Brain Activity Recognition using Deep Electroencephalography Representation

Johri, Riddhi; Pandey, Pankaj; Miyapuram, Krishna Prasad; Lomas, Derek

DOI
10.1109/APSCON56343.2023.10100986

Publication date
2023

Document Version
Final published version

Published in
APSCON 2023 - IEEE Applied Sensing Conference, Symposium Proceedings

Brain Activity Recognition using Deep Electroencephalography Representation

Riddhi Johri
IIT Gandhinagar
India

Pankaj Pandey
IIT Gandhinagar
India

Krishna Prasad Miyapuram
IIT Gandhinagar
India

Derek Lomas
TU Delft
Netherland

2023 IEEE Applied Sensing Conference (APSCON) | 978-1-6654-6163-4/23/$31.00 ©2023 IEEE | DOI: 10.1109/APSCON56343.2023.10100986

***Abstract*—Advances in neurotechnology have enhanced and simplified our ability to research brain activity with low-cost and effective equipment. One such scalable and noninvasive technique is Electroencephalography (EEG), which detects and records electrical brain activity. Brain activity recognition is one of the emerging problems as EEG wearables become more readily available. Our research has modeled EEG signals to classify three states (i) music listening, (ii) movie watching, and (iii) meditating. The datasets incorporating the brain signals induced while performing these activities are NMED-T for music listening, SEED for movie watching, and VIP_SNY_HYT for meditating. EEG activity is transformed into deep representation using a convolutional neural network comprising three different types of 2D convolutions: Temporal, Spatial, and Separable, to capture dependencies and extract high-level features from the data. The Depthwise Convolution function is responsible for learning spatial filters within each temporal convolution, and combining these spatial filters across all temporal bands optimally is learned by the Separable Convolutions. EEGNet and EEGNet_SSVEP are specially designed for EEG Signal Processing and Classification, and the DeepConvNet has incorporated more convolution layers. Our finding demonstrates that increasing the number of layers in the Network provided a higher accuracy of 99.94% using DeepConvNet. In contrast, the accuracy of EEGNet and EEGNet_SSVEP resulted in 85.63% and 75.76%, respectively.**

***Index Terms*—Human-Centered Computing, EEG Sensor, Machine Learning, Brain Activity**

I. Introduction

Human Activity Recognition-based wearables like Fitbit have become quite common in everyday life. They detect our movement by segmenting multivariate data streams from multiple sensors like altimeter, accelerometer, bioimpedance sensor, and gyroscope and label each segment with an activity. The increasing surge in research in the healthcare domain has spiked with the advent of biosensors such as electroencephalography (EEG), electromyography (EMG), and electrocorticogram (ECoG). The high temporal resolution and sampling rate of EEG, as well as its robustness and wearability, have made it widely popular.

The majority of EEG studies [1] employ Deep Learning models to categorise EEG data, notably to analyse sleep stages, detect seizures, and monitor emotion and cognition. Other researchers explore new processes for producing data, enhancing feature learning, or handling artefacts. In addition, the studies on wearable-based HAR devices incorporating EEG emphasise detecting motor imagery activity based on a person's physiological activity [2] or disentangling the data from irrelevant noise in order to learn essential features for HAR tasks [3]. However, this research proposes to detect the neurophysiological activities of a subject in naturalistic events using an EEG sensor.

We use three EEG datasets that have the brain signals of participants recorded while they were listening to music, watching movies, and meditating. Three different Convolutional Neural Networks, namely EEGNet, EEGNet_SSVEP, and DeepConvNet, were employed to detect these activities, and the training and testing datasets consisted of entirely different subjects. In contrast to most previous studies in which either of these datasets were analysed separately to classify songs, emotions, movies, or meditation, we have combined all three datasets and determined the 55 common EEG channels present in all the datasets to recognise the type of activity.

II. Data Description

NMED-T: This study analyzed NMED-T [4], a publicly available dataset used for research in music processing. Behavioural responses from twenty participants engaged in a naturalistic song listening study were recorded in this dataset, along with their EEG signals. The pre-processed version of this dataset, which contains 125 channels of EEG data captured at 125 Hz, was primarily employed in our study.

SEED: SEED [5] is composed of electroencephalogram signals collected from fifteen subjects including eight females and seven males, who watched fifteen excerpts from Chinese films. Fifteen trials were conducted per subject, lasting 305 seconds, including a hint of starting for 5 seconds, a movie clip for 4 minutes, a self-assessment for 45 seconds, and a rest for 15 seconds. The data collected from the 62-electrode EEG cap was then downsampled to 200 Hz, and the pre-processed version available online was used for the experiment [6].

VIP_SNY_HYT: Datasets collected by the Meditation Research Institute (MRI), located in Rishikesh, India, were used in order to examine the electroencephalography (EEG) activity of meditation practitioners from two different meditation traditions [7]: Vipassana (VIP) and Himalayan Yoga (HT). For the data, a group of sixteen subjects was selected for each meditation technique, making it a total of 32 subjects. The EEG signals were sampled at 256Hz and contained 64

We thank to SERB and PlayPowerLabs for PMRF to PP. We thank FICCI for facilitating the PMRF.

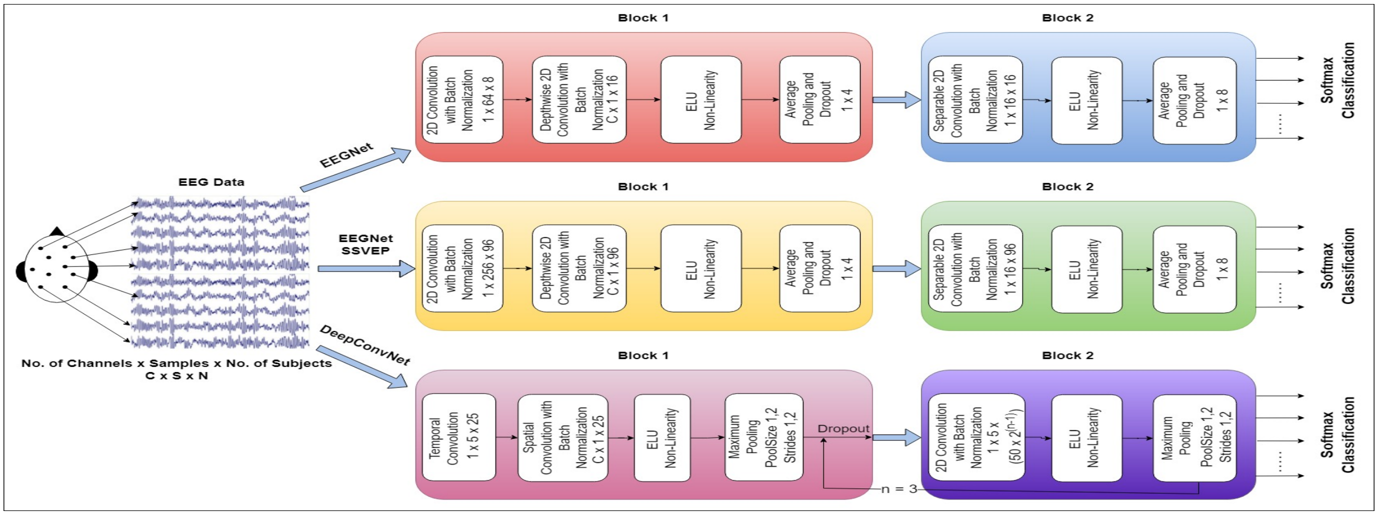

Fig. 1. The different architectures used for classification from EEG data. There were 55 common channels across all three datasets making the number for channels, C = 55,. The values of S and N vary between datasets.

channels. The study utilized the pre-processed version of these brain signals [7].

III. Methodology

EEGNet: EEG data consisting of C channels, S time samples, and N subjects were passed for 100 epochs to the EEGNet model [8] composed of three convolutions in sequence. The input was routed through eight 2D convolution filters of size (1,64), which generated feature maps at various bandpass frequencies in the first block to obtain temporal information. Then, D*8 depthwise convolutions of size (C,1) were employed to learn spatial information within each temporal filter. The depth parameter D determined how many spatial filters should be learned for each feature map. The model was regularised with a dropout rate of 0.5 after applying exponential linear unit (ELU) Non-Linearity and Average Pooling layer of size (1, 4).

In Block 2, there was a Separable 2D Convolution layer composed of sixteen filters of size (1,16). This helped to combine spatial filters across temporal bands optimally. After applying ELU Non-Linearity, dimensionality reduction was achieved using an Average Pooling layer of size (1, 8). All the convolutions were followed by Batch Normalization. Finally, features after dropout were passed to the Softmax Classification layer.

EEGNet_SSVEP: The SSVEP variant of EEGNet [9] was designed specifically for classifying Asynchronous Steady-State Visual Evoked Potentials signals. This differs from the above network in size and number of kernels utilised in each convolution layer, as illustrated in figure 1. In block 1, ninety-six 2D Convolution and Depthwise 2D Convolution (D=1) of size (1,256) and (C,1) respectively were used to obtain frequency-specific spatial filters. Furthermore, depthwise convolutions reduced the number of free parameters to fit compared to fully-connected convolutions.

In Block 2, ninety-six separable convolutions of size (1,16) were used, which reduced the number of parameters to fit while also explicitly decoupling the link between feature maps across and inside them. In turn, a kernel summarizing each feature map was learned, followed by the optimal merging of the outputs. After every Convolution layer, Batch Normalization was applied. The input was then processed via ELU non-linear activation, 2D average pooling, and dropout layers. Lastly, a dense layer and a softmax activation function were connected to the final layer.

DeepConvNet: The deep ConvNet architecture [10] to extract features and decode EEG signals was inspired by computer vision techniques. This architecture had four blocks, each consisting of a 2D convolution layer with max_norm constraint, batch normalization, ELU non-linearity activation, max pooling of size (1,2) with strides (1,2), and a dropout layer with a dropout rate of 0.5.

The convolution added in the first block was split into two convolution layers of 25 filters each, one temporal layer (1,5) and one spatial layer (C,1). These two layers helped in forcing a linear transformation into a blend of a temporal and a spatial filter, which implicitly regularized the overall convolution. Finally, the fifth layer was a dense layer with a softmax activation function for classification.

TABLE I
THE 55 CHANNELS FOUND TO BE COMMON IN ALL THE THREE DATASETS CONSIDERED

Region	Channels
Frontal	F1, F3, F5, F7, FZ, F2, F4, F6, F8, FP1, AF3, FPZ, FP2, AF4
Central	FC5, FC3, FC1, FC6, FC4, FC2, FCZ, C1, C3, C5, C2, C4, C6,
Parietal	CP5, CP3, CP1, CPZ, CP6, CP4, CP2, P1, P3, P5, P7, PZ, P2, P4, P6, P8,
Occipital	O1, OZ, POZ, PO7, PO3, PO8, PO4, O2
Temporal	TP7, TP8, FT7, FT8

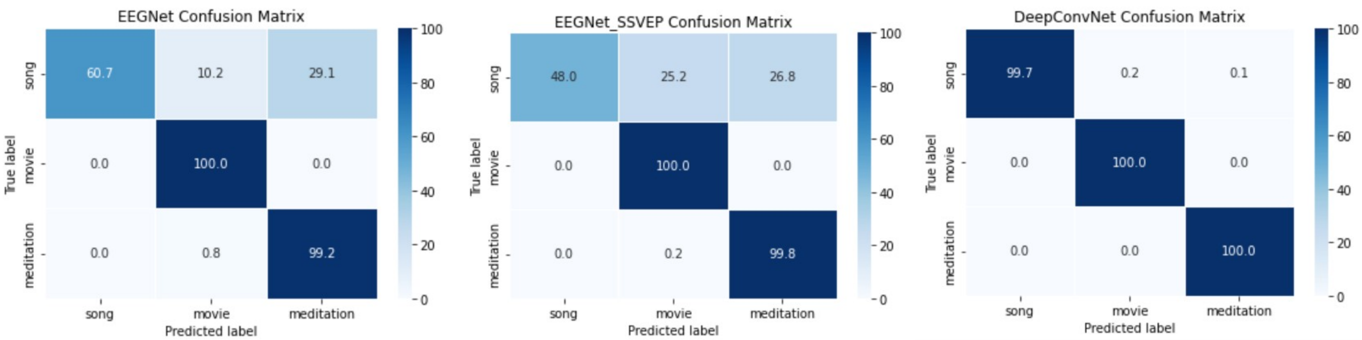

Fig. 2. Confusion Matrices of the brain activity classification using the three CNN architectures

TABLE II
NUMBER OF SAMPLES FOR TRAINING AND TESTING.

Dataset	Train	Test
NMED-T (Music) [4]	5448	1520
SEED (Movie) [5] [6]	8280	1990
VIP and HT (Meditation) [7]	3040	3024

TABLE III
CLASSIFICATION METRICS

Architecture	Accuracy	Precision	Recall	F1 Score
EEGNet [8]	84.63%	86.62%	87.36%	84.70%
EEGNet_SSVEP [9]	74.76%	82.60%	77.23%	74.43%
DeepConvNet [10]	99.94%	99.91%	99.94%	99.93%

IV. EXPERIMENTAL RESULTS

This research analysed three datasets for categorising brain signals according to the type of activity a subject was engaging in, such as music listening, movie watching, and meditation. As listed in Table I, 55 EEG channels were found to be shared across NMED-T, SEED, and VIP_SNY_ HYT datasets and the signals from these channels were analysed.

The Table II displays the number of samples from each dataset utilised for training and testing the model. One sample contained the brain signals for a duration of 10 secs from each of the 55 channels from every subject considered in the dataset. Models were never exposed to data from the subjects they were getting tested on because training and testing datasets included EEG signals of entirely different sets of subjects. Thus, the accuracies achieved in identifying the activities of the subjects as shown in Table III by the three CNN models indicates that they were able to learn the subject invariant representation of the EEG signals successfully.

While the high accuracy of 99.94% of DeepConvNet was attributed to incorporating two more convolution layers, all the models successfully identified a common neural signature for the naturalistic activities across several participants. This transfer learning of EEG data has been a prevalent challenge in terms of invariant representation since synaptic plasticity and interactions with the environment affect brain representation of experiences differently in different individuals.

V. CONCLUSION

Our research shows that the CNN models successfully learned the subject-independent features of brain signals specific to each of the activities. These findings can be leveraged in wearable EEG headsets to monitor and detect naturalistic scenarios using brain signals. The future possibility is to use a dataset collected at one place with the same participants performing different tasks.

REFERENCES

[1] Y. Roy, H. Banville, I. Albuquerque, A. Gramfort, T. H. Falk, and J. Faubert, "Deep learning-based electroencephalography analysis: a systematic review - iopscience," Aug 2019. [Online]. Available: https://iopscience.iop.org/article/10.1088/1741-2552/ab260c

[2] A. Salehzadeh, A. P. Calitz, and J. Greyling, "Human activity recognition using deep electroencephalography learning," *Biomedical Signal Processing and Control*, vol. 62, p. 102094, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1746809420302500

[3] J. Su, Z. Wen, T. Lin, and Y. Guan, "Learning disentangled behaviour patterns for wearable-based human activity recognition," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 6, no. 1, pp. 1–19, mar 2022. [Online]. Available: https://doi.org/10.1145%2F3517252

[4] S. Losorelli, D. T. T. Nguyen, J. P. Dmochowski, and B. Kaneshiro, "Naturalistic music eeg dataset - tempo (nmed-t)," 2017. [Online]. Available: https://exhibits.stanford.edu/data/catalog/jn859kj8079

[5] W.-L. Zheng and B.-L. Lu, "Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks," *IEEE Transactions on Autonomous Mental Development*, vol. 7, no. 3, pp. 162–175, 2015.

[6] R.-N. Duan, J.-Y. Zhu, and B.-L. Lu, "Differential entropy feature for EEG-based emotion classification," in *6th International IEEE/EMBS Conference on Neural Engineering (NER)*. IEEE, 2013, pp. 81–84.

[7] C. Braboszcz, B. R. Cahn, J. Levy, M. Fernandez, and A. Delorme, "Increased gamma brainwave amplitude compared to control in three different meditation traditions," *PloS one*, vol. 12, no. 1, p. e0170647, 2017.

[8] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, "Eegnet: a compact convolutional neural network for eeg-based brain–computer interfaces," *Journal of Neural Engineering*, vol. 15, no. 5, p. 056013, 2018. [Online]. Available: http://stacks.iop.org/1741-2552/15/i=5/a=056013

[9] N. Waytowich, V. J. Lawhern, J. O. Garcia, J. Cummings, J. Faller, P. Sajda, and J. M. Vettel, "Compact convolutional neural networks for classification of asynchronous steady-state visual evoked potentials," *Journal of Neural Engineering*, vol. 15, no. 6, p. 066031, 2018. [Online]. Available: http://stacks.iop.org/1741-2552/15/i=6/a=066031

[10] S. R. Tibor, S. J. Tobias, F. L. D. Josef, G. Martin, E. Katharina, T. Michael, H. Frank, B. Wolfram, and B. Tonio, "Deep learning with convolutional neural networks for eeg decoding and visualization," *Human Brain Mapping*, vol. 38, no. 11, pp. 5391–5420. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/hbm.23730