Electroencephalogram Based Emotion Detection Using Hybrid LongShort Term Memory

Thejaswini S,[1,] *Dr. K M Ravikumar ,[2]

[1]*DepartmentofElectronics and Telecommunication Engineering,B M S Institute of Technology and Management,Bengaluru, Karnataka,India.*
[2] *Principal and Professor Department of Electronics and Communication Engineering, S J C Institute of Technology, Chickballapur, Karnataka,India*

thejaswini79@gmail.com

Abstract

Emotion detection using physiological signals is an upcoming research extending applications in various domains. One important challenge in detection of inner emotion states is a good predictive rate in order to build any application. In our present work, a hybrid Long Short Term Memory (LSTM) algorithm is proposed based on channel fusion approach. Data is acquired by eliciting emotions using eight 3-D Virtual Reality (VR) videos for eight discrete emotion states. On preprocessed data, 8-level decomposition using Discrete Wavelet Transforms(DWT) is performed, wavelet features and time-domain features are extracted and fed to Hybrid LSTM. The hybrid algorithm is performing well for eight discrete emotion states (happy excited, calm, bored, fear, tensed, sad and relax), with an accuracy rate of 80.05% and 93.24 % for 4 states in categorical form (Valence- Arousal scale). Frequency domain features on various bands exhibited a good predictive rate than time domain features.

Keywords: EEG, Emotion states, LSTM, DWT, Channel reduction

1. Introduction

Human emotion states plays an important role in daily life and this has high impact on their daily routine activities [1]. To develop an emotional model which can measure human emotions states and interface these with computer is aim of affective computing which has enormously increased the applications in HCI. The inner emotions of person are subjective, it depends on feelings, experience both internals and externals of subjects [2]. There as several ways to detect and evaluate emotion states, to list a few is through speech, facial and physiological signals. The shortcomings of speech and facial approach is that subject can avoid or fake their emotion states which may give rise to wrong decision making. Analyzing using physiological signal have overwhelmed these flaws [3-4]. Detection of emotions using electroencephalogram (EEG) is rapidly growing due the fact like, no influencing of brain signals, availability of many portable devices for data acquisition. This has enabled to develop applications in medical and non- medical fields [5].

The categorical and dimensional models are used to represented emotions states. The Categorical / discrete model makes use of emotion states in discrete form namely happy, sad, anger, joy, fear etc. In dimensional model, common set of dimensions which links emotion states into two spaces-valence and arousal states and is depicted in Figure 1.

Based on 10-20 International standard electrode system, EEG signals are acquired. To stimulate emotions in subject's, video or audio stimulus is used. In this paper a hybrid LSTM algorithm is designed for real-time data acquired using VR stimuli. The main objective of our work is used to develop an efficient Deep learning algorithm to detect emotion states with respect to dimensional model (Four states) and Categorical model (eight states).

The paper is structured as sections. In section 2 related work in this area are explored. Step by step Methodology of the proposed work are briefed out in section 3. Results and discussions are described in section 4 followed by section 5 with conclusion and future scope.

2. Related Work

In [6], they developed a multi-modal attention-based BLSTM network of three layers to recognize temporal features and used DNN to classify future using AMIGOS database. They obtained an accuracy of 82.5 % for arousal and 77.8% for valance (third layer). LSTM algorithm using non-linear HOC for DEAP data was proposed [7]. They system accuracy rate of valence-arousal was 84.68% and 82.01% for 2 and 4 classes. Their system attained an accuracy for Valence-Arousal score for four classes as 82.01% and 84.68% for two-class. In [8], they proposed a channel fused dense CNN for SEED and DEAP database. They used images of Differential Entropy of 5 bands to CDCN model. The accuracies were an average of 90.63% for SEED and 92.58% DEAP datasets. For seed data base, the accuracy of different frequency bands was as follows: Delta 65.19, Theta 69.84, Alpha 72.16, Beta 80.83, and Gamma 82.63. For DEAP data base the valence and arousal accuracy rate were 92.24 & 92.92 respectively. In [9] PSD was used for pre and post 2-D and 2-D movie watching on five member groups using 20 channels device. The success of classifier rate was 73.36 % and 89.11% for PLSR and SVM using STFT. The same using DWT was 75.18% and 87.26%.

The following authors work is related to DEAP database. A [10] merged LSTM was used which consists of channel wise LSTM layer followed by dense layer for wavelet features. The classification rate of valence and arousal were 84.89% and 83.85% respectively. In [11], they used frequency band power on a combination of Stack auto encoder and LSTM frame work and got a valence accuracy as 81.10%, arousal rate as 74.38%. They proposed an R2G-STNNwhich consists of spatial and temporal neural network models for SEED data base and obtained an average accuracy of 92.9 %[12]. In [13], they used KNN classifier for different frequency bands and obtained 95% accuracy in gamma band. They showed increase in accuracy when number of channels were increased. In [14], they developed multi class SVM classifier using time domain as well as Gabor-IMF features and obtained an 93% classification rate for 7 and 3 channels. Deep learning frame work with arousal and valence performance of 75.92% and 76.83% explored [15]. In [16], they classified the data using SVM on frequency and time features with 72.6% and 70.3% accuracy rate. In [17]they, used DWT, energy and entropy features on SVM and KNN classifier with an efficiency of 86.75% for arousal and 84.05% for valance. In [18], they used 8 channels and SVM classifier on extracted features like non-linear, time and frequency domain. Their classifier performance is 74.43 % for valence and 80.01 % for arousal.

From the above review it was noted that work was carried out using different databases, used various machine learning and deep learning algorithms for classifying emotion states into basic emotion models discussed earlier with an accuracy of 75 % to 93%. Only a few work is carried out with discrete emotion states. In our proposed work, a modified LSTM network is built to classify emotions into valence- arousal and eight discrete emotion states. To elicit emotions VR video clips are used.

3. Methodology

Propose System: In the proposed work, to elicit emotion states, 3-D VR videos are used. The extracted time-frequency, Statistical features are applied for modified LSTM frame work to classify into eight (8) categorical states (calm, relax, excited, happy, tensed, fear, bored and sad) as well as four (4) dimensional states (HAHV, HALV, LAHV & LALV). The methodology proposed for our work is dissipated in Figure 2.

Protocol for Data Acquisition : The data acquisition protocol was developed based on the ground rules followed in DEAP [19] and SEED [20] databases. In the experiment set up, 66 subjects were shown eight 3-D VR videos and EEG signals are captured using Enobio- 32- channel device with 500 Hz sampling rate. Out of 66 subjects, 30 were female and 36 were male around 35 years average age. The channels used are CP2, C3, P7, FC1, CZ, CP1, P3, FC5, FC2, AF4, FPZ, C4, PO3, F7, AF3, CP5, PZ, O2, FP2, T7, P4, PO4, T8, FC6, F8, CP6, O1 P8, FP1, F3, F4, FZ [21-22].

The complete protocol is shown in Figure 3. The subject is briefed out with the complete protocol and with his/her concern the data acquisition is carried out. Once the subject wears EEG device they are asked to relax for 60 second [Relaxation state]. There after trail begins with presenting each video in the above said format. Trail number is displayed for 2 sec followed by video related to corresponding emotion state for 50 to 300 seconds. The trail ends with self-assessment.

Preprocessing: The acquired raw EEG data includes artifacts like muscle movements, blinking of eye etc. the device had inbuilt notch filter which remove 50Hs line interface noise. Further median filter, moving average referencing filter and a band pass FIR filter with 20th order is implemented using MATLAB command “filtflit” [23,24].

Feature Extraction: On the preprocessed signal, 8-Level DWT decomposition is applied to extract the desired band of frequencies. The extracted features are tabulated in Table I. Total 34 features are extracted for each subjects. The features are computed using MATLAB and channel wise feature extraction is implemented using Python.

Classification: Hybrid Deep Neural Network (HDNN) classifier is applied on the extracted features. In our previous work, [22] sequential ANN was implemented and classification accuracy for 4 states is 90.835% and 8 states is74.045 %. To increase the prediction rate we developed deep learning network using LSTM with a modified frame work on the features vector. For 66 subjects we have 8 trails each and for each trail 34 features were calculated. So the feature vector 34 * 264 which implies 34 features for 264 trails.

The hybrid LSTM [HLSTM] architecture has 10 layers is shown in Figure 4 . The layers are first layer is 1X1 Sequential layer which has 34 input nodes which represents number of features applied for the HDNN. Second layer is Fully connected layer. Third layer is BILSTM with 100 hidden units. Fourth layer is ReLU layer. Fifth and eight layer is Dropout with 0.2 % drop out of nodes occurs. Sixth layer is LSTM layer, which is followed by Leaky ReLu layer and drop out layer. Ninth layer is Softmax layer followed by classification layer which is the last layer of our network.

Hybrid LSTM is implemented as per the steps given below using MATLAB.

Step 1: Reading the channel wise feature vector.
Step 2: Splitting the 70% features for training 'XTrain' and 30% for testing 'XTest'
Step 3: Reading the label and splitting into 'YTrain' and 'YTest'
Step 4: Converting the labels into required form for network using categorical function
Step 5: building layers using deep network designer in MATLAB and analyzing for proper functionality of layers
Step 5: Initializing training options of the layers using "Training Options" command
Step 6: training the network designed using the command "train Network (XTrain, YTrain, layers, options)"
Step7: The trained network is tested using "classify (net, XTest)"
Step 8: Prediction accuracy is calculated and confusion matrix is plotted.

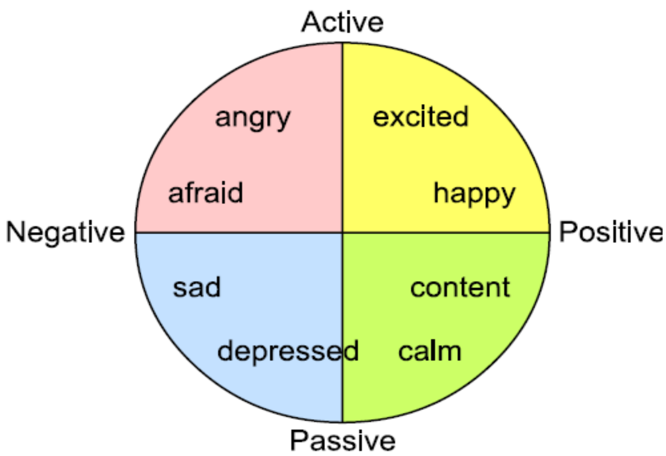

Figure 1: Dimensional model [1]

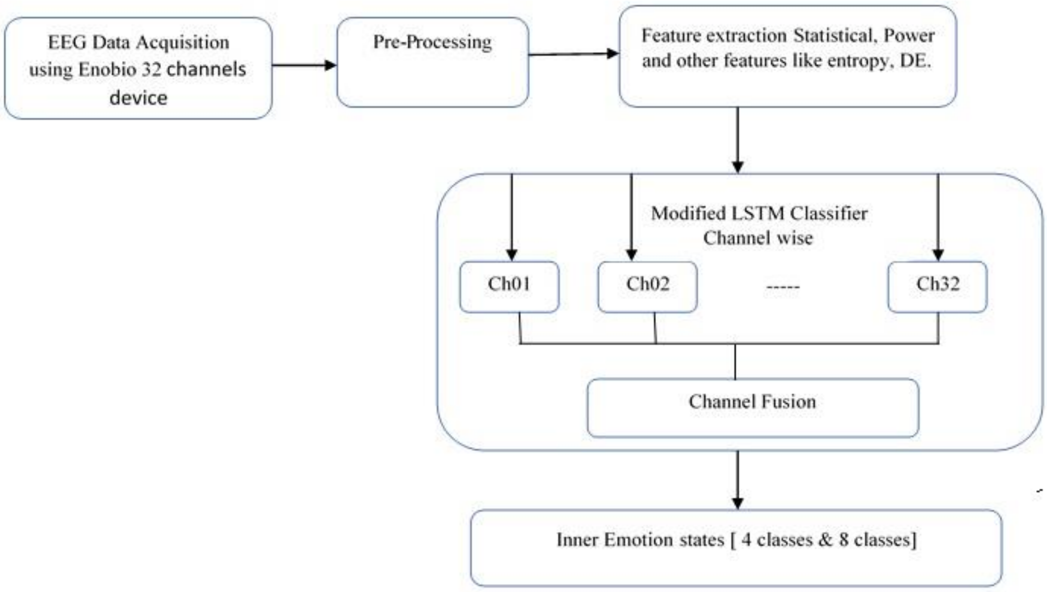

Figure 2: Methodology of proposed system

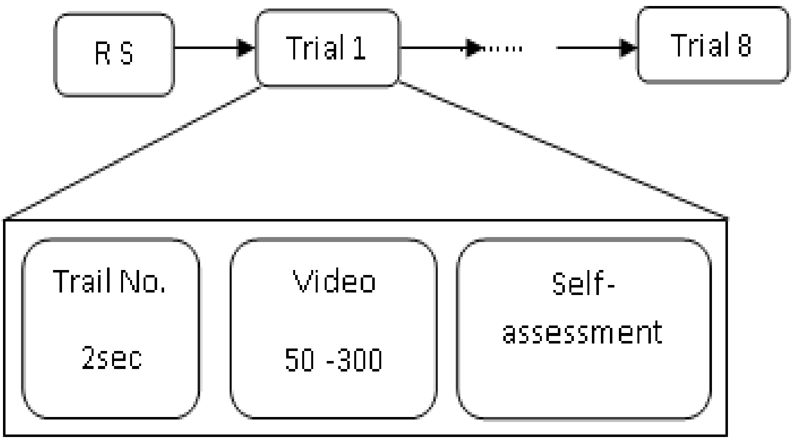

Figure 3: Protocol used for data acquisition [21, 22]

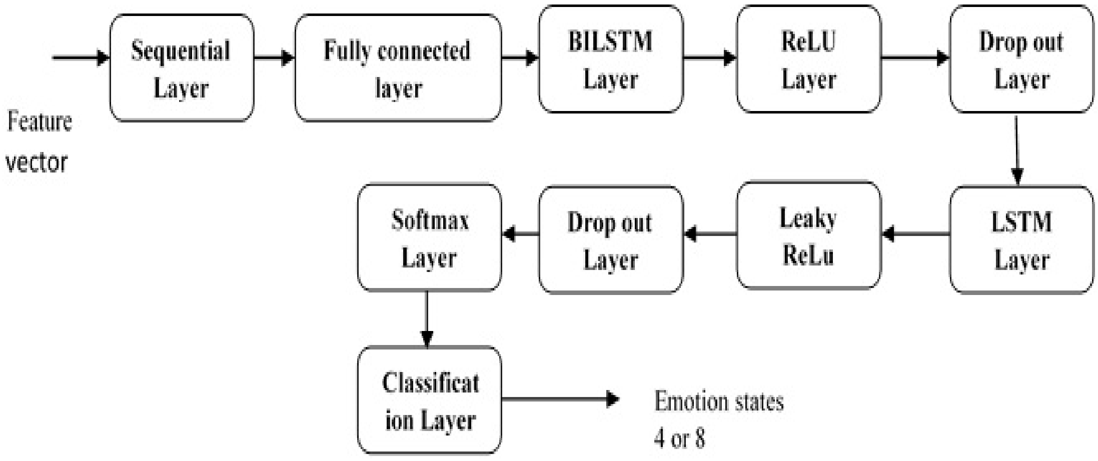

Figure 4: Hybrid LSTM Architecture

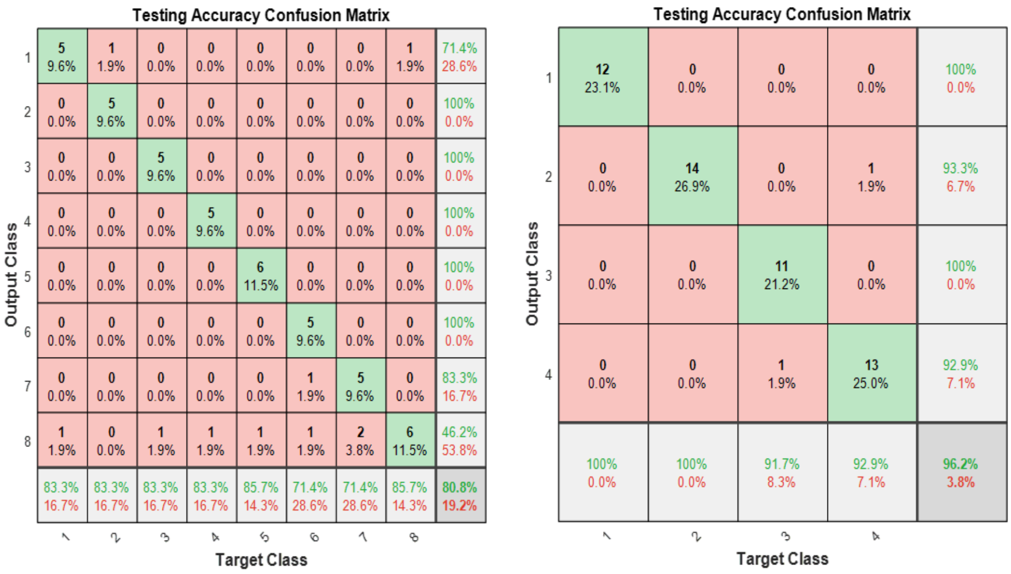

Figure 5: Confusion matrix of 4 & 8 classes of channel FP1

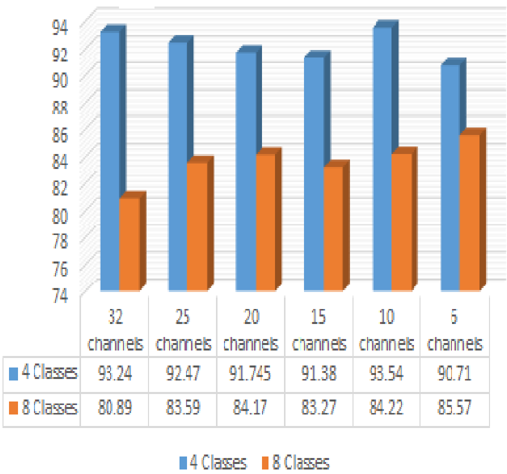

Figure 6: Performance of channel reduction

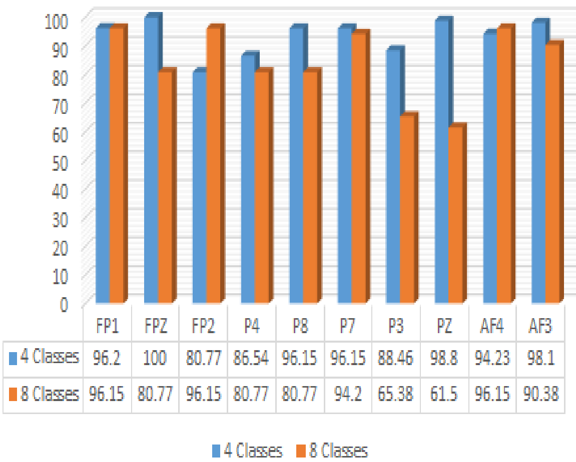

Figure 7: Classifier Accuracy of 10 Channels

Table I: Statistical and Frequency domain features [22]

Feature extracted	**List of Feature**
Statistical Features	Mobility and Complexity (Hijorth Parameters), Skewness ,Kurtosis Mean, Variance, First & second difference, Normalized
Frequency Domain features	Energy, DE, PSD, Average power
Other features	Hurst exponential and Permutation entropy

Table II: Performance of HLSTM classifier

Type of Features	**Four states**	**Eight states**
Combined features (Time and wavelet)	93.24%	80.05%
Wavelet Features	84.24%	76.85%
statistical	67.24%	53.9%

Table III: Classifier accuracy of HLSTM for channel reduction

Number of Channels selected	Four states	Eight states
32	93.24	80.89
25	92.47	83.59
20	91.745	84.17
15	91.38	83.27
10	93.54	84.22
6	90.71	85.57

4. Results and Discussion

Feature extracted form 66 subjects for eight trails are considered for analysis in the present work. Thirty-four combined features, fourteen statistical and twenty wavelet features are applied for H-LSTM network. Out of 264 samples, 70% of them i.e. 184 samples are used for training and reaming 30% i.e. 80 samples are used for testing the network. The prediction accuracy for 8 classes and 4 classes are 80.05% and 93.24% respectively and is tabulated in Table II. The confusion matrix for FP1channel is illustrated in Figure 5.

Further channel reduction is performed based on prefrontal, temporal, parietal and occipital channel to analyze the emotion states [22]. The performance analysis of channel reduction is shown in Table III and Figure 6. Ten predominant channels performance is shown in figure 6. It is observed that for four classes, when channel reduction is done, the performance is better for ten channels. To detect discrete emotions, channel reduction has shown better predication rate. From thirty-four features, differential entropy, PSD and power of 5 bands, Hurst exponential, kurtosis, skewness, 1st and 2nd normalized difference found to be predominant features. It was also keenly noted that HLSTM performed better for frequency domain features than time domain. Conclusion

In this paper, a hybrid LSTM [HLSTM] is proposed to detect the inner emotion states of 66 subjects whose emotions are elicited by projecting 8 3-D VR videos. The protocol is so designed that the emotions can be classified based on 2-dimensional (4 classes) and discrete model (8 classes). The time domain and frequency domain features are extracted after the removal of artifacts and applied to HLSTM network. The prediction accuracy for 4 classes is 93.24% and 80.05% for eight classes. The performance of our hybrid network is found to be better than others.

Further feature reduction can be used in order to improve the accuracy. The work can be compared with EEG signal acquired using normal video. Based on the emotions detected we can build an emotional DJ which plays the songs related to a detected state.

Acknowledgments

The Authors would like to thank, Dr. Manesh B Kokare, Dean Research and Development Associate Professor and PI of Center of Excellence in Signal and Image Processing and Dr. D D Doye, Professor, Department of Electronics and Telecommunication SGGS Institute of Engineering and Technology, Vishnupuri, Nanded, Maharastra, India, for permitting to acquire EEG data as well sharing the same.

References:

[1] Bos, D.O., 2006. EEG-based emotion recognition. The Influence of Visual and Auditory Stimuli, 56(3), pp.1-17.

[2] Alarcao, S.M. and Fonseca, M.J., 2017. Emotions recognition using EEG signals: A survey. IEEE Transactions on Affective Computing, 10(3), pp.374-393.

[3] Li, M., Chai, Q., Kaixiang, T., Wahab, A. and Abut, H., 2009. EEG emotion recognition system. In In-vehicle corpus and signal processing for driver behavior (pp. 125-135). Springer, Boston, MA.

[4] Liu, Y., Sourina, O. and Nguyen, M.K., 2010, October. Real-time EEG-based human emotion recognition and visualization. In 2010 international conference on cyberworlds (pp. 262-269). IEEE.

[5] Garcia-Molina, G., Tsoneva, T. and Nijholt, A., 2013. Emotional brain–computer interfaces. International journal of autonomous and adaptive communications systems, 6(1), pp.9-25.

[6] Li, C., Bao, Z., Li, L. and Zhao, Z., 2020. Exploring temporal representations by leveraging attention-based bidirectional lstm-rnns for multi-modal emotion recognition. Information Processing & Management, 57(3), p.102185...

[7] Sharma, Rahul, Ram Bilas Pachori, and Pradip Sircar. "Automated emotion recognition based on higher order statistics and deep learning algorithm." Biomedical Signal Processing and Control 58 (2020): 101867.

[8] Gao, Z., Wang, X., Yang, Y., Li, Y., Ma, K. and Chen, G., 2020. A Channel-fused Dense Convolutional Network for EEG-based Emotion Recognition. IEEE Transactions on Cognitive and Developmental Systems.

[9] Manshouri, N., Maleki, M. and Kayikcioglu, T., 2020. An EEG-based stereoscopic research of the PSD differences in pre and post 2D&3D movies watching. Biomedical Signal Processing and Control, 55, p.101642.

[10] Garg, A., Kapoor, A., Bedi, A.K. and Sunkaria, R.K., 2019, September. Merged LSTM Model for emotion classification using EEG signals. In 2019 International Conference on Data Science and Engineering (ICDSE) (pp. 139-143). IEEE..

[11] Xing, X., Li, Z., Xu, T., Shu, L., Hu, B. and Xu, X., 2019. SAE+ LSTM: A New framework for emotion recognition from multi-channel EEG. Frontiers in Neurorobotics, 13, p.37..

[12] Li, Y., Zheng, W., Wang, L., Zong, Y. and Cui, Z., 2019. From Regional to Global Brain: A Novel Hierarchical Spatial-Temporal Neural Network Model for EEG Emotion Recognition. IEEE Transactions on Affective Computing.

[13] 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.

[14] Zamanian, H. and Farsi, H., 2018. A new feature extraction method to improve emotion detection using EEG signals. ELCVIA: electronic letters on computer vision and image analysis, 17(1), pp.29-44.

[15] Chao, H., Zhi, H., Dong, L. and Liu, Y., 2018. Recognition of emotions using multichannel EEG data and DBN-GC-based ensemble deep learning framework. Computational intelligence and neuroscience, 2018.

[16] Liu, J., Meng, H., Li, M., Zhang, F., Qin, R. and Nandi, A.K., 2018. Emotion detection from EEG recordings based on supervised and unsupervised dimension reduction. Concurrency and Computation: Practice and Experience, 30(23), p.e4446.

[17] Mohammadi, Z., Frounchi, J. and Amiri, M., 2017. Wavelet-based emotion recognition system using EEG signal. Neural Computing and Applications, 28(8), pp.1985-1990.

[18] Li, X., Yan, J.Z. and Chen, J.H., 2017. Channel Division Based Multiple Classifiers Fusion for Emotion Recognition Using EEG signals. In ITM Web of Conferences (Vol. 11, p. 07006). EDP Sciences.

[19] DEAP data set (https://www.eecs.qmul.ac.uk/mmv/datasets/deap/readme.html)

[20] SEED dataset (http://bcmi.sjtu.edu.cn/home/seed/)

[21] Thejaswini, S. and Ravi Kumar, K.M., 2018. Detection of human emotions using features based on discrete wavelet transforms of EEG signals. International Journal of Engineering and Technology, 7(1.9), pp.119-122.

[22] Thejaswini S, K M Ravi Kumar," A Novel Method to Detect Inner Emotion States of Human Using Artificial Neural Networks", International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8, Issue-6, March 2020, PP : 5820 -5825

[23] Thejaswini, S., Kumar, K.R., Rupali, S. and Abijith, V., 2018. EEG based emotion recognition using wavelets and neural networks classifier. In Cognitive Science and Artificial Intelligence (pp. 101-112). Springer, Singapore.

[24] Thejaswini, S., Ravikumar, K.M., Jhenkar, L., Aditya, N. and Abhay, K.K., 2019. Analysis of EEG Based Emotion Detection of DEAP and SEED-IV Databases Using SVM. Int. J. Recent Technol. Eng, 8, pp.207-211.