A Deep Learning Algorithm for Classifying Grasp Motions using Multi-session EEG Recordings

Andy Partovi
Department of Biomedical Engineering
The University of Melbourne
apartovi@student.unimleb.edu.au

Farhad Goodarzy
Department of Medicine
The University of Melbourne
goodarzy@unimelb.edu.au

Milad Soleymani
Department of Electrical Engineering
KNTU University of Technology
msoleymani@email.kntu.ac.ir

Seyed Mehrshad Hosseini
Department of Computer Engineering
KNTU University of Technology
hosseini28@email.kntu.ac.ir

Soroush Ziaee
Department of Electrical Engineering
KNTU University of Technology
szsoroush@email.kntu.ac.ir

Kiana Liaghat
Department of Electrical Engineering
KNTU University of Technology
k_liaghat@email.kntu.ac.ir

Sahand Sadeghpour Vajdi
Department of Electrical Engineering
KNTU University of Technology
sahandss@email.kntu.ac.ir

Erfan Habibi Panah Fard
Department of Electrical Engineering
KNTU University of Technology
erfanhabibi@email.kntu.ac.ir

***Abstract*—The classification of motor imagery tasks using scalp EEG signals is a complicated procedure in BCI especially when the task comprises multiple gestures of the same hand. In this paper, we present a classification method to distinguish three grasp motion classes (cylindrical, spherical, and lumbrical) of one hand over two day training sessions in 15 subjects in a public dataset. We have developed Two ensemble methods consisting of (anomaly detection + fully connected neural network) and (anomaly detection + convolutional neural network) to classify grasp motion and have achieved more than 80% classification accuracy in 3 subjects and an average accuracy of 57% among the full cohort. Our results confirm the possibility of utilizing neural networks to decode motor movement intentions from scalp EEG in a complicated task.**

Keywords-component; BCI; Deep learning; Anomaly Detection; EEG decoding; Neural network; CNN

I. INTRODUCTION

Brain-Computer Interfaces (BCIs) have been used as a communication bridge between an individuals' brain and external electronic devices like computers. They can be a powerful tool for people who have lost communication or movement abilities due to illnesses or injuries. But the usefulness of BCI systems heavily rely on their accuracy in detecting what the user intentions are and conveying those to external devices. This accuracy is often affected by the complexity of the intended task and the limitations in BCI hardware and software. In this research, we explore statistical methods for increasing the accuracy of a complex motor imagery task of distinguishing three forms of grasp. Motor Imagery (MI) is the task of imagining or simulating an action. In order to enable the communication of user's intention, first brain signals must be acquired.

Scalp Electroencephalography (EEG) is a non-invasive method to record the brain's activity and is commonly used for its accessibility, mobility and relative ease of use. However, EEG has the major shortcoming of variations between days of recording due to slight changes in the electrode positions. This adds to the complexity of the attempted task as the training and test data belong to two separate recording sessions across two days.

In order to process EEG signals, for a BCI application, several preprocessing steps (e.g. filtering, feature extractions, and selection) are required before passing the results through a classifier. In previous studies, a novel feature extraction method was considered which is extracting statistical features from EEG signals that depict the variation of labeled EEG signals [1]. Recently, The Common spatial pattern (CSP) followed by the Support Vector Machine (SVM) or Linear Discriminant Analysis

978-1-7281-4707-9/20/$31.00 ©2020 Crown

(LDA) as classifiers was utilized for classifying motor imagery tasks [2]. A challenging characteristic of MI which makes their decoding difficult, is the unknown exact time of imagining movement. In this research, we are using statistical methods to detect the movement time.

Anomaly detection is often used for detecting outliers in a given dataset. Outliers are known as a data object that is different from the rest of the population in some regard. Anomaly detection paradigms have gained attention in a diverse range of studies [3]. As mentioned, in MI tasks, the exact time of action is not known a priori, however the non-action period can be considered and "anomaly" compared to the rest of the action period. We have used anomaly detection, which depends on the mean and variance of input features for distinguishing the Rest and Non-rest periods of an MI task. However, this alone is not sufficient in distinguishing multiple grasp types.

Recently, deep neural network approaches are used for decoding complex patterns such as brain activities [4, 5, 6]. Convolutional neural networks are one of the deep neural network architectures that derive from the idea of feed-forward networks and are mostly used for image classification. CNN models can extract features automatically which is one of the advantages of this method. In our research, we have used both a feed-forward fully-connected and a CNN architecture to classify different grasping motions. We have experimented different neuron density, number of hidden layers, kernel sizes, and dropout percentage in these architectures. The hyper-parameter tuning is performed on each model per subject to determine the best parameter for them.

The rest of this paper is structured as follows. First, in section II, we describe the dataset that was used in this study, the 2020 International BCI competition, Also, we explain our preprocessing and feature extraction methods as well as the neural network model in detail. In section III & IV, The result of each subject, concluding remarks and recommendations for future development are presented in section IV.

II. METHODS

A. EEG dataset

In order to test the performance of our methods, we used the publicly available 2020 International BCI competition dataset available at https://osf.io/pq7vb/. This dataset was chosen for two main reasons: 1) the attempted motor imagery task is a complex movement in only one hand, as opposed to the common eft vs right movements. 2) the recording was done over two days with natural variations in impedance and cap position.

In this dataset, 60 scalp EEG electrodes were placed on 15 people aged 20 to 30, who were all right-handed and they were asked to imagine three different grasps, Cylindrical, Spherical, and Lumbrical. The raw data was recorded using the Brain Vision (Brain Product Gmbh, Germany) with MATLAB 2019a (The MathWorks Inc., USA). Each subject performed 50 trials over two sessions on two different days. A single trial lasted 10 seconds and had three sub-stages; the first three seconds was relaxation, the second three seconds preparation and cue, and the last 4 seconds motor imagery. It is important to note that the participant was given a cue on which grasp to perform in the last 4 seconds. The procedure is depicted in Fig I. We used the data from the first day as the training data and the second day as the test data.

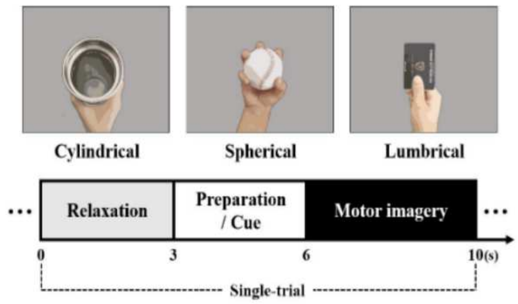

Figure I. Illustration of the three grasp tasks and experimental protocol.

B. Data preprocessing and preparation

As the first step of preprocessing, we applied a 60Hz notch filter to remove the power-line artifact. Next, two different approaches were taken to further process the data. One was to apply a 1Hz high pass filter to the notched signal to remove the DC component and the other was to pass on the signal without any further filtering.

Then the signal was binned into segments of 1 second length (equal to 250 samples) and then windowed to 0.1 sec (25 sample) bins. The binned data was passed on for feature extraction where the following statistical features and frequencies were considered [1, 2, 3]. The statistical features are collected from EEG signals to detect variation of grasp motions. In addition, due to nonlinearity and non-stationary characteristics of EEG signals, the statistical features are a simple way to determine variation tendency of signal.

1. **Statistical features:**

- **Coefficient of variation (CV):**

The CV is a statistical feature that defines a standardized measure of the dispersion of a probability distribution. The CV is the ratio of standard derivation(σ) and average of signal (μ).

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N} \tag{1}$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N}} \tag{2}$$

$$CV = \frac{\sigma}{\mu} \tag{3}$$

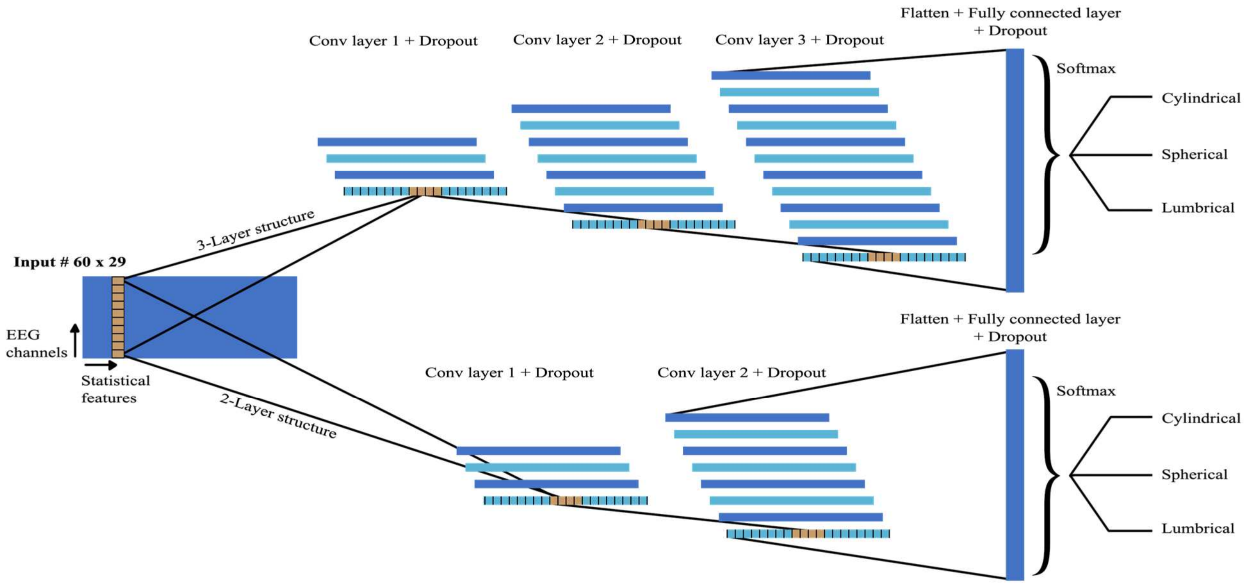

Figure II. CNN model architecture

- **Kurtosis:**

The kurtosis feature determines how the distribution tail varies from the normal distribution.

$$Kurt = \frac{1}{N} * \frac{\sum_{i=1}^{N}(x_i - \mu)^4}{\sigma^4} \qquad (4)$$

- **Skewness:**

The skewness feature is a measurement of Signal symmetry.

$$Skewness = \frac{1}{N} * \frac{\sum_{i=1}^{N}(x_i - \mu)^3}{\sigma^3} \qquad (5)$$

- **Hjorth:**

Hjorth included 3 components that define the activity, mobility, and complexity of the signal. Each component describes a specific property of signal e.g. activity measures the power of signal, mobility determines the mean frequency of signal, and complexity captures the change in the frequency.

$$Activity = Var(x) \qquad (6)$$

$$Mobility = \sqrt{\frac{Var(\dot{x})}{Var(x)}} \qquad (7)$$

$$Complexity = \sqrt{\frac{Mobility(\dot{x})}{Mobility(x)}} \qquad (8)$$

2. **Frequency features:**

- **Power Spectral Density:**

The classical method for computing PSD is via Fourier transform however, we used the Welch periodogram for calculating PSD Because of the non-stationary nature of the EEG signal and since EEG signals cannot be expressed as a combination of pure tones so calculating the PSD via Fourier transform might not be reliable.

- **Autoregression:**

Autoregression (AR) is a machine learning model that fits a model to the previous signal and predicts the feature behavior of the signal. There are several methods for calculating autoregression features; one of them is the burg method that fits an autoregressive model on the signal by minimizing the forward and backward prediction errors.
The burg AR method is a parametric scheme for frequency-domain and time-domain signal processing that fits an AR model to smooth spectra in comparison with frequency-based paradigm in order to figure out characteristic frequencies [10].

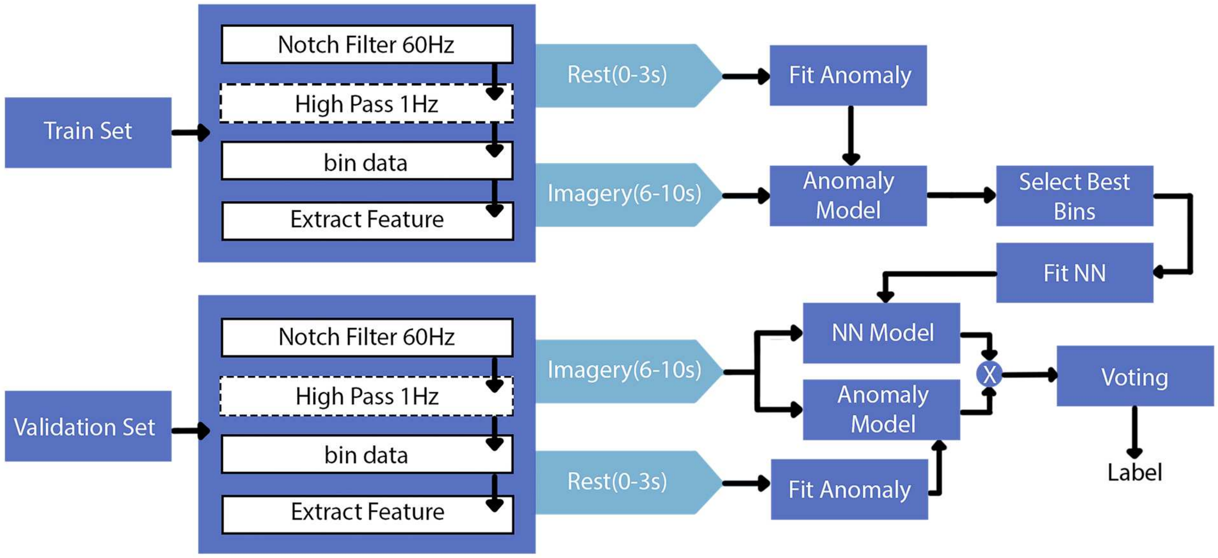

Figure III. Ensemble method for classifying grasp action

- **Entropy and Root Mean Square:**

We also calculate the entropy and RMS of the signal via a discrete Fourier transform (DWT). The original shape of EEG signal is 150*60*2501 (trials, channels, times), after applying segmentation to the signal the rest part of signal transformed to 150*21*60*250 (trials, bins, channels, times) and the imaginary part of signal changed to 150*31*60*250 (trials, bins, channels, times). Then the signal passed through our feature extractor and extracted 29 features from each channel. The final shape of the rest part is 3150 * 1740 and the imaginary part is 4650 * 1740.

There are 29 features per each bin of each electrode resulting in 1740 features in total. In order to reduce this feature set to avoid model overfitting, we utilized a principal component analysis (PCA) method and retained 12 principal components based on our variance analysis.

C. Experimental framework

In order to account for the variations between the two recording sessions, we designed an algorithm that uses a resting period's data to identify important time bins in the data and trained neural network models only on those bins (see Figure III). The algorithm has the following components:

1. Train the anomaly detection model with the rest features of the training set and set a threshold to distinguish rest and non-rest.
2. Apply anomaly detection model to the imagery features of the training set to select the non-rest bin of imagery features.
3. Feed the selected bins from the previous step to the classification model.
4. Train the anomaly detection model with the rest features of the validation set.
5. Apply anomaly detection model to the imagery features of validation set and get the probability of being rest or non-rest features.
6. Feed the imagery features of validation set to the trained classification model and get the probability of each class.
7. Multiply the probability of the anomaly detection model and the classification model summed over each bin of the trial to predict the label of that trial.

- **Anomaly detection**

We applied a Gaussian anomaly detection model to produce a threshold probability for detecting rest data from MI. The anomaly detection was trained on the rest section of each trial (the first 3 seconds) and then applied to the MI section to detect the most informative sections of those bins (the last 4 seconds of each trial).

- **Neural Network Classifier**

Once we detected the non-rest sections of the EEG signals in the MI section, we used two architectures for our neural network (NN) classifier and fine-tuned them with a grid search method for learning rate and drop out parameters. Models were trained with both the high passed data and the non-filtered data. The probability of each bin belonging to each class was measured using these models and the final prediction.

The first NN was two fully-connected feed-forward network with 2 and 3 hidden layers of activated neurons following by SoftMax output layer.

Table I. CLASSIFICATION ACCURACY OF GRASP MOTION ACTIONS
(NN1 2 HIDDEN LAYERS, NN2 3 HIDDEN LAYERS, CNN1 2 CONV LAYERS, CNN2 3 CONV LAYERS)

Participant	NN				CNN			
	NO Filter		*Hp Filter*		*NO Filter*		*Hp Filter*	
	NN1	*NN2*	*NN1*	*NN2*	*CNN1*	*CNN2*	*CNN1*	*CNN2*
1	62	58	56	62	**80**	75	72	73
2	64	63	75	69	89	**93**	81	77
3	48	43	47	46	49	49	50	**51**
4	92	92	75	77	**95**	**95**	87	87
5	43	44	47	43	**51**	46	46	47
6	40	38	44	39	**47**	**47**	45	44
7	41	39	41	42	47	47	**49**	47
8	45	44	42	41	**54**	51	49	47
9	42	44	41	40	**47**	45	**47**	46
10	41	44	42	43	**51**	49	47	49
11	40	40	40	42	**49**	**49**	47	47
12	54	**58**	42	42	57	56	49	50
13	57	58	52	51	**70**	67	59	58
14	67	69	61	63	73	**75**	66	65
15	60	61	50	50	**64**	62	56	55
Avg.	53.07	53	50.33	50	**61.55**	60.49	56.75	56.35
SD	14.44	14.71	11.7	11.99	16.23	16.83	13.65	13.26

- **Convolutional Neural Network**

The second NN architecture is two variations of the CNN architecture presented in [11, 12]. The Kernel of these NN cover the entire set of electrodes in each time step to preserve the interdependencies between the channels. One of the architectures is a two 2D Convolution layers with ReLU activated neurons, followed by a drop out (50% and 25%) and includes a flattened layer and at the end a Softmax layer to predict probability of each class. The other architecture is the same as the last one but the number of Conv2d layers on the second architecture is three. At first the MI features reshape to 4650*60*29*1 (trials * bins, channels, features, filter), then fed to CNN model and finally predicted the probability of each class. The model architecture depicted on Fig II and implemented using the Keras library.

The classification models are fine tuned using grid search to optimize the number of neurons in the dense and convolutional layers, the dropout percentage, the learning rate for the models, first layer , and the kernel window width.

III. RESULTS

As it's shown in Table I, an accuracy of 95% is achieved for the best performing participant in the cohort, using non-filtered data, and the both CNN models. The CNN models performed better than NN on all the subjects with the 2-layer CNN model showing the best performance on average. The results suggest that the model performs differently for different participants. However, on average, the non-filtered data works better. This can save processing time for larger data sets and online algorithms. Considering the extremely limited training data set and the fact that training set and validation set belong to 2 different days, we anticipate our models can achieve better performance once they are trained on longer datasets.

IV. CONCLUSION

In this experiment, we attempted a complex task of distinguishing three grasp motions in one hand, with the test and training data belonging to two separate recording sessions. We tested a combination of anomaly detection with feed forward and CNN models for this MI task. Despite the the small size of training data in this dataset, we achieved a maximum accuracy of 95% and an average of 61% with non-filtered data and the 2-layer CNN model. Different types of preprocessing, other methods of anomaly detection such as outlier or novelty detection and the use of more complex classification models such as RNN are planned fort further investigation and future work to improve the overall performance of the system.

REFERENCES

1. Hong He, Yonghong Tan, Jun Ying, Wuxiong Zhang, "Strengthen EEG-based emotion recognition using firefly integrated optimization algorithm", Applied Soft Computing, Volume 94, 2020

2. Zhichuan Tang, Chao Li, Shouqian Sun, "Single-trial EEG classification of motor imagery using deep convolutional neural networks" Optik, Volume 130, 2017.

3. Varun Chandola, Arindam Banerjee, and Vipin Kumar. 2009. Anomaly detection: A survey. ACM Comput. Surv. 41, 3, Article 15 (July 2009), 58 pages.

4. Yin Z and Zhang J 2017 Cross-session classification of mental workload levels using EEG and an adaptive deep learning model Biomed. Signal Process. Control 33 30–47.

5. Yin Z, Zhao M, Wang Y, Yang J and Zhang J 2017 Recognition of emotions using multimodal physiological signals and an ensemble deep learning model Comput. Methods Programs Biomed. 140 93–110.

6. Decoding EEG and LFP signals using deep learning: heading TrueNorth E Nurse, BS Mashford, AJ Yepes, I Kiral-Kornek... - Proceedings of the ACM International Conference, 2016

7. Roy, Y., Banville, H., Albuquerque, I., Gramfort, A., Falk, T. H., & Faubert, J. (2019). Deep learning-based electroencephalography analysis: a systematic review. Journal of neural engineering.

8. V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, B. J. Lance, Eegnet: A compact convolutional network for eeg- based brain-computer interfaces, CoRR abs/1611.08024.

9. Voulodimos, Athanasios, Nikolaos Doulamis, Anastasios Doulamis, and Eftychios Protopapadakis. "Deep learning for computer vision: A brief review." Computational intelligence and neuroscience 2018

10. Wei-Ling Chen, Chia-Hung Lin, Taisong Chen, Pei-Jarn Chen, Chung-Dann Kan, "Stenosis Detection using Burg Method with Autoregressive Model for Hemodialysis Patients", 2012

11. Andishe Partovi, Farhad Goodarzi, Kiana Liaghat, Seyed Mehrshad Hosseini, Soroush Ziaee, Milad Soleymani, Erfan Habibi Panah Fard, Sahand Sadeghpour Vajdi, "Optimizing a Convolutional Network Model in Hnad-squeeze Task Classification", 2020

12. Partovi, Andi, Farhad Goodarzy, Ewan S. Nurse, Philippa Karoly, Dean Freeston, Mark Cook, Anthony N. Burkitt, and David Grayden. "A Convolutional Neural Network Model for Decoding EEG signals in a Hand-Squeeze Task." In 2020 8th International Winter Conference on Brain-Computer Interface (BCI), pp. 1-6. IEEE, 2020.