HUMAN BEHAVIOUR ANALYSIS USING CNN

Dr. Anupama Budhewar[1*], Sanika Purbuj[1], Darshika Rathod[1], Mrunal Tukan[1], PalakKulshrestha[1]

[1]MIT-ADT University MIT ADT Campus, Rajbaugh Loni Kalbhor – 412201
+91 9764922888

***Corresponding Author:-** Dr. Anupama Budhewar
Email:-anupama.budhewar@mituniversity.edu.in

Abstract:
Emotion recognition has been the subject of extensive research due to its significant impact on various domains, including healthcare, human-computer interaction, and marketing. Traditional methods of emotion recognition rely on visual cues, such as facial expressions, to decipher emotional states. However, these methods often fall short when dealing with individuals who have limited ability to express emotions through facial expressions, such as individuals with certain neurological disorders.
This research paper proposes a novel approach to emotion recognition by combining facial expression analysis with electroencephalography (EEG) data. Deep learning techniques are applied to extract features from facial expressions captured through video analysis, while simultaneously analyzing the corresponding EEG signals. The goal is to improve emotion recognition accuracy by utilizing the complementary information offered by the interaction between facial expressions and EEG data.
Emotion recognition is a challenging task that has collected considerable recognition in the current years. Different and refined approaches to recognize emotions based on facial expressions, voice analysis, physiological signals, and behavioral patterns have been developed. While facial expression analysis has been a dominant approach, it falls short in instances where individuals cannot effectively express emotions through their faces. To overcome these limitations, there is a need to explore alternative methods that can provide a more accurate assessment of emotions. This research paper aims to investigate the collaboration and interaction between facial expressions and EEG data for emotion recognition. By combining the information from both modalities, it is expected to augment the accuracy and strength of emotion recognition systems. The proposed method can range from conducting literature reviews to designing and fine-tuning deep learning models for feature extraction, developing fusion models to combine features from facial expressions and EEG data, performing experimentation and evaluation, writing papers and documentation, preparing presentations for dissemination, and engaging in regular meetings and discussions for effective collaboration. Ethical considerations, robustness and generalizability, continual learning and skill development, and utilizing collaboration tools and platforms are also essential contributions to ensure the project's success.

Keywords: Affective Computing, Multimodal Emotion Recognition, Facial Expression Analysis, EEG Data Processing, Deep Learning Models, Feature Extraction, Real-time Emotion Detection, Individual Differences, Personalized Emotional Profiling, Human-Computer Interaction,Collaborative Environments, Mental Health Applications, Ethical Considerations, Privacy Protection, Informed Consent, Machine Learning Algorithms, Model Training and Optimization, Algorithmic Approaches, Virtual Reality Integration,

Neuro-feedback Systems.

1. Introduction

In the dynamic landscape of affective computing, the fusion of Expression-EEG Interaction with Machine Learning has emerged as a transformative paradigm in the realm of emotion recognition. This innovative approach leverages the synergistic potential of facial expression analysis and electroencephalography (EEG) data, fortified by the robust capabilities of Machine Learning algorithms. Through this integration, a multi-modal framework is established, affording a comprehensive understanding of human emotions. This nuanced comprehension is pivotal in revolutionizing a myriad of applications across diverse domains, from human-computer interaction to mental health support and beyond.
The advent of deep learning, and specifically CNNs, has provided a paradigm shift by enabling automated, data-driven approaches that can learn hierarchical representations directly from raw visual input.This paper explore the potential of CNNs in behavior analysis, with the comparative analysis by RNNs with a focus on their capacity to discern subtle and complex patterns within visual data. By leveraging the hierarchical feature extraction capabilities of CNNs, researchers can gain insights into the temporal and spatial dynamics of behaviors, surpassing the limitations of traditional methods. This study aims to contribute to the existing body of knowledge by presenting a systematic examination of CNN architectures, training strategies, and their efficacy in various behavior analysis scenarios

1.1 Applications in Human-Computer Interaction

In the of human-computer interaction (HCI), this integrated system stands poised to redefine user experiences. By endowing machines with the ability to discern and respond to human emotions, interfaces become more intuitive and adaptive. In virtual environments, this technology can facilitate seamless and responsive interactions, dynamically adjusting content and interfaces based on the user's emotional state. Similarly, in e-learning platforms, understanding the learner's emotional state can lead to personalized content delivery, optimizing comprehension and retention.

1.2 Mental Health Support and Therapy

The integration of Expression-EEG Interaction with Machine Learning holds profound implications for mental health interventions. Therapists can harness this technology to glean deeper insights into the emotional well-being of their clients. By analyzing facial expressions and EEG data, therapists can refine their diagnoses and treatment strategies. Wearable devices equipped with this technology offer continuous monitoring, providing timely interventions for individuals contending with conditions like anxiety or depression.

1.3 Ethical Considerations and Privacy Safeguards

While the potential applications of this technology are vast, ethical considerations loom prominently. The collection and analysis of facial expressions and EEG data necessitate vigilant safeguards to protect individual privacy and data security. Striking a harmonious balance between leveraging this technology for its transformative benefits while upholding the rights and privacy of individuals is imperative. In summation, the amalgamation of Expression-EEG Interaction with Machine Learning constitutes a watershed moment in emotion recognition technology. Its ramifications extend far beyond theoretical inquiry, reshaping how we perceive and engage with human emotions. Yet, it is incumbent upon us to approach this innovation with ethical acumen, ensuring that its deployment remains both responsible and beneficial in the myriad applications it promises revolutionize.

2. Literature Survey

Facial expression is the common signal for all humans to convey the mood. There are manyattempts to make an automatic facial expression analysis tool. as it has application in many fields such as robotics, medicine, driving assist systems, and lie detector. Since the twentieth century, Ekman et al. defined seven basic emotions, irrespective of culture in which a human grows with theseven expressions (anger, feared, happy, sad, contempt, disgust, and surprise). In a recent study on the facial recognition technology (FERET) dataset, Sajid et al. found out the impact of facial asymmetry as a marker of age estimation. Their finding states that right face asymmetry is better compared to the left face asymmetry. Face pose appearance is still a big issue with face detection. Ratyal et al. provided the solution for variability in facial pose appearance. They have used three-dimensional pose invariant approach using subject-specific descriptors. There are many issues like excessive makeup pose and expressions which are solved using convolutional neural networks and as well as recurrent neural networks. Recently, researchers have made extraordinary accomplishment in facial expression detection, which led to improvements in neuroscience and cognitive science that drive the advancement of research, in thefield of facial expression. Also, the development in computer vision and machine learning makes emotion identification much more accurate and accessible to the general population. As a result, facial expression recognition is growing rapidly as a sub-field of image processing. Some of the possible applications are human–computer interaction psychiatric observations, drunk driver recognition and the most important is lie detector. unit applied on a background-removed face image. In the proposed FERC model, we also have a non-convolutional perceptron layer as the last stage. Each of the convolutional layers receives the input data (or image), transforms it, and then outputs it to the next level. This transformation is convolution operation. All the convolutional layers used are capable of pattern detection. Within each convolutional layer, four filters were used. The input image fed to the first-part CNN (used for background removal) generally consists of shapes, edges, textures, and objects along with the face. The edge detector, circle detector, and corner detector filters are used at the start of the convolutional layer 1. Once the face has been detected, the second-part CNN filter catches facial features, such as eyes, ears, lips, nose, and cheeks. The edge detection filters used in this layer are shown in Fig. 3a. The second-part CNN consists of layers with 3 × 3 kernel matrix, e.g., [0.25, 0.17, 0.9; 0.89,0.36, 0.63; 0.7, 0.24, 0.82]. These numbers are selected between 0 and 1 initially. These numbers are optimized for EVdetection, based on the ground truth we had, in the supervisory training dataset. Here, we used mini-mum error decoding to optimize filter values. Once the filter is tuned by supervisory learning, it is then applied to the background-removed face (i.e., on the output image of the first-part CNN), for detection of different facial parts (e.g., eye, lips. nose, ears, etc.)

To generate the EV matrix, in all 24 various facial features are extracted. The EV feature vector is nothing but values of normalized Euclidian distance between each face part. Other side, recurrent neural layers worked same as the convolutional layers but then recurrent network layers struggle with capturing long-term dependencies in sequences.

3. Proposed Work

The libraries, mainly NumPy, OpenCV, and Keras are use in Python. NumPy is the open-source library that is used while working with arrays. OpenCV is a BSD licensed library that includes several hundred computer algorithms, it is also an open source. Keras is a high-level library for deep learning, built on top of TensorFlow. It is written in Python and provides a convenient way to create a range of deep learning models.

The experimental results show that a high accuracy is reached with a small amount of

training data, especially Kaggle dataset, which achieves an accuracy of 95.49%. It is supposed to give out 98% accuracy as per 2020 data census. The accuracy of the graph has only grown as compared to the losses over the years. According to dimensionality, the emotion model can also be categorized into two types, namely the 2D (two-dimensional) model and the 3D (three-dimensional) model. Mood detection often relies on sentiment lexicons or dictionaries containing words and phrases associated with different emotions. These resources help in assigning sentiments to text based on the presence of specific words or expressions.
The results confirmed the effectiveness of speech signals for mood detection but these signals are sensitive. The system possibly cannot make a correct judgment when the subject's inner trueemotions and external performance are inconsistent. The mood detection method based on physiological signals can achieve accuracy, while the collected data can objectively reflect the emotional state of the subjects. Conferring to the different signals recorded, existing mood recognition methods can be roughly divided into EEG signal-based, facial video feature-based, and multimodal emotion recognition.

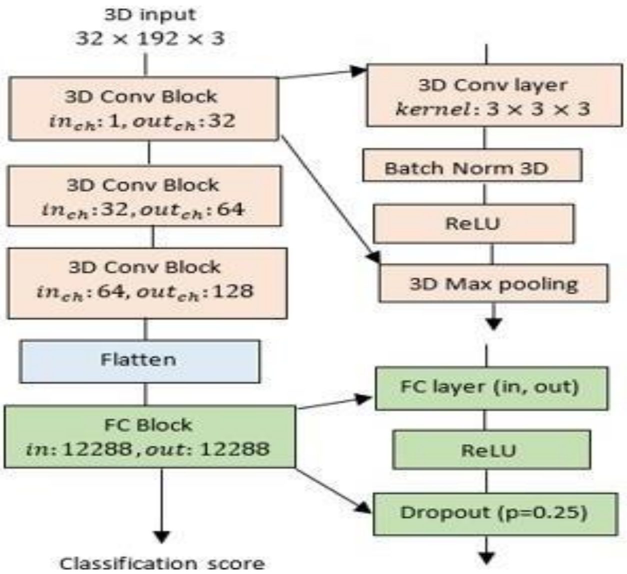

Fig.1. CNN model architecture for emotion detection

3.1 Method on Convolutional Neural Network

a. Data Collection

Data collection and preparation are crucial steps when creating a mood detection system using a convolutional neural network (CNN). The mood labels can include categories like "happy," "sad," "angry," "neutral," or any other mood classes relevant to your application. The dataset includes a diverse set of images that accurately represent the moods you want to detect. The dataset is used on the Kaggle Kernel server for data collection. It allows you to process machine learning operationson cloud computers. For more accuracy, the dataset is imported into the Keras library.

b. Importing the libraries

The importing deep learning method is used for building the model into a CNN model. Keras is the main library to import such models using Python in deep learning. To import the model in Keras, first load images, convert them into an array image, and then the image data generator is imported. After that, the main method of CNN is layering the models into libraries using Dense, Batch Normalization, Activation , Global Average Pooling

2D,Flaten, Convo 2D, , Input, Dropout, Activation, Max Pooling 2D, Global Average Pooling 2D etc.

c. Defining the Convolutional Neural Network Layer

In defining the first, second, and third layers are defined. In the first CNN layer, 64 filters and (3,3)size are used. Now, in the 2nd CNN layer, 128 filters and (5,5) size are used. 512 filter and (3,3) areused in the 3rd CNN layer. To input the 1D array, a flatten function is used that is fully connected toall layers of the CNN layers. After that, the ADAM optimizer is optimized with lr = 0.001.

The features used in traditional mood detection methods are external, such as facial expressions, body language, etc. The goal of mood detection is to automatically classify content as positive, negative, or neutral, or to categorize emotions such as happiness, anger, sadness, and more. Various ML and deep learning techniques are employed for mood detection, including Naive Bayes and convolutional neural networks (CNN).

The accuracy rate of the model between the facial language of different people is 64.3%, while testing for the same person has an accuracy rate of 93.2%, indicating that facial expressions can be adapted to fruitfully recognize emotions. It features and speech content for emotion recognition based on EEG signals.

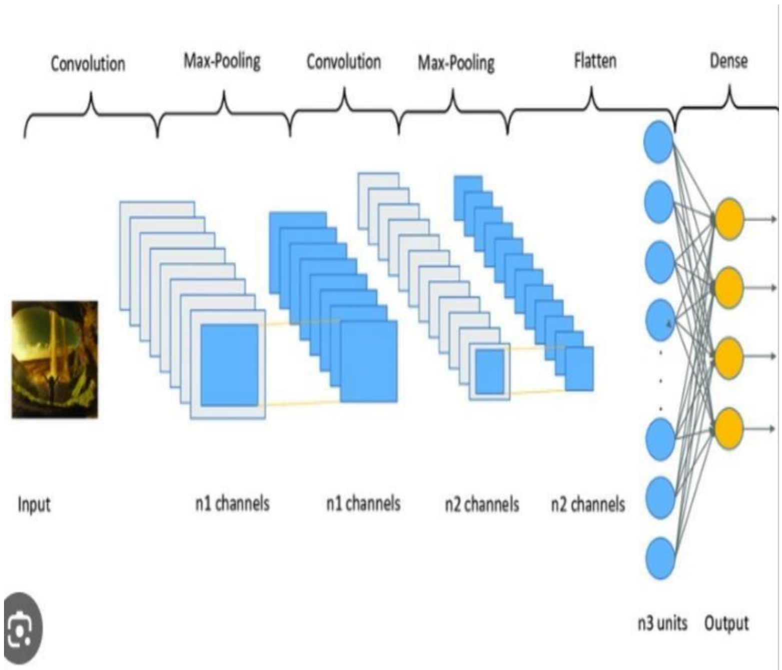

Fig.2. Block diagram of convolutional neural network

The fig .2 (taken from camera or) extracted from the video. The input image is then passed to the first-part CNN for background removal. After background removal, facial expressional vector (EV) is generated. Another CNN (the second-part CNN) is applied with the supervisory model obtained from the ground-truth database.

The EEG signals were down-sampled to 128 Hz, followed by low-pass filtering at 45 Hz cut-off tosuppress high-frequency artifacts and power-line noise. Channel re- referencing was performed withthe global average scheme. In the binary classification experiments, 5-fold cross validation was applied where 80% of the total samples at video trial level were used for training and 20% for testing. The models were trained for 200 epochs and outcome was the average accuracy of each fold. The training was conducted with minibatch size of 32 and Adam optimizer where the initial learning rate was 0.0001. A dropout of 0.25 was considered to avoid overfitting.

4. Experimental Analysis

The article discusses research on the recognition of speaker emotions from image signals, particularly image. It highlights the importance of processing methods that involve isolating image signals and extracting specific features for the purpose of classification.
The primary focus is on acoustics, emphasizing the valuable information derived from prosodic andspectral features in image processing techniques. The article also mentions the potential role of emotion recognition systems in aiding emotion classification through linguistic information.
The article classifies and surveys research papers related to emotion recognition from video channels, with an emphasis on feature extraction and classification methods, databases used for experimentation and performance issues. It also touches upon open trends and future research directions in this field.

IMF	Bandwidth (Hz)	Valence acc Train Test	Arousal acc Train Test
IMF1	24-45	97.45 85.37	97.45 80.56
IMF2	13-24	96.04 84.21	97.28 76.46
IMF3	8-13	94.02 77.25	96.43 71.73
IMF4	4-8	94.01 73.18	92.9 63.33
IMF5	2-4..	90.36 67.27	91.05 60.12

Table 1. Emotion Detection accuracy (%) for each IMF

11 IMFs are obtained for each channel by applying MEMD. The MHS of an IMF carries pertinent spectral information at a specific frequency range. The first five IMFs are selected according to theirmedian frequency. Table I shows the classification performance attained with individual IMFs computed by MEMD from all EEG channels. It is noted that the high oscillatory mode, i.e., IMF1,is most dis- criminative with accuracy of 85.37% and 80.56% for binary classification on valence and arousal state respectively. The classification accuracy declines significantly with the lower oscillatory modes. Using IMF4 and IMF5 alone attained 73.18% and 67.27% accuracy on valence state classification and 63.33% and 60.12% on arousal, respectively. On the whole, IMF1 IMF3 are found to be significant and the lower oscillating intrinsic modes are less important to this task. As IMF1 IMF3 provide the most discriminative features in an independent manner, spectral features delivered by frequency components ranged in 8-45 Hz are considered for emotion classification in this study.

Approach	Input	Model	Valence Acc(%)	Arousal Acc(%)
Ref [1]	Band	SAE+LSTM	81.10	74.38
Ref [2]	Raw EEG	LSTM	85.45	85.65
Ref [3]	DE	LSTM	69.06	72.97
Ref [4]	Spectrogram	2DCNN	80.46	76.56
Ref [5]	Raw EEG	3DCNN	82.32	84.12
Proposed	Images	3DCNN	89.25	86.23

Table 2. Performance comparison among recent approaches for subject independent emotion classification

Table 2 provides a performance comparison of the pro- posed system with other state-of-the-art systems of emotion detection reported on the DEAP dataset. 3D CNN with raw EEG as input features in [19] performs better than LSTM [12], [13], [14] and 2D CNN [16]

with different typesof features. Our proposed system with 3D CNN also outperforms the LSTM and 2D CNN models.
Using 3 dominant IMFs derived by the proposed system achieves accuracy of 89.25% and 86.23% on valence and arousal state detection respectively, which are significantly higher than the reported best results in [19].

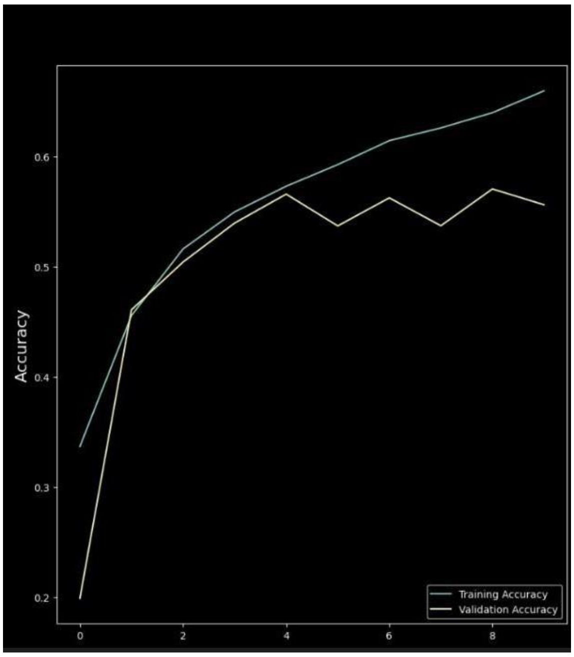

Fig.3. Comparative analysis of proposed method with existing method

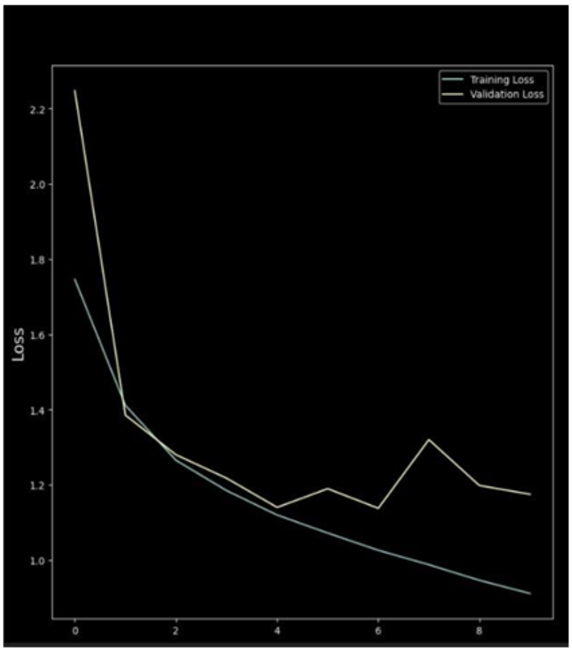

Fig.4. Comparative analysis of loss of proposed method with existing method

As shown in fig.4. A cross-database study indicates that proposed method has high generalization performance and the accuracy. In this case, since there is no measure associated with the edges of the resulting graph, their associated weights are fixed to 1.00 to indicate the existence of connections. It is supposed to give out 98% accuracy as per 2020 data census. The accuracy of the graph has only grown as compared to the losses over the year.

	Convolutional Neural Network	Recurrent Neural Network
Architecture	Feed- forward neural networks using filters and pooling.	Recurring network that feeds the result back into the network.
Data type	Relies on image data	Trained with sequence data
Input/output	The size of the input and the resulting output are fixed (i.e., receives images of fixed size and outputs them to the appropriate category along with the confidence level of its prediction)	The size of the input and the resulting output may vary (i.e., receives different text and output translations- the resulting sentences can have more or fewer words)
Commendable feature	Accuracy in recognizing images	Memory and self-learning
Ideal usage scenario	Spatial data (such as images)	Temporal/sequential data (such as text or video)
drawback	Large training data is required	Slow and complex training and gradient concern
Use cases	Image recognition and classification, face detection, medical analysis, drug discovery and image analysis.	Text translation, natural language processing, language translation, entity extraction, conversational intelligence, sentiment analysis, speech analysis.

Table 3. comparison between CNN and RNN

Another algorithm used for comparative analysis of emotion recognition is RNN (Recurrent Neural Network). RNNs are well-suited for tasks involving sequential data and capturing temporal dependencies. Training RNNs can be expensive, and they may suffer from the vanishing or exploding gradient problem. The statement that CNNs (Convolutional Neural Networks) show more accuracy than RNNs (Recurrent Neural Networks) is not universally true. The performance of these algorithms depends on the specific characteristics of the data, but the data for the analysis is used in this for training gives more accuracy by CNN. The difference between CNN and RNN for analysing it is between 20 to 25 %. RNNs might not capture spatial features as effectively as CNNs. If spatial arrangements of facial features are crucial for emotion recognition, RNNs may not perform as well as CNNs. RNNs can struggle with capturing long-term dependencies in sequences. If the emotional expression involves subtle changes that occur over an extended period. Because of the more difficulties in RNNs it might show more loss and less accuracy.

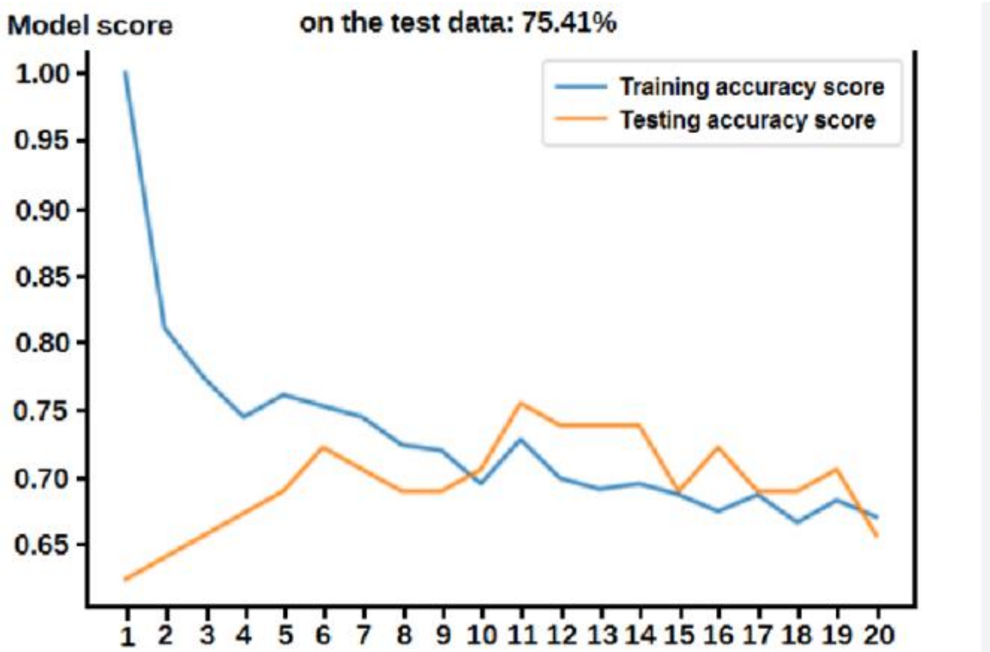

Fig.5. comparative analysis of RNN method

The experimental results show that a high accuracy is reached with a small amount of training data as shown in fig.3, especially Kaggle dataset, which achieves an accuracy of 95.49%. Accuracy calculated by CNN, RNN, CNN Figure.3 illustrate the accuracy calculated by CNN is 95.49% approximately, RNN is 75% approximately. This means the accuracy of CNN is better than RNN. Accuracy can be calculated as: True positive Rate (or sensitivity): TPR=TP/ (TP+FN) False positive Rate: FPR=FP/ (FP+TN) True negative rate (or specificity): TNR=TN/ (FP+TN) False negative Rate: FNR=1−TPRTrue Positive (TP): Correctly classified as class of Normal True Negative (TN): Correctly classifiedas not the class of Normal False Positive (FP): Incorrectly classified as class of Normal False Negative (FN): Incorrectly classified as not the class of Normal. The F1-score value for the CNN method is 0.95 and the RNN method is 0.75. The CNN allows the model to produce an average precision of 96.5%, the accuracy of 95.49%, recall of 95.49%, and F1-score of 95.49% (average).

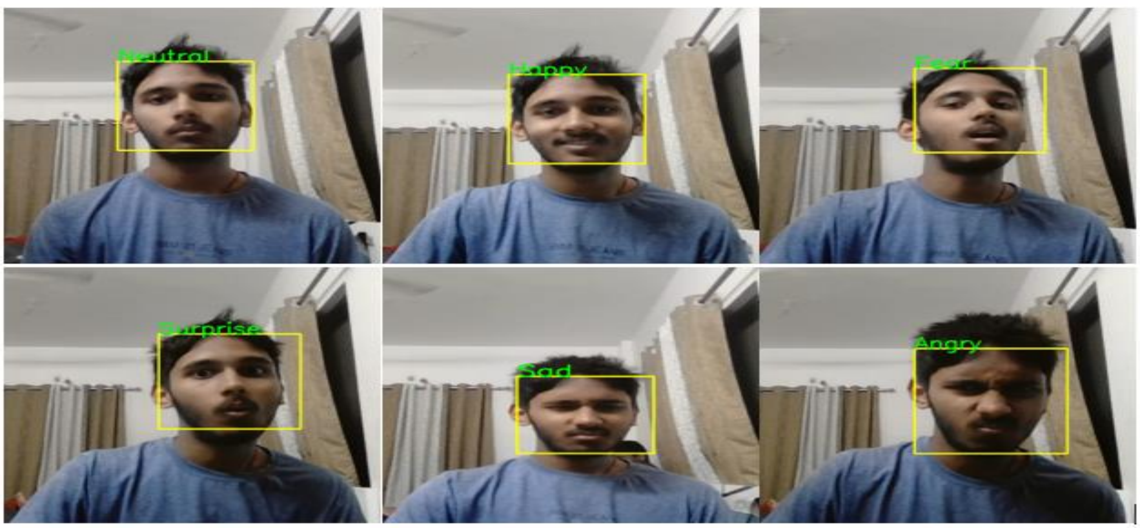

Fig.6. emotions captured by proposed model using CNN

Figure 6 shows the various emotions captured using CNN. The model analysed the facial features and recognized subtle patterns from the training data set to give the result with high accuracy in image processing tasks. CNN learns to recognize patterns and features from the training dataset and gives accurate predictions on new, unseen data.

5. Conclusion

The introduction outlines a research paper focused on Mood detection with Expression-EEG Interaction and Collaboration Using Machine Learning. It highlights the importance of this novel approach, which combines facial expression analysis and EEG data to enhance emotion recognition.The paper discusses the key elements of the research project, such as literature reviews, deep learning models, ethical considerations, and real-time emotion detection. It also emphasizes the potential applications in human-computer interaction, mental health support, and virtual environments, while underscoring the importance of privacy protection and informed consent. The summary encapsulates the paper's core concepts, making it clear that this innovative approach has the potential to revolutionize various domains while emphasizing the need for responsible and ethical deployment. This research is all about understanding people's feelings better. We're using a combination of looking at their facial expressions and brain activity to do this. It's a new and smart way to figure out how someone is feeling. This can help in many ways, like making computer programs that can understand and react to our emotions or helping people with their mental health.

In simple terms, we're looking at faces and brains to understand feelings, and this can be super helpful in many situations. But we also need to be careful and make sure people's privacy is protected when we use this technology. This research paper introduces an innovative approach to understanding people's emotions. It combines the analysis of facial expressions and brain activity (EEG data) using machine learning. The goal is to improve how we recognize emotions. This can have many practical uses, such as creating smarter computer programs that can respond to our feelings and helping people with their mental health. In summary, the paper is about using technology to better understand how people feel by looking at their faces and brains. It has great potential in various areas but also emphasizes the importance of protecting people's privacy and being ethical when using this technology.

References

1. Monira Islam, Tan Lee," MEMD-HHT based Emotion Detection from EEG using 3D CNN", July 11-15, 2022
2. HONGLI ZHANG,"Expression-EEG Based Collaborative Multimodal Emotion Recognition Using Deep AutoEncoder",September 7, 2020
3. GUOSHENG YANG , RUI JIAO , HUIPING JIANG , AND TING ZHANG,"Ground Truth Dataset for EEG-Based Emotion Recognition With Visual Indication",n October 13, 2020,
4. MD. RABIUL ISLAM, (Member, IEEE), MOHAMMAD ALI MONI,MD. MILON ISLAM," Emotion Recognition From EEG Signal Focusing on Deep Learning and Shallow Learning Techniques, June 22, 2021.
5. Dahua Li , Jiayin Liu , Yi Yang , Fazheng Hou , Haotian Song, Yu Song,"Emotion Recognitionof Subjects With Hearing Impairment Based on Fusion of FacialExpression and EEG Topographic Map",IEEE TRANSACTIONS ON NEURAL SYSTEMS AND REHABILITATION ENGINEERING, VOL. 31, 2023.
6. Salma Alhagry,Aly Aly Fahmy,Reda A. El-Khoribi,"Emotion Recognition based on EEG usingLSTM Recurrent Neural Network",Vol. 8, No. 10, 2017

7. Saeed Turabzadeh, Hongying Meng ,*ID , Rafiq M. Swash 1 ID , Matus Pleva 2 ID and Jozef Juhar ,"Facial Expression Emotion Detection for Real-Time Embedded Systems", 26 January 2018.
8. Saket S Kulkarni1, Narender P Reddy*1 and SI Hariharan,"Facial expression (mood) recognition from facial images using committee neural networks",5 August 2009.
9. J. A. Russell, A circumplex model of affect, Journal of Personality and Social Psychology, vol. 39, no. 6, pp. 1161, 1980.
10. Katie Moraes de Almondes1, Francisco Wilson Nogueira Holanda Júnior2,Maria Emanuela Matos Leonardo1 and Nelson Torro Alves3,"Facial Emotion Recognition andExecutive Functions",17April,2020.
11. E. S. Salama, R. A. El-Khoribi, M. E. Shoman, and M. A. W.laby, EEG-based emotionrecognition using 3D convolutional neural networks, Int. J. Adv. Comput. Sci. Appl, vol. 9, no. 8, pp. 329-337,2018.
12. P. Gaur, R. B. Pachori, H. Wang, and G. Prasad, An automatic subject specific intrinsic mode function selection for enhancing two-class EEG-based motor imagery-brain computer interface,IEEE SensorsJournal, vol. 19, no. 16, pp. 6938-6947, 2019.
13. S. B. Wankhade and D. D. Doye, ‘‘Deep learning of empirical mean curve decomposition- wavelet decomposed EEG signal for emotion recognition,’’ Int. J. Uncertainty, Fuzziness Knowl.-Based Syst., vol. 28, no. 01, pp. 153–177, Feb. 2020
14. S. Zhao, A. Gholaminejad, G. Ding, Y. Gao, J. Han, and K. Keutzer, ‘‘Personalized emotion recognition by personality-aware high-order learning of physiological signals,’’ ACM Trans. Multimedia Comput., Commun., Appl., vol. 15, no. 1s, pp. 1–18, Feb. 2019.
15. S. Aydin, "Deep learning classification of neuroemotional phase domain complexity levels induced by affective video film clips,"IEEE J. Biomed. Health Informat., vol. 24, no. 6, pp. 1695–1702,Jun. 2020.
16. H. Kawano, A. Seo, Z. G. Doborjeh, N. Kasabov, and M. G. Doborjeh,"Analysis of similarity and differences in brain activities between perception and production of facial expressionsusing EEG data and the NeuCube spiking neural network architecture," in Proc. Int. Conf. Neural Inf. Process., vol. 9950, 2016.
17. Z. Liang et al., "EEGFuseNet: Hybrid unsupervised deep feature characterization and fusionfor high-dimensional EEG with an application to emotion recognition," IEEE Trans. Neural Syst. Rehabil. Eng., vol. 29, pp. 1913–1925, 2021.
18. W. Huang, Y. Xue, L. Hu, and H. Liuli, S-EEGNet: Electroencephalgram Signal Classification Based on a Separable Convolution Neural Network With Bilinear Interpolation, IEEE Access, vol. 8, pp. 131636-131646, 2020.
19. E. S. Salama, R. A. El-Khoribi, M. E. Shoman, and M. A. W. Shalaby, A 3D-convolutional neural network framework with ensemble learning techniques for multi-modal emotion recognition, Egyptian Informatics Journal, vol. 22, no. 2, pp. 167-176, 2021.
20. J. Chen, H. Li, L. Ma, H. Bo, and X. Gao, Application of EEMDHHT method on EEG analysis for speech evoked emotion recognition, In 2020 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), IEEE, 2020
21. A. Graves and J. Schmidhuber, "Framewise Phoneme Classification with Bidirectional LSTM and Other Neural Network Architectures," Neural Networks, vol. 18, no. 5-6, pp. 602–610, June/July 2005.
22. V. Blanz, K. Scherbaum, and H. Seidel, "Fitting a morphable model to 3d scans of faces," in Proceedings of International Conference on Computer Vision, 2007.
23. I. Kotsia and I. Pitaa, "Facial expression recognition in image sequences using geometric deformation features and support vector machines," IEEE Transaction On

Image Processing, vol. 16, no. 1, 2007.
24. P. Ekman, “Universals and cultural differences in facial expressions of emotion,” in Nebraska Symposium on Motivation 1971, J. Cole, Ed., vol. 19. Lincoln, NE: University of Nebraska Press, 1972, pp. 207–283
25. M. Schuster and K. K. Paliwal, “Bidirectional recurrent neural networks,” IEEE Transactions on Signal Processing, vol. 45, pp. 2673– 2681, November 1997
26. J. Cohn, A. Zlochower, J.-J. J. Lien, and T. Kanade, “Feature-point tracking by optical flow discriminates subtle differences in facial expression,” in Proceedings of the 3rd IEEE International Conference on Automatic Face and Gesture Recognition, April 1998, pp. 396 – 401
27. S. Romdhani, “Face image analysis using a multiple feature fitting strategy,” Ph.D. dissertation, University of Basel, Computer Science Department, Basel, CH, January 2005.