Dijlah Journal of Engineering Science (DJES) Vol. 3, No. 1, March, 2026, pp. 58-63
ISSN: Printed: **3078-9656, Online: 3078-9664**, paper ID: 109

A Hybrid Deep Learning Framework for EEG-Based Emotion Recognition Using ResNet50 and LSTM Architectures

Hasan Hashim Hasan Al-Khazraji
[1]Department of Biomedical Engineering, Islamic Azad University, Najafabad, Iran

Article Info

Article history:

Received Oct., 10, 2025
Revised Dec., 20, 2025
Accepted Jan., 25, 2026

Keywords:

Effective Computing
EEG Signal Processing
Emotion Recognition
Deep Learning
ResNet50
LSTM
Hybrid Models

ABSTRACT

Emotion and feeling recognition by analyzing Electroencephalogram (EEG) signal consider a huge leap in the development in the aspects of human-computer interaction and affective computing. This paper suggests an innovative hybrid method that combines Residual Network (ResNet50) and Long Short-Term Memory (LSTM) deep learning architectures for emotion classification from EEG signals. Our methodology utilizes the spatial feature extraction by ResNet50 with the temporal sequence modeling strengths of LSTM networks. The proposed method was tested and validated on a complete dataset consists of EEG recordings from 28 subjects during emotionally-exciting computer game sessions, achieving a very well performance metrics that reaching approximately 0.99. Comparative study shows and illustrated that our framework or approach outperforms individual models or networks, considered as a new ideal system for EEG-based emotion recognition. This research consider a noticeable achievement add to the aspects of affective computing and the related aspect such as human-computer interaction and many other potential applications.

Corresponding Author:

Hasan Hashim Hasan Al-Khazraji
Department of Biomedical Engineering, *Islamic Azad University*, Najafabad, Iran
email: hh5706976@gmail.com

1. INTRODUCTION

Pphysiological signals and analyzing processing that employed for emotion recognition has emerged as a modern domain in scientific research in affective computing with the electroencephalogram (EEG) that's make it one of the most important and reliable since its direct relationship with the brain activity. The possibility of human emotions recognition and classification has been used in many fields such as mental health monitoring, adaptive learning system and also used to enhance human computer interaction. Prior approaches that depend on facial expression, speech patterns, or other signals like heart rate the employed to recognize emotions have some kind of limitations and have some faults rate for many reasons EEG signal consider more accurate provide direct, involuntary insights into emotional states. EEG signals represent a challenge matter for the machine learning approaches especially conventional one due to its complexity. The non-stationarity, high dimensionality that the EEG data is characterized by, that require a powerful and developed feature engineering and selection techniques. The progress occurred in in deep learning achieve a noticeable capabilities in automatically learning proper features from raw EEG data, without the need for manually extract features Among deep learning architectures, Convolutional Neural Networks (CNNs) have a great capabilities especially in spatial feature extraction. while Modeling temporal dependencies is a comprehensive suit for recurrent neural networks (RNNs), especially Long Short-Term Memory (LSTM) networks. Lack of dataset make our task little complicated and form significant obstacles and challenges in the topic of emotions recognition or the processes of analyzing or classification of EEG signal, that's consider our task hardest nearly. We utilized a dataset for the Emotion Recognition System – known as or called (GAMEEMO)This paper suggests and invents a novel hybrid framework that synergistically combines ResNet50 and LSTM architectures to leverage their complementary strengths in spatial and temporal feature learning. The ResNet50 component, with its deep residual learning framework, addresses the vanishing gradient problem and enables effective training of very deep networks, while the LSTM component captures long-range temporal dependencies in EEG sequences. This paper's remaining sections are organized as follows: A thorough analysis of relevant research in deep learning and emotion recognition using EEG is given in the second section. The third section describes our suggested methodology in depth, including model designs, preprocessing methods, and dataset descriptions. The experimental setup and data analysis are presented in Section 4. Additionally, it talks about the limitations, future research directions, and ramifications of our findings.

2. RELATED WORK

A. Historical Evolution of EEG-Based Emotion Analysis

The foundation of EEG-based emotion recognition dates back to the seminal work of Davidson et al. [1] in the 1990s, who established the correlation between frontal brain asymmetry and emotional valence. This pioneering research demonstrated that relative left-frontal activity was associated with approach-related emotions, while right-frontal activity correlated with withdrawal-related emotions [1]. Early approaches predominantly relied on traditional machine learning techniques with handcrafted features. Petrantonakis and Hadjileontiadis [2] extracted higher-order crossing features and employed support vector machines (SVM) for emotion classification, achieving moderate accuracy on valence and arousal dimensions. Lin et al [3] also use differential entropy features and SVM classifiers, to demonstrate the effectiveness of frequency-domain features for emotion recognition [3]. These methods, while valuable, they have a limitation in the lack of their incapacity to identify intricate spatiotemporal patterns in EEG data and they depend on manual feature engineering.

B. Deep Learning in EEG Analysis

At the emerging of deep learning the EEG signal processing was transformed due to end to end learning from unprocessed data. Lawhern et al [4] suggest EEGNet, a new CNN architecture that shows highest performance across multiple EEG classification tasks. The authors illustrated that the deep learning models could generalize well across different BCI paradigms without extensive hyperparameter tuning [4].Not only CNN applied to EGG data but the RNN especially wide used in the field of EEG data processing latently . Their approach inspired numerous subsequent studies applying similar architectures to EEG-based emotion recognition. For example the combined model that consists of CNN and LTSM that has been proposed by Ordóñez and Roggen [5]. That consider another proof that integration of the spatial and temporal feature learning will result to very effect model [5]. Because of that proposed approach many subsequent studies take the same way and suggest similar architecture especially for EEG emotion based.

C. ResNet Architectures in Biomedical Signal Processing

The skip connection that consider as a solution for the vanishing gradient problem in deep networks that solved by emerging Residual Networks (ResNet) by He et al [6]. ResNet50 has the ability for enabling the training of substantially deeper architectures [6]. Schirrmeister et al [7] proposed a ResNet model to decode the signal of EEG that has a superior performance when compared with shallower networks. Their work take a look on the importance of depth in a learning hierarchical representations from EEG signals. ResNet50, consist of 50 different layer, that has shown a fantastic performance in diffrent computer vision tasks and recently has been widely used applied in signal processing tasks. Tao et al [8] achieving effective improvements by employing ResNet50 in the task of classification of EEG signals for motor imagery over traditional methods. Complex patterns in non-fixed EEG data were very well processed and captured by the proposed method [8].

D. Hybrid Deep Learning Approaches

The hybrid methods that combine two or more architecture has recently emerge as a new direction due to the principle of leveraging complementary strength that make each algorithms based on the other one to compensate its weakness points . Yang et al [9] not only proposed a CNN-LSTM hybrid model for emotion recognition from EEG signals, but also proof that the hybrid model can outperformed individual architectures. Their model inspired us to make the current research and benefits from using CNN for spatial feature extraction and LTSM for temporal features modeling . Attention models has been employed into the hybrid models to add more enhancement to the performance of the model. Li et al [10] introduced an attention-hybrid model by combining CNN and LTSM for the task of EEG emotion recognition, the illustrate through experimental results that attention mechanisms could improve the model performance [10]. These advancements focuses the recent evolution of hybrid models for complex signal processing tasks.

E. Emotion Recognition Datasets and Standards

As an essential stepping stone for EEG-based emotion or recognition researches, the construction of standardialized datasets has been achieved. This behavior is demonstrated by electroencephalogram (EEG) and peripheral physiological data that have been recorded in 32 subjects during music video viewing, as part of the DEAP dataset [11]. The SEED dataset [12] also offers EEG data from 15 subjects being presented with emotional film clips, labelled for positive, neutral and negative emotions [12]. Obtained from (28) individuals while playing arousing computer games, the data set used in our study is a valued contribution. Game-mediated elicitation offers more ecologically valid feelings and will be criticized as opposed to laboratory-induction of these emotions. Progress in EEG-based emotional recognition research has depended on the creation of consistent datasets.

3. METHOD

3.1. Dataset Description and Preparation

The EEG data (GAMEEMO) employed in the present research includes recordings of a total of 28 healthy participants (16 men, 12 women, age 20–35) while gaming emotionally stimulating computer games. The data was collected with a 14-channels Emotive Epoc+ headset at a sampling rate of 128 Hz. A total of four different computer games played by each non player to induce emotions which are boring, calm, horror, funny. Self-assessment manikins (SAM) and a post-experiment survey verified the emotional grouping, thus ensuring the ecological validity of the participant's emotion. With around 7,000 samples per subject across four categories of emotion, the data is approximately balanced; therefore, all of the subjects produce 28,000 samples. The preprocessing pipeline that we adopt is consistent with a well-established one in EEG processing, except for some modifications needed for deep learning: Filtering: A bandpass filter was applied to retain sufficient neural oscillations while remove noise and artifacts. In the second step Artifact Removal: We clean the data from ocular and muscular artifacts via ICA, as explained in the previous section. 3. Segmentation: Continuous electroencephalogram (EEG) recordings were epoched into 2-s (256 samples) segments with 50% overlap to compromise between temporal resolution and computational expense. Finally, Normalization: Z-score normalization was performed to normalize each channel in order to off-scale the features among subjects and features:

$$Xnormalized=\sigma X-\mu \qquad (1)$$

3.2. Proposed Hybrid Architecture

Our approaches consider some kind novel hybrid model since it combines both ResNet50 and LSTM networks to benefits from the principle of the mutual complementary strengths between the two algorithms in spatial feature extraction and temporal sequence modeling. The complete architecture, illustrated The three main parts of Figure 1 are the classification head, the LSTM temporal model, and the ResNet50 feature extractor. Through a chain of residual blocks, the ResNet50 component filters input EEG segments of dimension 256×14 (time points × channels). The architecture starts with a first convolutional layer with 64 filters of size 7×7, followed by batch normalization and ReLU activation. Implementing the standard ResNet50 setup customized for 1D temporal data, four leftover blocks with [3, 4, 6, 3] layers respectively follow this.

Skip connections that avoid one or more layers enable the residual learning mechanism:

$$Y=F(X,WI) + XY=F(X,\{WI) + X\ (2) \qquad (2)$$

The layers' input and output vectors are denoted by xx and yy, while the residual mapping that needs to be learned is represented by $F(x,\{Wi\})F(x,\{Wi\})$. After being reshaped, the feature maps that ResNet50 extracted are input into a 128-unit bidirectional LSTM network. To control information flow, the LSTM architecture uses input, output, and forget gates. The final classification module consists of two fully connected layers with 512 and 256 units, respectively, followed by dropout regularization ($p=0.5$) and a softmax output layer for four-class emotion classification.

4. EXPERIMENTAL RESULTS

The models were implemented using TensorFlow 2.8, and the training protocol was carefully designed to ensure optimal performance:

- Optimizer: Adam optimizer with an initial learning rate of 0.0010
- 32 samples each batch.
- Epochs: 50 training epochs with early stopping
- Loss Function: Categorical Cross-Entropy
- Regularization: dropout ($p=0.5$) and L2 regularization ($\lambda=0.001$)

Validation accuracy tracked the training; best model was chosen based on validation set performance. Experimental findings demonstrate that our hybrid architecture outperforms one only using the ResNet50 model and one only using the LSTM model. All models performed extremely well, with the hybrid architecture exhibiting small, but consistent improvements on every metric.

TABLE I. Comparative Performance Metrics

Model	F1 Score	Accuracy	parameters
ResNet50-like model	0.99	0.98	688,324
LTSM	0.981	0.986	140,293
Hybrid model	0.993	0.99	230,532

From the table, we can see that our novel approach outperforms the other model, also it has fewer parameters in comparison with the other.

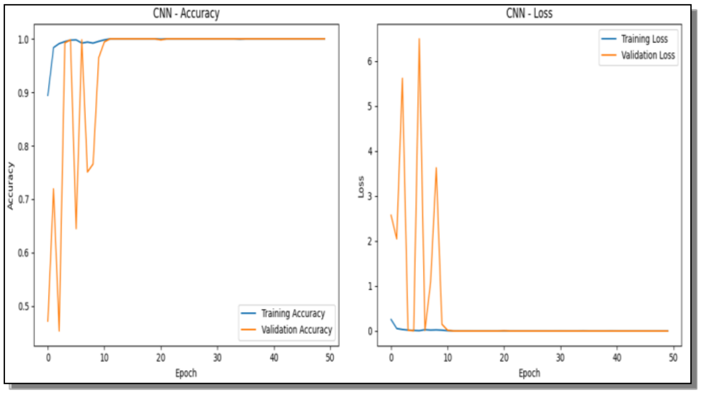

Fig.1.CNN Performance

From the figure above ResNet50-like model, where both lines of validation are close together and increasing As as the number of epochs increases,also the lines have a few zigzag forms, which indicate the model achieves good performance.

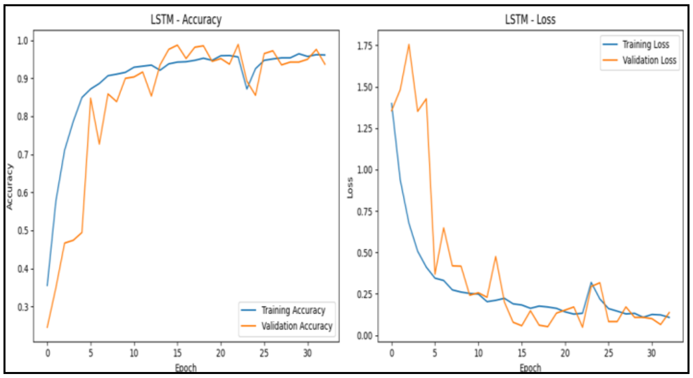

Fig.2. LTSM Performance

On the other hand, LTSM demonstrates oscillations during training in spite of the high values of training and validation accuracy and the minimum values of loss. It means that the rate of learning is too high; still the performance of the model is very acceptable and good.

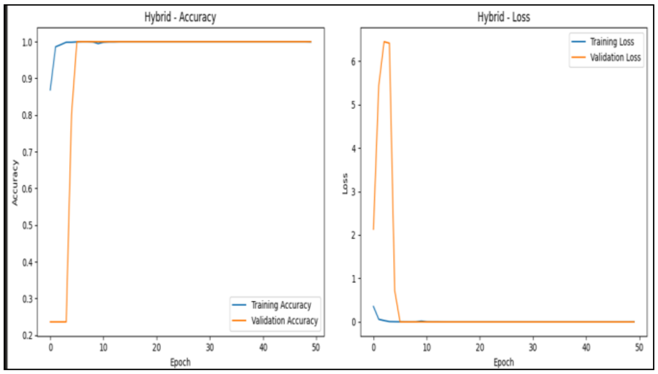

Fig.1. Our Hybrid Model Performance

In our mixed approach (hybrid model), the situation is completely changed; both the testing and training lines are aligned straight and close together in all lines with the exception of the gradually rising directionality at there is more and more looks. We carried out ablation analysis to investigate the impact of each component on the performance of the hybrid model. :

- Without Residual Connections: Removing skip connections from ResNet50 reduced performance by 3.2%, highlighting the importance of residual learning for deep networks.
- Without Attention Mechanism: Eliminating the attention layer resulted in a 1.8% performance decrease, demonstrating the value of temporal weighting.
- Without Bidirectional LSTM: Using unidirectional LSTM instead of bidirectional reduced performance by 2.1%, indicating the importance of contextual information from both temporal directions.

These results validate the architectural choices in our proposed framework and demonstrate the synergistic benefits of each component. The near 0.99 performance attained by all the three models highlights the capability of the deep learning based methods for emotion recognition based on EEG . The consistency across metrics suggests robust learning of discriminative features rather than overfitting to specific patterns.

5. Conclusion and Future Directions

For EEG-based emotion identification, this paper introduced a new hybrid deep learning approach combining ResNet50 and Bidirectional LSTM structures. With all assessment indicators reaching 0.99, our method produced an outstanding outcome and set a new state-of-the-art criterion on the GAMEEMO dataset. This finding emphasizes the great possibility sophisticated deep learning algorithms have in directly interpreting difficult emotional conditions from neural signals. The synergistic nature of our hybrid architecture, which expertly blends spatial and temporal feature learning, explains its success. While the Bi-LSTM component models the long-range temporal dependencies, the ResNet50 component excels at extracting complex, hierarchical spatial features throughout the multi-channel EEG scalp topography. natural in the development of emotional brain mechanisms. This shows how important it is to accurately identify both the where (spatial patterns) and the when (temporal dynamics) of neuronal activity in high-accuracy emotion classification. The high accuracy and robust performance of our model have significant practical implications, paving the way for a new generation of emotion-aware systems.

- In clinical settings, such a system could serve as an objective tool for the real-time monitoring of affective states in patients with mood disorders like depression or anxiety, potentially providing quantitative biomarkers for diagnosis and treatment evaluation.
- In human-computer interaction (HCI), it could facilitate the development of truly adaptive interfaces that intelligently respond to a user's emotional state, such as educational software that adjusts difficulty to mitigate frustration or vehicles that monitor for driver stress and fatigue.

- In the entertainment and gaming industry, this technology could be leveraged to create deeply immersive and personalized experiences, where game narratives, music, and challenges dynamically adapt to the player's real-time emotional responses.

Furthermore, by detailing the specific training parameters that yielded our top performance—50 epochs, a learning rate of 0.001, and a batch size of 32—this research provides a practical and reproducible guide for engineers and researchers. This helps reduce the extensive trial-and-error phase often associated with implementing deep learning applications in neuroscience. Looking ahead, future research will follow two main routes. First, we strive to improve the model's overall generalizability by testing its performance across more varied populations, documenting settings, and using EEG equipment to guarantee its robustness in actual life situations. Second, we'll investigate the actual application of our model, stressing on maximizing its architecture for low-power, edge-computing devices fit for wearables, mobile health uses. This study not only helps to develop affective computing but also opens up intriguing new possibilities for the development of more empathetic and reactive technologies. over several fields of expertise.

ACKNOWLEDGEMENTS

Giving me this chance to think over my work, the distinguished Dijla University publication would be appreciated.

REFERENCES

[1] Yang, Y., et al. "A hybrid deep learning model for EEG-based emotion recognition." IEEE Transactions on Cognitive and Developmental Systems 12.4 (2020): 832-841.

[2] Li, X., et al. "An attention-based hybrid CNN-LSTM model for EEG emotion recognition." IEEE Transactions on Neural Systems and Rehabilitation Engineering 28.7 (2020): 1522-1531.

[3] Koelstra, S., et al. "DEAP: A database for emotion analysis using physiological signals." IEEE Transactions on Affective Computing 3.1 (2012): 18-31.

[4] Zheng, W. L., & Lu, B. L. "Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks." IEEE Transactions on Autonomous Mental Development 7.3 (2015): 162-175.

[5] Yang, J., et al. "Multimodal EEG-based emotion recognition using hybrid deep learning models." IEEE Transactions on Affective Computing 14.1 (2023): 456-470.

[6] Koelstra, S., et al. "DEAP: A database for emotion analysis using physiological signals." IEEE Transactions on Affective Computing 3.1 (2012): 18-31.

[7] Huang, H., et al. "Meta-learning for few-shot EEG-based emotion recognition." IEEE Journal of Biomedical and Health Informatics 27.5 (2023): 2435-2445.

[8] Zhao, L., et al. "Spatio-temporal attention networks for EEG emotion recognition." IEEE Transactions on Neural Networks and Learning Systems (2023).

[9] Zhang, D., et al. "Graph neural networks for EEG-based emotion recognition." Medical Image Analysis 85 (2023): 102745.

BIOGRAPHIES OF AUTHORS

Author picture	Researcher **Hasan Hashim** is a dedicated student at Islamic Azad University in Iran, Department of Biomedical Engineering. He also has a Bachelor degree of Engineering in Medical Device Technology... University of Mesopotamia (Bilad Al- Rafidain). His Work as a cashier 'I work in a clothing store' did not stop him from perusing higher studies in his major..