International Research Journal of Multidisciplinary Scope (IRJMS), 2026;7(3):216-231 IRJMS

Original Article | ISSN (O): 2582-631X DOI: 10.47857/irjms.2026.v07i03.010171

Towards Low-resource Deep Learning for Biomedical Signal Classification: An ECG and EEG Based Approach

Thota Leela Venkata Umamahesh*, Veerraju Gampala

Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Guntur andhra Pradesh, India. *Corresponding Author's Email: maheshuma413@gmail.com

Abstract

This study presents two lightweight deep learning pipelines for ECG arrhythmia and EEG emotion classification. These pipelines are designed for resource-constrained environments and can be executed on standard CPU hardware. The proposed framework combines self-supervised representation learning, hybrid neural architectures and GAN-based data augmentation into a single workflow to address limited labelled data and class imbalance. Class-wise GANs are used to balance ECG beats from the MIT-BIH Arrhythmia Database. A ViT–TFN classifier receives the robust latent representations that the SimCLR and MoCo models learn from the enhanced data. This model covers long-range temporal dependencies as well as fine-grained morphological characteristics. The emotion dataset follows the same augmentation and self-supervised learning pipeline before being processed by a GCN–GAT classifier that makes use of attention-based weighting and channel-wise spatial correlations. The EEG model achieves 94.15% accuracy across negative, neutral and positive emotions, the ECG model obtains 92.14% accuracy, with good performance for normal and supraventricular beats. The observed class separability is supported by Principal Component Analysis and t-Distributed Stochastic Neighbour Embedding visualizations, which show distinct clusters in the learned feature space. Overall, the findings show that hybrid architectures, augmentation and self-supervised learning can provide competitive performance in real world low-resource settings.

Keywords: ECG Beat Classification, EEG Emotion Recognition, Generative Adversarial Networks (GANs), Hybrid Deep Learning Classifiers, Low-resource Biomedical Signal Processing, Self-supervised Learning.

Introduction

Two essential non-invasive biomedical signals used in routine diagnosis; monitoring and decision support are electrocardiogram (ECG) and electroencephalogram (EEG) signals. Accurate automatic classification can help cardiologists identify life-threatening arrhythmias early and improve patient outcomes, according to recent ECG arrhythmia studies based on deep learning applied to 2D spectrograms, one-dimensional ECG traces and digital clinical recordings (1-4). Simultaneously, research on EEG based emotion recognition using hybrid CNN–Long Short-Term Memory models, attention-based transformers and spatiotemporal attention networks has shown that EEG can support brain–computer interfaces and affective computing applications by carrying rich information about affective states (5-7).

Benchmark Datasets

This work is supported by carefully selected open datasets. A significant benchmark for ECG is still the MIT-BIH Arrhythmia Database, which provides long-term Holter recordings with beat-level annotations spanning a variety of normal and pathological rhythms (8). In this study, an emotion-based EEG dataset previously used for sentiment and affective-state categorization is employed. Multi-channel recordings labeled with positive, neutral and negative emotional states make up the dataset (9). The EEG Brainwave Dataset and the MIT-BIH Arrhythmia Database were chosen due to their appropriateness for the suggested framework, clinical and application relevance and high annotation quality (8, 9). Furthermore, these datasets are useful benchmarks for creating trustworthy classifiers in resource-constrained environments where effective learning from comparatively small but informative data is crucial. Due to inter-subject heterogeneity, class imbalance and noise, these datasets are realistic but difficult testbeds.

Deep Learning Advances for ECG Classification

Deep learning has transformed ECG arrhythmia identification by replacing manually designed

(Received 19th January2026; Accepted 09th June 2026; Published 03rd July 2026)

features with automatically learned representations. By directly learning temporal and morphological features from raw or processed ECG segments, models based on CNN and Bidirectional Long Short-Term Memory have demonstrated exceptional performance (1-4). These architectures show how end-to-end training on large annotated datasets may sometimes even surpass expert level performance. However, they also highlight shortcomings in the classification of clinically uncommon arrhythmias, dataset change and class imbalance. In order to increase sensitivity for minority classes and effectively utilize the information present in lengthy ECG sequences, subsequent research has investigated residual CNNs, CNN–Bidirectional Long Short-Term Memory hybrids and transfer learning strategies (10-12).

Deep Learning Advances for EEG Emotion Recognition

In particular, existing EEG emotion classification models exploit the spatiotemporal complexity of multi-channel recordings. CNN–Bidirectional Long Short-Term Memory networks with residual Bidirectional Long Short-Term Memory designs and attention mechanisms have gradually enhanced the recognition of positive, neutral and negative emotions by integrating convolutional, recurrent and attention operations (5-7). Classification performance and robustness have been improved by more recent transformer-based EEG models and multi-fusion CNN–Bidirectional Long Short-Term Memory frameworks (13-16). These tests demonstrate the significance of recording both temporal dynamics within trials and spatial correlations across electrodes for accurate emotion recognition. Nevertheless, they remain susceptible to insufficient labeled data, noisy measurements and dataset specificity, especially when there are few individuals or sessions available.

Self-supervised Learning and Generative Augmentation

Recent developments in generative modelling and representation learning have improved bio-signal analysis by addressing label scarcity and class imbalance. By maximizing agreement between variously augmented representations of the same sample while contrasting them against other samples, SimCLR and Momentum Contrast (MoCo) develop discriminative embeddings, making efficient use of vast pools of unlabeled data (17, 18). It has been suggested that generative adversarial networks (GANs) increase training sets and balance underrepresented classes in medical time series data by creating realistic examples from latent noise vectors (19). By combining contrastive self-supervised learning with GAN-based augmentation, both labelled and unlabelled ECG and EEG recordings can be used more effectively to enhance model robustness in low data regimes in a way that makes sense.

Advanced Model Architectures for Bio-signals

Additional capabilities for simulating complex bio-signals are provided by transformer and graph-based designs. The Vision Transformer (ViT) is motivated to be used to ECG segments for global context modeling because it interprets inputs as patch sequences and employs multi-head self-attention to capture long range relationships (20). By combining gating, variable selection and attention mechanisms to concentrate on the most instructive temporal patterns, Temporal Fusion Networks (TFNs) expand this concept to time series forecasting (21). Graph Convolutional Networks (GCNs) simulate electrodes as nodes in a brain graph for EEG and transmit data along edges that represent functional or anatomical connectivity, in contrast the Graph Attention Networks (GATs) highlight the most informative inter-electrode interactions by introducing learnable attention weights on graph edges (22, 23). The spatial, temporal and spectral structure included in ECG and EEG signals can be effectively exploited by combining these structures.

Evaluation and Interpretability

For possible clinical use, dependable evaluation and interpretability tools are crucial. In medical decision support evaluation, receiver operating characteristic (ROC) analysis is usual because it offers threshold independent characterization of classifier performance (24). Visualization of high dimensional feature embeddings is made possible by dimensionality reduction techniques like Principal Component Analysis and t-Distributed Stochastic Neighbor Embedding, which aid in examining class separability and comprehending how various arrhythmia or emotion classes cluster in the learned representation space (25, 26). By providing intuitive insight into the behavior of intricate deep models, these visualizations

supplement quantitative measures. Figure 1 illustrates the transition from isolated, manual feature-based processing to an integrated and resource-efficient learning technique by contrasting the conventional pipeline with the suggested framework.

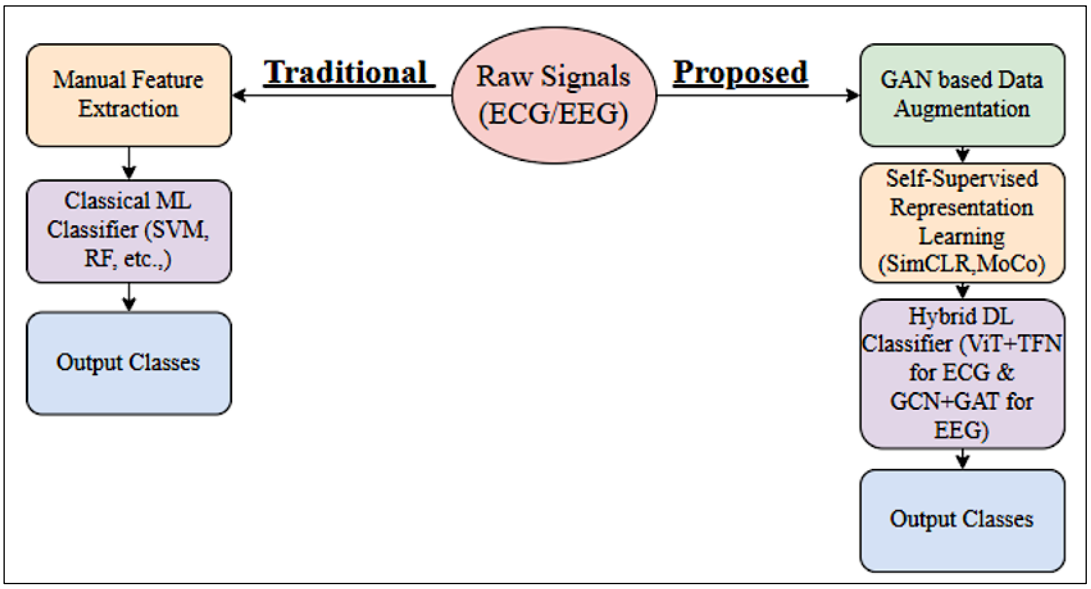

Figure 1: Traditional Vs Proposed Approach

Scopes

The diagnosis of ECG arrhythmias has improved thanks to deep learning, especially CNN and Bidirectional Long Short-Term Memory based models that directly learn temporal and morphological patterns from ECG signals (1-4). Because dataset shift, class imbalance and unusual arrhythmia types can still affect performance, further research explored residual CNNs, CNN and Bidirectional Long Short-Term Memory hybrids and transfer learning to improve minority-class sensitivity (10-12).

Additionally, CNN–Bidirectional Long Short-Term Memory, residual Bidirectional Long Short-Term Memory, transformer-based and multi-fusion models have enhanced the classification of positive, neutral and negative emotions (5-7, 13-16). Deep learning has also improved EEG emotion detection. However, these methods remain vulnerable to noise, dataset specificity, inadequate labeled data and high computing costs in low-resource contexts.

To bridge these gaps, the proposed architecture makes use of GAN-based augmentation, SimCLR and MoCo self-supervised learning, ViT–TFN for ECG and GCN–GAT for EEG (17-23). The study aims to create resource-efficient classifiers and evaluate them using ROC, Principal Component Analysis and t-Distributed Stochastic Neighbor Embedding visualizations on the MIT-BIH Arrhythmia Database and the EEG Brainwave Dataset (8, 9, 24-26).

Methodology

Data Collection

The MIT-BIH Arrhythmia Database (8), a popular benchmark dataset for cardiac signal processing, was used for the ECG classification task. The dataset contains 48 half-hour, two-channel ambulatory ECG recordings from 47 individuals, sampled at 360 Hz with 11-bit resolution. The dataset contains five different beat classes: Fusion Beats, Supraventricular Premature Beats, Ventricular Ectopic Beats, Normal Beats and Unknown Beats. The raw ECG waveform shows the characteristic P wave, QRS complex and T wave, reflecting the cardiac morphology required for arrhythmia classification. The waveform of the equivalent raw ECG signal is shown in the accompanying Figure 2. Accordingly, each ECG sample corresponds to approximately 2.78 msec, so the X-axis can be expressed in milliseconds. The Y-axis represents the ECG signal amplitude in millivolts (mV), corresponding to the electrical activity recorded over a 10-mV range. Its multi-class structure and clinical annotations make it appropriate for testing and developing arrhythmia classification algorithms.

The EEG Brainwave Dataset (9), which includes

EEG recordings from two people, one male and one female, under three emotional states: negative, neutral and positive; was used for the EEG categorization task. Both three minutes of recording for each emotional state and six minutes of neutral resting data were collected. The signals were recorded using a Muse EEG headgear and dry electrodes positioned at TP9, AF7, AF8 and TP10. The delta, theta, alpha, beta and gamma frequency bands all exhibit organized activity in the raw EEG waveform, suggesting significant brain-signal patterns appropriate for identifying emotions. Figure 3 displays the corresponding raw waveform of the EEG signal. The X-axis represents the frequency components obtained through Fast Fourier Transform-based spectral analysis and is expressed in Hertz (Hz), while the Y-axis represents normalized spectral power in arbitrary units (a.u.), reflecting the relative distribution of spectral energy across EEG frequency bands. The Y-axis is reported in arbitrary units because normalization removes the absolute physical scale, so the values are unitless and represent only relative magnitude.

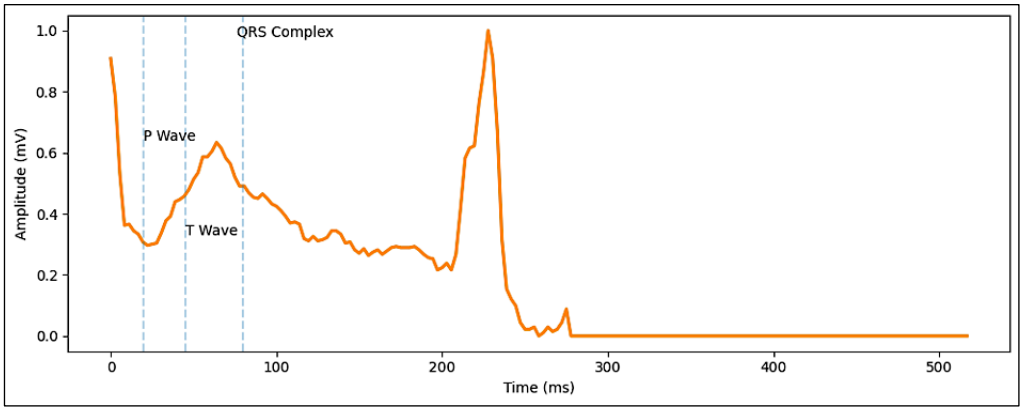

Figure 2: Raw ECG Waveform Representation

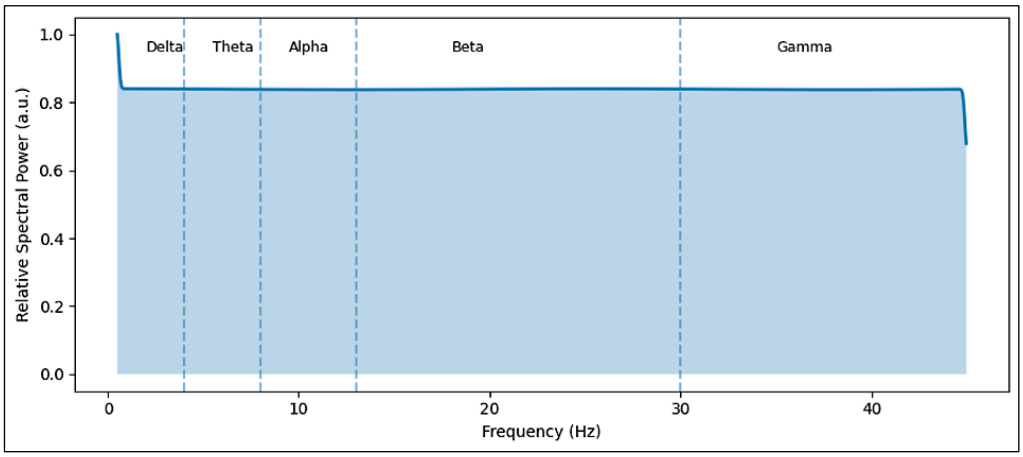

Figure 3: Raw EEG Waveform Representation

Data Pre-processing

The ECG dataset contains 21,981 samples, each of which is represented by 187 cardiac signal characteristics. The class designations correspond to the five arrhythmia categories: normal beat, supraventricular premature beat, ventricular ectopic beat, fusion of ventricular and normal beat and unknown beat. To place all variables on the same scale, feature normalization was done using Standard Scaler. Following that, an 80:20 stratified split was used to divide the dataset into 17,512 training samples and 4,379 testing samples while preserving the class distribution across the subsets (8).

The EEG dataset contains 2,132 samples with 2,548 features after removing one original feature. The labels represent three emotional states: negative, neutral and positive. Like ECG, an 80:20 stratified split yielded 1,705 training samples and 427 testing samples, which were normalized using Standard Scaler (9).

Data Augmentation

Generative Adversarial Networks (GANs) were used to create synthetic samples for both ECG and

EEG training data (19). The augmentation was implemented class-wise, preserving the original class structure while improving class balance by increasing each class to the nearest multiple of 500 samples. In this process, the generator produced synthetic samples from random noise and the discriminator learned to distinguish between generated and real data. Until the desired per-class augmentation target was reached, the adversarial training procedure continued.

The initial class distribution for the ECG training set was 14,493 Normal Beats, 445 Supraventricular Premature Beats, 1,158 Ventricular Ectopic Beats, 130 Fusion of Ventricular and Normal Beats and 1,286 Unknown Beats. A total of 18,500 samples made up the enhanced ECG training set after the class sizes were increased to 14,500, 500, 1,500, 500 and 1,500, respectively. There were initially 566 Negative, 573 Neutral and 566 Positive samples in the EEG training set. A total of 3,000 samples made up the enhanced EEG training set after each class was increased to 1,000 samples following augmentation.

The original training samples and the synthetic samples were combined to form the final augmented datasets, as shown in Equations [1, 2].

$$x_{aug\_ecg} = x_{train\_ecg} \cup \left(\bigcup_c \tilde{x}_{c,ECG}\right) \quad [1]$$

$$x_{aug\_eeg} = x_{train\_eeg} \cup \left(\bigcup_c \tilde{x}_{c,EEG}\right) \quad [2]$$

Figure 4 depicts the general GAN-based augmentation workflow. The next stage of the suggested pipeline is the self-supervised learning phase, which uses the augmented training sets created at this point as input.

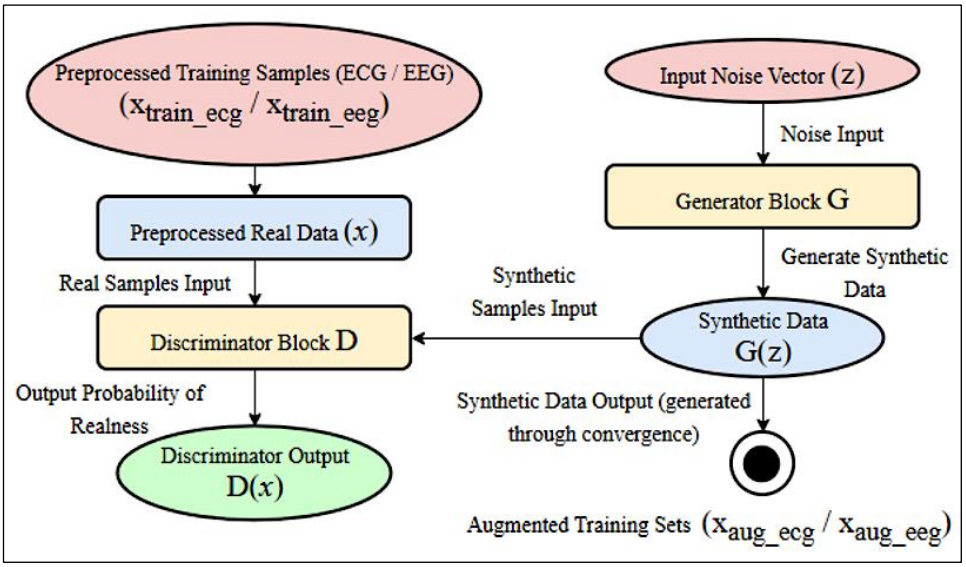

Figure 4: Framework for Enhancing ECG and EEG Signals Using Generative Adversarial Networks (GANs)

Self-supervised Learning

Due to differences in signal structure and representation requirements between the two modalities, the ECG and EEG training sets were handled independently in the self-supervised learning step following GAN-based augmentation (8,9,19). Let x_{aug_*} denote the augmented training set for either modality, where $*$ represents ECG or EEG. Using stochastic augmentations like noise injection and dropout, each sample in the augmented dataset was converted into two correlated views, enabling the model to learn resilient and invariant representations from the same input signal. $\hat{x}_k$ denotes the k^{th} sample in the augmented dataset, while t1 and t2 are the two stochastic augmentation functions applied to that sample as shown in the Equation [3].

$$\tilde{x}_{k1} = t1(\hat{x}_k), \tilde{x}_{k2} = t2(\hat{x}_k) \quad [3]$$

The two augmented views were then encoded independently to obtain latent representations, as expressed in Equation [4]. Before using contrastive optimization, the model must translate each converted view into a compact representation, which makes this step essential.

$$z_{k1} = \mathrm{Enc}(\tilde{x}_{k1}), z_{k2} = \mathrm{Enc}(\tilde{x}_{k2}) \quad [4]$$

Here, Enc(.) denotes the encoder network and z_{k1} and z_{k2} represent the corresponding latent embeddings of the two augmented views.

Equation [5] defines the SimCLR objective (17). This loss is used to pull positive pairs closer together while pushing apart negative pairs, thereby learning discriminative and invariant features from the augmented data.

$$l_{\mathrm{SimCLR}}(i,j) = -\left(\log \frac{\exp(\mathrm{sim}(z_i,z_j)/\tau)}{\sum_{k=1}^{2N} 1_{k\neq i}\exp(\mathrm{sim}(z_i,z_k)/\tau)}\right) \quad [5]$$

In this equation, z_i and z_j are embeddings of a positive pair, z_k represents embeddings of other samples in the batch, sim(·,·) is cosine similarity, τ is the temperature parameter and N is the number of original samples before augmentation.

To further strengthen contrastive learning, MoCo was incorporated using a momentum encoder and a dynamic queue of negative samples (18). This improves representation stability and expands the pool of negative examples available during training, which is why the MoCo loss is included (18).

$$\zeta_{\mathrm{MoCo}} = -\log \frac{\exp(q\cdot k^{+}/\tau)}{\exp(q\cdot k^{+}/\tau)+\sum_{k^-}\exp(q\cdot k^{-}/\tau)} \quad [6]$$

In Equation [6], q denotes the query embedding, k^+ denotes the positive key and k^- denotes the negative keys.

Equation [7] illustrates how the embeddings acquired from SimCLR and MoCo were concatenated to create a single fused feature representation for each modality (17,18). Here, z_{SimCLR} and z_{MoCo} are the feature vectors obtained from the two contrastive frameworks and X_{fused} is the final fused representation used for downstream classification and is verified by Random Forest classifier for interpretability.

$$X_{\mathrm{fused}} = [z_{\mathrm{SimCLR}}, z_{\mathrm{MoCo}}] \quad [7]$$

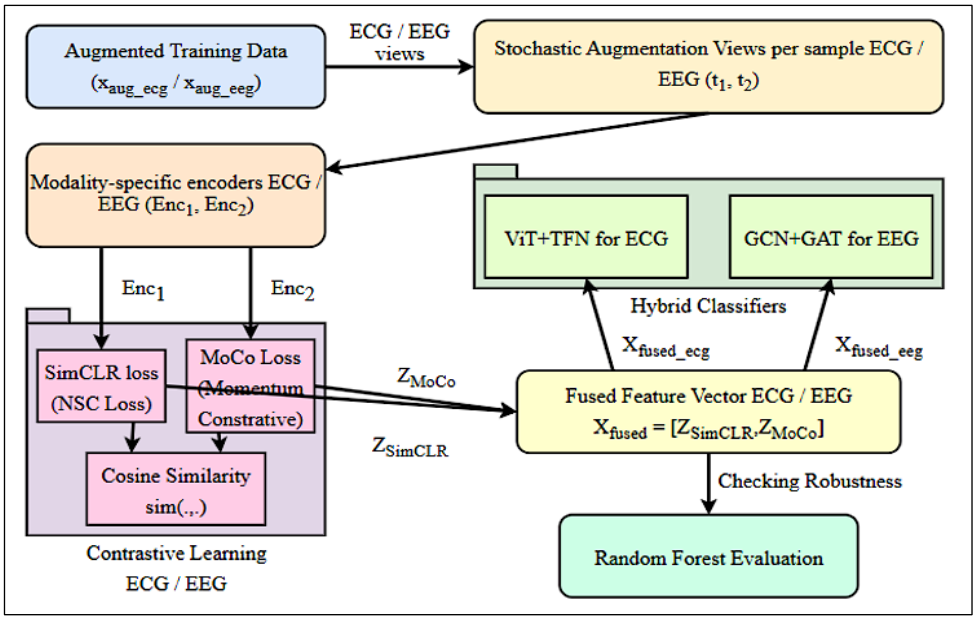

Figure 5: Self-supervised Learning Framework for ECG and EEG Classification that Demonstrates Contrastive Representation Learning and Feature Fusion

These fused representations were subsequently made available to the modality-specific hybrid classifiers, ViT–TFN for ECG and GCN–GAT for EEG, for final classification (20-23). The whole architecture of the self-supervised learning process is depicted in Figure 5, where the fused features generated in this step serve as the direct input for the subsequent hybrid classification stage.

Classification through Hybrid Classifiers

In order to capture the unique features of ECG and EEG signals, two modality-specific hybrid designs were used in the final classification stage. For ECG, the fused feature representation X_{fused_ecg} obtained from the self-supervised learning stage was projected into a token sequence suitable for transformer-based processing. A Temporal Fusion Network was subsequently used to describe temporal relationships in the ECG sequence after a Vision Transformer had processed these tokens to capture global dependencies (20, 21).

In the ECG branch, the attention mechanism computes the relevance between token embeddings through query, key and value matrices, denoted by Q, K and V, respectively.

Here, d_k represents the dimensionality of the key vectors and the softmax function normalizes the attention scores so that the model can focus on the most informative token relationships. This operation is expressed in Equation [8]:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \qquad [8]$$

A compact ECG representation was subsequently obtained by passing the attention output through bidirectional temporal modeling and pooling, which was finally transformed into class logits. The following step converts the learned representation into decision values for final ECG classification.

In Equation [9], W_{out} and b_{out} denote the learnable weight matrix and bias term of the output layer and z^{ECG} represents the unnormalized output scores for the ECG classes.

$$z^{ECG} = W_{out} \cdot \text{pool}(\text{bi\_outputs}) + b_{out} \qquad [9]$$

The fused EEG feature representation was treated as graph-structured input, with each node representing an electrode feature or EEG channel. During feature integration, local neighborhood information was first gathered using a Graph Convolutional Network and then surrounding nodes were given adaptive priority using a Graph Attention Network (22, 23). The model is able to extract useful inter-channel correlations and spatial dependencies from EEG data thanks to its architecture.

The graph attention mechanism computes normalized attention weights between neighbouring nodes. This weighting helps the model emphasize more informative neighboring nodes during graph-based feature aggregation.

In Equation [10], e_{ij} denotes the raw attention score between nodes i and j, α_{ij} is the normalized attention coefficient and $\mathcal{N}(i)$ represents the neighborhood of node i.

$$\alpha_{ij} = \frac{\exp\left(\text{LeakyRelu}(e_{ij})\right)}{\sum_{j \in \mathcal{N}(i)} \exp\left(\text{LeakyRelu}(e_{ik})\right)} \qquad [10]$$

After graph attention, the updated node embeddings were pooled into a graph-level representation and mapped to output logits for EEG classification. This step produces the final EEG prediction scores before label assignment.

In Equation [11], H' denotes the updated node embeddings after graph processing, while z^{EEG} represents the unnormalized output scores for the EEG emotion classes.

$$z^{EEG} = W_{out} \cdot \text{pool}(H') + b_{out} \qquad [11]$$

For both ECG and EEG, the predicted labels were compared with the ground-truth labels to evaluate classification performance. For both models, standard measures such as accuracy, precision, recall and F1-score were calculated. To give a more thorough evaluation of discriminative performance, especially for minority or clinically significant classes, ROC-AUC and class-specific sensitivity were also computed for each class of each model. Figure 6 depicts the modality-specific hybrid designs utilized in this phase and the Results and Discussion section presents the appropriate performance analysis.

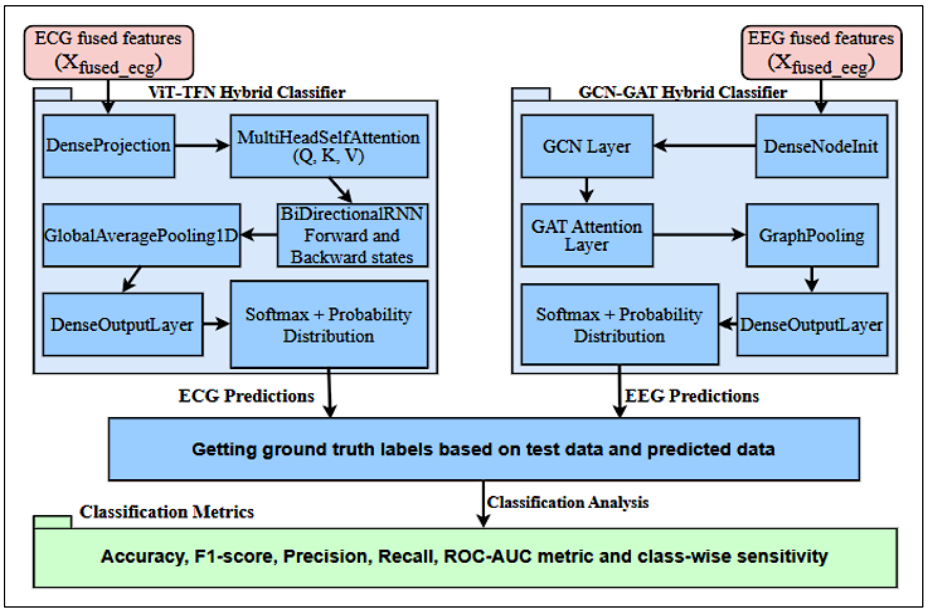

Figure 6: Modality-specific Hybrid Architecture Employing ViT–TFN for ECG Classification and GCN–GAT for EEG Emotion Recognition

Results and Discussion

Qualitative Analysis

The ViT and TFN models captured the essential morphological characteristics of the main cardiac beat types, as shown by the qualitative analysis of ECG data from the MIT-BIH Arrhythmia Database (8, 20, 21). While less common categories like Supraventricular Premature Beats and Fusion Beats exhibit morphological similarities, which may lead to misdiagnosis, Normal Beats have consistent waveform features that the models reliably identify. Due to their modest waveform fluctuations and overlap with neighboring patterns, Ventricular Ectopic Beats and Unknown Beats also pose classification issues. These results suggest that more robust feature representation and extraction techniques are required to enhance model interpretability and detection accuracy because it is difficult to notice minute changes in the waveforms.

GCN and GAT models are used to extract temporal and spatial information from the EEG Brainwave Dataset for the EEG emotion categorization task (9, 22, 23). Positive emotions are less distinct and overlap more with the other classes in the EEG feature space, while negative and neutral emotions show more pronounced signal differences that support stronger classification. This qualitative behavior emphasizes how difficult it is to decode affective states and demonstrates the necessity for more efficient graph-based feature learning techniques in order to increase classification accuracy.

Quantitative Analysis

The ECG hybrid model using ViT–TFN on the MIT-BIH Arrhythmia Dataset achieved 92.14% accuracy and a weighted ROC-AUC of 0.9263 (8, 20, 21, 24). Normal Beats were found to have the highest degree of dependability, exhibiting their dominating and constant waveform pattern with accuracy 0.93, recall 0.99, F1-score 0.96 and sensitivity 98.73%. Supraventricular Premature Beats exhibited poor sensitivity of 24.32%, poor F1-score of 0.39, poor recall of 0.24 and good precision of 0.93, indicating many missed detections despite precise positive predictions. Due to substantial overlap with other beat types,

the most difficult class was Fusion of Ventricular and Normal Beats, with precision 0.55, recall 0.19, F1-score 0.28 and sensitivity 18.75%. With accuracy of 0.75, recall of 0.54, F1-score of 0.63 and sensitivity of 54.48%, Ventricular Ectopic Beats had a mediocre performance. With accuracy of 0.97, recall of 0.83, F1-score of 0.89 and sensitivity of 82.61%, the detection of unknown beats was comparatively well. In general, dominant classes fared better with the ECG model than minority ones. The ECG categorization metrics are displayed in Table 1.

The Brainwave Dataset yielded a weighted ROC-AUC of 0.9866 and 94.15% accuracy for the GCN–GAT-based hybrid EEG model (9, 22-24). With accuracy 0.99, recall 0.99, F1-score 0.99 and sensitivity 98.60%, negative emotion had the strongest separability from the other groups. With accuracy of 0.89, recall of 0.97, F1-score of 0.93 and sensitivity of 96.48%, neutral emotion likewise demonstrated good identification and consistent categorization. With accuracy 0.95, recall 0.87, F1-score 0.91 and sensitivity 86.62%, positive feeling fared somewhat worse, indicating some overlap with adjacent emotional states. Compared to ECG classification, EEG classification was generally more evenly distributed among classes. EEG categorization metrics are shown in Table 2.

Table 1: Performance Metrics for ECG Beat Classification

Class	Precision	Recall	F1-score	Support
Normal beats	0.93	0.99	0.96	3624
Supraventricular premature beats	0.93	0.24	0.39	111
Ventricular ectopic beats	0.75	0.54	0.63	290
Fusion of ventricular and normal beats	0.55	0.19	0.28	32
Unknown beats	0.97	0.83	0.89	322
Accuracy			0.92	4379
Macro average	0.82	0.56	0.63	4379
Weighted average	0.92	0.92	0.91	4379

Table 2: Performance Metrics for EEG Emotion Classification

Class	Precision	Recall	F1-score	Support
Neutral	0.89	0.97	0.93	142
Negative	0.99	0.99	0.99	143
Positive	0.95	0.87	0.91	142
Accuracy			0.94	427
Macro average	0.94	0.94	0.94	427
Weighted average	0.94	0.94	0.94	427

Confusion Matrix Analysis

The MIT-BIH Arrhythmia Dataset's ECG confusion matrix for the ViT–TFN hybrid model reveals that Normal Beats are most accurately detected (8, 20, 21, 25), with very few data being incorrectly classified, primarily as Supraventricular or Ventricular kinds. The model's inability to identify this minority class is demonstrated by the fact that Supraventricular Premature Beats are often forecasted as Normal Beats. Confusion is also evident in ventricular ectopic beats, with a sizable percentage classified as normal. The model is biased toward the dominant class and is less successful for rarer or morphologically ambiguous arrhythmias, as evidenced by the increased likelihood of Fusion Beats and Unknown Beats being mistaken for Normal Beats. The ECG confusion matrix is shown in Figure 7.

The confusion matrix for EEG emotion classification using the GCN–GAT hybrid model on the EEG Brainwave Dataset shows that negative emotions are well recognized, whereas there is little misunderstanding regarding the positive class (9, 22, 23, 26). Additionally, neutral emotions are categorized with a high degree of consistency and minimal errors. However, because some samples are incorrectly labeled as Negative or Neutral, indicating some overlap among the affective states in feature space, positive feelings are relatively more challenging to detect. Overall, the EEG model performs somewhat worse for classes with higher overlap, but it does a better job of differentiating between distinct emotional states. The EEG confusion matrix is displayed in Figure 8.

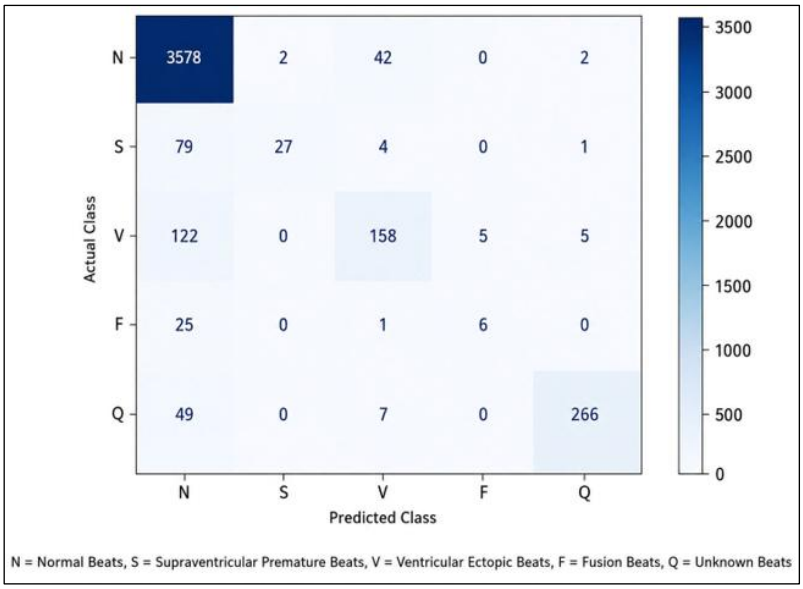

Figure 7: Confusion Matrix for ECG Arrhythmia Classification using the Proposed Hybrid ViT–TFN Classifier

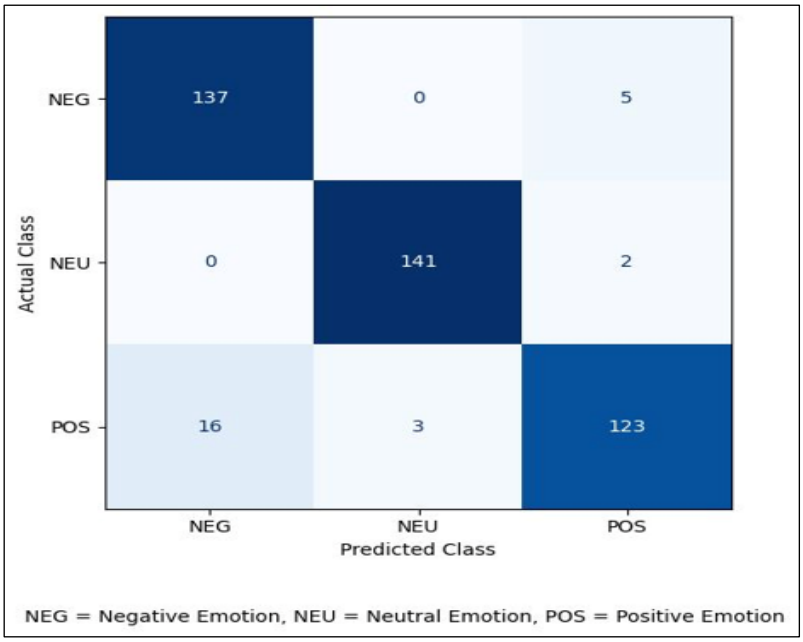

Figure 8: Confusion Matrix for the Hybrid GCN–GAT Classifier Proposed for EEG Emotion Classification

Embedding Space Analysis

The Principal Component Analysis of ECG characteristics on the MIT-BIH Arrhythmia Dataset predicted by the ViT and TFN models shows that Normal Beats (N) form a close-knit cluster that reflects similar feature representations (8, 20, 21, 25). It is difficult to identify Supraventricular Beats (S) from Normal Beats due to their greater dispersion and extensive overlap. Ventricular Beats (V) show considerable clustering but still overlap with Normal and Supraventricular beats, whereas Fusion (F) and Unknown (Q) beats are more dispersed and less identifiable, which is consistent with their rarity and morphological complexity. The t-Distributed Stochastic Neighbor Embedding projections of the same ECG features show a clearer separation: Fusion and Unknown beats appear as smaller, more dispersed groupings, Supraventricular and Ventricular beats form partially overlapping but more distinctive clusters and Normal beats stay firmly packed (8, 20, 21, 26). Figures 9 and 10 present these renderings.

For the negative, neutral and pleasant emotions that the GCN and GAT models learned for the EEG

Brainwave Dataset, Principal Component Analysis shows somewhat overlapping clusters. Neutral emotions are quite compact, negative emotions form a larger cluster and positive emotions have some overlap with the neighboring classes (9, 22, 23, 25). Although there is still considerable overlap at the borders, the t-Distributed Stochastic Neighbor Embedding image displays class separation with Negative, Neutral and Positive emotions forming tighter and more distinct clusters. This indicates that t-Distributed Stochastic Neighbor Embedding maintains local structure better for EEG embeddings (9, 22, 23, 26). The renderings are shown in Figures 11 and 12.

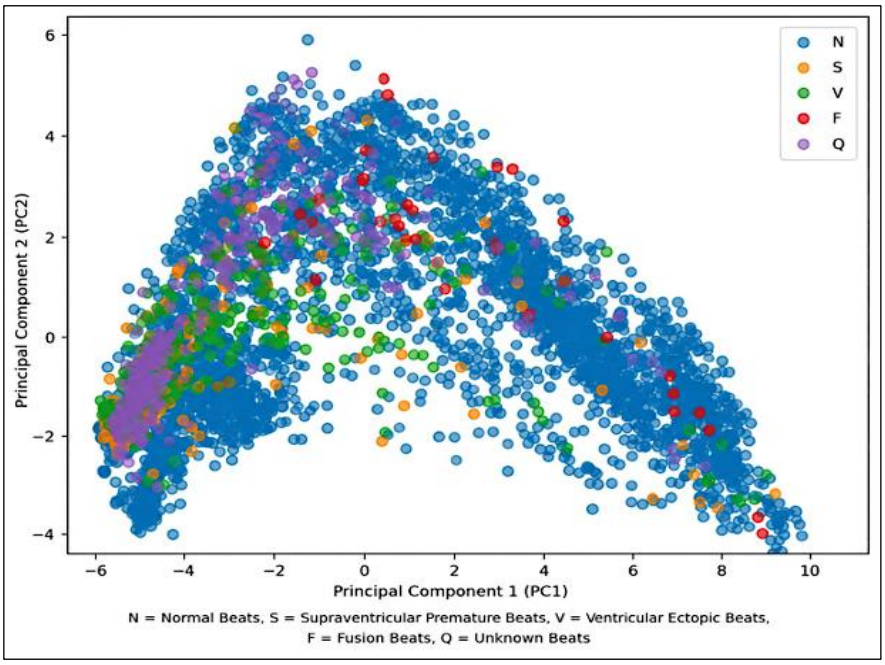

Figure 9: Visualization of ECG Feature Embeddings Generated by the Suggested Hybrid VIT–TFN Classifier using Principal Component Analysis

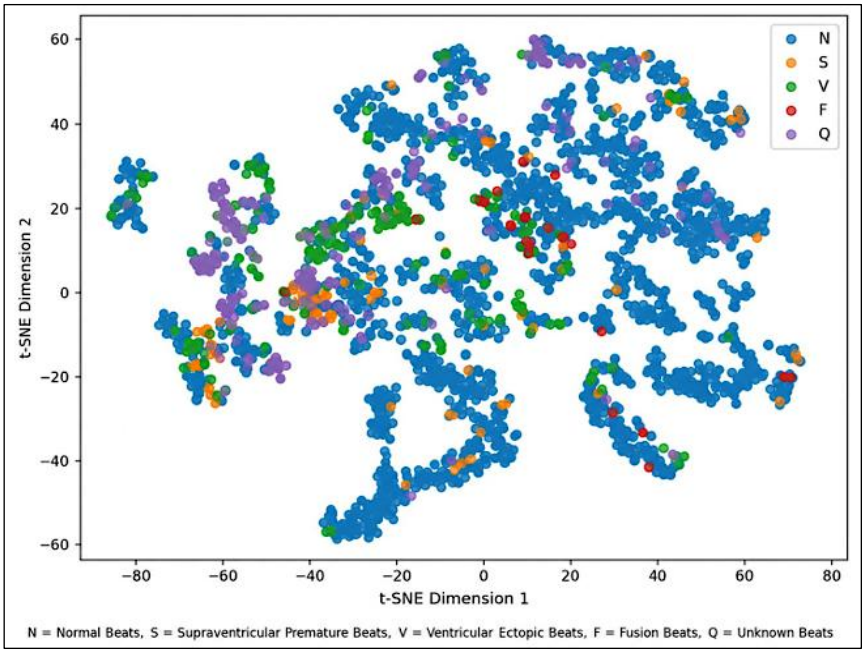

Figure 10: ECG Feature Embeddings Generated by the Suggested Hybrid VIT–TFN Classifier are Visualized using t-Distributed Stochastic Neighbor Embedding

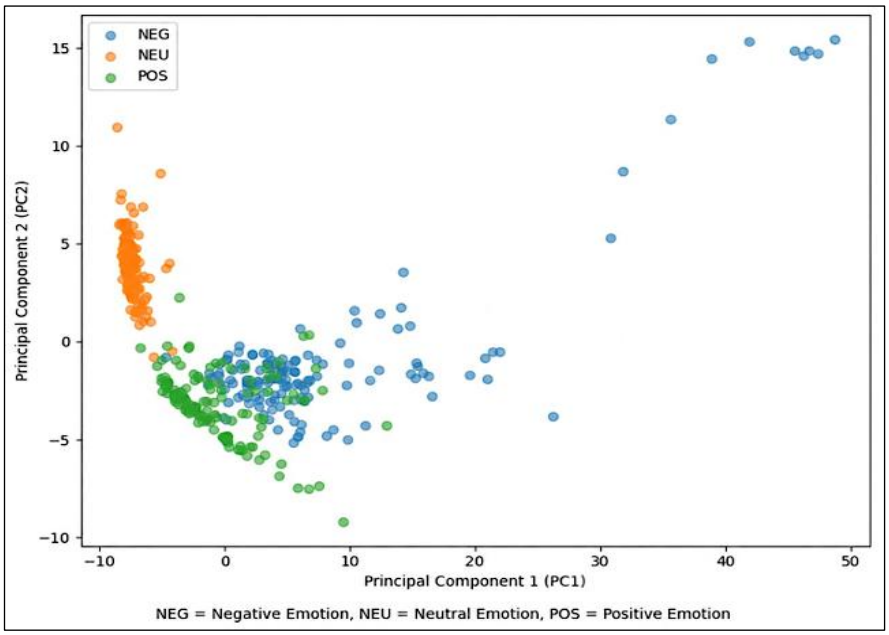

Figure 11: Visualization of EEG Feature Embeddings Produced by the hybrid GCN–GAT Classifier using Principal Component Analysis

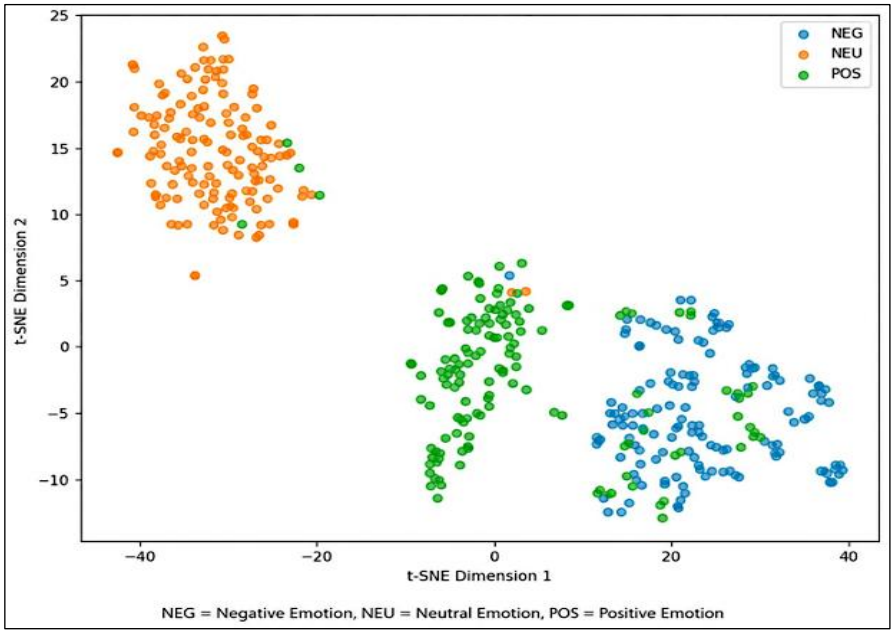

Figure 12: EEG Feature Embeddings Generated by the Suggested Hybrid GCN–GAT Classifier is Visualized using t-Distributed Stochastic Neighbor Embedding

Comparative Analysis

Ensemble networks built on AlexNet, GoogLeNet, VGG16 and SqueezeNet show the efficacy of deep CNN–based feature extraction in ECG classification, reporting an accuracy of 97.175% (1). Competitive performance has also been demonstrated by hybrid CNN models that incorporate temporal and frequency-domain information through additional handcrafted or transformed features. For instance, a CNN–based method utilizing hybrid features achieved 92%, while another sophisticated CNN–based arrhythmia classification framework reported 91.8% (2, 3). Multiscale residual CNN architectures obtained 99.6% accuracy among the best performing ECG techniques (4), while a CNN in conjunction with bidirectional long short-term memory and weighted loss achieved 99.33% accuracy (10). Furthermore, a deep CNN model showed 99.68% accuracy (11) and transfer learning using pretrained ResNet-18 and 2D CNN–

Short-Time Fourier Transform achieved 90.8% accuracy (12). Our proposed ViT–TFN hybrid model, which combines Temporal Fusion Network and Vision Transformer concepts (20, 21), achieved 92.14% on the MIT-BIH Arrhythmia Dataset (8), demonstrating competitive performance in a low-resource setting. The comparative analysis findings are displayed in Figure 13.

In EEG classification, CNN–Long Short-Term Memory based approaches have demonstrated strong performance, with the ResNet-152-based framework reporting 98% accuracy (5). Transformer-based techniques have also shown competitive effectiveness, as the AMDET model achieved 94.22% accuracy (6) and the spatiotemporal attention model with graph-smooth signals reached 95% (7). Among later EEG models, the CNN–Bidirectional Long Short-Term Memory with attention attained 99% accuracy (13), the residual-structure fusion Bidirectional Long Short-Term Memory also performed strongly (14), the spatial-temporal transformer framework reported 95.83% (15) and the Differential Entropy–CNN–Bidirectional Long Short-Term Memory multi-fusion model achieved 94.86% (16). Our proposed GCN–GAT hybrid model achieved 94.15% on the EEG Brainwave Dataset, confirming that the method remains competitive and suitable for low-resource settings. The comparative analysis is shown in Figure 14.

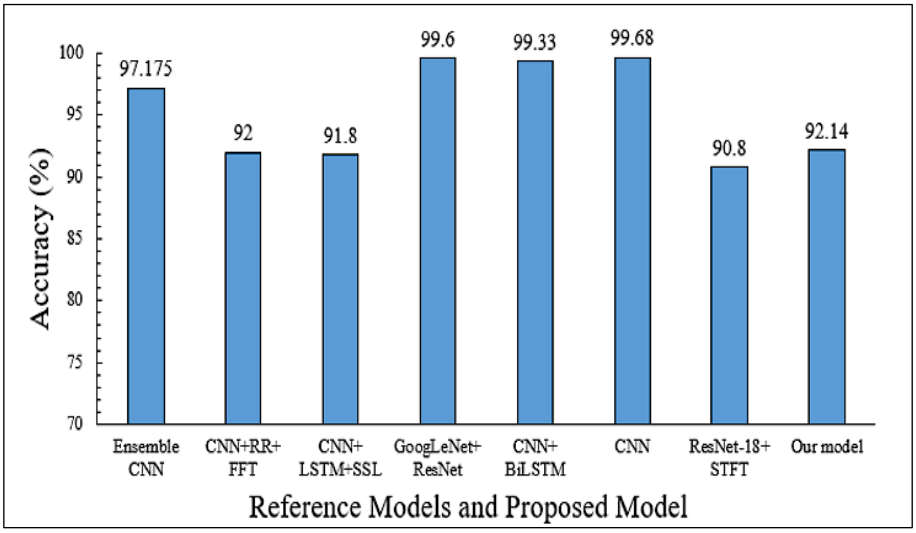

Figure 13: Comparative Evaluation of The Suggested Hybrid Model and Benchmark Architectures for ECG Classification Performance

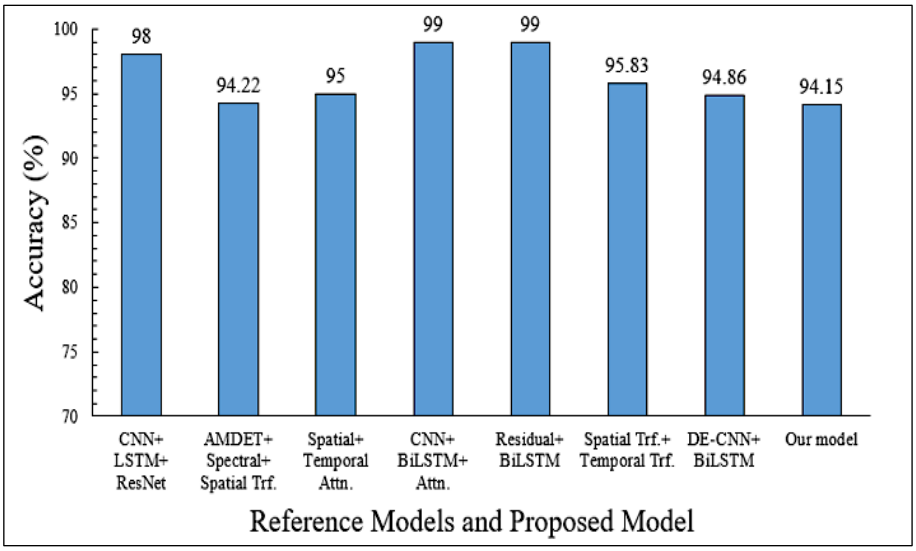

Figure 14: Comparative Evaluation of The Suggested Hybrid Model and Benchmark Architectures for EEG Classification Performance

Conclusion

This study proposed two low-resource deep learning pipelines for biological signal classification that used hybrid neural architectures, generative augmentation and self-supervised representation learning. The ECG pipeline uses a Vision Transformer in conjunction with a Temporal Fusion Network to capture both local waveform morphology and longer-term temporal dependencies, whereas the EEG pipeline uses a Graph Convolutional Network with a Graph Attention Network to model spatial relationships between electrodes and temporal dynamics in the signal. These models demonstrated that well-designed hybrid architectures can produce competitive ECG and EEG classification performance in environments with limited labelled data and limited computing resources. They achieved strong overall accuracies and class-wise performance across the MIT-BIH arrhythmia and EEG Brainwave datasets.

Concurrently, a number of restrictions were found. First, hardware and runtime constraints limited the complexity of architectures that could be trained and implemented. This likely limited performance compared with more resource-intensive models and made the combination of self-supervised learning and GAN-based augmentation computationally demanding for real-time embedded or point-of-care scenarios. By raising each class to the closest multiple of 500 samples, GAN-based augmentation improved class balance, but it only partially alleviated class overlap and the inherent difficulty of unusual classes. Second, under realistic acquisition conditions, noise-induced instability or signal variability could not be entirely eliminated by typical preprocessing and biomedical signals remained vulnerable to noise and artifacts, particularly in EEG recordings. Third, the transformer-based and graph-based components of the suggested hybrid architectures still have poor interpretability and explainability, which could have an impact on clinical trust and acceptance. Fourth, the framework is not meant to claim full applicability for extremely complicated or highly diverse clinical data without additional validation because it was purposefully optimized for small datasets, ultralightweight architecture design and low-resource execution in low-end systems. Lastly, there are unanswered problems about tight real-time performance and wider practical deployment since latency and throughput characteristics were not systematically measured.

These restrictions point to many avenues for further research. More hardware-aware and computationally efficient designs that preserve classification accuracy while lowering training and inference costs should be the focus of future research. For ECG and EEG, further development of GAN-based augmentation and self-supervised contrastive learning may enhance minority-class performance and representation quality while lowering the need for laborious manual annotation. To increase resilience to artifacts and non-stationary noise, more sophisticated denoising, adaptive filtering and spectral-temporal fusion techniques should be investigated. To increase the transparency of model decisions, more robust explainability mechanisms such as attribution techniques, saliency analysis and attention visualization should be added concurrently. To improve real-world applicability and translational value, future research should also incorporate systematic latency and throughput profiling, as well as external validation on bigger, more varied and clinically significant ECG and EEG datasets.

Abbreviations

CNN: Convolutional Neural Network, ECG: Electrocardiogram, EEG: Electroencephalogram, GAN: Generative Adversarial Network, GAT: Graph Attention Network, GCN: Graph Convolutional Network, MoCo: Momentum Contrast, SimCLR: Simple Framework for Contrastive Learning of Visual Representations, TFN: Temporal Fusion Network, ViT: Vision Transformer.

Acknowledgement

I would like to express my sincere gratitude to my co-author and research supervisor, Dr. Gampala Veerraju, for his continuous guidance, encouragement and valuable insights throughout this work. This research was carried out as part of my academic studies in the Department of Computer Science and Engineering at Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur andhra Pradesh, India.

Author contributions

Thota Leela Venkata Umamahesh: Conceptualization, Methodology, Software, Investigation, Formal analysis, Writing – Original Draft, Veerraju Gampala: Supervision, Writing – Review and Editing and guidance in comparative analysis and manuscript refinement.

Conflict of Interest

I confirm that neither I nor my co-author have any conflicts of interest related to this research.

Data Availability

I confirm that all datasets used in this study are publicly available. The ECG data were obtained from the MIT-BIH Arrhythmia Database and the EEG data were obtained from the EEG Brainwave Dataset. These datasets can be accessed through their respective public repositories.

Declaration of Artificial Intelligence (AI) Assistance

I confirm that AI-assisted tools were used solely for limited refinement purposes during the preparation of this manuscript. Perplexity was consulted only to improve technical clarity and Grammarly was used exclusively for language polishing. The core writing of the research, including the ideas, methodology, models, literature review, analysis, results and conclusions, was carried out entirely by me and my co-author without the use of generative AI. The authors take full responsibility for the content's originality, interpretation and accuracy.

Ethics Approval

I confirm that this study was conducted using only publicly available datasets and did not involve any direct interaction with human participants. Therefore, formal ethical approval was not required.

Funding

We confirm that this research did not receive any external financial support and was carried out by me and my co-author as part of our academic curriculum.

References

1. Malleswari PN, Odugu VK, Subrahmanyeswara Rao TJV, Aswini TVNL. Deep learning-assisted arrhythmia classification using 2-D ECG spectrograms. EURASIP J Adv Signal Process. 2024;2024:104. doi: 10.1186/s13634-024-01197-1
2. Kiranyaz S, Ince T, Gabbouj M. Real-time patient-specific ECG classification by 1-D convolutional neural networks. IEEE Trans Biomed Eng. 2016;63(3):664–675. doi: 10.1109/TBME.2015.2468589
3. Sattar S, Mumtaz R, Qadir M, Mumtaz S, Khan MA, De Waele T, De Poorter E, Moerman I, Shahid A. Cardiac arrhythmia classification using advanced deep learning techniques on digitized ECG datasets. Sensors (Basel). 2024;24(8):2484. doi: 10.3390/s24082484
4. Zhang F, Li M, Song L, Wu L, Wang B. Multi-classification method of arrhythmia based on multi-scale residual neural network and multi-channel data fusion. Front Physiol. 2023;14:1253907. doi: 10.3389/fphys.2023.1253907
5. Chakravarthi B, Ng SC, Ezilarasan MR, Leung MF. EEG-based emotion recognition using hybrid CNN and LSTM classification. Front Comput Neurosci. 2022;16:1019776. doi: 10.3389/fncom.2022.1019776
6. Xu O, Du Y, Li L, Lai H, Zou J, Zhou T, Xiao L, Liu L, Ma P. AMDET: Attention based multiple dimensions EEG transformer for emotion recognition. IEEE Trans Affect Comput. 2023;15(3):1067–1077. doi: 10.1109/TAFFC.2023.3318321
7. Sartipi S, Torkamani-Azar M, Cetin M. A hybrid end-to-end spatiotemporal attention neural network with graph-smooth signals for EEG emotion recognition. IEEE Trans Cogn Dev Syst. 2023;16(2):732–743. doi: 10.48550/arXiv.2307.03068
8. Moody GB, Mark RG. The impact of the MIT-BIH Arrhythmia Database. IEEE Eng Med Biol Mag. 2001;20(3):45–50. doi: 10.1109/51.932724
9. Bird JJ, Faria DR, Manso LJ, Ekárt A, Buckingham CD. A deep evolutionary approach to bioinspired classifier optimization for brain-machine interaction. Complexity. 2019;2019:4316548. doi:10.1155/2019/4316548
10. Yang M, Liu W, Zhang H. A robust multiple heartbeats classification with weight-based loss based on convolutional neural network and bidirectional long short-term memory. Front Physiol. 2022;13:982537. doi: 10.3389/fphys.2022.982537
11. Raza MA, Anwar M, Nisar K, Ibrahim AAA, Raza UA, Khan SA, Ahmad F. Classification of electrocardiogram signals for arrhythmia detection using convolutional neural network. Comput Mater Contin. 2023;77(3):3817–3834. doi: 10.32604/cmc.2023.032275
12. Hannun AY, Rajpurkar P, Haghpanahi M, Tison GH, Bourn C, Turakhia MP, Ng AY. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. Nat Med. 2019;25(1):65–69. doi: 10.1038/s41591-018-0268-3
13. Huang Z, Ma Y, Wang R, Li W, Dai Y. A model for EEG-based emotion recognition: CNN-Bi-LSTM with attention mechanism. Electronics. 2023;12(14): 3188. doi: 10.3390/electronics12143188
14. Xu Y, Gao Y, Zhang Z, Du S. Emotional recognition of EEG signals utilizing residual structure fusion in bi-

directional LSTM. Complex Intell Syst. 2025;11:62. doi: 10.1007/s40747-024-01682-y
15. Li M, Yu P, Shen Y. A spatial and temporal transformer-based EEG emotion recognition in VR environment. Front Hum Neurosci. 2025;19: 1517273. doi: 10.3389/fnhum.2025.1517273
16. Cui F, Wang R, Ding W, Chen Y, Huang L. A novel DE-CNN-BiLSTM multi-fusion model for EEG emotion recognition. Mathematics. 2022;10(4):582. doi: 10.3390/math10040582
17. Jaiswal A, Babu AR, Zadeh MZ, Banerjee D, Makedon F. A survey on contrastive self-supervised learning. Technologies. 2021;9(1):2. doi:10.3390/technologies9010002
18. Le-Khac PH, Healy G, Smeaton AF. Contrastive representation learning: A framework and review. IEEE Access. 2020;8:193907–193934. doi:10.1109/ACCESS.2020.3031549
19. Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y. Generative adversarial nets. In: Advances in Neural Information Processing Systems 27 (NeurIPS 2014). 2014; p. 2672–2680. doi: 10.48550/arXiv.1406.2661
20. Khan S, Naseer M, Hayat M, Zamir SW, Khan FS, Shah M. Transformers in vision: A survey. ACM Comput Surv. 2022;54(10s):1–41. doi:10.1145/3505244
21. Lim B, Arık SÖ, Loeff N, Pfister T. Temporal fusion transformers for interpretable multi-horizon time series forecasting. Int J Forecast. 2021;37(4):1748–1764. doi: 10.1016/j.ijforecast.2021.03.012
22. Zhou J, Cui G, Hu S, *et al*. Graph neural networks: A review of methods and applications. AI Open. 2020;1:57–81. doi: 10.1016/j.aiopen.2021.01.001
23. Wu Z, Pan S, Chen F, Long G, Zhang C, Yu PS. A comprehensive survey on graph neural networks. IEEE Trans Neural Netw Learn Syst. 2021;32(1):4–24. doi:10.1109/TNNLS.2020.2978386
24. Fawcett T. An introduction to ROC analysis. Pattern Recognit Lett. 2006;27(8):861–874. doi: 10.1016/j.patrec.2005.10.010
25. Jolliffe IT. Principal component analysis. 2nd ed. New York: Springer. 2002. doi: 10.1007/978-1-4757-1904-8
26. van der Maaten L, Hinton G. Visualizing data using t-SNE. J Mach Learn Res. 2008;9:2579–2605. https://www.jmlr.org/papers/v9/vandermaaten08a.html

How to Cite: Umamahesh TLV, Gampala V. Towards Low-resource Deep Learning for Biomedical Signal Classification: An ECG and EEG Based Approach. Int Res J Multidiscip Scope. 2026;7(3):216-231.
DOI: 10.47857/irjms.2026.v07i03.010171