Psychological Research

Original Article

EEG Signal Based Multi Class Emotion Recognition using Hybrid 1D-CNN and GRU

Harshil Gupta [1], Ojesh Sharma [2], Dhruv Bhardwaj [3], Jyoti Yadav[4] , Inderjeet Singh[5]

Abstract: In this study, a hybrid architecture combining a Convolutional Neural Network (1D-CNN) and Gated Recurrent Unit (GRU) is proposed for multi-class emotion recognition using EEG signals.Emotion recognition using EEG signals is a challenging task due to the ever-changing nature of EEG signals and the high dimensionality of the feature space. The proposed approach aims to address these challenges by utilizing a hybrid architecture that combines the strengths of both 1D-CNN and GRU. The 1D-CNN is used to retrieve relevant spatial features from the EEG signals, while the GRU is employed to capture the temporal dependencies in the signals. The models were used to classify multi-class emotions: four and sixteen emotions based on the valence-arousal and valence-arousal-liking-dominance planes, respectively, using the benchmark DEAP dataset. The experiment results showed that the proposed models achieved high accuracy in classifying emotions for both four and sixteen emotions as compared to state of art methods. The results of this research have significant implications for the development of affective computing systems in various fields, including healthcare, human-computer interaction, and education. In conclusion, this study demonstrates the potential of deep learning models in affective computing and provides a foundation for future research in this field. The use of reliable physiological signals and the combination of different architectures have shown to be effective in accurately classifying emotions.

Keywords: Electroencephalogram (EEG), Convolutional Neural Network (CNN), Emotion Recognition System, Gated Recurrent Unit (GRU), Emotion Classification

1. Introduction

The study of emotions has led to a variety of viewpoints among neuroscientists and academics, despite the fact that everyday experiences such as happiness, sorrow, anger, fear, contempt, and surprise are all recognized as universal emotions. Despite the fact that fundamental emotions are understood, there is an increasing demand for knowledge about the subtleties and complexity of emotional experiences; When it comes to emotions, there are two major schools of thought. One perspective views emotions as broad states of individuals, while the other regards them as physiological interactions.[1]. The first perspective, known as the categorical approach, sees emotions as universal and distinct categories of experiences. This view is based on the belief that emotions have a universal biological foundation and that the basic categories are the same across cultures and people. The second perspective, known as the dimensional approach, views emotions as complex and continuous experiences that cannot be reduced to a limited number of categories. Instead, emotions are seen as arising from the interaction between physiological, cognitive, and situational factors. This perspective puts more emphasis on the role of cognition and context in shaping emotional experiences. The development of Emotion Recognition Systems (ERS) has made it possible to identify human emotions in various settings, including computer science, retail, healthcare, and multimedia. Due to the advancement of AI and machine learning methods, ERS will soon be utilized extensively. To identify human emotions, several researchers have created ERS over the past ten years [2]. Emotion Recognition Systems

Netaji Subhas University of Technology (NSUT), Delhi - 110078;
harshil.ic19@nsut.ac.in[1] , ojesh.ic19@nsut.ac.in[2] , dhruv.bhardwaj.ic19@nsut.ac.in[3], , bmjyoti@gmail.com[4], inderjeet@dseu.ac.in[5]

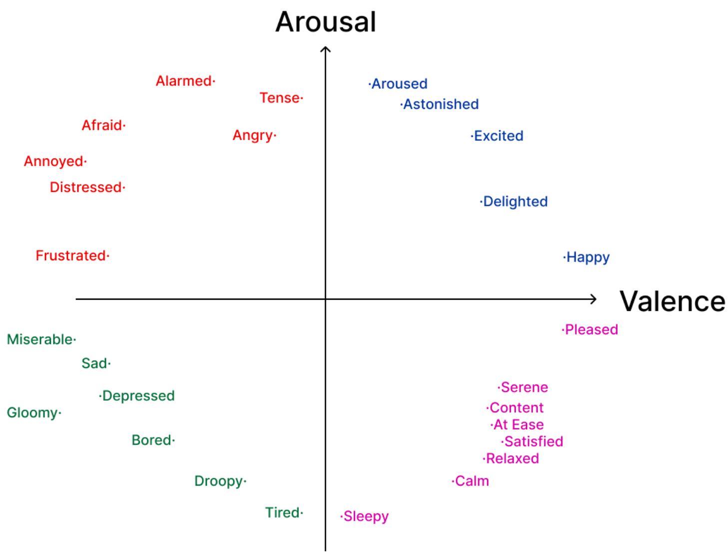
(ERS) can be created through various means, including analyzing facial expressions [3][4], speech patterns [5][6], and physiological responses [7][8]. Examples of physiological signals that have been used in empirical research to develop ERS include electroencephalogram (EEG) [9][10], temperature (TEMP) [11][12], electrocardiogram (ECG) [13][14], galvanic skin reaction (GSR) [15][16], and photoplethysmography (PPG) [17][18].

A physiological signal focused here is the EEG signal which has become a popular tool for emotion analysis due to its simplicity and accessibility. These signals provide a unique insight into the functioning of the brain and offer good time and spatial resolution, making them ideal for use in BCI (Brain Computer Interface) applications such as emotion recognition. EEG-based systems for recognizing emotions have become more precise and dependable due to developments in signal processing techniques and machine learning. Although there are obstacles to overcome, such as variations in EEG patterns between individuals and the influence of external factors, EEG signals are still useful for identifying emotions and have significant potential for advancement and use in various fields. So, a decision was made to incorporate EEG signals for emotion analysis because of these advantages and the promise they hold for advancing our understanding of emotions. The emotion elicitation problem is typically approached by measuring valence and arousal in the emotion circumplex model. "Valence" relates to whether an emotion is positive or negative, while "arousal" describes how intense the emotion is. In ERS, the predominant method is to categorize emotions as having high or low valence and high or low arousal. However, our research goes beyond this by classifying emotions into 16 categories, providing a more comprehensive understanding of human emotions.

This text discusses the problem of emotion elicitation and the common approach of measuring valence and arousal using the emotion circumplex model. While most affective computing research focuses on binary classification tasks for valence and arousal, this research classifies emotions into four categories: Excitement, relaxed pleasure, distress and Fear" and "Sad according to the extent of valence and arousal.

Additionally, this research classifies emotions into 16 emotions according to the level of Valence, Arousal, Liking, and Dominance to gain a more comprehensive comprehension of emotional experiences. This approach provides more nuanced information about emotional experiences and may be useful for future research in affective computing.

Figure 1: Circumplex model by James Russell



Russell is the creator of the circumplex model, which distributes emotions in a 2-D plane. In this model, the horizontal axis denotes valence, while the vertical axis represents arousal. The circumplex model enables the depiction of emotional states at varying degrees of valence and arousal [19]. While we have previously classified emotions using a two-dimensional framework of arousal and valence, a more thorough method takes into account multiple dimensions, such as dominance, liking, arousal, and valence. This approach acknowledges that emotions are complex and nuanced, and cannot be easily categorized into a restricted set of distinct categories. The addition of dominance as a dimension recognizes the role of control in emotional experience while liking is the degree of pleasantness or unpleasantness related to the emotional state. Thus, the multidimensional approach provides a more comprehensive framework for understanding the complexity of human emotions.

1.1 Related Work

Figures 2 and 3 illustrate the analysis of literature regarding the methods employed and outcomes achieved in the domain of EEG-based emotion recognition on the DEAP dataset. Advanced machine learning models incorporating methods such as LSTM and CNN have been shown to achieve high accuracy levels in valence and arousal classification, generally above 90% [20]. It should be mentioned that some methods exhibit superior performance on datasets other than DEAP, even though they still perform well on DEAP. An instance of this is demonstrated in the work of R.Zu et.al. [21], who attained a classification accuracy of 94.85% for valence and a classification accuracy of 93.40% for arousal on the DEAP dataset, while obtaining an average accuracy of 99.27% for classification and 99.20% for classification on valence and arousal respectively on the DESC dataset.

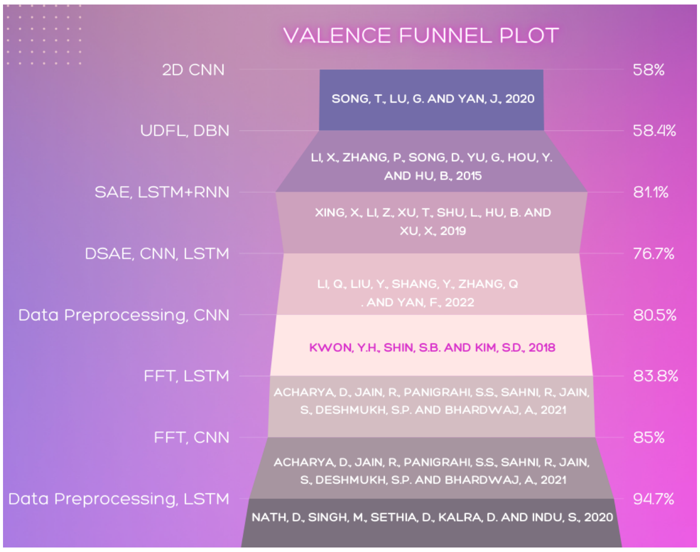


Figure 2: Valence Funnel Plot of the Literature Survey

[The abbreviations used in the funnel plots are UDFL - Unsupervised Deep Feature Learning, DBN - Deep Belief Networks, SAE- Stack AutoEncoder, DSAE - Deep Sparse AutoEncoder, FFT - Fast Fourier Transform]

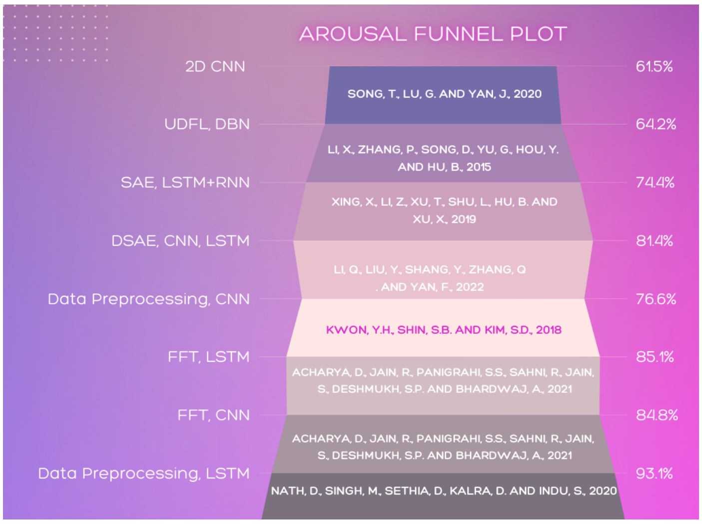


Figure 3: Arousal Funnel Plot of the Literature Survey

Dabas, H. et.al. [35] utilized deep learning techniques and attained a 93.5% accuracy on the DEAP dataset. IEEE Transactions on Affective Computing: In the research article by Liu, J., Wu, G. et. al. [36] deep neural networks were employed, resulting in a 92.1% accuracy on the DEAP dataset. Bazgir, O., Mohammadi et. al. [37] employed support vector machines (SVM) and EEG-based features, achieving a 90.8% accuracy on the DEAP dataset. The study by Donmez, H. et. al. [38] and Salama, E.S et. al. [39] employed an ensemble of deep convolutional neural networks (CNN), achieving a 91.5% and 91.3% accuracy on the DEAP dataset respectively.

In their 2020 study, Liu and colleagues[13] introduced an innovative method for classifying emotions based on EEG signals. They employed a combination of a convolutional neural network (CNN), sparse autoencoder (SAE), and deep neural network (DNN). The CNN was utilized to extract features, which were then encoded and decoded using SAE to reduce redundancy. The resulting data were fed into the DNN as input features. Remarkably, their approach demonstrated impressive recognition accuracies of 89.49% for valence and 92.86% for arousal when tested on the DEAP dataset.

Chowdary, M.K., and Anitha et. al.[40] examined three different architectures: recurrent neural network (RNN), long short-term memory network (LSTM), and gated recurrent unit (GRU). The experiment utilized the EEG Brain Wave Dataset:

Feeling Emotions and yielded impressive results. The RNN achieved an average accuracy of 95% for emotion detection, while LSTM achieved 97% and GRU achieved 96%. Another study by Ozdemir, M.A., and Degirmenci et. al. [41] analyzed three-channel topographical images in sequence. For the classification of negative and positive Valence, the network achieved a test accuracy of 90.62%. In distinguishing between high and low Arousal, the accuracy reached 86.13%. Additionally, for the classification of high and low Dominance, the network achieved an accuracy of 88.48%. Lastly, in the classification of like-unlike, the reported accuracy was 86.23%.

Studies show that the classification procedure in this field uses primitive machine-learning techniques, which may result in lower accuracy. In addition, many categorization systems only consider a select group of basic emotions, such as happiness, rage, anxiety, and sadness. This can overlook the complexity and diversity of emotional experiences. Thus, this research paper presents a hybrid model that merges a convolutional neural network that utilizes one-dimensional convolution with gated recurrent unit techniques to classify and recognize emotions based on EEG data. This novel approach produces outstanding accuracy rates, surpassing those of previous techniques. Specifically, the proposed model attains a remarkable accuracy rate of 99.72% for classifying four distinct emotions on the DEAP dataset. Additionally, a system for the classification of 16 emotions on a valence-arousal-liking-dominance plane, which is the first of its kind and therefore should not be directly compared to other literature. The accuracy of our model for the 16 emotions classification was found to be 99.34%. Our proposed approach has been demonstrated to improve the precision of identifying and categorizing emotions using EEG signals according to experimental findings, outperforming existing methods on the DEAP dataset.

The suggested technique has potential implications for enhancing the development of precise and dependable systems for recognizing and categorizing emotions in different fields, including affective computing, diagnosis of mental health, and interaction between humans and computers.

Section 2 provides details of the methodology. In Section 3, the obtained results are examined and compared to the current state-of-the-art methods along with the limitations. Section 4 covers the conclusions drawn from the research and outlines the possibilities for future work.

2. Methodology

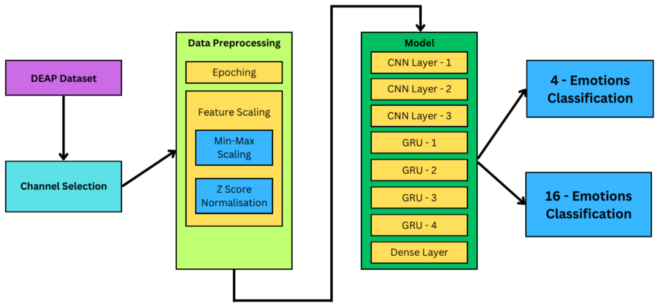


Figure 4:- Block diagram of multiclass emotion recognition system

This study is composed of two fundamental parts: first, the data will be pre-processed to make it ready for analysis, and then advanced deep learning methods will be employed to classify it. The detailed methodology is discussed in Figure 4 which is discussed in more detail in subsequent subsections.

2.1 Datasets and Label Classification

The DEAP dataset is composed of EEG recordings of 40 one-minute-long music videos that were watched by 32 individuals. The subjects evaluated each video based on various factors, including arousal, valence, like/dislike, dominance, and familiarity. The dataset includes information from 32 subjects, with EEG recordings taken from 32 channels and eight peripheral channels that captured other signals, giving a total of 40 channels per subject. For each subject, the data comprises two arrays - data and labels.

The dataset's signal was subjected to down-sampling to 128 Hz and band-pass filtering with a frequency range of 4-45 Hz. Furthermore, the elimination of EOG artifacts through electrooculography and standardization of data to a shared reference was also carried out. Ratings were collected on a scale of 1 to 9 for each video/trial, with higher ratings indicating stronger emotions and lower ratings indicating weaker emotions. To classify the data, four categories were created based on valence and arousal thresholds: High-Valence High-Arousal (HVHA) for Excitement, High-Valence Low-Arousal (HVLA) for relaxed pleasure, Low-Valence High-Arousal (LVHA) for distress, and Low-Valence Low-Arousal (LVLA) for Fear and Sadness, using a classification threshold of 5. Table 1 shows details about data labeling. Figure 5 shows emotion classification and description for four emotions based on Valence and Arousal. Table 2 shows the details corresponding to 16 emotions.

Table 1: Classification of Labels

Label	LVLA	LVHA	HVLA	HVHA
Valence	<=5	<=5	>5	>5
Arousal	<=5	>5	<=5	>5

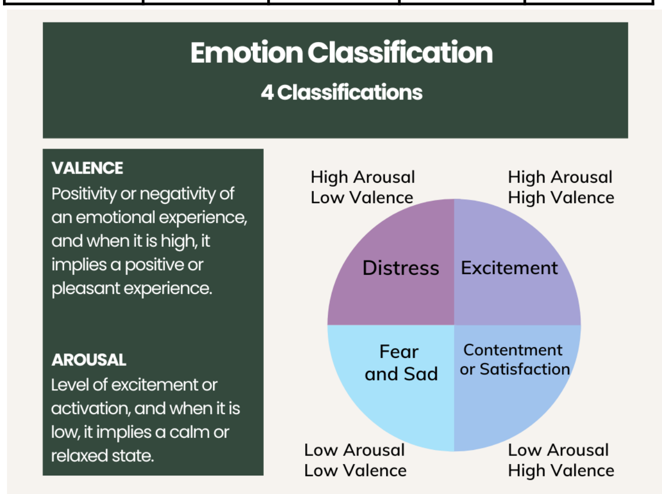


Figure 5:- Emotion Classification and Description for 4 emotions based on Valence and Arousal

Table 2:- Description of 16 Emotions based on Valence, Arousal, Dominance and Liking parameters

Valence	Arousal	Dominance	Liking	Emotion
HV	HA	HD	HL	Excitement
HV	HA	HD	LL	Frustration
HV	HA	LD	HL	Elation
HV	HA	LD	LL	Anxiousness
HV	LA	HD	HL	Contentment
HV	LA	HD	LL	Disappointment
HV	LA	LD	HL	Serenity
HV	LA	LD	LL	Melancholy
LV	HA	HD	HL	Anger
LV	HA	HD	LL	Jealousy
LV	HA	LD	HL	Fear
LV	HA	LD	LL	Despair
LV	LA	HD	HL	Relaxation
LV	LA	HD	LL	Boredom
LV	LA	LD	HL	Comfort
LV	LA	LD	LL	Sadness

To minimize computational expenses in this study, a subset of channels were utilized. Of all the channels, only 14 channels were shortlisted and further processed to accomplish the emotion classification task. Table 3 was used to identify the brain regions that exhibit emotional states, and the selection of channels was made based on their significance. Following the categorization of labels into four distinct categories - HVHA, HVLA, LVHA, and LVLA - the channel selection process was performed. The outcome of this undertaking led to an enhancement in the training accuracy and a decrease in the validation loss. Various channel readings provide valuable insights into different types of emotions [22]. Overall, the selection of these channels is based on their proximity to brain regions consistently implicated in the processing and regulation of specific emotions. By measuring EEG activity at these channels, it may be possible to capture neural activity associated with the experience and regulation of different emotions.

Table 3:- Channel Selection

Channel	Index	Channel	Index
AF3	01	AF4	17
F3	02	F4	19
F7	03	F8	20
FC5	04	FC6	21
T7	07	T8	25
P7	11	P8	29
O1	13	O2	31

2.2 Data Preprocessing

The process of epoching involves extracting short segments of the continuous EEG signal, which are then utilized as input to a model. In this particular study, the 2-second segments of the signal were extracted with a 2-second time step, resulting in 31 segments per channel. Each segment consists of 256 data points, representing a 2-second duration of the signal. This approach enables the input of brief, distinct signal segments into the model for subsequent analysis.To improve the model's effectiveness and decrease overfitting, the signals in the dataset will undergo two methods of standardization: z-score normalization and min-max scaling.

2.3 Classification

In this study, we utilized the pre-processed DEAP dataset, which contains information from 32 subjects. The dataset was subjected to additional pre-processing and categorized appropriately. Following the dataset into three sets: training, validation, and testing, using a 60:20:20 ratio.

2.3.1 Convolution Neural Network (CNN)

This research employs 1-D signals. One-dimensional convolutional neural networks (1D-CNNs) are robust techniques for examining time-series data, such as EEG signals. These models can identify and extract various features of EEG data that are essential for emotion classification tasks.

The main components of a CNN are

1. Convolutional Layer: It is the backbone of CNN, where feature learning and computation take place. Within this layer, the input image undergoes convolution with kernels or filters, which extract important characteristics from the image. Mathematically, the convolution operation is given by

   $$Output(i, j) = \Sigma(\Sigma(Input(i + m, j + n) * Kernel(m, n))) \text{ - } (1)$$

   in which summations are carried out over the dimensions of the kernel, m, and n are the spatial indices of the kernel, and i and j are the spatial indices of the output.

2. Activation Function: After the convolution operation, an activation function is applied to each element for the output. This introduces non-linearity into the model, which enables it to learn more complex patterns. Rectified Linear Unit or ReLU is the most frequently used activation function.

   $$ReLu(x) = max(0, x) - (2)$$

3. Pooling Layer: This layer lowers the computational difficulty and aids in

preventing overfitting by reducing the spatial dimensions of the data. There are other kinds of pooling, but the most popular one computes the highest value in a nearby neighborhood. It is mathematically defined as

$$Max - Pooling(i,\ j) =$$

$$max(Input(i * stride\ + m,\ j * stride\ +\ n))$$

- (3)

4. Fully Connected Layer: It makes up the end part of a CNN and is utilized to conduct high-level reasoning and provide the final output (such as class probabilities). This layer basically functions as a standard feed-forward neural network and accepts the preceding layers' flattened output as input. A completely linked layer's mathematical operation can be written as the following equation. The weight matrix is represented by W, the bias vector by b, and Activation refers to the activation function chosen, such as ReLU or softmax.

$$Output = Activation(W\ *\ Input + b) \text{ - (4)}$$

A CNN architecture is created by stacking these parts together in different ways. Backpropagation and gradient descent techniques are utilized during the training phase to adapt the weights of the filters and fully connected layers, in order to minimize the loss function that gauges the difference between the projected and real output.

2.3.2 Gated Recurrent Unit (GRU)

GRUs, or gated recurrent units, are a type of recurrent neural networks (RNN) that deal with the problem of vanishing gradients common to many RNNs. GRUs are appropriate for processing sequential data, such as text or time-series data, because they can capture and understand long-term dependencies within the data. A GRU has three gates:

1. Update gate (z): It governs the proportion of the previous hidden state that is preserved and how much of the new memory content should be added. It is calculated using a sigmoid activation function:

$$z\_t\ =\ \sigma(W\_z * x\_t + U\_z * h\_(t-1) + b\_z) \text{ - (5)}$$

Sigmoid function being represented by σ, Weight matrices by W_z and U_z , input at time step t by x_t, hidden state at the previous time step by h_(t-1), and bias vector being b_z.

2. Reset gate (r): It governs the proportion of the previous hidden state that is taken into account while calculating new memory content. It is also calculated using a sigmoid activation function:

$$r\_t = \sigma(W\_r * x\_t + U\_r * h\_(t-1) + b\_r) \text{ - (6)}$$

weight matrices being represented by W_r and U_r and bias vector by b_r.

3. New memory content (h'): The new memory content is a candidate for the next hidden state, which is calculated on the basis of previous hidden state and current time step, modulated by reset gate. New memory content is calculated using the hyperbolic tangent (tanh) activation function:

$$h't = tanh(W\_h * x\_t + U\_h * (r\_t \odot h(t-1)) + b\_h) \text{ - (7)}$$

Weight matrices are represented by W_h and U_h, ⊙ being the element-wise multiplication (Hadamard product), and bias vector being represented by b_h.

4. Hidden state update (h): The update gate controls how the current time step uses a linear

function to combine the prior hidden state and the new memory content to update the hidden state.

$$h\_t = (1 - z\_t) \odot h\_(t - 1) + z\_t \odot h'\_t \text{ - } (8)$$

The GRU equations are applied at each time step in a sequential manner, and the model is trained using backpropagation through time (BPTT) to update the weight matrices and bias vectors to minimize a loss function.

2.3.3 1D-CNN-GRU Hybrid Model

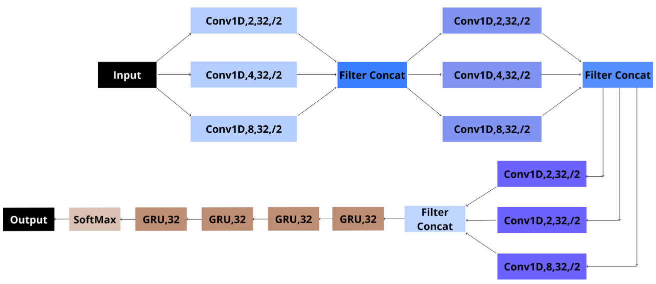


Figure 6:- Model Flowchart of 1D CNN-GRU

The model architecture includes 1D-CNN and GRU structures with an input size of 256 x 1 representing two seconds of epoched signal data. The first layer contains three 1D-CNNs with 32 filters each, kernel sizes of 2, 4, and 8, and the ReLU function. The output of the first layer is concatenated and followed by three more layers identical to the first one. The subsequent four layers consist of GRUs with 32 units and a tanh activation function. The output layer contains one dense filter with a sigmoid activation function. The model is composed of 127,489 trainable parameters. See Table 4 for details of the model configuration.

Table 4:- Model Description of 1DCNN - GRU

Layer (Type)	**Output Shape**	**No. Of Parameters**
Input Layer	[(None, 256, 1)]	0
CNN1_Layer1	[(None, 128, 32)]	96
CNN2_Layer1	[(None, 128, 32)]	160
CNN3_Layer1	[(None, 128, 32)]	288
Concatenate_1	[(None, 128, 96)]	0
CNN1_Layer2	[(None, 64, 32)]	6176
CNN2_Layer2	[(None, 64, 32)]	12320
CNN3_Layer2	[(None, 64, 32)]	24608
Concatenate_2	[(None, 64, 96)]	0
CNN1_Layer3	[(None, 32, 32)]	6176
CNN2_Layer3	[(None, 32, 32)]	1230
CNN3_Layer3	[(None, 32, 32)]	24608
Concatenate_3	[(None, 32, 96)]	0
GRU1	[(None, 32, 32)]	12480
GRU2	[(None, 32, 32)]	6336
Concatenate_4	[(None, 32, 64)]	0
GRU3	[(None, 32, 32)]	9408
Concatenate	[(None, 32, 96)]	0
GRU4	[(None, 32)]	12480
Dense	[(None, 1)]	33

3 Experimental Results and Discussion

The proposed hybrid CNN ID and GRU based emotion recognition system is proposed in present work. The experimentation is conducted on the Google Colab platform using various libraries such as numpy, pandas, sklearn, mne, Tensorflow, matplotlib, among others, in Python 3.7. libraries. EEG signals of 32 subjects were considered in the study. The EEG scalp signals are preprocessed in the following steps: downsampled to 128Hz and bandpass filtering with a frequency range of 4 to 45 Hz frequency is carried out. Further EEG signals are segmented in 2 seconds short segments which results in 31 segments per channel. Further EEG signals are labeled in different dimensions: Valance, Arousal, dominance, and liking. Each dimension is rated on a scale of 9. The values less than 5 are considered as weaker emotion and 5 and more than 5 values indicate high emotion.

The further dataset is classified into training, testing, and validation in the ratio 60:20:20. Table 5 shows the data shape description of 4/16 emotions.

Table 5:- Data Shape (4/16 Emotions Classification)

Set	**Data**	**Percentage Allotted**
Train	26664960 x 5 x1	60%
Validation	8888320 x 5 x 1	20%
Test	8888320 x 5 x 1	20%

Adequate hyperparameter tuning is crucial for achieving good results. Batch size of 256 and Adaptive Moment Estimation optimizer is used for architectures. Since the dataset is multi-labeled therefore, categorical cross-entropy is the loss function for backpropagation. In addition, we have implemented the callback method approach with the patience parameter value of 10 to track the validation loss value during the training process.

3.1 Four Emotions Classification

Differentiating emotions such as pleasure, distress, fear, and sadness requires considering valence and arousal as important factors. The model in this study is trained for 10 epochs using a batch size of 256 and the ADAM optimizer with a learning rate of 0.0001. The loss function employed is categorical cross-entropy.

The obtained performance measures are given in Table 6. These metrics offer an understanding of how well the model performed throughout the training process and can assist in assessing and enhancing the structure and hyperparameters of the model. The 4 emotions classification system gives an accuracy of 99.36%, 99.72%, and 99.72% for training, validation, and testing respectively using hybrid ID-CNN and GRU models. F1 score of 0.668720 for training, 0.668784 for validation, and 0.668644 for testing. Loss of 0.0297 for training, 0.0125 for validation, and 0.0125 for testing. The observed results show the efficacy of the suggested hybrid model. The accuracy and loss of the proposed model are shown and accuracy and loss curves are shown in Figures 7 and 8 respectively.

Table 6:- Accuracy and Loss Table for 4 Emotion Classification

Set	**Accuracy**	**F1 - Scores**	**Losses**
Train	99.36 %	0.668720	0.0297
Validation	99.72 %	0.668784	0.0125
Test	99.72 %	0.668644	0.0125

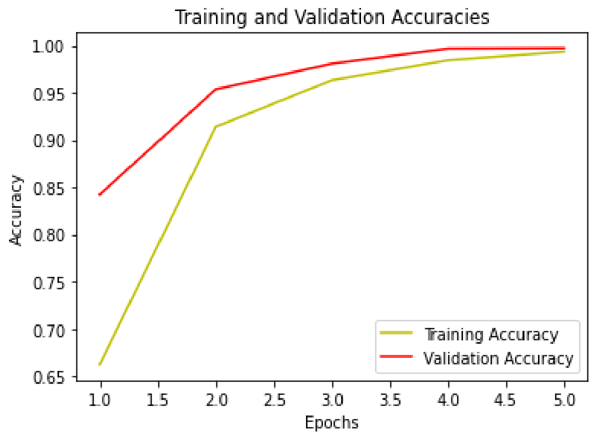


Figure 7:- Accuracies for 4 emotion classifications (Training and Validation)

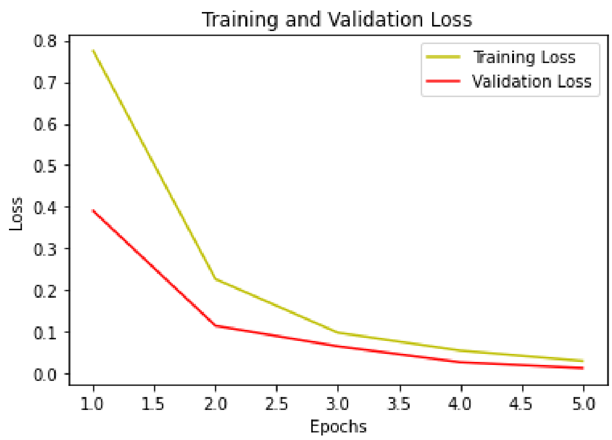


Figure 8:- Losses for 4 emotion classifications (Training and Validation)

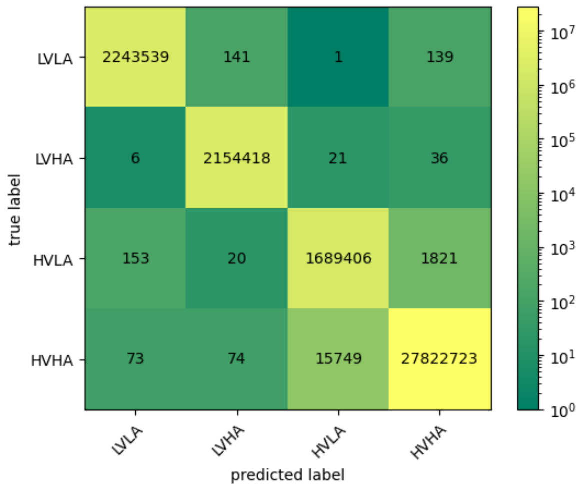


Figure 9:- Confusion matrix of 4 emotions classification system. It represents the class distribution of actual and predicted classes.

Figure 9 shows the confusion matrix of 4 emotion classification systems. It is observed that large numbers exist on the diagonals which shows that most of the predictions are correct.

Figure 10 and 11 shows Receiver operating characteristic(ROC) curve. One vs Rest (OvR) was used in the study to evaluate multiclass models by transforming them into binary classification problems. This allowed for a comprehensive evaluation using binary classification metrics. Table 7 shows the AUC values of Arousal and Valence are 0.940 and 0.974

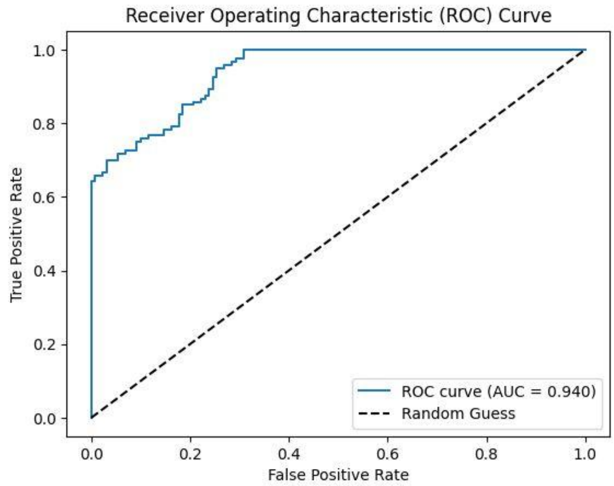


Figure 10:- Arousal Classification ROC Curve

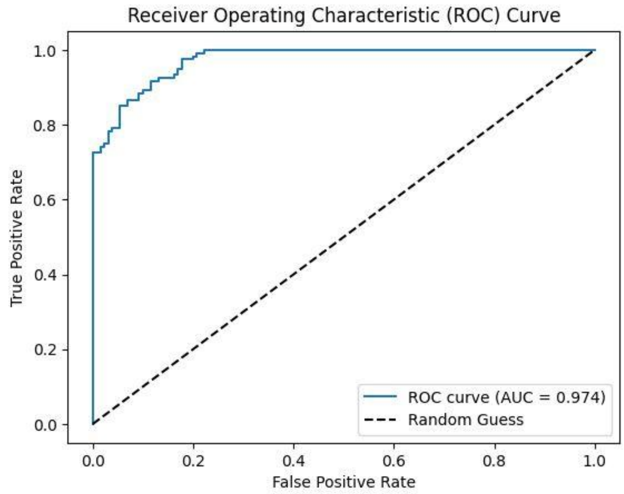


Figure 11:-Valence Classification ROC Curve

Table 7: AUC values for valence and arousal classification

Classification	**AUC Value**
Arousal	0.940
Valence	0.974

3.2 Sixteen Emotions Classification

The proposed method is further used for classification of the highest number of discrete emotions using the 3D VAD model among the state-of-the-art methods. Along with valence, arousal, dominance, liking parameters, liking is also considered to develop 16 emotion classification systems.

The model in this study is trained for 10 epochs using a batch size of 256 and the ADAM optimizer with a learning rate of 0.0001. The loss function employed is categorical cross-entropy.

Table 8 shows performance of 16 emotion classification systems with ID CNN and GRU method. Training, testing and validation accuracy is more than 99%. F1 score is 0.299377, 0.299363 and 0.299302 for training, validation and testing. Losses are 0.0306, 0.0231 and 0.0230 for training, validation and testing. ROC courses for valence, arousal, dominance and linking are shown in figure 12 to 15. Table.9 shows AUC values for valence, arousal, dominance and liking classification. The higher value of the AUC shows the model is performing well to distinguish between the positive and negative classes.

Table 8:- Accuracy and Loss Table for 16 Emotion Classification details for each class

Set	Accuracies	F1 - Scores	Losses
Train	99.09 %	0.299377	0.0306
Validation	99.33 %	0.299363	0.0231
Test	99.34 %	0.299302	0.0230

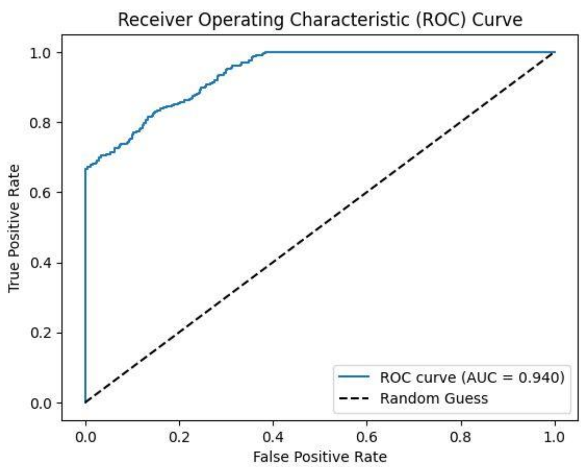


Figure 12:- Arousal Classification ROC Curve

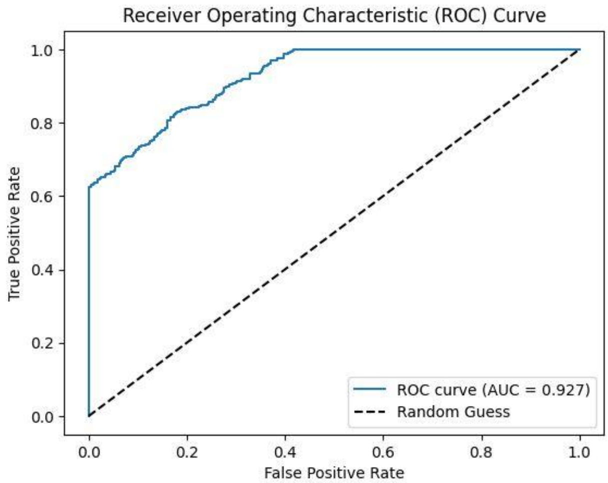


Figure 13:- Valence Classification ROC Curve

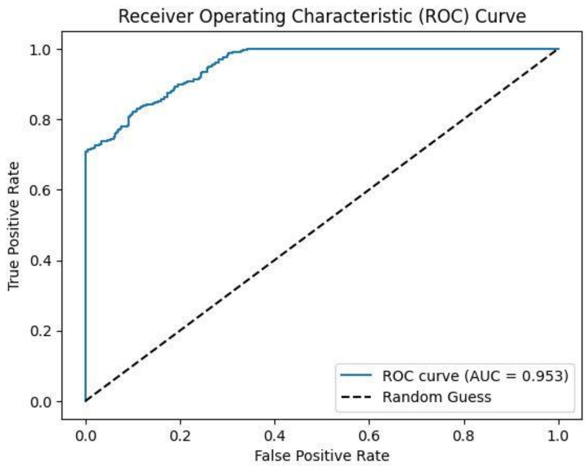


Figure 14:- Liking Classification ROC Curve

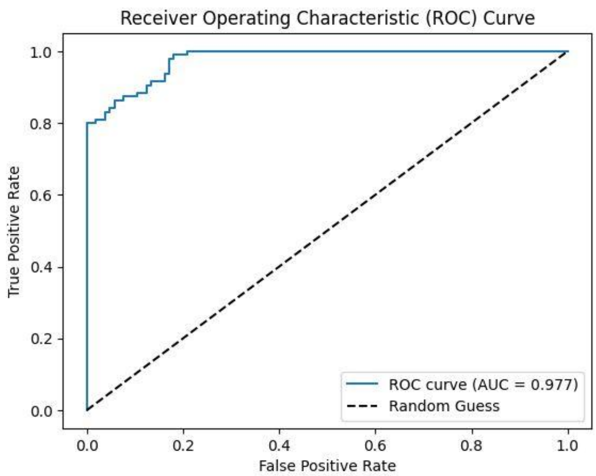


Figure 15:- Dominance Classification ROC Curve

Table 9: AUC values for valence, arousal, dominance and liking classification

Classification	AUC Value
Arousal	0.940
Valence	0.927
Dominance	0.953
Liking	0.977

These metrics offer understanding into how well the model performed throughout the training process and can assist in assessing and enhancing the structure and hyperparameters of the model.

An essential aspect that needs to be addressed in our research is the limited understanding of the underlying neural mechanisms despite the identification of EEG activity patterns linked with various emotions. The lack of comprehensive knowledge about the neural networks and brain regions involved in emotional processing hinders our ability to optimize channel selection and develop more accurate emotion classification models. Therefore, more research is necessary to investigate the underlying neural mechanisms, as this will significantly impact our current work's success.

Table 10: Comparison with state of art of art methods

Methods	TECHNIQUES USED	RESULTS
Multi-class emotion classification using EEG signals [42]	Fast Fourier Transform (FFT), LSTM	Valence - 83.83% Arousal - 85.07%
Multi-class emotion classification using EEG signals[42]	Fast Fourier Transform (FFT), CNN	Valence - 85.01% Arousal - 84.77%
An Efficient Approach to EEG-Based	Data Pre-processing, LSTM	Valence - 94.69% Arousal -
Emotion Recognition using LSTM Network [43]		93.13%
Differential Entropy Feature Signal Extraction Based on Activation Mode and Its Recognition in Convolutional Gated Recurrent Unit Network [44]	Differential Entropy (DE), Power Spectral Density (PSD), 2D CNN, BiGRU	Arousal - 87.89% Valence - 88.69%
EEG emotion recognition using fusion model of graph convolutional neural networks and LSTM [45]	Graph Convolutional Neural Networks (GCNN), LSTM.	Subject Dependent: Arousal - 90.60% Valence - 90.45% Subject Independent: Arousal - 85.27% Valence - 84.81%
EEG-based emotion recognition using 4D convolutional recurrent neural network [46]	Differential Entropy (DE), 4-D CRNN (CNN+LSTM MODEL)	Arousal - 94.58% Valence - 94.22%
EEG-Based Emotion Classification Using a Deep Neural Network and Sparse Autoencoder [36]	Convolutional Neural Network (CNN), Sparse Autoencoder (SAE), and Deep Neural Network (DNN)	Arousal - 92.86% Valence - 89.49%
Our Proposed Method	Data Pre-Processing, 1-D CNN + GRU Hybrid Model	4 emotion Classification 99.72 % 16 emotion Classification 99.33 %

However, building such a model requires careful consideration of data preprocessing, model architecture, and training and validation procedures. In future work, further research can be done to optimize the architecture of the hybrid model and investigate its performance on large-scale datasets. Additionally, it may be worthwhile to explore other types of deep learning models, such as attention-based models, and incorporate other modalities, such as facial expressions or speech, to enhance the accuracy of systems that identify emotions.

Our literature presents a versatile approach that can be applied to diverse datasets, including DREAMER, and SEED, among others. The proposed method, which combines CNN and GRU, demonstrates consistent and accurate classification results for both 4 and 16 emotions classification and therefore should be proposed for other datasets as well.

Enhancement of classification can be achieved by transitioning from a two-level (low and high) system to a three-level (low, middle, and high) system. This modification has the potential to elevate the accuracy of the classification process further.

In addition, we have included a comprehensive analysis wherein the existing literature has been compared in conjunction with our proposed methodology. Our findings have been assimilated in Table 10 to present a consolidated view. Along with the widely utilized DEAP dataset, which was used for data collection by every academic work in Table 10, two academic studies, [36] and [46], have also used data from a data source called SEAD.

As a future direction, it is essential to explore the effects of culture, gender, and age on emotion classification through EEG signals since previous studies have mainly concentrated on college-aged participants from Western cultures. The lack of variety among the individuals who participated emphasizes the necessity for more exploration and investigation in this particular field.

4 Conclusions and Discussions

This research aims to construct an emotion recognition system that employs EEG signals, utilizing a hybrid model that merges 1D-CNN and GRU. This hybrid model exploits the benefits of both convolutional and recurrent neural networks to grasp temporal dependencies and extract features from the EEG signals. Two emotion recognition systems are designed: 4 class emotions and 16 class emotions using EEG signals from the DEAP database. It is observed from the results that the 4-class emotions recognition system achieves a validation accuracy of 99.72% classification accuracy with the proposed method using valence and arousal levels.

Furthermore, the suggested method is implemented in recognizing 16 different emotions, resulting in an accuracy of 99.33%. The hybrid model has been shown to achieve high accuracy in multi-class emotion recognition tasks, which is important for practical applications such as affective computing, human-computer interaction, and healthcare.

Declaration

Ethical Approval

The data used in this study was acquired from the DEAP (Database for Emotion Analysis using Physiological Signals) dataset [10]. The DEAP dataset creators have previously obtained all the necessary ethical approvals for the collection and sharing of the data used in their dataset. Therefore, no additional ethical approval was required for this particular study.

Competing interests

Not Applicable. The authors declare that they have no competing interests.

Authors' contributions

(applicable for submissions with multiple authors)
Harshil Gupta, Ojesh Sharma, and Dhruv Bhardawaj contributed equally to this manuscript. They were responsible for writing the entire manuscript, including the development of the code, creation of corresponding figures and tables, and data analysis.

Jyoti Yadav and Inderjeet Singh served as supervisors and provided guidance throughout the research process. They reviewed and critically evaluated the manuscript, offering valuable insights and suggestions for improvement.

Funding

Not Applicable. This research received no specific funding.

Availability of Data and Materials

The data utilized in this study was acquired from the DEAP (Database for Emotion Analysis using Physiological Signals) dataset [10]. To access the dataset, interested researchers are required to follow the instructions provided by the dataset creators. This involves printing, signing, and scanning an EULA (End User License Agreement) and submitting it via the dataset request form. Upon approval, the dataset creators will provide a username and password for downloading the data.

For more detailed information regarding the dataset and its contents, please refer to the dataset description page available on the DEAP dataset website.

5. References

1. Thagard P. Mind: Introduction to Cognitive Science. Cambridge, MA: MIT press; 2005.
2. Ghali, A.L.I. and Kurdy, M.B., 2018. Emotion recognition using facial expression analysis. Journal of theoretical and applied information technology, 96(18), pp.6117-6129.
3. Kuruvayil, S. and Palaniswamy, S., 2022. Emotion recognition from facial images with simultaneous occlusion, pose, and illumination variations using meta-learning. Journal of King Saud University-Computer and Information Sciences, 34(9), pp.7271-7282.
4. Soleymani, M., Asghari-Esfeden, S., Fu, Y. and Pantic, M., 2015. Analysis of EEG signals and facial expressions for continuous emotion detection. IEEE Transactions on Affective Computing, 7(1), pp.17-28.
5. Mannepalli, K., Sastry, P.N. and Suman, M., 2018. Emotion recognition in speech signals using optimization-based multi-SVNN classifier. Journal of King Saud University-Computer and Information Sciences.
6. Özseven, T., 2019. A novel feature selection method for speech emotion recognition. Applied Acoustics, 146, pp.320-326.
7. Schmidt P., Reiss A., Duerichen R., Van Laerhoven K., Introducing WeSAD, a multimodal dataset for wearable stress and affect detection. In: ICMI 2018 - Proc. 2018 Int. Conf. Multimodal Interact., pp. 400–408, 2018.
8. Kim, B.H. and Jo, S., 2018. Deep physiological affect network for the recognition of human emotions. IEEE Transactions on Affective Computing, 11(2), pp.230-243.
9. Katsigiannis, S. and Ramzan, N., 2017. DREAMER: A database for emotion recognition through EEG and ECG signals from wireless low-cost off-the-shelf devices. IEEE journal of biomedical and health informatics, 22(1), pp.98-107.
10. Koelstra, S., Muhl, C., Soleymani, M., Lee, J.S., Yazdani, A., Ebrahimi, T., Pun, T., Nijholt, A. and Patras, I., 2011. Deap: A database for emotion analysis; using physiological signals. IEEE transactions on affective computing, 3(1), pp.18-31.
11. Ali, M., Al Machot, F., Haj Mosa, A., Jdeed, M., Al Machot, E. and Kyamakya, K., 2018. A globally generalized emotion recognition system involving different physiological signals. Sensors, 18(6), p.1905.
12. Sanches, C.L., Augereau, O. and Kise, K., 2016, December. Manga content analysis using physiological signals. In Proceedings of the 1st international workshop on coMics ANalysis, Processing and Understanding (pp. 1-6).
13. Sarkar, P. and Etemad, A., 2020. Self-supervised ECG representation learning for emotion recognition. IEEE Transactions on Affective Computing, 13(3), pp.1541-1554.
14. Wang, X., Guo, Y., Chen, C., Xia, Y. and Liu, Y., 2019. Analysis of female drivers' ECG characteristics within the context of connected vehicles. Journal of intelligent and connected vehicles, 2(2), pp.55-66.
15. Bachynskyi, A., 2018. Emotional State Recognition Based on Physiological Signals.
16. Raheel, A., Majid, M., Alnowami, M. and Anwar, S.M., 2020. Physiological sensors based emotion recognition while experiencing tactile enhanced multimedia. Sensors, 20(14), p.4037.
17. Goshvarpour, A. and Goshvarpour, A., 2018. Poincaré's section analysis for PPG-based automatic emotion recognition. Chaos, Solitons & Fractals, 114, pp.400-407.
18. Lee, M.S., Lee, Y.K., Pae, D.S., Lim, M.T., Kim, D.W. and Kang, T.K., 2019. Fast emotion recognition based on single pulse PPG signal with convolutional neural network. Applied Sciences, 9(16), p.3355.

19. Seo, Y.S. and Huh, J.H., 2019. Automatic emotion-based music classification for supporting intelligent IoT applications. Electronics, 8(2), p.164.
20. Soroush, M.Z., Maghooli, K., Setarehdan, S.K. and Nasrabadi, A.M., 2017. A review on EEG signals based emotion recognition. International Clinical Neuroscience Journal, 4(4), p.118.
21. Du, R., Zhu, S., Ni, H., Mao, T., Li, J. and Wei, R., 2022. Valence-arousal classification of emotion evoked by Chinese ancient-style music using 1D-CNN-BiLSTM model on EEG signals for college students. Multimedia Tools and Applications, pp.1-18.
22. Al-Qazzaz, N. K., Sabir, M. K., Ali, S., Ahmad, S. A., & Grammer, K. (2019, July). Effective EEG channels for emotion identification over the brain regions using differential evolution algorithm. In 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 4703-4706). IEEE.
23. Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1981). STAI manual for the state-trait anxiety inventory (form Y). Palo Alto, CA: Consulting Psychologists Press.
24. Buhle, J. T., Silvers, J. A., Wager, T. D., Lopez, R., Onyemekwu, C., Kober, H., Weber, J., & Ochsner, K. N. (2014). Cognitive reappraisal of emotion: A meta-analysis of human neuroimaging studies. Cerebral Cortex, 24(11), 2981-2990.
25. Silvers, J. A., McRae, K., Gabrieli, J. D. E., Gross, J. J., Remy, K. A., & Ochsner, K. N. (2017). Age-related differences in emotional reactivity, regulation, and rejection sensitivity in adolescence. Emotion, 17(2), 303-317.
26. McRae, K., Gross, J. J., Weber, J., Robertson, E. R., Sokol-Hessner, P., Ray, R. D., ... & Ochsner, K. N. (2010). The development of emotion regulation: An fMRI study of cognitive reappraisal in children, adolescents and young adults. Social Cognitive and Affective Neuroscience, 5(1), 47-55.
27. Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: attentional control theory. Emotion, 7(2), 336-353.
28. Ochsner, K. N., & Gross, J. J. (2005). The cognitive control of emotion. Trends in Cognitive Sciences, 9(5), 242-249.
29. Kross, E., Davidson, M., Weber, J., & Ochsner, K. (2009). Coping with emotions past: The neural bases of regulating affect associated with negative autobiographical memories. Biological Psychiatry, 65(5), 361-366.
30. Wicker, B., Keysers, C., Plailly, J., Royet, J. P., Gallese, V., & Rizzolatti, G. (2003). Both of us disgusted in my insula: The common neural basis of seeing and feeling disgust. Neuron, 40(3), 655-664.
31. Vytal, K. E., & Hamann, S. (2010). Neuroimaging support for discrete neural correlates of basic emotions: A voxel-based meta-analysis. Journal of Cognitive Neuroscience, 22(12), 2864-2885.
32. Hariri, A. R., Bookheimer, S. Y., & Mazziotta, J. C. (2000). Modulating emotional responses: Effects of a neocortical network on the limbic system. Neuroreport, 11(1), 43-48.
33. Davidson, R. J. (2002). Anxiety and affective style: Role of prefrontal cortex and amygdala. Biological Psychiatry, 51(1), 68-80.
34. Balconi, M., & Canavesio, Y. (2013). Emotional contagion and trait empathy in prosocial behavior in young people: the contribution of autonomic (facial feedback) and balanced emotional empathy scale (BEES) measures. Journal of Clinical and Experimental Neuropsychology, 35(1), 41-48.
35. Dabas, H., Sethi, C., Dua, C., Dalawat, M. and Sethia, D., 2018, December. Emotion classification using EEG signals. In Proceedings of the 2018 2nd International Conference on Computer Science and Artificial Intelligence (pp. 380-384).
36. Liu, J., Wu, G., Luo, Y., Qiu, S., Yang, S., Li, W. and Bi, Y., 2020. EEG-based emotion classification using a deep neural network and sparse autoencoder. Frontiers in Systems Neuroscience, 14, p.43.
37. Bazgir, O., Mohammadi, Z. and Habibi, S.A.H., 2018, November. Emotion recognition with machine learning using EEG signals. In 2018 25th national and 3rd international iranian conference on biomedical engineering (ICBME) (pp. 1-5). IEEE.
38. Donmez, H. and Ozkurt, N., 2019, October. Emotion classification from EEG signals in convolutional neural networks. In 2019

Innovations in Intelligent Systems and Applications Conference (ASYU) (pp. 1-6). IEEE.
39. Salama, E.S., El-Khoribi, R.A., Shoman, M.E. and Shalaby, M.A.W., 2021. A 3D-convolutional neural network framework with ensemble learning techniques for multi-modal emotion recognition. Egyptian Informatics Journal, 22(2), pp.167-176.
40. Chowdary, M.K., Anitha, J. and Hemanth, D.J., 2022. Emotion Recognition from EEG Signals Using Recurrent Neural Networks. Electronics, 11(15), p.2387.
41. Ozdemir, M.A., Degirmenci, M., Izci, E. and Akan, A., 2021. EEG-based emotion recognition with deep convolutional neural networks. Biomedical Engineering/Biomedizinische Technik, 66(1), pp.43-57.
42. Acharya, D., Jain, R., Panigrahi, S.S., Sahni, R., Jain, S., Deshmukh, S.P. and Bhardwaj, A., 2021. Multi-class emotion classification using EEG signals. In Advanced Computing: 10th International Conference, IACC 2020, Panaji, Goa, India, December 5–6, 2020, Revised Selected Papers, Part I 10 (pp. 474-491). Springer Singapore.
43. Nath, D., Singh, M., Sethia, D., Kalra, D. and Indu, S., 2020, February. An efficient approach to eeg-based emotion recognition using lstm network. In 2020 16th IEEE international colloquium on signal processing & its applications (CSPA) (pp. 88-92). IEEE.
44. Zhu, Y. and Zhong, Q., 2021. Differential entropy feature signal extraction based on activation mode and its recognition in convolutional gated recurrent unit network. Frontiers in Physics, 8, p.636.
45. Yin, Y., Zheng, X., Hu, B., Zhang, Y. and Cui, X., 2021. EEG emotion recognition using fusion model of graph convolutional neural networks and LSTM. Applied Soft Computing, 100, p.106954.
46. Shen, F., Dai, G., Lin, G., Zhang, J., Kong, W. and Zeng, H., 2020. EEG-based emotion recognition using 4D convolutional recurrent neural network. Cognitive Neurodynamics, 14, pp.815-828.