Optimizing a Convolutional Network Model in a Hand-squeeze Task Classification

Andi 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

Kiana Liaghat
Department of Electrical Engineering
KNTU University of Technology
k_liaghat@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

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

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

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

Abstract - **motor imagery based BCI can help individuals with movement disorders such as locked-in syndrome, tetraplegia, cerebral palsy by developing a correlation between brain and external devices, for example, speller, wheelchair, and prosthetics. inspired by recent BCI research, in this paper, a convolutional neural network is applied to EEG signals collected through hand squeeze tasks both in terms of classifying and evaluating the accuracy. In our approach, we reduce the number of channels, besides, practicing various hyperparameters such as learning rates, batch-size, number of epochs to investigate results on classification accuracy. All the above tasks applied on 2-labeled and 3-labeled datasets to compare the accuracies. The result proposes that a simple architecture and using a combination of influential channels rather than all channels indicates a proper classification accuracy and can be used in further researches.**

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

I.INTRODUCTION

the systems based on Brain Computer-interface (BCI) creates a correlation between human brain and external devices. In recent research, a non-invasive scheme such as an electroencephalogram (EEG) considered for BCI tasks. [1-4].

EEG requires inexpensive equipment to implement, besides since EEG provides a safe and comfortable way to use for almost all people and it enables disabilities using brain waves to control their bodies, that make the EEG method popular in BCI researches. [5-8].

As The Motor Imagery(MI) area is related to body movements that contain event-related synchronization/desynchronization and the effect of unilateral limb motions on mu and beta rhythms in the sensorimotor area of the contralateral hemisphere and the ipsilateral hemisphere, respectively [11-13], The MI has become a communication system to control computers by applying MI task of the brain without actual movement of subject's body [9,10].

Many feature extraction and classification schemes have been tested for motor imagery task recognition. The procedure of building a relationship between EEG features and MI tasks is an organized way. First of all, designing a proper input features and then applying machine learning algorithms to create the correlations [14,15].

The various algorithms for feature extracting and classifying MI-based BCI systems used in many cases. the common spatial pattern (CSP) is a popular feature extraction method applied in many types of research [16,17]. Besides, the other algorithms like support vector machine (SVM) [26], Linear discriminant analysis (LDA) [27], Independent component analysis (ICA) [19] and Bayesian classifier [28] implemented in different studies [29,30], Another approach in MI-based BCI systems are increasing the precision of classifiers. For instance, The Principal component analysis [20] is investigated to enhance classification accuracy [21]. In recent studies, algorithms like sparse Bayesian extreme learning machine [22], sparse group representation model [23]and Jaya based adaptive neuro-fuzzy classifier (NFC)[24] applied to improve classification accuracy, moreover, the tensor decomposition method is used for motor imagery feature extraction and classification to improve the classification efficiency [25].

The Convolutional neural network (CNN) is a new approach in feature extracting and classifying tasks like image and video recognition, Natural language processing (NLP), recommender systems, signal classifying, etc.

the convolutional neural network is biological-inspired variants of multilayer perceptron design to use the minimal amount of preprocessing [33]. recently a CNN-based classifier is used for detecting p300 waves which proposed the highest accuracy of 95% on a data set from BCI competition [34].

For instance, three CNNs with various architectures used for classifying MI EEG signals, which contain deep ConvNet and residual network (ResNet). recently a combination of the Spatio-temporal characteristic of EEG signal and deep neural network executed to extract features and classify MI tasks that shown 86.41% accuracy in classifying signals [36].

A new convolutional neural network scheme performed for classifying EEG signals captured from the motor imagery area by using a wavelet time-frequency image. Since CNN has high accuracy in image recognition, a 2-dimensional image created by applying a continuous wavelet transform (CWT) to feed into CNN. [37].

In recent research, a convolutional neural network applied to learn the activation pattern of different motor imagery signals, to improve the classification of a variational autoencoder (VAE) used with 5 hidden layers [38].

The CNN architecture used for classifying MI tasks that regarded as a series of images. The method considered the spatial correlations among neighboring EEG channels by accounting for the arrangement of electrodes on the scalp [39].

In recent researches, Wavelet neural network (WNN) is implemented to train and classify two classes of motor imagery signals by using complex Morlet wavelet to transform EEG data into tensors as inputs of the neural network. The wavelet neural network is a new type of neural network using wavelets instead of the convolutional layers [31].

The researchers have compared CNN and Long Short-Term Memory (LSTM) methods to extract features and classify MI BCI tasks [32] which are proposed CNN method is more efficient than the LSTM method.

II.METHODS

A.EEG data acquisition.

The data which collected from EEG where planted in the scalp of the five participants that we named them from A to E, They performed a hand squeeze task under the self-pace circumstances with ethics approval (ID 1339680, Human Research Ethics Committee, University of Melbourne) that we have. All participants had to be written consent. Except participant E, the rest of the participants were right-handed. Neither of the participants had an experience with BCIs. The data from participant A could not be recovered because of the corruption in the data, for that we did not include participant A result in our results.

The data acquired from EEG that using a SynAmps2 64 channel Quik-Cap and a SynAmps2 24-bit amplification system (Compumedics Ltd, Australia), All the EEG that we have just use for some participants and we didn’t use them for all participants.

Hand motion time is the same as EEGs with simultaneous motion electromyography (EMG) using the same access system. In their forearms we attached the surface EMG electrodes to the flexor digitorum. With using window method, Notch filtered at 50 Hz, bandpass filtered between 0.1-100Hz and the data was sampled at 1 kHz. EEG and EMG electrodes was implemented in CURRY 7(Compumedics Ltd) which were first tapered using a Hamming window and then filtered with a zero-phase delay FIR filter (Hann window).

We asked participants to alternate clenching their right and left hands at the constant pace which they have, with their eyes open and focusing their eyes on a point on the wall while they seated. No stimuli were given to stimulate the participant to shake hands in order to reduce the effect and minimize the influence of visual or auditory sensations and evoked potential in the training data. Participants held a foam squeeze ball in their both hands. Recordings were in four-minute sessions between each resting session. Generally, data was recorded for a total of 20-30 minutes. EEG data recorded before network training was reduced to 250 Hz.

B. Data preprocessing and preparation

In order to label EEG data, RMS power of EMG signals were used which computed in consecutive

500ms of the signal. Detection of movement depended on the RMS power crossing a certain threshold. This threshold which may vary subject to subject was determined by comparing baseline RMS power before the trial and during the clenches. To ensure minimal inconsistency in the threshold, participants were told to avoid moving their arms between actions and perform similar hand-squeezes.

After labeling, EEG data were divided into epochs with window width of 400ms, consisting 100ms before and 300ms after the start of hand-squeezing. The reason behind including 300ms after the start in each epoch was to avoid detecting a movement twice. Between clenches, times that weren't labeled as movement for at least 500ms were used as rest or no-movement signals with the same 400ms of window width in EEG epochs.

C. Classification model

Since one of our tasks is comparing accuracies among 2-labeled and 3-labeled datasets, we apply a specific classifier, based on our number of classifications, to our datasets.
A 2-dimensional matrix where the rows demonstrate time and the columns represent the number of channels is fed to the classifier.
CNNs are the biological-inspired variants of multilayer perceptron design and subtype of feed-forward neural network. CNNs consist of two types of layers: 1) The convolutional layers, which create numerous smaller matrices from the input matrix that use as a classification feature.
2) The fully connected layers, which perform the classification task on the output of the final convolutional layer in the shape of a flattened-out feature vector.
Alike the local receptive field in visual systems, the convolutional layers convolve a kernel window on the input with a certain step size(stride) to extract features from the input, besides, multiple kernels can implement directing to multiple features maps built from the same input in each convolution layer. To make convolution faster than the feed-forward layer that needs to learn weights for a separate group of neurons, the constant weights associated with kernels for each feature map.
The outputs of the convolutional layer combined, besides, the kernel size increased in each sequential convolutional layer to learn more high-level features from the output of each convolutional layer.
The fully connected layers that perform the classification and create a probability vector which demonstrates the probability of each classification labeled for the given input, connected to all output of their previous layer.
we applied a high-level CNN architecture that illustrated in figure1. For extracting features from our datasets, two sequential convolutional layers, with k and 2k of Relu-activated neurons implemented and for classifying a fully connected layer of m Relu-activated neurons followed by m SoftMax-activated neurons for 3-labeled datasets or Sigmoid-activated single neuron for 2-labeled datasets used in the output layer.
In contrast with image processing and computer vision tasks where use a small kernel window, we used a window of width **w** and height equal to the number of EEG channels to capture dependencies between all channels instead of the local dependencies between channels that happen to be close in the input data.
Both the order in which channels are presented to the classifier and optimal channel order do not affect this method. To examine our theory, we executed several shuffles on our input channels. The result approves that channel order does not affect the performance.
Pooling layers are sub-sampling layers in image recognition CNNs that aggregate a group of neurons (using operations like average or max) and use that as the representation of that group thereby reducing the size of the model after application. Our observation emphasized that add a pooling layer between the convolutional layers will dramatically decrease the classification performance so we decided not to use the pooling layer in our architecture.
We implemented and trained the model in the Keras framework. The training method used was backpropagation with stochastic gradient descent using various batch sizes, learning rates and epochs. Therefore, in each of the training epochs, 100 random samples from the input were selected for training.
To avoid over-fitting and reduce neuron co-dependency, we "shut down" 25% of each neuron randomly during the training (The weights of these neurons are not updated in that specific epoch) by implementing drop out in all of the three hidden layers of the neural network.

D. Experimental framework

A hypermeter is a not learned parameter whose value is used to control the learning process. And grid search is a comprehensive search in conjunction with a valuation framework, used to find a value for a hypermeter in which the network has optimal performance.

We used grid search with a 5-fold cross validation framework to find a 6member optimal hypermeter set for our CNN:

1. First layer neuron count changes in a range of 3 to 11
2. Density number changes in a range of 50 to 550 with a step of 100

3. The width of the kernel window has a value from 1 to 11 with a step of 2
4. Batch size changes in a range of 70 to 110 with a step of 10
5. Epoch size changes in a range of 70 to 120 with a step of 10
6. The learning rate has a value between the numbers 0.01, 0.03, 0.1, 0.3

In each iteration, 20% of the data for testing the performance -which is not used during the training - and measured the classification performance in terms of accuracy.

Then we fixed the best set to the network and started testing combinations with fewer channels involved. For this matter, we made a set of 10 most influential channels for each patient, selected combinations of 2 and more

III. RESULTS

All the experiments were accomplished once for 2 labeled data and once for 3 labeled data.

A. *Optimizing the architecture*

On the first phase of the experiment, we performed 2 grid searches for each participant, first grid search for finding the best set of first layer neuron count (f), density (d) and window width (w) and the second search for batch size (b), epoch size (e), and learning rate (l), in order to reduce runtimes.

In order to reduce runtimes, we ran a grid search for three hypermeters of first layer neuron count (f), density (d) and window width (w). after finding best set of (f,d,w), we ran another grid search to find the best batch size (b), epoch size (e), and learning rate (l).

TABLE I. HYPERPARAMETER SETS THAT YIELDED THE BEST AND WORST PERFORMANCE WITH 2 CLASSES

patient	Best case	Worst case
B	f=8, d=50, w=5, b=70, e=110, l=0.01	f=128, d=50, w=1, b=100, e=100, l=0.3
C	f=8, d=50, w=1, b=90, e=100, l=0.03	f=64, d=50, w=11, b=80, e=90, l=0.3
D	f=8, d=150, w=1, b=100, e=80, l=0.01	f=128, d=50, w=11, b=70, e=70, l=0.1
E	f=8, d=50, w=11, b=90 e=90, l=0.01	f=128, d=50, w=7, b=100, e=100, l=0.3

TABLE II. HYPERPARAMETER SETS THAT YIELDED THE BEST AND WORST PERFORMANCE WITH 3 CLASSES

patient	Best case	Worst case
B	f=1, d=250, w=9, b=90, e=110, l=0.01	f=10, d=50, w=3, b=110, e=110, l=0.01
C	f=3, d=150, w=4, b=110, e=120, l=0.01	f=6, d=50, w=11, b=70, e=100, l=0.1
D	f=1, d=450, w=6, b=70, e=120, l=0.03	f=8, d=250, w=11, b=80, e=110, l=0.3
E	f=4, d=350, w=11, b=100, e=120, l=0.01	f=3, d=50, w=1, b=110, e=70, l=0.3

B. *Reducing channels*

On the second phase, we selected 10 most influential channels for each patient (except for patient E whose total channel numbers were 10 and we chose 5 among them) which were the closest channels to motor imagery.

We calculated the accuracy for these sets and then repeated the experiment for combinations of two or more channels and kept on enlarging them until we got an accuracy equal or very close to the 10 channels state.

TABLE III. HIGHEST ACCURACYS FOR COMBINATIONS WITH 2 CLASSES

Patient / Channel numbers	B	C	D	E
2	%71	%81	%78	%78
3	%77	%84	%84	%83
4	%79	%86	%85	%84
5	%84	%85	%86	%85
10	%84	%83	%87	%86
All channels	%80	%88	%89	%86

TABLE IV. HIGHEST ACCURACYS FOR COMBINATIONS WITH 3 CLASSES

Patient / Channel numbers	B	C	D	E
2	%57	%59	%51	%67
3	%57	%60	%57	%74
4		%61	%60	%76
5		%63	%62	%76
10	%58	%61	%63	%81
All channels	%56	%66	%72	%81

Based on these findings, we can conclude that using fewer channels close to the Motor imagery reduces the computational cost and also gives a proper accuracy.

C. *Comparison between 2 and 3 classes classification*

As it was mentioned, all the experiments were accomplished once for 2 labeled data and once for 3 labeled data. Table V shows a comparison between the results of the experiments in part A.

TABLE V. COMPARISON BETWEEN 2 AND 3 CLASSES RESULTS

Accuracy / Patient	2 classes		3 classes	
	Average	Variation range	Average	Variation range
B	%61	%19	%52	%4
C	%78	%10	%57	%9
D	%68	%21	%63	%9
E	%73	%13	%67	%14

It can be construed from the table that a classification with 2 classes has shown a higher average performance than 3 classes which can be cause of uneven sample sizes of the data. But a 3 labeled classification has a lower variation range and the average accuracy is much closer to the maximum than a 2 labeled classification.

IV. CONCLUSION

In conclusion, the classification accuracy affected by changing hyperparameters indicates that the simpler our architecture is the more efficient our classification is, besides, in some cases using the different combination of channels has improved the classification precision but mostly no changes have been observed on precision. The result demonstrates some channels maintain critical data. All things considered, ‹The outcomes represent the accuracy of 2-labeled is approximately better than 3-labeled datasets, moreover, the difference between these two types of datasets may be the effect of unbalancing data in 3-labeled datasets. Finally, our research proposes that the classification accuracy of a simple architecture and using few influential channels can be as efficient as a complicated neural network and using the whole channels of EEG.

REFERENSES

[1] N. Birbaumer, L.G. Cohen, Brain–computer interfaces: communication and restoration of movement in paralysis, J. Physiol. 579 (3) (2007) 621–636.

[2] B. Varkuti, C. Guan, Y. Pan, K.S. Phua, K.K. Ang, C.W.K. Kuah, K. Chua, B.T. Ang, N. Birbaumer, R. Sitaram, Resting state changes in functional connectivity correlate with movement recovery for BCI and robot-assisted upper-extremity training after stroke, Neurorehabil. Neural Repair 27 (1)(2013) 53–62.

[3] L.M. McCane, S.M. Heckman, D.J. McFarland, G. Townsend, J.N. Mak, E.W. Sellers, D. Zeitlin, L.M. Tenteromano, J.R. Wolpaw, T.M. Vaughan,P300-based brain–computer interface (BCI) event-related potentials (ERPS): People with amyotrophic lateral sclerosis (ALS) vs. age-matchedcontrols, Clin. Neurophysiol. 126 (11) (2015) 2124–2131.

[4] B. Z. Allison, E. W. Wolpaw, and J. R. Wolpaw, “Brain-Computer Interfaces System: Progress and Prospects,” Expert Rev. Medical Device, vol.4, no. 4, pp. 463-474, 2007.

[5] I. Milovanovic, R. Robinson, E.E. Fetz, C.T. Moritz, Simultaneous and independent control of a brain–computer interface and contralateral limbmovement, Brain Comput. Interfaces 2 (4) (2015) 174–185.

[6] E. Yin, Z. Zhou, J. Jiang, F. Chen, Y. Liu, D. Hu, A novel hybrid BCI speller based on the incorporation of SSVEP into the P300 paradigm, J. Neural Eng. 10(2) (2013) 026012.

[7] M. Wang, I. Daly, B.Z. Allison, J. Jin, Y. Zhang, L. Chen, X. Wang, A new hybrid BCI paradigm based on P300 and SSVEP, J. Neurosci. Methods 244 (2015)16–25.

[8] S. Gao, Y. Wang, X. Gao and B. Hong, “Visual and Auditory Brain-Computer Interfaces,” in IEEE Transaction on Biomedical Engineering, vol 61, no 5, May 2014.

[9] J.S. Brumberg, J.D. Burnison, K.M. Pitt, Using motor imagery to control brain–computer interfaces for communication, in: International Conference onAugmented Cognition, Springer, 2016, pp. 14–25.

[10] Hyeon Kyu Lee, Young-Seok Choi,” A Convolution Neural Networks Scheme forClassification of Motor Imagery EEG based onWavelet Time-Frequecy Image”.

[11] Pfurtscheller, G.; Da Silva, F.L. Event-related EEG/MEG synchronization and desynchronization: Basic
principles. Clin. Neurophysiol. **1999**, 110, 1842–1857.

[12] Tang, Z.; Sun, S.; Zhang, S. A Brain-Machine Interface Based on ERD/ERS for an Upper-Limb Exoskeleton
Control. Sensors **2016**, 16, 2050.

[13] M. Pregenzer, G. Pfurtscheller, Frequency component selection for an EEG-based brain to computer interface, IEEE Trans. Rehabil. Eng. 7 (4) (1999)413–419.

[14] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, B. Arnaldi, A review of classification algorithms for EEG-based brain–computer interfaces, J. Neural Eng.4 (2) (2007) 1–24.

[15] N. Naseer, K.-S. Hong, FNIRS-based brain–computer interfaces: a review, Front. Hum. Neurosci. 9 (2015) 3.

[16] Ramoser, H.; Muller-Gerking, J.; Pfurtscheller, G. Optimal Spatial Filtering of Single Trial EEG During
Imagined Hand Movement. IEEE Trans. Rehabil. Eng. **2000**, 8, 441–446.

[17] Xygonakis, I.; Athanasiou, A. Decoding Motor Imagery through Common Spatial Pattern Filters at the EEG
Source Space. Comput. Intell. Neurosci. **2018**, 2018, 7957408.

[18] Lu, N.; Li, T.; Ren, X.; Miao, H. A Deep Learning Scheme for Motor Imagery Classification based on Restricted

Boltzmann Machines. IEEE Trans. Neural Syst. Rehabil. Eng. **2017**, 25, 566–576.

[19] P. Comon, ``Independent component analysis, a new concept?" *Signal Process.*, vol. 36, no. 3, pp. 287_314, Apr. 1994.

[20] P. J. García-Laencina, G. Rodríguez-Bermudez, and J. Roca-Dorda, ``Exploring dimensionality reduction of EEG features in motor imagery task classi_cation," *Expert Syst. Appl.*, vol. 41, no. 11, pp. 5285_5295, 2014.

[21] X. Guo, L. Wang, X. Wu, and D. Zhang, ``Dynamic analysis of motor imagery EEG using kurtosis based independent component analysis," in *Advances in Cognitive Neurodynamics ICCN*. Dordrecht, The Netherlands: Springer, 2008, pp. 381_385.

[22] Z. Jin, G. Zhou, D. Gao, and Y. Zhang, ``EEG classi_cation using sparse Bayesian extreme learning machine for brain_computer interface," *Neural Comput. Appl.*, pp. 1_9, 2018.

[23] Y. Jiao *et al.*, ``Sparse group representation model for motor imagery EEG classi_cation," *IEEE J. Biomed. Health Inform.*, to be published.

[24] Suraj, R. K. Sinha, and S. Ghosh, ``Jaya based ANFIS for monitoring of two class motor imagery task," *IEEE Access*, vol. 4, pp. 9273_9282, 2016. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7778152

[25] A. H. Phan and A. Cichocki, ``Tensor decompositions for feature extraction and classi_cation of high dimensional datasets," *IEICE Nonlinear Theory Appl.*, vol. 1, no. 1, pp. 37_68, 2010.

[26] A. B. Das, M. I. H. Bhuiyan, and S. M. S. Alam, ``Classi_cation of EEG signals using normal inverse Gaussian parameters in the dual-tree complex wavelet transform domain for seizure detection," *Signal, Image Video Process.*, vol. 10, no. 2, pp. 259_266, 2016.

[27] C.-Y. Chiu, C.-Y. Chen, Y.-Y. Lin, S.-A. Chen, and C.-T. Lin, ``Using a novel LDA-ensemble framework to classi_cation of motor imagery tasks for brain-computer interface applications," in *Proc. Int. Comput. Symp. (ICS)*, vol. 1, 2014, pp. 136_142.

[28] L. He, D. Hu, M. Wan, Y. Wen, K. M. von Deneen, and M. Zhou, ``Common Bayesian network for classi_cation of EEG-based multiclass motor imagery BCI," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 46, no. 6, pp. 843_854, Jun. 2016.

[29] K. K. Ang, Z. Y. Chin, C.Wang, C. Guan, and H. Zhang, ``Filter bank common spatial pattern algorithm on BCI competition IV datasets 2a and 2b," *Frontiers Neurosci.*, vol. 6, Mar. 2012, Art. no. 39.

[30] A. Schlögl, F. Lee, H. Bischof, and G. Pfurtscheller, ``Characterization of four-class motor imagery EEG data for the BCI-competition 2005," *J. Neural Eng.*, vol. 2, no. 4, pp. L14_L22, 2005.

[31] ZHIWEN ZHANG, FENG DUAN, JORDI SOLÉ-CASALS, JOSEP DINARÈS FERRAN, ANDRZEJ CICHOCKI, (Fellow, IEEE), ZHENGLU YANG, AND ZHE SUN," A Novel Deep Learning Approach with Data Augmentation to Classify Motor Imagery Signals"

[32] **ZijianWang, Lei Cao, Zuo Zhang, Xiaoliang Gong, Yaoru Sun, HaoranWang," Short time Fourier transformation and deep neural networks for motor imagery brain computer interface recognition"**