

# Four-Class Motor Imagery EEG Signal Classification using PCA, Wavelet and Two-Stage Neural Network

Md. Asadur Rahman<sup>1</sup>

Department of Biomedical Engineering  
Khulna University of Engineering & Technology (KUET)  
Khulna-9203, Bangladesh

Farzana Khanam<sup>2</sup>

Department of Biomedical Engineering  
Jashore University of Science and Technology (JUST)  
Jashore-7408, Bangladesh

Md. Kazem Hossain<sup>3</sup>, Mohammad Khurshed Alam<sup>4</sup>, Mohiuddin Ahmad<sup>5</sup>

Department of Electrical and Electronic Engineering, Khulna University of Engineering & Technology (KUET)  
Khulna-9203, Bangladesh

**Abstract**—Electroencephalogram (EEG) is the most significant signal for brain-computer interfaces (BCI). Nowadays, motor imagery (MI) movement based BCI is highly accepted method for. This paper proposes a novel method based on the combined utilization of principal component analysis (PCA), wavelet packet transformation (WPT), and two-stage machine learning algorithm to classify four-class MI EEG signal. This work includes four-class MI events by an imaginary lifting of the left hand, right hand, left foot, and Right Foot. The main challenge of this work is to discriminate the similar lobe EEG signal pattern such as left foot VS left hand. Another critical problem is to identify the MI movements of two different feet because their activation level is very low and show an almost similar pattern. This work firstly uses the PCA to reduce the signal dimensions of the left and right lobe of the brain. Then, WPT is used to extract the feature from the different class EEG signal. Finally, the artificial neural network is trained into two stages – 1<sup>st</sup> stage identifies the lobe from the signal pattern and the 2<sup>nd</sup> stage identifies whether the signal is of MI hand or MI foot movement. The proposed method is applied to the 4-class MI movement related EEG signals of 15 participants and found excellent classification accuracy (>74% on average). The outcomes of the proposed method prove its effectiveness in practical BCI implementation.

**Keywords**—Brain-computer interface; electroencephalogram; motor imagery; principal component analysis; wavelet packet transformation; artificial neural network; classification

## I. INTRODUCTION

Brain-computer interface (BCI) creates a communication system between computer and brain functionality to control the other devices. BCI can be used to control devices like a wheelchair, room light, fan, etc. that assists physically challenged people. There are many modalities such as Electroencephalogram, Electroencephalogram (EEG), Functional near-infrared Spectroscopy, Magneto-encephalogram, Functional Magnetic Resonance Imaging, etc. to read the brain signal. Among these modalities, EEG is the most familiar and cheapest technique to record brain signal [1].

EEG is a non-invasive way to record neuro-electric signal which is often called an EEG signal. Since a number of brain stimuli have been proposed for implementing the BCI system. Among all other stimuli, motor imagery (MI) movement is the highest choice for the researchers [2]. MI has a special benefit because it needs no additional setup like visual stimuli [3].

An MI movement is a process where a candidate imagines the real movement execution and corresponding neuro-electric activities are recorded by the EEG modality. There are different types of EEG-based BCI like as simple and compound limb motor imagery [4], continuous arm movement from EEG signals [5] and individual finger movements from one hand using human EEG signals [6], etc. In the present works of literature of MI, most of the research works [7-12] are related to two-class or three-class such as left hand vs. right hand and left hand, right hand, and foot, respectively. Here both feet are considered as a single class.

Multiclass MI movement *i.e.*, more than 3 class classification task is always a challenging issue when the discrimination is necessary between two-foot movement. Two hands and one-foot MI movement classification provide acceptable classification accuracy though, four-class MI movement *i.e.*, imagery left hand (iLH), right hand (iRH), left foot (iLF), and right foot (iRF) classification is still not an acceptable range. To meet this challenge, a very handful research works [13-15] have been proposed those considered the fourth class as tongue movement, not an approach to discriminate between two imagery feet-based EEG signal along with the other two imagery hand. The approach of excluding the two-foot is due to having an almost similar pattern and neuro-activation of imagery feet movements. The precise difference between the patterns of two-foot movements is the main challenge to attain the acceptable classification accuracy regarding 4-class MI movement classification. Therefore, very wise feature selection and classifier models are two important tasks to meet this 4-class EEG signal classification regarding MI movements.

A number of feature extraction methods from EEG signal have been proposed in recent years. Within these feature extraction techniques, there are autoregressive (AR) methods [16, 17], wavelet transforms (WT) method [17-21], and phase-space reconstruction approach [10], CSP based methods [13, 14, 22], empirical mode decomposition [23-25], etc. For a wide range of pattern recognition, wavelet packet transformation (WPT) provides excellent time-frequency features. The WPT coefficients during EEG signal decomposition are widely used in EEG signal classification. But, for multiple class, the WPT based feature extraction has two important limitations: *i*) Structuring the features and *ii*) Selection of the bases [26]. The features are structured by WPT coefficients those are considered to yield the significant pattern of the different classes EEG signal. Besides the feature structuring, the proper base selection is the other step by which the structured features can show the highest discriminative characteristics among the classes. As a result, the WPT based features of the four classes are not enough criteria for satisfying classification accuracy. Therefore, there is a scope to add some innovative approach to meet the challenge.

In this paper, we have proposed a novel method for the aforementioned multiple class EEG signal classification utilizing a principal component analysis (PCA), WPT, and two-stage artificial neural network (ANN). The proposed method utilizes the concept of PCA to reduce the signal dimension. The reduced dimension EEG signals are fed to the WPT algorithm to extract its features. After that, we have modeled two ANN sequentially for identification of the lobe-origin and limb-origin of the signal. The proposed ANN-based two-stage model finds the signal's origin at first, *i.e.*, is the signal comes from either left lobe or right lobe and secondly, it decides whether the signal is of either lower limb or upper limb, *i.e.*, foot or hand. In this work, four-class motor imagery EEG data were collected in our laboratory from 15 participants. The data were preprocessed and separated according to the tasks. Then the dimensionality reduction and feature extraction were conducted by PCA and WPT, respectively. Finally, the two-stage ANN model was trained with the training data set. The proposed method was finally applied on the testing data set and we found that the proposed algorithm can improve classification performance (on average =11.25%) than the WPT-single ANN model.

The rest of the article is organized as: The materials used in this research work are described in Section II. The proposed methodology is elaborately explained in Section III. The results are presented with related discussions in Section IV. Finally, the whole work has been concluded briefly in Section V.

## II. MATERIALS

### A. Participants

Fifteen healthy adult male subjects (age=23±2.5 years) participated in this experiment. The participants did not have any psychological disorder and they were visually corrected person. In addition, all participants were not in any medication last one month. The participants claimed themselves having no pain in muscle and no mental disorder. All the subjects were right-handed based on Edinburg Handedness Inventory to avoid the variation of EEG signal pattern in left and right lobes during data collection. The participants were verbally informed about the protocol and they also practiced the protocol several times prior to proceed the actual data acquisition session. Their written consent was taken from all the participants prior to the EEG data collection. Since it was a volunteer-based contract, no participant was paid for the data collection.

### B. Data Acquisition Protocol

The subject was asked to sit on a chair before a computer screen at a convenient distance. Then he was told to take rest for a while to reduce the physiological artifacts which can be accumulated with the electrical activities of the brain. There were four sequential tasks that every participant had to do. These tasks were: iLH, iRH, iLF, and iRF movements. The experimental protocol is sketched in Fig. 1. The protocol was started with 5s rest. It was followed by 10s task and 20s rest. The tasks were performed, sequentially as given in Fig. 1 by the participants. One participant performed 5 trials per each session and 4 sessions per day. Each participant participated in the data acquisition in two different data in a single week. Eventually, forty trials of each task were acquired from each participant. The protocol does not violate the Declaration of Helsinki [27].

### C. Data Acquisition and Data Description

The EEG data were acquired utilizing the 9-channel B-Alert X10 wireless device in the Neuroimaging Laboratory, Department of Biomedical Engineering, Khulna University of Engineering & Technology (KUET). The data logging in computer memory was conducted by Acqknowledge 4.4 software [28]. The acquisition sampling rate of the data was 256 Hz. The collected 9 channel EEG data covers the following position of the scalp F3, Fz, F4, C3, Cz, C4, P3, Pz, and P4 according to the international 10/20 system. Therefore, this device covers the frontal, central, and parietal lobe of the brain during EEG data acquisition. The data acquisition procedure of this research work with the 9-channel B-Alert X10 wireless device is graphically presented in Fig. 2.

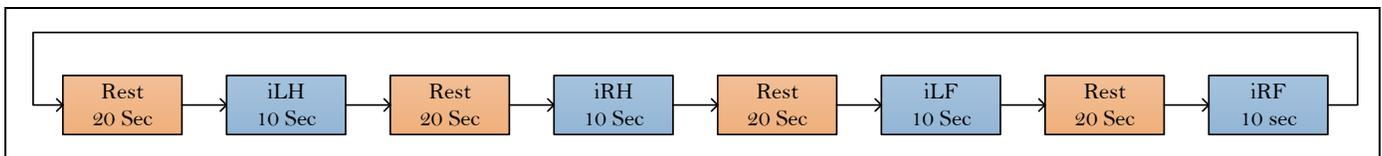


Fig. 1. Data Acquisition Protocol with a Time Schedule of the Tasks.



Fig. 2. MI EEG Data Acquisition Procedure Utilizing the B-Alert X-10 Wireless Device with a Computer-Assisted Command.

### III. METHODS

#### A. Proposed Method

Our proposed method is an offline EEG signal classification method. After the raw EEG signal acquisition, several processing steps were conducted to prepare the signal for machine understanding format. At first, the raw EEG signals were filtered with a 50 Hz notch filter to remove the power line noise from the signal. A 2<sup>nd</sup> order band stops IIR filter was designed considering its cut off band from 49 to 51 Hz and utilized as a notch filter for the EEG signal. Secondly, the signals were filtered with 3<sup>rd</sup> order bandpass IIR filter with a passband from 4 to 100 Hz. The eye blink effect was removed from the signal utilizing the method described in [29, 30] with their Matlab based automatic EEG artifact rejection toolbox, EAWICA.

After filtering the data, the EEG signals were separated according to the tasks. These filtered data were used as input of our proposed method to model the 4-class MI movement prediction. The proposed method to construct the ANN-based predictive model can be presented by the following block diagram given in Fig. 3.

A complex with several steps is shown in Fig. 3 where the input EEG signal is considered as preprocessed and separated for further analysis. Then the dimension on the signals is reduced applying the PCA. Since the EEG signals were acquired nine channel device, the left lobe channels F3, C3, and P3 are reduced along with the neutral positions Fz, Cz, and Pz to a single channel with the combination of these six channels with their correlation weight utilizing the PCA concept. On the other hand, the right lobe channels F4, C4, and P4 are also reduced in the same way with PCA. Therefore, applying the PCA the 9-channel EEG signal was reduced to two channel signals. During PCA, we considered only the highest PCA (1<sup>st</sup> principal components) factors.

According to the proposed method, WPT was used to extract the features from these two reduced signals of every task, separately. After that the features of 4-class have been broadly divided into two classes based on the lobe of operation *i.e.*, *i*) iLH & iLF and *ii*) iRH & iRF. These wider two classes of features were used to train the ANN to predict the input signal whether it is of left or right. Additionally, another two classes were formed by the features of iLH+iRH and iLF+iRF and trained the ANN so that it can predict whether the signal is off hand or foot.

Therefore, two separate ANN were utilized to make a hybridized two-stage ANN predictive model that would be able to predict a scratch signal of any of the four classes. For any MI EEG signal, this model will predict the left or right limb signal, firstly and then in the second stage, the model will find it as a signal of hand or foot. The classification procedure of the two-stage ANN is given in Fig. 4. In this figure, the block preprocessing includes the filtering and dimensionality reduction by PCA as explained before.

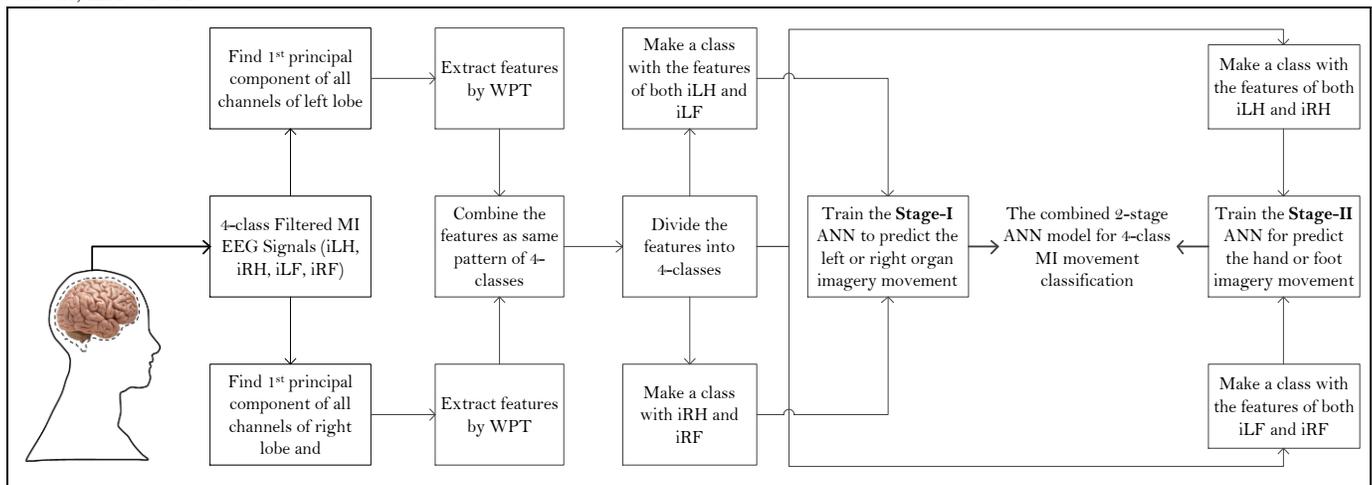


Fig. 3. The Block Diagram of the Proposed Method for Training and Constructing the Two-Stage ANN-based Predictive Model Along with PCA and WPT based Feature Extraction.

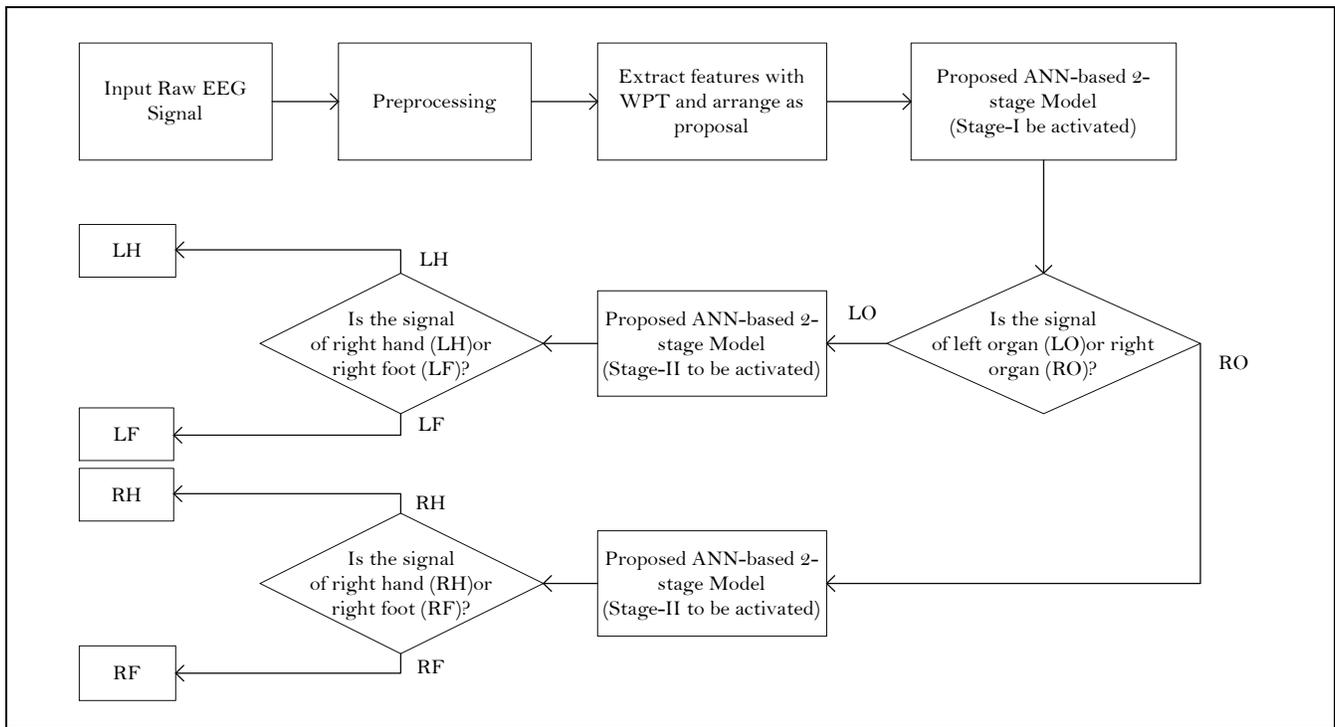


Fig. 4. The Proposed Two-Stage ANN-based Predictive Model and its Classification Technique from a Raw Signal to Decision.

### B. Principal Component Analysis

A matrix  $\Delta$  consists of data of  $n$  dimension. Now, a matrix  $H$  can be designed which characterizes the eigenvectors arranged as the eigenvalues of the covariance matrix of  $\Delta$ . Then we get the PCA of the data  $\Delta$  in the form of  $P$  as,

$$P = H^T \Delta \quad (1)$$

The eigenvectors can also be named as the principal components. To project the data, if first rows ( $R$ ) of  $P$  are selected, the data becomes of  $R$  dimensional from  $d$  dimensions. This transformation is performed by singular value decomposition (SVD). Matrix decomposition can describe the procedure to perform PCA by SVD. Suppose, the matrix,  $\Delta$  can be decomposed using SVD as

$$\Delta = \Psi \kappa \theta^T \quad (2)$$

Here,  $\Psi$  is an  $n \times m$  matrix with orthonormal columns ( $\Psi^T \Psi = I$ );  $\theta$  is an  $m \times m$  orthonormal matrix ( $\theta^T \theta = I$ ), and  $\kappa$  is an  $m \times m$  diagonal matrix with positive or zero elements which is recognized as a singular value. Besides, the covariance matrix can be calculated,  $C$  of  $\Delta$  as,

$$C = \frac{1}{N} \Delta \Delta^T = \frac{1}{N} \Psi \kappa^2 \Psi^T \quad (3)$$

As the singular values are organized in descending order and if  $n < m$ , the first  $n$  columns in  $\Psi$  corresponds to the sorted eigenvalues of matrix  $C$  and if  $m \geq n$ , the first  $m$  corresponds to the sorted non-zero eigenvalues of  $C$ . Therefore, eventually the transformed data can be written as,

$$P = H^T \Delta = H^T H \kappa \theta^T \quad (4)$$

### C. Wavelet Packet Transformation for Feature Extraction

The WPT is a concept that is different from conventional wavelet transformation (WT). WPT decomposes both the approximate coefficients and the detailed coefficients. The WPT may be considered as a subspace tree. We can present the original signal as  $\Pi_{0,0}$  which reflects the root mood of the tree in the original signal space. Generally, the notation  $j$  and  $k$  in  $\Pi_{j,k}$  denotes the scale and sub-band space. The WPT decomposes an original signal  $\Pi_{j,k}$  into two different subspaces: an approximation space  $\Pi_{j,k} \rightarrow \Pi_{j+1,2k}$  and a detailed space,  $\Pi_{j,k} \rightarrow \Pi_{j+1,2k+1}$ . This space decomposition utilizes the concept of dividing the orthogonal basis function  $\{\phi_j(t-2^j k)\}_{k \in \mathbb{Z}}$  of the original signal space into two new orthogonal bases, i)  $\{\phi_{j+1}(t-2^{j+1} k)\}_{k \in \mathbb{Z}}$  of approximate space  $\Pi_{j+1,2k}$  and ii)  $\{\psi_{j+1}(t-2^{j+1} k)\}_{k \in \mathbb{Z}}$  of detailed space  $\Pi_{j+1,2k+1}$ . Here  $\Pi_{j,k}(t)$  and  $\Psi_{j,k}$  represents the scaling and wavelet functions, respectively. These functions are equated as [26]:

$$\phi_{j,k}(t) = \frac{1}{\sqrt{12^j}} \phi\left(\frac{t-2^j k}{2^j}\right) \quad (5)$$

$$\psi_{j,k}(t) = \frac{1}{\sqrt{12^j}} \psi\left(\frac{t-2^j k}{2^j}\right) \quad (6)$$

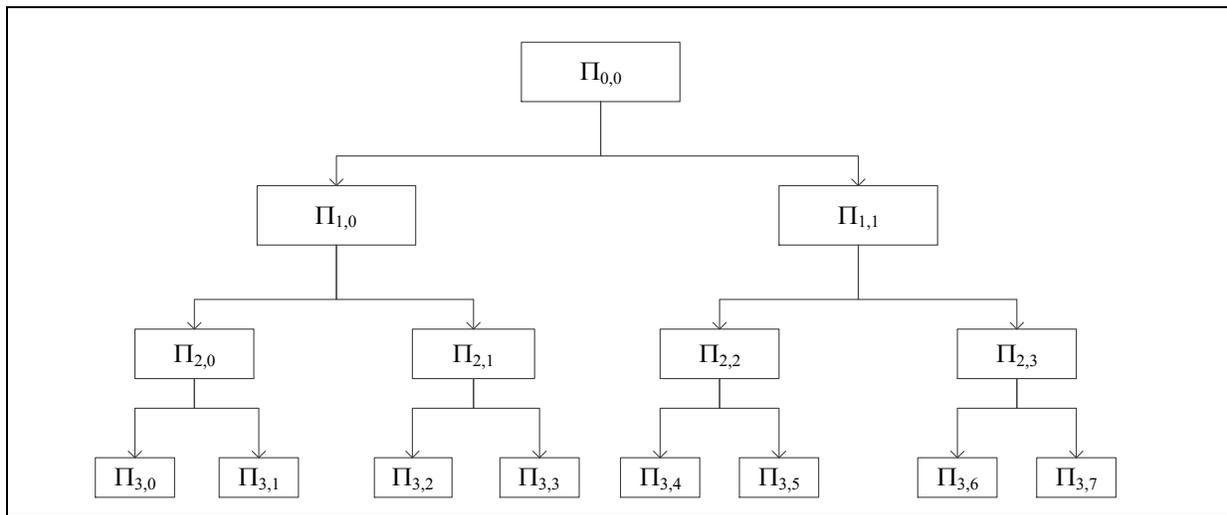


Fig. 5. Graphical Representation of Wavelet Packets Decomposition Method that Decomposes  $\Pi_{0,0}$  into Tree-Structured Subspaces.

Here  $2^j$  is the scaling parameter that measures the scaling or compression degree of the original signal. In addition,  $2^j k$  is the location parameter or translation parameter that indicates the time location of the wavelet. The aforesaid process can be repeated  $J$  times, where  $J$  must be less than  $\log_2 N$ . Here,  $N$  is the total number of samples in the original signal. This process of WPT  $J \times N$  finds coefficients. Therefore, at any level of transformation  $j [j = 1, 2, \dots, J]$ , the tree has  $N/(2^j)$  coefficient blocks. This iterative process in a WPT can be treated as a tree-like structure, where the tree nodes represent the subspaces of different frequency localization characteristics. The corresponding decomposition procedure can be presented as Fig. 5 [26, 31].

**D. Artificial Neural Network**

ANN replicates the functional concept of the human brain. The multilayer feedforward ANN has three basic layers: an input layer, an output layer, and a hidden layer. A typical model of a feed-forward network with its prominent layers is given in Fig. 6. The four outputs are chosen in the output layer because total data are to be classified by this work into four classes.

In a supervised neural network, the input layer is formed based on the size of the features and the output layer is chosen upon its number of classes to be classified. The significant layer is a hidden layer which is connected to the input and output with single or multiple layers while in case of multiple layers are often referred to multilayer neural network. The connected manners are actually a mathematical function with a predefined function which is also known as a neuron. Generally, a neuron at  $j$  label receives an input  $p_j(t)$  from the previous neuron at a discrete time,  $t$ . Suppose, the activation function of the neuron at this level is  $a_j(t)$  where a threshold,  $\theta_j$  is also chosen. Therefore, the activation function,  $\zeta$  computes the next level activation,  $a_j(t+1)$  from the current information as [32],

$$a_j(t+1) = \zeta(a_j(t), p_j(t), \theta_j) \tag{7}$$

Inside an ANN, the output of a neuron  $i$  performs as an input to a neuron  $j$  and each connection is assigned with a weight  $w_{ij}$  with a bias term  $w_{0j}$  (sometimes). A propagation function calculates the input of the neuron  $j$  from the output,  $o_i(t)$  of the neuron  $i$  with the assigned bias value as [33],

$$p_j(t) = \sum_i o_i(t)w_{ij} + w_{0j} \tag{8}$$

In a supervised learning method of an ANN, a set of given pairs of input features,  $x$  and output,  $y (x \in X, y \in Y)$  are fed to find a function,  $\Lambda: X \rightarrow Y$  subject to the assumed class of functions. By this predictive function, it is expected that it could be able to infer the applied data mapping. A cost function is used to relate the expected and actual mismatch of the inferring data map that has prior knowledge about the domain of interest. A commonly used cost function for the ANN is a mean-square error that aims to minimize the mean square error between  $\Lambda(x)$  and the target  $y$ . In a multilayer perceptron, ANN network utilizes the gradient descent algorithm to optimize the cost function by the backpropagation algorithm.

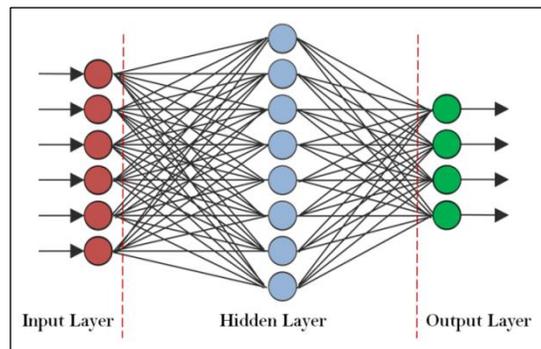


Fig. 6. A Typical Model of an ANN with the Structure of its basic Layers.

#### IV. RESULTS AND DISCUSSIONS

The results regarding the proposed analytical methods were demonstrated by Matlab 2018a [34]. Utilizing the Acqknowledge software the raw EEG data were converted to .mat file to make it compatible with Matlab. The filtering, feature extraction, classification, etc. processing was conducted utilizing the different toolboxes of Matlab 2018 in an offline fashion. The raw signal was filtered with several steps as notch filtering, bandpass filtering, and eye blink removal. The

following procedures were applied and the resulting effects on the EEG signal are presented in Fig. 7. Then, the dimensionality reduction of the six-channel EEG signal was performed and one-dimensional EEG signal is prepared from a lobe using PCA. PCA converts the signal taking the maximum variations in the signal and avoiding the similarity among the signal which helps to get maximum frequency contents without taking the curse of dimensionality. The resulting signals are graphically presented in Fig. 8.

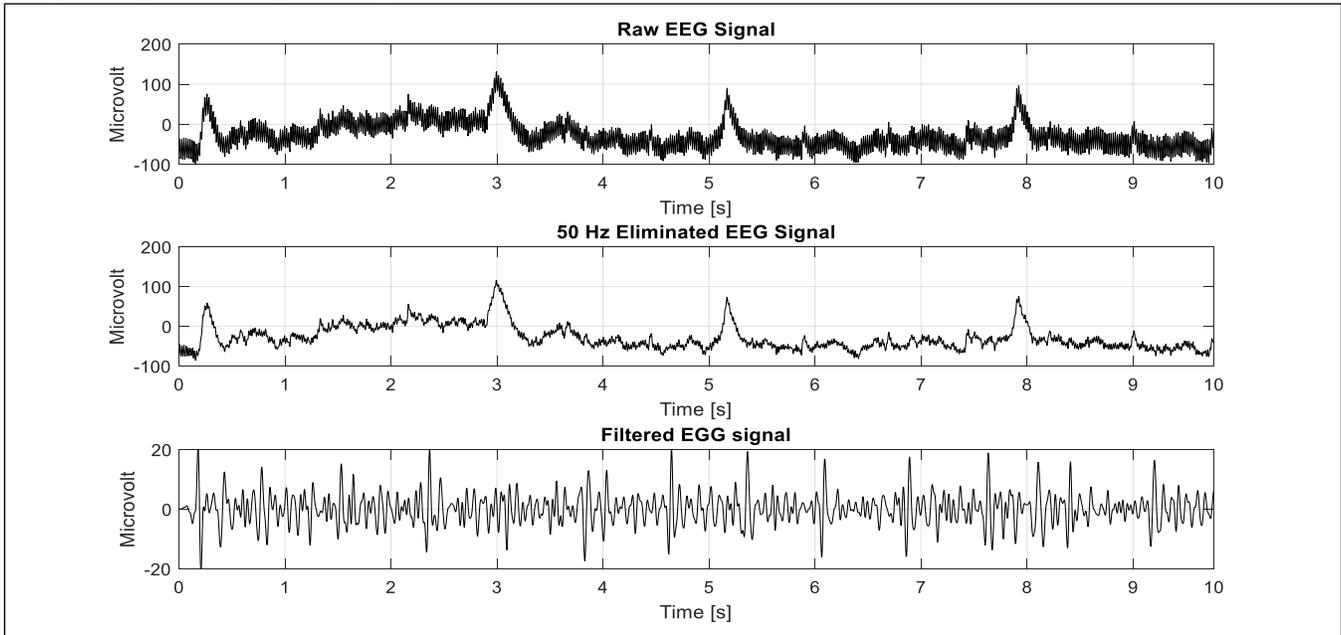


Fig. 7. Preprocessing Steps.

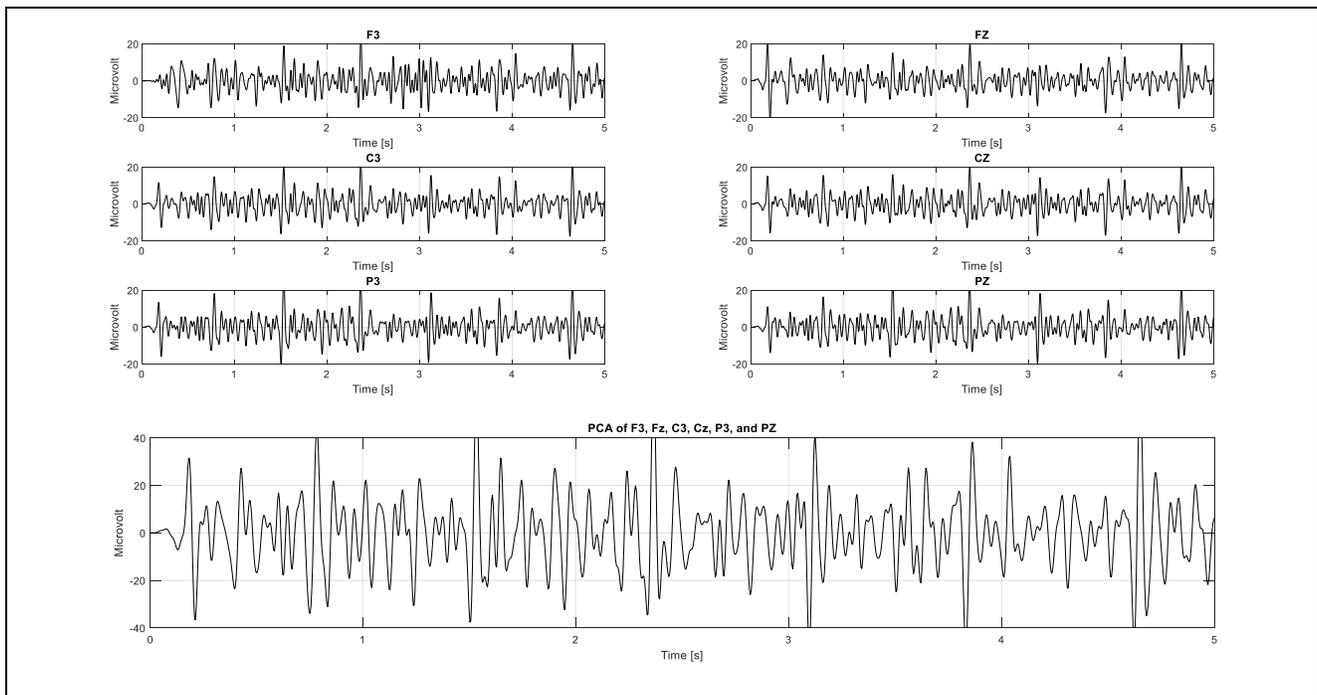
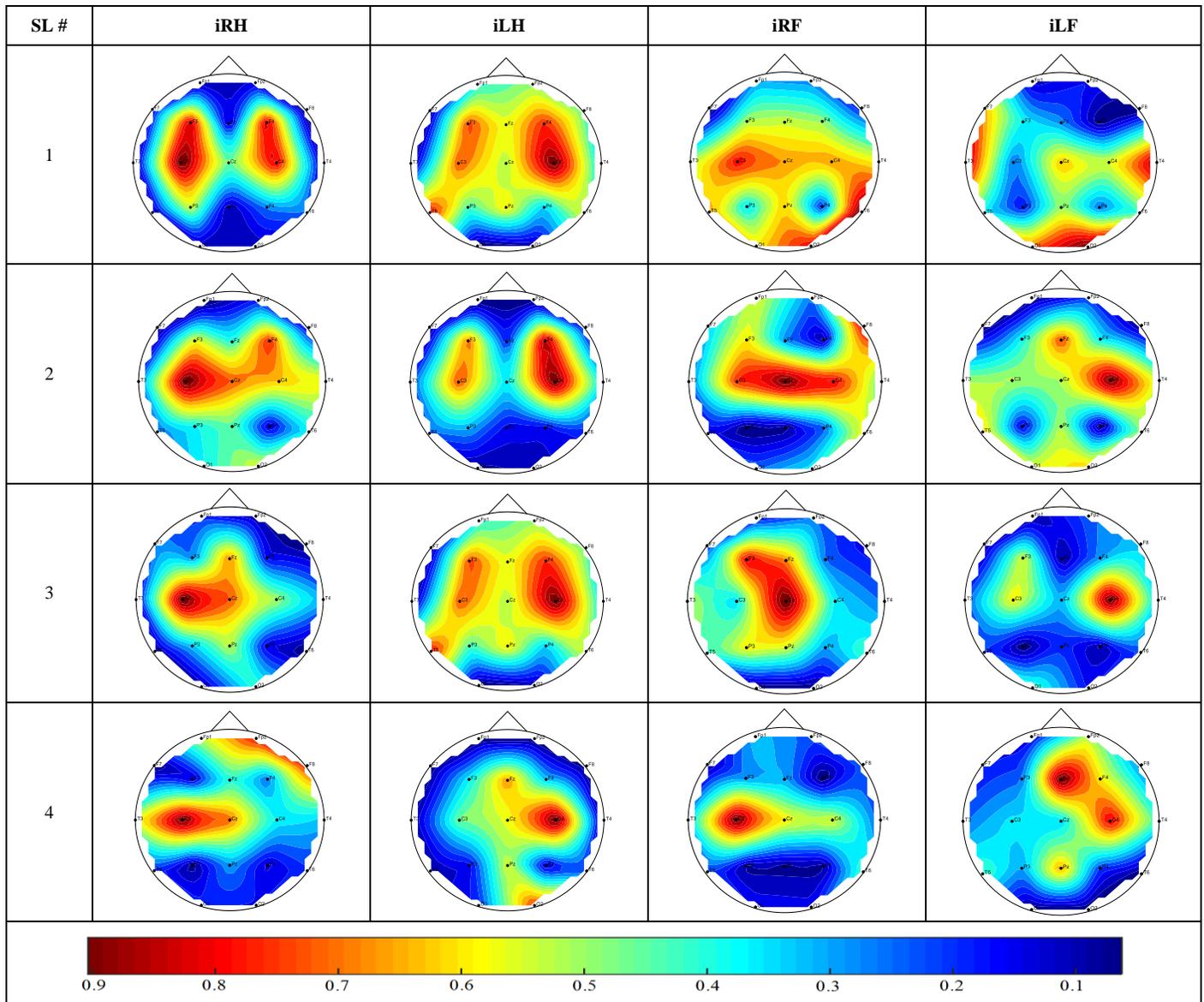


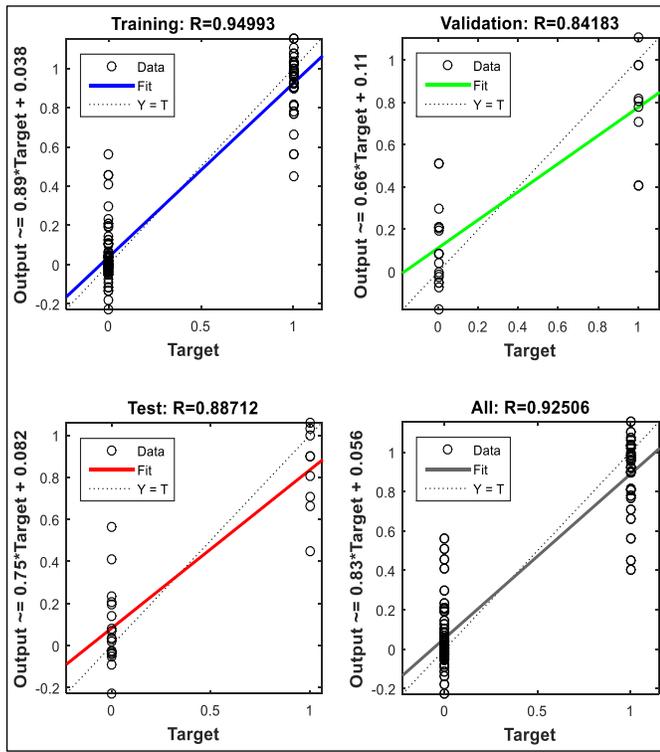
Fig. 8. The Proposed Method of Dimensionality Reduction of the Channels by PCA.

TABLE I. TOPOPLOTS OF THE FOUR PARTICIPANTS REGARDING DIFFERENT IMAGERY STIMULI

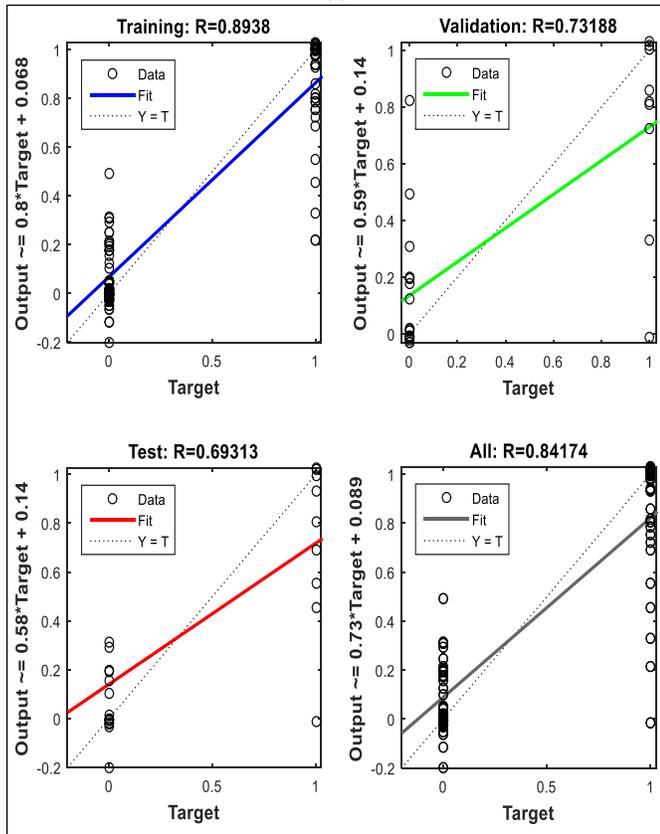


From the strength of the signal power regarding four imagery movements, we found various kinds of pattern. The EEG signals of four randomly chosen participants were analyzed to observe the neural activation pattern due to the different applied stimuli. The neural activations of the aforesaid conditions were prepared based topographic plot or topoplot on demo human scalp. Utilizing the open-source Matlab based function of topoplot (which is solely designed for the 9 channel B-alert EEG data and available in [35]); the activation based topoplots of aforesaid stimuli are given in Table I. The resulting topoplots demonstrate that the imagery hands movements (both iLH and iRH) show consistent pattern in their neural activations. On the other hand, there are some discrepancies in the neural activations for imagery feet movements. This is the cause we utilized the two stages ANN training method to recognize this pattern with discrepancies of the feet movements.

In the training session of the ANN, the wavelet tree coefficients were used as features. The training and testing features were separated as a 5-fold cross-validation technique. The classification accuracies were calculated utilizing the proposed method as a subject dependent approach. It is found during the training process that the first stage training performance is better than the second stage training. It may occur due to the similarity of the EEG based neural activation of the same lobe either from hand or foot. The training sessions of the features of subject 1 are given in Fig. 9. In this figure, the stage I training and validation performances are given in Fig. 9(a) and the stage II training and validation performances are given in Fig. 9(b). It is clear from the figures that, the performances in case of stage II are inferior to that of the stage I. It has a significant impact on the classification accuracy. Such training and testing were conducted for an individual subject to present the classification accuracy.



(a)



(b)

Fig. 9. Training and Testing Performances of the Proposed Network in its Stage I (a) and Stage II (b) for Participants 1.

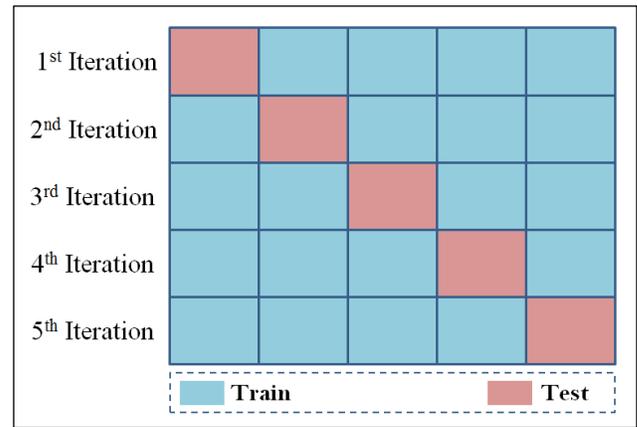


Fig. 10. The Training and Testing Ratio of the 5-Fold Cross-Validation Technique.

Since there are 40 trials of each imagery task, according to the 5-fold cross-validation technique 32 trials were used to train and validation of the network and the rest 8 trials of each task were used to test the accuracy of the trained network. The selection of the training and testing trials from the 40 trials were performed 5 different sets as the presentation in Fig. 10. The final classification accuracy was estimated from the average value of the five testing results. Using this consideration, the classification accuracy of the subject one was calculated and the classification performance of the proposed method is given by a confusion matrix in Fig. 11. This result is of the 1<sup>st</sup> iteration of the 5-fold cross-validation technique where we found that the classification accuracy is 81.3%. Similar four more iterations were performed to get the average classification accuracy. We found the classification accuracy for rest four trials as 68.75%, 78.2%, 71.9%, 65.7%, respectively. Therefore, the average four-class classification accuracy for participants is 73.17%.

		Confusion Matrix				
		IRH	ILH	IRF	ILF	IRH
Output Class	IRH	7 21.9%	1 3.1%	2 6.3%	0 0.0%	70.0% 30.0%
	ILH	0 0.0%	7 21.9%	0 0.0%	1 3.1%	87.5% 12.5%
	IRF	1 3.1%	0 0.0%	6 18.8%	1 3.1%	75.0% 25.0%
	ILF	0 0.0%	0 0.0%	0 0.0%	6 18.8%	100% 0.0%
		87.5% 12.5%	87.5% 12.5%	75.0% 25.0%	75.0% 25.0%	81.3% 18.8%
		IRH	ILH	IRF	ILF	IRH

Fig. 11. Confusion Matrix of Classification Accuracy of 1<sup>st</sup> Iteration Among the 5 Iterations.

TABLE II. CLASSIFICATION ACCURACIES FOUND UTILIZING THE SINGLE-STAGE AND TWO-STAGE ANN CLASSIFIERS

Sub. Num.	Classification Accuracies (%)	
	One-stage ANN	Two-stage ANN
1	62.24	73.17
2	63.50	74.65
3	52.5	68.25
4	60.25	81.46
5	58.24	75
6	45.8	65.5
7	50.8	71.25
8	60.45	75.50
9	62.50	78.25
10	58.90	72
11	58.90	72
12	61.45	82.17
13	62.45	86.25
14	60.50	73.40
15	55.5	70.25
Average $\pm$ std	58.26 $\pm$ 5.04	74.60 $\pm$ 5.50

REFERENCES

[1] L. F. Nicolas-Alonso and J. Gomez-Gil, "Brain computer interfaces, a review," *Sensors*, vol. 12, pp. 1211-1279, 2012.

[2] K. Hong, M. J. Khan, and M. J. Hong, "Feature extraction and classification methods for hybrid fNIRS-EEG brain-computer interfaces," *Frontiers in Human Neuroscience*, vol. 12, no. 246, 2018.

[3] K. Kitahara, Y. Hayashi, S. Yano, and T. Kondo, "Target-directed motor imagery of the lower limb enhance event-related desynchronization," *PLOS One*, vol. 12, no. 9, pp. 1-15, 2017.

[4] Y. Weibo, S. Qiu, H. Qi, L. Zhang, B. Wan, and D. Ming, "EEG feature comparison and classification of simple and compound limb motor imagery," *J. Neuroeng. Rehabil* 10 (2013).

[5] W. J. Seok, K. R. Muller, and S. W. Lee, "Classifying directions in continuous arm movement from EEG signals." 3<sup>rd</sup> International Winter Conference on BrainComputer Interface (BCI), 2015, pp. 1-2.

[6] K. Liao, R. Xiao, J. Gonzalez, and L. Ding, "Decoding individual finger movements from one hand using human EEG signals." *PLoS One*, vol. 9, no. 1, 2014.

[7] M. Li, W. Zhu, H. Liu, and J. Yang, "Adaptive feature extraction of motor imagery EEG with optimal wavelet packets and SE-isomap," *Applied Sciences*, vol. 7, no. 390, pp. 1-18, 2017.

[8] Y. Ma, X. Ding, Q. She, Z. Luo, T. Potter, and Y. Zhang, "Classification of motor imagery EEG signals with support vector machines and particle swarm optimization," vol. 2016, Article ID. 4941235, pp. 1-8, *Computational and Mathematical Methods in Medicine*, 2016.

[9] Z. Tang, C. Li, J. WU, P. LIU, and S. Cheng, "Classification of EEG-based single-trial motor imagery tasks using a B-CSP method for BCI," *Frontiers of Information Technology & Electronic Engineering*, 2018.

[10] R. Djemal, A. G. Bazyed, K. Belwafi, S. Gannouni, and W. Kaaniche, "Three-class EEG-based motor imagery classification using phase-space reconstruction technique," *Brain Sciences*, vol. 6, no. 36, pp. 1-19, 2016.

[11] L. Cao, B. Xia, O. Maysam, J. Li, H. Xie, and N. Birbaumer, "A synchronous motor imagery based neural physiological paradigm for brain-computer interface speller," *Frontiers in Human Neuroscience*, vol. 11, no. 274, pp. 1-9, 2017.

[12] J. Petersen, H. K. Iversen, and S. Puthusserypady, "Motor imagery based brain-computer interface paradigm for upper limb stroke rehabilitation," *International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Honolulu, HI, 2018, pp. 1960-1963.

[13] L. Wang and X. Wu, "Classification of four-class motor imagery EEG data using spatial filtering," *International Conference on Bioinformatics and Biomedical Engineering*, Shanghai, 2008, pp. 2153-2156.

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

[15] S. Ge, R. Wang, and D. Yu, "Classification of four-class motor imagery employing single-channel electroencephalography," *PLOS One*, vol. 9, no. 6, pp. 1-7, 2014.

[16] Y. Zhang, X. Ji, and Y. Zhang, "Classification of EEG signals based on AR model and approximate entropy," *International Joint Conference on Neural Networks (IJCNN)*, Killarney, 2015, pp. 1-6.

[17] Y. Zhang, B. Liu, X. Ji, D. Huang, "Classification of EEG signals based on autoregressive model and wavelet packet decomposition," *Neural Processing Letter*, vol 45, no. 2, pp. 365-378, 2017.

[18] H. U. Amin, A. S. Malik, R. F. Ahmad, N. Badruddin, N. Kamel, M. Hussain, and W-T. Chooi, "Feature extraction and classification for EEG signals using wavelet transform and machine learning techniques," *Australasian Physical & Engineering Sciences in Medicine*, vol. 38, no. 1, pp. 139-149, 2015.

[19] F. Sherwani, S. Shanta, B. S. K. K. Ibrahim, and M. S. Huq, "Wavelet-based feature extraction for classification of motor imagery signals," *IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, Kuala Lumpur, 2016, pp. 360-364.

Furthermore, the similar feature extraction method was applied for classification utilizing stage ANN classifier. The results due to the conventional one-stage classifier and the proposed two-stage classifiers for 15 participants are given in Table II. Here the average classification accuracies from 5-fold cross-validations are considered. The outcomes suggest that the two-stage classification accuracies are better than that of the conventional one-stage classification method with the same features. In average, the two-stage classification technique provides 74.60% accuracies whereas, the single-stage classifiers give use 58.26 % classification accuracies regarding the four-class motor imagery EEG signal.

V. CONCLUSIONS

By this proposed work, it has been found that not only the innovative feature extraction is mandatory but also the classifier setup with an appropriate approach is a concerning issue. The feature extraction method utilizing the PCA based wavelet packet transformation is although an excellent approach to find the properties of the EEG signal; it proves failure to classify the four-class motor imagery signals with satisfactory accuracy. On the other hand, the proposed two-stage classifier improves the classification accuracy at 16.34% on average which is a remarkable outcome. This approach can be applied in any higher-class classifier than two-class. Therefore, this proposal hopefully outtakes the conventional approach in practical BCI implementation.

One potential drawback of the proposed multi-stage training-based classifier is its time requirement for the training stage. If the classification accuracy becomes the first priority, the said limitation can be avoided in the implementation.

- [20] A. Sharmila and P. Mahalakshmi, "Wavelet-based feature extraction for classification of epileptic seizure EEG signal," *Journal of Medical Engineering & Technology*, vol. 41, no. 8, pp. 670-680, 2017.
- [21] M. A. Rahman, M. M. O. Rashid, F. Khanam, M. K. Alam, and M. Ahmad, "EEG based brain alertness monitoring by statistical and artificial neural network approach," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 1, 2019.
- [22] P. K. Saha, M. A. Rahman, and M. N. Mollah, "Frequency domain approach in CSP based feature extraction for EEG signal classification," *International Conference on Electrical, Computer and Communication Engineering (ECCE)*, 7-9 February 2019, Cox's Bazar, Bangladesh.
- [23] P. A. Munoz-Gutierrez, E. Giraldo, M. Bueno-Lopez, and M. Molonas, "Localization of Active Brain Sources From EEG Signals Using Empirical Mode Decomposition: A Comparative Study," vol. 12, no. 55, pp. 1-14, 2018.
- [24] J. Sokhal, B. Garg, S. Aggarwal, and R. Jain, "Classification of EEG signals using empirical mode decomposition and lifting wavelet transforms," *International Conference on Computing, Communication, and Automation (ICCCA)*, Greater Noida, 2017, pp. 1197-1202.
- [25] T. Nazneen, M. A. Rahman, and M. N. Mollah, "Towards the effective intrinsic mode functions for motor imagery EEG signal classification," *International Conference on Electrical, Computer and Communication Engineering (ECCE)*, 7-9 February 2019, Cox's Bazar, Bangladesh.
- [26] R. N. Khushaba, S. Kodagoda, S. Lal, and G. Dissanayake, "Driver drowsiness classification using fuzzy wavelet-packet-based feature-extraction algorithm," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 1, pp. 121-131, 2011.
- [27] World medical association declaration of Helsinki-ethical principles for medical research involving human subjects, Adopted by 64th WMA General Assembly, Fortaleza, Brazil, Special Communication: Clinical Review & Education. (2013).
- [28] Acqknowledge Software (v 4.4). Available in: <https://www.biopac.com/demo/acqknowledge-4-demo/>
- [29] N. Mammone and F. C. Morabito, "Enhanced automatic wavelet independent component analysis for electroencephalographic artifact removal," *Entropy*, vol. 16, no. 12, pp. 6553-6572, 2014.
- [30] N. Mammone, F. L. Foresta, and F. C. Morabito, "Automatic artifact rejection from multichannel scalp EEG by wavelet ICA," *IEEE Sensors Journal*, vol. 12, no. 3, pp. 533-542, 2012.
- [31] M. K. Wali, M. Murugappan, and B. Ahmmad, "Wavelet packet transform based driver distraction level classification using EEG," *Mathematical Problems in Engineering*, vol. 2013, Article ID 297587, 10 pages, 2013.
- [32] A. Zell, *Simulation Neuronaler Netze [Simulation of Neural Networks] (in German) (1st ed.)*. 1994, Addison-Wesley. ISBN 978-3-89319-554-1.
- [33] C. W. Dawsin, "An artificial neural network approach to rainfall-runoff modelling," *Hydrological Sciences Journal*, vol. 43, no. 1, pp. 47-66, 1998.
- [34] Matlab Product (Matlab 2018a). Available in: <https://www.mathworks.com/company/newsroom/mathworks-announces-release-2018a-of-the-matlab-and-simulink-product-families.html>
- [35] M. A. Rahman, "Topoplot for B-Alert X-10 9-Channel EEG Signal," <https://www.mathworks.com/matlabcentral/fileexchange/69991-topoplot-for-b-alert-x-10-9-channel-eeeg-signal>, MATLAB Central File Exchange. Retrieved April 2, 2019.