

# Application of Deep Learning for Emotion Recognition

Mrs. Varsha R.Toshniwal<sup>1</sup>, Mrs. Pooja S.Puri<sup>2</sup>

*Department of Textile, D. K. T. E. Society's Textile & Engineering Institute, and Ichalkaranji*

[<sup>1</sup>modani.varsha18@gmail.com](mailto:modani.varsha18@gmail.com)

*Department of Electronics, D. K. T. E. Society's Textile & Engineering Institute, and Ichalkaranji*

[<sup>2</sup>poojapuri553@gmail.com](mailto:poojapuri553@gmail.com)

## ABSTRACT

The electroencephalogram (EEG) gained a lot of importance in recent years because of its property to depict the nature and actions of human perception. EEG signals are good at capturing the emotional state of a person by measuring the neuronal activities in different regions of the brain. Lots of EEG-based brain-computer interfaces with a different number of channels ( 62 channels, 32 channels, etc.) are being used to capture neuronal activities which can be segmented into different frequency ranges (delta, theta, alpha, beta and gamma). This paper puts forward a neural network architecture for the recognition of emotion from EEG signals and a study providing the set of brain regions and the frequency type associated with the corresponding brain region which contributes most for the detection of emotion through EEG signals. For experimentation, SEED-IV dataset has been used

## KEYWORDS

EEG, Emotion Recognition, Deep Neural Network, Brain Region Analysis, SEED-IV

## 1. INTRODUCTION

Human emotion recognition has gained a lot of popularity and boom across the research field due to the need for diagnosing emotion related diseases and psychological needs. The emotion of a person can be recognized from a number of factors like his/her facial expressions, body language, audio language, psychological signals (electroencephalography (EEG), electrocardiogram (ECG), electromyography (EMG) and other), etc. As compared to these factors, psychological signals are the factors that a person cannot hide or alter and hence they are more useful to examine the emotional state of a person. Other factors like facial expression, body language, etc can be altered to hide the current emotional state of a person and hence actual emotional state cannot be depicted using these factors. A lot of research is being done to recognize human emotions using the non-psychological signals but the systems built using psychological signals are more reliable in terms of depicting the true emotional state of a person.

EEG signals measure voltage fluctuations from the cortex in the brain which is achieved by setting up the brain computer interfaces (BCI) which includes a number of sensors and machines. The device used to record the signal comes with a different number of channels or nodes like 62 channels, 32 channels, etc. Increasing the number of channels helps in measuring the signals more accurately over a greater area on the cortex. For the experimentation we are

using the SEED-IV [8] dataset which contains the EEG signal values recorded from a 62 channel device. After the signals are recorded the signal data goes through preprocessing phases because EEG signals at the initial stage after recording contain a lot of noise such as AC power lines, lighting and a large array of electronic equipment (from computers, displays and TVs to wireless routers, notebooks and mobile phones) [1]. To deal with noise in SEED-IV dataset [8], the raw EEG data are first down-sampled to a 200 Hz sampling rate and then the EEG data are processed with a bandpass filter between 1 Hz to 75 Hz [8]. In the frequency domain, the most important frequency bands present in EEG signals are delta (1–3 Hz), theta (4–7 Hz), alpha (8–13 Hz), beta (14–30 Hz) and gamma (31–50 Hz) [2]. Since EEG signals contain the time domain values so to extract the frequency domain values we have to transform time-domain values to the frequency domain values. This can be done by using methods like Fast Fourier Transform (FFT), Wavelet Transform (WT), eigenvector and autoregressive[3].

The EEG signals are recorded across different regions of the brain but the signals from every region do not contribute to the actual detection of the emotional state of the person. In this paper proposed framework for the classification of the emotional state of a person by interpreting the EEG signals and also provide the study regarding the findings of the brain regions and the frequency type (delta, theta, alpha, beta and gamma) associated with the corresponding brain region which contributes most for the depiction of emotional state using EEG signals.

## 2. RELATED WORK

The research work conducted so far in the recognition of emotions using EEG signals has spanned over a wide range of fields and techniques like graph neural network, deep canonical correlation analysis, etc.

Peixiang Zhong and Di Wang in their paper, have described a graph neural network for the classification of emotions on EEG data. Their proposed neural network architecture also captures both local and global relations among different EEG channels. Specifically, they have modelled the inter-channel relations in EEG signals with the help of an adjacency matrix [4]. Their method resulted (on SEED-IV dataset [8]) in an accuracy of 79.37% with a standard deviation of 10.54 for subject independent classification and 73.84% accuracy with a standard deviation of 8.02 for subject independent classification. Wenming Zheng proposed a canonical correlation analysis (GSCCA) method for emotion recognition [5]. The author makes use of conventional CCA method to model the linear correlations between emotional EEG class label vectors and the corresponding EEG feature vectors. Their method resulted in an accuracy of 69.08% with a standard deviation of 16.66 for subject independent classification. Authors from “A Novel Bi-hemispheric Discrepancy Model for EEG Emotion Recognition” [6] proposed a bi-hemispheric discrepancy model (BiHDM) that presents a study for asymmetric differences between two hemispheres for electroencephalograph (EEG) emotion recognition. Their model achieved an accuracy of 74.35% with a standard deviation of 14.09 for subject independent classification and 69.03% accuracy with a standard deviation of 8.66 for subject independent classification.

Some simple machine learning algorithms like SVM were also used on the dataset but the resulting accuracy was less as compared to other complex models. On using SVM the resulting accuracy was 56.61% with a standard deviation of 20.05 for subject dependent classification and accuracy of 37.995 with a standard deviation of 12.52 for subject independent classification.

### 3. EXPERIMENTAL SETTINGS

#### 3.1 DATASET

We are using the SEED-IV [8] dataset for the experimentation. The SEED dataset comprises EEG data of 15 subjects. For each participant, 3 sessions are performed on different days and each session contains 24 trials. In one trial, the participant watches one of the film clips, while his/her EEG signals and eye movements are collected with the 62-channel ESI NeuroScan System and SMI eye-tracking glasses [8]. The clips shown to the participants contain a tendency of four emotions i.e. happy, sad, fear or neutral emotion. The dataset also contains processed data (noise removed) in the form of power spectral density (PSD) and differential entropy (DE) features within each segment at 5 frequency bands of delta (1~4 Hz), theta(4~8 Hz), alpha(8~14 Hz), beta(14~31 Hz) and gamma (31~50 Hz) [8].

$$\mathbf{PSD} = E[x^2]$$

$$\mathbf{DE} = - \int_{-\infty}^{\infty} P(x) \ln(P(x)) dx$$

Figure 1. The calculation of PSD and DE of a random variable X.

For the experimentation we are directly using differential entropy(DE) features smoothed by linear dynamic systems (LDS) so that we can make a comparison with the existing methods which also make use of the same.

##### 3.1.1 DATASET PROCESSING

The SEED-IV dataset contains the data in the format of (62, w, 5) where 62 is the number of channels, w is the number of time windows in the trial and 5 represents the five frequency bands i.e. delta, theta, alpha, beta and gamma. Firstly we extracted the data associated with each frequency which resulted in five sets of (62, w, 1) data (one for each frequency). Then taking transpose and combining data of each frequency across column resulted in a dataset of format (w, 62\*5). The input features for the model will be the values of de\_LDS of all 62 channels across all 5 frequency bands. Normalization techniques are also being used in the preprocessing phase of the dataset.

#### 3.2 CLASSIFICATION SETTING

We have followed the procedure of prior studies to conduct both subject dependent and subject independent classification.

##### 3.2.1 SUBJECT DEPENDENT CLASSIFICATION

For subject dependent classification, we have followed the experimental setup used by [6], [7] and [4] i.e. to use the first 16 trials for the training set and remaining 8 trials containing all emotions (two trials per emotion class) for testing. Data from all the three sessions were used in the training and the test sets [8].

##### 3.2.2 SUBJECT INDEPENDENT CLASSIFICATION

For subject independent classification, we have followed the experimental setup used by [6] and [4] i.e. using leave-one-subject-out cross-validation in which during each fold, we use 14

subjects for training purpose and remaining 1 for the test set [4]. The final accuracy is calculated by averaging accuracy across all test subjects. Data from all the three sessions were used in the training and the test sets [8].

#### 4. PROPOSED APPROACH FOR EMOTION RECOGNITION

In this work a Deep Neural Network with 4 hidden layers of size 256, 256, 256 and 32 is being used. A dropout rate of 70% is being used and is accompanied by L1 and L2 regularizers. The learning rate of 0.001 is being used with a decay of  $1e-6$ . We have used Stochastic gradient descent as an optimizer and the ReLU activation function is being utilized in the input and hidden layers. To accelerate gradient vectors in the right direction, we have also used the momentum which helps in faster converging.

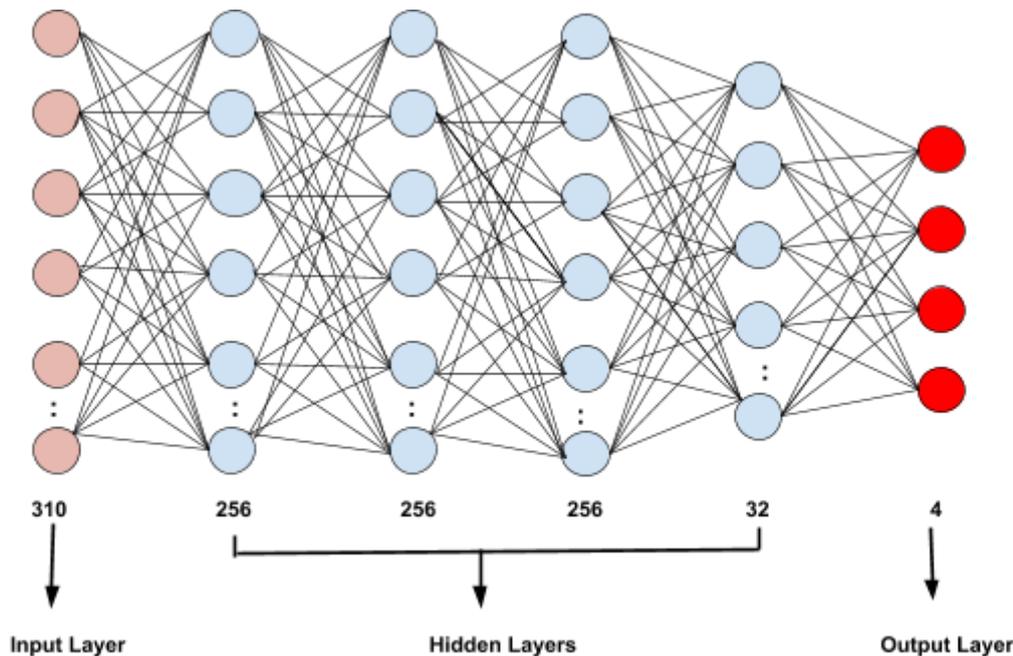


Figure 2. Proposed Neural Network Architecture

Using our model we have achieved an accuracy (mean) of 68.01% with a standard deviation of 8.87 for subject dependent classification and accuracy (mean) of 60.04% with a standard deviation of 7.72 for subject independent classification on SEED-IV database [8].

For subject dependent classification our model outperforms the results of the DAN model [10] and results from machine learning algorithms (like using SVM accuracy achieved was 56.61 with a standard deviation of 20.05). For subject independent classification our model outperforms the results from [10] and other machine learning algorithms (like using SVM accuracy achieved was 37.99 with a standard deviation of 12.52).

#### 5. IMPORTANT FREQUENCY BAND AND CHANNEL INVESTIGATION

After training a neural network, the “Connection Weights” algorithm is one of the prominent ways to find which input feature contributes most or is important for a classification task. The

connection weights algorithm calculates the sum of products of final weights of the connections from input neurons to hidden neurons and hidden neurons to output neurons for all input neurons [9]. According to the connection weights algorithm the relative importance of an input variable X is defined as[9]:

$$RI_X = \sum_{Y=1}^m W_{XY} W_{YZ}$$

Where  $RI_X$  is the relative importance of input neuron X,  $\sum_{Y=1}^m W_{XY} W_{YZ}$  is the sum of product of final weights of the connection from input neuron to hidden neurons with the connection from hidden neurons to output neuron, Y is the total number of hidden neurons, and Z is output neurons[9].

In case of our architecture with 4 hidden layers the importance of an input variable X will be defined by:

$$RI_X = \sum_{Y=1}^{256} \sum_{Z=1}^{256} \sum_{K=1}^{256} \sum_{L=1}^{32} \sum_{M=1}^4 W_{XY} W_{YZ} W_{ZK} W_{KL} W_{LM}$$

Where Y is the total number of neurons in the first hidden layer, Z is the total number of neurons in the second hidden layer, K is the total number of neurons in the third hidden layer, L is the total number of neurons in the fourth hidden layer and M is the total number of neurons in the final output layer. For the ease of computation we have used matrix multiplication of matrices containing weight connectivity to the next layer. After the matrix multiplication the output vector was of the shape (no. input features, no. of output features) i.e. (310,4). So taking the summation along the axis 1 will yield a vector of size (no. input features) i.e. (310) containing the relative importance of each inputfeature.

After investigating, we found the following set of frequency bands and channels that have the most relative importance:

Frequency Band	Channel Name
Gamma	F6
Gamma	FT8
Beta	FP1
Delta	F3
Theta	PZ
Gamma	AF4
Gamma	P08
Gamma	C4

Gamma	FP1
Delta	P03
Theta	CB2

Table 1. Set of channels and their frequency bands.

Among all the 5 frequency bands the channels from the gamma band have more relative importance as compared to other bands. Our results are in alignment with the results from [11], [12], [13] and [14] which claims that gamma frequency has more relative importance for emotion recognition as compared to the other frequency bands.

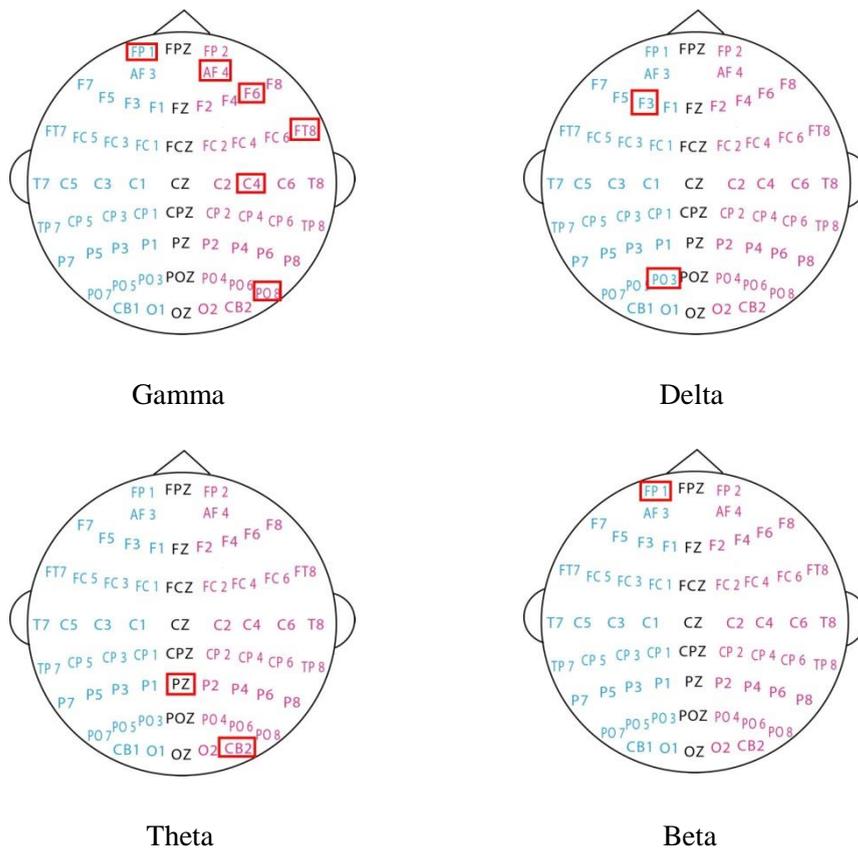


Figure 3. Top 11 channels and their associated frequency bands.

### 5. CONCLUSION

In this paper we propose a deep learning-based neural network architecture and a study which provides the set of 11 input channels and the frequencies associated with them which have the most relative importance for the recognition of emotion using EEG data. Our method outperforms some of the previously implemented architectures for the emotion recognition on the SEED-IV dataset and also provides a study on important input channels and frequency

bands which are essential for emotion recognition. Our results are also in alignment with the previously published studies.

## REFERENCES

- [1] Repovš, G., (2010) “Dealing with Noise in EEG Recording and Data Analysis”.
- [2] The McGill Physiology Virtual Lab.[Online]. Available online: [https://www.medicine.mcgill.ca/physio/vlab/biomed\\_signals/eeg\\_n.htm](https://www.medicine.mcgill.ca/physio/vlab/biomed_signals/eeg_n.htm)
- [3] Amjed S. & Ausilah A., (2014) “Methods of EEG Signal Features Extraction Using Linear Analysis in Frequency and Time-Frequency Domains”, Hindawi Publishing Corporation, ISRN Neuroscience, Vol. 2014, Article ID730218.
- [4] Zhong, P., Wang, D., & Miao, C., (2019), “EEG-Based Emotion Recognition Using Regularized Graph Neural Networks”, *ArXiv, abs/1907.07835*.
- [5] W. Zheng, (2016), “Multichannel EEG-based emotion recognition via group sparse canonical correlation analysis”, *IEEE Transactions on Cognitive and Developmental Systems*, vol. 9, no. 3, pp.281–290.
- [6] Y. Li, W. Zheng, L. Wang, Y. Zong, L. Qi, Z. Cui, T. Zhang, and T. Song, (2019), “A novel bi-hemispheric discrepancy model for EEG emotion recognition,” *arXiv preprint arXiv:1906.01704*.
- [7] Zheng, W., Liu, W., Lu, Y., Lu, B., & Cichocki, A. (2019), “EmotionMeter: A Multimodal Framework for Recognizing Human Emotions”, *IEEE Transactions on Cybernetics*, 49, 1110-1122.
- [8] Zheng, W.L., Liu, W., Lu, Y., Lu, B.L. and Cichocki, A., 2018. Emotionmeter: A multimodal framework for recognizing human emotions. *IEEE transactions on cybernetics*, 49(3), pp.1110-1122.
- [9] Julian D., Donald A., “Illuminating the “black box”: a randomization approach for understanding variable contributions in artificial neural networks”, *Ecological Modelling*. 2002. August 15; 154(1): 135–50.
- [10] He L., Yi-Ming J., Wei-Long Z., Bao-Liang L., “Cross-subject emotion recognition using deep adaptation networks,” in *International Conference on Neural Information Processing*. Springer, 2018, pp.403–413.
- [11] Yang K., Tong L., Shu J., Zhuang N., Yan B., Zeng Y., High Gamma Band EEG Closely Related to Emotion: Evidence From Functional Network. *Frontiers in Human Neuroscience*. 2020;14.
- [12] Li, M., Xu, H., Liu, X. and Lu, S., 2018. Emotion recognition from multichannel EEG signals using K-nearest neighbor classification. *Technology and Health Care*, 26(S1), pp.509-519.
- [13] Li, M. and Lu, B.L., 2009, September. Emotion classification based on gamma-band EEG. In *2009 Annual International Conference of the IEEE Engineering in medicine and biology society* (pp. 1223-1226). IEEE.
- [14] Duan, R.N., Zhu, J.Y. and Lu, B.L., 2013, November. Differential entropy feature for EEG-based emotion classification. In *2013 6th International IEEE/EMBS Conference on Neural Engineering (NER)* (pp. 81-84). IEEE