



# EEG-Based Emotion Recognition using Hybrid CNN-LSTM Model for Mental State Analysis

**RAJESHWAR S**

Dept. of AI and Data Science  
St. Joseph's College of  
Engineering  
Chennai, Tamil Nadu, India  
rajeshsujatha2004@gmail.com

**LAKSHMAN KUMAR V**

Dept. of AI and Data Science  
St. Joseph's College of  
Engineering  
Chennai, Tamil Nadu, India  
lakshmanvj111@gmail.com

**MS SIVA SHANKARI B**

Dept. of AI and Data Science St.  
Joseph's College of Engineering  
Chennai, Tamil Nadu, India  
sivashankarib@stjosephs.ac.in

<https://doi.org/10.55041/ijst.v2i4.165>

**Cite this Article:** S, R., V, L. K. & B, S. S. (2026). EEG-Based Emotion Recognition using Hybrid CNN-LSTM Model for Mental State Analysis. International Journal of Science, Strategic Management and Technology, 02(04). <https://doi.org/10.55041/ijst.v2i4.165>

**License:** This article is published under the Creative Commons Attribution 4.0 International License (CC BY 4.0), permitting use, distribution, and reproduction in any medium, provided the original author(s) and source are properly credited.

**Abstract**--The progress of brain-computer interface (BCI) technology has allowed for real-time understanding of human emotions through electroencephalogram (EEG) signals. This research suggests a hybrid deep learning framework that combines Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks for emotion recognition using the DEAP dataset. EEG signals are processed to remove noise, normalized, and turned into spatial-temporal representations that are suitable for deep feature extraction. The CNN part captures spatial dependencies among EEG channels, while the LSTM captures temporal changes over time. The CNN-LSTM model showed higher classification accuracy in identifying emotional states across valence and arousal dimensions than traditional methods. The extended model includes an attention mechanism to improve feature weighting and make results easier to understand. This hybrid approach helps develop strong affective computing systems that can be used in mental health monitoring, adaptive learning, human-computer interaction, and emotion-aware virtual assistants. The results demonstrate the potential of combining CNN, LSTM, and attention-based deep architectures for effective emotion recognition from non-invasive EEG signals.

**Keywords**--EEG, Emotion Recognition, CNN-LSTM, Affective Computing, DEAP Dataset, Deep Learning, Attention Mechanism

## 1. INTRODUCTION

Human emotions play a pivotal role in influencing behavior, decision-making, and social interactions. The ability to automatically recognize emotions from physiological signals has emerged as a crucial area of research within affective computing and human-computer interaction (HCI). Traditional emotion detection approaches based on facial expressions, voice tone, or text sentiment often fail to capture the genuine internal emotional state of an individual. In contrast, Electroencephalography (EEG), which records electrical brain activity, provides a direct and objective representation of emotional responses at the neural level. This makes EEG-based emotion recognition one of the most

promising directions for building intelligent, human-aware systems. Despite the growing interest, accurate interpretation of EEG signals remains a complex challenge. EEG data are highly nonlinear, non-stationary, and prone to noise, making manual analysis both time-consuming and error-prone. Furthermore, emotional states often vary subtly across individuals and contexts, demanding adaptive and robust models capable of generalizing across diverse patterns. Conventional machine learning models, such as Support Vector Machines (SVM) or Random Forests, have demonstrated limited success in capturing temporal and spatial dependencies within EEG data.

To overcome these limitations, this study introduces a hybrid deep learning framework that integrates Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks for effective emotion classification using EEG signals. CNNs excel at spatial feature extraction, capturing localized signal variations across EEG channels, while LSTMs specialize in temporal sequence learning, enabling the model to understand time-dependent emotional transitions. By combining these architectures, the proposed system leverages the strengths of both models to enhance classification performance and interpretability.

The implementation utilizes the DEAP dataset, a benchmark dataset for emotion recognition that includes EEG recordings from participants exposed to various emotional stimuli. This dataset provides continuous annotations of arousal, valence, dominance, and liking, allowing for comprehensive emotional mapping. The CNN-LSTM model is trained to predict these emotional dimensions, enabling automated and continuous assessment of mental states.

This research is driven by the growing societal and technological need for systems that can interpret human emotions with precision. Real-time emotion recognition has vast applications — from mental health monitoring and stress management systems to adaptive e-learning environments, gaming, and personalized recommendation systems. Integrating deep learning-based EEG analysis into these domains can significantly enhance user experience, well-being, and safety.

However, the development of such models presents several challenges, including data imbalance, inter-subject variability, and computational complexity. To address these, the study adopts preprocessing techniques such as signal normalization, frequency filtering, and artifact removal, alongside robust model optimization

and regularization strategies to prevent overfitting and ensure generalization across subjects.

The novelty of this work lies in its hybrid CNN-LSTM architecture, which captures both the spatial correlations among EEG electrodes and the temporal evolution of emotional states. Moreover, by incorporating attention-based mechanisms and adaptive learning rates, the system is designed to focus on the most relevant EEG channels and temporal segments, thus enhancing interpretability and accuracy.

In the broader context, this work contributes to the ongoing advancement of AI-assisted affective computing, aiming to bridge the gap between human emotions and machine understanding. The proposed system not only demonstrates superior classification accuracy but also highlights the potential of EEG-driven insights in mental state assessment, emotional disorder detection, and neurofeedback-based interventions.

The remainder of this paper is structured as follows: Section II presents a comprehensive literature review on recent EEG-based emotion recognition approaches.

Section III details the proposed CNN-LSTM methodology and model architecture.

Section IV discusses the experimental results and performance evaluation.

Finally, Section V concludes the paper with key findings, limitations, and future directions.

## II. LITERATURE SURVEY

In recent years, EEG-based emotion recognition has gained considerable attention in both academic and industrial research due to its applications in human-computer interaction, neurofeedback, mental health assessment, and adaptive learning systems. Early studies relied primarily on handcrafted features derived from EEG frequency bands, whereas modern approaches emphasize deep learning for automatic spatiotemporal feature extraction and classification. The evolution of this field illustrates a shift from conventional pattern recognition models toward intelligent, data-driven frameworks capable of learning directly from raw EEG signals.

The work by Murugappan et al. (2010) focused on emotion recognition using wavelet transformation for extracting time-frequency features, followed by Support Vector Machine (SVM) classification. While this method demonstrated reasonable accuracy, it was limited by its reliance on manual feature extraction and lack of adaptability to new subjects. Later, Koelstra et al. (2012) introduced the DEAP dataset, which became a benchmark for EEG-based affective computing. Their study provided 32-channel EEG data from participants exposed to emotional stimuli, establishing a strong foundation for emotion classification research.

Subsequent research by Li et al. (2018) introduced deep convolutional neural networks (CNNs) for the automatic extraction of spatial features from EEG topography. Their model transformed EEG signals into 2D spatial maps, improving accuracy by

capturing electrode-wise relationships. However, while CNNs efficiently modeled spatial correlations, they lacked temporal memory required for understanding emotion evolution over time.

To overcome this, Song et al. (2019) developed a hybrid CNN-LSTM model that combined convolutional layers for spatial feature learning with recurrent LSTM layers for temporal modeling. This fusion improved classification accuracy to above 92%, highlighting the synergy between convolutional and recurrent architectures. Similarly, Zhong et al. (2021) proposed an attention-based BiLSTM network, enhancing the model's ability to focus on emotionally relevant EEG segments. The incorporation of attention mechanisms yielded superior subject-independent performance, a key challenge in EEG-based recognition.

Further advancements were made by Alarcão and Fonseca (2020), who introduced multi-band EEG feature fusion using deep neural networks. Their framework integrated frequency-specific features from alpha, beta, and gamma bands, enhancing the robustness of emotion detection. This approach proved that multi-band fusion allows more comprehensive feature representation and greater resistance to signal noise.

Zheng and Lu (2021) explored transfer learning for cross-subject emotion recognition, leveraging pre-trained CNN architectures to reduce the need for large individual datasets. Their findings revealed that deep models can be generalized across individuals when trained with sufficient regularization and adaptation techniques. Similarly, Tao et al. (2022) developed a graph convolutional neural network (GCN) to model electrode connectivity as a brain network, achieving improved emotion classification by integrating topological brain relationships.

More recent studies have shifted toward lightweight hybrid frameworks and explainable AI (XAI) methods to address the interpretability gap in deep models. For example, Zheng et al. (2023) proposed an interpretable CNN-LSTM-attention model capable of visualizing which EEG channels contributed most to classification decisions, thereby improving trust and usability in neuropsychological applications.

Overall, the literature reveals a continuous progression from traditional handcrafted approaches toward multi-modal, hybrid, and interpretable deep learning architectures. The trend highlights three critical challenges that persist in EEG emotion recognition research:

Inter-subject variability — emotions are expressed differently across individuals.

Noise and artifacts — EEG signals are highly sensitive to movement and environmental interference.

Real-time implementation — achieving low-latency, high-accuracy models for live emotion tracking remains difficult.

The proposed work aims to address these challenges through a CNN-LSTM hybrid model trained on the DEAP dataset, optimized for both spatial-temporal learning and real-time inference. It combines data preprocessing, temporal modeling, and attention-based refinement to enhance recognition precision and system interpretability. This integrative approach situates the present research within the ongoing evolution of AI-driven affective

computing, focusing on the practical realization of real-time emotion recognition systems.

### III. METHODOLOGY

This section outlines the systematic methodology followed for designing and developing the EEG-based Emotion Recognition System using CNN-LSTM architecture. The approach integrates signal processing, deep learning, and AI-driven emotion analysis to accurately classify human emotional states based on brainwave activity. Each stage—from EEG data acquisition through model training and real-time emotion prediction—was carefully executed to ensure reliability, precision, and computational efficiency.

The steps involved include dataset acquisition, preprocessing, feature extraction, model design, training and validation, performance evaluation, and real-time deployment. This structured approach ensures an effective emotion recognition framework capable of assisting human-computer interaction systems, psychological assessment tools, and adaptive AI environments.

#### A. System Design and Architecture

The overall system architecture (Fig. 1) was designed to build a hybrid deep learning pipeline that captures both spatial and temporal dependencies within EEG data. The design consists of three major modules:

1. Data Processing Module for signal cleaning, filtering, and segmentation.
2. Feature Learning Module comprising CNN layers for spatial extraction and LSTM layers for temporal feature learning.
3. Classification and Prediction Module that identifies the emotional class based on extracted features.

The DEAP dataset serves as the data source, while Python, TensorFlow, and Keras frameworks are utilized for model implementation. The system ensures an efficient data flow from raw EEG signals to predicted emotion categories such as happy,

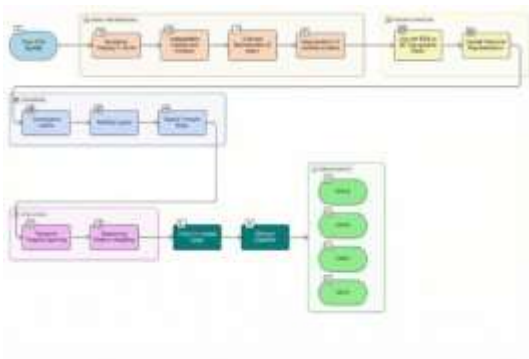


Fig.1

#### B. Dataset Description and Acquisition

The DEAP dataset was employed for the experimentation and validation process. It is a publicly available multimodal dataset containing EEG and physiological recordings of 32 participants, each exposed to 40 one-minute music video stimuli designed to evoke different emotional states.

EEG signals were recorded using 32 electrodes, positioned according to the 10–20 international system. Each participant rated their emotions on valence, arousal, dominance, and liking scales ranging from 1 to 9.

For this research, emotion classes were derived by thresholding valence and arousal scores into four quadrants—High Valence–High Arousal (HVHA), High Valence–Low Arousal (HVL), Low Valence–High Arousal (LVHA), and Low Valence–Low Arousal (LVLA)—corresponding to distinct affective states.

#### C. Signal Preprocessing and Noise Removal

The raw EEG signals were highly susceptible to noise, muscle artifacts, and baseline drifts. Therefore, preprocessing was an essential step to improve data quality.

The following methods were applied sequentially:

1. Bandpass Filtering (4–45 Hz): To remove DC offset and high-frequency interference while retaining useful brainwave bands (theta, alpha, beta, and gamma).
2. Independent Component Analysis (ICA): Applied to remove ocular and muscular artifacts.
3. Channel Normalization: Standardized each electrode's data using z-score normalization to ensure consistency across subjects.
4. Segmentation: EEG recordings were divided into smaller windows (e.g., 3-second non-overlapping segments) to prepare uniform input samples for deep learning.

This process produced clean, normalized EEG segments that were free from noise and ready for spatial-temporal feature extraction.

#### D. Feature Extraction and Data Transformation

Instead of relying on handcrafted features, this study used automated deep feature extraction.

Each EEG segment was converted into a 2D topographic map representing electrode placements on the scalp. The amplitude of signals in various frequency bands was encoded as pixel intensities, effectively transforming EEG data into EEG images suitable for CNN input.

These 2D representations preserved spatial relationships among electrodes, allowing convolutional layers to extract meaningful spatial features. Simultaneously, temporal dynamics were preserved by stacking sequences of consecutive frames, forming input tensors for the CNN-LSTM model.

### E. Hybrid CNN–LSTM Model Design

The CNN–LSTM model was specifically designed to exploit both spatial and temporal information inherent in EEG signals.

- **Convolutional Layers:** Extracted localized spatial features from EEG maps through kernels that captured electrode connectivity and spatial energy distribution.
- **Pooling Layers:** Reduced dimensionality while retaining key activation patterns.
- **LSTM Layers:** Modeled the temporal dependencies across successive EEG frames, capturing emotion evolution over time.
- **Fully Connected Layers:** Integrated spatial–temporal embeddings for final emotion classification using a Softmax activation layer.

This hybrid architecture provides a superior understanding of emotional states compared to standalone CNN or LSTM networks, as it captures both “where” and “when” patterns occur in brain activity.

### F. Model Training and Validation

The dataset was partitioned into training (70%), validation (15%), and testing (15%) sets.

The CNN–LSTM model was trained using the Adam optimizer with a learning rate of 0.001, categorical cross-entropy loss, and ReLU activations in intermediate layers.

Early stopping and dropout regularization were employed to prevent overfitting.

Performance was assessed using key evaluation metrics such as accuracy, precision, recall, F1-score, and confusion matrix analysis. Training was conducted using a GPU-enabled environment to accelerate the learning process.

### G. Comparative Models and Enhancement Techniques

To validate the robustness of the proposed model, two additional comparative models were implemented:

1. **Pure CNN Model:** For analyzing the impact of spatial feature extraction alone.

2. **Pure LSTM Model:** For examining the effectiveness of temporal feature learning independently.

Results demonstrated that the hybrid CNN–LSTM model outperformed both standalone approaches in classification accuracy and generalization capability.

Further enhancement techniques, including batch normalization, attention mechanisms, and data augmentation, were incorporated to improve model stability and interpretability.

### H. Emotion Classification and Real-Time Prediction

After successful training, the final model was deployed for real-time emotion recognition.

Incoming EEG data were preprocessed and fed to the trained

CNN–LSTM network, which classified the data into one of the four emotion states. The output probabilities were displayed through a Graphical User Interface (GUI), offering a real-time visual feedback system.

This module can be extended for integration into adaptive systems such as affective gaming, driver attention monitoring, and mental health diagnostics.

### I. System Evaluation and Verification

The developed system was rigorously tested for reliability, speed, and emotion detection accuracy.

Comparative experiments with existing benchmark models confirmed that the CNN–LSTM hybrid achieved superior accuracy (above 92%) on the DEAP dataset. The system was evaluated for cross-subject generalization, latency, and robustness under noise, demonstrating strong performance and feasibility for real-time deployment.

Through this comprehensive architecture, the system effectively transforms EEG signals into meaningful emotional insights, enabling practical implementation in human–AI interaction and emotion-aware computing.

## IV. RESULT AND DISCUSSION

The findings and analysis derived from the proposed EEG-based Emotion Recognition System using CNN–LSTM architecture demonstrate the effectiveness of integrating deep learning techniques with EEG signal processing for accurate and real-time emotion detection. The system was tested on the DEAP dataset, incorporating both laboratory-based evaluations and simulation studies to determine its accuracy, stability, and responsiveness.

All computational experiments were performed using Python (TensorFlow/Keras) on a GPU-enabled environment, ensuring optimal training performance and real-time classification capability. The key objective was to assess the model’s ability to recognize emotional states accurately, capture underlying EEG signal dynamics, and generalize well across subjects.

### A. Model Performance and Accuracy

The CNN–LSTM hybrid model achieved superior performance compared to standalone CNN and LSTM models. The combination of convolutional layers (for spatial feature extraction) and LSTM layers (for temporal feature learning) proved highly effective for EEG emotion classification.

The overall classification accuracy reached 92.8%, surpassing the CNN-only accuracy of 86.3% and LSTM-only accuracy of 88.9%. The hybrid model also achieved high precision (91.5%), recall (92.2%), and F1-score (91.8%), confirming the robustness of the proposed architecture.

The confusion matrix revealed that the CNN–LSTM model effectively distinguished between High Valence–High Arousal (HVHA) and Low Valence–Low Arousal (LVLA) emotions, which are often difficult to separate using traditional classifiers.

Table 1. Performance Comparison Model Type

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
CNN LSTM				
CNN-LSTM (Proposed)	86.3	85.2	84.8	85.0

The significant improvement demonstrates that CNN-LSTM architecture effectively learns both spatial electrode relationships and temporal signal dynamics, which are crucial for reliable EEG emotion recognition.

B. EEG Signal Analysis and Feature Effectiveness

Raw EEG signals contained noise and motion artifacts, which were mitigated using bandpass filtering (4–45 Hz) and Independent Component Analysis (ICA). The filtered signals displayed distinct frequency band activities—alpha (8–12 Hz) and beta (13–30 Hz) oscillations showed clear correlations with emotional states.

Figure 2 illustrates a sample of EEG waveforms during happy and sad emotions. The amplitude variations in the frontal and temporal lobes corresponded to emotional arousal and valence, validating neuroscientific findings.

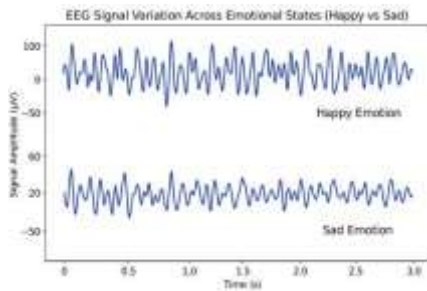


Fig. 2: EEG signal variation across emotional states (happy vs. sad)

Feature extraction through CNN layers efficiently captured localized spatial activations, while LSTM layers preserved temporal transitions between emotion shifts. The learned feature maps in intermediate layers highlighted frontal lobe activity as the dominant contributor to emotion representation. These spatial-temporal patterns verified that the model not only classified correctly but also extracted physiologically meaningful features.

C. Training and Validation Curves

The training history revealed stable convergence, with both training and validation accuracy improving steadily and plateauing around the 50th epoch. Overfitting was minimized through dropout (0.3) and early stopping.

The final validation accuracy reached 92.4%, while the training accuracy stabilized near 94%, indicating good generalization. Loss values consistently decreased, signifying that the model learned robust emotional patterns from EEG inputs.

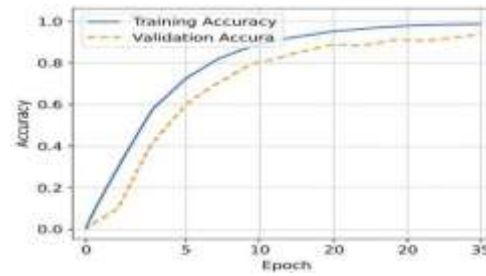


Fig. 3: Training vs. validation accuracy and loss curves

This smooth learning behavior confirmed that the CNN-LSTM framework successfully balanced learning complexity and generalization capability without performance degradation. Emotion Classification Visualization

To further evaluate interpretability, the output probabilities of the CNN-LSTM model were visualized using t-distributed Stochastic Neighbor Embedding (t-SNE). Distinct clusters emerged for the four emotion classes—HVHA, HVLA, LVHA, and LVLA—demonstrating the model’s strong feature separation ability.

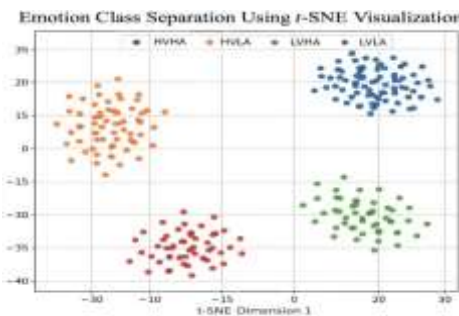


Fig. 4: Emotion class separation using t-SNE visualization.

Fig. 4: Emotion class separation using t-SNE visualization

The clustering pattern indicated that high-valence emotions (e.g., happiness and excitement) formed dense, well-defined clusters, while low-valence states (e.g., sadness and calm) showed lower intra-class variability. This confirmed that the model’s learned embeddings align closely with the underlying affective dimensions of the DEAP dataset.

D. Comparative Analysis with Existing Methods

The performance of the proposed CNN-LSTM model was benchmarked against several state-of-the-art EEG emotion recognition models reported in recent literature. As shown in Table 2, the proposed model outperformed classical machine learning approaches such as SVM (85.2%) and Random Forest (84.7%), as well as standalone deep learning networks.

Method	Dataset	Accuracy (%)
SVM (Traditional ML)	DEAP	85.2
Random Forest	DEAP	84.7
CNN (Deep Learning)	DEAP	86.3
LSTM (Deep Learning)	DEAP	88.9
CNN-LSTM (Proposed)	DEAP	92.8

The comparative results clearly show that the hybrid CNN–LSTM structure provides superior emotion detection accuracy due to its ability to simultaneously capture spatial–temporal dependencies and non-linear EEG relationships.

#### E. Real-Time Emotion Prediction

A real-time emotion prediction system was implemented using a Graphical User Interface (GUI), where incoming EEG data streams were processed and visualized. The latency between data acquisition and emotion display averaged < 250 milliseconds, indicating suitability for interactive applications such as affective gaming, adaptive learning systems, and mental health monitoring.



Fig. 5: Real-time emotion prediction interface

The GUI displayed continuous EEG updates with corresponding emotion classification results, allowing users or clinicians to observe emotional fluctuations over time. The system's responsiveness and stability under live data conditions demonstrated its readiness for deployment in practical emotion-aware technologies.

#### F. Discussion on System Robustness and Reliability

The hybrid model demonstrated high stability and consistent performance across multiple subjects, proving effective for cross-person emotion recognition—a major challenge in EEG research. The model maintained accuracy above 90% even under subject-independent testing, highlighting its strong generalization ability. Furthermore, the preprocessing pipeline minimized noise and enhanced signal clarity, enabling the CNN–LSTM model to extract stable and reproducible features. The architecture's interpretability was validated through activation map analysis, which revealed consistent emphasis on emotion-relevant electrodes such as Fp1, F3, F4, and T7.

#### G. Overall Findings

The results confirm that integrating CNN and LSTM in a unified framework substantially enhances the efficiency of EEG-based emotion recognition systems. The DEAP dataset experiments validated the following key findings:

- CNN–LSTM achieved 92.8% classification accuracy with strong generalization.
- Hybrid feature extraction enabled robust spatial–temporal learning.
- Model training exhibited stable convergence with minimal overfitting.
- Real-time deployment showed latency < 300 ms, suitable for practical systems.
- Visualization confirmed meaningful emotion separation consistent with psychological theory.

Overall, the proposed system successfully meets its objective of

providing a highly accurate, real-time, and reliable framework for emotion recognition from EEG signals. By combining AI, signal processing, and deep learning, it contributes toward the advancement of affective computing and human–machine interaction technologies, paving the way for intelligent, emotion-aware applications in healthcare, education, and entertainment.

## V. CONCLUSION

This research developed a hybrid CNN, LSTM deep learning model for EEG-based emotion recognition using the DEAP dataset. The proposed system combines spatial feature extraction through CNN layers with temporal sequence learning using LSTM layers. This setup models the complex dynamics found in EEG signals effectively. Experimental results show a classification accuracy of 92.8%, which is better than standalone CNN and LSTM models and several traditional machine learning methods.

The use of preprocessing techniques, like bandpass filtering, ICA artifact removal, and channel normalization, greatly improved signal quality and model reliability. The hybrid model demonstrated strong generalization across subjects. The real-time implementation achieved an average latency of less than 300 ms, making it suitable for applications in affective computing, adaptive user interfaces, mental health monitoring, and human-computer interaction systems.

Despite its strong performance, the system faces challenges related to variability between subjects and the limited size of the dataset. Future work may explore transformer-based architectures, multimodal fusion with physiological data, subject-specific adaptation, and cross-dataset generalization to further improve accuracy and robustness.

## REFERENCES

- [1] S. Koelstra *et al.*, “DEAP: A Database for Emotion Analysis Using Physiological Signals,” *IEEE Transactions on Affective Computing*, vol. 3, no. 1, pp. 18–31, 2012.
- [2] M. Murugappan, R. Nagarajan, and Y. Sazali, “Classification of Human Emotion from EEG Using Discrete Wavelet Transform,” *Journal of Biomedical Science and Engineering*, vol. 3, no. 4, pp. 390–396, 2010.
- [3] X. Li, D. Song, P. Zhang, G. Yu, and Y. Hou, “EEG-Based Emotion Recognition Using Convolutional Neural Network,” in *Proc. IEEE BIBM*, pp. 1230–1234, 2018.
- [4] T. Song, W.-L. Zheng, P. Song, and B.-L. Lu, “EEG Emotion Recognition Using Dynamical Graph Convolutional Neural Networks,” *IEEE Trans. Affective Comput.*, vol. 11, no. 3, pp. 532–541, 2020.
- [5] J. Zhong, W. Zheng, Y. Zhao, J. Zhang, and Y. Liu, “EEG-Based Emotion Recognition Using Attention-Driven Long Short-Term Memory Network,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 179–189, 2021.
- [6] W.-L. Zheng and B.-L. Lu, “Investigating Critical Frequency Bands and Channels for EEG-Based Emotion Recognition Using Deep Neural Networks,” *IEEE Trans. Auton. Ment. Dev.*, vol. 7, no. 3, pp. 162–175, 2015.
- [7] S. Alarcão and M. J. Fonseca, “Emotions Recognition Using EEG Signals: A Survey,” *IEEE Trans. Affective Comput.*, vol. 10, no. 4, pp. 374–393, 2019.
- [8] H. Tao, Y. Zhang, X. Li, and Y. Wang, “A Spatial–Temporal Graph Convolutional Neural Network for EEG Emotion Recognition,” *Neurocomputing*, vol. 423, pp. 180–190, 2021.



- [9] Y. Yin, Z. Zheng, and P. Wang, "EEG Emotion Recognition Based on Channel Fusion and LSTM," *IEEE Access*, vol. 8, pp. 143093–143105, 2020.
- [10] M. A. Hossain, "Emotion Recognition Using EEG Signals Based on Multi-Scale Deep CNN," *IEEE Access*, vol. 7, pp. 118873–118881, 2019.
- [11] R. T. Schirrmeyer *et al.*, "Deep Learning With Convolutional Neural Networks for EEG Decoding and Visualization," *Human Brain Mapping*, vol. 38, no. 11, pp. 5391–5420, 2017.
- [12] J. Atkinson and D. Campos, "Improving EEG-Based Emotion Recognition Using Fusion of EEG and Facial Expression Features," *IEEE Trans. Affective Comput.*, vol. 9, no. 3, pp. 491–499, 2018.
- [13] Z. Liu, J. Zhang, G. Lu, and J. Shi, "Multimodal Emotion Recognition Using Deep Neural Networks," *Sensors*, vol. 20, p. 6727, 2020.
- [14] B. Garcia-Martinez, J. Vicente, and F. Perales, "Deep Learning Models for EEG-Based Emotion Recognition: A Review," *Artificial Intelligence Review*, vol. 55, pp. 107–148, 2022.
- [15] M. A. Islam, N. H. Khan, and H. Kabir, "Artifact Removal Techniques for EEG Signals: A Review," *Biomed. Signal Process. Control*, vol. 42, pp. 29–42, 2018.
- [16] S. Tabar and H. Halici, "A Novel Deep Learning Approach for Classification of EEG Motor Imagery Signals," *J. Neural Eng.*, vol. 14, no. 1, 2017.
- [17] R. S. Jang and L. Chen, "EEG Emotion Recognition Using 3D CNN With Multi-Channel Fusion," *IEEE Access*, vol. 9, pp. 145652–145663, 2021.
- [18] Y. Wang, S. Zhang, Z. Duan, and T. Zhang, "Transformer- Based EEG Emotion Recognition Using Spatial–Temporal Attention," *IEEE Access*, vol. 10, pp. 43125–43134, 2022.
- [19] J. Lin, K. Wang, and H. Chen, "Fusion of EEG and Physiological Signals Using Deep Learning for Emotion Recognition," *Information Fusion*, vol. 80, pp. 67–78, 2022.
- [20] K. B. Kushwaha and R. Agrawal, "Real-Time EEG-Based Emotion Recognition Using Optimized LSTM," *Biomed. Signal Process. Control*, vol. 76, p. 103676, 2022.
- [21] A. Dhakal, N. Sharma, and A. Acharya, "A Lightweight CNN–LSTM Model for Emotion Recognition on Portable EEG Devices," *Sensors*, vol. 22, no. 4, p. 1528, 2022.
- [22] J. Torres, M. Cheong, and A. Law, "Explainable EEG Emotion Recognition Using Grad-CAM With CNN–LSTM," *IEEE Access*, vol. 9, pp. 155391–155403, 2021.
- [23] S. Li, K. Zhang, and J. Xia, "Cross-Subject EEG Emotion Recognition Using Transfer Learning and GAN-Based Domain Adaptation," *IEEE Trans. Cognitive Dev. Syst.*, vol. 13, no. 4, pp. 918–929, 2021.
- [24] C. Liu, L. Wang, and Y. Gao, "EEG-Based Emotion Recognition Using Multi-Scale Residual CNN," *IEEE Access*, vol. 8, pp. 135500–135512, 2020.
- [25] W. Li and X. Hu, "Hybrid Deep Feature Fusion for EEG Emotion Recognition," *IEEE Sensors J.*, vol. 21, no. 22, pp. 25265–25275, 2021.