

Received 5 October 2025, accepted 2 November 2025, date of publication 10 December 2025, date of current version 22 January 2026.

Digital Object Identifier 10.1109/ACCESS.2025.3642249



RESEARCH ARTICLE

Emotion Recognition From EEG Signals Using a Hybrid Deep Learning Approach

NATASHA PIEDRABUENA^{ID1}, (Student Member, IEEE), WAFA M. ELMANNAI², (Member, IEEE), AND YI WANG², (Senior Member, IEEE)

¹Department of Computer Science, Manhattan University, Riverdale, NY 10471, USA

²Department of Electrical and Computer Engineering, Manhattan University, Riverdale, NY 10471, USA

Corresponding author: Wafa M. Elmannai (wafa.elmannai@manhattan.edu)

ABSTRACT In recent years, the rise of advanced machine learning techniques has led to an increase in research on brain-computer interfaces. It's considered a multifaceted challenge to develop applications that can analyze EEG signals into domains such as mental health assessment, real-time emotional monitoring, and personalized wellness recommendations. In the proposed approach, it will display the effectiveness of a hybrid deep learning neural network architecture that integrates Bidirectional Gated Recurrent Units (BiGRU) and Bidirectional Long-Short-Term Memory (BiLSTM) layers for the classification of multiple emotions based on EEG data. The publicly available EEG Brainwave Dataset: Feeling Good, provided by the School of Engineering and Applied Sciences at Aston University, Birmingham, was used in this study. The central hypothesis we propose combines the BiGRU and BiLSTM layers, enhancing the model's capacity to capture temporal dependencies and manage the complex sequential patterns inherent in EEG signals. To evaluate this hypothesis, three distinct model architectures were assessed: BiGRU-only model, BiLSTM-only model, and combined BiGRU + BiLSTM model. These models were then trained and tested on a labeled dataset of EEG recording. The results demonstrated that the hybrid BiGRU + BiLSTM model achieved the highest classification accuracy, reaching approximately 98.59%, compared to 98.13% for the BiGRU model and 98.1% for the BiLSTM model. This result suggests potential future applications in assistive technologies, where individuals with severe motor impairments—such as those with ALS or spinal cord injuries—could benefit from EEG-based emotion detection systems to support non-verbal communication and adaptive interaction. Further analysis of precision, recall, and F1 scores in the three emotion classes provided deeper insights into the sensitivity and specificity of the model. This underscores its potential for implementation in real-time affective computing applications in future works.

INDEX TERMS Emotion in brain-computer interaction, bidirectional long- and short-term memory (BiLSTM), bidirectional gated recurrent unit (BiGRU), machine learning.

I. INTRODUCTION

In recent years, there has been a significant leap forward in the way humans interact with machines mainly when it comes to Brain Computer Interface (BCI). BCI systems enable direct interaction between the human brain and external devices without the use of traditional motor pathways. BCI has its roots in neuroscience and engineering with the potential to transform lives through the unconventional method of

The associate editor coordinating the review of this manuscript and approving it for publication was Hasan S. Mir.

communication via EEG biosignals. For instance, individuals with severe motor impairments such as those affected by amyotrophic lateral sclerosis (ALS), progressive muscular atrophy (PMA), spinal cord injuries, etc... depend on various systems to be able to interact with others. With the use of biosignals and BCIs it can unlock control over computers, prosthetics, and even robotic systems. Therefore providing alternatives for communication and offering a new level of independence for those with physical limitations.

The human brain is a highly complex organ that operates through intricate interactions between neurons by sending

electrical signals through their axons, that is then converted to chemical signals that are carried by neurotransmitters through the synapse to the other neurons. These neurons communicate via electrical impulses also known as action potentials which travel along the neural pathway to transfer information throughout the body's nervous system. Similarly, the human brain's ability to process and transmit information is adapted into artificial neural networks in machine learning. For instance, neurons inside the brain form connected networks that process information as sensor inputs and generate responses based on actions. While, artificial neural connections form layers that also process data and produce outputs. Additionally, both systems learn through the strengthening or weakening of connections between nodes(neurons) also known as synaptic plasticity in the brain. This adaptive mechanism allows the human brain to learn from experiences and adjust its responses over time which is similar to how artificial neural networks adjust weights during training to improve performance on a given task. The development of advanced artificial neural networks has inspired researchers to develop more advanced BCIs that can learn and adapt in real-time, improving their accuracy and responsiveness. Additionally, the parallel between neurons and neural networks highlights the importance of the brain's role as both a biological information processor and a model for developing advanced computational systems.

When it comes to BCI technology there are various types of systems including and not limited to invasive, semi-invasive and non-invasive approaches. Invasive BCIs involve direct implantation of electrodes into the brain tissue which can provide high-resolution signals and more precise control (reduced noise in signals). Meanwhile, semi-invasive BCIs strike a balance by placing electrodes on the scalp's surface, reducing surgical complication compared to invasive BCIs. Alternatively, non-invasive BCIs is a non-surgical approach more commonly known as Electroencephalography (EEG) due to its ability to record electrical activity from the brain via scalp electrodes. The signals are generated by synchronous activity of large groups of neurons and can be classified into various frequency bands such as Δ Delta, Θ Theta, α Alpha, β Beta, and γ Gamma. These frequency bands are then associated with different cognitive states. For instance, alpha waves are linked to a relaxed state, beta waves are associated with concentration and active learning. On the other hand, EEG signals are also considered unreliable due to interference with external noise that can affect the signal quality. The inherent noise can make it difficult to distinguish between relevant and irrelevant activity within the raw biosignal that was collected. To counteract these limitations it requires sophisticated signal processing techniques to extract the necessary information needed for the system. This challenges the pre-processing steps in machine learning systems as the raw data needs to be cleaned and processed into an acceptable format for the analysis step.

Among the diverse applications of BCI, emotion recognition has emerged as a particularly promising area, as

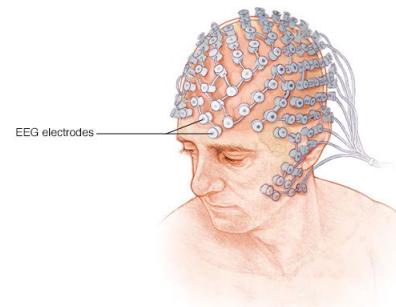
emotional states strongly influence decision-making, learning, and communication. EEG-based emotion recognition can support clinical applications such as monitoring patients with neurodegenerative disorders, as well as non-clinical uses like adaptive gaming, human-computer interaction, and affective computing systems. Establishing accurate recognition of emotions from EEG signals is therefore a critical step in advancing practical BCI systems, directly connecting neuroscience insights to real-world applications.

This proposal aims to address these limitations by introducing a comprehensive design that integrates advanced neural network architectures to improve its accuracy. The goal is to develop a hybrid deep learning model that combines Bidirectional Long Short-Term Memory (BiLSTM) and Bidirectional Gated Recurrent Unit (BiGRU) layers to accurately classify and interpret distinct emotional states. The combination is expected to outperform standalone models, as BiLSTM captures long-term dependencies while BiGRU efficiently models short-term dynamics, making their integration complementary. While alternative architectures such as Transformers and CNN-LSTM hybrids have been applied in related domains, there has been limited exploration of BiLSTM-BiGRU hybrids for EEG-based emotion recognition. This study therefore aims to address that gap by leveraging their complementary strengths within a hybrid framework. By implementing a machine-learning-based EEG emotion recognition framework, this project envisions transformative applications in medical fields, particularly for individuals with neurodegenerative diseases who face communication challenges.

II. BACKGROUND

A. ELECTROENCEPHALOGRAM

An electroencephalogram (EEG) is a device utilized to record electrical activity throughout the human brain via electrodes placed on the scalp as seen in Fig 1. This device can be used to detect and monitor various neurological conditions like epilepsy, sleep disorders, brain tumors, and head injuries.



© MAYO FOUNDATION FOR MEDICAL EDUCATION AND RESEARCH. ALL RIGHTS RESERVED.

FIGURE 1. Electroencephalogram (mayo clinic) [16].

However, emotion recognition using EEG signals remains challenging due to susceptibility to noise, inter-subject variability, and other sources of uncertainty that can hinder

reliable classifications. In this study we will utilize data collected from a muse EEG device [11], [12].

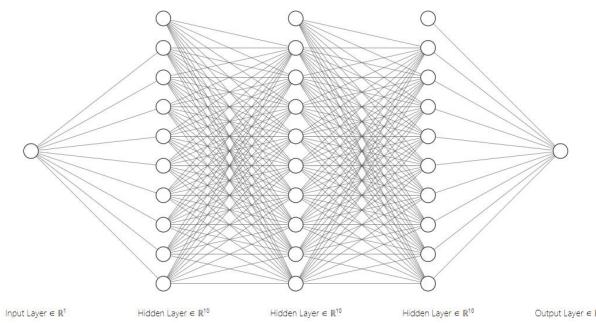


FIGURE 2. Example diagram of a fully-connected deep neural network [13].

B. MACHINE LEARNING ALGORITHM

There are machine learning algorithms such as deep neural networks (DNNs) that are designed to mimic complex patterns similar to the structure and functionality of the human brain [15] as seen in Fig 2. These neural networks are made up of interconnected neurons that are able to process information [15]. For instance, each neuron receives input and applies activation functions & weights to then pass the output as subsequent layers [15]. Then these layered architecture allows DNNs to learn hierarchical representations of data, capturing intricate patterns and relationships [15]. In the context of EEG emotion recognition, DNNs are particularly well suited for handling the multi-dimensional nature of EEG raw data. They integrate spectral, spatial, and temporal dimensions into a unified representation for emotion classification. DNNs can also optimize performance based on tasks such as image recognition, natural language processing, and etc... [15].

III. DATASET

An overview of the dataset utilized in this proposal to assist in achieving an accurate emotion classification using EEG signals is the EEG brainwave dataset from Aston University [12]. It's a widely recognized and publicly available dataset for emotion analysis [12]. However, the limited sample size constitutes a major limitation that could restrict the generalizability of this investigation.

In the experiment various film clips were selected to rank positive, neutral, and negative emotions based on specific criteria [12]. Each state clip lasted about 3 minutes, with well-edited content to maximize emotional impact [12]. The experiment involved 6 minute rest between clips [12]. The sequence of clips was arranged to avoid consecutive emotions of the same type throughout the experiment [12]. Participants reported their emotional responses through a questionnaire after each clip [12]. The statistical extraction was transformed into waves that can be described in a temporal fashion [12].

In this proposal the EEG brainwave dataset developed by Aston university was used for developing a sophisticated hybrid neural network system aimed to improve current emotion classification systems.

A. DATASET INFORMATION

The dataset consists of 36 minutes of EEG data collected from two subjects (one male and one female, aged ≥ 20) while they watched six film clips that trigger emotional responses. The clips used are seen in Table 1. Each clip generated 60 seconds of data, resulting in a total of 12 minutes for the emotional states and an additional 6 minutes of neutral data collected before the emotional responses to avoid contamination. The data was recorded at a frequency of 150 Hz, yielding 324,000 data points.

Participants were instructed to watch the clips without making conscious movements to minimize the influence of Electromyographic (EMG) signals on the EEG data [11]. Facial expressions during the clips were noted as they reflect real-world emotional responses and were included in the classification model [11]. The dataset includes a total of 2400 statistical features that allows for a detailed analysis of the emotional states elicited by the film clips [11].

TABLE 1. Source of film clips utilized as stimuli for EEG data [11].

Stimulus	Valence	Studio	Year
Marley and Me	Neg	Twentieth Century Fox, etc.	2008
Up	Neg	Walt Disney Pictures, etc.	2009
My Girl	Neg	Imagine Entertainment, etc.	1991
La La Land	Pos	Summit Entertainment, etc.	2016
Slow Life	Pos	BioQuest Studios	2014
Funny Dogs	Pos	MashupZone	2015

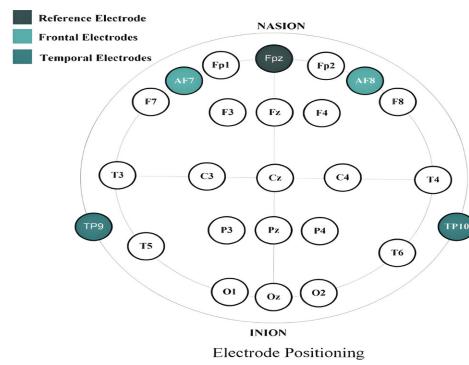


FIGURE 3. Electrode distribution on scalp [14].

B. SIGNAL ACQUISITION & FREQUENCY RESAMPLING

The signal acquisition is through utilizing the Muse EEG headband which is recorded via dry electrodes on TP9, AF7, AF8, TP10 as seen on Fig 3 & Fig 4.

Each electrode captures brainwave activity which is downsampled to 150 Hz to reduce computational complexity

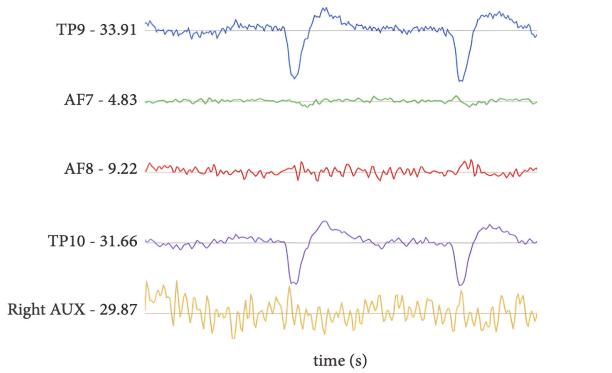


FIGURE 4. Raw EEG data stream from the muse EEG headband (Y-axis represents measured brainwave activity in microVolts (mV) & X-axis is the time the data is recorded [12].

while retaining critical signal information. It was further extracted to 2400 features through a sliding window of 1 second in a $t = 0$ to $t = 5$ interval [12]. This total was obtained by applying multiple statistical and spectral descriptors (e.g., mean, variance, skewness, kurtosis, entropy, FFT coefficients) to each segmented window, resulting in approximately 2400 features overall. EEG signals in general are often decomposed into Delta δ (0.5–4 Hz), Theta θ (4–8 Hz), Alpha α (8–13 Hz), Beta β (13–30 Hz), and Gamma γ (30–50 Hz) each associated with different cognitive and emotional states [11], [12].

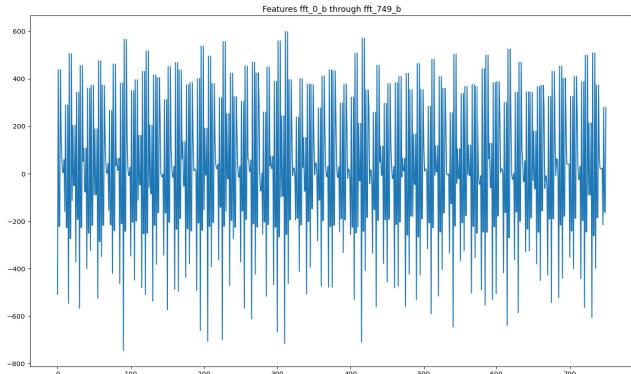


FIGURE 5. Frequency domain representation of EEG data (fft_0_b through fft_749_b).

C. FEATURE EXTRACTION

In the initial investigation, they employed a sliding window technique to segment continuous EEG data into overlapping one-second time intervals with a 0.5-second increment [12]. This approach was to capture dynamic changes in brain activity, essential for analyzing mental states [12]. For instance, for each window 2,549 features were initially extracted and analyzed utilizing variety of statistical features including (mean, standard deviation, skewness/kurtosis, min/max, median, euclidean distance between min/max, Shannon

Entropy, Log-energy entropy, fast fourier's transforms and several derived metrics (like derivatives and entropies) [11], [12]. The high-dimensional EEG data was then applied to four selection methods OneR, BayesNet, InfoGain, and Symmetrical Uncertainty—to reduce the features to a more manageable subset [11], [12]. The breakout varies by algorithm: OneR (52 features), BayesNet (67), InfoGain (63), and Symmetrical Uncertainty (72) [12]. These selected features as seen in Fig 5. improved classification efficiency while preserving the information necessary for predicting emotional states [12].

IV. RELATED WORKS

The progression of research on human-computer interaction(HCI) led to emphasis on emotion recognition utilizing brain-computer interfaces(BCI) [1]. This led to the inspiration from existing work on Electroencephalogram emotion recognition classification systems utilizing deep learning methods. One of the proposed approaches [1] involved in BiLSTM based deep feature extraction and classification into binary classes. Firstly, the raw data would be extracted onto a pre-processing step in which the baseline signal can be excluded to improve the performance [1]. Then the raw EEG data is sampled based on the frequency Hz into its specific channel [1]. Once the pre-processing step is completed it's inserted onto the BiLSTM deep feature extraction [1]. When brain activity shifts over time, temporal features are reflected through the time dimension [1]. BiLSTM architecture can learn temporal information of time-domain signals due to the temporal dependencies from EEG data [1]. Then the data is organized to classify emotion into binary classes, the data is randomly divided into 90% training data and 10% test data [1].

Another method that was explored utilized CNN-LSTM architecture to identify spatial and temporal features [2]. Initially, the raw data is split into sections of similar length and separated based on channels [2]. It excludes the baseline signal from the matrix by taking the difference between raw data matrix and baseline mean matrix [2]. Once extracted the matrices are merged into a large matrix of the same size as the raw EEG signal, this is applied to enhance performance [2]. The pre-processed data is in 1D signal format, hence it can be directed into an LSTM module to extract the temporal features [2]. Then the spatial features are extracted by converting the 1D signal into 2D frames [2].

Another deep learning model utilized convolutional neural network(CNN) and graph convolutional network(GCN) [3]. For signal pre-processing the fast fourier transform(FFT), differential entropy, hjorth parameters, applied welch's method was used to form power spectrum to find the power spectral density that compares the signal relative to the frequency domain [3]. The proposed model is a 1D CNN and GCN, with a total of three convolution layers [3]. The training epoch was set to 300 for the training sessions, the trend on the proposed model loss/accuracy graph reached a validation

score above 95% [3]. Accuracy of the proposed hybrid neural network (CNN+GCN) is 97.7% and weighted F1 score is 98% [3].

This proposed method adopts a CNN architecture that consists of three consecutive layers of (3×3) kernel size to operate in both spatial and temporal dimensions to extract useful information for the classification phase [4]. The signal processing methods utilized included MEMD, IMF, HHT, Marginal Hilbert Spectrum extractions [4]. For the input preparation, the signal must be divided into sections of 1 second long, then the signal is agitated from all EEG channels to construct a 2D image map [4]. To convert the 2D image into 3D input for the 3D CNN, the 2D image map must be stacked in three consecutive segments [4]. Comparison between the 3D-CNN with raw EEG data has a valence accuracy of 82.32% and arousal accuracy of 84.12% [4]. While the 3D-CNN with the pre-processed data reach an valence accuracy of 89.25% and an arousal accuracy of 86.23% [4]. This unique technique of applying non-linear analysis to multi-channel EEG let to effective spatial-temporal feature representation and higher accuracy compared to other based linear models such as LSTM [4].

To achieve a feasible system for learning in temporal and spatial domains a wavelet-based detection on multi-channel EEG was implemented and a hybrid 2D CNN-LSTM model was proposed. The proposed 2D CNN-LSTM model architecture consists of four various convolutional blocks, with each block containing a convolutional layer of size (3×3) kernel [5]. The reason for this is due to the fact that the convolutional layer is performed in a spatial domain in which each input has to be represented as a feature block of $E \times W \times C$ [5]. The CNN layer investigates the spatial features by performing the convolutional operation on the spatial domain, while the LSTM layer is used to extract the temporal features from the temporal domain [5]. The proposed architecture includes the four convolutional blocks with convolutional layers, 2D max pooling layer, LSTM layer flattened, and three fully connected layers [5]. In this framework the DEAP dataset was utilized, measuring the valence and arousal levels to establish various emotions [5]. The valence accuracy is approximately 94.36% and arousal accuracy is 94.07% [5]. Overall accuracy over an evaluation of 5-fold is approximately 94.41% [5].

In this proposed emotion detection system it suggested an advanced CNN model using wavelet transform and classification on the SEED dataset [6]. Utilizing discrete Wavelet Transform feature extraction reduces noise and maintains the key information from the raw signals [6]. In the investigation the process included was a pre-processing phase, feature signal extraction, feature signal selection, classification, and pattern recognition [6]. The implementation of the CNN model consists of pairs of convolutional pooling layers and an output layer [6]. The accuracy across 12 channels overall is 95.67% [6].

Another method that was explored is a hybrid model BiGRU and LSTM based ensemble learning emotion detection system [7]. In the CNN model architecture it consists of four convolutional blocks and the network is tied with three dense connective layers [7]. The LSTM model consists of four layers and the output is tied with connective layers [7]. The hybrid model improves the performance of the spatial and temporal features, leading to a higher accuracy in the emotion detection system [7]. In the models each trained for 60 epoch with a batch size of 64 [7]. The model architecture consists of data input, CNN model, LSTM model, hybrid model and ensemble model [7]. The accuracy of the proposed model emotion detection system is approximately 97.16% with the ensemble stack [7].

Another proposal suggested utilizing CNN and BiLSTM hybrid models for a multimodal emotion recognition based system [8]. The features are extracted and cleaned the two modalities are merged to express the final emotional state [8]. The proposed model will be assessed utilizing the DEAP dataset, a publicly available EEG dataset [8]. However, in this proposal the model is compared to single modal systems rather than multimodal systems [8]. The 2D EEG data frames after the pre-processing step are input into the CNN and the audio signals are input into BiLSTM [8]. For the BiLSTM layer the structure contains 128×2 nodes and the CNN is composed of 3 convolution layers of size 32, 64, and 128 filters [8]. Once the feature is extracted from both structures it is inserted into the softmax layer for the final emotional state [8]. The CNN+BiLSTM the arousal accuracy is 93.18% and the valence is 93.20% [8]. The proposed framework has led to a new multimodal emotion recognition architecture that utilizes CNN+BiLSTM that can learn both spatial and temporal characteristics from EEG data [8].

In this proposal, the EEG-based emotion recognition is developed using graph neural networks(GNN) [9]. The dataset utilized is the publicly available SEED and SEED-IV [9]. The classification setting in this proposal conducts both subject-dependent and subject-independent for both SEED and SEED-IV [9]. For the SEED dataset the experimental settings to evaluate the RGNN (Regularized Graph Neural Network) was to use subject-dependent classification [9]. Then to train the model the first 9 segments are utilized [9]. The remaining trials are used for validation [9]. While the SEED-IV subject-independent is applied as it performs marginally better compared to the SEED [9]. In the confusion matrix for SEED the subject-independent accuracy across negative is 90.4%, neutral is 97.60%, and positive is 94.80% [9]. While confusion matrix analysis for subject-dependent the accuracy across negative is 79.14%, neutral is 84.83% and positive is 91.67% [9]. However, in the confusion matrix for SEED-IV the subject-independent accuracy across negative is 75%, sad is 91.92%, fear is 71.85%, happy is 74.35% [9]. This demonstrates that the accuracy of subject-independent for SEED outperformed that of subject-dependent while

subject-independent for SEED-IV outperformed that of subject-dependent of the SEED-IV [9].

Another proposal for an EEG-based emotion detection system utilized a model that contains three main components CL + GAN + GNN [10]. The MAHNOB-HCI is a publicly available dataset that contains 27 subjects (ages between 19 to 40 years old) with the EEG and peripheral physiological sensors [10]. In the proposed model the trials were shuffled, 80% of the trials will be used for training and the remaining 20% is used for testing [10]. One of the main components for the proposed system is the contrastive learning (CL) algorithm architecture [10]. The second component is the generative adversarial network (GAN) and the third component is the GNN model [10]. When comparing the various benchmarks of the proposed CL+GAN+GNN in the DEAP dataset the valence accuracy is 64.84%, and arousal is 66.40% [10].

Another investigation for a mental emotional sentiment classification with EEG-based Brain-Computer interface proposed a random forest model [11]. The accuracy achieved by this proposed model is 97.89% [11]. The paper utilized a combination of testing various models with alternative versions of the pre-processed & feature extracted data [11]. Hence, the conclusion was that the random forest performed the highest accuracy and the multilayer perception(MLP) was the most consistent model with an accuracy of 94.89% [11].

In this research investigation a Deep Evolutionary (DEvo) approach to classify complex signals using bio-inspired computing methods uses evolutionary algorithms for feature selection and to optimize the neural network [12]. The investigation tested their models on these dataset which include mental state, emotional state, and the MindBigData digits dataset, comparing a DEvo optimized Multilayer Perceptron (MLP), Long Short-Term Memory (LSTM), an AdaBoosted DEvo MLP, and an AdaBoosted LSTM [12]. For the mental and emotional state datasets, the AdaBoosted LSTM model achieved the highest accuracies of 84.44% and 97.06% respectively [12]. While the AdaBoosted DEvo MLP was slightly less accurate at 79.7% and 96.23% respectively, it trained significantly faster [12]. On the MindBigData digits dataset, the AdaBoosted DEvo MLP was the most accurate, achieving 31.35% accuracy [12]. In this case, the LSTM models did not achieve useful results [12]. The authors concluded that the DEvo approach produced a model that achieved high classification ability with reduced resource usage [12].

V. METHODOLOGY

This methodology outlines the systematic approach to developing an EEG emotion-based classification utilizing deep learning techniques. In this investigation we propose an approach of implementing a hybrid neural network with Bidirectional Long Short-Term Memory (BiLSTM) and Bidirectional Gated Recurrent Unit (BiGRU) model for EEG-based emotion classification. The hybrid design combines BiLSTM and BiGRU layers to capture temporal structure at multiple scales. The BiLSTM layer models

long-range bidirectional dependencies, while the subsequent BiGRU layer focuses on short-term temporal fluctuations with lower computational overhead, allowing the network to more effectively separate subtle emotional patterns embedded in EEG time sequences. The strategy employed is to shuffle and partition the dataset, allocating 80% for training and 20% for validation. The proposed model and implementation will be further broken down in the subsequent sections.

A. DATASET PREPARATION

The EEG dataset comprises multi-channel recordings labeled with three emotion classes: **NEGATIVE**, **NEUTRAL**, and **POSITIVE**. Each EEG sample is represented as a time-series sequence of feature vectors

$$X = \{x_1, x_2, \dots, x_T\}$$

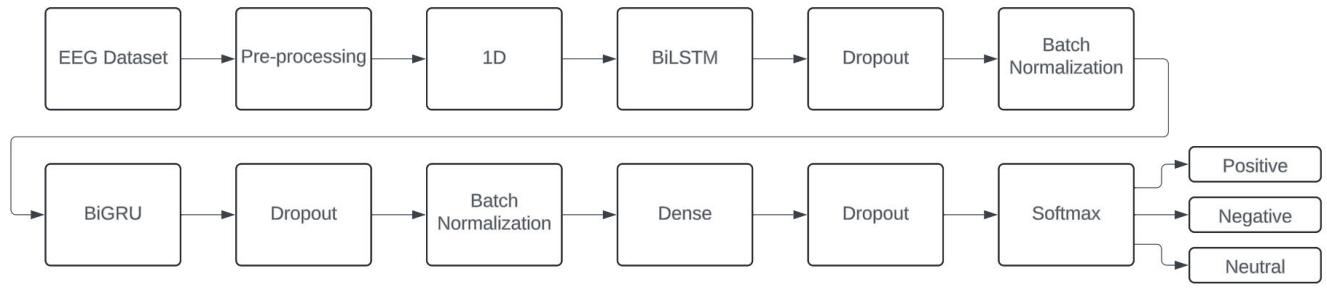
where T denotes the length of the sequence and $x_t \in \mathbb{R}^n$ is an n -dimensional feature vector at time t . To enhance convergence stability, each feature is standardized to zero mean and unit variance. The preprocessing steps included converting the emotion labels into numeric classes, casting all feature values to `float32`, reshaping the feature vectors into a three-dimensional format ($\text{samples} \times \text{timesteps} \times \text{features}$) required by the model, and splitting the data using an 80/20 stratified partition; the same procedure was applied for preparing the full dataset used in the 10-fold cross-validation.

TABLE 2. Model architecture of the proposed hybrid neural network.

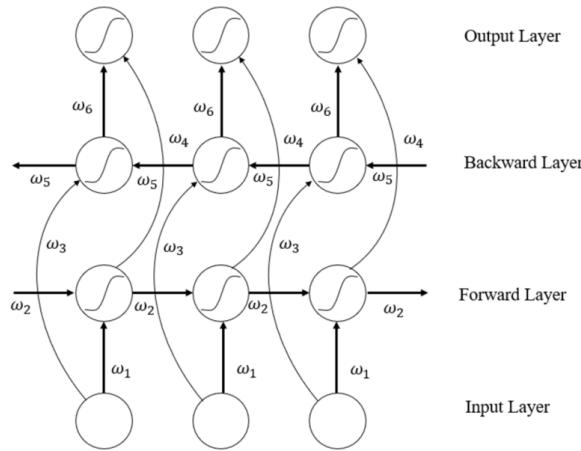
Layer (Type)	Output Shape
Input Layer	(None, 2548, 1)
BiLSTM (128 units)	(None, 2548, 256)
Dropout (Rate: 0.5)	(None, 2548, 256)
Batch Normalization	(None, 2548, 256)
BiGRU (128 units)	(None, 256)
Dropout (Rate: 0.5)	(None, 256)
Batch Normalization	(None, 256)
Dense (64 units, Activation: ReLU)	(None, 64)
Dropout (Rate: 0.5)	(None, 64)
Dense (Output, 3 units, Activation: Softmax)	(None, 3)

B. MODEL ARCHITECTURE

The proposed model as seen in Fig 6. combines Bidirectional Long Short-Term Memory (BiLSTM) and Bidirectional Gated Recurrent Unit (BiGRU) layers to capture forward and backward temporal dependencies in the EEG sequences, leveraging the BiLSTM's capability for modeling long-range dependencies and the BiGRU's efficiency to balance computational complexity as seen in Table 2. The input layer processes the standardized EEG sequence X and passes it to the first BiLSTM layer. During development, we experimented with different hidden unit sizes (64, 128, 256) and dropout rates (0.3–0.5) across multiple training seeds to evaluate stability and generalization. The configuration with 128 units and a 0.5 dropout rate consistently

**FIGURE 6.** Proposed model implementation.

achieved the best balance between classification accuracy and overfitting mitigation, which motivated its selection in the final architecture.

**FIGURE 7.** Representation of BiLSTM model [17].

C. BIDIRECTIONAL LSTM LAYERS

The BiLSTM layer processes the sequence in both forward and backward directions to retain comprehensive temporal dependencies Fig 7. This allows the BiLSTM to capture long-range contextual information through prior steps and future steps. At each time step t , each LSTM layer computes a hidden state h_t and cell state c_t , as follows:

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \quad (1)$$

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \quad (2)$$

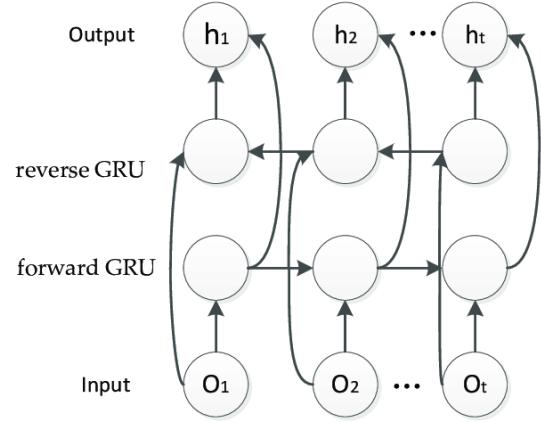
$$\tilde{c}_t = \tanh(W_c x_t + U_c h_{t-1} + b_c) \quad (3)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \quad (4)$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \quad (5)$$

$$h_t = o_t \odot \tanh(c_t) \quad (6)$$

Here, f_t , i_t , \tilde{c}_t , and o_t represent the forget, input, candidate cell, and output gates, respectively; σ denotes the sigmoid activation function, and \odot represents element-wise multiplication.

**FIGURE 8.** Representation of BiGRU model [18].

D. BIDIRECTIONAL GRU LAYERS

After the BiLSTM layer, a Bidirectional Gated Recurrent Unit (BiGRU) layer is used to capture additional temporal patterns Fig 8. The GRU computes the hidden state h_t at each time step as follows:

$$z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z) \quad (7)$$

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r) \quad (8)$$

$$\tilde{h}_t = \tanh(W_h x_t + U_h(r_t \odot h_{t-1}) + b_h) \quad (9)$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \quad (10)$$

Here, z_t and r_t represent the update and reset gates, respectively, controlling the flow of information in each direction.

Additionally, in the proposed hybrid model, the BiLSTM and BiGRU layers complement each other. The BiLSTM captures the long-range contextual dependencies in both forward and backward directions. On the other hand, the BiGRU efficiently models shorter-term temporal dynamics with fewer parameters. Therefore, by combining these capabilities, the hybrid architecture provides a balanced representation that captures both long-range dependency and short-term dynamics while maintaining computational efficiency.

E. BATCH NORMALIZATION

Batch Normalization (BN) is employed in the model to stabilize and accelerate the training process by normalizing the inputs to each layer. This technique reduces internal covariate shift, allowing for higher learning rates and improving the overall performance.

For each mini-batch, BN standardizes the inputs to have a mean of zero and a variance of one:

$$\hat{x}^{(t)} = \frac{x^{(t)} - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \quad (11)$$

where $x^{(t)}$ is the input of the mini-batch, μ_B is the batch mean, σ_B^2 is the batch variance, and ϵ is a small constant added for numerical stability. After normalization, the output is scaled and shifted using learnable parameters γ and β :

$$y^{(t)} = \gamma \hat{x}^{(t)} + \beta \quad (12)$$

The inclusion of Batch Normalization layers after each LSTM and GRU layer enhances the convergence rate and allows for deeper network architectures by mitigating issues related to vanishing gradients.

This normalization process helps the model generalize better by reducing overfitting, making it a valuable component in our architecture.

F. DENSE LAYER AND SOFTMAX OUTPUT

The final hidden states from the BiGRU layers pass through a fully connected dense layer with a softmax activation function, yielding class probabilities:

$$\hat{y}_k = \frac{\exp(W_k^\top h + b_k)}{\sum_{j=1}^C \exp(W_j^\top h + b_j)}, \quad k \in \{1, \dots, C\} \quad (13)$$

where $C = 3$ denotes the number of classes, W_k and b_k represent the weights and bias for the k -th class, and \hat{y}_k is the predicted probability for the k -th class.

G. TRAINING PROTOCOL

The model is trained over 30 epochs using the Adam optimizer, which computes parameter updates as follows:

$$\theta_t = \theta_{t-1} - \alpha \hat{v}_t + \epsilon \hat{m}_t \quad (14)$$

where θ_t denotes the model parameters at iteration t , α is the learning rate, and \hat{m}_t and \hat{v}_t are the bias-corrected first and second moment estimates of the gradient, defined as:

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t}, \quad \hat{v}_t = \frac{v_t}{1 - \beta_2^t} \quad (15)$$

with

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) \nabla_\theta L(\theta_t) \quad (16)$$

and

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) \nabla_\theta L(\theta_t)^2. \quad (17)$$

The categorical cross-entropy loss L used to optimize the model is defined as:

$$L = - \sum_{k=1}^C y_k \log(\hat{y}_k) \quad (18)$$

where y_k represents the true label for the k -th class. Early stopping is employed to terminate training if validation accuracy does not improve for a predefined number of epochs, minimizing overfitting.

H. EVALUATION METRICS

The model's performance is assessed on a hold-out test set using accuracy, precision, recall, and F1-score metrics. These measures are particularly appropriate for emotion classification tasks, as they not only evaluate overall accuracy but it also highlights the model's ability to correctly identify subtle differences between emotional states, even in the presence of potential class imbalances. Additionally, the area under the receiver operating characteristic (ROC AUC) and precision-recall curve (PRC AUC) are calculated to provide insights into model performance across varying classification thresholds. Accuracy is formally defined as:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (19)$$

where TP, TN, FP, and FN denote true positives, true negatives, false positives, and false negatives. A confusion matrix further visualizes the model's classification performance per class will be shown in the results section.

Algorithm 1 Hybrid BiGRU–BiLSTM With 10-Fold Cross-Validation

```

Require: EEG feature matrix  $X$ , labels  $y \in \{0, 1, 2\}$ 
Require: batch size  $B = 32$ , learning rate  $\eta = 0.001$ , dropout  $p = 0.4$ 
Require: epochs  $E = 30$ , L2 regularization  $\lambda = 0.001$ , folds  $K = 10$ 
1: Convert labels to numeric classes.
2: Normalize features and reshape input to  $(N, F, 1)$ .
3: Initialize stratified  $K$ -fold splitter.
4: Initialize accuracy list  $\mathcal{A} = []$ .
5: for each fold  $k = 1 \dots K$  do
6:   Split data into training and validation sets.
7:   Input layer  $\rightarrow$  BiLSTM(128, L2( $\lambda$ ), return seq.)
8:   Dropout( $p$ )  $\rightarrow$  BatchNorm
9:   BiGRU(128, L2( $\lambda$ ), no return seq.)
10:  Dropout( $p$ )  $\rightarrow$  BatchNorm
11:  Dense(64, ReLU, L2( $\lambda$ ))  $\rightarrow$  Dropout( $p$ )
12:  Output Dense(3, softmax)
13:  Compile with Adam( $\eta$ ) and sparse categorical loss.
14:  Use EarlyStopping (patience=5) and ReduceLROnPlateau (factor=0.5).
15:  Train for up to  $E$  epochs with validation on fold  $k$ .
16:  Compute accuracy  $a_k$  on validation set.
17:  Append  $a_k$  to  $\mathcal{A}$ .
18: end for
19: Compute mean accuracy  $\bar{a}$  and standard deviation  $\sigma_a$  from  $\mathcal{A}$ .
20: return  $\bar{a}, \sigma_a$ 

```

Note: Algorithm 1 summarizes the complete training pipeline, including preprocessing, hybrid model construction, and 10-fold cross-validation.

VI. RESULTS AND EXPERIMENTAL ANALYSIS

A. EXPERIMENTAL SETUP

A triad of evaluative methodologies was the approach to this investigation to prove its accuracy and efficiency. First, we evaluated the model that integrates three layers of Bidirectional Long Short-Term Memory (BiLSTM) with batch normalization and dropout. Then, we evaluated the model that integrates three layers of Bidirectional Gated Recurrent Units (BiGRU) with batch normalization and dropout. Finally, our proposed novel hybrid model that integrates BiGRU and BiLSTM networks for the task of emotion recognition from electroencephalogram (EEG) signals. The primary dataset utilized in this research comprises EEG recordings sourced from EEG Brainwave Dataset which encompasses data from individuals subjected to a series of emotional stimuli [11]. In the system the data was sequenced and randomized 80% training and 20% for validation for all three models.

The model architecture includes an in-depth design for sequential EEG emotion classification using three bidirectional Long Short-Term Memory (BiLSTM) layers, each with distinct dropout and normalization steps to control overfitting and promote robust learning. The input layer processes EEG samples of shape (timesteps, features) to initialize sequential data flow through the architecture. The first layer is a Bidirectional LSTM with 128 units, which operates by capturing sequential dependencies from both past and future time steps, enabling the model to leverage bi-directional temporal information. A dropout layer with a 40% rate follows to mitigate overfitting, while batch normalization stabilizes the learning process by normalizing outputs from this layer. In the subsequent stacked layers, the second and third layers apply BiLSTM with progressively smaller unit sizes (64 and 32 units) to gradually reduce the feature space while preserving learned temporal features. Each BiLSTM layer is followed by identical dropout and batch normalization layers, further supporting generalization by reducing co-adaptations of neurons. The output of the last BiLSTM layer is processed through a fully connected layer with 64 units and a ReLU activation function, allowing for non-linear combinations of features before reaching the final softmax layer for classification. The softmax output layer provides a probability distribution over the three emotion classes (Negative, Neutral, Positive), facilitating a multi-class classification approach for EEG emotion recognition. This layered structure is optimized for both learning stability and adaptability to EEG signal complexities.

Similarly, the other model architecture employs a sequential EEG classification design using three Bidirectional Gated Recurrent Unit (BiGRU) layers, each followed by dropout and batch normalization layers to improve learning stability and generalization. The input layer accepts EEG samples shaped as (timesteps, features), setting up a structured flow for processing sequential data through the network. The first layer is a Bidirectional GRU with 128 units, which extracts temporal dependencies in both directions (past and future),

leveraging bi-directional context within EEG sequences. A dropout layer with a 40% rate follows, helping prevent overfitting by randomly setting neuron activations to zero during training, while batch normalization regulates output stability by standardizing this layer's activations. In the stacked GRU layers, the second and third layers progressively reduce the feature space using smaller unit counts (64 and 32 units, respectively), retaining critical temporal features while simplifying the data representation. Each layer also includes dropout and batch normalization, fostering robust learning by preventing reliance on specific neuron interactions. The output from the final GRU layer is passed through a dense layer with 64 units and a ReLU activation, enabling non-linear transformation of features prior to the output layer. The softmax output layer produces a probability distribution over the three emotion classes (Negative, Neutral, Positive), supporting a multi-class classification strategy. This approach is tailored to capture complex EEG signal patterns effectively, offering adaptability to temporal intricacies within emotional states.

The proposed hybrid model begins with an input layer that accepts data in the shape of (2, 64) which represents 2 seconds of EEG recordings across 64 channels. This data format ensures that temporal and spatial information from the EEG signals is effectively preserved for subsequent processing. The first processing layer is a BiGRU Layer with 128 units, designed to capture sequential dependencies within the EEG time series. BiLSTM Layer with 128 units is incorporated to further enhance the model's temporal understanding by processing the data in both forward and backward directions, capturing contextual information from both past and future states. Following the BiLSTM, BiGRU layers were implemented to capture time-dependent data as they are efficient in capturing patterns in long sequences, helping the model recognize and retain important sequential information within each EEG segment. To address overfitting, a dropout layer with a 0.3 dropout rate is applied, randomly deactivating 30% of neurons during training to encourage more generalized feature learning. Lastly, the output layer employs a softmax-activated dense layer to produce probability scores for each emotional category. Training was conducted with a batch size of 32, a learning rate of 0.001, and the adam optimizer, selected for its adaptive learning capabilities, over 30 epochs. An early stopping mechanism based on validation loss further mitigates overfitting and enhances model generalization. After defining the architecture, the model was further evaluated using 10-fold cross-validation to ensure that its performance was stable across different data partitions.

B. PERFORMANCE COMPARISON

To further examine model stability the training and validation loss curves need to be observed for each model. The loss function trends provide insight into convergence rates and any potential issues with overfitting.

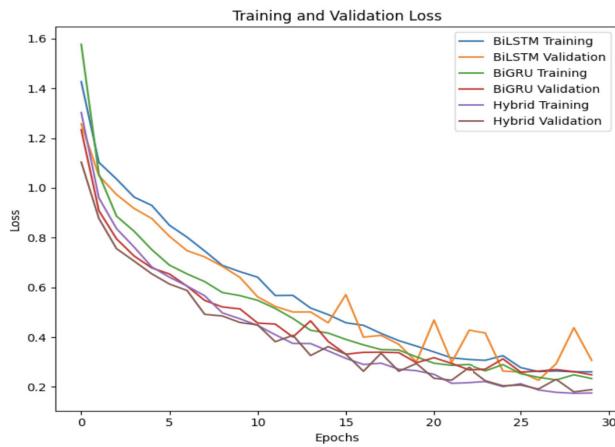


FIGURE 9. Loss evaluation.

Across all models as seen in Fig 9, the loss function decreased steadily, indicating successful learning. The hybrid BiGRU + BiLSTM model achieved the lowest final validation loss, with a more stable convergence pattern compared to the standalone BiGRU and BiLSTM models. This lower loss indicates that the hybrid model effectively minimized prediction errors, maintaining a closer match between training and validation performance. Regularization techniques, such as dropout and batch normalization, played a role in achieving this stability, allowing the hybrid model to generalize effectively across new data while avoiding overfitting.

The accuracy of each model—BiGRU, BiLSTM, and the hybrid BiGRU + BiLSTM—was tracked across training epochs to assess their learning progression and final classification performance. Both training and validation accuracy curves were analyzed to determine each model’s capacity to generalize beyond the training data.

The hybrid BiGRU + BiLSTM model demonstrated the highest test accuracy at approximately 98.59%, compared to 98.1% for BiGRU and 98.13% for BiLSTM. The 10-fold validation ranged from 96% to 98% and resulted in an average accuracy of about $97.4\% \pm 0.8\%$. This improvement suggests that the hybrid architecture more effectively leverages temporal dependencies by combining GRU’s memory efficiency with LSTM’s strength in handling longer sequences, leading to enhanced classification accuracy. The accuracy curves for the hybrid model as seen in Fig 10 & Table 3, stabilized at a higher level with less fluctuation, indicating consistent generalization across the emotional classes. However, given the relatively small dataset size (two subjects), steps were taken to minimize the risk of overfitting. Regularization techniques, including dropout layers (rate = 0.5) and L2 penalties, were applied throughout the architecture. Batch normalization was incorporated to stabilize training and improve generalization. In addition, an early stopping strategy was employed to prevent over-training on the validation set. To further ensure robustness, experiments were seeded

and repeated with different hyperparameter variations, which demonstrated stable performance across runs.

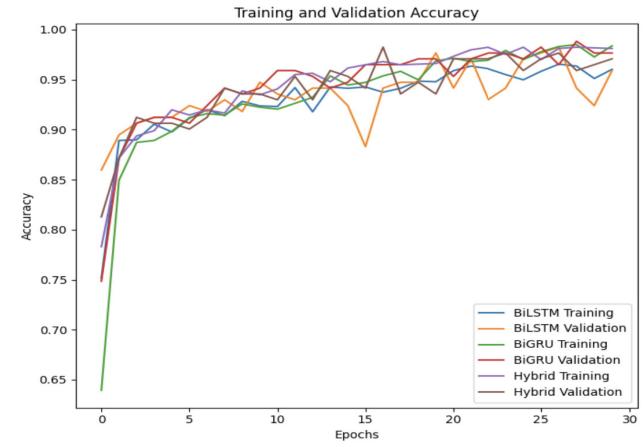


FIGURE 10. Accuracy evaluation.

TABLE 3. Summary of training and validation metrics across key epochs.

Epoch	Train Accuracy	Train Loss	Val Accuracy	Val Loss
1	0.689	1.510	0.848	1.077
5	0.920	0.704	0.930	0.670
10	0.935	0.488	0.959	0.468
15	0.959	0.360	0.971	0.360
20	0.980	0.246	0.965	0.275
25	0.982	0.208	0.965	0.265
30	0.983	0.200	0.971	0.254

While hold-out validation was adopted in this study, we acknowledge that future work should explore cross-validation and data augmentation techniques to further enhance generalizability across subjects.

The confusion matrices were generated, as seen in Fig 11, for each model to provide a detailed breakdown of classification performance across the three emotional categories: Negative, Neutral, and Positive.

While both the BiLSTM (A) and BiGRU (B) models performed well, they each exhibited challenges in distinguishing between certain classes, particularly between negative and positive emotions. Specifically, the BiLSTM (A) model misclassified a small number of negative samples as positive, while the BiGRU (B) model misclassified some positive samples as negative. On the other hand, the proposed hybrid (C) model revealed fewer misclassifications across all classes, particularly in the positive class. This reflects its improved capacity to differentiate emotional states compared to the standalone BiLSTM (A) and BiGRU (B) models. This enhanced accuracy demonstrates the hybrid model’s superior ability to capture subtle variations in contextual cues, leading to more accurate predictions. Furthermore, the hybrid model exhibited higher recall and precision in detecting neutral and positive emotions, underscoring its robustness in complex emotional classification tasks. The confusion matrix shows that most misclassifications occur between the

Neutral and Positive classes, which aligns with EEG literature indicating that these emotional states often share overlapping spectral-temporal patterns. This overlap leads to slightly higher false-positive tendencies, as the subtle EEG variations associated with positive affect can resemble neutral cognitive activity.

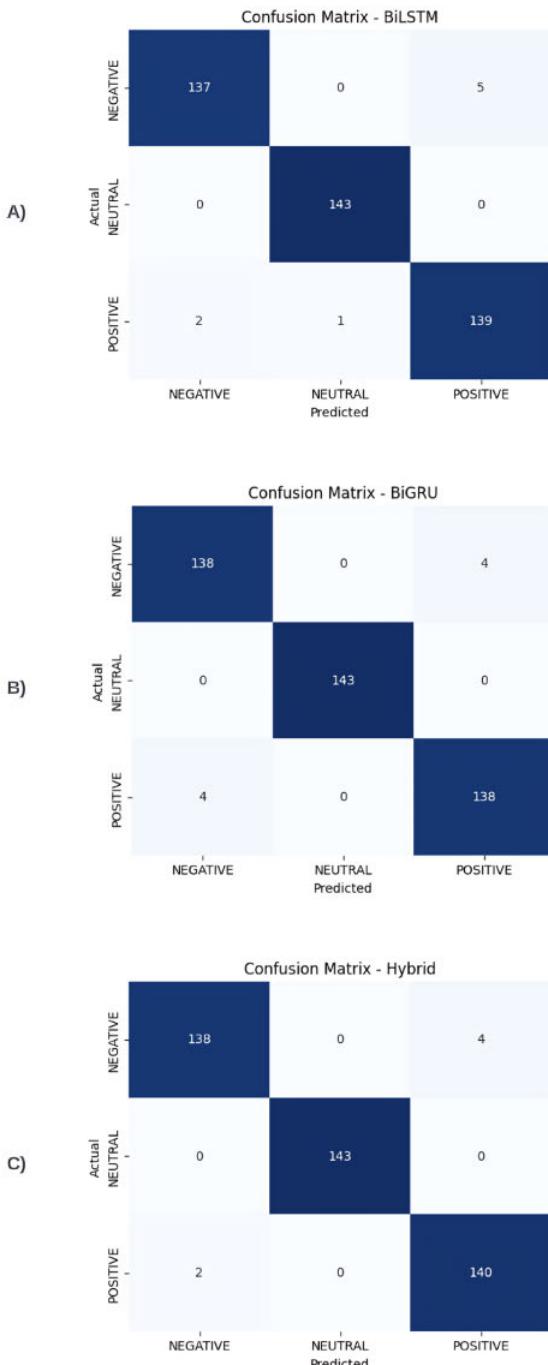


FIGURE 11. Confusion matrix evaluation.

The BiLSTM model showed a slight improvement in separating these classes over BiGRU, likely due to LSTM's superior ability to capture long-term dependencies in sequen-

tial data. However, the hybrid model outperformed both, achieving higher precision and recall across all classes. These results highlight the hybrid model's ability to capture nuanced emotional distinctions, benefiting from the complementary strengths of both BiGRU and BiLSTM layers.

The proposed model in this research demonstrates a higher performance in emotion recognition, achieving an accuracy of 98.59%, which surpasses all other models included in the comparison as seen in Table 4 & Table 5. In contrast, Table 4 presents a comparison using a different dataset, while Table 5 provides the analysis for models with the same dataset (InfoGain). The next best-performing model, Random Forest combined with Feature Engineering (FE) [11], achieved an accuracy of 97.89%, followed closely by the CNN + GCN model [3] at 97.7%. While these models approach the performance of the proposed model, the 0.7% improvement signifies a noteworthy advancement, particularly in the context of emotion recognition, where even marginal increases in accuracy can substantially enhance real-world applicability. Additionally, models such as 3D CNN and MLP [4] yielded lower accuracies of 86.72% and 87.97%, respectively, while the CL + GAN + GNN [10] approach recorded an even lower performance of 65.62%. This highlights the robustness of the proposed model, which consistently outperforms various architectures, including hybrid approaches and traditional deep learning models, across the board. The results suggest that the proposed model captures the nuanced features essential for accurate emotion recognition more effectively than both simpler and more complex architectures, establishing it as a new benchmark in the field. However, it should be noted that many of the comparative models referenced in this study were evaluated on different datasets, which limits the direct comparability of performance. It is important to also note that most of the compared studies reported only accuracy as their primary evaluation metric, with precision and recall not consistently provided; therefore, these measures are included only for the proposed model where available. Although a full complexity comparison was beyond the scope of this study, the proposed BiGRU-BiLSTM model achieves 98.59% accuracy while remaining relatively lightweight. The limited number of subjects and lack of diversity in the dataset highlight the need for future work. While the improvement over the single BiLSTM and BiGRU models is modest (0.4–0.5%), such gains are still meaningful in EEG-based emotion recognition,

TABLE 4. Comparison of emotion recognition models.

Model	Dataset	Overall Accuracy
CNN + GCN [3]	Dreamer & GAMEEMO	97.7 %
3D CNN [4]	DEAP	86.72 %
Wavelet + CNN + LSTM [5]	DEAP	94.41 %
Wavelet + CNN [6]	SEED	95.67 %
BiGRU + LSTM [7]	SEED & DEAP	97.16 %
CNN + BiLSTM [8]	DEAP	93.19 %
GNN [9]	SEED & SEED-IV	94.27 %
CL + GAN + GNN [10]	MAHNOB-HCI & DEAP	65.62 %
Proposed Model	InfoGain [11, 12]	98.59 %

where even small increases in accuracy can indicate better separation of subtle emotional states.

TABLE 5. Comparison of emotion recognition models on InfoGain.

Model	Recall	F1 Score	Overall Accuracy
Random Forest + FE [11]	—	—	97.89 %
LSTM [12]	—	—	90.75 %
MLP [12]	—	—	87.97 %
Proposed Model	0.99	0.99	98.59 %

VII. CONCLUSION AND FUTURE WORK

In summary, the hybrid BiGRU + BiLSTM model consistently outperformed the standalone BiGRU and BiLSTM models across all evaluation metrics, including accuracy, loss, precision, recall, and F1-score. The accuracy and loss function trends highlight the hybrid model's enhanced convergence and stability, while the confusion matrix analysis confirms its superior class differentiation capability. By combining GRU's computational efficiency with LSTM's long-range dependency handling, the hybrid model demonstrates a balanced approach for EEG-based emotion classification, achieving both high performance and generalizability. We also acknowledge that cross-subject validation was not possible due to the limited number of participants, and future work involving larger datasets with more than ten subjects will be essential for establishing stronger statistical validity.

This hybrid architecture shows potential for real-time affective computing applications where accurate emotion recognition is critical. The results emphasize the advantages of integrating GRU and LSTM architectures to capture complex temporal patterns in EEG data effectively. Real-time deployment were not examined in this study, the efficiency of the proposed architecture indicates that these applications may be feasible and merit further investigation in future work.

In the future, we plan to extend our model beyond basic affective states to capture a broader range of emotions. Our aim is to explore its applicability to imagined speech decoding, which shares overlapping neural dynamics with emotional processing and could further validate the generality of our approach. We plan to extend our work by evaluating the model on multimodal datasets that integrate EEG with complementary physiological and behavioral signals (e.g., ECG, GSR, facial expressions, and speech). Incorporating these additional modalities will allow us to explore multimodal emotion-recognition strategies. Additionally, we plan to develop a real-time emotion-aware system incorporating GAN-based EEG augmentation that integrates both hardware and software components toward practical applications in human-computer interaction.

REFERENCES

- [1] Md. S. Mahmud, O. Saha, and S. A. Fattah, "An efficient bidirectional LSTM-based deep neural network for automatic emotion recognition using EEG signal," in *Proc. 12th Int. Conf. Electr. Comput. Eng. (ICECE)*, Dec. 2022, pp. 417–420, doi: [10.1109/ICECE57408.2022.10088864](https://doi.org/10.1109/ICECE57408.2022.10088864).

- [2] J. M. Jose and Aravindh. J, "Frame work for EEG based emotion recognition based on hybrid neural network," in *Proc. 7th Int. Conf. Bio-Signals, Images, Instrum. (ICBSII)*, Mar. 2021, pp. 1–7, doi: [10.1109/ICB-SII51839.2021.9445130](https://doi.org/10.1109/ICB-SII51839.2021.9445130).
- [3] R. A. Nahin, M. T. Islam, A. Kabir, S. Afrin, I. A. Chowdhury, R. Rahman, and M. G. R. Alam, "Electroencephalogram-based emotion recognition with hybrid graph convolutional network model," in *Proc. IEEE 13th Annu. Comput. Commun. Workshop Conf. (CCWC)*, Mar. 2023, pp. 705–711, doi: [10.1109/CCWC57344.2023.10099220](https://doi.org/10.1109/CCWC57344.2023.10099220).
- [4] M. Islam and T. Lee, "MEMD-HHT based emotion detection from EEG using 3D CNN," in *Proc. 44th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2022, pp. 284–287, doi: [10.1109/EMBC48229.2022.9871012](https://doi.org/10.1109/EMBC48229.2022.9871012).
- [5] M. Islam and T. Lee, "Wavelet based emotion detection from multi-channel EEG using a hybrid CNN-LSTM model," in *Proc. TENCON IEEE Region Conf. (TENCON)*, Nov. 2022, pp. 1–6, doi: [10.1109/TENCON55691.2022.9978122](https://doi.org/10.1109/TENCON55691.2022.9978122).
- [6] B. Ramar, R. Ramalakshmi, V. Gandhi, and P. Pandiselvam, "Classification of EEG signals on SEED dataset using improved CNN," in *Proc. 2nd Int. Conf. Edge Comput. Appl. (ICECAA)*, Jul. 2023, pp. 1095–1102, doi: [10.1109/ICECAA58104.2023.10212279](https://doi.org/10.1109/ICECAA58104.2023.10212279).
- [7] A. Iyer, S. S. Das, R. Teotia, S. Maheshwari, and R. R. Sharma, "CNN and LSTM based ensemble learning for human emotion recognition using EEG recordings," *Multimedia Tools Appl.*, vol. 82, no. 4, pp. 4883–4896, Apr. 2022, doi: [10.1007/s11042-022-12310-7](https://doi.org/10.1007/s11042-022-12310-7).
- [8] Z. Li, G. Zhang, J. Dang, L. Wang, and J. Wei, "Multi-modal emotion recognition based on deep learning of EEG and audio signals," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2021, pp. 1–6, doi: [10.1109/IJCNN52387.2021.9533663](https://doi.org/10.1109/IJCNN52387.2021.9533663).
- [9] P. Zhong, D. Wang, and C. Miao, "EEG-based emotion recognition using regularized graph neural networks," *IEEE Trans. Affect. Comput.*, vol. 13, no. 3, pp. 1290–1301, Jul. 2022, doi: [10.1109/TAFFC.2020.2994159](https://doi.org/10.1109/TAFFC.2020.2994159).
- [10] S. S. Gilakjani and H. A. Osman, "A graph neural network for EEG-based emotion recognition with contrastive learning and generative adversarial neural network data augmentation," *IEEE Access*, vol. 12, pp. 113–130, 2024, doi: [10.1109/ACCESS.2023.3344476](https://doi.org/10.1109/ACCESS.2023.3344476).
- [11] J. Bird, A. Ekart, C. Buckingham, and D. Faria, "Mental emotional sentiment classification with an EEG-based brain-machine interface," *Tech. Rep.*, 2019.
- [12] J. Bird, D. Faria, L. Manso, A. Ekárt, and C. Buckingham, "A deep evolutionary approach to bioinspired classifier optimisation for brain-machine interaction," *Tech. Rep.*, 2019, doi: [10.48550/arXiv.1908.04784](https://doi.org/10.48550/arXiv.1908.04784).
- [13] K. Li, K. Tang, J. Li, T. Wu, and Q. Liao, "A hierarchical neural hybrid method for failure probability estimation," *IEEE Access*, vol. 7, pp. 112087–112096, 2019, doi: [10.1109/ACCESS.2019.2934980](https://doi.org/10.1109/ACCESS.2019.2934980).
- [14] B. G. Barsy, G. Győri, and P. T. Szemes, "Development of EEG measurement and processing system in LabVIEW development environment," *Int. Rev. Appl. Sci. Eng.*, vol. 11, no. 3, pp. 287–297, Nov. 2020, doi: [10.1556/1848.2020.00151](https://doi.org/10.1556/1848.2020.00151).
- [15] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016. [Online]. Available: <https://www.deeplearningbook.org/>
- [16] Mayo Foundation for Medical Education and Research. *Electroencephalography (EEG)*. [Online]. Available: <https://www.mayoclinic.org/tests-procedures/eeg/about/pac-20393875>
- [17] T. Sun, C. Yang, K. Han, W. Ma, and F. Zhang, "Bidirectional spatial-temporal network for traffic prediction with multisource data," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2674, no. 8, pp. 78–89, Aug. 2020, doi: [10.1177/0361198120927393](https://doi.org/10.1177/0361198120927393).
- [18] Z. Gao, Z. Li, J. Luo, and X. Li, "Short text aspect-based sentiment analysis based on CNN + BiGRU," *Appl. Sci.*, vol. 12, no. 5, p. 2707, Mar. 2022, doi: [10.3390/app12052707](https://doi.org/10.3390/app12052707).

NATASHA PIEDRABUENA (Student Member, IEEE) is currently pursuing the B.S. degree in computer science with Manhattan University, Riverdale, New York. Since 2024, she has been a Research Assistant with the HIAT Laboratory under Dr. Elmannai. She plans to continue her academic career by pursuing graduate studies and further contributing to research in artificial intelligence. Her research interests include agentic AI, assistive technology, computer vision, and deep learning models.



WAFA M. ELMANNAI (Member, IEEE) is currently an Associate Professor with the Department of Electrical and Computer Engineering, Manhattan University. She is also the Founder and the Co-Director of the Women's Research Mentorship Program, Manhattan University, established through the Women and Gender Resource Center and supported by the 2023–2024 AAUW Community Action Grant. This initiative promotes research engagement among undergraduate through mentorship and academic exploration. She has supervised and co-supervised numerous undergraduate and graduate research projects. She has published extensively in respected national and international venues and has served in various professional roles, including several IEEE conferences, the International Wireless Communications and Mobile Computing Conference, and American Society for Engineering Education (ASEE). She has contributed to the peer-review process and the development of high-quality technical programs in areas, such as artificial intelligence and smart systems. Her research interests include the intersection of emerging technologies and inclusive computing, encompassing assistive technologies, smart systems, artificial intelligence, the Internet of Things (IoT), mobile computing, computer architecture, and quantum computing.

Dr. Elmannai is an UDL Teaching Fellow. She is an Active Member of several professional and honor societies, including IEEE, the IEEE Communications Society, Eta Kappa Nu, Phi Kappa Phi, UPE, SWE, and

Arab American Association of Engineers and Architects (AAAEA). She was a recipient of the AAUW International Fellowship. During her graduate studies, she served as the Chapter President and the Vice President for the Upsilon Pi Epsilon (UPE) Honor Society, UB AAUW Ambassador, and the Vice President for the Society of Women Engineers (SWE), UB. She continues this leadership as the IEEE-HKN Chapter Advisor at Manhattan University.



YI WANG (Senior Member, IEEE) received the B.S. degree in information management and information system and the M.S. degree in computer science from Wuhan University of Science and Technology, Wuhan, China, in 2006 and 2009, respectively, and the Ph.D. degree in computer engineering from The University of Alabama in Huntsville, Huntsville, AL, USA, in 2015. He is currently an Associate Professor with the Department of Electrical and Computer Engineering, Manhattan University, Riverdale, NY, USA. His work has received funding from the National Science Foundation (NSF) and New York City Department of Transportation. His main research interests include image processing, machine learning, cybersecurity, and blockchain systems.

• • •