG based MI. Another motivation was to develop a feature selection method that suitable for inter-subject which gain excellent and consistent accuracy yet reduce features optimally. Thus, the proposed method suitable for reliable BCI systems for further implementation.

This study is organized as follows. In Section 2, we explain material and method which comprises dataset description and the proposed method from feature extraction, feature selection until classification technique. Following by the experimental result which describe the quantitative result, comparison to previous studies and further discussion in Section 3. Finally, Section 4 presents the conclusion of this study.

## 2. Material and method

### 2.1 Dataset

In this study, two publicly available datasets which provided by BCI competition were used to test the effectiveness of the proposed method. These two datasets taken from BCI competition III-2005 and BCI competition IV-2008.

In the first dataset, Dataset IVa from BCI competition III [35] were used in this study. This dataset was recorded from five healthy subjects (“aa”, “al”, “av”, “aw”, “ay”). All subjects performed three types of motor imageries i.e. right foot, left hand and

right hand. However, for the competition, only right foot and right hand were provided. The recording was made using BrainAmp amplifiers with 118 EEG channels were measured based on extended international 10/20-system. Signals were band-pass filtered between 0.05 and 200 Hz and then digitized at 1000 Hz with 16 bit (0.1  $\mu$ V) accuracy. However, 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: subject (training; testing), “aa” (168; 112), “al” (224; 56), “av” (84; 196), “aw” (56; 224) and “ay” (28; 252). 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.

In second dataset, BCI Competition IV–Dataset 2a were used in this study. This dataset consists of EEG data from 9 subjects [36]. The cue-based BCI paradigm consisted of four different motor imagery tasks, namely the imagination of movement of the left hand (class 1), right hand (class 2), both feet (class 3), and tongue (class 4). Two sessions on different days were recorded for each subject. Each session is comprised of 6 runs separated by short breaks. One run consists of 48 trials (12 for each of the four possible classes), yielding a total of 288 trials per session. Twenty-two Ag/AgCl electrodes (with inter-electrode distances of 3.5 cm) were used to record the EEG. All signals were recorded monopolarly with the left mastoid serving as reference and the right mastoid as ground. The signals were sampled with 250 Hz and band-pass filtered between 0.5 Hz and 100 Hz. The sensitivity of the amplifier was set to 100  $\mu$ V. An additional 50 Hz notch filter was enabled to suppress line noise. In this study, only left hand (class 1), right hand (class 2) were utilized due to test the consistency and robustness of the proposed method related to the first dataset, which is 2-class motor imagery.

In this study, only 17 channels were selected in the first dataset (FC3, FC1, FCz, FC2, FC4, C5, C3, C1, Cz, C2, C4, C6, CP3, CP1, CPz, CP2, and CP4). Since these channels related to motor activity as used by [37] and according to the homunculus theory that represents motor activity area [38]. Meanwhile in second dataset, all channels (22 channels) were used.

## 2.2 Feature extraction and channel instantiation approach

Before feature extraction is conducted, EEG signal is filtered using 4-th order Butterworth band-pass filter as commonly used in EEG signal processing [39,40]. EEG signal were filtered in certain frequency range 8-30 Hz. This frequency range is most selected in prior research [41,42]. After filtered, according to nature of recorded data, a time slot between 0.5 – 3.5 seconds is chosen for further process for first dataset. The selected time slot, which is range about 3 seconds, consists of 300 data points because 100 Hz down-sampled data is used in this study. In second dataset, time slot between 3 – 5 seconds was selected, which is range about 2 seconds, consists of 500 data points because the sampling rate is 250 Hz. These data points (300 and 500 data points for first dataset and second dataset, respectively) then divided into 10 windows for further feature extraction process.

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

- Mean Absolute Value (mav):

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

- Root Means Square (rms):

$$rms = \sqrt{\frac{1}{n} \sum_1^N x_n^2} \quad (2)$$

- Standard Deviation ( $\sigma$ ):

$$\sigma = \sqrt{\frac{1}{n} \sum_1^N (x_n - \mu)^2} \quad (3)$$

- Skewness:

$$skewness = \sqrt{\frac{1}{n} \sum_1^N \frac{(x_n - \mu)^3}{\sigma^3}} \quad (4)$$

- Kurtosis:

$$kurtosis = \sqrt{\frac{1}{n} \sum_1^N \frac{(x_n - \mu)^4}{\sigma^4}} \quad (5)$$

- Variance to Mean Ratio (vmr):

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

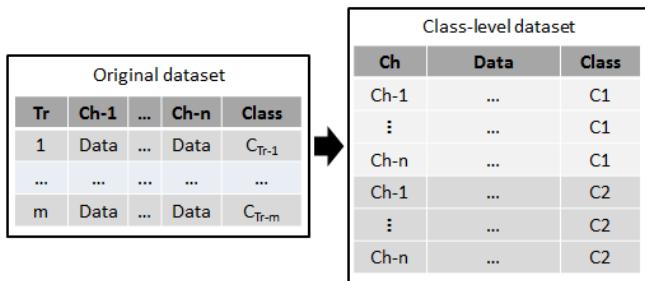


Figure. 1 Channel-class dataset is class-level dataset generation based on channel instantiation approach  
(Tr=Trial, Ch=Channel)

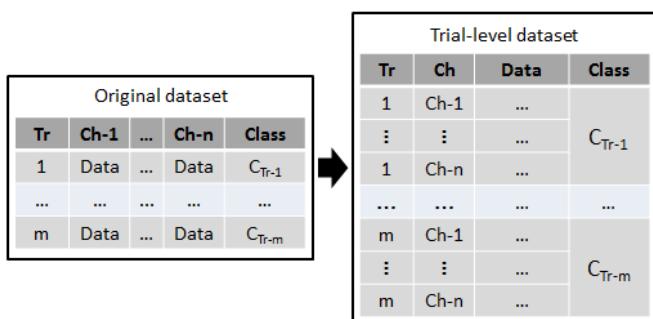


Figure. 2 Channel-trial dataset is trial-level dataset generation based on channel instantiation approach

- Coefficient of Variation (cv)

$$cv = \sigma/\mu \quad (7)$$

where  $x_n$  is n-th data point and  $\mu$  is mean value.

In this study, channel instantiation as similar to multi-instance learning (MIL) approach was used. In some applications, the objective is not to classify bags (a set of instances), but to classify individual instances [43]. MIL approach is compatible since EEG signal was recorded in multi channels, whereas each channel has its own signal. This approach transforms EEG raw signal and generate two kinds of new dataset i.e. Class-level dataset and Trial-level dataset. Class-level dataset was used to create feature selection pattern, meanwhile Trial-level dataset is used as dataset for classification task. In MIL approach, the classification task by using trial-level dataset is called instance-level prediction. Fig. 1 and Fig. 2 depict the class-level dataset and trial-level dataset generation, respectively.

Class-level dataset was used in feature selection scheme since it consists lower record hence faster in the execution time. Channel-Class dataset has lower number of record rather than Channel-Trial dataset with 236 (2 classes x 118 channels) and 4,760 (280 trials x 17 channels) for first dataset. In second dataset, channel-trial dataset has 2,508 records (114 trials x 22 channels).

The following steps for class-level dataset are:

- 1) Split data from each class
- 2) Create windowing: 10 windows (with same data points number for each window)
- 3) Calculate seven statistic measures from each window from each channel
- 4) Combine all feature vectors from each class become one dataset
- 5) Transform dataset so channel become record (instance) and statistic measures become column (attribute).

Meanwhile, trial-level dataset process little bit different where class in class-level dataset replace with trial. The following steps for trial-level dataset are:

- 1) Create windowing: 10 windows (with same data points number for each window)
- 2) Calculate seven statistic measures from each window from each trial
- 3) Combine all feature vectors from each trial become one dataset
- 4) Transform dataset so channel become record (instance) and statistic measures become column (attribute).

### 2.3 Feature selection

In this study, we use logistic regression (LR) as feature selection method by using LR ability in feature weighting as shown in Fig. 3. Two pre-processed dataset is used in this feature selection scheme, Channel-Class dataset and Channel-Trial dataset. Channel-Class dataset feed to LR as training dataset where LR weight its features then weight threshold is used to filter the dataset with features that pass the given threshold. In this study, weight threshold is set to  $\geq 0.5$  with absolute value which means a feature with weight value both positive and negative that surpass the threshold will be selected.

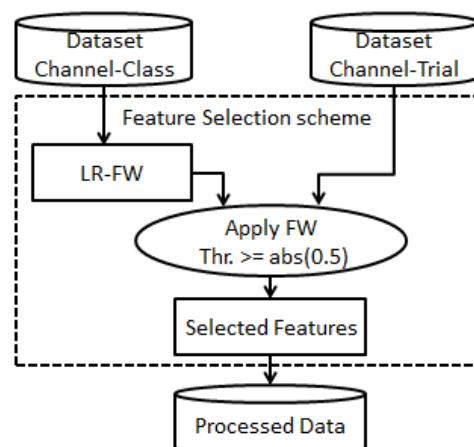


Figure. 3 Feature selection scheme

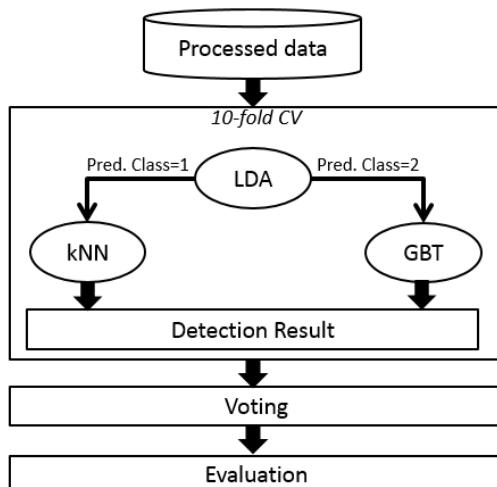


Figure. 4 Two stage detection and channel voting block diagram

This feature selection scheme then apply to Channel-Trial dataset for further processing so called processed data. The processed data consists of selected features from Channel-Trial dataset according to this feature selection scheme.

This feature selection scheme is apply to all five datasets (“aa”, “al”, “av”, “aw”, “ay”). So, different dataset may produce different selected features. In this study, all selected features from all five datasets then combined together and pick certain selected features with at least has 3 occurrence in all 5 datasets.

The feature that has at least 3 occurrence is selected because almost all dataset has agree about its selection and representative enough as selected feature. This second scheme is intended to improve inter-subject robustness of EEG based MI classification from subject-specific feature selection [44].

#### 2.4 Two-stage detection

Processed data from feature selection stage than feed to 10-fold cross validation (CV) scheme as shown in Fig. 4. 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. 4, three classifiers are used, i.e.: LDA, k-NN and GBT. In this study, k-parameter on k-NN is set to 3 and number of tree parameter on GBT is set to 600.

In first-stage detection, LDA is used since based on our empirical experiment, LDA is able to detect both classes with balance detection for both classes although in low accuracy. After detected by LDA, the second-stage detection is conducted where class 1 (Right Hand) detected data feed to k-NN meanwhile

class 2 (Right Foot) detected data feed to GBT. In 10-fold CV, the training stage will produce model that will used in testing stage. In testing stage, a testing dataset is feed to the model that built from the training stage. After 10-fold CV conducted, the detected data then we calculate the accuracy based on channel voting.

In this study, LDA was chosen since LDA generally provides good ability for classification [20] which find optimal discriminant features by maximizing the ratio of the between-class variance to the within-class variance of a given dataset [45]. Based on empirical result, LDA able to separates between classes equally and not tends to particular class. Meanwhile GBT was chosen, since GBT using boosting scheme and new decision tree model is trained to reduce the error [46]. Besides GBT, kNN was also chosen as second-stage classifier because kNN is a non-parametric learning algorithm method; thus, no need a priori assumptions about the distributions of data [47].

In this study, the two-stage detection is a form of hybrid classifier which also a form of ensemble learning. In ensemble, combining classifiers may provide a better choice and provide more acceptable approximations [25]. Thus, this ensemble technique helps derive stable classification results [48].

#### 2.5 Evaluation

In this study, accuracy is used as main evaluation since many prior researches used accuracy as main evaluation. In first dataset, 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 many prior researches. Meanwhile in second dataset, only testing dataset were used in evaluation by using transfer learning from first dataset (feature selection and modelling scheme).

Another important evaluation related to BCI systems is execution time. In BCI system, EEG based motor imagery classification evaluation regarding execution time has two main stage, i.e. feature extraction and classification stage. Several studies reported these two time measurements and then calculate the total time execution [26,28,49]. In addition, they achieved the total execution time below 1 second per trial, thus implementable for online BCI systems.

### 3. Experimental results

The experiments were 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, R programming was

Table 1. Classification accuracy and distribution of channel number detection grouped by trial on BCI competition III-Dataset IVa

#Range	Subject				
	aa	al	av	aw	ay
15 – 17	51	81	87	73	75
12 – 14	125	116	148	124	147
9 – 11	87	61	41	68	49
6 – 8	14	18	4	14	9
3 – 5	3	4	0	1	0
0 – 2	0	0	0	0	0
#correct	263	258	276	265	271
#wrong	17	22	4	15	9
Accuracy (%)*)	93.93	92.14	98.57	94.64	96.79

\*) Average Accuracy = 95.21% ± 2.51%

Table 2. Selected features from all subject with high occurrence on BCI competition III-Dataset IVa (14 of 70 Features)

No	Feature	#occurrence	Subject				
			aa	al	av	aw	ay
1	skewness1	4	v	v	v	v	-
2	skewness6	4	v	-	v	v	v
3	kurtosis10	3	v	-	v	v	-
4	kurtosis2	3	-	v	v	-	v
5	kurtosis3	3	-	v	-	v	v
6	kurtosis4	3	v	v	-	-	v
7	kurtosis7	3	v	-	v	-	v
8	kurtosis9	3	-	v	v	-	v
9	mav2	3	-	-	v	v	v
10	skewness10	3	-	v	v	-	v
11	skewness2	3	-	v	-	v	v
12	skewness3	3	-	-	v	v	v
13	skewness4	3	-	v	-	v	v
14	skewness8	3	v	v	v	-	-
#Selected Features each Subject		6	9	10	8	11	

used in feature extraction and evaluation; while Rapidminer 9.2 Educational License was used for modelling.

In this study, the proposed methods so called two-stage detection with feature selection based on feature weighting on LR algorithm (LRFS+TSD). Table 1 and Table 3 show the performance result from the proposed methods for BCI competition III-Dataset IVa and BCI competition IV-Dataset 2a,

respectively. As shown in Table 1, Range of channel with correct detection for each trial (#Range), number of correct detection (#correct) at trial level, number of wrong detection (#wrong) at trial level and accuracy is presented for each subject in every table. Since each trial consists of 17 channels, so true detection is made when #Range above 9. On the other hand, the detection is true for certain class when more than half of 17 channels, which mean at least 9 channels, are detected for its corresponding class.

As shown in Table 1, all subjects gained excellent accuracy. Number of correct detections mostly lies in range of 12-14 followed by range 15-17 and 9-11. Those range are majority detection that correct detection so the decision will be true detection for their respective trial. Subject “av” gained highest accuracy and subject “al” gained lowest accuracy. As a result, average accuracy for all subject yielded 95.21%. The result shows it’s consistent as indicated by low standard deviation value (2.51%).

The proposed method (LRFS+TSD) is performed using the same features for all subjects with selected 14 features. The selected features are used in this proposed method comes from a combination from all selected feature from all subjects based on the second proposed method with high occurrence (at least 3 occurrence). This feature selection approach able to reduce until 80% from existing features (14 of 70 features). Table 2 shows the detailed of the selected feature and its occurrence in all subjects.

Selected feature mostly comes from higher order statistic (skewness and kurtosis) that confirm several prior research about its effectiveness as features [4,32]. Skewness dominate the selected feature with 7 features followed by kurtosis and mav (mean absolute value) with 6 and 1 feature, respectively. As shown in Table 2, subject “av” and “ay” has higher occurrence in this selected feature with 10 and 11 selected features, respectively. This number of selected features affected the performance result of the proposed method where subject “av” and “ay” gained higher accuracy compared to subject “aa”, “al” and “aw”.

In order to evaluate the robustness of the proposed method regarding the consistency accuracy, another dataset comes from BCI competition IV-Dataset 2a were used. In this study, only left hand and right hand (2 class) were selected from this dataset. Selected features based on previous dataset were utilized and the same model, LRFS+TSD, was also investigated. Table 3 presents the performance of the proposed method.

As shown in Table 3, all nine subject (S1 until S3) gained excellent accuracy with average accuracy 94.83%. The performance of the proposed method is

Table 3. Classification accuracy and distribution of channel number detection grouped by trial on BCI competition IV-Dataset 2a

Range	Subject								
	S1	S2	S3	S4	S5	S6	S7	S8	S9
20 – 22	39	101	55	36	92	18	12	69	30
16 – 19	77	27	62	70	44	68	81	52	63
12 – 15	23	12	19	25	8	51	41	17	37
8 – 11	5	4	8	8	0	7	9	3	14
4 – 7	0	0	0	5	0	0	1	3	0
3 – 0	0	0	0	0	0	0	0	0	0
#correct	139	140	136	131	144	137	134	138	130
#wrong	5	4	8	13	0	7	10	6	14
Accuracy (%)*	96.53	97.22	94.44	90.97	100	95.14	93.06	95.83	90.28

\*) Average Accuracy = **94.83% ± 3.07%**

Table 4. Comparison to prior research for BCI competition III-Dataset IVa

Sub ject	M1 [16]	M2 [20]	M3 [21]	M4 [50]	M5 [22]	PM
aa	69.64	73.9	98.93	82.59	81.43	93.93
al	96.42	94.5	93.21	97.19	92.41	92.14
av	60.57	59.5	81.79	75.94	70	98.57
aw	70.53	80.7	93.21	98.2	83.57	94.64
ay	78.57	79.9	97.5	99	85	96.79
<b>avg.</b>	<b>75.15</b>	<b>77.70</b>	<b>92.93</b>	<b>90.58</b>	<b>82.48</b>	<b>95.21</b>
<b>sd.</b>	13.49	12.67	6.73	10.62	8.11	2.51

Note: M1=DSAA, M2=CSP+SGRM, M3=LRCFSP+SVM, M4=GQCSP+SVM, M5=CSPRMF+LDA, PM=LRFS+TSD

also consistent with low standard deviation among nine subjects (3.07%). Five subjects gained accuracy above 95% meanwhile the rest subject gained below 95% but still above 90%. This result confirms and validated the proposed method regarding its consistency and effectiveness.

To evaluate the effectiveness and competitiveness of the proposed method, a comparison between the proposed method and prior research were conducted and presented in Table 4. Five prior research were selected according to the similar dataset, similar the using of 10 fold cross validation and the newer publication time. Note that, font bold indicates highest accuracy among all methods and underlined font italic indicates lowest accuracy among each method.

As depicted in Table 4, the proposed method has highest accuracy in subject “av” compared to prior method. In addition, subject “av” is subject with lowest accuracy for each prior method. In matter of average accuracy, the proposed method has highest

average accuracy (95.21%) yes lowest standard deviation (2.51%) compared to prior research. All subject have high accuracy in the proposed method (above 90%). Contradict to prior research, at least one subject in prior research has accuracy below 90%. These findings indicate that the proposed method is effective and competitive and outperforms prior research. Another elaborative finding of the proposed method presented in boxplot and violin plot as depicted in Fig. 5. From this plot, the effectiveness and consistency of the proposed method depicted clearly.

As shown in Fig. 5, the proposed method has highest average accuracy and smaller range between minimum and maximum value compared to prior research on BCI competition III-Dataset IVa. Based on the violin shape, the proposed method has narrow range performance among all subjects; meanwhile in prior research, almost all of them have wide range performance. In BCI competition IV-Dataset 2a plot,

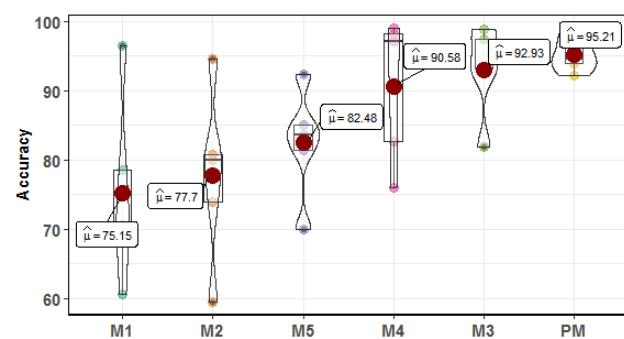


Figure. 5 Combination of boxplot and violin plot of prior research and the proposed method accuracy on BCI competition III-Dataset IVa (M1=DSAA, M2=CSP+SGRM, M3=LRCFSP+SVM, M4=GQCSP+SVM, M5=CSPRMF+LDA, PM=LRFS+TSD)

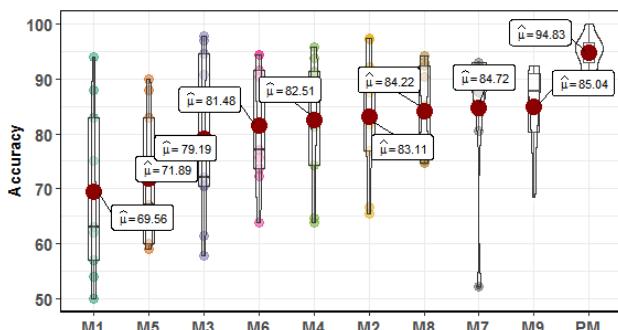


Figure. 6 Combination of boxplot and violin plot of prior research and the proposed method accuracy on BCI competition IV-Dataset 2a (M1=STR, M2=CCSP+SVM, M3=MEMDBF+CSP, M4=TSGSP+SVM, M5=ARFD+LDA, M6=CSE+UAEL, M7=DSAA, M8=SJGDA+KNN, M9=DBFCSP+CNN, PM=LRFS+TSD)

as shown in Fig. 6, the proposed method has similar result with its performance plot. The proposed method has highest average accuracy and narrow range performance compared to prior research. The detailed comparison between the proposed method and prior research for BCI competition IV-Dataset 2a is presented in Table 5.

As presented in Table 5, the proposed method has better performance compared to prior research in matter of higher accuracy and lower standard deviation, 94.83% and 3.07% respectively. The proposed method achieves 5 of 9 highest accuracy among all prior research. Based on these findings, the proposed method indicates consistence performance for both datasets. The accuracy and standard deviation are almost similar between two datasets.

Another comparison is made in this study. A multi-comparison is made by using Bonferroni-Dunn test. For this purpose, nine aggregate value based on statistical measures were composed from each method from its performance for all subject, such as: mean, median, range, standard deviation, 1<sup>st</sup> quartile, 3<sup>rd</sup> quartile, inter-quartile range, coefficient of variation and mean absolute deviation. Then, we calculate rank between methods for each measure (where lower rank means better than higher rank) and calculate the average rank for each method. This average ranks from all methods then used by Bonferroni-Dunn test. Fig. 7 and 8 present critical difference plot according to Bonferroni-Dunn test for BCI competition III-Dataset IVa and BCI competition IV-Dataset 2a, respectively. Bonferroni-Dunn test were selected since it more suitable when compare many methods with control (the proposed method as control).

As shown in Fig. 7, according to Bonferroni-Dunn test, critical difference (CD) value is 2.27. So,

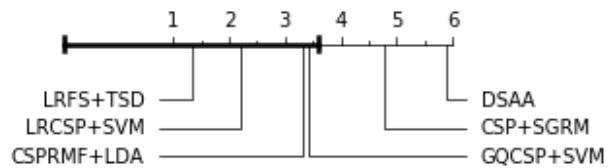


Figure. 7 Critical difference (CD) plot based on Bonferroni-Dunn test for prior research and the proposed method (LRFS+TSD) on BCI competition III-Dataset IVa (CD=2.27)

there will be difference between the proposed method and other method where the ranks are more than CD value. There is no difference between the proposed method with LRCSP+SVM, CSPRMF+LDA and GQCSP+SVM. Meanwhile with DSAA and CSP+SGRM, the proposed method has significant difference since there no black line covered these 2 methods. According to the ranks, the proposed method has lowest rank compared to other methods. This means, the proposed method better than other prior method although there is no significant different with LRCSP+SVM, CSPRMF+LDA and GQCSP+SVM.

In BCI competition IV-Dataset 2a performance, as shown in Fig. 8, the proposed method has lowest rank and no significant different with only three other methods, i.e. DFBCSP+CNN, SJGDA+KNN and DSAA. Meanwhile, compared to 6 other methods, there is significant difference with the proposed method, since their ranks higher than CD value (3.96). Again, from these two CD plots, the proposed method has lowest rank which mean better than prior research.

According to the CD plot as depicted in Fig. 7 and 8, the proposed method has first rank, which mean the performance of the proposed method is better than prior research. The proposed method also has significant difference with many prior researches, which means the proposed method is outperform many prior researches.

Beside accuracy, the proposed method was measured regarding its time execution. Table 6 presents time execution of the proposed method in matter of feature extraction and classification task.

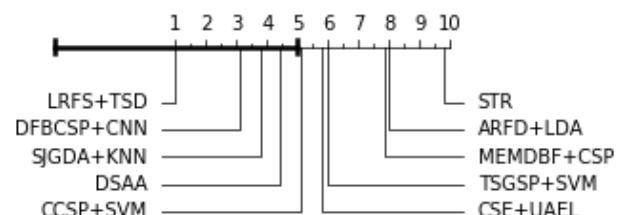


Figure. 8 Critical difference (CD) plot based on Bonferroni-Dunn test for prior research and the proposed method (LRFS+TSD) on BCI competition IV-Dataset 2a (CD=3.96)

Table 5. Comparison to prior research for BCI competition IV-Dataset 2a (left/right hand only)

<b>Method</b>	<b>Subject</b>									<b>Average</b>
	<b>S1</b>	<b>S2</b>	<b>S3</b>	<b>S4</b>	<b>S5</b>	<b>S6</b>	<b>S7</b>	<b>S8</b>	<b>S9</b>	
STR [28]	83	50	94	62	57	63	54	88	75	$69.56 \pm 15.93$
CCSP+SVM [23]	92.1	66.7	97.3	82.8	65.5	76.9	81.8	97.4	87.5	$83.11 \pm 11.86$
MEMDBF+CSP [53]	90.78	57.75	97.08	70.69	61.48	70.37	72.14	97.76	94.62	$79.19 \pm 15.85$
TSGSP+SVM [31]	87	64.7	93.8	74.3	90.4	63.9	91.4	95.8	81.3	$82.51 \pm 12.24$
ARFD+LDA [12]	74	59	83	60	60	66	67	88	90	$71.89 \pm 12.36$
CSE+UAEL [25]	91.67	63.89	94.44	72.22	77.08	75.69	73.61	94.44	90.28	$81.48 \pm 11.33$
DSAA [16]	88.88	80.55	93.05	52.09	87.5	90.27	92.36	85.41	92.36	$84.72 \pm 12.87$
SJGDA+kNN [27]	90.29	74.62	93.05	75.02	75.31	78.15	85.08	92.35	94.14	$84.22 \pm 8.46$
DFBCSP+CNN [26]	87.8	68.41	91.68	80.41	88.89	78.45	86.52	92.32	90.89	$85.04 \pm 7.88$
LRFS+TSD	96.53	97.22	94.44	90.97	100	95.14	93.06	95.83	90.28	$94.83 \pm 3.07$

Table 6. Execution time of the proposed method

<b>Task</b>	<b>Execution Time</b>
Feature Extraction	333.74 ms
Classification	2.5 ms
Total	336.25 ms (~ 0.34 s)

Time execution is important since EEG based motor imagery classification is part of BCI systems which mostly online system where low time execution is main consideration after its accuracy. Table 6 presents the execution time of the proposed method.

As shown in Table 6, the total time execution is about 0.34 second. This is still below 0.5 second. This achievement is indicated that the proposed method is implementable in BCI systems [16], especially in 2-class classification.

The effectiveness and efficiency of the results from the proposed method are generated by four components. These four components comprise of feature extraction scheme, channel instantiation approach, ensemble learning, and feature selection scheme. The first three components mostly affect the excellent and consistent accuracy which means the detection effectiveness. Meanwhile the last component affects the low execution time which means the detection efficiency.

In first component, feature extraction scheme plays a crucial role since effective feature extraction will lead effective classification. In this scheme, narrow window feature extraction approach was chosen to handle the nature of EEG signal which is non-stationary. A combination of first order statistic and higher order statistic work effectively. According to the most selected feature, three kind of statistic measures proved give discriminative and its importance in EEG based MI classification, i.e. mean absolute value (mav), skewness, and kurtosis (see

Table 2). Mean absolute value effectiveness as feature extraction consistent with previous study [4,51]. Similar to mean absolute value, skewness and kurtosis also confirm the findings of previous works which conducted by [4,32].

The second component is the using of channel instantiation approach [32]. By using channel instantiation, one single trial will be decompose into many instances which as many as number of channels. So, the detection will conduct in channel which act as instance complete with its own features as generated in feature extraction scheme. This detection, so called multi-instance learning (MIL) with instance-level prediction [43]. Since the detection in instance or channel level, so the final decision will be conducted based on voting scheme which conducted after two-stage detection.

In the third component, ensemble technique was employed in form of hybrid classifier so called to-stage detection. In first-stage, LDA was utilized. Several studies were utilized LDA and shown its effectiveness in classifying a 2-class EEG based MI classification [16,30]. In second-stage detection, GBT and kNN were used. At the decision stage, a voting scheme were implemented by using maximum probability for each detection. This voting technique is the most popular way of combining classifiers [16]. This ensemble technique helps derive consistent and better classification results [52].

Finally, the fourth component is feature selection scheme based on LR by utilizing LR feature weighting ability. From 70 available features, only 14 features (see Table 2) were selected by using this scheme. However, the performance of the detection still excellent and reduce execution time which make this study promising for BCI system. The combination of these four components explains the

underlying of the effectiveness of the proposed method.

#### 4. Conclusion

In this study, LR based feature selection and two step detection were used for classification task of MI based EEG signal. A multi-channel with taken from BCI competition III-Dataset IVa and BCI competition IV-Dataset 2a were used for this purpose. We conducted and evaluated the proposed method by using promising approach so called channel instantiation as data transformation, two-stage detection and simple yet effective feature selection scheme. From the experimental results, the proposed method gained excellent accuracy for BCI competition III-Dataset IVa and BCI competition IV-Dataset 2a with 95.21% and 94.83%, respectively. The proposed method gained consistent detection for both datasets with low standard deviation about 2.51% and 3.07%, respectively. According to the Bonferroni-Dunn test, the proposed method proved its competitiveness and consistency with gained first rank for both datasets. The feature selection scheme able to reduce 80% from available features and able to maintain consistent excellent accuracy. These results show that the proposed method has the potential to obtain a reliable EEG based MI classification and can be used practically in controlling a wheelchair, rehabilitation therapies for the stroke rehabilitation or improve motor rehabilitation outcomes. However, several issue still open for future works such as: channel reduction as a feature selection according the original dataset and apply model in multi-class EEG based MI classification to test the robustness of the proposed method.

#### Conflicts of Interest

The authors declare no conflict of interest.

#### Author Contributions

Conceptualization, AW, TBA, and NAS; methodology, AW, TBA, and NAS; software, AW; validation, AW, TBA, and NAS; formal analysis, AW and NAS; investigation, AW and NAS; resources, AW; data curation, AW; writing—original draft preparation, AW; writing—review and editing, TBA and NAS; visualization, AW; supervision, TBA and NAS; project administration, TBA and NAS; funding acquisition, AW and NAS.

#### Acknowledgments

The authors would like to thanks to the Ministry of Education and Culture of the Republic of Indonesia (formerly the Ministry of Research, Technology, and Higher Education) for the BPPDN scholarship while doing Doctoral education at Universitas Gadjah Mada, Indonesia. The Authors would like to thanks to Universitas Gadjah Mada for Final Project Recognition 2020 Grant.

#### References

- [1] X. Song and S. C. Yoon, “Improving brain-computer interface classification using adaptive common spatial patterns”, *Comput. Biol. Med.*, Vol. 61, pp. 150–160, 2015, doi: 10.1016/j.combiomed.2015.03.023.
- [2] 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, doi: 10.1016/j.bspc.2016.10.015.
- [3] C.-I. Hung, P.-L. Lee, Y.-T. Wu, L.-F. Chen, T.-C. Yeh, and J.-C. Hsieh, “Recognition of Motor Imagery Electroencephalography Using Independent Component Analysis and Machine Classifiers”, *Ann. Biomed. Eng.*, Vol. 33, No. 8, pp. 1053–1070, 2005, doi: <https://doi.org/10.1007/s10439-005-5772-1>.
- [4] 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, doi: 10.1016/j.bspc.2016.09.007.
- [5] Y. Yu, Z. Zhou, E. Yin, J. Jiang, J. Tang, Y. Liu, and D. Hu., “Toward brain-actuated car applications: Self-paced control with a motor imagery-based brain-computer interface”, *Comput. Biol. Med.*, Vol. 77, pp. 148–155, 2016, doi: 10.1016/j.combiomed.2016.08.010.
- [6] 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, doi: 10.1016/j.eswa.2014.02.043.
- [7] 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, p. 478–489.e1, 2016, doi: 10.1016/j.apmr.2015.07.012.
- [8] T. Zhang, T. Liu, F. Li, M. Li, D. Liu, R. Zhang, H. He, P. Li, J. Gong, C. Luo, D. Yao, and P. Xu,

- “Structural and functional correlates of motor imagery BCI performance: Insights from the patterns of Fronto-Parietal Attention Network”, *Neuroimage*, 2016, doi: 10.1016/j.neuroimage.2016.04.030.
- [9] H. Burianová, L. Marsteller, P. Sowman, G. Tesan, A. N. Rich, M. Williams, G. Savage, and B. W. Johnson, “Multimodal functional imaging of motor imagery using a novel paradigm”, *Neuroimage*, Vol. 71, pp. 50–58, 2013, doi: 10.1016/j.neuroimage.2013.01.001.
- [10] 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, doi: 10.1016/j.neucom.2015.05.133.
- [11] 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, doi: 10.1049/iet-smt.2016.0443.
- [12] R. Salazar-Varas and R. A. Vazquez, “Evaluating the effect of the cutoff frequencies during the pre-processing stage of motor imagery EEG signals classification”, *Biomed. Signal Process. Control*, Vol. 54, p. 101592, 2019, doi: 10.1016/j.bspc.2019.101592.
- [13] 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.
- [14] 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, 2010, doi: 10.1016/j.patrec.2010.04.009.
- [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, doi: 10.1016/j.bspc.2017.11.014.
- [16] K. Belwafi, S. Gannouni, H. Aboalsamh, H. Mathkour, and A. Belghith, “A dynamic and self-adaptive classification algorithm for motor imagery EEG signals”, *J. Neurosci. Methods*, Vol. 327, 2019, doi: 10.1016/j.jneumeth.2019.108346.
- [17] 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, doi: 10.1016/j.procs.2016.07.422.
- [18] 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, doi: 10.1049/iet-smt.2015.0200.
- [19] J. S. Kirar and R. K. Agrawal, “A combination of spectral graph theory and quantum genetic algorithm to find relevant set of electrodes for motor imagery classification”, *Appl. Soft Comput. J.*, no. xxxx, p. 105519, 2019, doi: 10.1016/j.asoc.2019.105519.
- [20] Y. Jiao, Y. Zhang, X. Chen, E. Yin, J. Jin, X. Wang, and A. Cichocki, “Sparse Group Representation Model for Motor Imagery EEG Classification”, *IEEE J. Biomed. Heal. Informatics*, Vol. 23, No. 2, pp. 631–641, 2019, doi: 10.1109/JBHI.2018.2832538.
- [21] Y. Park and W. Chung, “Frequency-Optimized Local Region Common Spatial Pattern Approach for Motor Imagery Classification”, *IEEE Trans. Neural Syst. Rehabil. Eng.*, Vol. 27, No. 7, pp. 1378–1388, 2019, doi: 10.1109/TNSRE.2019.2922713.
- [22] J. K. Feng, J. Jin, I. Daly, J. Zhou, Y. Niu, X. Wang, and A. Cichocki, “An Optimized Channel Selection Method Based on Multifrequency CSP-Rank for Motor Imagery-Based BCI System”, *Comput. Intell. Neurosci.*, Vol. 2019, 2019, doi: 10.1155/2019/8068357.
- [23] L. Li, G. Xu, J. Xie, and M. Li, “Classification of single-trial motor imagery EEG by complexity regularization”, *Neural Comput. Appl.*, vol. 31, no. 6, pp. 1959–1965, 2019, doi: 10.1007/s00521-017-3174-6.
- [24] Y. Zhang, C. S. Nam, G. Zhou, J. Jin, X. Wang, and A. Cichocki, “Temporally constrained sparse group spatial patterns for motor imagery BCI”, *IEEE Trans. Cybern.*, Vol. 49, No. 9, pp. 3322–3332, 2019, doi: 10.1109/TCYB.2018.2841847.
- [25] H. Raza, D. Rathee, S. M. Zhou, H. Cecotti, and G. Prasad, “Covariate shift estimation based adaptive ensemble learning for handling non-stationarity in motor imagery related EEG-based brain-computer interface”, *Neurocomputing*, Vol. 343, pp. 154–166, 2019, doi: 10.1016/j.neucom.2018.04.087.
- [26] B. E. Olivas-Padilla and M. I. Chacon-Murguia,

- “Classification of multiple motor imagery using deep convolutional neural networks and spatial filters”, *Appl. Soft Comput. J.*, Vol. 75, pp. 461–472, 2019, doi: 10.1016/j.asoc.2018.11.031.
- [27] S. Guan, K. Zhao, and S. Yang, “Motor Imagery EEG Classification Based on Decision Tree Framework and Riemannian Geometry”, *Comput. Intell. Neurosci.*, Vol. 2019, pp. 1–13, 2019, doi: 10.1155/2019/5627156.
- [28] P. G. Rodrigues, C. A. S. Filho, R. Attux, G. Castellano, and D. C. Soriano, “Space-time recurrences for functional connectivity evaluation and feature extraction in motor imagery brain-computer interfaces”, *Med. Biol. Eng. Comput.*, pp. 1709–1725, 2019, doi: 10.1007/s11517-019-01989-w.
- [29] P. Gaur, R. B. Pachori, H. Wang, and G. Prasad, “An Automatic Subject Specific Intrinsic Mode Function Selection for Enhancing Two-Class EEG-Based Motor Imagery-Brain Computer Interface”, *IEEE Sens. J.*, Vol. 19, No. 16, pp. 6938–6947, 2019, doi: 10.1109/JSEN.2019.2912790.
- [30] R. Salazar-Varas and R. A. Vazquez, “Evaluating the effect of the cutoff frequencies during the pre-processing stage of motor imagery EEG signals classification”, *Biomed. Signal Process. Control*, Vol. 54, p. 101592, 2019, doi: 10.1016/j.bspc.2019.101592.
- [31] Y. Zhang, C. S. Nam, G. Zhou, J. Jin, X. Wang, and A. Cichocki, “Temporally constrained sparse group spatial patterns for motor imagery BCI”, *IEEE Trans. Cybern.*, Vol. 49, No. 9, pp. 3322–3332, 2019, doi: 10.1109/TCYB.2018.2841847.
- [32] 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, doi: 10.1016/j.measurement.2016.02.059.
- [33] 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, doi: 10.1016/j.eswa.2017.07.033.
- [34] D. Liang, C.-F. Tsai, and H.-T. Wu, “The effect of feature selection on financial distress prediction”, *Knowledge-Based Syst.*, Vol. 73, pp. 289–297, 2015, doi: 10.1016/j.knosys.2014.10.010.
- [35] B. Blankertz, K. R. Müller, D. J. Krusienski, G. Schalk, J. R. Wolpaw, A. Schlögl, G. Pfurtscheller, J. D. R. Millán, M. Schröder, and N. Birbaumer, “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, doi: 10.1109/TNSRE.2006.875642.
- [36] C. Brunner, R. Leeb, G. R. Muller-Putz, A. Schlogl, and G. Pfurtscheller, “BCI Competition 2008 – Graz data set A”, 2008.
- [37] 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.
- [38] 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, doi: 10.1016/j.dcn.2015.08.004.
- [39] 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, doi: 10.1016/j.bspc.2017.10.012.
- [40] 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, doi: 10.1162/NECO\_a\_00963.
- [41] 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, doi: 10.1016/S1005-8885(17)60215-2.
- [42] 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. 2016, pp. 1–10, 2017, doi: 10.1016/j.neucom.2016.12.039.
- [43] M. A. Carboneau, V. Cheplygina, E. Granger, and G. Gagnon, “Multiple instance learning: A survey of problem characteristics and applications”, *Pattern Recognit.*, Vol. 77, pp. 329–353, 2018, doi: 10.1016/j.patcog.2017.10.009.
- [44] Y. Yang, S. Chevallier, J. Wiart, and I. Bloch, “Subject-specific time-frequency selection for multi-class motor imagery-based BCIs using few Laplacian EEG channels”, *Biomed. Signal Process. Control*, Vol. 38, pp. 302–311, 2017, doi: 10.1016/j.bspc.2017.06.016.
- [45] R. Fu, Y. Tian, T. Bao, Z. Meng, and P. Shi, “Improvement Motor Imagery EEG Classification Based on Regularized Linear

- Discriminant Analysis”, *J. Med. Syst.*, Vol. 43, No. 6, 2019, doi: 10.1007/s10916-019-1270-0.
- [46] T. Yang, W. Chen, and G. Cao, “Automated classification of neonatal amplitude-integrated EEG based on gradient boosting method”, *Biomed. Signal Process. Control*, Vol. 28, pp. 50–57, 2016, doi: 10.1016/j.bspc.2016.04.004.
- [47] S. K. Bashar and M. I. H. Bhuiyan, “Classification of motor imagery movements using multivariate empirical mode decomposition and short time Fourier transform based hybrid method”, *Eng. Sci. Technol. an Int. J.*, Vol. 19, No. 3, pp. 1457–1464, 2016, doi: 10.1016/j.jestch.2016.04.009.
- [48] 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. 26, No. 2, pp. 498–505, 2018, doi: 10.1109/TNSRE.2017.2757519.
- [49] S. Razi, M. R. Karami Mollaei, and J. Ghasemi, “A novel method for classification of BCI multi-class motor imagery task based on Dempster–Shafer theory”, *Inf. Sci. (Ny)*, Vol. 484, pp. 14–26, 2019, doi: 10.1016/j.ins.2019.01.053.
- [50] J. S. Kirar and R. K. Agrawal, “A combination of spectral graph theory and quantum genetic algorithm to find relevant set of electrodes for motor imagery classification”, *Appl. Soft Comput.*, p. 105519, 2019, doi: 10.1016/j.asoc.2019.105519.
- [51] P. K. Pattnaik and J. Sarraf, “Brain Computer Interface issues on hand movement”, *J. King Saud Univ. - Comput. Inf. Sci.*, Vol. 30, No. 1, pp. 18–24, 2018, doi: 10.1016/j.jksuci.2016.09.006.
- [52] 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. 26, No. 2, pp. 498–505, 2018, doi: 10.1109/TNSRE.2017.2757519.
- [53] P. Gaur, R. B. Pachori, H. Wang, and G. Prasad, “An Automatic Subject Specific Intrinsic Mode Function Selection for Enhancing Two-Class EEG-Based Motor Imagery-Brain Computer Interface”, *IEEE Sens. J.*, Vol. 19, No. 16, pp. 6938–6947, 2019, doi: 10.1109/jsen.2019.2912790.