International Journal of Engineering & Technology, 7 (2.6) (2018) 163-167

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research Paper

Multiclass classification of motor imagery EEG signals using ensemble classifiers & cross-correlation

D.Hari Krishna [1]*, I.A. Pasha [2], T.Satya Savithri [3]

[1,2]ECE Dept,BVRIT, Narsapur, TS, India
[3] ECE Dept, JNTUCEH, Hyderabad, TS, India
**Corresponding author E-mail: hari_012@yahoo.com*

Abstract

To communicate without any muscle movement and purely based on brain signal has been the goal of Brain computer interfacing (BCI). Recent BCI based studies reported more and more accurate detection of brain states. This paper proposes a study that detects EEG signal belonging todifferent imaginary motor activities (Right leg, right hand, left leg and left hand). The Electroencephalogram (EEG) signal has been conditioned by band pass filter (BPF) to improve signal to noise ratio (SNR). The proposed method is based on similarity between signals to extract features. For measuring the similarity between signals, Cross correlation (CC) is used. An ensemble set of five classifiers (Linear Discriminant Analysis (LDA), K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Naïve Bayes (NB) and Binary Decision Tree) was used collectively. As the similarity measurement was binary in nature, one versus rest (OVR) approach was used for multi class classification. Random subset of features was used to train the ensemble of classifiers. The classification label was obtained by using majority voting. An average accuracy of 89.57% was observed among all 10 subjects.

Keywords: *Brain computer interfacing (BCI); Cross correlation; EEG classification; Ensemble; Motor Imagery (MI)*

1. Introduction

Motor imagery (MI) or imagery motor activity is a type of mental task used in the brain computer interfacing (BCI) system where in EEG activity will be different for activities before and during the imagery movement[1]. It is one of the major fundamentals ofmental tasks, which is based on the fact that the imagination of the movement of any of the body part has similar variations of the sensorimotor rhythms of the brain than the movement comprehension itself. This phenomenon iscalled as event-related synchronization (ERS) and event-related desynchronization (ERD). In general, this kind of task is attained only from two electrodes situated over the brain motor cortex area, named as C3 and C4. BCI research based on motor imagery activities mainly focuses on the discrimination of given imaginary tasks. (Example: movements of either left hand, right hand, any finger and/or foot). Researchers are continuously developing algorithms to improve the performance of classification. Classifiers such as artificial neural network (ANN), K nearest neighbor (KNN), Linear discriminant analysis (LDA) and Support vector machine (SVM)have already achieved superior performance [1]–[3]. Feature extraction technique which is used inclassifying motor imagery also plays major role inimproving the accuracy [1]–[5]. A combination of superior features with a pertinent classification scheme can enhance the performance.Researchers are able to successfully decode the gestures of tongue imagery, foot, left and right hands [6] into commands to control the movement of a humanoid [7], mobile robots [8,9] and wheelchairs [10,11] for persons with disability. Electroencephalography (EEG) is the most regularly used technique by the BCIresearch community to measure the brain activity as it is non-invasive, ease of accessibility, easy-to-use, good quality temporal resolution and portable.[12, 13]. Any BCI technology mainly consists of three stages: i) pre-processing of signals ii) extractingthe features and finally iii) classification[14].For any BCI device, the prior condition for getting superior performance is based on recognition rate obtained from extracted features of the classifiers.EEG is a complex signal with non Gaussian and non stationary properties [12]. It entails the use of various method like frequency, time, time-frequency anddiverse algorithms for processing of non linear signals using wavelet transforms [15], estimation of band powers [16] hjorth parameters [17], adaptive autoregressive parameters [18], and approximate entropy [19], to extort significant features from the original raw data. These featuresare input to the classifiers, which will produce an output signalthat corresponds to the brain states.Generally, classifiers like SVM, nave Bayesian (NB), LDA, kNN [20] and neural networks (NN) are the popular for two class classifications as they are good in yielding high-quality recognition accuracy, but their performancefor multi-class classification is not adequate. In order to get improved results using multiclass classification, many techniques are used. Few of them areone-against-one (OAO), one-against-all (OAA), and error correction code (ECC) [21]. These techniques when combined with above mentioned classifiers are able to produce better results.In spite of many efforts in the search of good number of passable methods like OAO, OAA and ECC combined with many different classifiers, it is still not viable to come to a conclusion that any of the algorithm or combination of various algorithms is better for the given type of biological data, as the accomplished conclusions are limited to only one dataset or few data sets. However, for a BCI system that is real in nature, there is notaconsent about an algorithm that coalescesless processing cost, high performance and robustness with dominance over the others.Generally, in a multi-class problem due to increase in the number of classes, the number of training sets (of high dimension-

Copyright © 2018 Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ality) turns out to be relatively smaller. The classifiers trained on small number of training sets turn out to be biased and has huge difference due to inadequateassessment of related parameters.Hence this type of classifiers is termed as '*weak*' classifiers. Many such 'weak' classifiers are combined by Ensemble classifier to furnish the result [22]. *Boosting [23], Bagging [24] and Random Subspaces [25] are few of the frequently used ensemble schemes.*Ensembles methods are nothing but grouping of individual classifiers that carry out the same task with the goal of engendering a better model. As affirmed in [26] and [27], fusion of the classifiers has shown potential results when they are applied to BCI data. Additionally, they are capable enough in providing apromising solution to the "curse of dimensionality" and at the same time they are dynamic in nature to the variance of EEG data in the time domain as well asfor inter-subject (dataset), because of this reason only they have become more attractive for BCI system applications.There are two reasons behind how ensemble methods are being useful/suitable for EEG classification. 1. Since the primary pre-requisite for any BCI system is training the classifier at the earliest or as early as possible, to fulfill this requirement, training sets must be small. High dimensionality of EEG signal is an additional benefit in this aspect. 2. In case of time varying signal like EEG, it is not at all advisable to use a classifier which isjust trained once to recognize the class with the help of unknown features [22].Ensemble studies are not popular in BCI research despite the above mentioned advantages and only very few studies exists on this.In general, an ensemble of classifiers can be defined as"a group of classifiers that mingles in some way their individual results in order to classify new instances" [28]. In order to create a strong ensemble model, there are two factors of critical importance: the diversity and accuracy of the ensemble components [29]. A classifier is accurate if it's performance over the random guessing and the miscellany is assured if the components of the ensemble produce different outcomes to all the new instances.

According to the literature, there are three main categories of classifier fusion procedures [30], based on three aspects of the ensemble modeling. The first aspect is related to combination of classifiers and their type, being separated into classifier assortment or classifier fusion. In classifier assortment, every section is trained in portions of the feature space, the objective being an expert in the models of that easy task. Whereas in caseof classifier fusion, everymodule in the ensemble is trained on the complete feature space. Second aspect engrosses the decision whether the ensemble from fusion method can be trained or not,depending on the inevitability of extra training for the ensemble fusion subsequent to training every classifier. Statistics such as maximum, minimum, mean, median, or product are used as non-trainable methods for fusion operator. In case, if it is trainable, the combination rules are applied by the algorithms as stacked simplification or weighted average, which establishes the parameters during an extra learning progression. Third aspect,after taking into count how the occurrences cooperate with the conformation of the ensemble, separates ensembles as either static or dynamic. The ones that make use of the full set of classifiers for every instanceare static ensembles. The dynamic ensembles are converse; they dynamically select classifiers for each instance.

2. Experiment

2.1.Data Collection

The EEG data was collectedwith 24-channeld EEG apparatusmanufactured by NCC medical corp., China. The 24 EEG channels were prearranged in 10-20 system in the region of the scalp. The EEG signals were amplified and sampled at 128 Hz. A notch filter at 50 Hz was applied for line noise removal. The original data was piled up in system native in .NED format which were latertransformed to a .txt file for more processing using software supplied by the NCC manufacturer.

2.2. Experimental Setup

The subject is seated in a comfortable chair wearing an EEG cap as shown in the figure 1. We have total ten subjects; normal healthy young adults (mean age 21 years). The subject is shown a stimuli on the LCD monitor placed approx. 2 feet in front of him/her. The stimulus consists of images of hands and legs which appear in four boxes. The imagesof hands will appear in top row of boxes and images of legs willappear in the bottom row of the boxes.



Fig.1: Experimental Setup

Only one image will appear at a time. When an image of hand appear in the left box (Top row), the subject has to imagine the movement of left and when an image of leg appears on the left box (Bottom row), the subject has to imagine the movement of left leg. The subject has to imagine movement of hand and leg of right side when stimuli appear on the right side of boxes. The figure.2 shows the stimuli shown to the subjects.

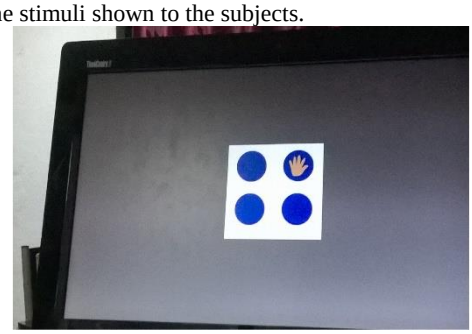


Fig.2: Stimulus

Each of this imagery activity is called a trial and there are total 50 trials for each of the hand and leg movement making total of 200 trials. Each trial is of the duration of 3 seconds followedby means of 3 seconds of rest or relaxation period. The experiment was designedusingPsychtoolbox in Matlab. Consequent data processing was done using Matlab used for all.

3. Methodology

The process of identifying which imagery motor activity the EEG signal belongs to can be separated into three parts; pre-processing, feature extraction and classification. The raw EEG signals are filtered and segmented by their respective mental motor activity in the pre-processing part. The feature extraction part calculates the different parameters of the signal through which the signals are classified into the mental motor activities in the classification part. To get better estimation of accuracy of the classification model, K-Fold cross validation technique is employed in this study.

3.1. Pre-processing

Band pass Butterworth filter in the frequency range 8-30 Hz is used to filter the acquired 24-channel data. The EEG signals containing imagerymotor signals are in α and β bands which ranged

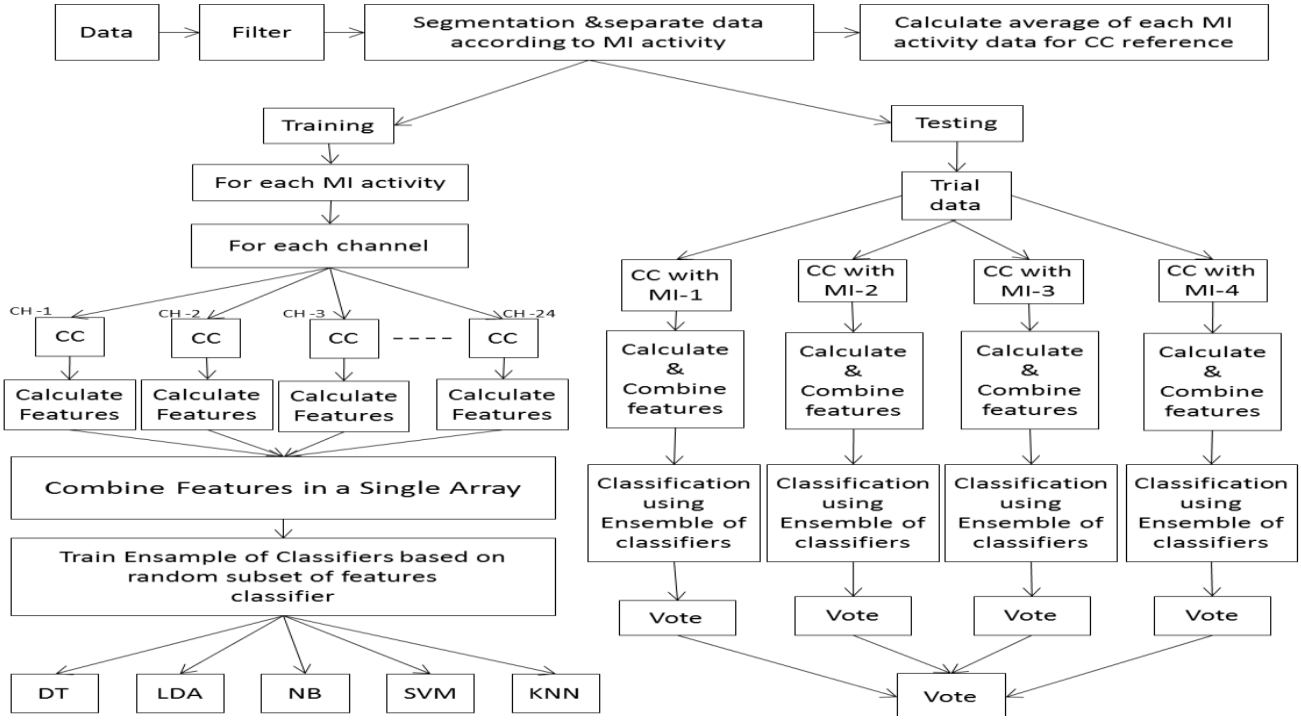


Fig. 3: EEG Data Training and Testing Process

from 8-12 Hz and 12-28Hz respectively. A moving average filter is also used to reduce the artifacts effects like eye movement/ blinks and movements of muscles. The EEG data is segmented based on the imagery motor activity based on the timestamps measured during data collection.

3.2. Feature extraction

The earlier studies reported that while EEG signals from diverse places from the scalp does not contain same amount of information, there are interrelated and in coherence with each other[31,32]. So, the EEG signals collected while performing same types of activity will have more similarity then the signals produced fromdifferent motor imagery activities. This difference in similarity can be detected and used to differentiate EEG signals belongs to different motor activities. The cross-correlation can be employed as a method of similarity measure among signals using the formula

$$R_{xy}[m] = \sum_{k=0}^{N-|m|-1} x[k]y[k-m];$$
$$where\ m = -(N-1), -(N-2), \ldots 0,1,2 \ldots (N-2), (N-1) \quad (1)$$

Rxy [m] is the cross-correlated sequence at m lag. If each of the signals, x and y, consists of M finite number of samples, the resultant cross-correlation sequence has 2M-1 samples. The features can be extracted from the result signal. To calculate the features, it was decided to use statistical features through which it can be found whether the two cross-correlated signals are similar or not. The five numbers summery was reported to be providing better estimation of the data distribution. The parameters proposed in five number summery are first quartile (Q1), smallest (Minimum) observation, median (Q2), 3rd quartile, and the largest (Maximum) observation. In addition to these features we also included five more features which are Mean, Standard Deviation, Mode, Inter-quartile range and RMS value making it total ten features per signal.

3.3. Classification

The features used to classify imagery motor activity are based on similarity between two signals and hence this method of feature extraction is binary in nature. In order to classify between four types of imagery motor activity, ‘one verses rest’ approach is used which reduces the multiclass classification into sets of binary class classifications. The final classification label will be based on results of all the classifiers. This study focus on ensemble of classifier trained on randomly selected subset of features. The final classification label will be based on majority voting and one verses rest voting. To calculate cross-correlation signal, a reference signal is required, to determine the reference signal, a median signal of all the trial is calculated for all the channels, generating twenty-four signals to be used as reference for the particular channels. The reference signals are calculated for all the four types of imagery motor activities. As we compute cross-correlated signals for each channel, ten features are generated for each channel, making it total 240 features for each trial. These features are reshaped into a signal array and a subset of it will be selected randomly to train classifiers. An ensemble of 100 classifiers of each of these five classifiers, LDA, SVM, KNN,Binary Decision Tree and NBwere trained. These total 500 classifiers were trained for each of the four motor imagery activities. To decide the accuracy of the trained classifier model, a set of trial data which was kept separately for the testing purpose was used. As per one verses rest model, the trial data has to be processes into four parallel streams, each stream will have features calculated by cross-correlated using reference belonging to the respective motor imagery activity. In each of the stream, the calculated features are classifies using 500 trained classifiers and final classification label is calculated using majority voting. The final class labelwill be based on ‘one verses rest’ approach.

4. Results

K-Fold cross validation was used to determine the classification accuracy of the classifier model with K = 10. The 3 depicts the

classification accuracies for ten subjects are depicted in the figure 4. The performance of the system was observed to be consistent overthe several subjects. The classification accuracies were also calculated for each of the classifiers separately as well. The individual performance of LDA was observed to be best at 85.84 ± sion matrix can be used for better estimation and analysis of the classification performance of the model. The finer details of the classification results such as correct classifications and misclassifications as well as which classes they are misclassified as can be observed from the confusion matrix. Table 1 depicts the classification performance of one of the subject. In the confusion matrix, the actual class will be represented in rows while the classified class will be represented in columns. For example, in table 1, the second row of the column titled "No Activity" depicts that the data belongs to Left Hand activity has been classified as No Activity for total 4 times by classification algorithm.

Table 1: Confusion Matrix

	No Activity	Left Hand	Right Hand	Left Leg	Right Leg	Total
No Activity	0	0	0	0	0	0
Left Hand	4	45	0	0	1	50
Right Hand	6	2	42	0	0	50
Left Leg	5	0	1	44	0	50
Right Leg	4	1	0	0	45	50

While manymisclassifications can be seen from the confusion matrix, majority of them falls into the category of "No Activity". This is due to the tie breaker rule which classifies the trial as belonging to "No Activity" whenever there is tie between two classes. The authors believed that not identifying the activity is better than incorrect identification which is the main reason behind the tie-breaker rule.

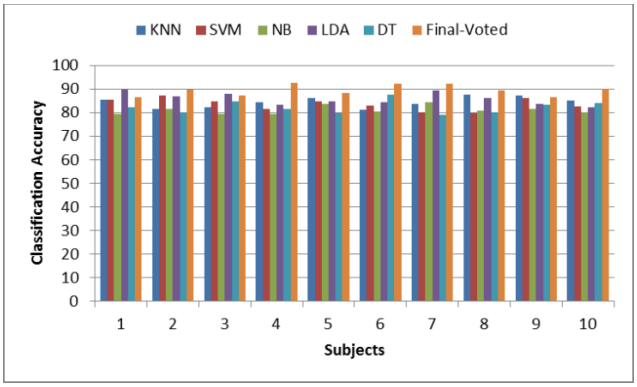


Fig. 4: Classification Accuracies

The Figure 5 depicts the comparison of classification accuracies from some of the recent studies [33 - 35].

2.43 while NB performed the least with average classification accuracy of 81.09± 1.64. However, the combined results from all five sets of ensemble classifiers were observed to be higher with the average classification accuracy was 89.57± 2.20%.The confu

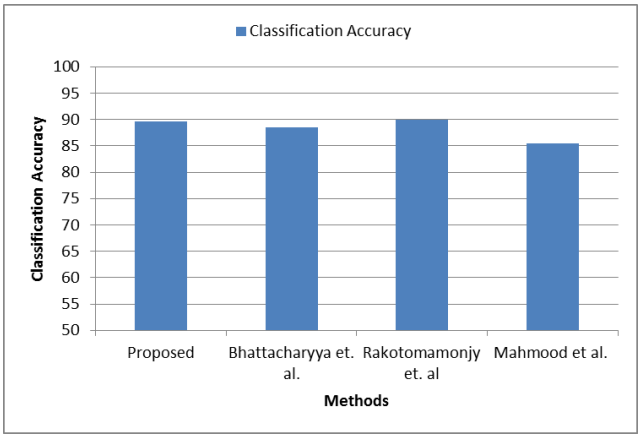


Fig. 5: Comparison of Classification Accuracies

5. Conclusion

The ensemble classifier based on five classifiers with total 500 classifiers was implemented in this study. The multiple types of classifiers performed at an average classification accuracy of 89.57 %. The comparison with recent studies demonstrates that the model performs at least as well as the other multiclass classification models. The observations from confusion matrix show that multiple classifications in the category of "No activity" which is most likely as a result of tie between two classes. The future scope can be to implement feature selection or boosting methods to improve classification accuracies.

References

[1] S. Sanei and J.A. Chambers, EEG Signal Processing, Wiley, 2007

[2] S. Bhattacharyya, A. Khasnobish, A. Konar, D.N. Tibarewala, A.K. Nagar, Performance analysis of left/right hand movement classification from EEG signal by intelligent algorithms, Computational Intelligence,Cognitive Algorithms, Mind, and Brain (CCMB), 2011 IEEE Symposium on , vol., no., pp.1,8, 11-15 April 2011

[3] B. Blankertz, K. R. Muller, G. Curio, T. M. Vaughan, G. Schalk, J. R. Wolpaw, A. Schlogl, C. Neuper, G. Pfurtscheller, T. Hinterberger, M.Schroder, and N. Birbaumer, The BCI competition 2003: progress and perspectives in detection and discrimination of EEG single trials, IEEE Transactions on Biomedical Engineering, vol. 51, pp. 1044-1051, 2004.

[4] B. Xu, A. Song, J. Wu, Algorithm of Imagined Left-Right Hand Movement Classification Based on Wavelet Transform and AR Parameter Mode, Bioinformatics and Biomedical Engineering, 2007. ICBBE 2007. The 1st International Conference on , pp.539,542, 6-8 July 2007

[5] L. Ming-Ai, W. Rui, H. Dong-Mei, Y. Jin-Fu, Feature Extraction and Classification of Mental EEG for Motor Imagery, Natural Computation, 2009. ICNC '09. Fifth International Conference on , vol.2, no.,pp.139,143, 14-16 Aug. 2009

[6] C. Brunner, R. Scherer, B. Graimann, G. Supp, and G. Pfurtscheller, "Online control of a brain-computer interface using phase synchronization,"*IEEE Trans. Biomed. Engg.*, vol. 53, no. 12, pp. 2501-06, 2006

[7] Y. Chae, J. Jeong, and S. Jo, "Toward brain-actuated humanoid robots: Asynchronous direct control using an EEG-based BCI," *IEEE Trans. Robotics*, vol. 28, no. 5, pp. 1131-1144, 2012

[8] S. Bhattacharyya, A. Sengupta, T. Chakraborti, D. Banerjee, A. Khasnobish, A. Konar, D.N. Tibarewala, R. Janarthanan, "EEG controlled remote robotic system from motor imagery classification," In: *3rd Int.conf on Com., Comm. & Networking Tech. (ICCCNT) 2012*, pp. 1-8, 26-28 July 2012

[9] D. H. Krishna, I. A. Pasha and T. S. Savithri, "Autonomuos robot control based on EEG and cross-correlation," 2016 10th International Conference on Intelligent Systems and Control (ISCO), Coimbatore, 2016, pp. 1-4.doi:10.1109/ISCO.2016.7727098

[10] J. Long, Y. Li, H. Wang, T. Yu, J. Pan, and F. Li, "A hybrid brain computer interface to control the direction and speed of a simulated or real wheelchair," *IEEE Trans. Neural Sys. Rehab. Eng.*, vol. 20, no. 5,pp. 720-729, Sept. 2012

[11] F. Galan, M. Nuttin, E. Lew, P.W. Ferrez, G. Vanacker, J. Philips, and J. del R. Millan, "A brain-actuated wheelchair: asynchronous and noninvasive brain-computer interfaces for continuous control of robots," *Clin. Neurophy.*, vol. 119, pp. 2159-2169, 2008

[12] S. Sanei, and J.A. Chambers, EEG Signal Processing, John Wiley & Sons, 2008

[13] G. Townsend, B. Graimann, and G. Pfurtscheller, "Continuous EEG classification during motor-imagery simulation of an asynchromous BCI," *IEEE Trans. Neural Sys. Rehab Engg.*, vol. 12, no. 2, pp. 258-265, 2004

[14] S.G. Mason, and G.E. Birch, "A General Framework for Brain-Computer Interface Design", *IEEE Trans. Neural Sys. & Rehab. Eng.*,vol. 11, no. 1, pp. 70-85, 2003

[15] T. Yamaguchi, M. Fujio, K. Inoue, and G. Pfurtscheller, "Wavelet Analysis of EEG signals during Motor Imagery," In: *Int Conf. Wavelet Analysis & Pattern recog. 2008*, pp. 454-59, Aug 2008

[16] P. Herman, G. Prasad, T.M. McGinnity, and D. Coyle, "Comparative Analysis of Spectral Approaches to Feature Extraction for EEG based Motor Imagery Classification," *IEEE Trans. Neural Sys. Rehab. Engg.*,vol. 16, no. 4, pp. 317-326, 2008

[17] C. Vidaurre, N. Kramer, B. Blankertz, and G. Pfurtscheller, "Time Domain Parameters as a feature for EEG-based brain-computer interfaces," *Neural Networks*, vol. 22, pp. 1313-1319, 2009

[18] A. Schlogl, "Dynamic Spectral Analysis based on Autoregressive model with time-varying coefficients," *IEEE-EMBC and CMBEC*, 1995

[19] Y. Fusheng, H. Bo, and T. Qingyu, Approximate entropy and its application to bio-signal analysis, In: Nonlinear Biomedical Signal Processing, Vol 2: Dynamic Analysis and Modeling, M.Akay, Ed., IEEE Press Series on Biomedical Engineering, NY (USA), pp. 72-92

[20] E. Alpaydin, Introduction to machine learning, MIT Press, 2004

[21] G. Townsend, B. Graimann, and G. Pfurtscheller, "Continuous EEG classification during motor-imagery simulation of an asynchromous BCI," *IEEE Trans. Neural Sys. Rehab Engg.*, vol. 12, no. 2, pp. 258-265,2004

[22] S. Sun, C. Zhiang, D. Zhang, "An experimental evaluation of ensemble methods for EEG signal classification," *Pattern Recognition Letters*, vol.28, pp. 2157-2163, 2007

[23] P. Xiaomei, Z. Chongxun, X. Jin, and B. Guangyu, "Adaboost for improving classification of left and right hand motor imagery tasks," In: *1st Int. Conf. on Neural Interface and Control*, Wuhan, China, pp.10-13, May 2005

[24] L. Breiman, "Bagging Predictors," *Machine Learning*, vol. 24, no. 2, pp. 123-140, 1996

[25] T.K. Ho, "The random subspace method for constructing decision forests," *IEEE Trans Pattern Analysis Mach. Int.*, vol. 20, no. 8, pp.832-844, 1998

[26] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche and B. Arnaldi, "A review of classification algorithms for EEG-based brain–computer interfaces", *Journal of Neural Engineering*, Volume 4, 2007

[27] A. Soria-Frisch, "A Critical Review on the Usage of Ensembles for BCI" in *Towards Practical Brain-Computer Interfaces*, 1st Edition, B. Z. Allison, Ed. Berlin: Springer Berlin Heidelberg, 2012, pp. 41-65

[28] T. G. Dietterich, "Ensemble Methods in Machine Learning" in *Multiple Classifier Systems: First International Workshop, MCS 2000 Cagliari, Italy, June 21-23, 2000 Proceedings*, Ed. Berlin: Springer Berlin Heidelberg, 2000, pp. 1-15

[29] L. Hansen and P. Salamon, "Neural Network Ensembles", *IEEE Transactions on Patern Analysis and Machine Intelligence*, Volume 12, 1990, pp. 993-1001

[30] L. I. Kuncheva, *Combining Pattern Classifiers: Methods and Algorithms*, Wiley-Interscience, 2004

[31] A. J. Meng, G. Liu, G. Huang, and X. Zhu, "Automated selecting subset of channels based on CSP in motor imagery brain-computer system," in Proc. 2009 IEEE Int. Conf. Robot. Bioinformat., Guilin, China, Dec. 19–23, 2009, pp. 2290–2294

[32] Siuly, S., & Li, Y. (2012).Improving the separability of motor imagery EEG signals using a cross correlation-based least square support vector machine for brain–computer interface. Neural Systems and Rehabilitation Engineering, IEEE Transactions on, 20(4), 526-538

[33] Bhattacharyya, S., Konar, A., Tibarewala, D.N., Khasnobish, A. and Janarthanan, R., 2014, January.Performance analysis of ensemble methods for multi-class classification of motor imagery EEG signal.In *Control, Instrumentation, Energy and Communication (CIEC), 2014 International Conference on* (pp. 712-716)

[34] Mahmood, A., Zainab, R., Ahmad, R.B., Saeed, M. and Kamboh, A.M., 2017, July. Classification of multi-class motor imagery EEG using four band common spatial pattern. In *Engineering in Medicine and Biology Society (EMBC), 2017 39th Annual International Conference of the IEEE* (pp. 1034-1037)

[35] Rakotomamonjy, A. and Guigue, V., 2008. BCI competition III: dataset II-ensemble of SVMs for BCI P300 speller. *IEEE transactions on biomedical engineering*, *55*(3), pp.1147-1154