

Emotion Recognition Based on EEG using DEAP Dataset

Rama Chaudhary¹, Ram Avtar Jaswal², Sunil Dhingra³

M.Tech. Scholar¹, Department of Electrical Engineering¹, University Institute of Engineering and Technology, Kurukshetra¹
 Assistant Professor², Department of Electrical Engineering², University Institute of Engineering and Technology, Kurukshetra²
 Professor³, Institute of Instrumentation Engineering, Kurukshetra University, Kurukshetra³
 ramachaudhary772@gmail.com¹, ramavtar.jaswal@gmail.com², sdhingra@kuk.ac.in³

Abstract: Recognizing emotions at better accuracy is very challenging task. Therefore, in recent time, the human-machine interaction technology has gained so much success for recognizing the emotional states depending on physiological signals. The human emotional states can be detected by using facial expressions, but sometimes the accurate results are not achieved. Therefore in proposed work, the emotions are recognized using Electroencephalogram (EEG) which work on the basis of brain signal. Here, the human emotional states data is collected using DEAP Dataset and Artificial Neural Network (ANN) is used as classifier. Five time domain features namely correlation, average, variance, kurtosis and skewness are calculated for three frequency bands theta, alpha and beta. The data for two emotional dimensions valence and arousal is taken from DEAP Dataset. The proposed work gives better recognition results for valence and arousal dimensions which are 85.60 % and 87.36 % respectively. So we get the success in achieving significant accuracy.

Keywords: Emotion Recognition, Electroencephalogram (EEG), Brain Signal, DEAP Dataset, Features Extraction, Time domain features, frequency bands, Artificial Neural Network (ANN)

1. INTRODUCTION

Emotions play an important role in the life of human beings. Our body works on the basis of emotions. Without emotions, the person can be considered as dead. Emotions are states of feeling that cause physical and psychological changes. There are lots of emotions such as fear, hate, happy, rage, proud, frustration, affection, confused, confidence, excited, panic, depression, sorrow, joy, tense, surprised, bore, etc. Emotions can be conveyed by facial expressions, by sign language, by talking which help people to interact with one another. There are 2 opinions about emotions: one approach considers emotions as general states of individuals and the other one knows emotions as physiological interactions [1]. Human emotions can be categorized depending on two common principles as discussed: the discrete basic emotion classification and the dimension approaches. According to the discrete basic emotion approach, emotions can be classified as, negative and positive emotions.



Figure 1: Types of Emotions

Negative emotions give us bad feelings. These are those emotions that make us feel frustrated, sad, embarrassed and helpless. Some negative emotions are anger, guilt, fear, embarrassment, shame, depression, tired, panic, frustration, nervous, regret, irritation, etc. Positive emotions give us good feelings. These are those emotions that make us happy, excited, loving and respecting others, and also make us proud of ourself. Some positive emotions are joy, love, faith, excited, hopeful, happy, delighted, pride, etc.

According to dimension approach, the emotions are categorized into three dimensions (valence, arousal & dominance). Valence describes the extent of an emotion to be negative or positive. Therefore, it is further classified into two categories: Low valence or negative valence emotion (fear, tense, anger, sadness, bore, etc.) and High valence or positive valence emotion (happy, calm, delighted, excited, etc.).

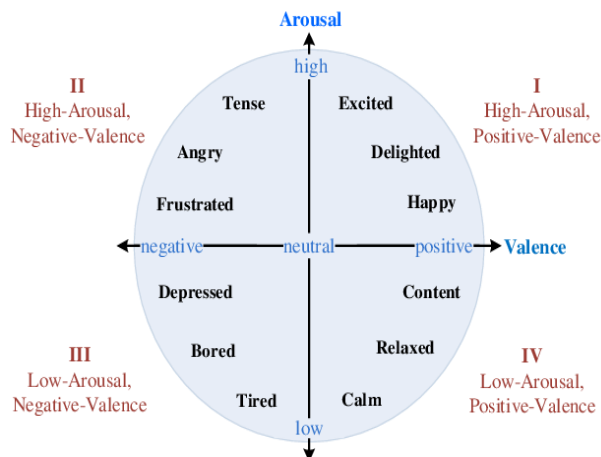


Figure 2: Two dimensional valence-arousal space [2]

Arousal refers to the strength of a related emotional state and indicates the level of enthusiasm or apathy. It is further classified into two categories: Negative or low arousal emotion (relaxed, tired, bored, calm, depressed, etc.) and Positive or high arousal emotion (angry, excited, tense, frustrated, happy, etc.)

Submissive (no control) to dominance (empowered) are the different levels of dominance. The Dominance-Submissiveness Scale assesses your control and dominance, as well as your control and submissiveness. While both fear and anger are negative emotions, anger is the dominant emotion, while fear is the submissive emotion.

Brain is the command centre of human nervous system [3]. It controls all the body functions and activities. It consists of mainly three parts: cerebellum, cerebrum and brain stem. Cerebrum is the main and greatest component of human brain and it is divided into left and right hemispheres. Each hemisphere consists of four lobes: frontal, temporal, parietal and occipital lobe. The frontal lobe controls the body movement, personality, concentration, etc. The temporal lobe controls emotions like thirst, hunger, recognizing faces [3] and also controls hearing. The parietal lobe controls the sensation feeling of touch, taste, etc. With the help of occipital lobe, human beings are able to recognize colors, shape, size, depth, distance, etc. The sectional part of human brain is shown in figure 3.

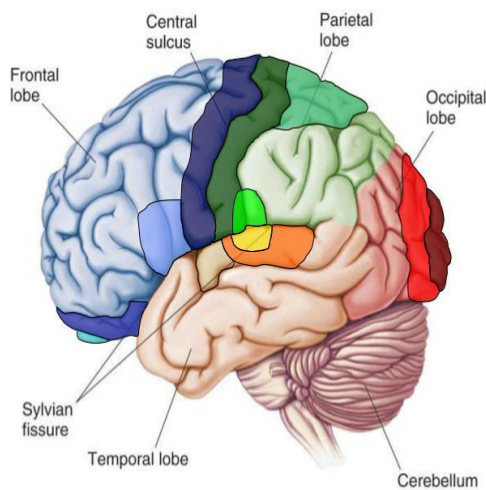


Figure 3: Human Brain

This paper is organized as follows: firstly, the information about Electroencephalography has been provided. In next section, the methodology of proposed work is described which includes information about DEAP Dataset and Artificial Neural Network (ANN). Then, result of proposed work has been discussed and also our work is compared with some existing works on DEAP Dataset. At last, the proposed work is concluded.

2. ELECTROENCEPHALOGRAPHY

Electroencephalography is a method of using electrophysiological monitoring to capture electrical activity in the brain. The monitoring of the brain's spontaneous electrical activity over time using multiple electrodes mounted on the scalp is known as an electroencephalogram (EEG) [4]. The EEG measures the voltage fluctuations caused by the flow of ionic current in the brain's neurons. An EEG equipment records the electrical potential produced by neurons when they are active. Small flat metal discs called electrodes are attached to the scalp with wires. These electrodes monitor electrical impulses in the brain and send signals to a computer for analysis. Electrical impulses appear as wavy lines with peaks and valleys in an EEG recording, as illustrated in figure 4. EEG signals are generated by the central nervous system (CNS), and they react to emotional changes faster than other peripheral neural signals. One of the most significant instruments for assessing brain activity is the electroencephalogram (EEG). The EEG, on the other hand, does not always represent the electrical activity of a single neuron, but rather the electrical activity of a group of neurons in the brain region where the EEG measurement electrode is placed. As a result, the EEG signal contains a variety of psychophysiological data that is both relevant and valuable. Although the low frequency region of 0.5–30 Hz is most relevant to cognition, the EEG signal has a modest amplitude of 50 to 200 μV and a frequency spectrum of 0.5-100 Hz. Researchers commonly divide it into five frequency sub-bands, each of which correlates to a particular cognitive function: δ (1-4 Hz), θ (4-8 Hz), α (8-13 Hz), β (13-30 Hz), γ (>30 Hz) [4].

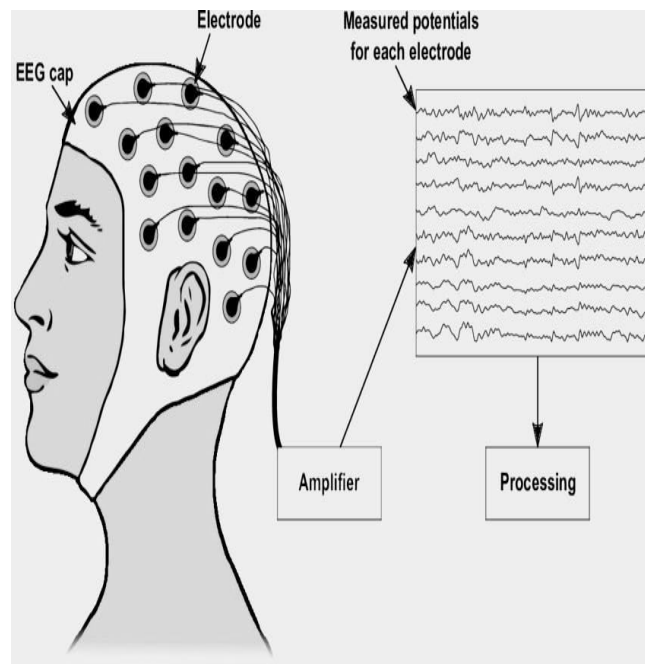


Figure 4: Electroencephalogram

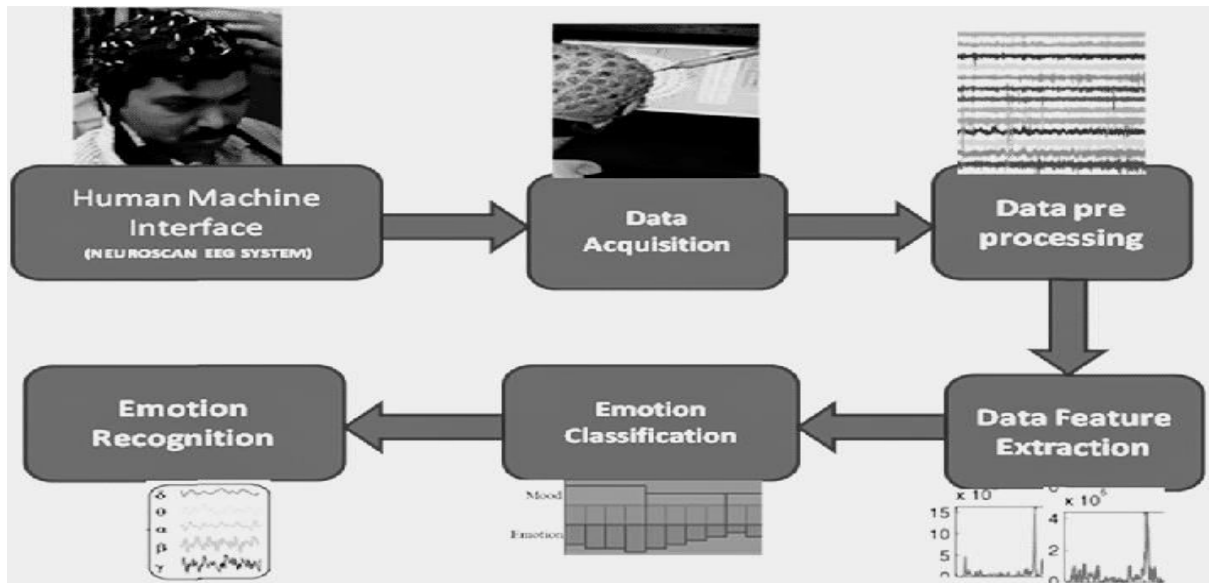


Figure 5: Block diagram for Emotion Recognition using EEG Signal

3. METHODOLOGY OF PROPOSED WORK

To recognize human emotional states on better accuracy, the DEAP Dataset [5] and Artificial Neural Network is used. So firstly, the data is collected from DEAP dataset, then it is pre-processed. After pre-processing, five time domain features are extracted for three frequency bands. Then, artificial neural network is used to classify emotions and finally the desired accuracy is achieved. The block diagram for this whole process is shown in figure 5. The methodology for proposed work is discussed below:

3.1 DEAP Dataset

The DEAP dataset is a publicly available multimodal dataset [5] that is used for studying human affective states. It contains the data for four types of states: valence, arousal, dominance & liking. Since different sampling rates were used in data collection and different types of tests were performed, the DEAP Dataset is an aggregate of different types of data. The large volume of data is obtained by combining this data into a single comprehensive dataset. The researchers use it for testing the emotion states. For using this dataset, EULA (End User License Agreement) is to be printed, signed and then upload it via dataset request form. After following all these steps, Username & Password is to be provided by them. The DEAP dataset [5] was collected from 32 subjects when they were watching 40 sets of 1-min music and video clips. After finishing each video, each participant was asked to give the rating from 1 to 9. Therefore, the total video ratings collected were 1280, which is obtained by multiplying the number of videos (i.e. 40) to the number of volunteers (i.e. 32). For 22 participants, frontal face video was also recorded. Various signals, such as EEG, electromyograms, breathing region, plethysmographs, temperature, and so on, were reported as 40 channel data during each subject's 40 trials. EEG data is stored in 32 channels out of 40, which are selected for experiments. There are 32 files, one file for each participant's recordings. Each trial contains 63s signals, the first 3s are the base line signals. The data were down-sampled from 512 Hz to 128 Hz after a series of filtering operations and a bandpass frequency filter from 4.0–45.0 Hz was applied. The frequency bands are decomposed into δ (1–4 Hz), θ (4–8 Hz), α (8–13 Hz), β (13–30 Hz), γ (> 30 Hz) [4]. After collecting and pre-processing the data, features are to be extracted for different frequency bands. In proposed work, the accuracy for two emotional dimensions, valence and arousal, is calculated. Valence is divided into low (ranging from one to five) and high (ranging from five to nine) valence scale, and similarly, arousal is divided into low (ranging from one to five) and high (ranging from five to nine) arousal scale. Then the mean of valence and arousal ratings is collected and recorded these values to the data file S01–S32.dat. Four arrays are made in MATLAB, named by HVD, LVD, HAD and LAD, to collect the data of high valence, low valence, high arousal and low arousal respectively.

3.2 Feature Extraction

The signal's attributes or qualities are represented by the feature. Feature extraction is the process of extracting those features from raw data, which increases the model's performance. EEG signal is a kind of chaotic time series. Therefore, time domain features are the most direct reflection of EEG signals [6]. Here, we have calculated the five time domain features as discussed below:

- (a) Correlation: It is a statistic that measures the degree to which two variables move in relation to each other. It refers to the statistical relationship between two entities. The correlation coefficient can take any values from -1 to +1. The correlation coefficient (r_{xy}) that indicates the relation between two variables ($x(t)$, $y(t)$) can be found using the following formula:

$$r_{xy} = \frac{\sum(x(t) - \mu(x))(y(t) - \mu(y))}{\sqrt{\sum(x(t) - \mu(x))^2 \sum(y(t) - \mu(y))^2}}$$

- (b) Average: Average refers to arithmetic mean, which is calculated by dividing the sum of the values in the set by their number. It is denoted by μ . In mathematical form, average is expressed as

$$\mu = \frac{1}{n} \sum_{t=1}^n x(t)$$

- (c) Variance: It is a measure of statistical dispersion, indicating the degree of variability in certain situations. It is denoted by σ^2 .

$$\sigma^2 = \frac{1}{n} \sum_{t=1}^n (x(t) - \mu)^2$$

- (d) Kurtosis: This is a calculation that determines the degree of flatness of a distribution, investigating whether it is more tapered or flattened compared to the pattern characterized as normal. Briefly, the higher the kurtosis the greater the presence of values that are distant from average. Kurtosis is defined as:

$$kurtosis = \frac{1}{n} \sum_{t=1}^n \frac{(x(t) - \mu)^4}{\sigma^4}$$

- (e) Skewness: It has the purpose of verifying and calculating the data symmetry. The value of skewness is defined as:

$$skewness = \frac{1}{n} \sum_{t=1}^n \frac{(x(t) - \mu)^3}{\sigma^3}$$

The above described time domain features, correlation, average, variance, kurtosis and skewness are separately extracted for three frequency bands (α , β and θ bands) and four rating scales (high valence, low valence, high arousal and low arousal). For extracting these features, each trial of 60 seconds is divided into 7 segments, having 12 seconds time window moving every 8 seconds with an overlapping of 4 seconds. Finally, from a total of 32 participants, 8,960 (seven segments \times 40 trials \times 32 participants) samples are obtained. All time domain features are calculated for every segment and channel. Therefore the matrix for correlation, average, variance, kurtosis and skewness will be of order 7×32 . Here, 7 represents the number of time segments and 32 represents the number of channels. Then, all these features are extracted for three frequency bands (α , β and θ bands). Therefore the matrix will be of order 21×32 . Here, 21 is the product of 7 segments and 3 frequency bands. Then the range of valence and arousal is decided, ranging from one to nine, depending upon these features. Valence is divided into low (for less than five range) and high (for more than five range) valence scale, and similarly, arousal is

divided into low (for less than five range) and high (for more than five range) arousal scale. In proposed work, 20 are high valence and 20 are low valence videos. Similarly, there are 25 high arousal videos and 15 low arousal videos. After extraction of above discussed five time domain features for three frequency bands, these are extracted for every emotional state. Therefore, for 20 high valence video ratings, the matrix will be of order 420×32 . Same matrix order will be for 20 low valence video ratings. Here, 420 is obtained by the multiplication of 21 (as calculated above) and 20. Similarly, for 25 high arousal video ratings, the matrix will be of order 525×32 and for 15 low arousal video ratings, it will be of order 315×32 . All the values of high valence, low valence, high arousal and low arousal for all time domain features are stored in their separate matrices.

3.3 Artificial Neural Network

An artificial neural network (ANN) is a component of a computational system that simulates how the human brain analyses and processes data. Artificial intelligence (AI) is built on this basis, and it solves problems that would be impossible or difficult to solve by human or statistical standards. An artificial neural network (ANN) is made up of hundreds or thousands of artificial neurons called processing units that are linked by nodes [7]. Input and output units make up these processing units. Input unit contains the given information and then the artificial neural network learn about the information provided in in order to generate output. The network is made up of connections, each of which serves as an input to another neuron by passing the output of one neuron.

The input layer, hidden layer, and output layer are the three layers that make up a neural network, as shown in figure 6. The number of hidden layers can be any number, and the number of layers is raised to improve the accuracy and identification rate. The output of feature extraction can be applied to input layers with N nodes [8]. The number of output nodes can be as many as the number of output classes required.

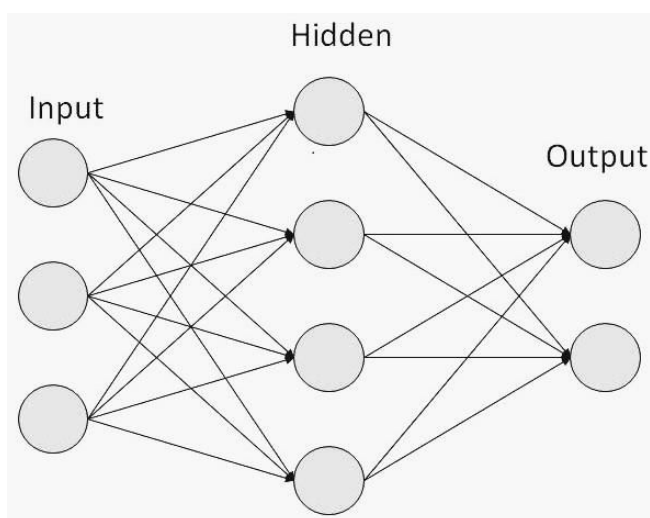


Figure 6: Artificial Neural Network Structure

The emotions are classified using 160 input layers, 10 hidden layers and 2 output layers. Due to 32 channel and 5 time-domain features, the number of input layers is 160. The data for ANN classifier is prepared by making separate target matrix for high valence, low valence, high arousal and low arousal. For high valence, the data stored in target matrix is in the form of one and zero, for low valence it is in the form of zero and one. Similarly, for high arousal the data stored in target matrix is in the form of one & zero and for low arousal it is in the form of zero and one. There are three phases in neural network: training phase, validation phase and testing phase.

Training Phase: Training phase is used to train the data and it also reduces the data size to few selected important features required from entire data [9]. In this model, emotional states data and target data is given as an input to the training phase. Input to the training phase is 70 % collected data of valence and arousal states.

Validation phase: Validation phase is a separate section of dataset that is used during training to check the performance of model on those data which are not used in training. Here, 15 % data is used for validation phase.

Testing phase: Testing phase is used to test the final performance of model i.e. how well will it do in production. Here, the 15 % data is used for testing phase.

After all these procedures, the performance accuracy is calculated. In proposed work, the performance function used is cross entropy. The model's prediction is compared to the label, which is the true probability distribution. As the forecast becomes more accurate, the cross-entropy decreases. If the prediction is perfect, it becomes zero. As a result, the cross-entropy can be used to train a classification model as a loss function.

4. RESULTS AND DISCUSSION

In proposed work, the data for human emotional states is collected from DEAP dataset [5] and Artificial Neural Network is used to get the better results as compared to already existing works. As above mentioned that DEAP dataset contains the data of 40 videos which were watched by 32 volunteers and they rated the data according to high (ranges from 5 to 9) and low (ranges from 1 to 5) scales, So obtained numbers of high valence, low valence, high arousal and low arousal video ratings are shown in Table 1. Out of 40 videos, the number of high valence and low valence video ratings is 20 and 20 respectively. Similarly, the obtained numbers of high arousal and low arousal video ratings is 25 and 15 respectively.

After extracting the five time domain features for all three frequency bands, four emotional scales, and applying on classifier, the resulted accuracy for valence and arousal is 85.60 % and 87.36 % respectively as shown in Table 2. From the results, we can find that the performance of the proposed method has been improved on all two emotional dimensions of valence and arousal.

The results of the proposed method have been compared with some different existing methods that used the DEAP data set, as shown in Table 3. The proposed method represents a significant improvement over the existing methods on all two emotional dimensions.

When we compared the proposed work with some existing works using same dataset but different classifiers, then we have found that the obtained results are far better than the previous works and the resulted accuracies on arousal and valence are 87.36 % and 85.60 % respectively.

Table 1: Number of Video Ratings for High and Low States (out of 40 videos)

Video Ratings	Valence	Arousal
High	20	25
Low	20	15

Table 2: Percentage Accuracy of proposed work

Emotion states	Accuracy (%)
Valence	85.60
Arousal	87.36

Table 3: Comparison of proposed work with existing works

Reference	Dataset	Classifier	Accuracy (%)	
			Valence	Arousal
Zhuang et al.[10]	DEAP	SVM	69.10	71.99

Vaishali and Rajesh [11]	DEAP	BiLSTM	73.5	75
You and Liu[12]	DEAP	SAE+LSTM	81.10	74.38
Chao et al.[13]	DEAP	SVM	70.21	71.85
Huang et al.[14]	DEAP	SVM	64.30	67.30
X. Chen et al. [15]	DEAP	H-ATT-BGRU	67.9	66.5
Pandey and Seeja[16]	DEAP	CNN	61.50	58.50
Zhan et al.[17]	DEAP	CNN	82.95	84.07
Parui et al.[18]	DEAP	XGBoost+LightGBM+ Random forest	77.19	79.06
Proposed Work	DEAP	ANN	85.60	87.36

5. CONCLUSION

Human emotional states based on EEG have been recognized by so many researchers. Different analysis methods, datasets and classifiers are used for obtaining the better accuracy. After analyzing all works, it is decided to use DEAP Dataset for collecting the data of human emotional states and Artificial Neural Network (ANN) for classifying the emotions so that better accuracy can be achieved without increasing so much complexity. The proposed work is based on three main steps: processing, feature extraction, and classification. Here, five time domain features correlation, average, variance, kurtosis and skewness are extracted on three frequency bands theta (4-8 Hz), alpha (8-13 Hz) and beta (13-30) Hz. The work also includes a section for comparing the proposed work with some existing works, which revealed that the proposed method has better accuracy for predicting arousal and valence dimensions.

REFERENCES

- [1] Thagard P. *Mind: Introduction to Cognitive Science*. Cambridge, MA: MIT press; 2005.
- [2] Liang-Chih Yu, Lung-Hao Lee, Shuai Hao, Jin Wang, Yunchao He, Jun Hu, K. Robert Lai and Xuejie Zhang, 2016, "Building Chinese Affective Resources in Valence-Arousal Dimensions," DOI: 10.18653/v1/N16-1066 <https://www.researchgate.net/publication/304124018>
- [3] Ayan Seal, Puthi Prem Nivesh Reddy, Pingali Chaithanya, Arramada Meghana, Kamireddy Jahnavi, Ondrej Krejcar, and Radovan Hudak, 2020, "An EEG Database and Its Initial Benchmark Emotion Classification Performance," *Hindawi Computational and Mathematical Methods in Medicine* Volume, Article ID 8303465, 14 pages <https://doi.org/10.1155/2020/8303465>
- [4] P. Nunez and R. Srinivasan, *Electric Fields of the Brain: The Neurophysics of EEG*, Oxford University Press, 2006.
- [5] Koelstra, S.; Muhl, C.; Soleymani, M.; Lee, J.S.; Yazdani, A.; Ebrahimi, T.; Pun, T.; Nijholt, A.; Patras, I., "DEAP: A Database for Emotion Analysis Using Physiological Signals" *IEEE Trans. Affect. Comput.* 2012, 3, 18–31.
- [6] S. Blanco, H. Garcia, R. Q. Quiroga, L. Romanelli, and O. A. Rosso, "Stationarity of the EEG series," *IEEE Engineering in Medicine and Biology Magazine*, vol. 14, no. 4, pp. 395–399, 1995.
- [7] Yafei Sun, "Neural Networks for Emotion Classification", August 2013.
- [8] Devi Arumugam and Purushothaman, "Emotion Classification Using Facial Expression," *International Journal of Advanced Computer Science and Applications*, Vol. 2, No.7, 2011.
- [9] Miss Pooja A Choudhari, Dr. M .R. Dixit, "Human Emotion Recognition using Neural Network Technique" *International Journal of Engineering Research & Technology (IJERT)* ISSN: 2278-0181 Vol. 5 Issue 04, April-2016, <http://www.ijert.org>.
- [10] N.Zhuang, Y.Zeng, L.Tong, C.Zhang, H.Zhang, and B.Yan, "Emotion recognition from EEG signals using multidimensional information in EMD domain," *Biomed Research International*, vol. 2017, Article ID 8317357, 9 pages, 2017.
- [11] Vaishali M. Joshi, Rajesh B. Ghongade, 2020 "IDEA: Intellect Database for Emotion Recognition" *Journal of King Saud University- Computer and Information Sciences*, <https://doi.org/10.1016/j.jksuci.2020.10.007>.
- [12] S. D. You and C. Liu, "Classification of user preference for music videos based on EEG recordings," in *Proceedings of the IEEE 2nd Global Conference on Life Sciences and Technologies (LifeTech)*, Kyoto, Japan, March 2020.
- [13] Chao et al., "Improved Deep Feature Learning by Synchronization Measurements for Multi-Channel EEG Emotion Recognition" *Hindawi Complexity* Volume 2020, Article ID 6816502, 15 pages <https://doi.org/10.1155/2020/6816502>

- [14] D. Huang, S. Zhang, and Y. Zhang, "EEG-based emotion recognition using empirical wavelet transform," in Proceedings of the 2017 4th International Conference on Systems and Informatics (ICSAI), pp. 1444–1449, Hangzhou, China, November 2017.
- [15] X. Chen et al., "A Hierarchical Bidirectional GRU Model With Attention for EEG-Based Emotion Classification" IEEE Access 7, 118530-118540, 2019
- [16] Pandey P., Seeja K., "Subject independent emotion recognition system for people with facial deformity: an EEG based approach" J. Ambient Intell. Hum. Comput., 1-10, 2020
- [17] Y. Zhan et al., "A computation resource friendly convolutional neural network engine for EEG-based emotion recognition," in Proceedings of the IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA), Tianjin, China, June 2019.
- [18] Parui et al., "Emotion recognition from EEG signal using XGBoost algorithm," in Proceedings of the IEEE 16th India Council International Conference (INDICON), Rajkot, India, December 2019.