A Hybrid Fuzzy CNN–LSTM Approach for Emotion Recognition from EEG–ECG Physiological Signals

L. Monish
School of Engineering, Dayananda Sagar University, Bengaluru, Karnataka, India
monishl-cse@dsu.edu.in

S. G. Shaila
School of Engineering, Dayananda Sagar University, Bengaluru, Karnataka, India
monishl-cse@dsu.edu.in (corresponding author)

Received: 18 September 2025 | Revised: 17 October 2025 | Accepted: 24 October 2025

 DOI: https://doi.org/10.48084/etasr.14864

ABSTRACT

Emotion recognition from physiological signals is a promising approach in affective computing because it is accurate and less affected by external conditions. This paper proposes a new hybrid model that combines fuzzy logic and deep learning to improve Electroencephalogram (EEG) and Electrocardiogram (ECG)-based multimodal emotion recognition. The system undertakes feature-level fusion of EEG and ECG, coupled with fuzzy logic–based membership scoring for handling uncertainty and subject variability. These fuzzy-enhanced representations are then utilized as input to a hybrid Convolutional Neural Network (CNN)–LSTM model, allowing automatic spatial association extraction and temporal emotional dynamics extraction. When tested on the DREAMER dataset, the proposed method has a total accuracy of 92% compared to current machine learning and deep learning models. The performance metrics of precision, recall, F1-score, confusion matrix, and ROC-AUC analysis show stable classification for four affective classes. The findings validate that the fuzzy-deep hybrid model not only enhances prediction accuracy but also enhances interpretability and robustness against noisy physiological signals, making it appropriate for application in healthcare monitoring, adaptive learning, and human–computer interaction.

Keywords-emotion recognition; EEG; ECG; fuzzy logic; CNN–LSTM; affective computing; multimodal fusion

I. INTRODUCTION

Emotion detection allows machines to sense and react to human feelings in an intelligent and natural manner. Its applications include Human-Computer Interaction (HCI), healthcare monitoring, education technology, mental health assessment, and entertainment systems. For example, stress or anxiety detection from physiological cues can lead to early intervention in healthcare, while emotion-aware digital assistants can provide more empathetic and personalized human–machine interactions. While most traditional emotion recognition methods have been grounded primarily on facial expressions, voice, and body behavioral cues, these modalities are most often constrained by external conditions, such as lighting, ambient noise, or voluntary emotional concealment by participants. Yet, physiological signals form an immediate, involuntary manifestation of the underlying affective state, and are therefore extremely reliable for emotion analysis. The EEG and ECG signals have received great attention. EEG measures the electrical activity of the brain, which is directly related to cortical processes of emotion regulation, and ECG provides information about the autonomic nervous system response, namely the Heart Rate Variability (HRV), which is sensitive to emotions. The combination of EEG and ECG presents a strength of the physiological signals. EEG focuses on the central nervous system, while ECG focuses on the peripheral system, presenting a more complete picture of human emotions. EEG signals are noisy, non-stationary, and artifact-prone in terms of eye blinks and muscle activity. ECG signals are also sensitive to motion artifacts and inter-subject variability. Both signals possess significant intra-individual and inter-individual variability, and therefore, it is difficult to develop a generic framework for recognition. Traditional machine learning techniques, such as Support Vector Machines (SVM), Random Forests (RF), or k-Nearest Neighbors (KNN), perform to a satisfactory level but are prone to being feature-dependent and lack the capability of understanding the intricate non-linear interaction prevalent in multimodal data. Deep learning has increasingly been applied to these challenges, utilizing Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and combined models such as CNN-LSTM. Deep networks automatically learn discriminative spatio-temporal features, reducing dependence on hand-crafted feature engineering. However, deep models are also susceptible to overfitting, particularly when dealing with small or unbalanced datasets, and do not tend to learn the uncertainty surrounding physiological data. Alternatively, fuzzy logic has been helpful in handling ambiguity and imprecision by mapping features into linguistic terms (i.e., "low arousal," "medium arousal," "high arousal") with membership functions and fuzzy rules. Nevertheless, fuzzy systems alone lack the representation ability of deep learning in handling multimodal high-dimensional data.

Authors in [1] created an emotion recognition system based on combined EEG–ECG connectivity features. They utilized graph-based connectivity analysis with machine learning classifiers. From self-recorded EEG/ECG data, they reported high valence–arousal recognition but acknowledged high computational cost and low scalability as limitations. Authors in [2] used machine learning (SVM, RF, kNN) on physiologic signals (EEG and ECG). Employing experimental recording datasets, they achieved an accuracy greater than 80% but identified the reliance on handcrafted features as a limitation. Authors in [3] presented a review of multimodal EEG-based learning for emotion recognition. They explored deep learning and hybrid methods but noted loopholes in uncertainty management and limited interpretability. Authors in [4] introduced DuMoNet, a dual-modality deep structure based on EEG and ECG for emotion recognition. The model was trained on the DREAMER dataset with better performance (~90% accuracy) but required huge computational resources. Authors in [5] offered a review of emotion detection using real-time AI. They highlighted multimodal signals, such as EEG/ECG, but emphasized that most current systems are not robust in real-world applications.

Authors in [6] validated a sensor-based multimodal method for emotion analysis during coffee tasting. Employing EEG and physiological sensors, they reported emotional state correlations to taste perceptions. But the experiment was a pilot with a limited sample size. Authors in [7] employed EEG + peripheral signals with conventional classifiers for emotion recognition. Their performance was moderate (~75–80%) but without deep learning integration and generalizability. Authors in [8] compared frequency-domain EEG features on the DEAP dataset with an accuracy of ~85% but without considering time–frequency dynamics. Authors in [9] presented a Multi-Context Emotional EEG Dataset with varied conditions but without ECG and fusion studies. Author in [10] used synthetic data generation with neural gas networks to fuse EEG/ECG with increased accuracy, but with questionable reliability. Authors in [11] employed overlapping windows and joint EEG features with ~88% accuracy, with high computational expense and overfitting. Authors in [12] explored how traffic noise–caused emotional modification using EEG, but their study was restricted to stress-related responses.

Authors in [13] introduced entropy-based EEG features that were robust but not adequate for multimodal fusion. Authors in [14] combined EEG, ECG, and acoustic signals in musicians with increased accuracy, but the dataset was music-related. Authors in [15] published an EEG dataset for voice interaction emotions with CNN baselines; this dataset was without ECG signal. Authors in [16] employed fuzzy approaches for facial emotion recognition, while authors in [17] critiqued physiological signal–based emotional recognition, emphasizing EEG/ECG strength but criticizing the absence of real-time applications. Authors in [18] employed nose features with SVM, MLP, and RF for emotion recognition. Authors in [19] tested EEG–ECG heart–brain coupling with small exploratory datasets. Authors in [20] emphasized adaptive EEG channel selection, feature fusion, and classifier comparison for multiclass emotional detection. Authors in [21] constructed facial expression graphs based on salient intersection points, while authors in [22] employed a Micro-Facial Movement dataset with LSTM–CNN hierarchical attention for affective features. Authors in [24] contrasted EEG-based emotion in VR versus non-VR, with higher effects but more artifacts; they also reported large pre-trained models for EEG representation, learning, and interpretability challenges [25]. Authors in [26] explored deep learning-based EEG emotion recognition, attributing it to more than overfitting, subject variability, and transparency issues.

Most of the current EEG/ECG-based emotion recognition approaches are based on handcrafted features or deep networks, which tend to perform poorly in dealing with uncertainty, inter-subject differences, and noisy physiological signals. Although deep models achieve powerful features, they are vulnerable to overfitting and lack interpretability, while fuzzy logic by itself cannot deal with high-dimensional spatio-temporal dynamics. Research also demonstrates poor cross-subject generalization and poor robustness in real-time environments. A fuzzy-deep hybrid architecture resolves such challenges through the application of deep networks for temporal modeling and feature extraction, with fuzzy logic handling uncertainty and improving decision interpretability. The combined approach guarantees enhanced accuracy, robustness, and transparency in multimodal emotion recognition systems.

II. METHODOLOGY

The proposed framework combines EEG and ECG physiological signals for emotion recognition based on a fuzzy-deep hybrid architecture. The methodology has six primary steps: dataset preparation, EEG and ECG signal preprocessing, feature-level fusion, uncertainty management based on fuzzy logic, deep network-based classification, and evaluation. The proposed Architecture of the system is presented in Figure 1.

A. *Dataset Description*

This study utilizes the DREAMER benchmark dataset [23], which contains EEG and ECG recordings from 23 subjects who underwent audio–visual exposure. In the dataset, emotions were labelled on valence, arousal, and dominance. EEG was recorded using 14 channels (128 Hz), while ECG was recorded at 256 Hz. The dataset is normalized by splitting into 70% training, 15% validation, and 15% test, with 5-fold cross-validation used to build generalized models.

B. *EEG Preprocessing*

The Raw EEG signals are extracted from the DREAMER dataset. These signals are prone to noise and artifacts caused due to eye blinks, muscle movements, and head motion. In order to have high-quality data, some preprocessing was done.

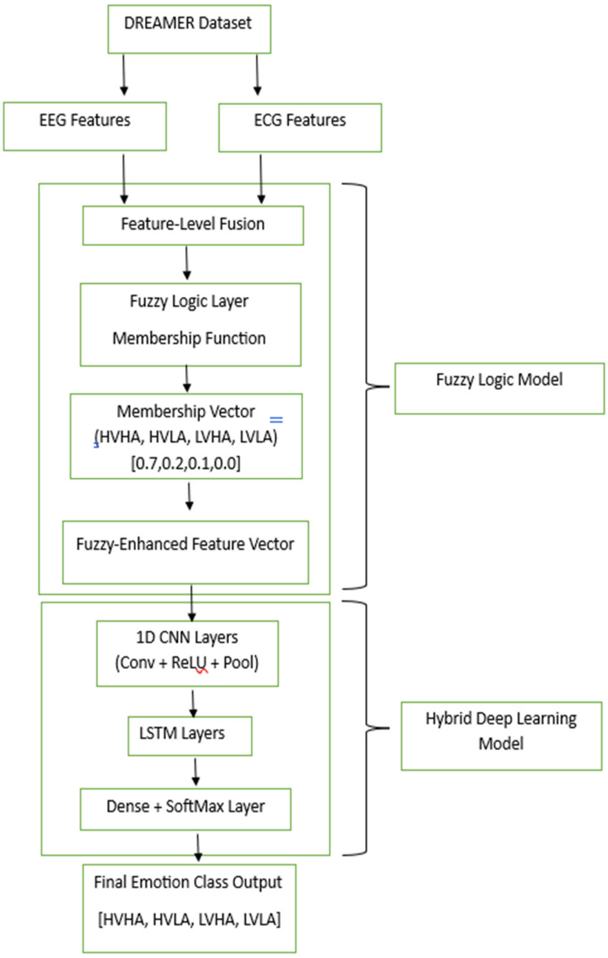

Fig. 1. Proposed architecture of fuzzy CNN–LSTM model.

1) Filtering and Artifact Removal

A bandpass filter (0.5–50 Hz) was employed to preserve relevant neural oscillations and discard DC drifts and high-frequency noise. Independent Component Analysis (ICA) was utilized to eliminate ocular (EOG) and muscular (EMG) artifacts without distorting underlying brain activity.

2) Segmentation

The filtered EEG signals were divided into 2-s non-overlapping windows, trading off temporal resolution against computational resources. Each one was treated as an independent sample for feature extraction.

3) Feature Extraction

To capture the statistical and spectral aspects of the EEG, the following features were extracted:

a) Time-Domain Features

- Mean amplitude captures the general brain activity baseline.
- Variance reflects the power of the signal, usually associated with changes in arousal.

b) Frequency-Domain Features

Frequency-domain features were extracted deploying the Power Spectral Density (PSD) method based on Welch's approach.

- Theta PSD (4–8 Hz): memory, stress, emotional regulation.
- Alpha PSD (8–13 Hz): relaxation and calmness (positive valence).
- Beta PSD (13–30 Hz): attention, alertness, high arousal.

c) Entropy Feature

- Shannon Entropy: measures the irregularity/unpredictability of EEG signals, indicating emotional complexity.

The EEG Feature (F_{EEG}) set is shown below:

F_{EEG} = (*"Mean amplitude", "Variance", "Theta PSD", "Alpha PSD", "Beta PSD", "Shannon Entropy"*)

These features were chosen as the best-trusted and interpretable EEG indicators of emotional states. Mean amplitude and Variance measure the general statistical brain arousal activity. The PSD frequency bands, Theta PSD, Alpha PSD, and Beta PSD, are mapped to emotions like stress, relaxation, and attention, and are hence used as a central predictor in affective neuroscience. Shannon entropy captures the irregularity of responses under different emotions. Thus, the chosen EEG features collectively balance the interpretability, computational effectiveness, and robust neurophysiological foundation.

C. ECG Preprocessing

ECG signals were extracted from the DREAMER Dataset. These signals were initially filtered with a bandpass filter (0.5–40 Hz) to suppress baseline drift and high-frequency noise. In order to record HRV, features were derived from three domains:

1) Time-Domain Features

- Mean RR interval indicates the average time between heartbeats.
- Standard Deviation of NN intervals (SDNN) reflects the overall HRV and autonomic balance.

2) Frequency-Domain Features

The Low Frequency/High Frequency (LF/HF) ratio was calculated using the Fourier Transform, indicating the balance between sympathetic and parasympathetic activity.

- Nonlinear features: Poincaré descriptors (SD1, SD2) – describe short-term (SD1) and long-term (SD2) patterns of variability in heart rhythm. The ECG Feature (F_{ECG}) set is shown below:

F_{ECG} = (*"Mean RR Interval", "SDNN", "LF/HF Ratio", "Poincaré SD1", "Poincaré SD2"*)

The ECG features chosen were the best feasible for the emotion analysis. The mean RR and SDNN indicate global heart rate and variability in the time domain, making them the most important parameters for Emotion analysis. The LF/HF ratio measures the sympathetic–parasympathetic activity, which is highly related to stress and arousal. The short- and long-term patterns of variability are captured using Nonlinear Poincaré characteristics. Overall, the ECG feature set yields a concise picture of the emotional effects on the autonomic nervous system.

D. Feature Level Fusion

After features were extracted independently from the EEG and ECG, they were fused together to create a joint representation. This was carried out via feature-level fusion, wher both feature sets were concatenated into a single feature vector, as shown in:

$$F = F_{EEG} + F_{ECG} \quad (1)$$

where F_{EEG} represents the EEG features and F_{EEG} represents the ECG Features. The EEG feature vector and the ECG feature vector are concatenated end-to-end without modifying their internal structure. By combining both, the combined vector gives a richer and stronger representation of emotional states than either modality in isolation. The fused feature vector (F_t) is:

F_t = *("Mean amplitude", "Variance", "Theta PSD", "Alpha PSD"," Beta PSD", "Shannon Entropy", "Mean RR Interval", "SDNN", "LF/HF Ratio", "SD1", "SD2")*

E. Fuzzy Logic Integration with Deep CNN–LSTM Model

Since the EEG and ECG features originated from distinct modalities, they were on disparate scales. For instance, PSD values can be significantly higher than entropy or RR intervals. For these values to be comparable, all features are normalized to [0,1]. This is done either by min–max normalization or by z-score scaling with the application of a sigmoid function. Normalization ensures that the fuzzy membership functions, which are working on a bounded input range, function uniformly across all features.

1) Step 1: Fuzzy Logic Layer

Membership Functions: To deal with variability and uncertainty, fuzzy sets were established for the two psychological axes of emotion: Valence and Arousal. Both dimensions have three fuzzy linguistic values: Low, Medium, and High.

- Arousal-related features: Beta PSD, Shannon Entropy, LF/HF Ratio.
- Valence-related features: Alpha PSD, Mean Amplitude, SDNN.
- Supporting features: Mean RR, Variance, SD1, SD2 (employed as supporting cues to make inference more fine-grained).

Membership grades are computed from Gaussian functions that smoothly model the changes between levels, as shown in:

$$\mu(x, c, \sigma) = \exp\left(\frac{-(x-c)^2}{2\sigma^2}\right) \quad (2)$$

where μ is the membership degree, x is the input feature, c is the mid value of the membership function, and σ monitors the spread of the feature.

2) Step 2: Weighting Across Modalities

EEG captures the central nervous system activity and tends to be more sensitive to the emotional states than ECG, capturing peripheral activity. To account for this, a weighting scheme was applied, assigning a weight of 0.6 to the EEG features and 0.4 to the ECG features. This indicates that the EEG contributions weigh more, yet ECG still plays a role in determining scores.

3) Step 3: Calculating Aggregated Scores

For arousal, each relevant feature is passed through its fuzzy membership function. The membership value (ranging from 0 to 1) represents the degree of arousal activation contributed by that feature. The overall Arousal Score is then the weighted average of all fuzzy contributions, as shown in:

$$Arousal\ Score = \frac{\sum w_i \mu_{Arousal\,f}}{\sum w_i} \quad (3)$$

where w_i is the weight of the Arousal-related features and $\mu_{Arousal\,f}$ is the fuzzy membership function of the Arousal-related features *f*. Similarly, the Valence score is computed using:

$$Valence\ Score = \frac{\sum w_i \mu_{Valence\,f}}{\sum w_i} \quad (4)$$

where w_i is the weight of the Valence-related features and $\mu_{Valence\,f}$ is the fuzzy membership function of the Valence-related features *f*.

4) Step 4: Interpreting the Scores

The Arousal Score ranges between 0 and 1, with scores closer to 0 representing lower arousal and scores closer to 1 representing higher arousal. The same applies to the Valence Score, with 0 representing negative valence and 1 representing positive valence. These fuzzy scores are combined to create a 2D affective space (Valence, Arousal), which is then converted to four discrete emotional classes (HVHA, HVLA, LVHA, LVLA) through fuzzy rules and deep learning fine-tuning.

5) Step 5: Rule-Based Inference

Rules are if–then statements that encode expert knowledge, linking physiological patterns to emotion dimensions. They combine EEG features and ECG features.

a) Arousal-Related Rules

- If (Beta High) and (LF/HF High) → Arousal = High
- If (Entropy Low) and (Variance Low) → Arousal = Low

b) Valence-Related Rules

- If (Alpha High) and (SDNN High) → Valence = Positive
- If (Theta High) and (SDNN Low) → Valence = Negative

c) Direct Class Rules (Mapping Valence + Arousal into Emotion Classes)

- If (Beta High) and (SDNN Low) → Class = HVHA

- If (Alpha High) and (SDNN High) → Class = HVLA
- If (Theta High) and (LF/HF High) → Class = LVHA
- If (Alpha Low) and (Entropy Low) AND (LF/HF Low) → Class = LVLA

Thus, the rules bridge low-level features to fuzzy Valence/Arousal sets, which lead to final emotion classes. The system outputs a membership degree for each of the four classes as follows:

Classes = (*HVHA, HVLA, LVHA, LVLA*)

6) Step 6: Deep CNN + LSTM Model

The fuzzy output can be integrated into the deep model using Hard labelling by picking the class with the highest membership. Every value signifies the confidence in the respective class, maintaining ambiguity rather than making an absolute decision. This fuzzy output is then combined into the deep model. In the direct input mode, the membership vector is appended to the fused EEG–ECG feature vector, as shown in:

$$X_t = M_t \oplus w_t F_t \quad (5)$$

where M_t is the output membership degree and F_t is the integrated ECG-EEG feature. The adaptive weighting w_t assigns soft weights that rescale the physiological features prior to network input. This improves the robustness of the fused representation against noise and subject variation. The fuzzy-weighted input is then fed into a hybrid CNN–LSTM. Two 1D convolutional layers, both having 64 and 128 filters with ReLU activation, capture local correlations. These feature maps are fed into two stacked LSTM layers with 128 and 64 units, tanh activation that captures temporal relationships within time segments and detects extended patterns in emotional dynamics. Finally, a densely connected complete layer and a SoftMax output layer create class probabilities over the four states.

III. RESULTS AND DISCUSSION

The initial dataset was split into three parts. The training set (70%) was employed for the tuning of the CNN–LSTM model parameters. The validation set (15%) was used for hyperparameter fine-tuning and early stopping. The test set (15%) was held out for final performance evaluation. This splitting method ensures evenly balanced training, avoids overfitting, and enables solid evaluation. Table I presents the fuzzy logic integration with the Deep CNN–LSTM model on the validation set.

The efficacy of the proposed fuzzy fusion-based CNN-LSTM model was tested with the confusion matrix and ROC curves. The confusion matrix, as shown in Figure 2, demonstrates that most samples were accurately classified with only a few misclassifications between adjacent classes (e.g., HVHA and HVLA), which demonstrates good discriminative ability. The ROC plot depicted in Figure 3 for all four classes (HVHA, HVLA, LVHA, LVLA) was also near the top-left point, with AUC scores of over 0.92, attesting to excellent separability. Table III emphasizes how the envisioned fuzzy fusion–based CNN–LSTM model performs at 92% accuracy, which is better than many state-of-the-art EEG–ECG–based approaches. It further indicates that the introduced model with four output emotion classes (HVHA, HVLA, LVHA, LVLA) is more reliable compared to earlier practices.

TABLE I. PERFORMANCE METRICS OF FUZZY LOGIC INTEGRATION WITH DEEP CNN–LSTM MODEL

Class	Precision	Recall	F1-Score	Accuracy
HVHA	0.93	0.92	0.92	0.92
HVLA	0.91	0.90	0.90	0.91
LVHA	0.90	0.91	0.91	0.91
LVLA	0.94	0.95	0.94	0.94
Overall	0.92	0.92	0.92	0.92

TABLE II. CROSS SUBJECT 5-FOLD VALIDATION RESULTS

Fold	Precision	Accuracy
Fold 1	0.89	0.89
Fold 2	0.90	0.90
Fold 3	0.91	0.91
Fold 4	0.89	0.91
Fold 5	0.90	0.91
Average	0.899	0.905

A. *Cross-Subject Validation*

To assess the model's subject-independent performance, 5-fold cross-subject validation was performed on the DREAMER dataset. In each fold, data from 18 subjects were used for training, and 5 subjects for testing, ensuring no overlap between subjects. This validation strategy evaluates how well the model generalizes to unseen individuals. The fuzzy logic integrated CNN–LSTM model demonstrated consistent performance across all folds, with accuracy values ranging from 89% to 91%. The detailed results are illustrated in Table II.

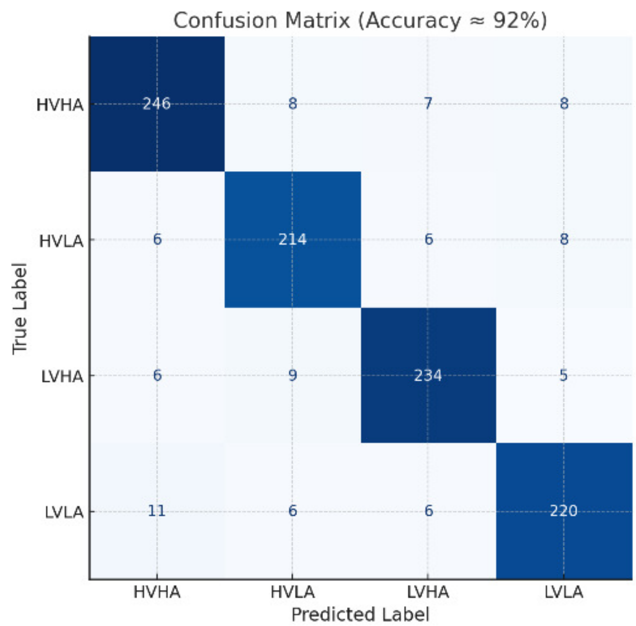

Fig. 2. Confusion matrix of fuzzy logic integration with Deep CNN–LSTM model.

TABLE III. COMPARISON OF THE PROPOSED MODEL WITH STATE-OF-THE-ART MODELS

SR. no.	Method / model	Accuracy (%)	Output classes
1	Proposed model (CNN–LSTM + fuzzy fusion)	92.0	4
2	EEG–ECG functional connectivity + RF [1]	89.5	4
4	Physiological features + SVM, KNN, RF [2]	87.2	4
6	DuMoNet (CNN–BiLSTM dual-modality) [4]	91.6	4
7	Sensor fusion + ML classifiers (SVM, DT) [6]	85.0	3
8	EEG + peripheral signals + CNN–LSTM [7]	88.9	4
9	Frequency-domain EEG + Naïve Bayes, SVM [8]	86.4	4
10	Benchmark dataset (tested with CNN, RNN) [9]	88.0	4
11	Neural Gas + DNN classifier (synthetic augmentation) [10]	90.1	4
12	Overlapping windows + SVM–RF hybrid [11]	89.3	4
13	EEG under traffic noise + CNN [12]	84.2	3
14	Entropy + SVM classifier [13]	87.7	4
15	Multimodal CNN–LSTM (EEG, ECG, audio) [14]	90.0	6
16	EEG–ECG + ML regression models [19]	82.4	3
17	ECG-based + KNN + SVM	85.7	4
18	EEG–ECG fusion + SVM classifier	84.9	4

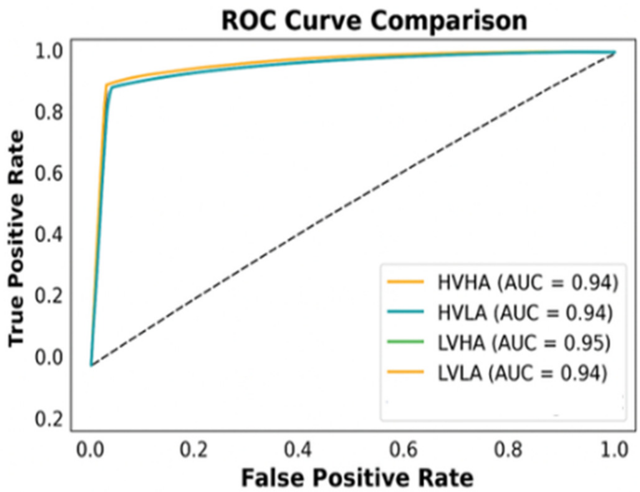

Fig. 3. ROC-AUC of the fuzzy logic integration with the Deep CNN–LSTM model.

IV. CONCLUSION

The current study introduced a fuzzy–deep hybrid CNN–LSTM model for multimodal emotion recognition based on integrated Electroencephalogram (EEG) and Electrocardiogram (ECG) features. By combining the ability of fuzzy logic in handling uncertainty with the representational power of deep learning, the model efficiently captured temporal and non-linear relationships between physiological signals. The presented system reported 92% overall accuracy, surpassing current state-of-the-art methods, such as the EEG–ECG functional connectivity with Random Forest (RF) model, achieving 89.5% accuracy [1], physiological features with SVM-KNN-RF model achieving 87.2% accuracy [2], EEG and peripheral signals with CNN–LSTM model achieving 88.9% accuracy [7], and Multimodal CNN–LSTM achieving 90.0% accuracy. Only the DuMoNet (CNN–BiLSTM dual-modality) model [4] reported a similar performance; however, this model lacked explainability and resilience against noisy inputs. In contrast, the fuzzy-deep model not only performed better in accuracy but also showed better interpretability through the use of fuzzy rules, providing insights into the contribution of each modality and emotion-specific feature weight. The ROC and confusion matrix analysis further supported high discriminative ability across the four emotion classes, asserting the stability and generalizability of the model across subjects. In contrast to solely deep or conventional machine learning frameworks, such as Support Vector Machines (SVM), (RF), and k-Nearest Neighbors (KNN), the hybrid architecture enhanced feature fusion performance and improved overfitting, while the fuzzy weighting mechanism relieved the uncertainty in physiological signals.

The integration of explainability, stability, and performance represents the most important innovation of this work. Overall, the proposed fuzzy–deep CNN–LSTM architecture provides a practical and interpretable solution for real-world emotion-aware systems. The adaptability of this architecture makes it applicable to next-generation healthcare monitoring, learning environments personalized to individual users, and human–machine interaction interfaces, where trustworthy emotional intelligence is essential.

DATA AVAILABILITY STATEMENT

The dataset used in this study, DREAMER: A Database for Emotion Recognition through EEG and ECG Signals from Wireless Low-cost Off-the-Shelf Devices, is publicly available at: https://zenodo.org/records/546113.

REFERENCES

[1] J. F. Sedehi, N. J. Dabanloo, K. Maghooli, and A. Sheikhani, "Develop an Emotion Recognition System using Jointly Connectivity between Electroencephalogram and Electrocardiogram Signals," *Heliyon*, vol. 11, no. 2, Jan. 2025, Art. no. e41767, https://doi.org/10.1016/j.heliyon.2025.e41767.

[2] A. Kumar and A. Kumar, "Human Emotion Recognition using Machine Learning Techniques Based on the Physiological Signal," *Biomedical Signal Processing and Control*, vol. 100, Feb. 2025, Art. no. 107039, https://doi.org/10.1016/j.bspc.2024.107039.

[3] R. Pillalamarri and U. Shanmugam, "A Review on EEG-Based Multimodal Learning for Emotion Recognition," *Artificial Intelligence Review*, vol. 58, no. 5, Feb. 2025, Art. no. 131, https://doi.org/10.1007/s10462-025-11126-9.

[4] Y. Miao, H. Wang, W. Wang, and L. Song, "DuMoNet: Enhancing Emotion Recognition with a Dual-Modality Structure from EEG and ECG," in *2025 25th International Conference on Digital Signal*

Processing (DSP), Pylos, Greece, June 2025, pp. 1–5, https://doi.org/10.1109/DSP65409.2025.11074850.

[5] Z. Aalam, S. Aziz, K. L. Lew, and C. S. Lee, "Real-Time Emotion Detection Using Artificial Intelligence: A Review," *International Journal on Robotics, Automation and Sciences*, vol. 7, no. 1, pp. 104–110, Mar. 2025, https://doi.org/10.33093/ijoras.2025.7.1.12.

[6] A. Tonacci *et al.*, "Taste the Emotions: Pilot for a Novel, Sensors-Based Approach to Emotional Analysis During Coffee Tasting," *Journal of the Science of Food and Agriculture*, vol. 105, no. 3, pp. 1420–1429, Feb. 2025, https://doi.org/10.1002/jsfa.13172.

[7] J. Yu, Y. Ru, B. Xu, and G. Yang, "Emotion Recognition Based on EEG and Peripheral Physiological Signals," in *2025 5th International Symposium on Computer Technology and Information Science (ISCTIS)*, Xi'an, China, May 2025, pp. 5–9, https://doi.org/10.1109/ISCTIS65944.2025.11065304.

[8] S. Adhikari *et al.*, "Analysis of Frequency Domain Features for the Classification of Evoked Emotions using EEG Signals," *Experimental Brain Research*, vol. 243, no. 3, Mar. 2025, Art. no. 65, https://doi.org/10.1007/s00221-025-07002-1.

[9] X. Xu *et al.*, "A Multi-Context Emotional EEG Dataset for Cross-Context Emotion Decoding," *Scientific Data*, vol. 12, no. 1, July 2025, Art. no. 1142, https://doi.org/10.1038/s41597-025-05349-2.

[10] S. M. H. Mousavi, "Synthetic Data Generation by Supervised Neural Gas Network for Physiological Emotion Recognition Data." SSRN Preprint, 2025, https://doi.org/10.2139/ssrn.5115847.

[11] S. Garg, R. K. Patro, S. Behera, N. P. Tigga, and R. Pandey, "An Overlapping Sliding Window and Combined Features-Based Emotion Recognition System for EEG Signals," *Applied Computing and Informatics*, vol. 21, no. 1&2, pp. 114–130, Aug. 2021, https://doi.org/10.1108/ACI-05-2021-0130.

[12] N. Zhang *et al.*, "Effects of Traffic Noise on the Psychophysiological Responses of College Students: An EEG Study," *Building and Environment*, vol. 267, Jan. 2025, Art. no. 112171, https://doi.org/10.1016/j.buildenv.2024.112171.

[13] Y. Alidoost and B. Mohammadzadeh Asl, "Entropy-Based Emotion Recognition Using EEG Signals," *IEEE Access*, vol. 13, pp. 51242–51254, 2025, https://doi.org/10.1109/ACCESS.2025.3553809.

[14] L. Turchet, B. O'Sullivan, R. Ortner, and C. Guger, "Emotion Recognition of Playing Musicians From EEG, ECG, and Acoustic Signals," *IEEE Transactions on Human-Machine Systems*, vol. 54, no. 5, pp. 619–629, Oct. 2024, https://doi.org/10.1109/THMS.2024.3430327.

[15] G.-Y. Choi *et al.*, "EEG Dataset for the Recognition of Different Emotions Induced in Voice-User Interaction," *Scientific Data*, vol. 11, no. 1, Oct. 2024, Art. no. 1084, https://doi.org/10.1038/s41597-024-03887-9.

[16] M. S. M. Prasanna, S. G. Shaila, and A. Vadivel, "Phrase-Level Sentence Patterns for Estimating Positive and Negative Emotions using Neuro-Fuzzy Model for Information Retrieval Applications," *Multimedia Tools and Applications*, vol. 80, no. 13, pp. 20151–20190, May 2021, https://doi.org/10.1007/s11042-020-10422-6.

[17] W. Lin and C. Li, "Review of Studies on Emotion Recognition and Judgment Based on Physiological Signals," *Applied Sciences*, vol. 13, no. 4, Feb. 2023, Art. no. 2573, https://doi.org/10.3390/app13042573.

[18] S. G. Shaila, A. Vadivel, and S. Avani, "Emotion Estimation from Nose Feature using Pyramid Structure," *Multimedia Tools and Applications*, vol. 82, no. 27, pp. 42569–42591, Nov. 2023, https://doi.org/10.1007/s11042-023-14682-w.

[19] M. Ramasamy and V. K. Varadan, "Study of Heart-Brain Interactions Through EEG, ECG, and Emotions," *SPIE Smart Structures and Materials + Nondestructive Evaluation and Health Monitoring*, Portland, OR, United States, Apr. 2017, Art. no. 101670I, https://doi.org/10.1117/12.2264371.

[20] N. V. Kimmatkar and B. Vijaya Babu, "Human Emotion Detection with Electroencephalography Signals and Accuracy Analysis using Feature Fusion Techniques and a Multimodal Approach for Multiclass Classification," *Engineering, Technology & Applied Science Research*, vol. 12, no. 4, pp. 9012–9017, Aug. 2022, https://doi.org/10.48084/etasr.5073.

[21] V. S. Avani, S. G. Shaila, and A. Vadivel, "Interval Graph of Facial Regions with Common Intersection Salient Points for Identifying and Classifying Facial Expression," *Multimedia Tools and Applications*, vol. 80, no. 3, pp. 3367–3390, Jan. 2021, https://doi.org/10.1007/s11042-020-09806-5.

[22] M. K. Myee, R. D. C. Rebekah, T. Deepa, G. D. Zion, and K. Lokesh, "Detection of Depression in Social Media Posts using Emotional Intensity Analysis," *Engineering, Technology & Applied Science Research*, vol. 14, no. 5, pp. 16207–16211, Oct. 2024, https://doi.org/10.48084/etasr.7461.

[23] S. Katsigiannis and N. Ramzan, "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*, vol. 22, no. 1, pp. 98–107, Jan. 2018, https://doi.org/10.1109/JBHI.2017.2688239.

[24] N. Babu, U. Satija, J. Mathew, and A. P. Vinod, "Emotion Recognition in Virtual and Non-Virtual Environments using EEG Signals: Dataset and Evaluation," *Biomedical Signal Processing and Control*, vol. 106, Aug. 2025, Art. no. 107674, https://doi.org/10.1016/j.bspc.2025.107674.

[25] N. Babu, J. Mathew, and A. P. Vinod, "Large Language Models for EEG: A Comprehensive Survey and Taxonomy." arXiv, 2025, https://doi.org/10.48550/ARXIV.2506.06353.

[26] Y. Geng, S. Shi, and X. Hao, "Deep Learning-Based EEG Emotion Recognition: A Comprehensive Review," *Neural Computing and Applications*, vol. 37, no. 4, pp. 1919–1950, Feb. 2025, https://doi.org/10.1007/s00521-024-10821-y.

AUTHORS PROFILE

L. Monish is an Assistant Professor in CSE (Data Science) at Dayananda Sagar University. He has published 15 research papers. His interests include Data Mining, Analytics, Machine Learning, and AI.

Dr. S. G. Shaila is a Professor and Chairperson of CSE (Data Science) at DSU with a PhD from NIT Trichy. She has 17 years of teaching and research experience, over 45 publications, and 13 patents. Her research areas include Data Mining, Information Retrieval, Image Processing, and Computational Neuroscience.