Journal of Information Systems Engineering and Management
2025, 10(7s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Feature and Decision Levels Fusion for the Synergistic Analysis of Facial Expressions and EEG Signals in the Context of Discrete Emotion Recognition

Mr. Amit S. Ufade[1], Dr. V. J. Gond[2], Ms. Madhuri D. Kawade[3]
[1]Research scholar Sandip Foundations, SITRC Nashik,
Email: amitufade13@gmail.com,
[2]Professor,HOD,E&TC Engineering MET IOE BKC Nashik,
Email: vitthalg_ioe@bkc.met.edu
[3]Assistant professor, Computer Engineering, Dept SNJB'S COE Chandwad Nashik,
Email: kawade.mdcoe@snjb.org

ARTICLEINFO

Received: 08 Oct 2024
Revised: 10 Dec 2024
Accepted: 24 Dec 2024

ABSTRACT

The undertaking of brain-computer interface emotion recognition represents a challenging task that demands meticulous thinking in order for machines to discern human emotions and respond appropriately. This work aims to improve the efficacy of individual modalities by incorporating multimodality for emotion recognition, which employs two distinct modalities. Electroencephalogram (EEG) data and facial expressions are used as independent modalities for emotion identification, with each modality evaluated separately. To combine the modalities, decision- and feature-level fusion algorithms are used. While EEG-derived emotions are primarily classified as continuous domains of valence and arousal, facial emotions are primarily classified as discrete emotions, which presents a considerable hindrance to the fusion process. Ashford Bird[10]. present the dataset used for EEG experiments at UKCI-2019.the CFEE dataset serves as the foundation for face emotion recognition for the proposed work[9]. EEG signals are analyzed for statistical properties such as mean, standard deviation, skewness, and kurtosis, while the face emotion dataset is used to identify various action units. Given the large number of features, redundant feature removal approaches are used to determine feature efficacy. Using the product rule, decision-level fusion obtains an accuracy of 80%. Accuracy in feature-level fusion is 94.48% for KNN and 98.66% for SVM classification which are comparatively higher than individual average accuracy of facial expression 88.04% and 90.73% for EEG signal.

Keywords: Emotion recognition,EEG-Electroencephalography,Feature level fusion,decision level fusion,facial expression

INTRODUCTION

Humans use a variety of sensory modalities to perceive the world, including but not limited to touch, taste, aural perception, and vision. Recent advances in deep learning algorithms enable the processing of multimodal data, which reduces computing complexity. Emotion detection can be divided into three broad categories based on the number of modalities used: single modality, bimodal, and multimodal emotion recognition.Depending on the sensor type used, three methods of emotion identification can be distinguished: non-contact emotion recognition, as demonstrated by facial expressions; peripheral physiological signals such as heart rate, EEG, and ECG; and central nervous system reactions.One captivating field within affective computing is emotion recognition for brain-computer interfaces, which predominantly utilizes a singular modality for the identification of emotions, drawing upon various human sensory inputs. Face,voice,different biomedical signals can be used to recognize emotions .Emotion, sentiment, and mood possess distinct characteristics and can often be deceptive. In many instances, reliance on a singular modality may lead to erroneous interpretations; for instance, the human visage and its emotional expressions can be misleading, a principle that holds true for other modalities such as speech. Biomedical signals, including EEG, ECG, and EMG, cannot be easily concealed like facial expressions; however, the acquisition and processing of such signals present formidable challenges.Recognizing emotion across any modality is a fundamental pattern recognition task that requires a thorough understanding of the modality, as well as skilled feature engineering and algorithmic development. The combination of various modalities is increasingly acknowledged for its ability to address potential constraints and improve the overall accuracy of individual modalities through data fusion from diverse sources, constituting an exciting and difficult task. Techniques for fusion, such as sensor fusion and feature fusion, are employed in this context. This experimentation is directed towards the establishment of a robust emotion recognition system utilizing two distinct modalities, specifically facial expressions and EEG signals. Presently, there exists a dearth of datasets containing individually measured and recorded signals from disparate sources, highlighting a critical demand for such datasets that could significantly advance research within this specialized domain.Emotion constitutes a psychological state and an affective response to an occurrence, predicated upon individual subjective experience.

EMOTION CLASSIFICATION

Discrete emotion theory claims that there are few primary emotions.As per discrete emotion theory, these distinctive core emotions represent inherent emotional responses that are universally expressed and recognized by all individuals, irrespective of their origins or social context.In 1972, Paul Ekman and his associates conducted a cross-cultural study, ultimately concluding that the six fundamental discrete emotions which are anger, disgust, fear, happiness, sadness, and surprise.

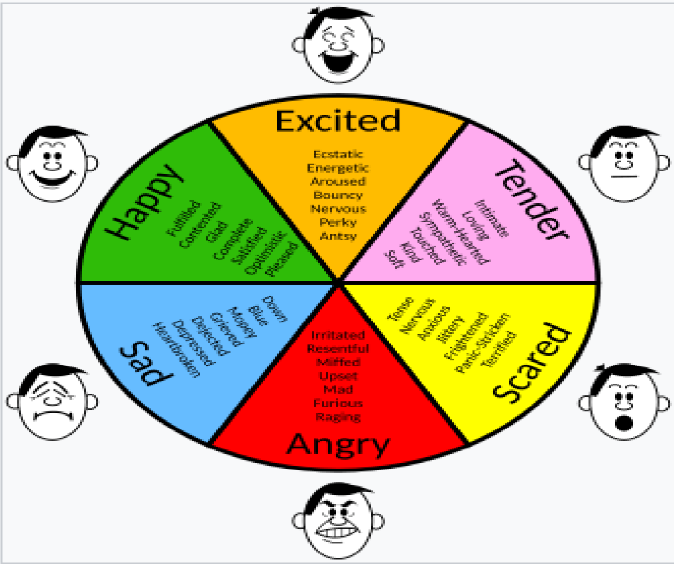

Fig.1.Discrete emotion classes[Source:wikipedia.org]

Dimensional frameworks of emotion attempt to elucidate human affectivity by ascertaining their localization within two or three spatial dimensions. Predominantly, these models encompass the factors of valence and arousal/intensity.James Russell established the circumplex paradigm of emotion.This conceptual framework posits that emotions are arrayed within a two-dimensional circular realm characterized by arousal and valence dimensions. Arousal is epitomized by the vertical axis, while valence is embodied along the horizontal axis; the center of the circle signifies a state of neutral valence coupled with moderate arousal.

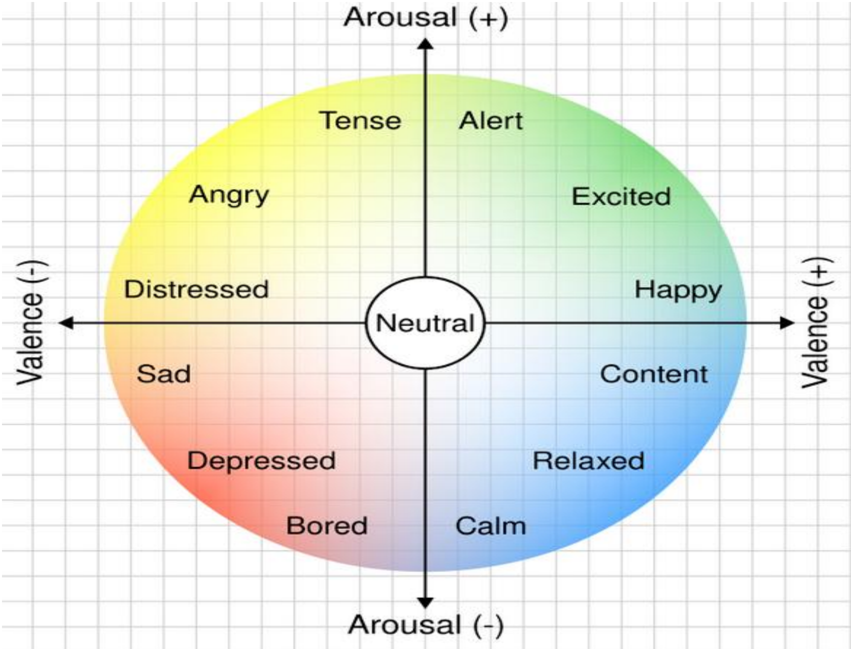

Fig.2.Dimensional model of emotion.[Source:wikipedia.org]

MATERIAL AND METHODS

A. EEG Database Description: The electroencephalogram (EEG) dataset utilized in this experimental attempt was derived from the Advances in Computational Intelligence Systems conference of 2019, curated by Ashford J. Bird et al., in their seminal work entitled 'Classification of EEG Signals Based on Image Representation of Statistical Features.' Within this scholarly pursuit, various statistical attributes are meticulously extracted from the dataset, subsequently employed to generate images. Convolutional Neural Networks (CNNs) are harnessed for the classification of these images. The authors employed a MUSE EEG headband, a commercial EEG sensing apparatus, featuring four electrodes—TP9, AF7, AF8, and TP10—strategically positioned in accordance with the 10-20 system. Recognizing that emotional states necessitate external or internal stimuli, the authors incorporated video stimuli to elicit emotional responses. During the actual signal recording phase, data is captured at a variable frequency, reaching up to 300 Hz and down to a few microvolts. The presented research leverages the same dataset to identify various significant statistical features which may enhance the classifier's accuracy while mitigating the risk of model overfitting.

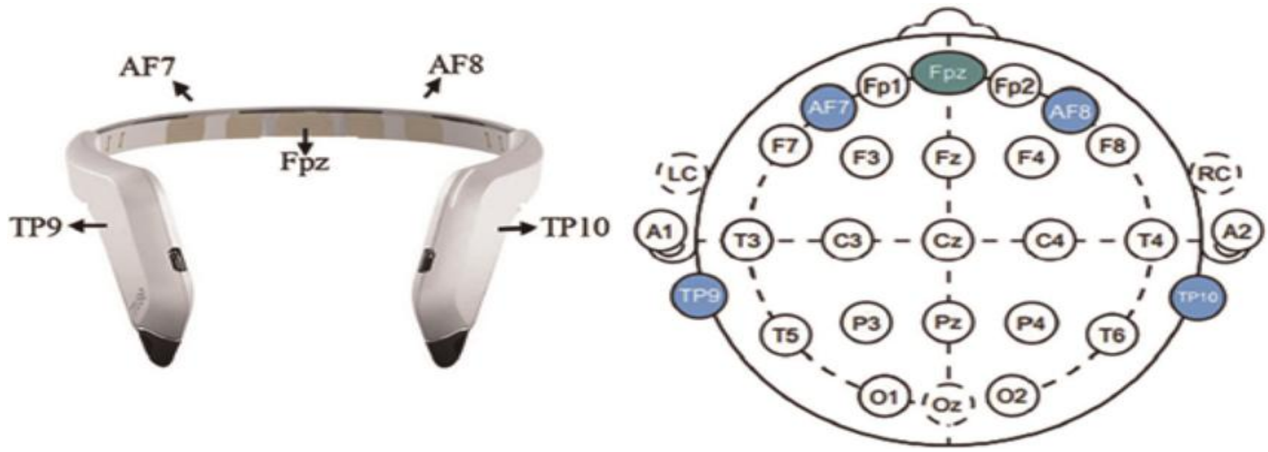

Fig.3.Muse Electrode placement for TP9, AF7, AF8, and TP10[Source:Teixeira, Ana & Gomes, Anabela & Brito-Costa, Sonia. (2023). An Overview of Mindwave Applications: Study Cases. 10.5772/intechopen.112736.]

B. Facial Database Description: for the emotion recognition predicated on facial expressions and the subsequent extraction of features for feature level integration, the Compound Facial Expressions of Emotion (CFEE) dataset is employed. This dataset encompasses approximately 1,605 images, meticulously categorized into seven distinct classes. Each image is a color photograph, boasting a resolution of 1000 x 750 pixels.

RELATED WORK

Fusion of Facial Expressions and EEG for Multimodal Emotion Recognition by Yongrui Huang et al. published in Computational Intelligence and Neuroscience,presents a novel approach to emotion recognition by integrating facial expressions and electroencephalogram (EEG) signals. The study aims to enhance the accuracy of emotion detection by leveraging the strengths of both modalities, addressing the limitations inherent in using either source independently.The study utilized a set of movie clips designed to evoke four specific emotional states: happiness, neutrality, sadness, and fear.Facial expressions were analyzed using a neural network classifier, while EEG signals were processed through two support vector machine (SVM) classifiers to detect both emotional states and intensity levels,two decision level fusion operations, a sum and a production rule, were able to combine the results from both classifiers. The fusion techniques were able to achieve 81.25% and 82.75% accuracy, which were higher than the separate accuracies of facial expression detection (74.38%) and EEG detection (66.88%). The statistical analysis conducted demonstrated that these multimodal fusion techniques could enhance the ability to accurately recognize emotions. The results illustrate that the combination of the facial expression and EEG data helps to neutralize the weaknesses that each of the modalities has, such as how mobile the facial expressions can be in comparison to the more stable, yet less informative, EEG measures of emotion. The work underlines the potential of decision level fusion as a practical approach of dealing with different physiological signals which increases the robustness of emotion recognizers. In this paper, the authors are aware of the constraints that come with their current dataset and use of one electrode sensor.

A Multi-Modal Emotion Recognition Approach Using Facial Expressions And Electroencephalography, authored by Ying et al, provides solutions to specific challenges that Human-Robot Interaction(HRI) systems face such as a clash of emotions that negatively impacts the interaction and greatly hinders effective communication between humans and robots. The authors propose an innovative approach for multi-modal emotion recognition through facial expressions and EEG (Electroencephalogram) to solve the problem of weakened emotions in HRI systems. The two methods are the combination of face expressions, image classification algorithms and the EEG signals feature extraction algorithms. The study first employs public datasets to train the model as it was later tested on data collected from subjects. When facing the problem of sparse data sets, the combination recognition techniques are prepared using the Monte Carlo method that strengthens the recognition process. The results of combining different approaches gave a recognition rate of 83.33%.Additionally, a perceptual assessment conducted with participants yielded an average satisfaction score of 7 out of 10, indicating a positive user experience.the work lacks further research to refine the system, particularly in real-world applications.

Power Spectral Density Based Discrete Emotional State Recognition System Using Electroencephalography Signals by Ufade et al.presents a system for recognizing discrete emotional states through the analysis of electroencephalography (EEG) signals, utilizing power spectral density (PSD) as a key feature. The study emphasizes the importance of emotion recognition in human-machine interaction and investigates several approaches, including the use of classifiers to improve accuracy.While applying k-fold cross validation, the SVM classifier attained the highest accuracy of 92.76 percent on the DREAMER dataset followed by KNN classifier 81.90%. Logistic Regression performed the worst because it is a linear classifier. For the SEED-IV dataset, the LSTM classifier achieved a maximum of 74 percent accuracy, while GRU classifier achieved 75%. The results depict that the choice of algorithm had a huge difference in recognition accuracy rates. The research places emphasis on the implementation of multimodal approaches, stating that using EEG along with other modalities can increase recognition accuracy even further. The authors also consider the problems of EEG signal drifting including the need for normalization techniques in order to boost classification results.

Petrantonakis et al. (2006), in their paper, Emotion Recognition from EEG Using Higher Order Crossings, published in IEEE Transactions on Information Technology in Biomedicine, suggested new methods of emotion recognition utilizing electroencephalogram (EEG) signals and placed particular focus on higher order crossings (HOC) used for feature extraction. The paper studies the potential of this method in recognizing the six basic emotions: happiness, surprise, anger, fear, disgust and sadness. The classification accuracy came out to be 62.3% with QDA for single channel case, and 83.33% with SVM for the combined-channel case.The study suggests that the elaborated HOC-EC approach is more effective in the recognition of emotions from EEG signals than the already existing methods. The authors indicate that their findings have an impact on the design of an affective computer, especially in the field of medicine.The HOC-EC's ability to accurately classify emotions from EEG signals suggests its potential for integration into human-machine interfaces, enhancing the emotional responsiveness of such systems.

METHODOLOGY

1.EEG based Emotion Recognition System: The primary goal of this work is to develop a system capable of detecting emotions through independent analysis of EEG signals and facial expressions. Furthermore, the aim is to create a system capable of doing feature- and decision-level fusion. The ultimate goal of creating these independent systems and combining them is to improve the total accuracy of each modality.emotion recognition is fundamentally a pattern recognition task ,as shown in fig.4 the subsequent sensor data pertaining to object class classification principally encompasses the phases of feature extraction and the selection of salient features. The identification of pertinent features for the task at hand presents a formidable challenge, yet remains a crucial undertaking for ascertaining both the model and classifier efficacy. Numerous algorithms have been established, with many currently in development, to facilitate the identification of relevant features.

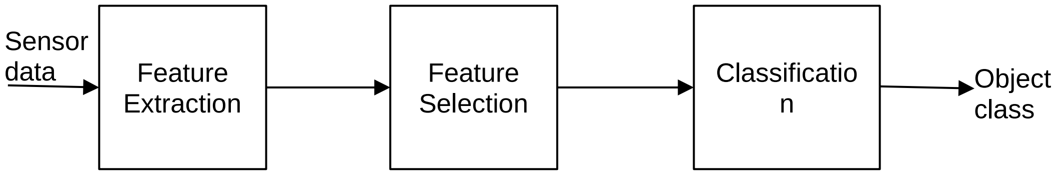

Fig.4.Pattern Recognition Basic steps

The proposed undertaking encompasses the availability of electroencephalogram (EEG) recordings within the provided dataset. EEG signals collected from four distinct electrodes are accessible, accompanied by thorough preprocessing attempts. A multitude of features may be derived from EEG signals, which predominantly fall into three principal categories: 1. time domain, 2. frequency domain, 3. time-frequency domain. Each category possesses its unique advantages and disadvantages. Time domain feature extraction is executed directly on the original signal and yields significant results. As illustrated in Figure 5, an EEG-based emotion recognition system can be developed independently.

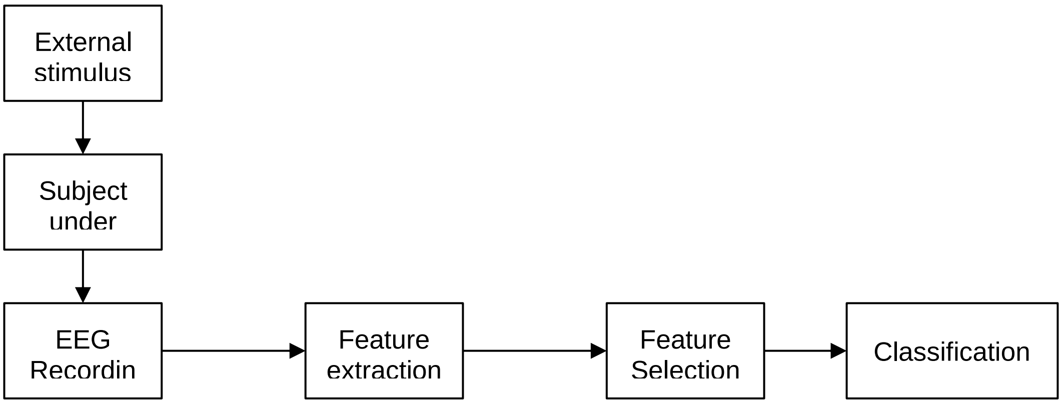

Fig.5.EEG based emotion recognition System.

An analogous framework is employed for the recognition of facial emotions through external stimuli, with independent execution of feature extraction. For the purpose of decision-level fusion, the following schematic diagram is proposed, as illustrated in fig. 6. In this context, various decision-making strategies, such as the sum rule or product rule, assume significant importance and must be judiciously selected.

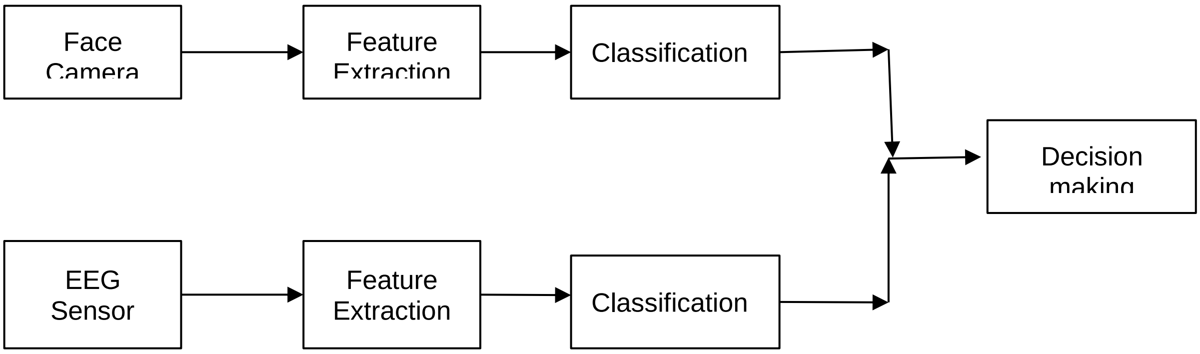

Fig.6.Proposed system block diagram for Decision level fusion

Experiments are additionally conducted to explore feature-level fusion by amalgamating attributes sourced from both modalities. A unified vector is synthesized from the selected features of facial expressions and electroencephalogram (EEG) signals. The creation of a dataset that encompasses both EEG signals and facial expressions from the same individual poses substantial challenges due to the intrinsic characteristics of EEG data. Given their extreme susceptibility to artifacts during recording, subjects undergoing testing are required to remain motionless to mitigate the introduction of extraneous artifacts.

A surplus of statistical characteristics are derived from the preprocessing of EEG signals. The meticulous extraction of features from the signal constitutes a fundamental challenge in the realm of brain-computer interfacing. The extracted features are systematically organized in tabular form for each column, encompassing metrics such as mean, standard deviation, skewness, kurtosis, and the minimum and maximum values. The mean indicates the central tendency of the data set, while the standard deviation indicates the variation or dispersion around the mean. Skewness measures the asymmetry of the data distribution; Kurtosis also determines the tails of the distribution, with higher kurtosis values indicating a higher proportion of outliers.

Recursive feature elimination(RFE) : RFE is a powerful algorithm for feature engineering that plays an important role. This method is known for its methodical selection process.The suggested study extracts a wide range of information from EEG signals. The task of feature selection becomes a lengthy endeavor when training and evaluating the classifiers in question.. Recursive Feature Elimination (RFE) assumes a pivotal role in such contexts, wherein a systematic process of identifying a subset of pertinent features for integration into model construction is employed. In summary, RFE serves to mitigate model complexity and elevates performance by pinpointing the requisite features aligned with the specific task at hand.

Algorithm of RFE:

$Input: D = \{a_1, a_{2,} a_3 \ldots\ldots\ldots\ldots\ldots\ldots a_n\}$

$Output\ Feature\ Rank:\ R = \{r_{a1}, r_{a2}, r_{a3} \ldots\ldots\ldots\ldots r_{an}\}$

1.$Built\ a\ classifier\ by\ \ from\ the\ training\ dataset\ from\ D.$

2.$Observe\ the\ performance\ of\ the\ classifier.$

3.$for\ each\ feature\ \ a_i\ in\ D.$

4.$D = D - a_i.$

5.$Train\ the\ classifier\ with\ D.$

6.$Compute\ the\ classification\ accuracy.$

7.$Find\ out\ the\ accuracy\ loss\ of\ the\ classifier\ due\ to\ elimination\ of\ a_i.$

8.$Compile\ the\ loss\ profile\ of\ features\{a_1, a_{2,} a_3 \ldots\ldots\ldots\ldots\ldots\ldots a_n\}$

9.$Compute\ the\ feature\ rank\ R = \{r_{a1}, r_{a2}, r_{a3} \ldots\ldots\ldots\ldots r_{an}\}.from\ the\ loss\ profiles.$

The EEG dataset used for the experimentation purpose is a labeled dataset with positive,negative,neutral as a label,and is a balanced dataset.

Following tables show the result achieved for features selected after following the algorithm of RFE.

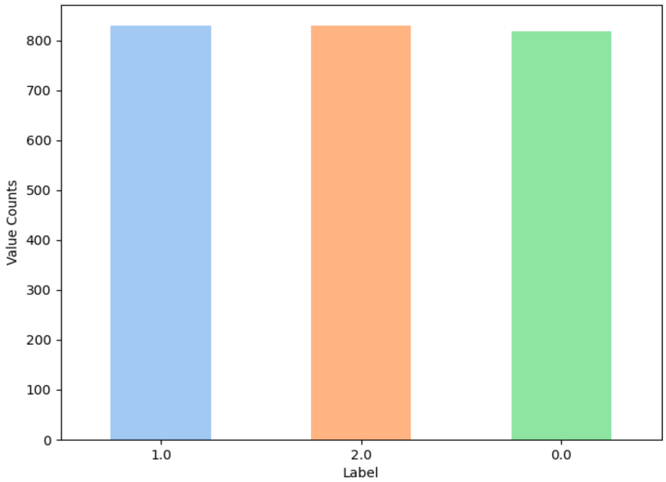

Fig. 7.Value count for 3 label vs number of features

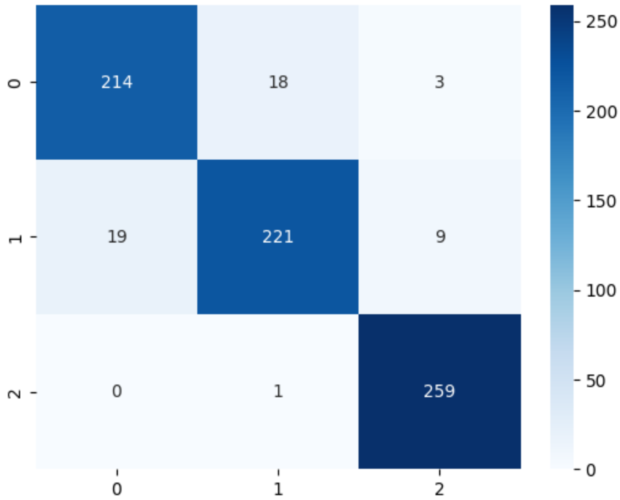

Fig.8.Confusion matrix for SVM classification of EEG based emotion classification System

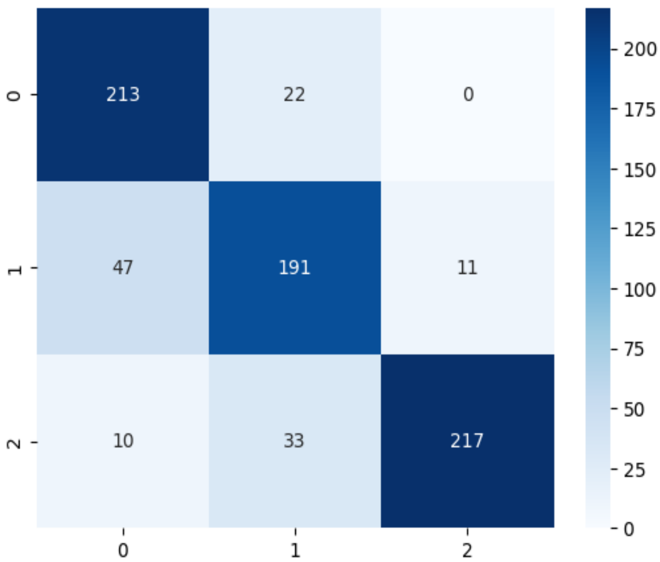

Fig.9.Confusion matrix for KNN classification of EEG based emotion classification System

The feature set is split into training and validation sets with 70-30 ratio.in order to check the performance of the classification.

Table 1.Results of EEG based emotion classification System.

Sr.no	Name of the classifier	Training Accuracy in %	Testing Accuracy in %	K-Fold validation Accuracy in %
1	Support Vector Machine(SVM)	93.68	93.28	90.73
2	K-nearest Neighbour (KNN)	89.42	84.02	83.46

2.Facial Emotion Recognition System: As previously noted, the Compound Facial Expressions of Emotion (CFEE) dataset serves as the foundation for the identification of facial features. The Facial Action Coding System (FACS), conceived by Paul Ekman and Wallace V. Friesen in 1978, constitutes a thorough framework for categorizing facial movements predicated on muscular activity. FACS dissects facial expressions into Action Units (AUs), wherein scholarly literature posits that particular combinations of these Action Units correspond to distinct emotions. A plethora of deep learning methodologies can be utilized to autonomously extract Action Units from facial images and videos.various packages are available in python for the same in the work proposed py-feat library is used for extracting Action units and corresponding emotions.alternative option to extract emotion from facial expressions is MTCNN Library.

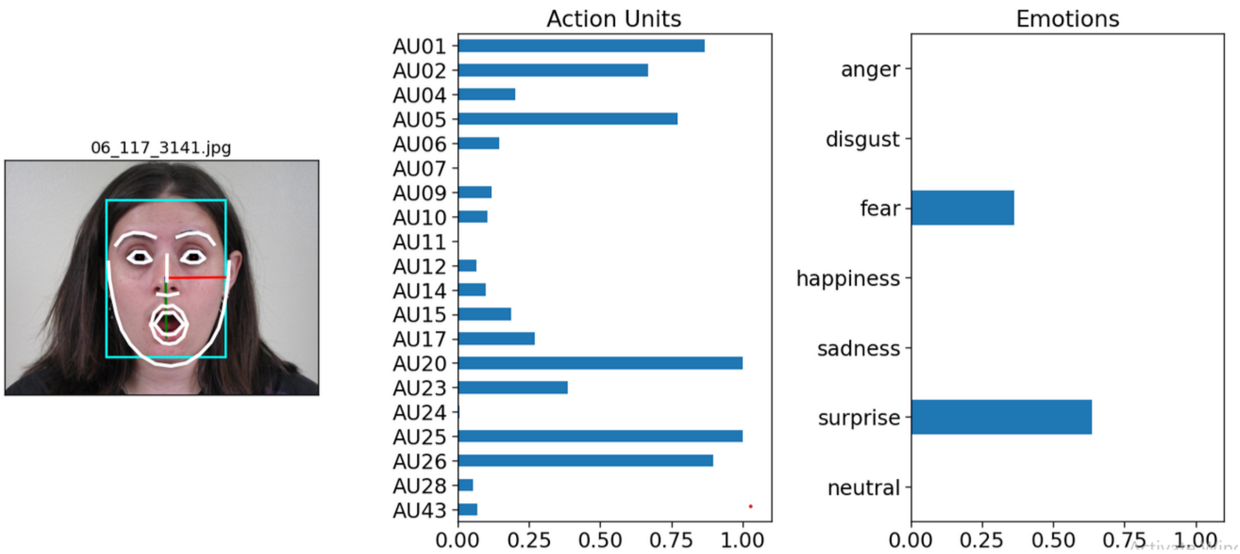

Fig.10.Sample Image from CFEE database

Following are the results obtained for the CFEE dataset .

Table 2.Results of Facebased emotion classification System

Sr.No.	Emotion class	Accuracy in %
1	Anger	80.00
2	Disgust	100.0
3	Happy	97.81
4	Neutral	95.91
5	Sadness	66.48

Decision level Fusion: As illustrated in Figure 6, the decision-level fusion entails a facial dataset comprising five distinct annotated classes of emotion, whereas our EEG dataset encompasses only three emotional categories: Neutral, Positive, and Negative. To mitigate the intricacies associated with decision-making, we have opted to utilize only the three corresponding classes from the EEG dataset. The ensuing rule-based decision-level fusion is executed thusly: if the true annotated label of an image is Neutral and the prediction rendered by the facial emotion recognition system also yields Neutral, the output of the Facial Emotion Recognition (FER) is deemed Neutral. Similarly, in the context of the EEG-based emotion recognition system, such as Support Vector Machine (SVM), if the actual label is Neutral and the predicted EEG label aligns as Neutral, then the conclusive emotional state is regarded as Neutral.following table shows the sample results of decision level fusion.

Fig.11. *Sample images of CFEE dataset [9]*

Table 3.: Sample Results of Decision Level fusion

Sr. No	Image ID of CFEE dataset	Annotated Label_Image	Predicated (FER)_Image	Row Index of EEG dataset	Annotated Label_image	Predicated(SVM)_image	Final Decision level Fusion by simple sum rule
1	2527	Neutral	Neutral	20	Neutral	Neutral	Neutral
2	3350	Happy	Happy	591	Positive	Positive	Positive
3	3756	Anger	Anger	2	Negative	Negative	Negative
4	2721	Sad	Fear	28	Neutral	Positive	Neutral
5	3245	Happy	Happy	57	Positive	Positive	Positive
6	3123	Neutral	Neutral	7	Neutral	Neutral	Neutral
7	3141	Surprise	Surprise	13	Positive	Positive	Positive
8	3370	Disgust	Disgust	22	Negative	Negative	Negative
9	2532	Sad	Sad	5	Negative	Negative	Negative
10	5850	Anger	Sad	21	Neutral	Neutral	Negative

From the above table it is evident that the dataset available for facial expressions and EEG signals must be from the same source and essentially have the same label too .

Feature level Fusion: As shown in the block diagram below for feature level fusion it is important to form a feature vector which will essentially have features from both the modalities.for the work presented facial features and then EEG features are concatenated ,using labels of EEG features as a annotation to from the combined feature vector

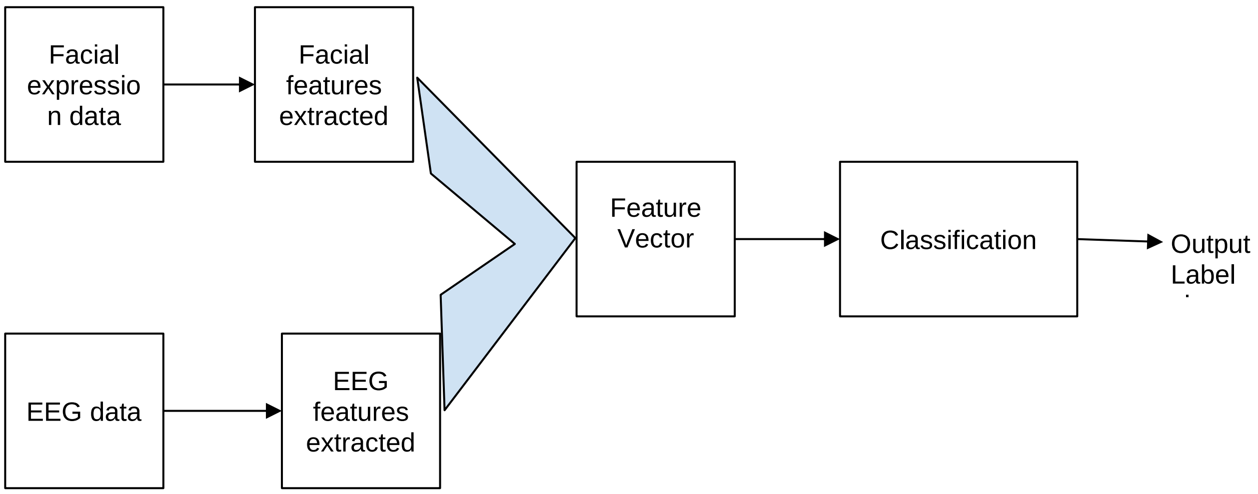

Fig.12. Block diagram for Feature level fusion process.

Following confusion matrix represents the results obtained on a combined feature vector.

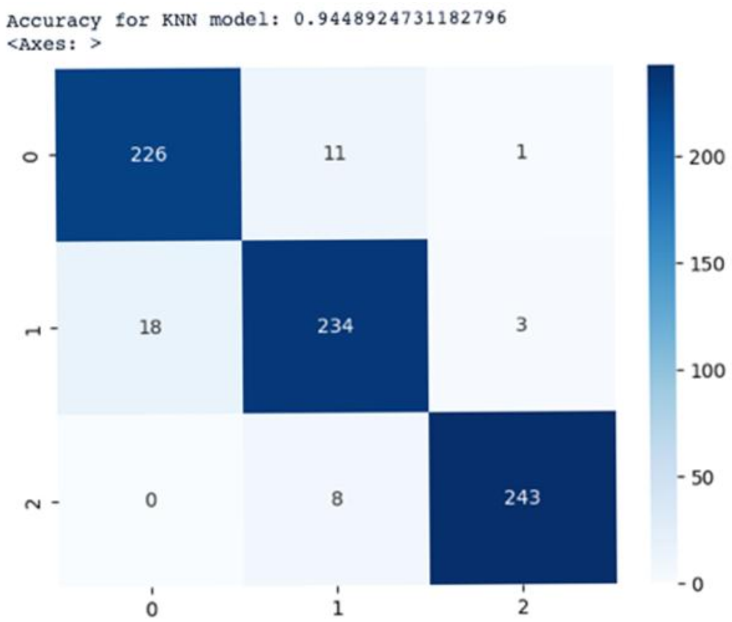

Fig.13.*Confusion matrix for KNN classification* of feature level fusion based emotion classification System.

Model accuracy score with default hyper parameters: 94.48% for KNN Classifier

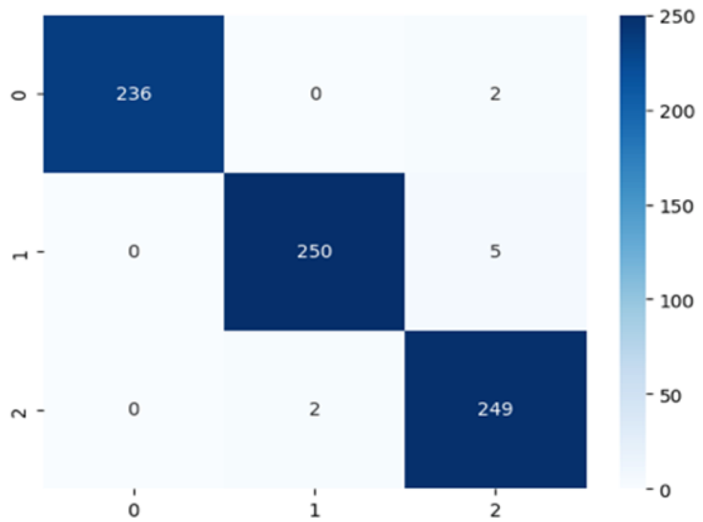

Fig.13.*Confusion matrix for SVM classification* of feature level fusion based emotion classification System.

Model accuracy score with default hyper parameters: 98.66% for SVM Classifier.

CONCLUSION

The experimental findings pertaining to individual modalities and the fusion strategies employed reveal that feature-level fusion yields commendable outcomes, achieving a peak classification accuracy of 98.66% with the support vector machine (SVM) classifier, while evading model overfitting. In contrast, the decision-level classification strategy attains a maximum accuracy of 80%. When juxtaposed with the fusion techniques implemented for a specific dataset across various experimental conditions, individual modalities reach a maximum of 90%. Although this indicates that the amalgamation of diverse modalities surpasses the efficacy of singular modalities, it remains contingent upon several factors, including the dataset utilized, extracted features, nature of said features, feature engineering, and classifiers with optimized hyperparameters, among others. It is also imperative to underscore that more robust fusion strategies are vital, and the incorporation of deep learning algorithms may yield promising results when utilized with a well-annotated, combined dataset of both the modalities.

REFERENCES

[1] Huang, J. Yang, P. Liao, and J. Pan, "Fusion of Facial Expressions and EEG for Multimodal Emotion Recognition," *Computational Intelligence and Neuroscience*, vol. 2017, pp. 1–8, 2017, doi: https://doi.org/10.1155/2017/2107451.

[2] M. Soleymani, Sadjad Asghari-Esfeden, Maja Pantic, and Y. Fu, "Continuous emotion detection using EEG signals and facial expressions," *International Conference on Multimedia and Expo*, Jul. 2014, doi: https://doi.org/10.1109/icme.2014.6890301.

[3] Y. Huang, J. Yang, S. Liu, and J. Pan, "Combining Facial Expressions and Electroencephalography to Enhance Emotion Recognition," *Future Internet*, vol. 11, no. 5, p. 105, May 2019, doi: https://doi.org/10.3390/fi11050105.

[4] Mr. A. S. Ufade, "Power Spectral Density Based Discrete Emotional State Recognition System Using Electroencephalography Signals," *African Journal of Biomedical Research*, pp. 59–66, Dec. 2024, doi: https://doi.org/10.53555/ajbr.v27i6s.5147.

[5] Y. Tan, Z. Sun, F. Duan, J. Solé-Casals, and C. F. Caiafa, "A multimodal emotion recognition method based on facial expressions and electroencephalography," *Biomedical Signal Processing and Control*, vol. 70, p. 103029, Sep. 2021, doi: https://doi.org/10.1016/j.bspc.2021.103029.

[6] J. Pan, W. Fang, Z. Zhang, B. Chen, Z. Zhang, and S. Wang, "Multimodal Emotion Recognition based on Facial Expressions, Speech, and EEG," pp. 1–8, Jan. 2023, doi: https://doi.org/10.1109/ojemb.2023.3240280.

[7] S. Wang, J. Qu, Y. Zhang, and Y. Zhang, "Multimodal Emotion Recognition From EEG Signals and Facial Expressions," vol. 11, pp. 33061–33068, Jan. 2023, doi: https://doi.org/10.1109/access.2023.3263670.

[8] Y. Alotaibi and Veera Ankalu. Vuyyuru, "Electroencephalogram based face emotion recognition using multimodal fusion and 1-D convolution neural network (ID-CNN) classifier," *AIMS mathematics*, vol. 8, no. 10, pp. 22984–23002, Jan. 2023, doi: https://doi.org/10.3934/math.20231169.

[9] Compound Facial Expressions of Emotion (CFEE) dataset. (n.d.). The Ohio State University, Center for Computational Bioengineering. Retrieved from https://cbcsl.ece.ohio-state.edu/compound.html.

[10] J. Ashford, J. J. Bird, F. Campelo, and D. R. Faria, "Classification of EEG Signals Based on Image Representation of Statistical Features," *Advances in Intelligent Systems and Computing*, pp. 449–460, Aug. 2019, doi: https://doi.org/10.1007/978-3-030-29933-0_37.

[11] Amit Konar and Aruna Chakraborty, *Emotion recognition : a pattern analysis approach*. Hoboken, New Jersey: John Wiley & Sons, Inc, 2015.

[12] P. Dutta and A. Barman, *Human Emotion Recognition from Face Images*. Springer Nature, 2020.

[13] J. A. Russell and A. Mehrabian, "Evidence for a three-factor theory of emotions", *J. Research in Personality*, vol. 11, no. 3, pp. 273-294, 1977

[14] Ekman P. and Friesen W. V., Facial action coding system: a technique for the measurement of facial movement, *Rivista di Psichiatria*. (1978) 47, no. 2, 126–138.

[15] N. Ahmed, Z. A. Aghbari, and S. Girija, "A systematic survey on multimodal emotion recognition using learning algorithms," *Intelligent Systems with Applications*, vol. 17, p. 200171, Feb. 2023, doi: https://doi.org/10.1016/j.iswa.2022.200171.

[16] Petrantonakis PC, Hadjileontiadis LJ. Emotion recognition from EEG using higher order crossings. IEEE Trans Inf Technol Biomed. 2010 Mar;14(2):186-97. doi: 10.1109/TITB.2009.2034649. Epub 2009 Oct 23. PMID: 19858033.

[17] Wagh, K. P., & Vasanth, K. (2019). Electroencephalograph (EEG) Based EmotionRecognition System: A Review. In Lecture Notes in Networks and Systems (Vol. 33, pp.37–59). Springer Singapore. https://doi.org/10.1007/978-981-10-8204-7_5

[18] Matlovic, T., Gaspar, P., Moro, R., Simko, J., & Bielikova, M. (2016). Emotions detection using facial expressions recognition and EEG. Proceedings - 11th International Workshop on Semantic and Social Media Adaptation and Personalization, SMAP 2016,18–23. https://doi.org/10.1109/SMAP.2016.7753378

[19] Y. Lu, H. Zhang, L. Shi, F. Yang, and J. Li, "Expression-EEG Bimodal Fusion Emotion Recognition Method Based on Deep Learning," *Computational and Mathematical Methods in Medicine*, vol. 2021, pp. 1–10, May 2021, doi: https://doi.org/10.1155/2021/9940148.

[20] S. Wang, J. Qu, Y. Zhang, and Y. Zhang, "Multimodal Emotion Recognition From EEG Signals and Facial Expressions," vol. 11, pp. 33061–33068, Jan. 2023, doi: https://doi.org/10.1109/access.2023.3263670.

[21] D. Li *et al.*, “The Fusion of Electroencephalography and Facial Expression for Continuous Emotion Recognition,” *IEEE Access*, vol. 7, pp. 155724–155736, 2019, doi: https://doi.org/10.1109/access.2019.2949707.

[22] N. Thammasan, K. Fukui, and M. Numao, “Multimodal Fusion of EEG and Musical Features in Music-Emotion Recognition,” *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 31, no. 1, Feb. 2017, doi: https://doi.org/10.1609/aaai.v31i1.11112.

[23] H. Huang *et al.*, “An EEG-Based Brain Computer Interface for Emotion Recognition and Its Application in Patients with Disorder of Consciousness,” *IEEE Transactions on Affective Computing*, vol. 12, no. 4, pp. 832–842, Oct. 2021, doi: https://doi.org/10.1109/TAFFC.2019.2901456.

[24] H. Jiang, R. Jiao, D. Wu, and W. Wu, “Emotion Analysis: Bimodal Fusion of Facial Expressions and EEG,” *Computers, Materials & Continua*, vol. 68, no. 2, pp. 2315–2327, 2021, doi: https://doi.org/10.32604/cmc.2021.016832.

[25] N. Lubis *et al.*, “Construction of Japanese Audio-Visual Emotion Database and Its Application in Emotion Recognition,” *Language Resources and Evaluation*, pp. 2180–2184, May 2016.

[26] Silva and Pei Chi Ng, “Bimodal emotion recognition,” Nov. 2002, doi: https://doi.org/10.1109/afgr.2000.840655.

[27] M. Wozniak, M. Sakowicz, Kacper Ledwosinski, J. Rzepkowski, P. Czapla, and Szymon Zaporowski, “Bimodal Emotion Recognition Based on Vocal and Facial Features,” *Procedia Computer Science*, vol. 225, pp. 2556–2566, Jan. 2023, doi: https://doi.org/10.1016/j.procs.2023.10.247.