**EEG Based Emotion Recognition Using SpinalLSTM**

1Shreedatta Nasik, 2Shreyas C P, 3Snehit Kinagi, 4Suhas Reddy S

*1,2,3,4Information Science and Engineering*, Bengaluru Institute of Technology, Bengaluru, India

1[shreedattanasik@gmail.com](mailto:shreedattanasik@gmail.com), 2[shreyascp387@gmail.com](mailto:shreyascp387@gmail.com), [3snehitkinagi@gmail.com](mailto:3snehitkinagi@gmail.com), 4suhasreddys2001@gmail.com

***Abstract*—This paper presents an innovative approach for emotion recognition utilizing Electroencephalography (EEG) signals, employing a two-stage framework comprising a Power Spectral Density (PSD) feature extractor and a SpinalLSTM classifier. The PSD feature extractor transforms raw EEG signals into frequency-domain representations, capturing the underlying neural oscillations associated with emotional states. Subsequently, the SpinalLSTM classifier, incorporating both spatial and temporal dependencies through its architecture, effectively learns discriminative patterns for valence and arousal classification. The proposed model is rigorously evaluated on the DEAP database, adopting a subject noncontingent experimental paradigm. Results demonstrate an average accuracy of 70% and 62% for valence and arousal classification, respectively, with peak accuracies reaching 85% for valence and 82% for arousal. Comparative analysis against existing methodologies underscores the efficacy of the PSD-SpinalLSTM network in enhancing emotion detection accuracy from EEG signals, highlighting its potential to advance research in affective computing.**

***Keywords—Electroencephalography (EEG); DEAP dataset; Power Spectral Density; SpinalLSTM***

1. INTRODUCTION

Emotion detection poses a significant challenge because of its intricate interplay of psychophysiological and neurobiological factors, complicated further by humans' adeptness at concealing their true emotional states behind external expressions [1]. Despite considerable efforts in facial, vocal, and gestural analysis, the reliability of such methods remains questionable. Consequently, there is a critical need for a more reliable and authentic approach to discerning emotions.

Electroencephalography (EEG) signals offer a promising avenue for addressing this challenge, providing an unbiased window into the mind's state, impervious to external influences [2]. Zohreh Gholami Doborjeh et al. [3] underscored the importance of EEG as spatio-temporal brain data (STBD), reflecting the brain's evoked neuronal activity. Emotion detection through EEG often involves subject dependent and subject independent approaches, with the former comparing an individual's current state to their own baseline data, typically yielding higher accuracy. However, practical limitations may hinder its widespread applicability, necessitating the development of subject-independent models. Experimental evaluations are conducted on the 'Database for Emotion Analysis using Physiological signals' (DEAP) database employing a subject noncontingent methodology. The performance of the proposed method is juxtaposed against existing techniques, showcasing its efficacy in extracting emotions from EEG signals.

The varied origin of EEG signals across different regions of the brain significantly influences emotion recognition. Leveraging both spatial and temporal information inherent in EEG signals is crucial for enhancing recognition accuracy. A key challenge lies in effectively extracting discriminative features and constructing deep learning models for classification in EEG-based emotion detection systems. Numerous researchers have addressed these challenges, proposing diverse methodologies and approaches to tackle this complex task [4].

The structure of the remainder of this paper is as follows: Section I surveys the state-of-the-art approaches and Section III introduced the data used in this paper. The proposed method is presented in Section IV. Section V contains the results and their discussion. Finally, the conclusion is presented in Section VI.

1. LITERATURE SURVEY

Chen et al. [5] employed the DEAP database and hierarchical bidirectional gated recurrent unit (GRU) models with attention mechanisms for classification. Their investigation revealed that 1-second segmented EEG sequences yielded optimal results, achieving average classification accuracies of 67.9% for valence and 66.5% for arousal using the subject noncontingent approach. Qing et al. [6] proposed a feature extraction approach incorporating differential entropy (DE), first-order, and second-order derivative features coupled with autoencoder classifiers. Their study, utilizing the DEAP and SEED databases, achieved accuracy rates of 63.09% and 75% for noncontingent approaches, respectively. Pandey et al. [7] explored empirical mode decomposition (EMD) and variational mode decomposition (VMD) features with a deep neural network (DNN) classifier in subject contingent experiments using the DEAP database, attaining average accuracy levels of 61.25% for arousal and 62.5% for valence states. Arjun et al [8]. proposed an unsupervised Long Short-Term Memory (LSTM) with channel-attention autoencoder for subject-invariant latent

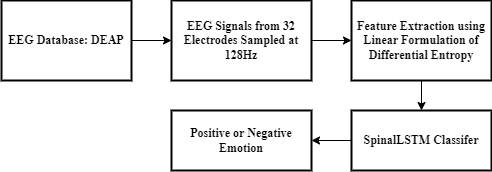


Figure 1. The proposed system of emotion recognition.

vector subspace extraction, followed by a convolutional neural network (CNN) with attention framework for subject-independent emotion recognition. Their approach, evaluated on both SEED and DEAP datasets, achieved average accuracies of 76.7% on SEED and 65.9% and 69.5% for arousal and valence on DEAP, respectively. Joshi and Ghongade [9] introduced a LF-DE feature extractor combined with a BiLSTM classifier for subject dependent and independent experiments, achieving 80.4% accuracy on SEED subject interdependent and 76% and 75.5% accuracy on DEAP valence and arousal states, respectively. Chao et al [10]. proposed a novel method integrating synchronization dynamics and deep learning for multichannel EEG-based emotion recognition. Their approach, utilizing Mutual Information Clustering (MIC) features and

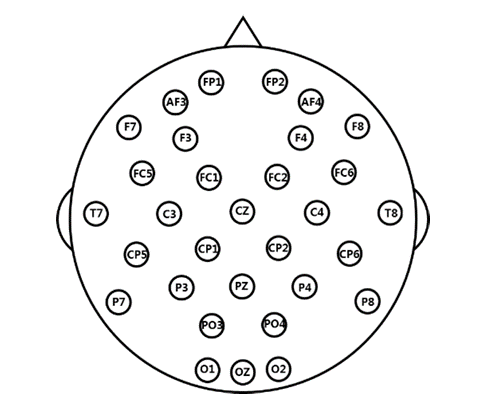
MIC gray images along with a PCA network-based deep learning model and linear SVM classifier, yielded promising average accuracies of 71.85% for arousal and 70.21% for valence on the DEAP dataset.

The current paper proposes an innovative approach to emotion recognition using EEG signals, harnessing the Power Spectral Density (PSD) feature extractor to capture pertinent frequencies and the SpinalLSTM network to discern long-term dependencies and spatial information across diverse brain regions. Specifically, the PSD extractor encapsulates both localized fluctuations and overarching dependencies within EEG signals, while the SpinalLSTM network adeptly learns temporal and spatial patterns crucial for enhancing classification accuracy. The proposed method is shown in the Figure [1]. It has two important phase the feature extraction, and classification.

1. DATA PREPERATION

The experimentation in this research was conducted utilizing the widely recognized dataset: the DEAP dataset [11], a commonly used benchmark in research. The Database for Emotion Analysis using Physiological Signals (DEAP) serves as a benchmark for affective EEG analysis, collected under controlled laboratory conditions. It comprises 32-channel EEG and 8-channel peripheral physiological signals from 32 subjects, with additional positive video recordings for 22 subjects. Electrode placements for EEG are illustrated in Figure [2]. Emotional states were elicited through 40 1-minute music videos, each corresponding to a different emotional state. Subjects rated arousal, valence, like/dislike, dominance, and familiarity of each trial on a scale of 1-9 using Self-Assessment Manikin (SAM). Emotions were defined according to the valence-arousal emotion model, dividing the two-dimensional emotional space into four regions: high valence-high arousal (HVHA), high valence-low arousal (HVLA), low valence-high arousal (LVHA), and low valence-low arousal (LVLA).

The DEAP dataset offers two versions of physiological signal data: raw and pre-processed.

Figure 2. Location of EEG electrodes in DEAP.

While raw data may yield varied results due to pre-processing, such as noise reduction, this study ensures consistency by utilizing pre-processed data. The pre-processed DEAP data comprises 32 channels of EEG signals (at 128Hz) and 8 channels of peripheral physiological signals.

1. METHODOLOGY

As it was mentioned earlier the whole approach has two phases which are feature extraction, and classification. In this section we will deep dive into the methods we have used for the respective phases.

1. *Feature Extraction*

Electroencephalography (EEG) features are crucial in designing human brain-computer interfaces, with various features discussed in the literature. This paper focuses on evaluating emotion detection performance using three specific features: power spectral density, differential entropy, and the linear formulation of differential entropy.

* 1. *Power Spectral Density:* It is a fundamental technique in signal processing, particularly in the analysis of EEG signals. It provides valuable understanding into the frequency content of a signal, offering a detailed understanding of the underlying physiological processes. PSD represents the distribution of power in a signal across different frequency components. It quantifies the contribution of each frequency to the overall signal power, making it an essential tool for studying the spectral characteristics of EEG signals. In our research, we utilized the Welch method for calculating PSD from EEG signals. In our analysis, we considered individual frequencies ranging from 0Hz to 45Hz, covering a wide spectrum relevant to EEG signals. Unlike grouping frequencies into traditional EEG frequency bands, we examined the PSD of each individual frequency separately. This approach allows for a more detailed exploration of the frequency-domain characteristics of EEG signals, enabling us to capture subtle variations and nuances in the signal's spectral profile.
  2. *Differential Entropy:* It quantifies the uncertainty of a continuous random variable X. It's calculated by integrating the probability density function (PDF) f(x) multiplied by the natural logarithm of f(x) over X's support region S. For X following a Gaussian distribution N(μ, σ2), DE simplifies to
     1. , (1)

where μ is the mean and σ2 is the variance. DE is crucial for assessing randomness in data, particularly for time series governed by Gaussian distributions.

* 1. *Linear formulation of Differential Entropy:* The EEG signals are random in nature [9], and for analysing the nonlinearities in these signals, higher-order statistics (HOS) are utilized. HOS are employed to describe higher-order statistical characteristics of a random process, specifically the higher-order spectral moment. In this research, we focus on the fourth-order spectral moment (LF-DE). An EEG signal over a short time period can be represented as the time function g(t). Linear measures are more straightforward to integrate within both frequency and time domains. Consequently, a logarithmic differential equation is transformed into a linear form using a sigmoid equation.Using (1),

By using Sigmoid equation

(2)

Let k= h(X)=DE

(3)

(4)

(5)

(6)

The spectral moment of order is equal to the variance n of the derivative of order .

…. are the variances of the first and second order derivatives of the time function g(t). The fourth-order spectral moment (m4) is utilized because it provides valuable details about the EEG signal. Substitute

in (6)

(7)

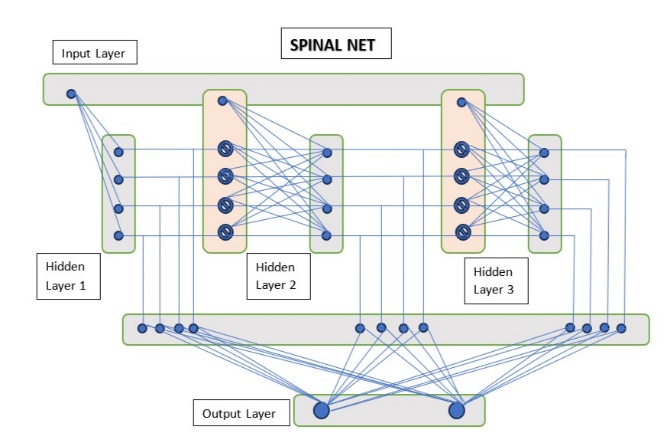


Figure 3. Architecture of SpinalNet

(7) depicts the mathematical equation of Linear Formulation of Differential Entropy with the sigmoid equation. The LF-DE is calculated by segmenting the data samples of each seconds.

1. *Classification Methods*

To validate the effectiveness of the proposed feature extraction method, multi- layer perceptron (MLP), SpinalNet, LSTM and SpinalLSTM classifiers are used.

1. *The Multi-layer perceptron (MLP*): MLP stands as a preeminent architecture within the realm of artificial neural networks (ANNs), particularly renowned for its prowess in classification tasks. The process of classification entails two pivotal phases: training and testing. During training, the MLP undergoes iterative optimization using the popular ADAM technique. This optimization methodology enables the MLP to adapt its parameters and learn from the training data, gradually reducing the training error towards an acceptable level. Training typically concludes once the model achieves a satisfactory level of accuracy on the training data. Subsequently, the trained MLP is subjected to the testing phase, wherein it encounters new, unseen data. This phase serves as a litmus test for the generalization ability of the MLP, evaluating its performance on data that it has not been previously exposed to.
2. *SpinalNet:*SpinalNet is a novel neural network structure inspired by the human somatosensory system, aimed at improving performance while reducing computational overhead. Traditional DNNs struggle with large input feature sets and vanishing gradients, hindering effective training. SpinalNet addresses these challenges by incorporating a gradual and distributed input mechanism, similar to the human spinal cord's processing. Comprising input, intermediate, and output sub-layers, SpinalNet efficiently handles split inputs, minimizing computational complexity [12]. By combining nonlinear activation functions

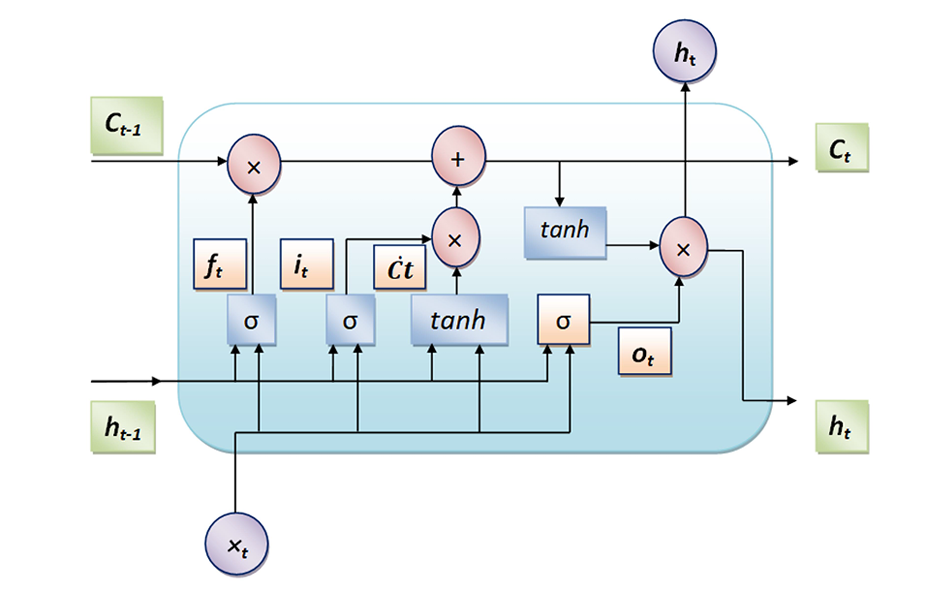


Figure 4. LSTM cell

in the intermediate layers with linear functions in the output layer, SpinalNet achieves robust performance with reduced overhead. Experimental results demonstrate SpinalNet’s superiority over traditional DNN architectures across various datasets. Integration of SpinalNet into popular Deep CNNs like VGG-16 further enhances performance, highlighting its adaptability and potential for future advancements in neural network design [13]. The architecture of the SpinalNet is shown in the Figure [3].As you can see from the figure how it distinguishes itself from the normal neural networks.

1. *LSTM:* The LSTM (Long Short-Term Memory) architecture comprises memory cells, input gates, forget gates, and output gates, as illustrated in Figure [4]. These components collaboratively manage the retention and discarding of information over time, enabling LSTMs to effectively capture patterns and dependencies in sequential data. At the heart of LSTM networks are the memory cells, which store information over time and control the flow of data through the network. This capability allows LSTMs to maintain relevant information while discarding irrelevant or redundant data. The input gate controls the inflow of new information into the memory cell, determining which parts of the current input should be stored. The forget gate, on the other hand, decides which information from the previous memory cell state should be discarded, preventing the network from being overloaded with irrelevant past data. Lastly, the output gate regulates the flow of information from the memory cell to the output, deciding which information from the current state of the memory cell should be used to generate the final output.
2. *SpinalLSTM:* Even though the given dataset has large number of data, most of the data will only be used for feature extraction and nothing else due to the shortage of labels. The number of class labels are 1280. The training data should also be 1280 for MLP or SpinalNet. In order to utilize the data we try to integrate LSTM to SpinalNet obtaining SpinalLSTM as shown inthe Figure [5] the Spinal LSTM architecture integrates LSTM layers with a structure resembling the incremental processing mechanism of the human spinal cord. It features multiple layers with gradual and distributed input reception, aimed at enhancing the model's performance in sequence data processing tasks. LSTMs feature previously mentioned cells. Together, they manage information, enabling effective pattern learning in sequential data. The core of LSTM networks, memory cells store

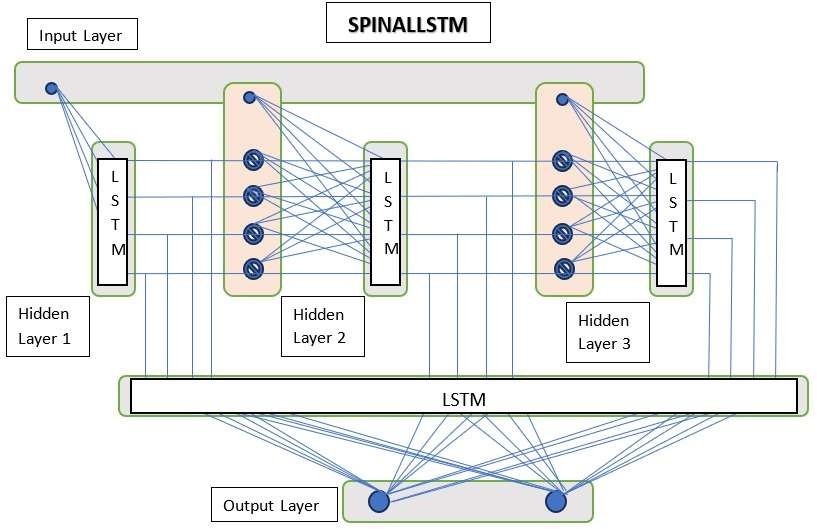


Figure 5. Architecture of SpinalLSTM

information over time and regulate the flow of information through the network as shown in the Figure [4]. This enables LSTMs to retain relevant information while discarding irrelevant or redundant data

1. RESULTS AND DISCUSSION

This section provides a thorough assessment of the proposed architecture, conducting a detailed comparison for both its key components:

1. The Power Spectral Density approach is extensively evaluated against alternative feature extraction methods such as Differential Entropy and Linear Formulation of Differential Entropy.
2. Furthermore, the effectiveness of the proposed SpinalLSTM classifier is rigorously assessed by comparing it with other time series classifiers, including SpinalNet, LSTM and Multiple Layer Perceptron.

Through meticulous analysis and comparison, this section aims to offer valuable insights into the performance and capabilities of the proposed architecture, shedding light on its strengths and potential areas for improvement. To ensure subject-independent accuracy across specified classes, a "leave-one-subject-out" approach was employed for training. This method involves reserving one subject for evaluation while using the remaining subjects for training. This evaluation process was repeated for each subject in the dataset, and the resulting accuracies were averaged to determine subject-independent accuracy. The implementation utilized Python 3.11.9 and TensorFlow 2.15.0.

All proposed models, including MLP, SpinalNet, LSTM and SpinalLSTM, were trained using the ADAM optimizer with a learning rate set to 0.001. For each method, detailed hyperparameters are presented in Tables [1], [2], [3] and [4].

In most of the approaches using DEAP dataset, emotional states, which are preferably discretized into various states such as excitement, anxiety, rage, pleasure, surprise, and so on, are broadly divided into two approximate dimensions: valence and

arousal. The valence variable determines the emotion's positive or negative effects, while the arousal dimension determines its intensity. The results obtained on the DEAP dataset by experimenting with various combinations of Feature extractor and classifier architectures is shown in Table [5].

Figure [6] compares the integration of Power Spectral Density (PSD) and Spinal LSTM with other state-of-the-art architectures.

Table 1. Hyperparameter details of MLP

|  |  |
| --- | --- |
| Hyperparameters | Value |
| Number of layers  Optimizer  Hidden units  Learning rate  Epochs | 2  Adam  256,256  0.001  100 |

Table 2. Hyperparameter details of SpinalNet

|  |  |
| --- | --- |
| Hyperparameters | Value |
| Number of hidden layers  Optimizer  Hidden units  Learning rate  Epochs  Features per layer | 11  Adam  20,30,30,30,30,30,30,30,30,30,30,30  0.001  100  1 |

Table 3. Hyperparameter details of LSTM

|  |  |
| --- | --- |
| Hyperparameters | Value |
| Number of hidden layers  Optimizer  Learning rate  Epochs | 1  Adam  0.001  100 |

Table 4. Hyperparameter details of SpinalLSTM

|  |  |
| --- | --- |
| Hyperparameters | Value |
| Number of hidden layers  Optimizer  Number of LSTM cells per layer  Learning rate  Epochs  Features per layer | 11  Adam  50,50,50,50,50,50,50,50,50,50,50  0.001  100  1 |

Table 5. Validation results for DEAP dataset with varied feature extraction and classifier

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Classifier |  | Feature Extractor  DE | LF\_DE | PSD |
| MLP | Valence  Arousal | 50.2%  51.78% | 52.1%  52.3 | 54.2%  52.% |
| SpinalNet | Valence  Arousal | 52.2%  52.8% | 53%  51.7% | 54.91%  52.61% |
| LSTM | Valence  Arousal | 55.7%  53.2% | 56.1%  53.3% | 70.%  62.2% |
| SpinalLSTM | Valence  Arousal | 53.4%  53% | 54.5%  53.2% | **72%**  **62.2%** |

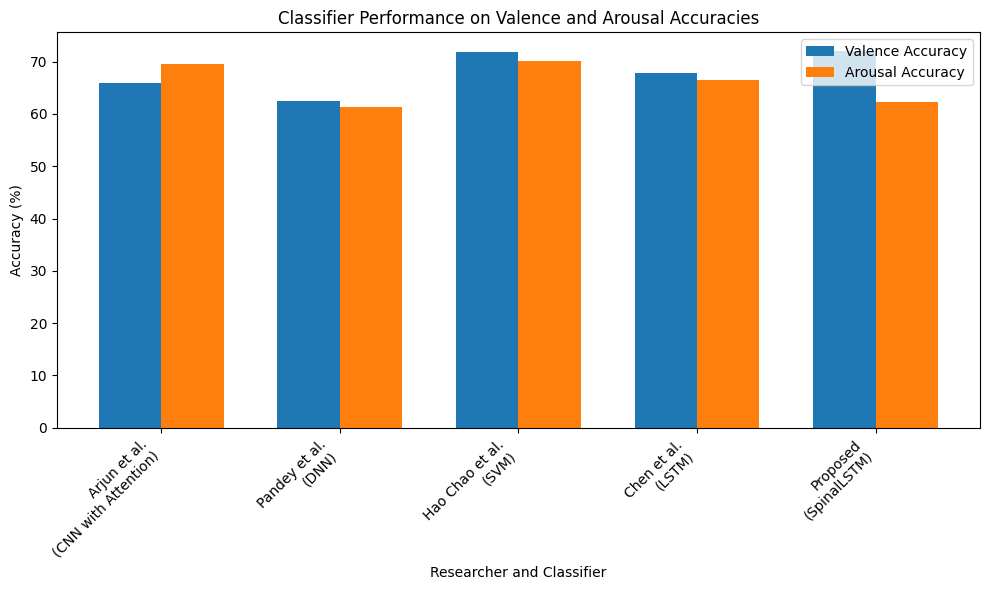


Figure 6. Comparison with state-of-the-art approaches

1. CONCLUSION

In conclusion, this research paper investigated the effectiveness of several feature extraction methods, including Power Spectral Density, Linear Formulation of Differential Entropy, and Differential Entropy (DE), alongside various classifiers such as Multi-Layer Perceptron (MLP), SpinalNet, and SpinalLSTM.

The SpinalLSTM model uniquely leverages both LSTM and SpinalNet architectures, incorporating the gradual input processing mechanism of SpinalNet alongside the memory retention capabilities of LSTM. This hybrid allows the model to capture both temporal dependencies and spatial information, contributing to its robust performance.

While our SpinalLSTM model did not outperform some state-of-the-art approaches, it demonstrated competitive performance, closely approaching their accuracy levels. This indicates that the integration of SpinalNet and LSTM architectures offers a promising avenue for improving classification accuracy in various applications

Table 5 presents a performance comparison with state-of-the-art methods on the DEAP database using a subject-independent approach.

REFERENCES

[1] Y. Liu, O. Sourina, M.K. Nguyen, Real-time EEG-based emotion recognition and its applications, in: Transactions on Computational Science XII, Springer, 2011, pp. 256–277.

[2] H. da Cunha Santiago, T.I. Ren, G.D. Cavalcanti, Facial expression recognition based on motion estimation, 2016 International Joint Conference on Neural Networks (IJCNN) (2016) 1617–1624.

[3] Z.G. Doborjeh, M.G. Doborjeh, N. Kasabov, Attentional bias pattern recognition in spiking neural networks from spatio-temporal EEG data, Cognit. Comput. 10 (2018) 35–48.

[4] C.J. de Naurois, C. Bourdin, A. Stratulat, E. Diaz, J.-L. Vercher, Detection and prediction of driver drowsiness using artificial neural network models, Accid. Anal. Prev. 126 (2019) 95–104.

[5] J. Chen, D. Jiang, Y. Zhang, A hierarchical bidirectional GRU model with attention for EEG-based emotion classification, IEEE Access 7 (2019) 118530–118540.

[6] C. Qing, R. Qiao, X. Xu, Y. Cheng, Interpretable emotion recognition using EEG signals, IEEE Access 7 (2019) 94160–94170.

[7] P. Pandey, K. Seeja, Subject independent emotion recognition from EEG using VMD and deep learning, J. King Saud Univ. Comput. Inf. Sci. (2019).

[8] Arjun and Aniket Singh Rajpoot, "Subject independent emotion recognition using EEG signals employing attention driven neural networks," biomedical signal processing and control (2022)

[9] Vaishali M. Joshi, Rajesh B. Ghongade, Aditi M. Joshi, Rushikesh V. Kulkarni. “EEG based emotion detection using fourth order spectral moment and deep learning”. 2021

[10] H. Chao, L. Dong, Y. Liu, B. Lu, Improved deep feature learning by synchronization measurements for multi-channel EEG emotion recognition, Complexity 2020 (2020).

[11] <https://www.eecs.qmul.ac.uk/mmv/datasets/deap/>

[12] Kabir, HM Dipu, et al. "Spinalnet: Deep neural network with gradual input." IEEE Transactions on Artificial Intelligence (2022).

[13] Santoso, B. Eko, and G. Putra Kusuma. "Facial emotion recognition on FER2013 using VGGSPINALNET." Journal of Theoretical and Applied Information Technology 100.7 (2022): 2008-2102.