2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Development of Machine-Learning Algorithms for Recognition of Subjects' Upper Limb Movement Intention Using Electroencephalogram Signals

Fatima Al-Khuzaei, Leen Al Homoud, *Member, IEEE,* Dana Alyafei, Reza Tafreshi, *Senior Member, IEEE,* and Md Ferdous Wahid

***Abstract*— This study aims to classify rest and upper limb movements execution and intention using electroencephalogram (EEG) signals by developing machine-learning (ML) algorithms. Five different MLs are implemented, including k-Nearest Neighbor (KNN), Linear Discriminant Analysis (LDA), Naïve Bayes (NB), Support Vector Machine (SVM), and Random Forest (RF). The EEG data from fifteen healthy subjects during motor execution (ME) and motor imagination (MI) are pre-processed with Independent Component Analysis (ICA) to reduce eye-blinking associated artifacts. A sliding window technique varying from 1 s to 2 s is used to segment the signals. The majority voting (MV) strategy is employed during the post-processing stage. The results show that the application of ICA increases the accuracy of MI up to 6%, which is improved further by 1-2% using the MV ($p<0.05$). However, the improvement in the accuracies is more significant in MI (>5%) than in ME (<1%), indicating a more significant influence of eye-blinking artifacts in the EEG signals during MI than ME. Among the MLs, both RF and SVM consistently produced better accuracies in both ME and MI. Using RF, the 2 s window size produced the highest accuracies in both ME and MI than the smaller window sizes.**

***Index Terms*— Electroencephalogram, hand gestures, machine learning, post-processing, majority voting, independent component analysis**

I. Introduction

Electroencephalogram (EEG) measures brain electrical activity. The ability to correlate human movements to brain signals using EEG has tremendous potential in the biomedical field. Classifying EEG signals helps understand how the brain controls the human body and is vital for neuroprosthesis [1]. Recognizing motion intention is important for limb rehabilitation and the improvement of prosthetic and exoskeleton devices' performance [2].

Analyzing the EEG signals is challenging as they are noisy and can easily be contaminated by motion artifacts such as eye-blinking or head movement. Independent Component Analysis (ICA) is usually used to remove eye-blinking artifacts from the EEG signals. The ICA uses blind source separation (BSS) methods to separate the EEG signals into statistically independent sources. The independent components representing eye artifacts are then removed from the EEG signals [3]. In [4], the authors reported that the ICA-based classification could outperform the time/frequency-based classification by 13%.

Different MLs have been previously investigated to predict the subject's motion intention. Bandara et al. [5] classified different states of arm movement, including rest position and grabbing a cup of water, using both artificial neural network (ANN) and support vector machine (SVM) algorithms. The authors reported an overall accuracy of 71.3-72.6% and 81.9-82.1% for the ANN and SVM algorithms, respectively. Other studies have used Linear Discriminant Analysis (LDA) to detect motion intention of different hand and arm movements with accuracy ranges between 47% and 76% [2], [6].

Movement-vs-rest and movement-vs-movement were classified using shrinkage LDA with accuracy of 55% for the movement-vs-movement and 87% for the movement-vs-rest [1]. Using the same EEG data from [1], other studies [7]–[10] used deep learning-based models to automate the feature extraction and selection steps. In [7], the authors used Deep Convolutional Neural Network to classify motor planning, referred to as a pre-movement, using the 1 s window segment preceding motion onset and reported an average accuracy of 90.3% for rest vs. pre-movement. Applying Regional Attention Convolutional Neural Network for multi-class classification of six movement and rest classes showed an average accuracy of 42.6% and 33.1% for motor execution (ME) and motor imagination (MI), respectively [8]. In [9], the authors used Common Spatial Patterns (CSP)-based features for the Convolutional Neural Network (CNN) and reported an average accuracy of 87.92% for palm extension vs. hand grasp of ME. However, the authors used the same CSP algorithms, which itself is a supervised method for feature extraction, for both training and testing dataset, and subsequently biased the CNN's outcomes. Finally, in [10], the authors used Wavelet–spatial filters ConvNet and discriminative spatial pattern to classify six movements for MI and reported an average of 31% accuracy.

Notably, the first half of the signals after the cue produces higher accuracy; however, the model may not perform well in real-time applications as it is not known when and for how long a subject may imagine the task. In summary, none of the above mentioned studies used a) the full duration of ME or MI EEG signals after the cue (previous studies have used either one second before the motion begins or only up to the

This work was supported by UREP grant from the Qatar National Research Fund (a member of Qatar Foundation), Grant No. 25-116-2-044.
F.A., D.A., R.T., and M.F.W., are with the Mechanical Engineering Program, and L.A. is with the Electrical Engineering Program, Texas A&M University at Qatar, Doha, Qatar (e-mail: f.khuzai, leen.alhomoud, dana.alyafei, rtafreshi, wahid @{tamu.edu}).

first half of the signals after the cue), b) post-processing such as majority voting (MV), and c) varying window size from 1 to 2 s. In another study using electromyography signals [11], we have demonstrated the importance of MV to improve the accuracy of MLs in classification.

Using the open-access EEG data [1], this study aims to investigate the influence of ICA, window size, and MV for both ME and MI motion intention recognition, areas that have not been attempted previously. In addition, the performance of five popular ML is compared.

II. METHODOLOGY

A. Data Description

An open-access EEG data [1] were used in this study. The EEG signals were acquired using 61 electrodes following the 10-20 electrode location system. For the present study, only nine channels were considered: P4, Pz, P3, C4, Cz, C3, F4, Fz, and F3, as most of the brain activities related to movement are contained within these channels [12]. The data were collected from 15 healthy subjects (27±5 years old, nine females); all subjects were right-handed except for one. The subjects performed both ME and MI tasks. The tasks consisted of six motion classes and one rest class. The total length of the acquired EEG signals was five seconds, where at zero second, the subjects were alerted by a 'beep' sound, and at two seconds, a cue was presented on a computer screen to start performing a specific task. Once the cue was shown, the subjects were performed either ME or MI for three seconds. A total of 10 runs per subject and 60 trials per class were recorded. The sampling rate was 512 Hz. An 8^{th}-order Chebyshev bandpass filter from 0.01 Hz to 200 Hz was applied to the dataset. A notch filter was also used to suppress the 50 Hz power line interferences. The dataset is described in more detail in [1].

B. Data Preprocessing

The EEG signals were mean-centered, band passed between 1-40 Hz, and removed from any linear trends. The ICA was used to remove electrooculogram (EOG) artifacts. First, the signals were decomposed into 30 components using fast-ICA algorithm [13]. Each of the components was then compared with three EOG signals. The ICA-component is assumed to be corrupted with eye-blinking artifacts if the correlation value between the component and EOG channels is greater than the mean plus two times the standard deviation. Figure 1 illustrates an example of automatically detecting corrupted components and the reconstructed EEG signals after removing the components.

C. Feature Extraction

The filtered EEG signals were segmented using the sliding window with an 80% overlapping strategy. The window size was varied from 1-2 s. Since the EEG signals are more active in the alpha (8-12 Hz) and beta (13-20 Hz) bands for motor activities [14], the band powers of these frequency ranges for each of the channels were computed. Additionally, two prominent time-domain features were extracted, waveform length (WL) and mean absolute value (MAV). As there are nine EEG channels and four features, the total number of features is 36.

D. Machine Learning

Five different MLs were used for classification, including KNN, LDA, SVM, RF, and NB. The algorithms' tuning parameters were optimized using a grid search approach. The values were 3-NN, 2.63 for kernel scale and 1 for box constraint when using SVM with radial basis function (RBF) kernel, 100 trees for RF, and normal distribution for NB. All the algorithms were trained using nine trials and tested with the tenth trial, and repeated ten times to achieve 10-fold cross-validation. Each algorithm's performance was evaluated by calculating the mean accuracy.

Majority voting (MV) as a post-processing technique was also employed to investigate if the classification accuracy can be improved further. MV considers more than one outcome from a classifier to reduce spurious results [11]. For example, the outcome of a 3-votes MV would be 1 for a predicted class of {1, 2, 1} due to the majority class of 1. A student's paired t-test was used to compare between with and without MV. All analyses were conducted using custom-written scripts in MATLAB.

III. RESULTS AND DISCUSSION

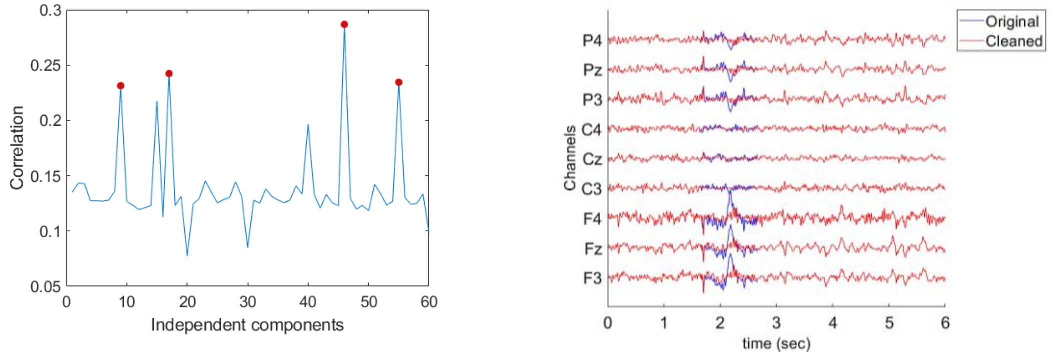
Unlike the previous studies, we consider each of the sliding windows as a representative of rest or movement class after the cue presented to the subjects. The previous studies have used only the first half of the signals after the cue in their MLs, which is not realistic in the BCI application. Different strategies for both pre- and post-signal processing were considered. Results show that the sliding window with 2 s produces the best accuracy for all the ML algorithms (Fig. 2). The improvement in classification accuracies is much greater in MI (>5%) than in ME (<1%) when employing ICA. The MV significantly improves the classification accuracies in both MI (>1%) and ME (>2%) with $p<0.05$ except when used the NB algorithm (Table 1). All the results are further discussed in the subsequent subsections.

A. Classification accuracies with different window sizes

The classification accuracies were increased when the window size was increased (Fig. 2). In general, the accuracy improvement is more prominent in all the ML algorithms between the window sizes of 1.5 s and 2 s. The two best algorithms are RF and SVM that consistently produce better accuracies than the other algorithms in both ME and MI. The ME classification accuracy was improved from 68.9% to 71.5% and 70.8% to 73.8% using SVM and RF, respectively, with MV ($p<0.05$). A similar improvement trend was also observed in both RF and SVM when using with or without ICA and with or without MV in MI (Table 1).

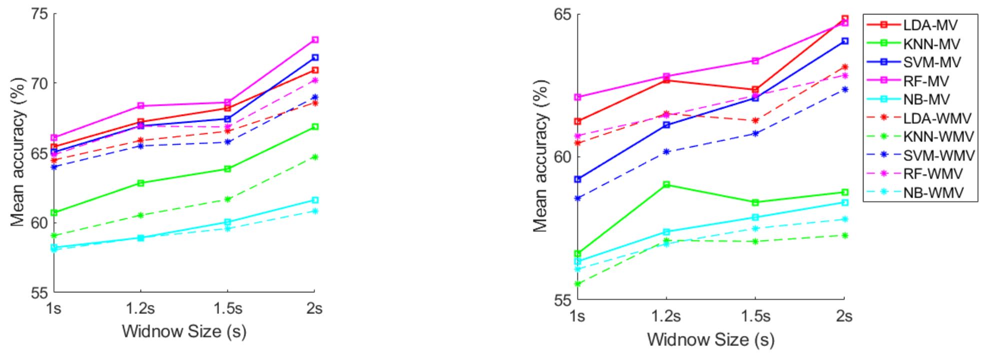
B. Classification accuracies with or without ICA

In most cases, the ICA improves the accuracy values in both ME and MI. However, the accuracy improvement is more prominent in MI than in ME. With ICA, the mean accuracy improvement for all algorithms is 5.6% in MI ($p<0.1$) but negligible in ME (Table 1). This improvement indicates that eye-blinking artifacts are more prominent in the



(a) (b)

Figure 1 EOG artifact removal using ICA, a) identified corrupted components shown in red dots, and b) reconstructed EEG signals (red) plotted on top of original (blue) EEG signals



(a) (b)

Figure 2 Classification accuracies after ICA using different machine-learning algorithms and window sizes, a) motor execution, and b) motor imagination. ICA-Independent Component Analysis, LDA-Linear Discriminant Analysis, KNN-k Nearest Neighbor, SVM-Support Vector Machine, RF-Random Forest, NB-Naïve Bayes, MV-majority voting, WMV-without majority voting.

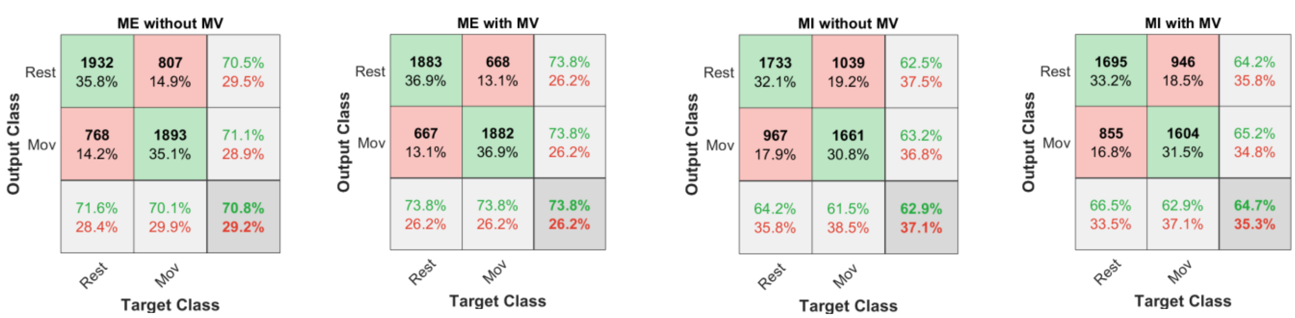


Figure 3 Confusion matrices for RF with a window size of 2 s. ME-motor execution, MI-motor imagination, MV-majority voting, Mov-movement.

TABLE 1 MEAN CLASSIFICATION ACCURACIES FOR MLS WITH AND WITHOUT MV AND ICA. ABBREVIATIONS ARE DEFINED IN THE CAPTION OF FIGURE 2

ML	ME				MI				p values				
	Without ICA		With ICA		Without ICA		With ICA						
	A	B	C	D	E	F	G	H					
	Without MV	With MV	Without MV	With MV	Without MV	With MV	Without MV	With MV	B vs. D	C vs. D	F vs. H	G vs. H	D vs. H
LDA	67.83	70.18	68.59	70.92	56.17	57.37	63.15	64.82	0.590	<0.01	0.024	<0.01	0.049
KNN	64.04	66.25	64.72	66.84	53.06	53.90	57.26	58.76	0.410	0.013	0.065	0.004	0.032
SVM	68.93	71.51	68.98	71.80	57.44	58.69	62.35	64.04	0.630	<0.01	0.051	0.008	0.008
RF	70.83	73.82	70.20	73.10	57.67	59.00	62.85	64.69	0.350	<0.01	0.117	<0.01	0.015
NB	61.63	62.16	60.83	61.63	53.43	53.84	57.81	58.41	0.420	0.170	0.114	0.065	0.065

EEG signals during an imagination task than an execution task.

C. *Classification accuracies with or without MV*

The MV significantly improves the MLs' performance in both ME and MI datasets (Table 1). Overall, the RF with MV had the highest overall mean accuracy (73.8%) with the window size of 2 s (Fig. 3). NB without MV achieved the lowest accuracy of 59% with a window size of 1 s. MV implementation improves the prediction accuracy for both rest and movement classes (Fig. 2). Such improvement is expected as the MV considers successive windows to make a final decision. The reason is that the ML may fail in predicting the true class for a specific segment of EEG signal that may represent abrupt changes. Such a false result usually stands out in a successive prediction. The result indicates that the abrupt changes in EEG signals are higher in ME that may have confounded with the actual limb movement.

D. Classification accuracies for ME and MI

The mean classification accuracies of ME were higher than those of MI for all the classifiers. Without ICA and MV, the initial accuracy for ME was much higher than the accuracy for MI. With ICA, the accuracy improvement in MI (>5%) was much higher than in ME (<1%). With MV, the improvement was greater than 2% in ME and 1.5% in MI except using NB. Higher accuracy for ME are due to the fact that it is much easier to classify actual hand movements than imagined movements, as proven in other studies [1].

E. Classification accuracies from different algorithms

RF consistently performed better among all the MLs (73.8% in ME without ICA and 64.7% in MI with ICA). The worst performance in classifying motion intention was observed for NB (<63% in ME and <59% in MI). Nonetheless, the prediction accuracy for most of the ML was observed to be higher for the movement class than the rest class.

F. Comparing results with previous studies

Although the results of this study cannot be directly compared with the previous studies due to differences in types of classes, the accuracy values presented here are higher than other studies. For example, [1] reported the accuracy as a time series curve and found the peak accuracy of 85% and 73% for ME and MI, respectively, which were drastically reduced to less than 60% just after 0.75 s of movement onset. Such a reduction of accuracies indicated a much lower average accuracy than the present study. Notably, a higher accuracy is achievable only just after the cue-stimulation. However, such a scenario cannot be applicable in a real-time application as there will be no 'cue-stimulation' for a given task. For any EEG-based BCI application, it is important to estimate the expected (average) accuracy of employed ML rather than the peak accuracy since the location of the peak may not be known. In [7], the authors used 1 s of EEG data preceding the actual movement, and in [8], they classified all seven classes and reported 42.6% and 33.1% accuracy for ME and MI, respectively. In addition, the above-mentioned studies used all 61 EEG channels. Conversely, the present study achieved accuracies of 73.8% and 64.7% using only nine channels and RF for ME and MI, respectively, which are representative of motion intention at any moment after the onset, not only within the first second.

IV. Conclusions

This paper investigated the performance of five different MLs (KNN, SVM, LDA, RF, and NB) with window sizes ranging from 1 s to 2 s and MV to classify rest and upper limb movements using only nine EEG channels. The results showed that the ME could be recognized with 73.8% accuracy, while MI could be classified with 64.7% using RF along with MV. The mean accuracy improvement for all algorithms is 5.6% in MI ($p<0.1$) when ICA was employed. These accuracies are significantly higher than the previous studies as it represents the subject's motion intention at any point after the movement onset rather than only within the first second of the ME/MI. The increase of window size improved the accuracies in both the ME and MI where the maximum accuracies occurred at 2 s. A paired t-test showed that the usage of MV significantly improves the accuracy in both ME and MI for all MLs except NB ($p<0.05$). Both RF and SVM performed consistently well for both ME and MI classifications, where the former algorithm yielded maximum accuracies. Classification of more movements will be investigated in future work.

V. References

[1] P. Ofner, A. Schwarz, J. Pereira, and G. R. Müller-Putz, "Upper limb movements can be decoded from the time-domain of low-frequency EEG," *PLoS One*, vol. 12, no. 8, 2017.

[2] M. Salvaris and P. Haggard, "Decoding intention at sensorimotor timescales," *PLoS One*, vol. 9, no. 2, p. e85100, 2014.

[3] J.-H. Jeong, B.-H. Lee, D.-H. Lee, Y.-D. Yun, and S.-W. Lee, "EEG Classification of Forearm Movement Imagery Using a Hierarchical Flow Convolutional Neural Network," *IEEE Access*, vol. 8, pp. 66941–66950, 2020.

[4] I. Navarro, B. Hubais, and F. Sepulveda, "A comparison of time, frequency and ICA based features and five classifiers for wrist movement classification in EEG signals," in *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*, 2006, pp. 2118–2121.

[5] D. S. V Bandara, J. Arata, and K. Kiguchi, "A noninvasive brain–computer interface approach for predicting motion intention of activities of daily living tasks for an upper-limb wearable robot," *Int. J. Adv. Robot. Syst.*, vol. 15, no. 2, 2018.

[6] E. Y. L. Lew, R. Chavarriaga, S. Silvoni, and J. del R. Millán, "Single trial prediction of self-paced reaching directions from EEG signals," *Front. Neurosci.*, vol. 8, p. 222, 2014.

[7] N. Mammone, C. Ieracitano, and F. C. Morabito, "A deep CNN approach to decode motor preparation of upper limbs from time–frequency maps of EEG signals at source level," *Neural Networks*, vol. 124, pp. 357–372, 2020.

[8] Z. Fang *et al.*, "Learning Regional Attention Convolutional Neural Network for Motion Intention Recognition Based on EEG Data."

[9] Y. Jiang, X. Zhang, C. Chen, Z. Lu, and Y. Wang, "Deep Learning Based Recognition of Hand Movement Intention EEG in Patients with Spinal Cord Injury," in *2020 10th Institute of Electrical and Electronics Engineers International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER)*, 2020, pp. 343–348.

[10] D. Zhao, F. Tang, B. Si, and X. Feng, "Learning joint space–time–frequency features for EEG decoding on small labeled data," *Neural Networks*, vol. 114, pp. 67–77, 2019.

[11] M. F. Wahid, R. Tafreshi, and R. Langari, "A Multi-Window Majority Voting Strategy to Improve Hand Gesture Recognition Accuracies Using Electromyography Signal," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 2, pp. 427–436, 2020, doi: 10.1109/TNSRE.2019.2961706.

[12] J. Sleight, P. Pillai, and S. Mohan, "Classification of executed and imagined motor movement EEG signals," *Ann Arbor Univ. Michigan*, vol. 110, 2009.

[13] A. Hyvärinen and E. Oja, "Independent component analysis: algorithms and applications," *Neural networks*, vol. 13, no. 4–5, pp. 411–430, 2000.

[14] K. C. Veluvolu, Y. Wang, and S. S. Kavuri, "Adaptive estimation of EEG-rhythms for optimal band identification in BCI," *J. Neurosci. Methods*, vol. 203, no. 1, pp. 163–172, 2012.