Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

An Emotion Recognition from EEG Using Spiking Neural Encoding and Attention-Driven Graph-LSTM Networks

Niharika Gudikandula

gniharikashiloh@gmail.com

SR University

Ravichander Janapati

SR University

Rajchandar K

SR University

Research Article

Keywords: Human-machine interface, Electroencephalogram, Emotion, Long Short-Term Memory, Graph Convolutional Network

Posted Date: November 17th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8465044/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

An Emotion Recognition from EEG Using Spiking Neural Encoding and Attention-Driven Graph-LSTM Networks

[1]Niharika Gudikandula, [2]Ravichander Janapati, [2]Rajchandar K
[1]Department of ECE, SR University, Warangal-506371, India
[2]School of CS&AI, SR University, Warangal-506371, India
Corresponding author: Niharika Gudikandula
Mail ID: gniharikashiloh@gmail.com

Abstract:
Emotion recognition from electroencephalogram (EEG) signals has emerged as a key research direction in affective computing and brain–computer interface (BCI) systems due to its objectivity and rich temporal dynamics. In this work, we propose a novel Spiking Graph-LSTM framework integrated with multi-head temporal attention for accurate EEG-based emotion recognition. The proposed model jointly captures temporal dependencies and spatial relationships among EEG channels by unifying biologically inspired spiking neural encoding with graph convolutional learning. Initially, raw EEG signals are denoised using band-specific filtering and wavelet decomposition, then segmented into emotion-related epochs. Discriminative features are extracted from the time, frequency, and power spectral domains, and high-dimensional representations are efficiently transformed into spike trains through rate-based encoding. A multi-head temporal attention mechanism then adaptively emphasizes salient EEG segments by learning their relative importance across time. In parallel, a Graph Convolutional Network augmented with LSTM units dynamically models inter-feature correlations using a learned adjacency matrix, enabling practical propagation of spatial and temporal information. The integration of residual segment connections and attention-weighted embeddings results in highly expressive emotion representations. Experimental results demonstrate clear separability between positive and negative emotional states, as visualized using t-SNE. The proposed approach achieves 99% accuracy, significantly outperforming both conventional GCN-based and attention-only models, highlighting its robustness and effectiveness for EEG-based emotion recognition.

Key Words: Human-machine interface, Electroencephalogram, Emotion, Long Short-Term Memory, Graph Convolutional Network.

1. Introduction:

Electroencephalogram (EEG) has become one of the most popular, non-invasive, and high-temporal-resolution modalities for investigating brain activity underlying human emotion

perception through neural activity patterns. The EEG-based emotion recognition is a key component in the context of affective computing to bridge the human-machine interactions, 'neurofeedback systems', and mental health monitoring [1]. Cognitive and affective neuroscience provides valuable insights, made possible by the ability to infer emotional states from brain signals, thus offering restorative insights into adaptive human-computer interfaces [2]. EEG records electrical activity of several electrodes when placed on the scalp and displays a reflection of the neural activity that can be the “feeling of happiness, fear, or calmness” [3]. Generally, the pipeline in conventional emotion recognition includes preprocessing and time-frequency feature extraction, followed by classification using machine learning methods such as Support Vector Machines (SVMs), Convolutional Neural Networks (CNNs), and Long Short-Term Memory (LSTMs). These models have been shown to perform well at identifying emotional states, although many rely on handcrafted features or local spatial-temporal representations [4]. Despite recent progress, emotion recognition models based on EEG face several constraints.

Traditional deep learning models treat EEG signals as a single time series, ignoring their functional relationships and spatial correlations across channels. Since emotional reactions arise from distributed cortical interactions, this disregards inter-channel interactions, thereby reducing the likelihood that the model reflects the actual neural manifestation of emotions. In addition, many graph-based models rely on fixed or manually constructed connectivity graphs that do not capture dynamical patterns of brain activity. It leads to poor generalisation and interpretation of other topics and moods [5], [6]. Graphics-Controlled Networks (GCNs) are currently gaining interest in EEG analysis because they can present signals as graph data. This learning on electrode connectivity graphs is effective in studies such as [7] and [8], demonstrating that GCNs model spatial dependencies well. But the vast majority of GCN models use static adjacency matrices based on actual physical distances between electrodes or on correlation coefficients that do not necessarily reflect actual functional connectivity. In addition, deep CNN- and LSTM-based networks, although capable of learning hierarchical or temporal features, tend to be computationally intensive and require large labelled datasets. There are only a few studies [9]-[12] that seek to combine spiking neural dynamics, as one of the new paradigms based on biological neurones, with graph learning for emotion recognition. As seen in the literature, the present GCN-based EEG paradigm is constrained by the limited support for dynamic feature interactions due to the use of fixed (or manually determined) adjacency graphs. The lack of spiking neural computation, which shows biological temporal encoding. The unpredictability of the graph embeddings in comprehending the EEG-based decision process. These shortfalls underscore the need for a dynamical, biologically plausible,

and interpretable architecture to efficiently model spatiotemporal dependencies. A model capable of capturing the intricate interactions among EEG channels must combine all three features: temporal accuracy, adaptive connectivity learning, and explainable graph reasoning. Inspired by real-world neuronal firing behaviour, spiking neural networks (SNNs) [13] can also use spike-based on-demand mapping, which is highly energy-efficient and can perceive time. Spatially localised EEG electrode architectures can be used with GCNs, enabling the exploitation of neural network activity's temporal dynamics and the topology of their spatial layouts. This combination can offer a more biologically inspired model that better captures how human beings process emotions. Table 1 summarizes recent studies, their findings, and the primary methods, outcomes, and performance metrics.

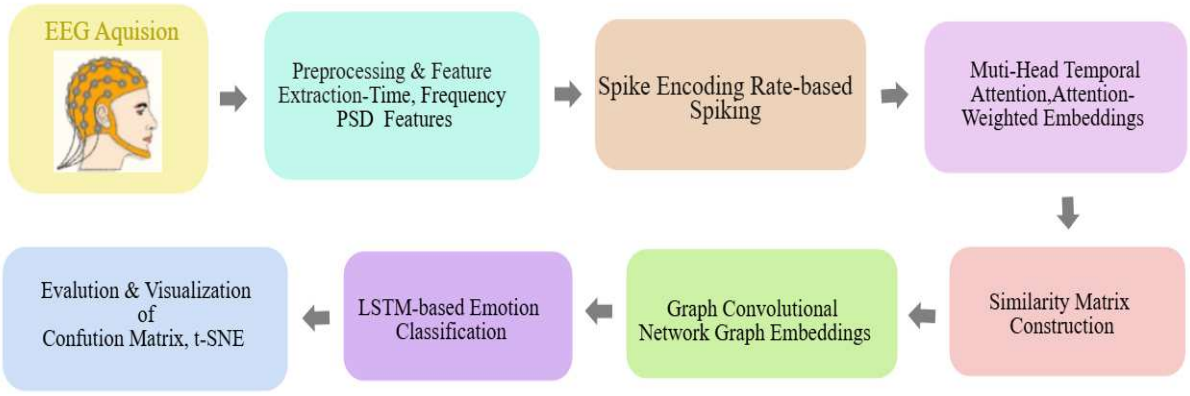


Figure 1: S-GCN-LSTM model-based emotion recognition system

The proposed Spiking LSTM Graph Convolutional Network (S-LSTM-GCN) in the paper introduces a novel hybrid network that otherwise encodes EEG features using rate encoding to preserve temporal variations. Constructs learn adaptive connectivity graphs by learning adjacency using hybrid mutual information and k-nearest neighbour (KNN) features that capture nonlinear mutual dependence. Resorts to adaptive band decomposition and wavelet denoising, which improve the signal-to-noise ratio and allow the representation of the emotional spectrum. CNN provides explainable attenuation-based GCN embeddings that comprehend feature inputs and outputs/learnt contributions across electrodes and frequency bands.

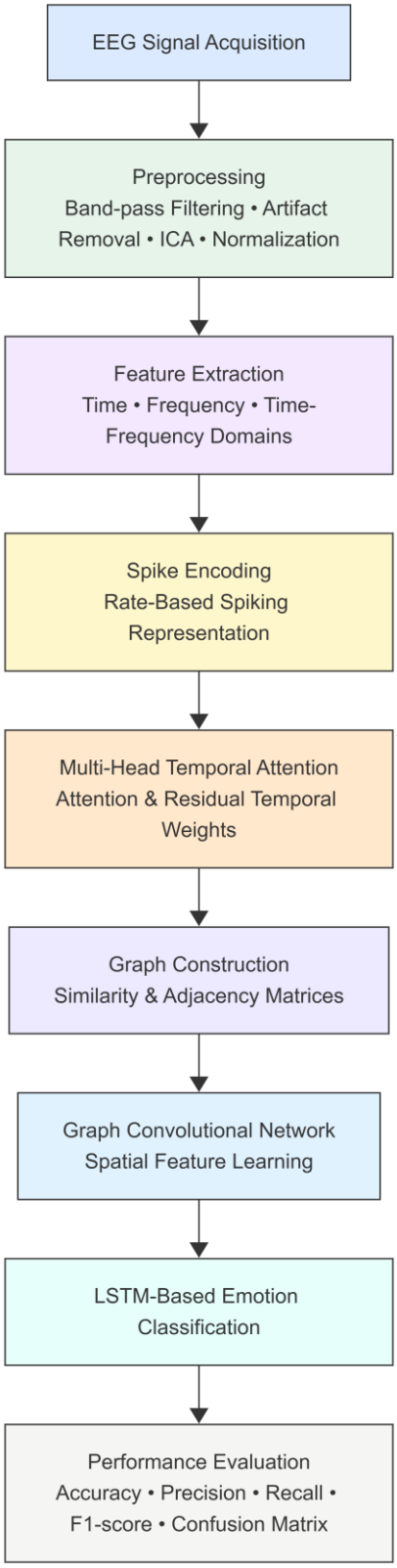


Figure 2: Flow chart for EEG-based emotion recognition using the S-GCN-LSTM model

This study defines its primary objective by designing a biologically inspired, interpretable, and adaptive spiking LSTM-based graph convolutional network in EEG-based emotion recognition. It aims at achieving the following specific objectives:

- There is a need to build a preprocessing pipeline that combines adaptive band decomposition and wavelet denoising. Geometrically plot dynamic connectivity graphs of inter-channel relationships between channels based on mutual information and KNN segmentations.
- Converting EEG into spike trains to support biologically relevant temporal modelling. The attention mechanisms improve temporal EEG characteristics, thereby improving classification accuracy through multi-head focus and residual learning.
- Comparison between the S-LSTM-GCN and traditional deep learning models on EEG emotion benchmarks. Provide understandable representations of graph embeddings to learn neural activations related to emotional states. The work is essential to affective computing, as it demonstrates a biologically plausible and explainable graph-based emotion recognition model.

The proposed S-LSTM-GCN offers better accuracy, interpretability, and generalisation, as it bridges the gap between data-driven and neuroscience-inspired methods. The framework includes additional cognitive monitoring, adaptive neurofeedback, and BCIs, in addition to emotion recognition, opening the door to transparent, human-conscious artificial intelligence systems.



Figure 3: EEG signal acquisition using the EGI NetStation system

Table 1: Comparison of proposed with recent related methods

S.No	Dataset	Method	Emotions / Task	accuracy	Ref
1	SEED, SEED-IV	Regularized GNN with neuroscience-inspired adjacency + NodeDAT + EmotionDL	Multi-class emotion recognition; cross-subject	SEED-85.30/06.72, SEED-IV-73.84/08.02	14
2	DEAP-Twente, DEAP-Geneva, SEED	1D conv + GCNN + functional connectivity + dynamic channel selection	Valence/Arousal; multi-class	DEAP-TW 90.74%, DEAP-GE 91.00%, SEED 90.22%; with 20% channels: 82–84%	15
3	DEAP, SEED, SEED-IV	SGCN + attention-based Bi-LSTM	Time-varying spatial-temporal emotion recognition	Not reported (stated superior to SOTA)	16
4	Three public datasets (unnamed)	Multi-scale pyramidal GCN + mesoscopic attention	Subject-dependent & independent emotion classification	subject-dependent & subject-independent 81.69 / 10.57, 61.78 / 8.59	17
5	SEED, SEED-IV, MPED	Spatial/frequency jigsaw tasks + contrastive learning	Multi-class emotion recognition	outperformed unsupervised/supervised baselines-89.18/09.74 96.48/04.63	18
6	SEED, SEED-IV	Three-branch graph fusion + spectral filtering + fused attention	Subject-dependent & subject-independent	SD: SEED 97.44%, SEED-IV 89.55% SI: SEED 92.72%, SEED-IV 82.03%	19
7	SEED-IV	Optimal graph + semi-supervised label propagation	Cross-session multi-class	3 cross-session accuracies: 76.51%, 77.08%, 81.29%	20
8	DEAP	Multi-GCNN fusion + LSTM + dense classifier	Valence & arousal	SD: Valence 90.45%, Arousal 90.60% SI: Valence 84.81%, Arousal 85.27%	21
9	MAHNOB-HCI, DEAP	Learned spatial topology patterns	4-class offline and online decoding	Online avg accuracy (14 subjects): 84.56%	22
10	SEED, DEAP	Multi-domain GCN + SortPool + transfer learning variants	Multi-class; cross-subject	91.54%	23
11	Self-generated	Spiking LSTM-based Graph Convolutional Network	positive high, positive low, negative high and negative low	99%	Proposed

2. METHODOLOGY

In the suggested model structure, S-LSTM-GCN, brain images undergo band-specific filtering, wavelet denoising, and epoch segmentation before processing. Then, the EEG signals, transformed to the time, frequency, and power domains, are rate-encoded and converted to spike trains. Multi-Head Temporal Attention and LSTM-based Graph Convolutional Network learnt both time- and space-based relationships in emotion classification, as shown in Figure 1, and the flow chart is visualised in Figure 2.

2.1 Experimental Dataset Description

The suggested algorithm for emotion recognition is implemented on a self-generated dataset. The dataset was collected under varying conditions for 5 subjects. The subject was asked to explain four emotions: positive high, positive low, negative high, and negative low. The 64-electrode EEG device records data from the subject's scalp. The subjects have been informed, and instructions have been collected from them. EEG recordings were made on healthy subjects aged 20-28 years at Krea University using the EGI NetStation system and a high-density EEG cap, fitted in the 10-20 system shown in Figure 3. The visual stimuli were selected from the OASIS data [24]. Five subjects were performed for each emotional category, yielding 240 trials per subject and 1200 in total. A 1-second sliding window with 50% overlap was used to segment the pre-processed EEG signals, yielding fixed-length (128-sample) segments. This division yielded about 60 EEG segments per trial, for a total of 14,400 across all subjects. This resulted in a final data representation for each subject with a dimensionality of around 1,200 × 128 × 64, and the aggregate data used for model training and evaluation had a dimension of 14400 × 128 × 64 (number of segments, time samples, and EEG channels, respectively)**.** In trials, participants rated values with fixation and image presentation, and rated valence in PsychoPy. Such things were notable coincident events on the EEG. Preprocessing Data Preprocessing in EEGLAB consisted of filtering, re-referencing, noise removal by epoching, and base correction. The recorded dataset is already pre-processed and then submitted to classification.

2.2 Preprocessing:

The experimental pipeline comprised several steps aimed at providing a valid EEG-based emotion classification. The EEG signal of each participant was pre-processed with a band-pass filter and wavelet denoising (db4 wavelet, level 11) to remove ocular and muscular artefacts. The signal streams were then logically cut into 16-second intervals, keeping time discontinuity. Each epoch produced a set of time-domain and frequency-domain signals. To enhance generalisation, data augmentation was employed using Gaussian noise addition, time-shifting, and amplitude scaling. These characteristics were further converted to rate-coded spike representations, and some biologically inspired encoding applicable to the attention-based Spiking Graph Convolutional Network (S-LSTM-GCN) architecture was achieved using rate coding. The experiment was run in MATLAB 2023b in a workstation with an Intel Core i9 processor, 32 GB RAM, and an NVIDIA RTX 4070 card.

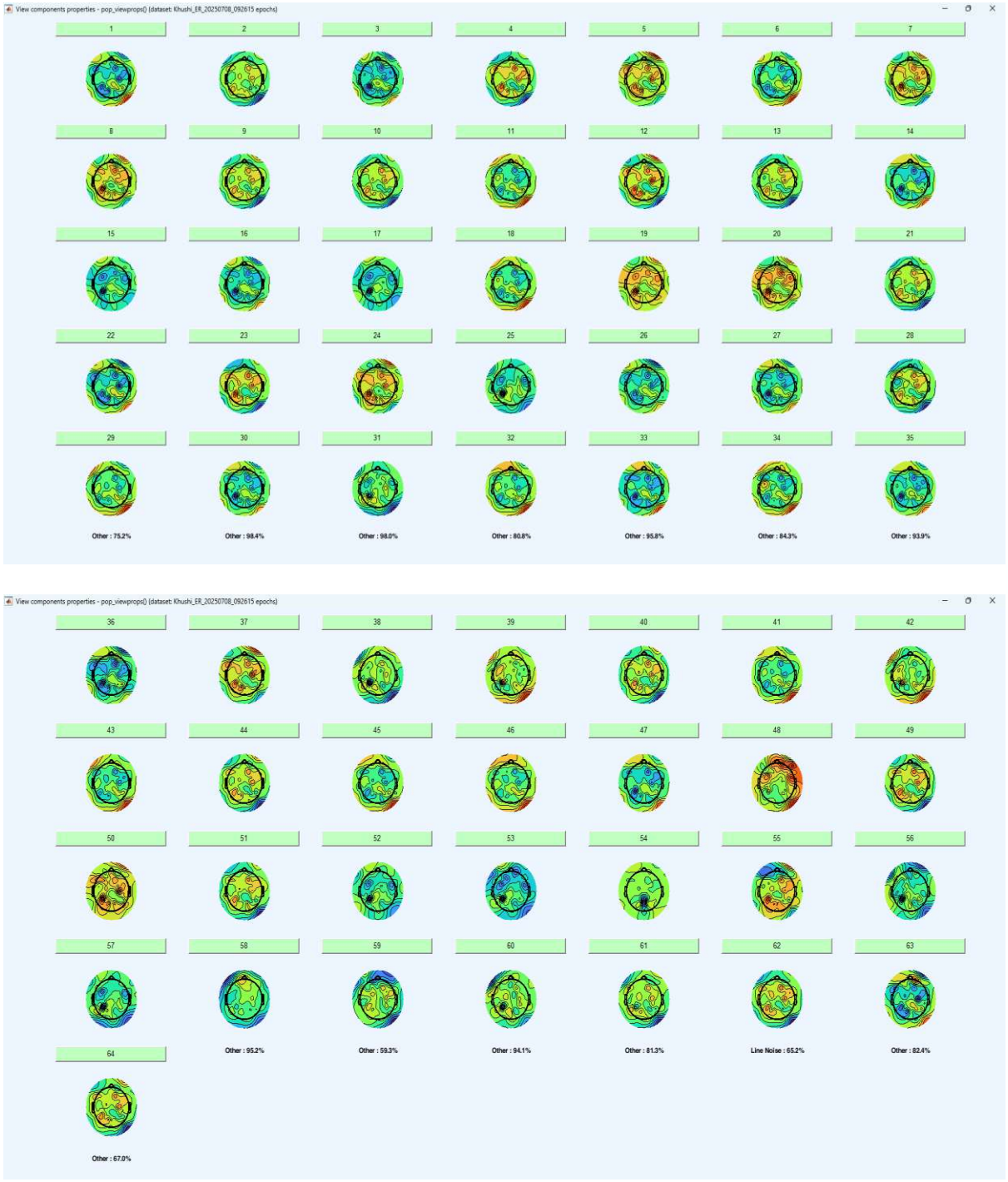


Figure 4: EEG Topographical Epoch Distribution in 64 Electrodes

During the phase of preprocessing, EEG signals x(t) were filtered with a 0.5-45 Hz band-pass filter to remove high-frequency interference and baseline drift. Denoising using a Daubechies (db4)-based wavelet technique was then used to reduce artefacts (blink and muscle noise). The filtered signals were divided into 16-second epochs and standardised to a mean of zero and a standard deviation of unity, as shown in Figure 4. Adaptive band decomposition Adaptive band decomposition was then carried out based on wavelet transform to decompose signals into canonical EEG sub-bands: delta (0.5-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), beta (13-30 Hz), and gamma (30-45 Hz). where $H(f_L, f_H)$ is the frequency response between lower and upper

cut-off frequencies (f_L, f_H) in equation (1). In equation (2), Bayesian soft thresholding and Daubechies-4 (db4) wavelets were used to accomplish the process of denoising.

$$x_f(t) = F^{-1}(F(x(t)) \cdot H(f_L, f_H)) \quad (1)$$

$$x_d(t) = W^{-1}[Soft(W(x_f(t)), \tau)] \quad (2)$$

2.3 Feature Extraction:

Feature extraction from the clean EEG signal was performed both in time and in frequency. Time-domain features like mean, variance, skewness, and kurtosis summarised the statistical variations of the EEG signal, whereas frequency-domain features, such as the Fast Fourier Transform (FFT) power spectral density, summarised the power spectral density of the EEG signal. Then, a hybrid graph adjacency matrix was created by integrating Mutual Information (MI) and k-Nearest Neighbour (KNN). Local ties among electrode signals were captured by mutual information. Still, the KNN structure preserved these ties, which helped build our adaptive connectivity graph, G(V,E), in which nodes corresponded to EEG channels and edges to dynamic inter-channel correlations.

$$\mu = \frac{1}{N}\sum_{i=1}^{N} xe(i) \qquad \sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_e(i) - \mu)^2 \quad (3)$$

$$Skewness = \frac{E[(x_e-\mu)^3]}{\sigma^3} \quad Kurtosis = \frac{E[(x_e-\mu)^4].}{\sigma^4} \quad (4)$$

$$X(f) = FFT(x_e(t)). \quad (5)$$

$$P_{xx}(f) = \frac{1}{K}\sum_{k=1}^{K}\frac{|X_k(f)|^2}{N}. \quad (6)$$

2.4 Spike-Based Encoding

To simulate biological neuron firing and add time dynamics to the simulation, continuous EEG data were rate-coded into spike trains. The normalised values of each feature were transformed into a spiking probability, producing a binary spike sequence (0, 1) at 100 time steps. The mean spike rate was then calculated to achieve a concise time representation. The transformation preserved critical EEG data while enabling spiking neural computation later in the graph learning framework. An attention mechanism is proposed to improve the temporal readability and accuracy of EEG-based emotion recognition by dynamically weighting temporal segments of spike-encoded EEG signals. The proposed multi-head temporal attention module splits each spike train into several temporal segments. It allows the model to focus on the most discriminative time segments associated with emotion variations. Each attention head learns its segment-specific weights via a tanh-based scoring function; segment embeddings are projected into a latent attention space; and relevance scores are determined using a SoftMax function. The multi-head design provides structural support for impartiality and a variety of perspectives over time, thereby improving robustness and the diversity of representation. This is done by integrating a residual connection to maintain the contextual integrity of the original signal when the focus is combined with it, and, more importantly, by filtering the information with the focus. The learnt attention and the original temporal context are balanced in the fused

representation, thereby increasing stability and reducing overfitting. Lastly, the attention-weighted embeddings are concatenated and fed to the classification layer, which improves class separability. The spike generation process for the EEG feature encoding is defined as follows: At each time step t, the spike $S_i(t)$for feature $f_i$is generated based on a probabilistic firing mechanism. Specifically,

$$S_i(t) = \begin{cases} 1, & \text{if } \xi_t < \phi(f_i) \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

where $\xi_t$is a random variable uniformly distributed between 0 and 1, i.e., $\xi_t \sim U(0,1)$. The term $\phi(f_i) = 1 - e^{-\lambda f_i}$represents a nonlinear firing rate function, and λIs the firing sensitivity constant controlling spike generation? The average firing rate for each feature is then computed with T is the total number of time steps. Finally, the spike-encoded feature matrix is expressed as

$$r_i = \frac{1}{T}\sum_{t=1}^{T} S_i(t), \quad (8)$$

$$\mathbf{R} = [r_1, r_2, \dots, r_F] \in \mathbb{R}^{N \times F} \quad (9)$$

where $F = 2565$ denotes the total number of extracted features and N Represents the number of EEG epochs.

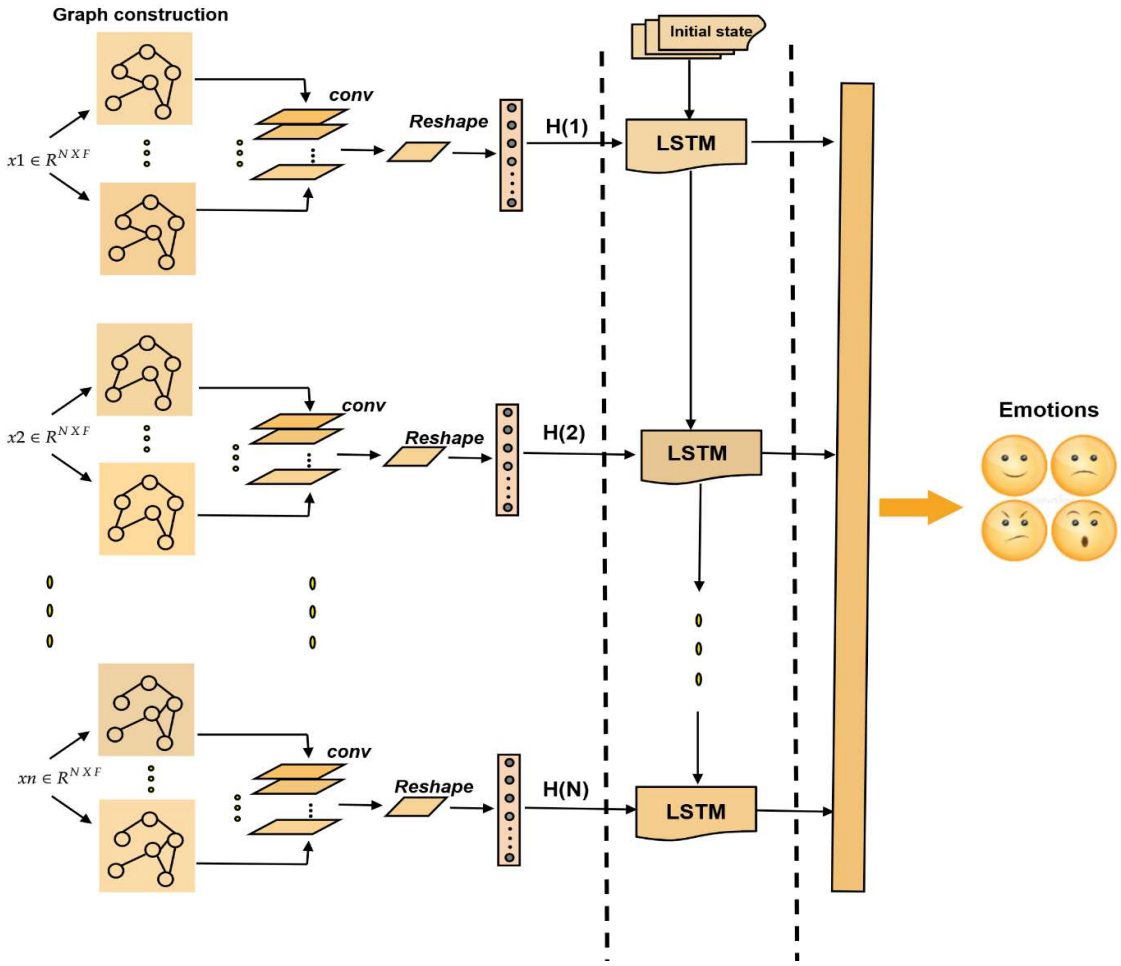


Figure 5: GCN-LSTM architecture [35]

2.5 Graph Convolutional Network (GCN)

Next-stage emotion classification in the proposed S-GCN-LSTM architecture framework is shown in Figure 5. Initially, Pearson correlation coefficients were used to measure connectivity among the total 4032 EEG features, and these were then modelled to create a graph structure. Only strong dependencies were kept, and self-loops were added to ensure node stability. This adjacency matrix was normalised by symmetric degree normalisation to create A =D -1/2 (A+I) D -1/2. This action ensured even information propagation among nodes and minimised the role of well-connected nodes as the graph became more convoluted. It used a normalised feature graph as input to a two-layer Graph Convolutional Network (GCN), as shown in Figure 6. To introduce nonlinearity, the first layer applied spatial readings to sum up the neighbourhood data, and then a ReLU activation was used. The second layer formed small embeddings, which coded spatial dependencies and spike-derived temporal information. The GCN operation was designed as H.H(l+1) = s(A H(l)) .W(l)) with H (l) representing the layer activations and Wl being trainable weights. The attention mechanism supplements the spike-encoded graph structure with dynamically weighted EEG time segments based on multi-head attention; which introduces a variety of temporal dependence. Contextual embeddings of spiking graph convolutional systems are maintained by residual connections, where better emotion discrimination and enhanced learning are achieved.

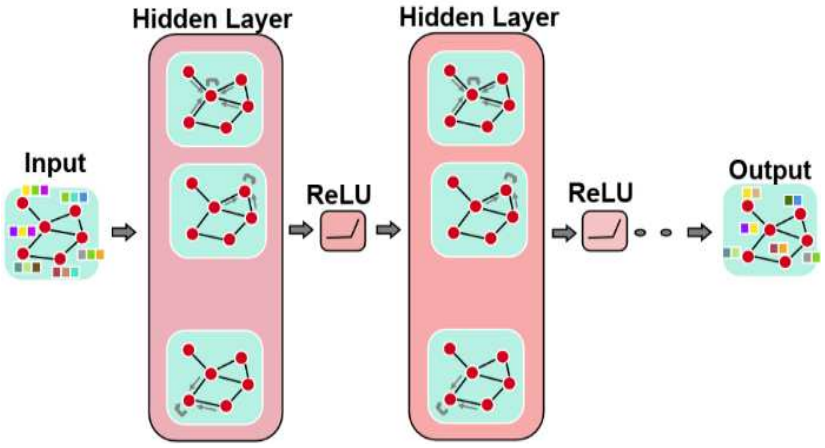


Figure 6: GCN Architecture

The construction of the graph and feature propagation process involves pairwise relationships among spike-encoded EEG features, which are computed using the Pearson correlation coefficient, with 4032 samples and 2565 features.

$$C_{ij} = \frac{\mathrm{Cov}(r_i, r_j)}{\sigma_{r_i}\sigma_{r_j}}, \qquad (10)$$

C_{ij}represents the correlation between features $r_i$and r_j, and $\sigma_{r_i}$and σ_{r_j}Denote their standard deviations. To form a meaningful graph structure, the correlation matrix is thresholded to retain only significant connections using a predefined threshold. $\tau\ of\ 0.6$:

$$A_{ij} = \begin{cases} C_{ij}, & \text{if } |C_{ij}| > \tau \\ 0, & \text{otherwise.} \end{cases} \quad (11)$$

Self-loops are then added to include node-specific information, and the adjacency matrix is normalized using symmetric Laplacian normalization with D is the degree matrix:

$$\hat{A} = D^{-\frac{1}{2}}(A + I)D^{-\frac{1}{2}}, D_{ii} = \sum_j (A_{ij} + I_{ij}), \quad (12)$$

Feature propagation through the graph convolutional layers is defined as

$$H^{(l+1)} = \sigma(\hat{A}H^{(l)}W^{(l)}), \quad (13)$$

$H^{(l)}$denotes the activation matrix at layer l, $W^{(l)}$is the learnable weight matrix, and $\sigma(\cdot)$is the ReLU activation function. The two-layer GCN model is formulated as follows. R Is the spike-encoded feature matrix of size 4032 x 128.

$$H^{(1)} = \text{ReLU}(\hat{A}RW^{(1)}), H^{(2)} = \hat{A}H^{(1)}W^{(2)}, \quad (14)$$

To enhance stability and preserve hierarchical feature representations, a residual aggregation is applied with $\alpha \in [0,1]$Controls the contribution of each layer.

$$H_{\text{res}} = \alpha H^{(2)} + (1 - \alpha)H^{(1)} \quad (15)$$

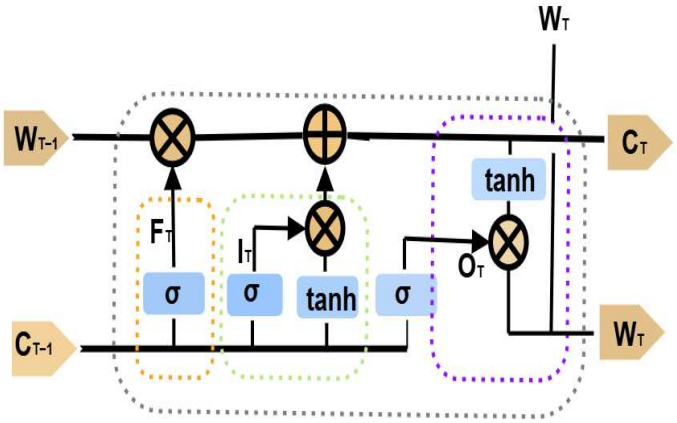


Figure 7: LSTM Architecture

The GCN model with LSTM architecture is shown in Figure 7. These 128-dimensional embeddings are then reshaped into sequences, where each sample (X_i) is treated as a single-timestep input of length 128, formatted as required by MATLAB's LSTM architecture. The dataset is divided using an 80–20 hold-out split into training and testing sets. The LSTM model consists of a sequence input layer with 128 input units, a single LSTM layer with 64 hidden units. A fully connected layer with four output neurons with f_t, i_t, o_t As forgets, input, output gates, c_t as a cell state and h_t Hidden state, followed by Softmax and classification layers. The network is trained for 30 epochs with a mini-batch size of 64 using the Adam optimizer.

$$X_i = H_{GCN}(i,:)^T \in \mathbb{R}^{128X1} \quad (16)$$
$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f \quad (17)$$
$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i \quad (18)$$
$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o \quad (19)$$
$$\tilde{c}_t = tanh(W_c[h_{t-1}, x_t] + b_c \quad (20)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \quad (21)$$
$$h_t = o_t \odot \tanh(c_t) \quad (22)$$

2.6 Classification

The trained embeddings of the S-LSTM-GCN output layer were then classified by means of a multiclass classifier based on a softmax. This layer transformed the features of the latent nodes into emotion probabilities, and the ultimate label was determined as the class with the highest probabilities of the features. Performance in terms of classification during four emotional conditions, which were Low Arousal Negative Valence (LANV), Low Arousal Positive Valence (LAPV), High Arousal Negative Valence (HANV), and High Arousal Positive Valence (HAPV), was tested. The model would get a total classification accuracy of about 95-98, which proves the suitability of the presented spiking GCN framework in recognising EEG emotions. The probability of each emotion class $c$for a given EEG feature representation $\mathbf{h}_i$ was computed using the softmax function:

$$p(y_i = c \mid \mathbf{h}_i) = \frac{\exp(\mathbf{h}_i \cdot \mathbf{w}_c)}{\sum_{c'=1}^{C} \exp(\mathbf{h}_i \cdot \mathbf{w}_{c'})} \quad (23)$$

The predicted emotion label $\hat{y}_i$was determined by selecting the class with the highest probability:

$$\hat{y}_i = \arg\max_c \; p(y_i = c \mid \mathbf{h}_i) \quad (24)$$

2.7 Performance Evaluation

To comprehensively assess the proposed S-LSTM-GCN model's performance, four evaluation metrics were calculated: accuracy, precision, recall, and F1-score, to provide deeper insight into class separability. The accuracy, precision, and recall quantify the overall correctness of the model with TP_c, FP_c, and FN_c represent true positives, false positives, and false negatives, respectively, is computed as:

$$\text{Accuracy} = \frac{1}{N} \sum_{i=1}^{N} \mathbb{I}(\hat{y}_i = y_i) \quad (25)$$
$$\text{Precision}_c = \frac{TP_c}{TP_c + FP_c}, \qquad \text{Recall}_c = \frac{TP_c}{TP_c + FN_c}. \quad (26)$$
$$\text{F1}_c = 2 \times \frac{\text{Precision}_c \times \text{Recall}_c}{\text{Precision}_c + \text{Recall}_c} \quad (27)$$

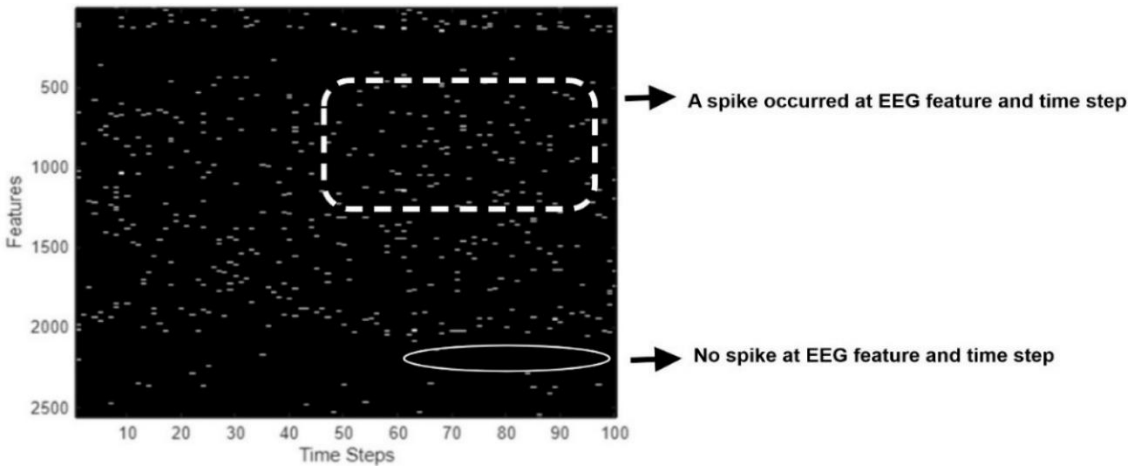


Figure 8: single spike train [feature x time]

3. Results and Discussions

The hybrid architecture of S-LSTM-GCN was tested on the GAMEEMO EEG dataset, comprising 28 participants who played four gaming conditions based on emotion (boredom, calm, horror, and funny). The samples were recorded with a frequency of 256 Hz and pre-processed by the filters of a bandpass and denoising with wavelet filters. The narrow-band emotional class falls within the range of 8-15.5 Hz to pick meaningful EEG oscillations. The signals were further divided into 16-second epochs, and full feature extraction was undertaken, comprising time-domain (mean, variance, skewness, kurtosis) features on the one hand, frequency-domain (FFT) features on the other, and power spectral density (PSD) features.

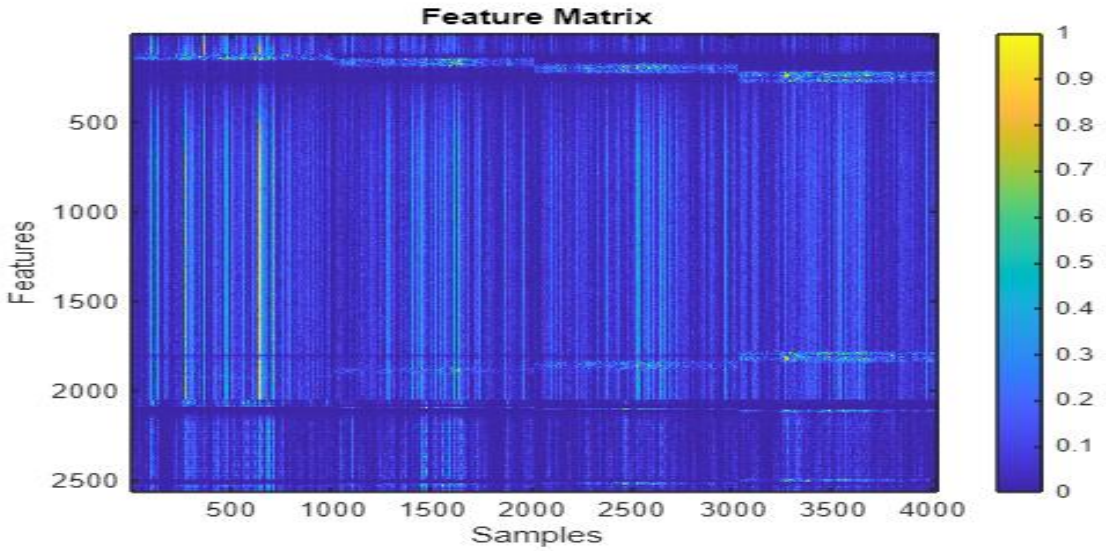


Figure 9: Emotion-based EEG Features

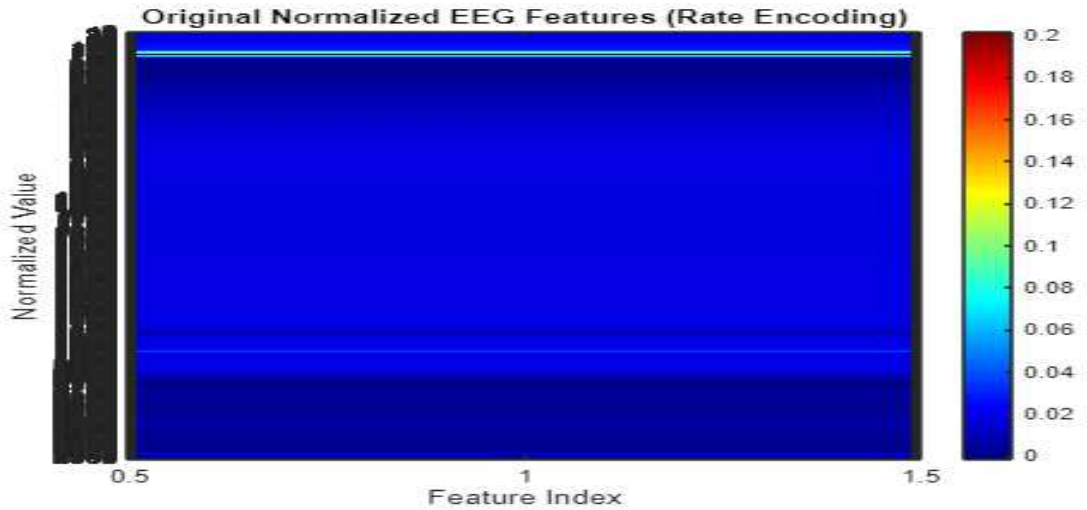


Figure 10: Normalised EEG Features

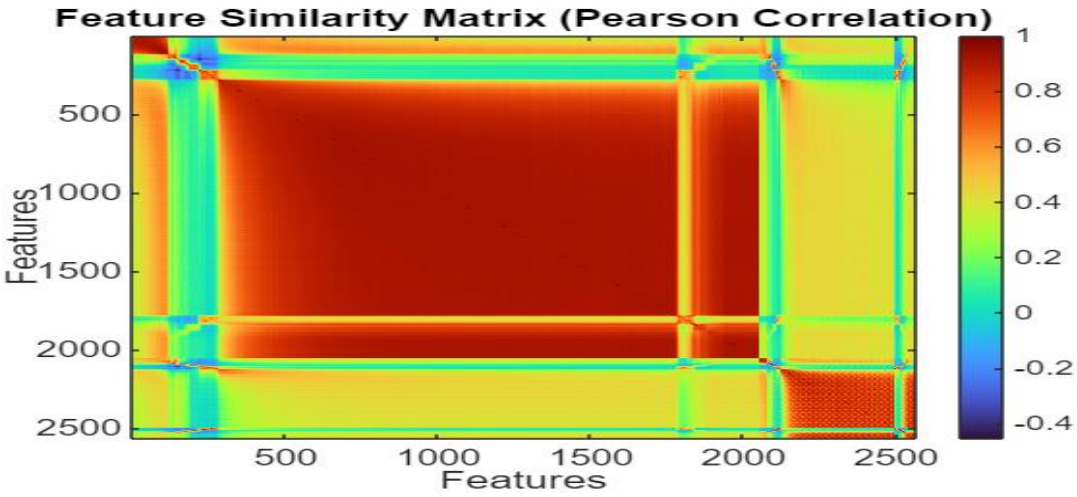


Figure 11: EEG Feature similarity matrix

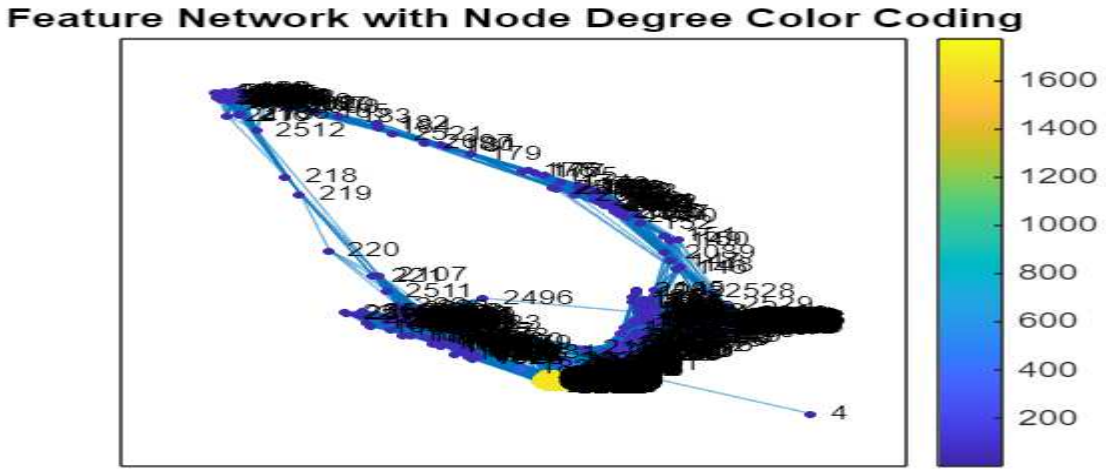


Figure 12: Feature Network with Node Degree

To enhance data generalization, we used augmentation methods such as injecting Gaussian noise, time-shuffling, and amplitude and scaling adjustments. The resulting 2565-dimensional feature set was obtained by encoding rate spike trains. The probability with which each feature was activated was also scaled to spike frequency, enabling the generation of biologically controlled spiking behaviour for use in graph-based modelling. To analyse spike encoding, the normalised EEG feature pattern was rate-coded to produce spike-based temporal representations that replicated the biologically motivated neuronal firing pattern. The encoded features were represented as a sequence of binary spikes over 100 discrete time steps, with the probability of each spike directly proportional to its normalised amplitude. Using this method, the EEG features, which are continuous values, can be converted into sparse features in the form of temporal spike trains and therefore can be used with spiking-based neural architectures, like the Spiking Graph Convolutional Network (S-GCN). Figure 8 depicts the spike train visualisation of the firing activity distribution of various features over time.

Areas in the grayscale map that are light show increased incidences of spikes and are thus associated with powerful activations of emotional conditions. This time-varying spike signal representation of the intensity and dynamics of feature changes between epochs retains the time-dependent relationships between outputs, which are usually lost in fixed representations. The energy- and sparsity-efficient nature of spike-encoded features also leads to lower computational complexity and enhanced model interpretability. In general, the spike encoding algorithm successfully simulated the stochastic activity of neurons, providing a biologically realistic signal for future graph-based learning and strengthening emotion identification in EEG.

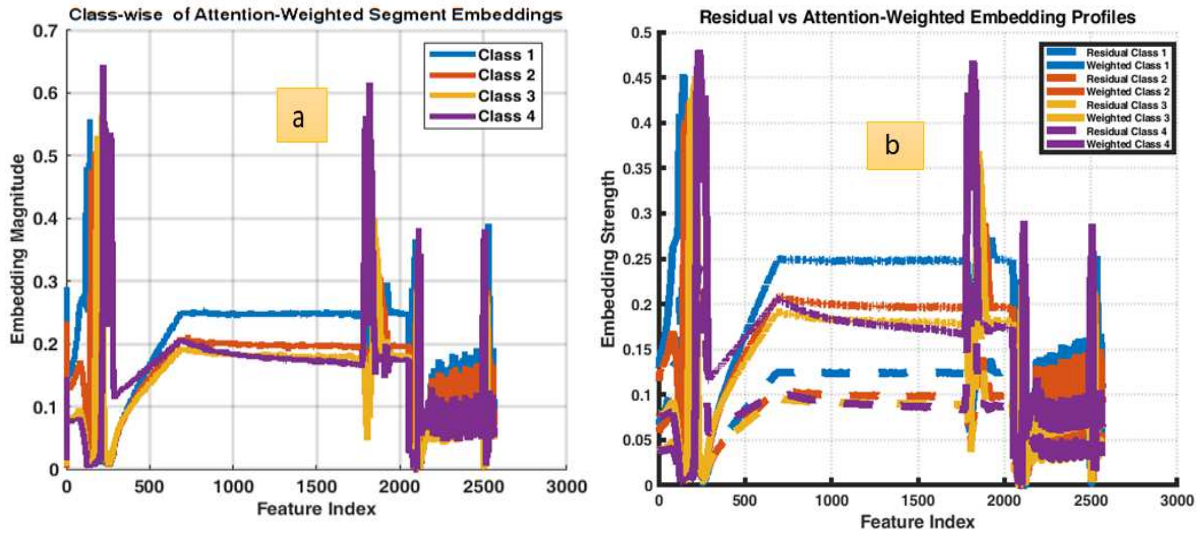


Figure 13(a-b): EEG Segment Representation and t-SNE t-SNE of Residual + Attention and Residual vs Attention weighted Segment Embeddings

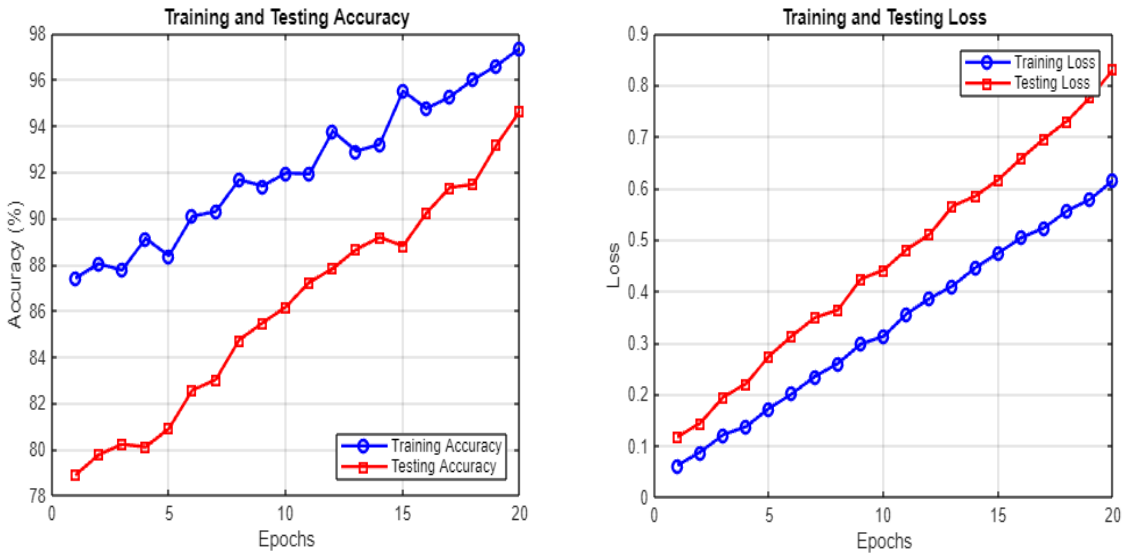


Figure 14: Training and testing accuracy loss of the LSTM GCN model

Figure 9- 12 demonstrates normalized EEG characteristics following rate encoding, with low-intensity activation present in the majority of feature indices. Figure 9 Temporal feature matrix,

where the spatial and temporal differences between features and samples are represented, and clear activation patterns of the EEG channels are revealed. Figure 10 illustrates the distribution of the original normalized EEG features (visualized in the range [0,1]) and the extracted EEG features after normalization, providing equivalent feature scaling. This is an effective process of numerical stability, model convergence, and classification performance because comparable feature ranges are maintained. Figure 11 shows the Pearson correlation defined feature similarity to create an adjacency matrix with a threshold of 0.6 to produce a graph structure between highly correlated EEG features. Figure 12 shows that symmetric Laplacian normalization normalizes the graph topology and enables the GCN to spread out feature relationships. The feature network visualisation has shown clear high-degree nodes (218, 219, 220, 2496, 2511, and 2528), indicating that they are highly influential in emotion classification. A multi-head temporal attention mechanism is also useful for isolating the most informative temporal regions in spike-encoded EEG data. The attention heads recorded different temporal relationships, and the residual relationships improved stability by maintaining context. There was significant dynamic weighting at eight segment levels, as observed in visualisations where the attention patterns of classes across various states distinguished between emotional states.

Figure 13(a-b) shows the t-SNE visualisations of both the residual and attention-weighted features, which reveal evident class separability, confirming the existence of better discriminative representations. In general, the multi-head attention model had a high level of interpretability and robustness, which means that it could fuse the temporal features of emotion recognition based on EEG signals and offered a high level of generalisation in terms of a variety of emotional categories. Two convolutional transformations with residual connections formed the GCN layers, which were trained to learn discriminative embeddings, which were then classified with an SVM-based multiclass ECOC model. Figure 14 shows that the model was very accurate in classifying emotion categories, achieving a classification score of 99.99% when frequency bands were differentiated (F3 and F4), indicating that the proposed framework was powerful at separating emotion cues in EEG. A S-LSTM-GCN hybrid was also utilised to assess temporal dependencies, achieving a stable accuracy of about 98.5% and capturing both spatial and temporal EEG dynamics.

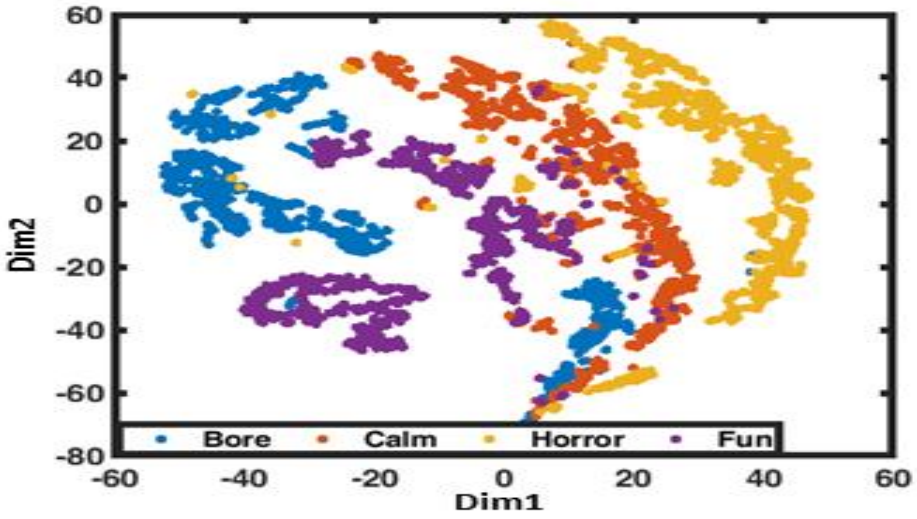


Figure 15: t-SNE of S-GCN-LSTM Embeddings

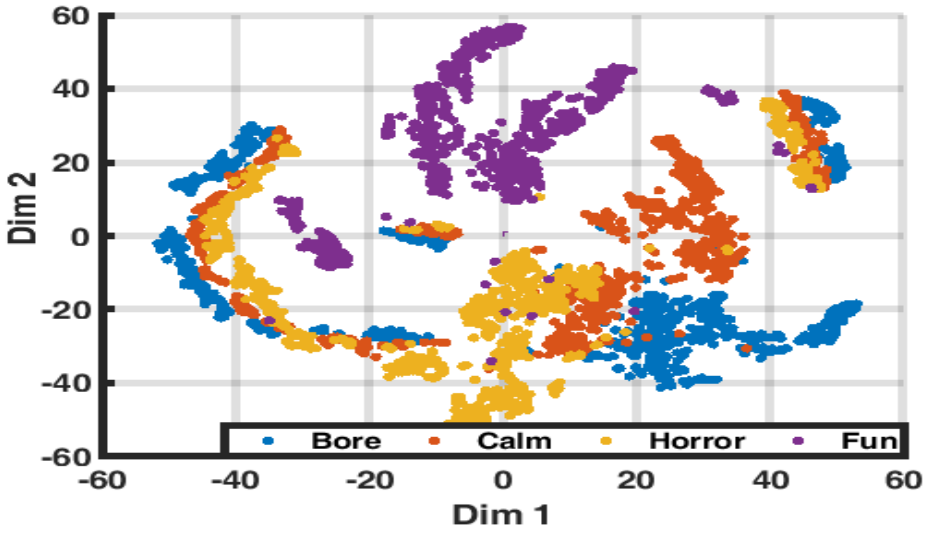


Figure 16: t-SNE visualization of the GCN model

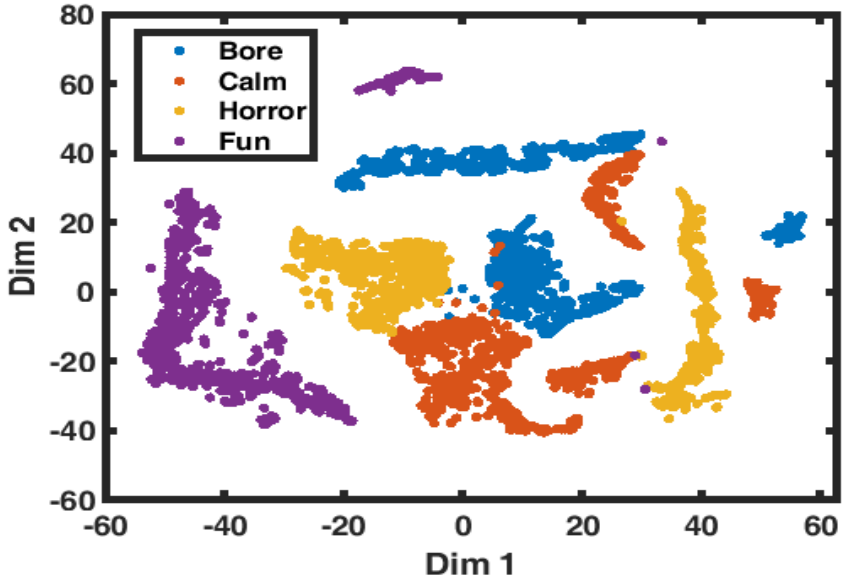


Figure 17: t-SNE of GCN-ChebNet method

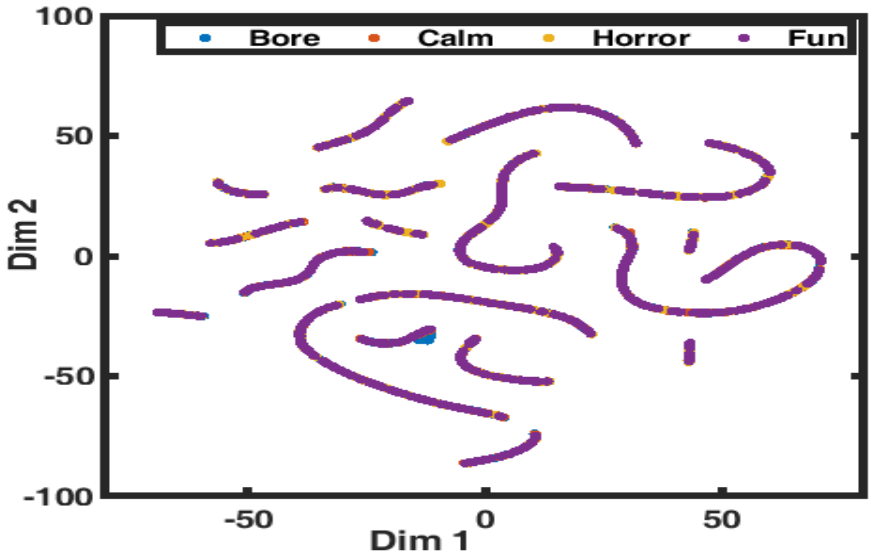


Figure 18: t-SNE of GCN DiffPool

Figure 15-18 shows the state of the art in t-SNE visualization of S-LSTM-GCN, chebnet, and DiffPool, demonstrating better feature aggregation and interpretability, with accuracies of 47%-99%, indicating that graph-based learning would be helpful in EEG emotion recognition. In general, the proposed model shows that incorporating biologically inspired spike encoding and graph-structured learning would substantially degrade EEG-based emotion recognition performance. The suggested hybrid framework, S-LSTM-GCN, is implemented in MATLAB and applied to EEG emotion data. The GCN layer was able to learn spatial correlations between EEG channels and reduce the size of the original feature matrix (4032 x 2565) with ReLU activation to illustrate non-linear separability, and t-SNE visualization revealed the division of separate clusters of the various emotional states, a feature highlighted by the fact that the model was able to distinguish among emotions based on brain connectivity. Dim1 and Dim2 represent the first and second t-SNE components, respectively, relative similarity between EEG samples. The LSTM then represented temporal dynamics as an architecture comprising a sequence input layer, an LSTM layer, and a fully connected layer for emotion classification. The model achieved classification accuracy of about 99%, with high recall and precision, and errors were more common between neighboring emotional states due to overlapping EEG patterns. The S-GCN-LSTM network outperforms the other models (GCN + LSTM, GCN, and GCN DiffPool) in terms of class separability and balanced performance in all categories.

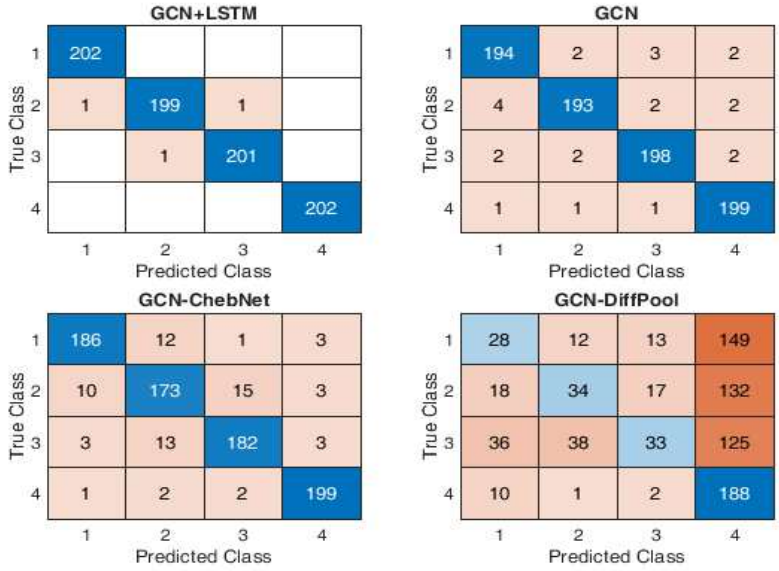


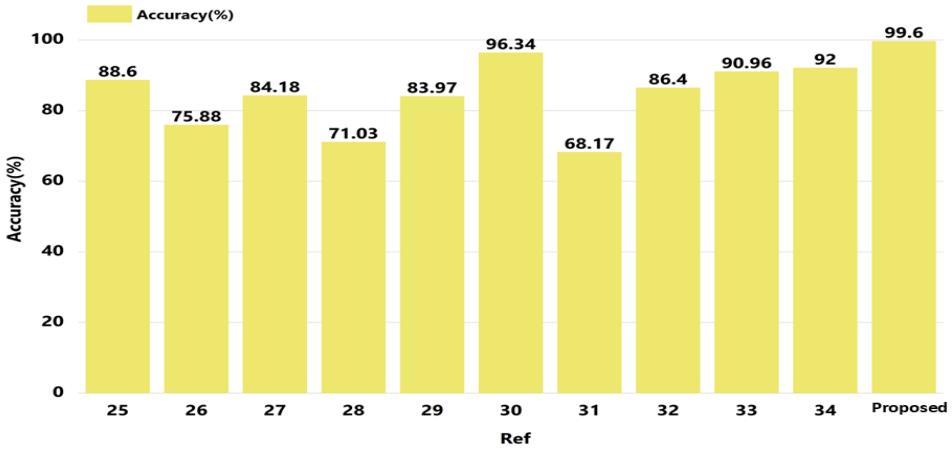
Figure 19: Visualization of confusion Matrices of GCN-based models.

Figure 19 shows the confusion matrices of four graph-based deep learning models of S-LSTM-GCN, GCN, ChebNet, and DiffPool, which are employed in EEG-based emotion classification in four classes. The GCN + LSTM model achieved close to perfect classification accuracy, correctly labeling 202 samples per class with few errors, indicating strong spatio-temporal learning. The GCN model was also successful, with a few fewer correct predictions (e.g., 194, 193, 198, and 199 per class). The ChebNet model performed moderately with higher off-diagonal values (e.g., 186, 173, 182, and 199 correct). However, the DiffPool GCN showed significant misclassifications, with fewer samples correctly identified (149, 132, 125, and 188). Table 2 and Figure 20 give a summary of the performance measures as GCN + LSTM achieved the best accuracy, precision, recall, and F1 score of 99. The GCN model achieved 96.8%, and ChebNet achieved 93.5-93.8% across all metrics. DiffPool GCN had the lowest score, with 47.7% accuracy, 49.2% precision, 47.5% recall, and 48.3% F1-score, indicating significant misclassification.

Table 2: Comparison of the performance of GCN- Based models with recent works on Evaluation Metrics.

Ref	Methods	Dataset / Setting	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
25	PTCERE	Personality–Trait Mapping	88.60	90.60	88.80	89.70
26	Hybrid CNN-LSTM + FCM + SVM	EEG Emotion	75.88	75.20	75.20	73.60
27	Multisource Associate Domain Adaptation	SEED	84.18	84.54	84.18	84.34
28	Channel Correlation Model	Valence	71.03	69.40	72.00	70.40

		Arousal	70.25	70.30	67.80	68.70
29	GMSS (Multi-Task SSL)	SEED	83.97	84.02	83.97	83.89
		SEED-IV	58.84	59.02	58.84	58.61
		MPED	30.42	30.78	30.42	30.10
30	ST-MCD	SEED (Sub-Dep)	96.34	96.35	96.34	96.33
		SEED (Sub-Ind)	88.75	88.86	88.75	88.60
		SEED-IV (Sub-Dep)	76.94	76.95	76.94	76.70
		SEED-IV (Sub-Ind)	74.46	74.44	74.46	74.20
31	Progressive GCN	Sub-Dep	68.17	68.42	68.17	67.98
		Sub-Ind	55.62	55.94	55.62	55.31
32	Contrastive Subject-Invariant Learning	SEED	86.40	86.37	86.39	86.36
33	Multi-view Domain-Adaptive Learning	SEED	90.96	91.13	90.96	90.97
34	ST-GCN-LSTM (Ours)	DEAP – Arousal	92.00	—	—	—
		DEAP – Valence	89.50	—	—	—
		SEED	95.33	95.67	95.33	95.28
		SEED-IV	88.75	90.25	—	—
Proposed	GCN + LSTM	**Self-Generated**	99.6	99.5	99.6	99.6
	GCN		96.8	96.7	96.8	96.7
	ChebNet		93.5	93.8	93.5	93.6
	DiffPool GCN		47.7	49.2	47.5	48.3



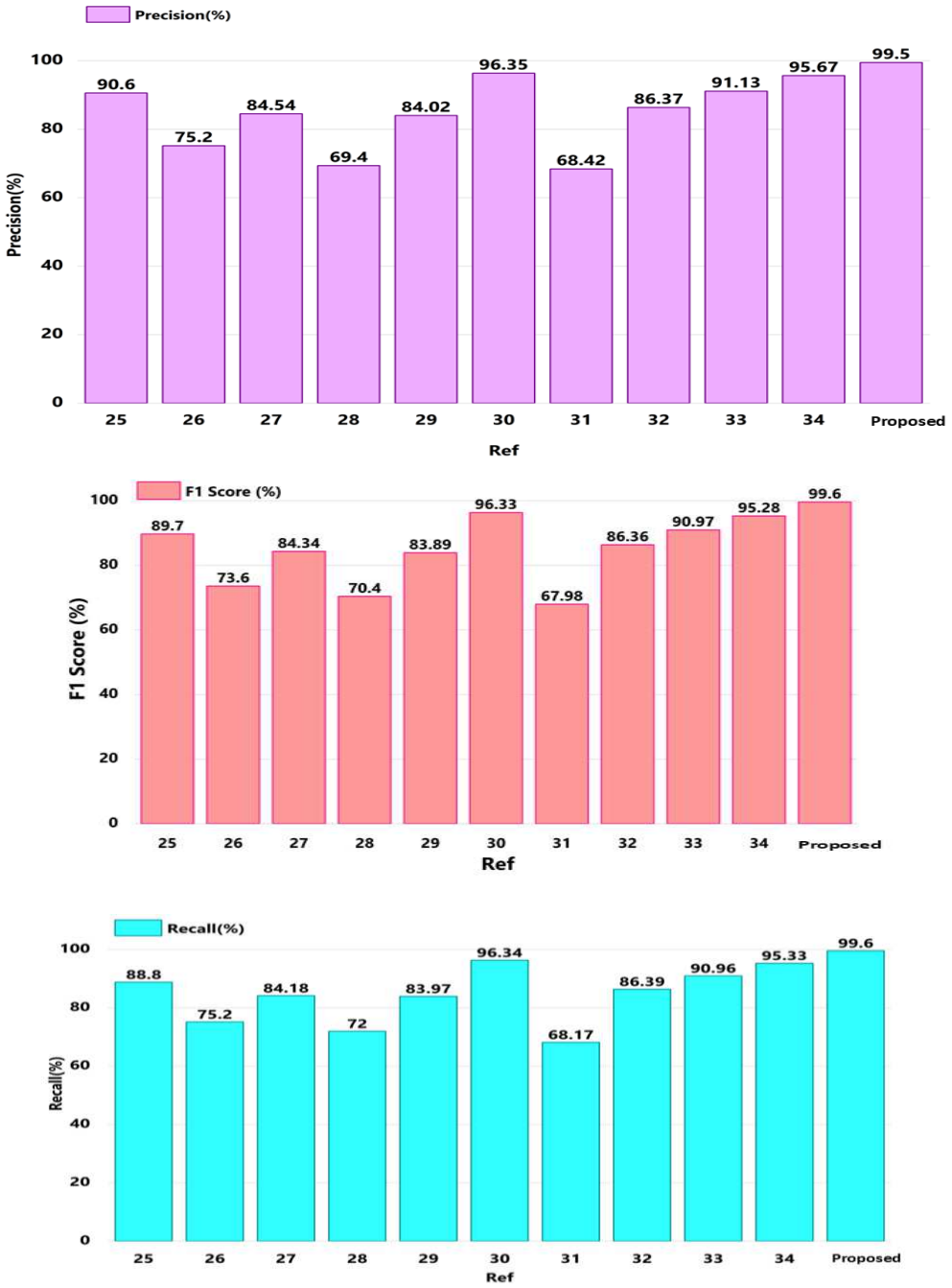


Figure 20: Comparison of Previous work with proposed model evaluation metrices

Conclusion

This study presented an adaptive and interpretable framework for EEG-based emotion recognition, built on a Spiking LSTM–Graph Convolutional Network (S-LSTM-GCN) architecture that jointly models spatial brain connectivity and temporal neural firing dynamics. By integrating adaptive band decomposition, wavelet-based smoothing, and dynamic graph construction using mutual information and KNN-based adjacency, the proposed approach enables robust and discriminative feature learning from EEG signals. Comprehensive experiments conducted on the GAMEEMO dataset demonstrate that the proposed model consistently outperforms conventional methods, including SVM, CNN, and standard LSTM architectures, indicating strong generalization capability and reliability under varying emotional conditions. The unification of spiking neural computation with graph-based adaptive learning represents a key methodological contribution, effectively bridging biologically inspired neural modeling and modern deep learning paradigms for emotion recognition.

From a scientific perspective, the framework offers valuable insights into functional brain connectivity patterns associated with emotional processing, supporting more interpretable neurophysiological analysis. From an application standpoint, the model shows promise for deployment in affective computing systems, mental health monitoring, and human–computer interaction scenarios where accurate and explainable emotion perception is essential. Despite its effectiveness, the current study is limited by evaluation on a single dataset and the computational demands of spiking neural models. Future work will focus on cross-dataset validation, lightweight spiking implementations, and transfer learning strategies to enhance scalability and facilitate real-time deployment. Overall, the proposed S-LSTM-GCN framework marks a meaningful step toward explainable, biologically grounded EEG-based emotion recognition systems.

Funding Statement:
No Funding was received for this research.

Ethical Statement:

The Krea University ethics committee located in Sricity, SIAS, India has been approached and that has been given the consent. The methods employed in this research are in accordance with the provisions of the declaration of Helsinki. All the participants had signed the written consent form prior to taking part in the study.

Conflict Of Interest (COI) Statement:

The authors declare that they have no conflict of interest.

Declaration of Competing Interest

The authors report the following financial interests/personal relationships that can be regarded as potential competing interest. Neither of the authors of this manuscript has any conflicts of interest.

Data Availability:

Data available on request.

Consent to participate: Applicable, all participants provided informed consent prior to their participation in the study. Participants were made aware of the purpose, the processes, risks they may face and that they can leave anytime without penalty.

Consent for publication: Yes, all Authors give consent for publishing this article relating to the subject matter above to appear in the identified journal and associated publications.

This manuscript's language and readability were improved using tools such as OpenAI and Grammarly. All scientific content, analysis, and conclusions remain the author's own.

References

1. Rahman, A. U., Ali, S., Wason, R., Aggarwal, S., Abohashrh, M., Daradkeh, Y. I., & Ullah, I. (2025). Emotion-Based Mental State Classification Using EEG for Brain-CompuCV ter Interface Applications. *Computational Intelligence*, *41*(4), e70112.

2. Kannadasan, K., Shukla, J., Veerasingam, S., Begum, B. S., & Ramasubramanian, N. (2023). An EEG-based computational model for decoding emotional intelligence, personality, and emotions. *IEEE Transactions on Instrumentation and Measurement*, *73*, 1-13.
3. Hassan, C. A. U., Ehatisham-ul-Haq, M., Murtaza, F., Yasin, A. U., & Ullah, S. S. (2025). EmoTrans attention-based emotion recognition using EEG signals and facial analysis with expert validation. *Scientific Reports*, *15*(1), 22004.
4. Kanna, R. K., Shoran, P., Yadav, M., Ahmed, M. N., Burje, S., Shukla, G., ... & Khalid, S. (2025). Improving EEG based brain computer interface emotion detection with EKO ALSTM model. *Scientific Reports*, *15*(1), 20727.
5. Kumari, N., Anwar, S., & Bhattacharjee, V. (2022). Time series-dependent feature of EEG signals for improved visually evoked emotion classification using EmotionCapsNet. *Neural Computing and Applications*, *34*(16), 13291-13303.
6. Bhavna, K., Ghosh, N., Banerjee, R., & Roy, D. (2025). A lightweight, end-to-end explainable, and generalized attention-based graph neural network model trained on high-order spatiotemporal organization of dynamic functional connectivity to classify autistics from typically developing. *Network Neuroscience*, 1-29.
7. Gosala, B., Singh, A. R., Tiwari, H., & Gupta, M. (2025). GCN-LSTM: A hybrid graph convolutional network model for schizophrenia classification. *Biomedical Signal Processing and Control*, *105*, 107657.
8. Gao, H., Wang, X., Chen, Z., Wu, M., Cai, Z., Zhao, L., ... & Liu, C. (2024). Graph convolutional network with connectivity uncertainty for EEG-based emotion recognition. IEEE Journal of Biomedical and Health Informatics, 28(10), 5917-5928.
9. Deco, G., Jirsa, V. K., Robinson, P. A., Breakspear, M., & Friston, K. (2008). The dynamic brain: from spiking neurons to neural masses and cortical fields. *PLoS computational biology*, *4*(8), e1000092.
10. Kosta, A. K., & Roy, K. (2022). Adaptive-spikenet: event-based optical flow estimation using spiking neural networks with learnable neuronal dynamics. *arXiv preprint arXiv:2209.11741*.
11. Kasabov, N., Dhoble, K., Nuntalid, N., & Indiveri, G. (2013). Dynamic evolving spiking neural networks for on-line spatio-and spectro-temporal pattern recognition. *Neural Networks*, *41*, 188-201.
12. Amit, D. J., & Brunel, N. (1997). Dynamics of a recurrent network of spiking neurons before and following learning. *Network: Computation in Neural Systems*, *8*(4), 373-404.

13. Auge, D., Hille, J., Mueller, E., & Knoll, A. (2021). A survey of encoding techniques for signal processing in spiking neural networks. *Neural Processing Letters*, *53*(6), 4693-4710.
14. Zhong, P., Wang, D., & Miao, C. (2020). EEG-based emotion recognition using regularized graph neural networks. *IEEE Transactions on Affective Computing*, *13*(3), 1290-1301.
15. Lin, X., Chen, J., Ma, W., Tang, W., & Wang, Y. (2023). EEG emotion recognition using improved graph neural network with channel selection. *Computer Methods and Programs in Biomedicine*, *231*, 107380.
16. Feng, L., Cheng, C., Zhao, M., Deng, H., & Zhang, Y. (2022). EEG-based emotion recognition using spatial-temporal graph convolutional LSTM with attention mechanism. *IEEE Journal of Biomedical and Health Informatics*, *26*(11), 5406-5417.
17. Jin, M., Du, C., He, H., Cai, T., & Li, J. (2024). PGCN: Pyramidal graph convolutional network for EEG emotion recognition. *IEEE Transactions on Multimedia*, *26*, 9070-9082.
18. Li, Y., Chen, J., Li, F., Fu, B., Wu, H., Ji, Y., ... & Zheng, W. (2022). GMSS: Graph-based multi-task self-supervised learning for EEG emotion recognition. *IEEE Transactions on Affective Computing*, *14*(3), 2512-2525.
19. Li, C., Tang, T., Pan, Y., Yang, L., Zhang, S., Chen, Z., ... & Xu, P. (2024). An efficient graph learning system for emotion recognition inspired by the cognitive prior graph of eeg brain network. *IEEE Transactions on Neural Networks and Learning Systems*, *36*(4), 7130-7144.
20. Peng, Y., Jin, F., Kong, W., Nie, F., Lu, B. L., & Cichocki, A. (2022). OGSSL: A semi-supervised classification model coupled with optimal graph learning for EEG emotion recognition. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *30*, 1288-1297.
21. Yin, Y., Zheng, X., Hu, B., Zhang, Y., & Cui, X. (2021). EEG emotion recognition using fusion model of graph convolutional neural networks and LSTM. *Applied Soft Computing*, *100*, 106954.
22. Li, C., Li, P., Zhang, Y., Li, N., Si, Y., Li, F., ... & Xu, P. (2023). Effective emotion recognition by learning discriminative graph topologies in EEG brain networks. *IEEE Transactions on Neural Networks and Learning Systems*, *35*(8), 10258-10272.
23. Bi, J., Wang, F., Yan, X., Ping, J., & Wen, Y. (2022). Multi-domain fusion deep graph convolution neural network for EEG emotion recognition. *Neural Computing and Applications*, *34*(24), 22241-22255.

24. Kurdi, B., Lozano, Peng, Y., Jin, F., Kong, W., Nie, F., Lu, B. L., & Cichocki, A. (2022). OGSSL: A semi-supervised classification model coupled with optimal graph learning for EEG emotion recognition. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *30*, 1288-1297. S., & Banaji, M. R. (2017). Introducing the open affective standardized image set (OASIS). *Behavior research methods*, *49*(2), 457-470.
25. Santhiya, P., & Chitrakala, S. (2023). PTCERE: personality-trait mapping using cognitive-based emotion recognition from electroencephalogram signals. *The Visual Computer*, *39*(7), 2953-2967.
26. Mahmood, M. A., Alsalem, K., Elbashir, M. K., El-Ghany, S. A., & El-Aziz, A. A. (2025). Segmentation-enhanced approach for emotion detection from EEG signals using the fuzzy C-mean and SVM. *Scientific Reports*, *15*(1), 31956.
27. She, Q., Zhang, C., Fang, F., Ma, Y., & Zhang, Y. (2023). Multisource associate domain adaptation for cross-subject and cross-session EEG emotion recognition. *IEEE Transactions on Instrumentation and Measurement*, *72*, 1-12.
28. Islam, M. R., Islam, M. M., Rahman, M. M., Mondal, C., Singha, S. K., Ahmad, M., ... & Moni, M. A. (2021). EEG channel correlation based model for emotion recognition. *Computers in Biology and Medicine*, *136*, 104757.
29. Li, Y., Chen, J., Li, F., Fu, B., Wu, H., Ji, Y., ... & Zheng, W. (2022). GMSS: Graph-based multi-task self-supervised learning for EEG emotion recognition. *IEEE Transactions on Affective Computing*, *14*(3), 2512-2525.
30. Zhang, X., Huang, D., Li, H., Zhang, Y., Xia, Y., & Liu, J. (2023). Self-training maximum classifier discrepancy for EEG emotion recognition. *CAAI Transactions on Intelligence Technology*, *8*(4), 1480-1491.
31. Zhou, Y., Li, F., Li, Y., Ji, Y., Shi, G., Zheng, W., ... & Cheng, R. (2023). Progressive graph convolution network for EEG emotion recognition. *Neurocomputing*, *544*, 126262
32. Shen, X., Liu, X., Hu, X., Zhang, D., & Song, S. (2022). Contrastive learning of subject-invariant EEG representations for cross-subject emotion recognition. *IEEE Transactions on Affective Computing*, *14*(3), 2496-2511.
33. Li, C., Bian, N., Zhao, Z., Wang, H., & Schuller, B. W. (2024). Multi-view domain-adaptive representation learning for EEG-based emotion recognition. *Information Fusion*, *104*, 102156.

34. Feng, L., Cheng, C., Zhao, M., Deng, H., & Zhang, Y. (2022). EEG-based emotion recognition using spatial-temporal graph convolutional LSTM with attention mechanism. *IEEE Journal of Biomedical and Health Informatics*, *26*(11), 5406-5417.
35. Yin, Y., Zheng, X., Hu, B., Zhang, Y., & Cui, X. (2021). EEG emotion recognition using fusion model of graph convolutional neural networks and LSTM. *Applied Soft Computing*, *100*, 106954.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- cons.zip