Optimal Number of Electrode Selection for EEG Based Emotion Recognition using Linear Formulation of Differential Entropy

Vaishali M. Joshi* and Rajesh B.Ghongade

Electronics Engineering, Bharati Vidyapeeth (Deemed to be University)
College of Engineering, Pune: 411043, Maharashtra, India.
*Corresponding Author E-mail : joshivaishali7@ gmail.com

Anxiety, nervousness and stress are daily challenges for mankind. These challenges are very severe specifically for students of the age group between years 14 to 25. Therefore it very important to develop the simplest, low cost, accurate and handy process which will be helpful for the society to gauge the anxiety levels and take necessary corrective actions to avoid health and psychological   issues. It is of extreme importance to have regular checks on change in behavior and to ensure correct emotion analysis and take the corrective action. Article elaborates unique feature extraction method called as "linear formulation of differential entropy". With this method we have significantly reduced number of (Electroencephalography) EEG channels for emotion detection. This work has discovered new approach in neuroscience. It's proved that, single-channel EEG contains sufficient information for emotion recognition. The performance of the newly proposed technique is based on the "Database for Emotion Analysis using Physiological Signals" (DEAP) benchmark database using single channel FP1, prefrontal channel [FP1, AF3, FP2, AF4], and all 32 channel. Bidirectional long short term memory (BiLSTM) is used as classifier. The performance shows that, accuracy achieved using proposed method of single channel (FP1) is almost equivalent to the accuracy of 32 channels.

Keywords: Bidirectional Long Short Term Memory; Electroencephalography;
Linear Formulation of Differential Entropy.

Day today challenges, success, failures, personal and professional relationship has huge impact on emotions. Little deviations from routines activities make either positive or negative impact on human mind set. Such deviations lead to generate positive, negative or neutral emotions within mind. Every emotion generates impact. Such impact can be noticed through body language, voice modulations, facial changes[7]. Human beings can hide such visible impacts and therefore it is pertinent to have scientific, economical and easiest way to recognise the emotions and provide necessary support to the object for its own behavioural development. Such support is extremely useful to handle pressure due to fears competition and pressure situations. Scientific research work on this subject is helpful for analysing mental health of human.

In this research work portable and wearable single channel device has been proposed. This method is very unique and varies from conventional practice to use bulky multichannel EEG devices.

Zhong Min et al. had explained significance of use of less number of electrodes for EEG based emotion detection. He had

used normalized mutual information (NMI) for enhancing accuracy[15]. W.Zheng proposed group sparse canonical correlation analysis for EEG based emotion recognition. He addressed importance of reduction of number of channels for lowering complexity[20]. K. Ansari Asl et al. highlighted a problem of large number electrode for EEG based emotion detection. He used synchronization likelihood method for emotion detection[1]. J. Zhang et al. proposed method known as Relief to analyze emotions[17]. His findings were issues with use of massive set of EEG electrode and complexity to handle data to while working on arithmetic computations. He also highlighted issue of great inconvenience to the user due to time consuming process of recording various signals from bulky and multi channel device. V. Gupta et al. worked on EEG based emotion recognition method called as ûexible analytic wavelet transform (FAWT) and his findings were also in line with above referred research work[4].

While we discussed on use of less number of channel for better accuracy various researchers suggested number of feature extraction and classification methods. Few of them are as elaborated. Ruo Nan Duan et al. suggested feature extractions method known as differential entropy (DE). DE is further classified as differential asymmetry (DASM) and rational asymmetry (RASM) method. He conclusions suggested that, differential entropy method generates much better results compare to traditional energy spectrum method[3]. L.C. Shi et al. further worked on Differential entropy method and proved that differential entropy is equal to logarithm power spectrum[12]. D.W. Chen et al. presented fusion of differential entropy with linear discriminant analysis (LDA) and explained that , how to reduces time complexity using DE method[2]. W.L. Zheng et al. proposed deep neural network method in which precise frequency bands were selected[18].

The deep learning is very advanced technique used in machine learning. Different deep learning architectures are introduced in the literature. The models with deep architectures attain excellent results which are better than old methods[16, 19 8, 9].

METHODOLOGY

Figure 1 present the work flow. There are four stages viz. database collection, band separation, feature extraction, and an emotion classification.

Collection of the EEG signal database

Established benchmarked, DEAP[6] open data base has been used for analysis. Existing open data base has been generated through experiment
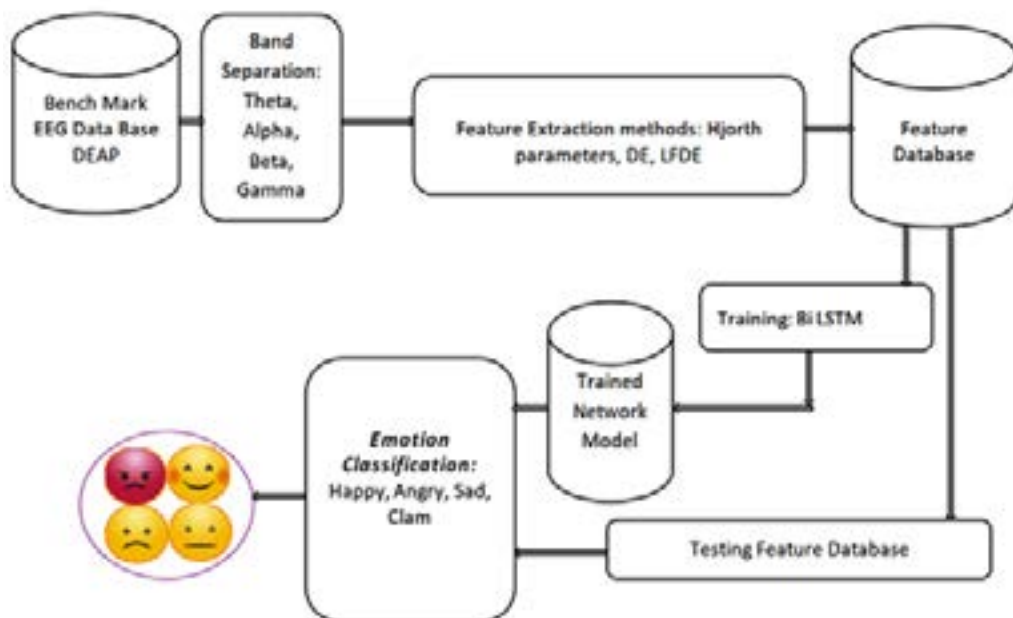

Fig. 1. Work flow block diagram

on 32 subjects. To perform experiment 40 Channels have been used. Out of this 32 Channels were EEG and remaining was peripheral. Each subject was asked to watch 40 clips of one minute. This recording was in line with 10-20 system using sampling frequency of 512 Hz. This 512 Hz frequency has been down sampled to 128 Hz. Such EEG signals are further band passed to 4 to 45 Hz.

Feature Extraction Method

Feature Extraction process plays important role to develop an EEG based brain computer interface (BCI) system. Available literatures have reported different feature extraction methods i.e. Hjorth parameters, wavelet entropy, mean variance, kurtosis, power spectral density (PSD) and differential entropy.

In this research paper, instead of using raw EEG signals, different features extraction methods have been used. Due to use of different features extraction methods, disadvantage of incredibly large size data handling and memory size requirement has been eliminated.

While developing analysis model three types of features namely Hjorth parameters, differential entropy and linear formulation of differential entropy have been used. The features are extracted in each of the frequency range i.e. theta, alpha, beta, and gamma. From Figure 2 it is evident that, EEG signals are consistent with frequency ranges: theta (θ) :- (4-8 Hz), alpha (α) :- (8-14 Hz), beta (β) :- (14-30 Hz), and gamma (γ) :-(30-44 Hz).

Hjorth Parameters[5]

Hjorth Parameters are used in signal processing in time domain analysis. These parameters are described by B. Hjorth in 1970. There are three types of parameters i.e. Activity, Mobility and Complexity.

Activity

The activity parameter measures the signal power and the variance of a time function. This is computed by using the following equation[10]:

$$Activity = var\left(k(t)\right) \qquad ...(1)$$

Where k *(t)* is the EEG signal.

Mobility

The mobility parameter indicates the mean frequency of the power spectrum.

Mobility is the ratio of the square root of variance of the first derivative of the signal k (t) to variance of the signal   k (t)[10].

$$Mobility = \sqrt{\frac{var\frac{dk(t)}{dt}}{var(k(t))}} \qquad ...(2)$$

Complexity

Complexity parameter represents frequency change. It is given by following equation[10],

$$Complexity = \frac{Mobility\frac{dk(t)}{dt}}{Mobility\left(k(t)\right)} \qquad ...(3)$$

Hjorth parameters are helpful to analyze the signals in time domain.

Differential entropy[3]

To determine the complexity of a continuous random variable, differential entropy is used. For normal distribution N (μ, $ä^2$), the differential entropy can be described below:

$$DE = \Box(z) = \frac{1}{2}log_e 2\pi e\delta^2 \qquad ...(4)$$

Where, z is a random variable, δ and e are constants. In a particular band, for a fix length EEG signal, DE is correspondent to the logarithmic spectral energy[12].

Linear Formulation of Differential entropy (LF-DE) using sigmoid functions for fourth order spectral moment of EEG signal

The main purpose of this work is to propose novel feature for EEG signal analysis using EEG trace in terms of slope spread. As fourth order spectral moment compute excessive feature correlation with the sine wave i.e. the "softest" possible curve.

It has been shown, by Hjorth (1970), that the moment of order 2n identical to the variance δ_n^2 of the derivative of order n [5], i.e.,

$$m_0 = \delta_0^2, m_2 = \delta_1^2, m_4 = \delta_2^2, m_6 = \delta_3^2 - - - m_{2n} = \delta_n^2$$

From Eq. 4,

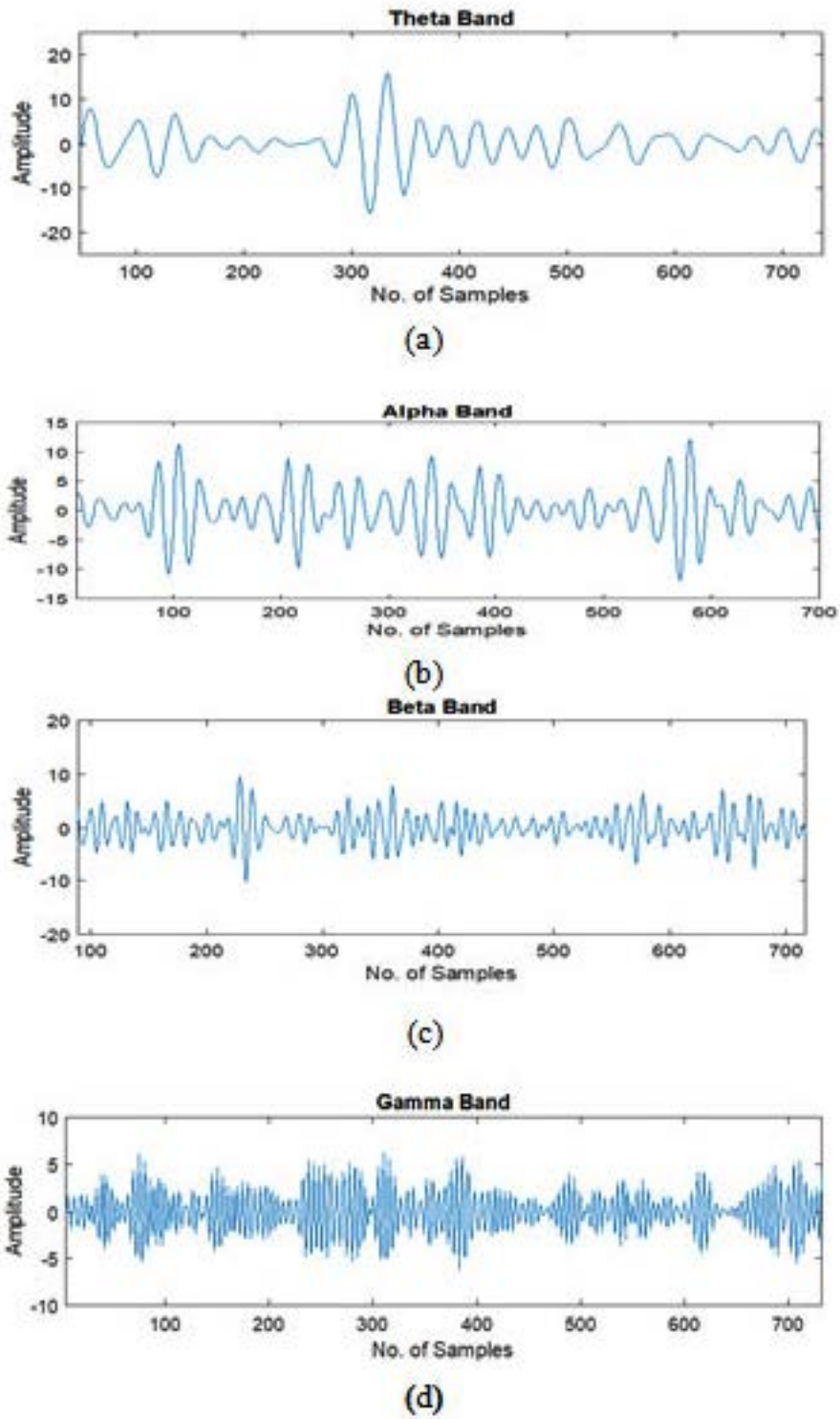$$h(z) = \frac{1}{2}log_e 2\pi e\delta^2$$

Using Sigmoid Function

Fig. 2. EEG Signals of different bands

$$v = \frac{e^t}{1 + e^t}$$ ...(5)

$$v = \frac{e^{\log_e \sqrt{2\pi e \delta^2}}}{1 + e^{\log_e \sqrt{2\pi e \delta^2}}}$$ ...(8)

Let $t = h(z) = DE$

$$\therefore v = \frac{e^{DE}}{1 + e^{DE}}$$ ...(6)

$$v = \frac{\sqrt{2\pi e \delta^2}}{1 + \sqrt{2\pi e \delta^2}}$$ ...(9)

$$v = \frac{e^{\frac{1}{2}\log_e 2\pi e \delta^2}}{1 + e^{\frac{1}{2}\log_e 2\pi e \delta^2}}$$ ...(7)

$$v = \frac{\delta \sqrt{2\pi e}}{1 + \delta \sqrt{2\pi e}}$$ ...(10)

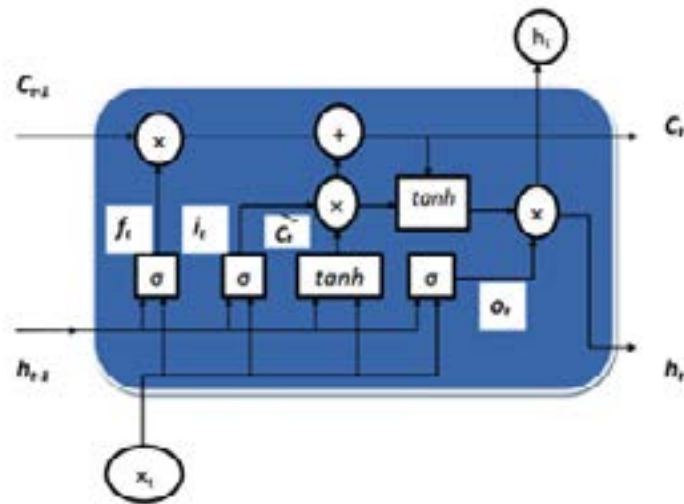$\delta = $ standard deviation of fourth moment of EEG signal

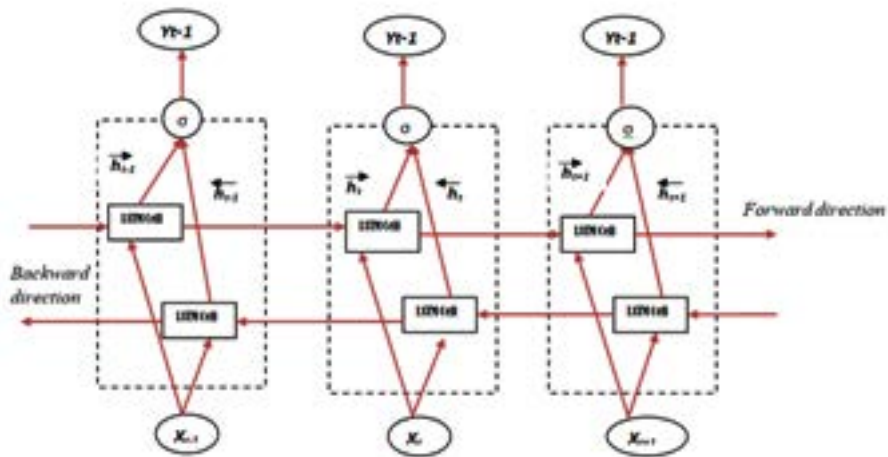**Fig. 3.** LSTM architecture

Fig. 4. Architecture of bidirectional LSTM

Eq. 10 shows, linear formulation of differential entropy using sigmoid function for fourth order spectral moment of EEG signal. Fourth order spectral moment effectively detects nonlinearity and non-Gaussianity in EEG signal. In the frequency and in the time domains, linear formulations are more easily merged[5].

Classification

In this paper, BiLSTM is used as classiûers for examine the performance of proposed feature extraction method.

Long-Short Term Memory (LSTMS)[11]

Long-Short Term Memory Networks (LSTMs) are a particular class of recurrent neural networks (RNN). Same is invented by Hochreiter
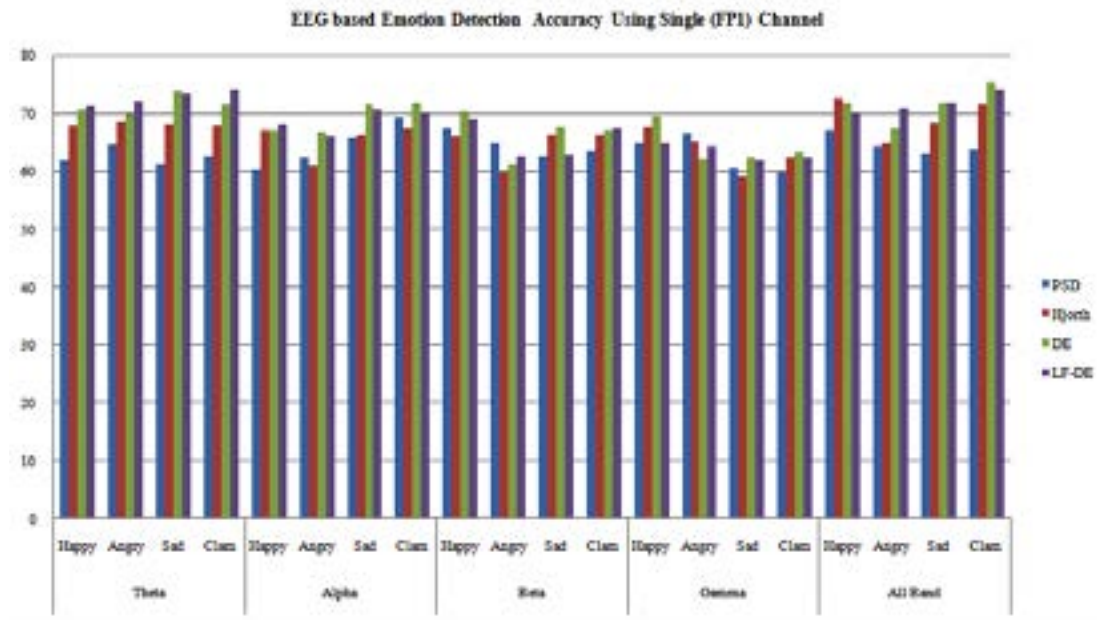

Fig. 5. EEG based Emotion Detection Using Single Channel
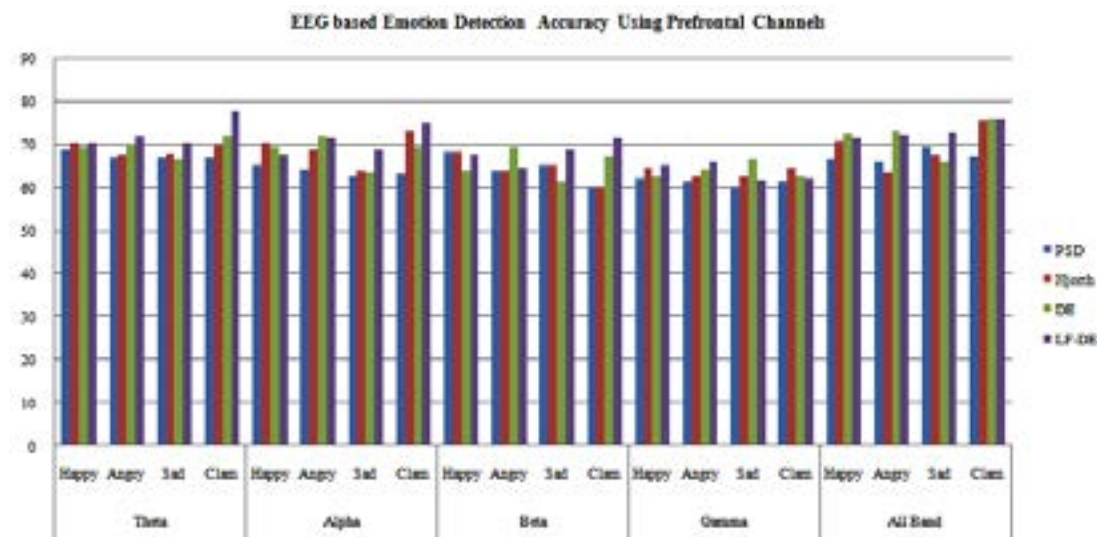

Fig. 6. EEG based Emotion Detection Using Prefrontal Channels

and Schmidhuber. To address the problem of long-term dependency in RNN, LSTM is introduced as shown in Figure 3. Long sequences pose challenges to learn from standard RNN and cause the problem of vanishing/exploding gradient. Lengthy sequences are memorized by the LSTM. LSTM can very efficiently extract the temporal and spatial features from EEG signal. Figure 4 depicts the structure of a BiLSTM layer, comprising a forward and a reverse LSTM layer. The LSTM comprises of three control units, namely input gate (5ØVÜ5ØaÜ), output gate (5Ø\Ü5ØaÜ) and forget gate (5ØSÜ5ØaÜ) at time 5ØaÜ, respectively.

A bidirectional LSTM network is used to advance systems performance. Bidirectional LSTM network has the ability to store more information as compared with unidirectional networks. The proposed design is implemented using BiLSTM.

Bi-LSTMs have two hidden layers, one in forward and other in a backward direction[14]. The forward hidden layer train input sequences the same as conventional unidirectional LSTM and backward hidden layer train input sequence in reverse order. The purpose of using a backward and forward hidden layer is to observe past and future data and learns its weights accordingly. Hence Bi-LSTM can use past and future context of a given input sequence [14].

Implementation details

The parameters used for BiLSTM are: Layer: BiLSTM, Gradient Decay Factor: 0.9000, Squared Gradient Decay Factor: 0.9990, Epsilon: 1.0000e-08, Initial Learn Rate: 1.0000e-03, BiLSTM with 200 hidden units and Adam (adaptive moment estimation) optimizer is used.

Table 1. Performance comparison of related work using DEAP Database

| Sr. No.	Name of Author	Method	No. of Channels	Name of class	Average Accuracy (%)
1.	Sander Koelstra et al 6	Power spectral	32	Valance, Arousal	59.8
2	Gyanendra K Verma et al 13	MLP multilayer perceptron	32	Valance, Arousal	66.55
3	Proposed	Deep learning	32	Happy, Angry, Sad and Clam	74
4	Proposed	Deep learning	4	Happy, Angry, Sad and Clam	73.37
5	Proposed	Deep learning	1	Happy, Angry, Sad and Clam	71.87
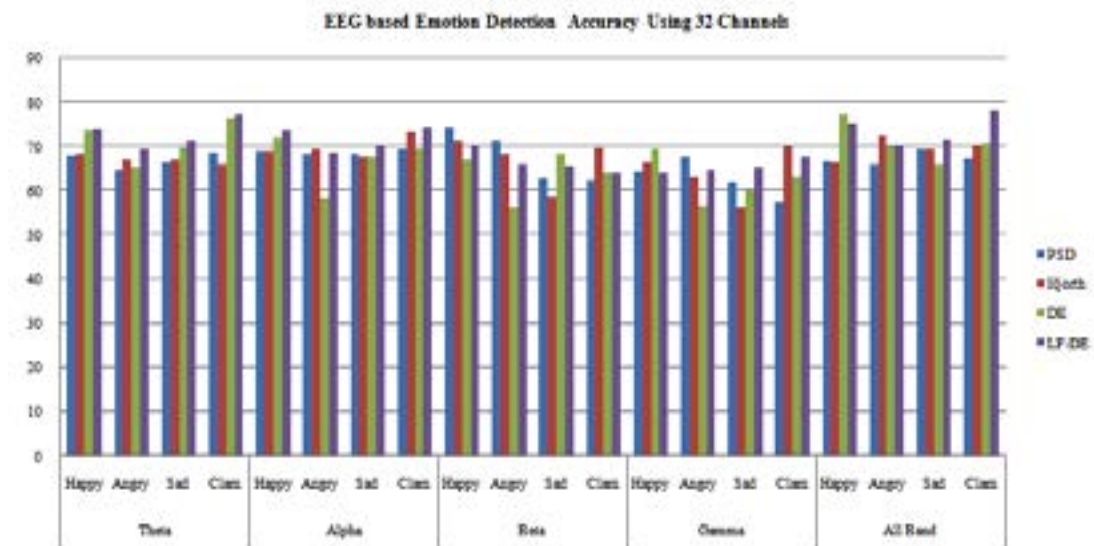
Fig. 7. EEG based Emotion Detection Using 32 Channels

RESULT AND DISCUSSION

From research work, it is found that currently there is no robust single channel feature extractor available with an equivalent performance of multichannel EEG device. Apart from this conventional classifier does not address the need of temporal property of EEG signal while extracting feature. Deep learning algorithm of BiLSTM provides solution on this issue. BiLSTM can learn both spatial and temporal feature and therefore enhances performance and reduces the noise. Before arriving at this conclusion, we have evaluated three feature extraction methods namely LF-DE, DE and Hjorth parameters. Findings of proposed with BiLSTM classifier with single channel, prefrontal and 32 channels are shown in Figure 5, 6, and 7 respectively. Comparison of results of proposed LF-DE feature extraction method with other existing method is shown in Table 1. It is evident that average accuracy achieved using single channel (FP1) is 71.87% which is very close to the 32 channels average accuracy of 74%. Same is be seen graphically in Figure 5, 6, and 7.

CONCLUSION

In this paper, a new feature extraction method "Linear Formulation of differential Entropy" (LF-DE) is proposed. The result obtained here using proposed method for single channel are almost equivalent to that of multichannel devices. The proposed work will help the society to a great extent. The frequent hospital visits can be reduced and diagnosis can be done by the individual using single channel EEG device and with same accuracy level. Single channel EEG device with dry electrode FP1 are commercially available at economical prices and are very convenient to use traditional approach to use full-channel EEG signals has huge disadvantage as this approach calls for handling of excessive and unwanted data which also leads to higher memory requirements and hence costly hardware and firmware requirements. Use of BiLSTM to classify emotional state helps to reduce handling of EEG channels from 32 to 1. The results shown above substantiate that, the proposed method of single channel use with the help of BiLSTM effectively improves the accuracy of emotion recognition.

ACKNOWLEDGEMENT

REFERANCES

1 Karim Ansari-Asl, Guillaume Chanel, and Thierry Pun, 'A Channel Selection Method for Eeg Classification in Emotion Assessment Based on Synchronization Likelihood', in *2007 15th European Signal Processing Conference* (IEEE, 2007), pp. 1241-45.

2 Dong-Wei Chen, Rui Miao, Wei-Qi Yang, Yong Liang, Hao-Heng Chen, Lan Huang, Chun-Jian Deng, and Na Han, 'A Feature Extraction Method Based on Differential Entropy and Linear Discriminant Analysis for Emotion Recognition', *Sensors,* **19**: 1631 (2019).

3 Ruo-Nan Duan, Jia-Yi Zhu, and Bao-Liang Lu, 'Differential Entropy Feature for Eeg-Based Emotion Classification', in *2013 6th International IEEE/EMBS Conference on Neural Engineering (NER)* (IEEE, 2013), pp. 81-84.

4 Vipin Gupta, Mayur Dahyabhai Chopda, and Ram Bilas Pachori, 'Cross-Subject Emotion Recognition Using Flexible Analytic Wavelet Transform from Eeg Signals', *IEEE Sensors Journal,* **19**: 2266-74 (2018).

5 Bo Hjorth, 'Eeg Analysis Based on Time Domain Properties', *Electroencephalography and clinical neurophysiology,* **29**: 306-10 (1970).

6 Sander Koelstra, Christian Muhl, Mohammad Soleymani, Jong-Seok Lee, Ashkan Yazdani, Touradj Ebrahimi, Thierry Pun, Anton Nijholt, and Ioannis Patras, 'Deap: A Database for Emotion Analysis; Using Physiological Signals', *IEEE transactions on affective computing,* **3**: 18-31 (2011).

7 Prashant Lahane, and Mythili Thirugnanam, 'A Novel Approach for Analyzing Human Emotions Based on Electroencephalography (Eeg)', in *2017 Innovations in Power and Advanced Computing Technologies (i-PACT)* (IEEE, 2017), pp. 1-6.

8 Xiang Li, Dawei Song, Peng Zhang, Guangliang Yu, Yuexian Hou, and Bin Hu, 'Emotion Recognition from Multi-Channel Eeg Data through Convolutional Recurrent Neural Network', in *2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* (IEEE, 2016), pp. 352-59.

9 Yang Li, Wenming Zheng, Lei Wang, Yuan Zong, and Zhen Cui, 'From Regional to Global Brain: A Novel Hierarchical Spatial-Temporal Neural Network Model for Eeg Emotion Recognition',

IEEE Transactions on Affective Computing (2019).

10   Anita Patil, Chinmayee Deshmukh, and AR Panat, 'Feature Extraction of Eeg for Emotion Recognition Using Hjorth Features and Higher Order Crossings', in *2016 Conference on Advances in Signal Processing (CASP)* (IEEE, 2016), pp. 429-34.

11   Jürgen Schmidhuber, and Sepp Hochreiter, 'Long Short-Term Memory', *Neural Comput,* 9 (1997), 1735-80.

12   Li-Chen Shi, Ying-Ying Jiao, and Bao-Liang Lu, 'Differential Entropy Feature for Eeg-Based Vigilance Estimation', in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (IEEE, 2013), pp. 6627-30.

13   Gyanendra K Verma, and Uma Shanker Tiwary, 'Multimodal Fusion Framework: A Multiresolution Approach for Emotion Classification and Recognition from Physiological Signals', *NeuroImage,* **102**: 162-72 (2014).

14   Li Wang, Weijian Huang, Zhao Yang, and Chun Zhang, 'Temporal-Spatial-Frequency Depth Extraction of Brain-Computer Interface Based on Mental Tasks', *Biomedical Signal Processing and Control,* **58**: 101845 (2020).

15   Zhong-Min Wang, Shu-Yuan Hu, and Hui Song, 'Channel Selection Method for Eeg Emotion Recognition Using Normalized Mutual Information', *IEEE Access,* **7**; 143303-11 (2019).

16   Yilong Yang, Qingfeng Wu, Ming Qiu, Yingdong Wang, and Xiaowei Chen, 'Emotion Recognition from Multi-Channel Eeg through Parallel Convolutional Recurrent Neural Network', in *2018 International Joint Conference on Neural Networks (IJCNN)* (IEEE, 2018), pp. 1-7.

17   Jianhai Zhang, Ming Chen, Shaokai Zhao, Sanqing Hu, Zhiguo Shi, and Yu Cao, 'Relieff-Based Eeg Sensor Selection Methods for Emotion Recognition', *Sensors,* **16**: 1558 (2016).

18   Wei-Long Zheng, and Bao-Liang Lu, 'Investigating Critical Frequency Bands and Channels for Eeg-Based Emotion Recognition with Deep Neural Networks', *IEEE Transactions on Autonomous Mental Development,* **7**: 162-75 (2015).

19   Wei-Long Zheng, Jia-Yi Zhu, Yong Peng, and Bao-Liang Lu, 'Eeg-Based Emotion Classification Using Deep Belief Networks', in *2014 IEEE International Conference on Multimedia and Expo (ICME)* (IEEE, 2014), pp. 1-6.

20   Wenming Zheng, 'Multichannel Eeg-Based Emotion Recognition Via Group Sparse Canonical Correlation Analysis', *IEEE Transactions on Cognitive and Developmental Systems,* **9**: 281-90 (2016).