Journal of Robotics and Control (JRC)
Volume 6, Issue 3, 2025
ISSN: 2715-5072, DOI: 10.18196/jrc.v6i3.26016

Optimizing Gated Recurrent Unit Architecture for Enhanced EEG-Based Emotion Classification

Yuri Pamungkas [1*], Stralen Pratasik [2], Made Krisnanda [3], Padma Nyoman Crisnapati [4]
[1] Department of Medical Technology, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
[2] Department of Informatics Engineering, Universitas Negeri Manado, Manado, Indonesia
[3] School of Information and Physical Sciences, University of Newcastle, Callaghan, NSW 2308, Australia
[4] Department of Mechatronics Engineering, Rajamangala University of Technology Thanyaburi, Bangkok, Thailand
Email: [1] yuri@its.ac.id, [2] stralente@unima.ac.id, [3] made.krisnanda@uon.edu.au, [4] crisnapati@rmutt.ac.th
*Corresponding Author

***Abstract*—Emotion recognition using EEG signals has gained significant attention in affective computing and brain-computer interface (BCI) applications. However, achieving high classification accuracy remains a challenge due to the complexity and variability of EEG signals. This study aims to optimize the Gated Recurrent Unit (GRU) model for improving the performance of EEG-based emotion classification. The approach involves feature selection and architectural modifications to the GRU model. Selected EEG features include mean, standard deviation, statistical moments (skewness and kurtosis), min-max values, logarithmic covariance matrix, covariance matrix, Shannon entropy, log-energy entropy, Fast Fourier Transform, and autocorrelation, extracted from alpha and beta frequency bands. The proposed GRU model consists of four stacked GRU layers with decreasing hidden state sizes, ensuring efficient temporal feature extraction while reducing computational complexity. Experimental results demonstrate the superiority of the proposed GRU model compared to Simple RNN, LSTM, and traditional Machine Learning models (Naïve Bayes, SVM, Random Forest, and Linear Regression). The GRU model achieves high recall (98.81%), specificity (99.42%), precision (98.82%), accuracy (99.22%), and F1 Score (98.81%), outperforming alternative models in all evaluation metrics. These findings indicate that the GRU model effectively captures temporal dependencies in EEG signals, making it a robust and efficient approach for EEG-based emotion classification. In conclusion, this research confirms that GRU is an optimal deep learning model for emotion recognition using EEG. Future research could explore multi-modal emotion recognition, attention-based architectures, and real-time deployment in wearable EEG devices to further enhance classification accuracy and real-world applicability.**

Keywords*—EEG-based Emotion Classification; BCI; Affective Computing; Gated Recurrent Unit; Temporal Feature Extraction.

I. INTRODUCTION

EEG signal-based emotion classification has become a major focus in Affective Computing and human-machine interaction research [1]. Emotions are a fundamental aspect of human life that influences decision-making, social behavior, and mental health [2]. However, measuring and classifying emotions remains a challenge due to its subjective and complex nature [3]. EEG, as a technique for recording brain electrical activity, offers a more objective approach to recognizing emotions compared to methods based on facial expressions or voice analysis [4]. With its ability to capture brain wave patterns in real-time, EEG is a promising tool for detecting and classifying a person's affective state, such as happy, sad, angry, or neutral [5]. The main advantage of EEG in emotion classification is its ability to capture brain activity that cannot be consciously controlled, thus providing a more accurate picture of a person's emotional state [6]. EEG also has high temporal resolution, allowing the detection of changes in brain activity in milliseconds, much faster than other brain imaging methods such as fMRI [7]. In addition, modern EEG devices are increasingly portable and non-invasive, allowing for use in a variety of environments, including psychology research, Human-Computer Interaction (HCI), and neurofeedback therapy [8]. Despite these advantages, EEG signals face several notable limitations. EEG is highly susceptible to noise from various sources, such as muscle movements, eye blinks, and environmental interference, which can significantly affect signal quality and classification accuracy [9]. In recent years, Artificial Intelligence, especially Machine Learning and Deep Learning, have played a significant role in improving the accuracy of EEG-based emotion classification [10]. AI models are able to extract complex patterns from EEG signals that are difficult to interpret manually. Techniques such as SVM, Random Forest, and k-NN have been widely used in EEG feature processing [11].

However, Deep Learning-based approaches are increasingly being adopted due to their ability to process EEG data directly without the need for manual feature extraction [12]. These models have shown significant improvements in emotion classification accuracy and open up broad opportunities in various applications [13]. Several recent studies have proposed various methods to enhance the accuracy of EEG-based emotion detection or classification. One promising approach is the integration of EEG with audiovisual signals using contrastive learning, as done by Lee et al. (2024) [14]. This model combines cross-modal attention mechanisms to improve the accuracy of emotion classification, demonstrating that the use of multimodal data can improve the accuracy of EEG-based classification systems. In the context of deep learning, several architectures have been explored to improve the performance of emotion classification. Kulkarni et al. (2024) developed a model based on DCNN and Bi-GRU, combined with Fourier Transform and Common Spatial Pattern for feature extraction [15]. This model achieved an accuracy of 96.24%, indicating that the combination of transform-based feature extraction and deep learning can improve the effectiveness of EEG-based emotion classification systems. In addition, Dai et al. (2022) proposed a Cross-Connected Convolutional Neural Network (C-CNN) that integrates information from multiple convolutional layers, achieving an accuracy of 93.7%, demonstrating the superiority of this model in capturing complex patterns in EEG signals [16].

Classical machine learning-based approaches are also still used for EEG-based emotion classification. Reis et al. (2025) evaluated SVM, MLP, and Random Forest in human-machine interaction to predict valence and arousal [17]. This study highlights that with proper feature selection, classical machine learning models can still provide competitive performance in EEG-based emotion classification systems. Meanwhile, Zuo et al. (2025) compared the effectiveness of various machine learning models in classifying EEG signals triggered by 2D and 3D VR stimuli, with the result that the Common Spatial Patterns (CSP) method was superior to Power Spectral Density (PSD) [18]. The Random Forest model achieved the highest accuracy of 95.02%, demonstrating the effectiveness of machine learning methods for VR-based emotion analysis. In addition to deep learning and machine learning, fuzzy inference-based approaches have begun to be developed to handle ambiguity in EEG signals. Li et al. (2025) developed a model based on adaptive fuzzy rule generation and fuzzy rule interpolation, which is able to improve the accuracy of emotion classification better than conventional methods [19]. This study shows that fuzzy inference can provide robustness to uncertainty in EEG signals, making it a promising alternative to deep learning-based methods.

In addition, Cruz-Vazquez et al. (2025) explored the use of quantum transforms and Fourier Neural Networks to improve EEG signal processing [20]. With this approach, their model achieved 95% accuracy, indicating that advanced signal transforms can significantly improve EEG classification accuracy. On the other hand, recent studies have also focused on adaptive electrode selection techniques to improve the efficiency of EEG-based emotion classification. Gannouni et al. (2021) developed a Zero-Time Windowing (ZTW) method for epoch estimation in EEG signals, which allows the system to select the most relevant parts of the EEG signal for emotion classification [21]. With this method, the classification accuracy can increase up to 89%, indicating that appropriate epoch estimation and electrode selection techniques can significantly improve the performance of EEG classification models. Therefore, this study focuses on optimizing GRU to improve the accuracy and efficiency of EEG-based emotion classification. To achieve this goal, this study will implement various optimization techniques, such as more effective feature selection, GRU hyperparameter tuning, and the use of attention mechanisms to improve EEG feature representation. In addition, this study will also compare the performance of GRU with other deep learning models to assess the superiority of GRU in capturing temporal patterns of EEG signals. With this approach, it is expected that the developed model can improve the accuracy of EEG emotion classification and accelerate the inference process, so that it can be applied in real-time applications, such as Brain-Computer Interface (BCI), mental health, and more adaptive human-machine interaction systems.

II. METHODOLOGY

A. EEG Brainwave Dataset

The EEG Brainwave dataset was collected from two participants, one male and one female, to record their brain activity in three main emotional states, namely positive, neutral, and negative [22]. Each state was recorded for 3 minutes per session, with an additional 6 minutes of neutral data during rest. The use of the same duration in each session aims to ensure that the data obtained is sufficient to analyze the pattern of changes in brain activity when individuals experience different emotions. With this approach, research can be more accurate in identifying differences in brain wave patterns that arise due to different emotional responses. The device used in data collection is the Muse EEG headband, a wearable EEG device that uses dry electrode sensors, allowing EEG signal recording without complex conductive procedures. This device records signals from four main points, namely TP9, AF7, AF8, and TP10, which have high relevance in emotional processing. Electrodes AF7 and AF8 are located in the frontal area of the brain that plays a role in cognition and emotional regulation, while electrodes TP9 and TP10 are located in the temporal region related to sensory processing and emotional memory. With this configuration, the device can capture patterns of brain activity that are directly related to emotional changes, allowing for deeper analysis of how the brain responds to emotional stimuli. The Fig. 1 are EEG signal measurement points (electrode channels) on the scalp.

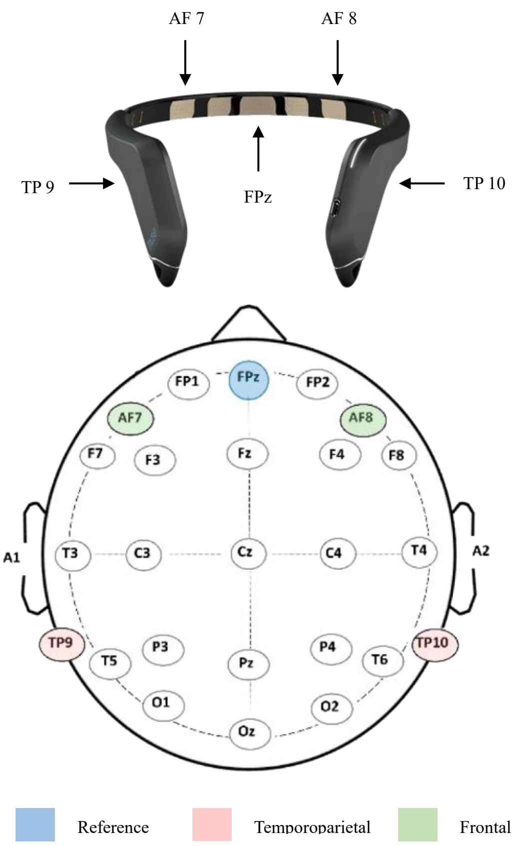


Fig. 1. Placement of electrode channels on the scalp [23]

To induce emotional reactions that can be observed in EEG signals, this study used movie clips as stimuli [24]. Movies were selected based on their ability to consistently evoke positive and negative emotions, so that they can produce stable responses from participants during the EEG recording process. The selection of movies with strong emotional content aims to ensure that the changes in brain activity recorded truly reflect the natural emotional responses of the individuals being tested. In evoking negative emotions, three movie clips with emotionally strong scenes were used. The first movie is Marley and Me (Twentieth Century Fox), which shows the death of Marley's pet dog. This scene is often associated with deep sadness and emotional attachment to pets. The second movie, Up (Walt Disney Pictures), presents a scene of losing a partner in a married life, which can trigger emotions of grief and deep feelings of loss. Meanwhile, the clip from My Girl (Imagine Entertainment) shows a funeral ceremony, which can evoke strong empathy and sadness for the loss of a loved one. Conversely, to evoke positive emotions, this study used three movie clips that have been proven to be able to trigger responses of happiness and joy. The first film, La La Land (Summit Entertainment), features an opening sequence with upbeat music and dancing, creating a joyful and energetic atmosphere. The second film, Slow Life (BioQuest Studios), is a documentary clip about the beauty of nature, which is known to create a sense of calm and happiness. Finally, the Funny Dogs (MashupZone) clip features a collection of funny dog videos, which are often used in affective neuroscience research to evoke positive emotions and laughter. By selecting the right stimuli, this study ensures that the emotional responses recorded via EEG truly reflect natural emotional experiences.

B. Features of EEG Signal

Electroencephalography (EEG) signal features are an important aspect in the analysis and classification of brain activity, especially in areas such as emotion recognition, neurological disorder diagnosis, and Brain-Computer Interface (BCI) development [25]. EEG signals recorded from scalp electrodes consist of non-stationary and complex brain electrical patterns, requiring feature extraction techniques to obtain interpretable information [26]. These features can be grouped into time domain, frequency domain, time-frequency domain, and connectivity-based features, each of which has its own advantages in characterizing brain activity [27]. In the EEG Brainwave dataset, the signal is extracted into several features, including:

1) Mean Values

Mean is the average amplitude of the EEG signal over a certain time interval [28]. This feature is used to determine the general trend of the EEG signal and can indicate the overall level of brain activity.

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{1}$$

where x_i is the EEG signal value at time point i; and N is the total number of samples.

2) Standard Deviation Values

Standard deviation (STD) measures the degree of variation or dispersion in the EEG signal from its mean [29]. A high STD indicates fluctuating brain activity, while a low STD indicates signal stability.

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2} \tag{2}$$

where μ is the mean EEG value; x_i is the EEG signal value at point i; and N is the total number of samples.

3) Statistical Moments

Statistical moments are used to analyze the distribution of EEG data. The commonly used moments include Skewness and Kurtosis [30]. Skewness is used to measure the asymmetry of the EEG data distribution. Meanwhile Kurtosis is used to measure the sharpness or flatness of the EEG signal distribution compared to a normal distribution.

$$S = \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^3}{\sigma^3} \tag{3}$$

$$K = \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^4}{\sigma^4} \tag{4}$$

A positive skewness value indicates a right-skewed distribution, while a negative skewness value indicates a left-skewed distribution. A high kurtosis value signifies a sharp peak, whereas a low value suggests a flatter distribution.

4) Max Value

Represents the highest amplitude value recorded in the EEG signal during a given interval [31]. This feature helps detect extreme brain activity, such as during emotional surges or high cognitive responses.

$$X_{max} = \max(x_1, x_2, \dots x_N) \tag{5}$$

5) Min Value

Represents the lowest amplitude value recorded in the EEG signal over a given interval [32]. This feature can indicate relaxation or inactivity in brain regions.

$$X_{min} = \min(x_1, x_2, \dots x_N) \tag{6}$$

6) Logarithmic Covariance Matrix

Used to measure statistical relationships between EEG electrodes. The logarithm of the covariance matrix enhances feature stability for emotion classification or neurological disorder detection [33].

$$C_{log} = \log(\det(C)) \tag{7}$$

where $C$is the EEG covariance matrix.

7) Covariance Matrix

The covariance matrix measures the linear relationship between pairs of EEG electrodes [34]. It is commonly used in brain connectivity analysis and synchronization studies between brain regions.

$$C_{ij} = \frac{1}{N-1}\sum_{k=1}^{N}(x_{ik} - \mu_i)(x_{jk} - \mu_j) \tag{8}$$

where C_{ij} is the covariance between electrodes i and j; and μ_i, μ_j are the mean values of EEG signals at each electrode.

8) Shannon Entropy

Shannon entropy measures the diversity or uncertainty of the EEG signal [35]. A high value indicates high variability, while a low value suggests a more regular signal.

$$H = -\sum_{i=1}^{N} P(x_i)\log P(x_i) \tag{9}$$

where $P(x_i)$ represents the probability of the EEG signal value x_i.

9) Log-Energy Entropy

Log-Energy Entropy measures the strength of the EEG signal while considering energy variations [36]. It is commonly used in epilepsy detection and emotion classification.

$$H_{log} = \sum_{i=1}^{N} \log(x_i^2) \tag{10}$$

10) Fast Fourier Transform (FFT)

FFT converts the EEG signal from the time domain to the frequency domain, allowing the analysis of frequency components such as delta, theta, alpha, beta, and gamma waves [37].

$$X(f) = \sum_{n=0}^{N-1} x(n)e^{-j2\pi fn/N} \tag{11}$$

where $x(n)$ is the EEG signal in the time domain; and $X(f)$ is the transformed signal in the frequency domain. FFT is widely used for detecting specific brain activity based on frequency bands.

11) Autocorrelation

Autocorrelation measures how similar an EEG signal is to a time-shifted version of itself [38]. It is useful for detecting brain rhythms and repeating patterns in EEG signals.

$$R(\tau) = \sum_{n=0}^{N-\tau} x(n)x(n+\tau) \tag{12}$$

where τ is the time lag; and $R(\tau)$ is the autocorrelation coefficient. A high autocorrelation at a certain lag indicates a repeating pattern in the EEG signal, which can be used for brain rhythm analysis and wave synchronization.

C. Gated Recurrent Unit

The Gated Recurrent Unit (GRU) is a type of Recurrent Neural Network (RNN) introduced as an alternative to Long Short-Term Memory (LSTM) [39]. GRU is designed to process sequential data, such as text, EEG signals, and other time-series data, by enhancing the model's ability to handle long-term dependencies without suffering from the vanishing gradient problem, which often occurs in conventional RNNs [40]. Compared to LSTM, GRU has a simpler structure as it only uses two main gates, namely the reset gate and update gate, making it computationally lighter and faster to train [41]. GRU plays a crucial role in sequential data processing, especially in applications such as Natural Language Processing (NLP), signal analysis, and time-series classification like EEG signals [42]. One of its primary functions is maintaining long-term memory in sequential data, enabling the model to retain relevant information from previous steps within a sequence [43]. This feature is particularly beneficial in EEG signal analysis, where temporal patterns and interdependencies are critical for recognizing emotions or detecting neurological disorders [44].

Additionally, GRU serves as a filter for relevant information in a sequence using the update gate, which decides whether information from the previous step should be retained or updated [45]. Meanwhile, the reset gate allows the model to discard irrelevant information when necessary [46]. By leveraging both mechanisms, GRU efficiently retains long-term dependencies better than standard RNNs while remaining computationally lighter than LSTMs, making it ideal for real-time applications [47]. The Fig. 2 is the architecture of the Gated Recurrent Unit.

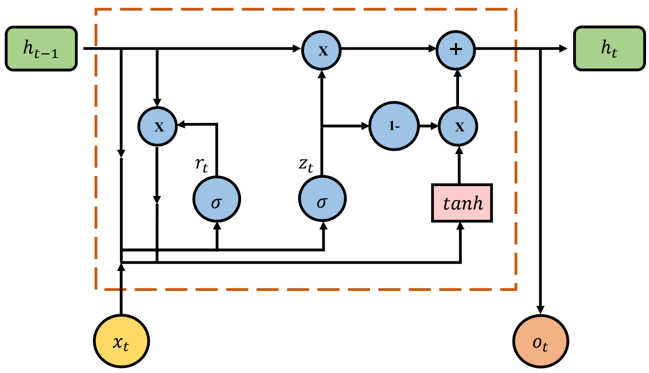


Fig. 2. Architecture of the Gated Recurrent Unit

1) Reset Gate (r_t)

The reset gate determines how much past information should be forgotten [48]. A small reset gate value causes the model to ignore previous information, allowing it to focus on new inputs.

$$r_t = \sigma(W_r.[h_{t-1}, x_t]) \tag{13}$$

where W_r represents the reset gate weight; h_{t-1} is the previous hidden state; and x_t is the input at time step t.

2) Update Gate (z_t)

The update gate determines how much information from the previous state should be carried forward [49]. If the update gate value is large, most of the previous information is retained; if it is small, the new input dominates.

$$z_t = \sigma(W_z.[h_{t-1}, x_t]) \tag{14}$$

where W_z represents the update gate weight.

3) New Candidate State Calculation ($\tilde{h}_t$)

After applying the reset gate, GRU computes the candidate new state $\tilde{h}_t$, representing the potential value for the current hidden state [50].

$$\tilde{h}_t = \tanh(W_h . [r_t \odot h_{t-1}, x_t]) \quad (15)$$

where the reset gate (r_t) controls how much information from the previous hidden state h_{t-1} is carried forward.

4) Final Hidden State Calculation (h_t)

A combination of past and new information is determined using the update gate, producing the final hidden state [51]:

$$h_t = z_t \odot h_{t-1} + (1 - z_t) \odot \tilde{h}_t \quad (16)$$

where the update gate (z_t) regulates the balance between the old state and the new state in determining the next hidden state.

The following is a proposed GRU model for EEG-based emotion classification.

The Fig 3. illustrates the architecture of a GRU-based model for EEG-based emotion classification. It consists of multiple Gated Recurrent Unit (GRU) layers followed by a dense layer with a softmax activation function to classify emotions into three categories. This structured pipeline ensures the effective extraction of temporal dependencies in EEG signals, enabling accurate emotion recognition. The process starts with EEG Data, which consists of raw brainwave signals collected from EEG sensors. These signals contain time-series information that reflects different emotional states. Since raw EEG data is highly complex and contains noise, the next step is EEG Feature Extraction, where various signal processing techniques are applied to derive meaningful features from the data. Following feature extraction, the first GRU layer is applied with 128 units. This layer processes the sequential EEG features and captures long-term dependencies in brainwave activity. The high number of units allows the model to learn rich temporal patterns that are crucial for distinguishing different emotional states. The second GRU layer (GRU_1) reduces the number of units to 64, which helps in refining the learned features while reducing computational complexity.

The third (GRU_2) and fourth (GRU_3) GRU layers further reduce the hidden state size to 32 units, enabling the model to progressively refine high-level temporal features. These layers ensure that the most relevant information is retained while minimizing redundant patterns, making the network more efficient and generalizable for real-world EEG-based emotion classification. After the sequential processing through the GRU layers, the model includes a Dense Layer with 3 neurons, corresponding to the three emotion classes (e.g., positive, neutral, and negative emotions). The softmax activation function is then applied, converting the output into probability distributions, ensuring that each instance is classified into one of the three emotion categories. Finally, the model produces the Emotion Classification output, which provides the final prediction based on the learned EEG features.

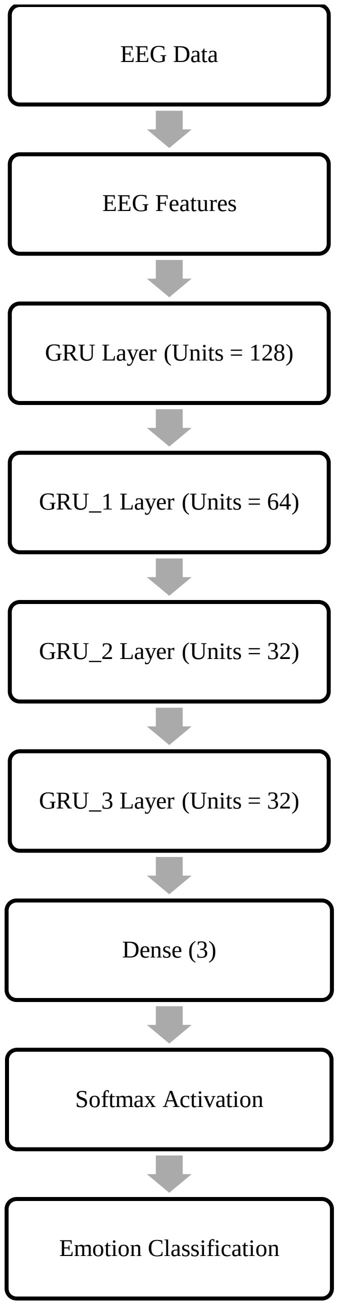


Fig. 3. The proposed GRU architecture

Table I shows the Keras implementation of the proposed GRU architecture.

TABLE I. KERAS IMPLEMENTATION OF THE PROPOSED GRU

Layer (type)	Output Shape	Param #
gru (GRU)	(None, 1, 128)	1028352
gru_1 (GRU)	(None, 1, 64)	37248
gru_2 (GRU)	(None, 1, 32)	9408
gru_3 (GRU)	(None, 32)	6336
dense (Dense)	(None, 3)	99
Total params: 1,081,443		
Trainable params: 1,081,443		
Non-trainable params: 0		

The Table I presents a deep learning model designed for EEG-based emotion classification, utilizing Gated Recurrent Unit (GRU) layers followed by a fully connected dense layer. The model consists of four GRU layers with progressively decreasing hidden state sizes, which help in capturing temporal dependencies in EEG signals while gradually refining the learned features. The total number of trainable parameters in this architecture is 1,081,443, indicating a complex model capable of handling high-dimensional EEG data. The first GRU layer has 128 hidden units and contains 1,028,352 parameters, making it the most computationally expensive component of the model. This layer processes the raw EEG signals, extracting initial features related to brainwave activity and emotional states. The large number of parameters arises from the weight matrices used for input-to-hidden and hidden-to-hidden connections, which are essential for learning sequential patterns in the EEG signals.

As the data flows through the model, the second GRU layer (gru_1) reduces the hidden state size to 64 units, significantly decreasing the number of parameters to 37,248. This layer continues refining time-dependent patterns extracted from EEG signals, ensuring that relevant information is passed while reducing computational complexity. The third (gru_2) and fourth (gru_3) GRU layers further reduce the hidden state to 32 units, containing 9,408 and 6,336 parameters, respectively. These layers help extract high-level temporal representations that are useful for distinguishing emotional states. Finally, the dense layer, with 99 parameters, maps the GRU output to a 3-class classification, indicating that the model is trained to classify EEG signals into three emotional states (positive, neutral, and negative emotions). This fully connected layer is responsible for making the final prediction based on the learned representations from the GRU layers. Fig 4. shows the proposed GRU architecture with input and output vector shapes.

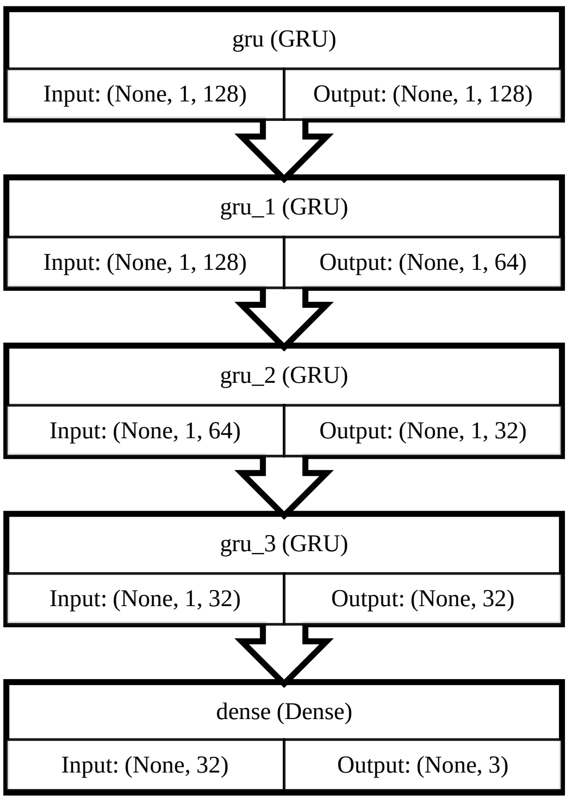


Fig. 4. Proposed GRU architecture with input and output vector forms

D. Performance Evaluation

Evaluating the performance of EEG-based emotion classification is crucial to ensure the reliability and effectiveness of the model [52]. Several key metrics are commonly used, including Sensitivity (Recall), Specificity, Precision, Accuracy, and F1 Score. These metrics help assess how well the model distinguishes between different emotional states and provide insights into false positives and false negatives, which are critical in brain-computer interface (BCI) and affective computing applications [53].

1) Sensitivity (Recall)

Sensitivity (also known as Recall) measures the ability of the model to correctly identify positive instances among all actual positive samples [54]. In EEG-based emotion classification, it quantifies how well the model detects a particular emotional state (e.g., positive emotion) when it truly occurs.

$$Sensitivity = \frac{TP}{TP + FN} \tag{17}$$

where: TP (True Positive) = correctly classified positive samples; and FN (False Negative) = misclassified positive samples. A high Sensitivity value means that the model rarely misses a positive case, which is crucial in applications like mental health monitoring, where failing to detect distress or anxiety could be problematic.

2) Specificity

Specificity measures the model's ability to correctly classify negative instances while minimizing false positives [55]. In EEG-based emotion recognition, it indicates how well the model avoids incorrectly labeling a neutral or negative emotion as positive.

$$Specificity = \frac{TN}{TN + FP} \tag{18}$$

where: TN (True Negative) = correctly classified negative samples; and FP (False Positive) = misclassified negative samples. A high Specificity ensures that the model is not overly biased toward predicting positive emotions, which is essential in maintaining balanced classification performance.

3) Precision

Precision evaluates the proportion of correctly classified positive instances among all predicted positive cases [56]. This metric is essential in EEG emotion classification, especially in scenarios where false positives need to be minimized, such as when diagnosing stress levels.

$$Precision = \frac{TP}{TP + FP} \tag{19}$$

High Precision means that when the model predicts a specific emotional state, it is likely to be correct, reducing false alarms in emotion classification.

4) Accuracy

Accuracy represents the overall correctness of the model by measuring the proportion of correctly classified instances across all categories [57]. It is a general measure of model performance in multi-class EEG emotion classification.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (20)$$

While Accuracy is a useful metric, it may not always be reliable if the dataset is imbalanced. For instance, if one emotional category (e.g., neutral) dominates the dataset, the model may have high accuracy but poor performance in minority classes.

5) F1 Score

The F1 Score is the harmonic mean of Precision and Recall (Sensitivity), providing a balanced measure that considers both false positives and false negatives. This metric is particularly useful in EEG-based emotion classification when dealing with imbalanced datasets, ensuring that the model performs well across all emotional categories [58].

$$F1\,Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (21)$$

A high F1 Score means the model balances both identifying positive cases correctly (Recall) and reducing false positives (Precision), making it a reliable metric for evaluating EEG emotion classification models.

In addition, a 5-fold cross-validation method is applied to test the performance of the EEG-based emotion classification model using a Gated Recurrent Unit (GRU). The EEG dataset is divided into five parts, where alternately four parts are used as training data and one part as testing data. This approach was taken to ensure that the evaluation results of the GRU model are objective, consistent, and can be generalised well to unfamiliar data [59].

III. Results and Discussions

This study focuses on optimizing the GRU model to improve the accuracy and efficiency of EEG-based emotion classification. To achieve this goal, there is a selection of EEG signal features for the classification process and modifications to the GRU architecture layer used. The signal features used include mean, standard deviation, statistical moments (skewness and kurtosis), min-max value, logarithmic covariance matrix, covariance matrix, shannon entropy, log-energy entropy, FFT, and autocorrelation. These features are extracted in the alpha and beta bands of the EEG waveform. In addition, the proposed GRU model for the classification process consists of four GRU layers with decreasing hidden state sizes, which helps capture temporal dependencies in the EEG signal while gradually refining the learned features [60]. The total number of trainable parameters in this architecture is 1,081,443, indicating a complex model capable of handling high-dimensional EEG data. Meanwhile, the dense layer maps the GRU output to a 3-class classification, which shows that this model was trained to classify EEG signals into three emotional states (positive, neutral and negative emotions).

To ensure the robustness and generalizability of the model, 5-fold cross-validation was conducted during the training and testing phases [61]. This technique helps reduce overfitting and provides a more accurate estimate of the model’s performance on unseen data [62]. Each fold serves as a test set once while the remaining folds are used for training, ensuring that all data points contribute to both training and evaluation [63]. The Table II is a confusion matrix resulting from EEG-based emotion classification testing using the GRU model.

TABLE II. Confusion Matrix of the Proposed GRU

Emotion	**Confusion Matrix**			
	TP	TN	FP	FN
Negative	133	289	2	3
Neutral	142	281	3	1
Positive	147	279	0	1

The confusion matrix of the proposed GRU model demonstrates its effectiveness in EEG-based emotion classification, with high True Positive (TP) and True Negative (TN) values across all emotion categories (Negative, Neutral, and Positive). Meanwhile, the low FP and FN values across all categories indicate that the model has high accuracy, precision, and recall, making it a reliable tool for EEG-based emotion recognition with minimal misclassification errors [64]. From the previous confusion matrix table, the performance values of the proposed GRU model are obtained as Table III.

TABLE III. Performance Values of the Proposed GRU

Emotion	**Performance Evaluation**				
	Recall	Specificity	Precision	Accuracy	F1 Score
Negative	97.79	99.31	98.52	98.83	98.15
Neutral	99.30	98.94	97.93	99.06	98.61
Positive	99.32	100.00	100.00	99.77	99.66
Average	98.81	99.42	98.82	99.22	98.81

From the table, the model demonstrates high recall values across all emotion categories, with the Negative class at 97.79%, Neutral at 99.30%, and Positive at 99.32%, resulting in an average recall of 98.81%. This indicates that the model correctly identifies most instances of each emotion, with very few false negatives. The specificity values are also notably high, with the Positive class achieving 100% specificity, meaning it has no false positives, while the Negative and Neutral classes have 97.31% and 98.94% specificity, respectively. This shows that the model is highly effective at correctly identifying non-target classes, reducing misclassification. The precision values further confirm the model’s reliability, with the highest precision observed in the Positive class (100%), meaning every detected positive emotion was classified correctly. The Negative and Neutral classes also exhibit high precision scores of 98.52% and 97.93%, respectively, leading to an overall precision of 98.82%. Additionally, the accuracy of the model is consistently high across all emotion categories, with an average of 99.22%, demonstrating that the model can generalize well across different emotions. Finally, the F1 Score, which balances precision and recall, is also significantly high across all classes, with an average F1 Score of 98.81%, indicating that the model maintains strong performance without sacrificing either precision or recall.

In addition, the Fig. 5 is a comparison of the performance values of the proposed GRU with Deep Learning (DL) models such as RNN and LSTM.

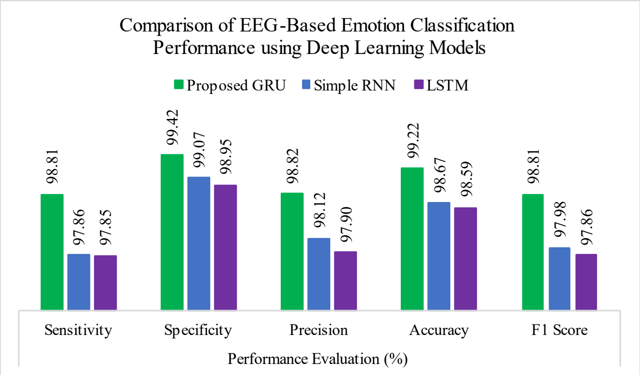
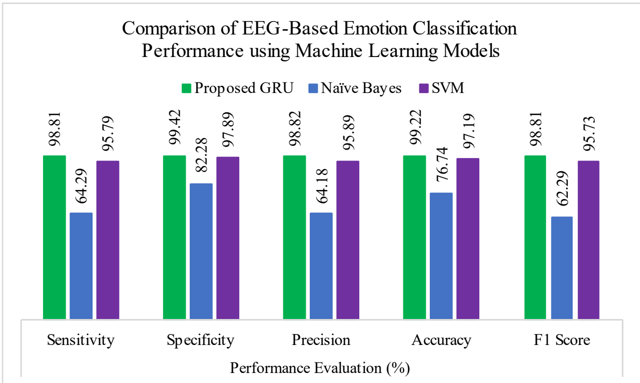


Fig. 5. Comparison of Emotion Classification Performance using DL

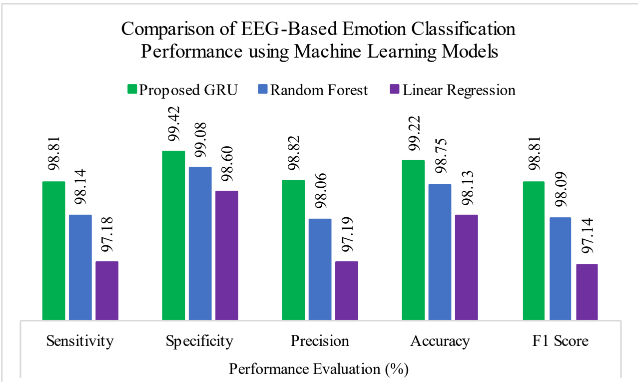
The comparison of EEG-based emotion classification performance using different deep learning models (Proposed GRU, Simple RNN, and LSTM) demonstrates the superiority of GRU across all key performance metrics (Sensitivity, Specificity, Precision, Accuracy, and F1 Score). The results indicate that the Proposed GRU model consistently outperforms Simple RNN and LSTM, making it a more reliable approach for emotion recognition based on EEG signals [65]. The Proposed GRU model achieves the highest Sensitivity (98.81%), indicating that it effectively detects emotional states while minimizing false negatives. In comparison, Simple RNN and LSTM both achieve 97.86%, suggesting that these models are slightly less effective at correctly identifying emotional states. Similarly, Specificity, which measures the ability to correctly classify non-target classes, is highest for GRU (99.42%), while Simple RNN and LSTM both score 98.95%. This suggests that GRU has fewer misclassifications, making it more precise in distinguishing between different emotional states, a crucial aspect in affective computing and mental health monitoring applications [66].

The Precision metric further confirms GRU’s superiority, achieving 98.82%, compared to LSTM (97.90%) and Simple RNN (98.12%). A higher precision score means that GRU produces fewer false positives, making it more trustworthy in emotion recognition tasks, particularly in real-time applications such as brain-computer interfaces (BCI), where accurate and immediate classification is essential [67]. Additionally, Accuracy, which measures the proportion of correctly classified instances across all emotion categories, is highest for GRU (99.22%), compared to Simple RNN (98.67%) and LSTM (98.59%). This highlights GRU’s ability to generalize better across different emotional states, reducing errors and increasing the overall reliability of the model. The F1 Score, which balances Precision and Recall, further validates GRU’s performance, scoring 98.81%, compared to Simple RNN (97.98%) and LSTM (97.86%). This confirms that GRU maintains a strong balance between correctly identifying emotions and minimizing misclassifications, making it a robust choice for EEG-based emotion classification. The higher F1 Score and accuracy of GRU suggest that it is better suited for capturing temporal dependencies in EEG signals, making it more effective than Simple RNN and computationally more efficient than LSTM [68].

The superior performance of GRU can be attributed to its gated architecture, which effectively handles long-term dependencies in sequential EEG data [69]. Unlike Simple RNN, which suffers from the vanishing gradient problem, GRU incorporates reset and update gates, enabling it to retain relevant information over long sequences without excessive computational cost [70]. Compared to LSTM, which has three gates (input, forget, and output gates), GRU’s two-gate mechanism makes it computationally lighter while maintaining similar or better performance [71]. Studies such as Rivas et al. (2025) and Glenn et al. (2023) support this finding, showing that GRU balances efficiency and performance better than LSTM, particularly in sequential data tasks such as EEG analysis [72, 73]. Furthermore, recent studies on EEG-based emotion recognition (Chowdary et al., 2022; Abgeena et al., 2023) confirm that GRU-based architectures outperform both RNN and LSTM due to their ability to extract highly relevant temporal features while reducing unnecessary complexity [74, 75]. This makes GRU particularly advantageous for real-time emotion recognition applications, where models must be both accurate and computationally efficient to provide immediate feedback in affective computing, mental health monitoring, and neuroscience applications [76]. Meanwhile, if the proposed GRU model's classification performance value is compared with several Machine Learning (ML) models such as SVM, K-NN, Random Forest, Decision Trees, and Naive Bayes, the Fig. 6 comparison results are obtained.



(a) Proposed GRU vs Naïve Bayes and SVM



(b) Proposed GRU vs Random Forest and Linear Regression

Fig. 6. Comparison of Emotion Classification Performance using ML

The performance comparison analysis between the Gated Recurrent Unit (GRU) and several traditional Machine Learning models, such as Naïve Bayes, Support Vector Machine (SVM), Random Forest, and Linear Regression, in EEG-based emotion classification demonstrates the significant superiority of GRU in handling sequential data. The evaluation is based on five key metrics, such as Sensitivity, Specificity, Precision, Accuracy, and F1 Score. The results show that GRU outperforms all Machine Learning models in every evaluation metric, confirming its effectiveness in capturing temporal patterns in EEG signals. In the comparison between GRU and Naïve Bayes and SVM, GRU achieves the highest Sensitivity (98.81%), significantly outperforming Naïve Bayes (64.29%) and SVM (95.70%). Higher Sensitivity means that GRU better detects emotions accurately, with fewer False Negative (FN) cases. GRU also exhibits higher Specificity (99.42%) compared to Naïve Bayes (63.18%) and SVM (98.52%), indicating that this model is more accurate in avoiding false positive predictions. Additionally, GRU's Precision (98.82%) surpasses that of Naïve Bayes (58.18%) and SVM (98.50%), proving that GRU's predictions are more reliable and less prone to errors. The overall Accuracy of GRU (99.22%) also exceeds Naïve Bayes (76.49%) and SVM (99.06%), highlighting GRU's ability to consistently classify various emotion categories more effectively.

Meanwhile, in the comparison between GRU and Random Forest and Linear Regression, GRU once again demonstrates superior performance. GRU achieves the highest Specificity (99.42%), compared to Random Forest (98.08%) and Linear Regression (69.08%), reaffirming that GRU makes fewer errors in classifying non-target emotions. GRU's Precision (98.82%) is also higher than that of Random Forest (97.19%) and Linear Regression (59.08%), meaning GRU's predictions are more specific and accurate. GRU's Accuracy (99.22%) outperforms Random Forest (98.53%) and Linear Regression (79.50%), confirming that GRU can generalize better than rule-based models like Random Forest and regression-based models like Linear Regression. The superiority of GRU over traditional Machine Learning models can be explained by its ability to capture temporal dependencies in EEG signals [77], something that Naïve Bayes, SVM, Random Forest, and Linear Regression cannot do, as they rely on statistical or decision tree-based approaches [78]. Machine Learning models such as Random Forest and SVM perform better on static feature-based data but are not designed to analyze time-dependent patterns deeply [79]. Naïve Bayes assumes independence between features, making it unsuitable for handling EEG data that heavily relies on temporal context [80]. Linear Regression performs the worst, as it works with simple linear relationships and cannot capture the complex, non-linear patterns present in brain signals [81].

IV. CONCLUSION

The findings of this study demonstrate the effectiveness of the optimized Gated Recurrent Unit (GRU) model in improving the accuracy and efficiency of EEG-based emotion classification. By integrating a feature selection process and modifying the GRU architecture, the model successfully captured temporal dependencies in EEG signals, enabling highly accurate classification of positive, neutral, and negative emotions. The high recall (98.81%), specificity (99.42%), precision (98.82%), accuracy (99.22%), and F1 Score (98.81%) confirm the robustness of the proposed GRU model in handling high-dimensional EEG data. Comparisons with Simple RNN and LSTM reveal that GRU consistently outperforms both models in all key performance metrics, demonstrating its computational efficiency and superior classification accuracy. Additionally, GRU surpasses traditional Machine Learning models such as Naïve Bayes, SVM, Random Forest, and Linear Regression, highlighting its ability to effectively process sequential EEG data, which is crucial for real-time affective computing and brain-computer interface (BCI) applications. Despite its impressive performance, several future research directions can be explored to further enhance EEG-based emotion classification. Firstly, expanding the dataset by including a more diverse group of participants with varying emotional responses could improve the generalizability of the model. Secondly, integrating multi-modal fusion techniques, such as combining EEG with facial expression recognition or physiological signals (e.g., heart rate variability), may enhance the model's ability to detect emotions more comprehensively. Additionally, exploring attention mechanisms or hybrid architectures (such as Transformer-GRU models) could further refine feature extraction and classification accuracy.

ACKNOWLEDGMENT

The authors would like to acknowledge the Department of Medical Technology, Institut Teknologi Sepuluh Nopember, for the facilities and support in this research. The authors also gratefully acknowledge financial support from the Institut Teknologi Sepuluh Nopember for this work, under project scheme of the Publication Writing and IPR Incentive Program (PPHKI) 2025.

REFERENCES

[1] A. Prakash and A. Poulose, "Electroencephalogram-Based Emotion Recognition: A Comparative Analysis of Supervised Machine Learning Algorithms," *Data Science and Management*, Jan. 2025, doi: 10.1016/j.dsm.2024.12.004.

[2] D. S. Menefee, T. Ledoux, and C. A. Johnston, "The Importance of Emotional Regulation in Mental Health," *American Journal of Lifestyle Medicine*, vol. 16, no. 1, pp. 28–31, Jan. 2022, doi: 10.1177/15598276211049771.

[3] M. Magalhães, A. Coelho, M. Melo, and M. Bessa, "Measuring users' emotional responses in multisensory virtual reality: a systematic literature review," *Multimedia tools and applications*, vol. 83, no. 14, pp. 43377–43417, Oct. 2023, doi: 10.1007/s11042-023-16918-1.

[4] C. Yu and M. Wang, "Survey of emotion recognition methods using EEG information," *Cognitive Robotics*, vol. 2, pp. 132–146, 2022, doi: 10.1016/j.cogr.2022.06.001.

[5] K. Erat, E. B. Şahin, F. Doğan, N. Merdanoğlu, A. Akcakaya, and P. O. Durdu, "Emotion recognition with EEG-based brain-computer interfaces: a systematic literature review," *Multimedia tools and applications*, vol. 83, no. 33, pp. 79647-79694, 2024, doi: 10.1007/s11042-024-18259-z.

[6] E. H. Houssein, A. Hammad, and A. A. Ali, "Human emotion recognition from EEG-based brain–computer interface using machine learning: a comprehensive review," *Neural Computing and Applications*, vol. 34, no. 15, pp. 12527-12557, 2022, doi: 10.1007/s00521-022-07292-4.

[7] T. Warbrick, “Simultaneous EEG-fMRI: What Have We Learned and What Does the Future Hold?,” *Sensors*, vol. 22, no. 6, p. 2262, Mar. 2022, doi: 10.3390/s22062262.

[8] G. Papanastasiou, A. Drigas, C. Skianis, and M. Lytras, “Brain computer interface-based applications for training and rehabilitation of students with neurodevelopmental disorders. A literature review,” Heliyon, vol. 6, no. 9, p. e04250, Sep. 2020, doi: 10.1016/j.heliyon.2020.e04250.

[9] R. Rai, K. Rai, A S. Mamatha, and K. Sooda, “Advancements in epilepsy classification: Current trends and future directions,” *MethodsX*, vol. 14, p. 103257, Jun. 2025, doi: 10.1016/j.mex.2025.103257.

[10] P. Samal and M. F. Hashmi, “Role of machine learning and deep learning techniques in EEG-based BCI emotion recognition system: a review,” *Artificial intelligence review*, vol. 57, no. 3, Feb. 2024, doi: 10.1007/s10462-023-10690-2.

[11] H. Albaqami, G. M. Hassan, A. Subasi, and A. Datta, “Automatic detection of abnormal EEG signals using wavelet feature extraction and gradient boosting decision tree,” *Biomedical Signal Processing and Control*, vol. 70, p. 102957, Sep. 2021, doi: 10.1016/j.bspc.2021.102957.

[12] J.-G. Wang, H.-M. Shao, Y. Yao, J.-L. Liu, H.-P. Sun, and S.-W. Ma, “Electroencephalograph-based emotion recognition using convolutional neural network without manual feature extraction,” *Applied Soft Computing*, vol. 128, p. 109534, Oct. 2022, doi: 10.1016/j.asoc.2022.109534.

[13] V. Marques, Alexandria, J. Alexandre, D. Ferreira, P. C. Motta, and B. Riccelli, “Emotion Detection from EEG Signals Using Machine Deep Learning Models,” *Bioengineering*, vol. 11, no. 8, pp. 782–782, Aug. 2024, doi: 10.3390/bioengineering11080782.

[14] J.-H. Lee, J.-Y. Kim, and H.-G. Kim, “Emotion Recognition Using EEG Signals and Audiovisual Features with Contrastive Learning,” *Bioengineering*, vol. 11, no. 10, p. 997, Oct. 2024, doi: 10.3390/bioengineering11100997.

[15] D. Kulkarni and V. V. Dixit, “EEG-based emotion classification Model: Combined model with improved score level fusion,” *Biomedical Signal Processing and Control*, vol. 95, pp. 106352–106352, Apr. 2024, doi: 10.1016/j.bspc.2024.106352.

[16] J. Dai, X. Xi, G. Li, and T. Wang, “EEG-Based Emotion Classification Using Improved Cross-Connected Convolutional Neural Network,” *Brain Sciences*, vol. 12, no. 8, p. 977, Jul. 2022, doi: 10.3390/brainsci12080977.

[17] S. Reis, L. P.-Coelho, M. Sousa, M. Neto, and M. Silva, “Advancing Emotion Recognition: EEG Analysis and Machine Learning for Biomedical Human–Machine Interaction,” *BioMedInformatics*, vol. 5, no. 1, pp. 5–5, Jan. 2025, doi: 10.3390/biomedinformatics5010005.

[18] M. Zuo, X. Chen, and L. Sui, “Evaluation of Machine Learning Algorithms for Classification of Visual Stimulation-Induced EEG Signals in 2D and 3D VR Videos,” *Brain Sciences*, vol. 15, no. 1, pp. 75–75, Jan. 2025, doi: 10.3390/brainsci15010075.

[19] F. Li, F. Yu, L. Shen, H. Li, X. Yang, and Q. Shen, “EEG-Based Emotion Recognition with Combined Fuzzy Inference via Integrating Weighted Fuzzy Rule Inference and Interpolation,” *Mathematics*, vol. 13, no. 1, p. 166, Jan. 2025, doi: 10.3390/math13010166.

[20] J. A. Cruz-Vazquez, J. Y. Montiel-Pérez, R. Romero-Herrera, and E. Rubio-Espino, “Emotion Recognition from EEG Signals Using Advanced Transformations and Deep Learning,” *Mathematics*, vol. 13, no. 2, p. 254, Jan. 2025, doi: 10.3390/math13020254.

[21] S. Gannouni, A. Aledaily, K. Belwafi, and H. Aboalsamh, “Emotion detection using electroencephalography signals and a zero-time windowing-based epoch estimation and relevant electrode identification,” *Scientific Reports*, vol. 11, no. 1, Mar. 2021, doi: 10.1038/s41598-021-86345-5.

[22] J. T. Bird, L. Manso, E. P. Ribeiro, A. Ekárt, and D. R. Faria, “A Study on Mental State Classification using EEG-based Brain-Machine Interface,” *International Conference on Intelligent Systems*, pp. 795-800, Sep. 2018, doi: 10.1109/is.2018.8710576.

[23] B. Tripathi and R. K. Sharma, “EEG-Based Emotion Classification in Financial Trading Using Deep Learning: Effects of Risk Control Measures,” *Sensors*, vol. 23, no. 7, p. 3474, Mar. 2023, doi: 10.3390/s23073474.

[24] J. J. Bird, D. R. Faria, L. J. Manso, A. Ekárt, and C. D. Buckingham, “A Deep Evolutionary Approach to Bioinspired Classifier Optimisation for Brain-Machine Interaction,” *Complexity*, vol. 2019, pp. 1–14, Mar. 2019, doi: 10.1155/2019/4316548.

[25] Y. Pamungkas and U. W. Astuti, “Comparison of Human Emotion Classification on Single-Channel and Multi-Channel EEG using Gate Recurrent Unit Algorithm,” *2023 International Conference on Computer Science, Information Technology and Engineering (ICCoSITE)*, pp. 1-6, Feb. 2023, doi: 10.1109/iccosite57641.2023.10127686.

[26] A. Chaddad, Y. Wu, R. Kateb, and A. Bouridane, “Electroencephalography Signal Processing: A Comprehensive Review and Analysis of Methods and Techniques,” *Sensors*, vol. 23, no. 14, pp. 6434–6434, Jul. 2023, doi: 10.3390/s23146434.

[27] A. D. Wibawa, N. Fatih, Y. Pamungkas, M. Pratiwi, P. A. Ramadhani, and Suwadi, “Time and Frequency Domain Feature Selection Using Mutual Information for EEG-based Emotion Recognition,” *2022 9th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, pp. 19-24, Oct. 2022, doi: 10.23919/eecsi56542.2022.9946522.

[28] S. Morales and M. E. Bowers, “Time-frequency analysis methods and their application in developmental EEG data,” *Developmental Cognitive Neuroscience*, vol. 54, p. 101067, 2022, doi: 10.1016/j.dcn.2022.101067.

[29] H. Azami *et al.*, “EEG Entropy in REM Sleep as a Physiologic Biomarker in Early Clinical Stages of Alzheimer’s Disease,” *Journal of Alzheimer s Disease*, vol. 91, no. 4, pp. 1557–1572, Jan. 2023, doi: 10.3233/jad-221152.

[30] J. Mathew, N. Sivakumaran, and P. A. Karthick, “Automated Detection of Seizure Types from the Higher-Order Moments of Maximal Overlap Wavelet Distribution,” *Diagnostics*, vol. 13, no. 4, p. 621, Feb. 2023, doi: 10.3390/diagnostics13040621.

[31] A. Topic and M. Russo, “Emotion recognition based on EEG feature maps through deep learning network,” *Engineering Science and Technology, an International Journal*, vol. 24, no. 6, pp. 1442-1454, Apr. 2021, doi: 10.1016/j.jestch.2021.03.012.

[32] I. Stancin, M. Cifrek, and A. Jovic, “A Review of EEG Signal Features and Their Application in Driver Drowsiness Detection Systems,” *Sensors*, vol. 21, no. 11, p. 3786, May 2021, doi: 10.3390/s21113786.

[33] H. Liu, Y. Zhang, Y. Li, and X. Kong, “Review on Emotion Recognition Based on Electroencephalography,” *Frontiers in Computational Neuroscience*, vol. 15, Oct. 2021, doi: 10.3389/fncom.2021.758212.

[34] G. Chiarion, L. Sparacino, Y. Antonacci, L. Faes, and L. Mesin, “Connectivity Analysis in EEG Data: A Tutorial Review of the State of the Art and Emerging Trends,” *Bioengineering*, vol. 10, no. 3, pp. 372–372, Mar. 2023, doi: 10.3390/bioengineering10030372.

[35] A. Cacciotti, C. Pappalettera, F. Miraglia, P. M. Rossini, and F. Vecchio, “EEG entropy insights in the context of physiological aging and Alzheimer’s and Parkinson’s diseases: a comprehensive review,” *GeroScience*, May 2024, doi: 10.1007/s11357-024-01185-1.

[36] L. Abou-Abbas, I. Jemal, K. Henni, Y. Ouakrim, A. Mitiche, and N. Mezghani, “EEG Oscillatory Power and Complexity for Epileptic Seizure Detection,” *Applied Sciences*, vol. 12, no. 9, pp. 4181–4181, Apr. 2022, doi: 10.3390/app12094181.

[37] A. F. Pérez *et al.*, “Development of RelaxQuest: A Serious EEG-Controlled Game Designed to Promote Relaxation and Self-Regulation with a Potential Focus on ADHD Intervention,” *Applied Sciences*, vol. 14, no. 23, pp. 11173–11173, Nov. 2024, doi: 10.3390/app142311173.

[38] H. Zhang *et al.*, “The applied principles of EEG analysis methods in neuroscience and clinical neurology,” *Military Medical Research*, vol. 10, no. 1, p. 67, Dec. 2023, doi: 10.1186/s40779-023-00502-7.

[39] Y. Zhang and H. D. Fill, “TS-GRU: A Stock Gated Recurrent Unit Model Driven via Neuro-Inspired Computation,” *Electronics*, vol. 13, no. 23, pp. 4659–4659, Nov. 2024, doi: 10.3390/electronics13234659.

[40] X. Kong *et al.*, “Deep learning for time series forecasting: a survey,” *International Journal of Machine Learning and Cybernetics*, Feb. 2025, doi: 10.1007/s13042-025-02560-w.

[41] D. Rohan, G. P. Reddy, V. Pavan, K. P. Prakash, and C. P. Reddy, “An extensive experimental analysis for heart disease prediction using

artificial intelligence techniques," *Scientific Reports*, vol. 15, no. 1, Feb. 2025, doi: 10.1038/s41598-025-90530-1.

[42] K. Zhao and X. Guo, "PilotCareTrans Net: an EEG data-driven transformer for pilot health monitoring," *Frontiers in Human Neuroscience*, vol. 19, Jan. 2025, doi: 10.3389/fnhum.2025.1503228.

[43] I. Malashin, V. Tynchenko, A. Gantimurov, V. Nelyub, and A. Borodulin, "Applications of Long Short-Term Memory (LSTM) Networks in Polymeric Sciences: A Review," *Polymers*, vol. 16, no. 18, pp. 2607–2607, Sep. 2024, doi: 10.3390/polym16182607.

[44] S. M. Al-Selwi *et al.*, "RNN-LSTM: From applications to modeling techniques and beyond—Systematic review," *Journal of King Saud University - Computer and Information Sciences*, vol. 36, no. 5, p. 102068, Jun. 2024, doi: 10.1016/j.jksuci.2024.102068.

[45] I. D. Mienye, T. G. Swart, and G. Obaido, "Recurrent Neural Networks: A Comprehensive Review of Architectures, Variants, and Applications," *Information*, vol. 15, no. 9, pp. 517–517, Aug. 2024, doi: 10.3390/info15090517.

[46] X. Wu, C. Shi, and L. Yan, "Driving Attention State Detection Based on GRU-EEGNet," *Sensors*, vol. 24, no. 16, p. 5086, Aug. 2024, doi: 10.3390/s24165086.

[47] C. Jiang *et al.*, "EEG Emotion Recognition Employing RGPCN-BiGRUAM: ReliefF-Based Graph Pooling Convolutional Network and BiGRU Attention Mechanism," *Electronics*, vol. 13, no. 13, pp. 2530–2530, Jun. 2024, doi: 10.3390/electronics13132530.

[48] P. B. Weerakody, K. W. Wong, and G. Wang, "Cyclic Gate Recurrent Neural Networks for Time Series Data with Missing Values," *Neural Processing Letters*, vol. 55, no. 2, pp. 1527-1554, Jul. 2022, doi: 10.1007/s11063-022-10950-2.

[49] X. Wu, Y. Zhang, J. Li, H. Yang, and X. Wu, "FC-TFS-CGRU: A Temporal–Frequency–Spatial Electroencephalography Emotion Recognition Model Based on Functional Connectivity and a Convolutional Gated Recurrent Unit Hybrid Architecture," *Sensors*, vol. 24, no. 6, pp. 1979–1979, Mar. 2024, doi: 10.3390/s24061979.

[50] M. Çelebi, S. Öztürk, and K. Kaplan, "An emotion recognition method based on EWT-3D–CNN–BiLSTM-GRU-AT model," *Computers in Biology and Medicine*, vol. 169, pp. 107954–107954, Feb. 2024, doi: 10.1016/j.compbiomed.2024.107954.

[51] W. Jiang et al., "Fuzzy ensemble-based federated learning for EEG-based emotion recognition in Internet of Medical Things," *Journal of Industrial Information Integration*, pp. 100789–100789, Feb. 2025, doi: 10.1016/j.jii.2025.100789.

[52] R. Pillalamarri and U. Shanmugam, "A review on EEG-based multimodal learning for emotion recognition," *Artificial Intelligence Review*, vol. 58, no. 5, Feb. 2025, doi: 10.1007/s10462-025-11126-9.

[53] E. P. Torres, E. A. Torres, M. Hernández-Álvarez, and S. G. Yoo, "EEG-Based BCI Emotion Recognition: A Survey," *Sensors*, vol. 20, no. 18, p. 5083, Sep. 2020, doi: 10.3390/s20185083.

[54] S. A. Hicks *et al.*, "On evaluation metrics for medical applications of artificial intelligence," *Scientific Reports*, vol. 12, no. 1, p. 5979, Apr. 2022, doi: 10.1038/s41598-022-09954-8.

[55] T. F. Monaghan *et al.*, "Foundational Statistical Principles in Medical Research: Sensitivity, Specificity, Positive Predictive Value, and Negative Predictive Value," *Medicina*, vol. 57, no. 5, p. 503, May 2021, doi: 10.3390/medicina57050503.

[56] I. M. De Diego, A. R. Redondo, R. R. Fernández, J. Navarro, and J. M. Moguerza, "General Performance Score for classification problems," *Applied Intelligence*, vol. 52, no. 10, pp. 12049-12063, Jan. 2022, doi: 10.1007/s10489-021-03041-7.

[57] A. E. Maxwell, T. A. Warner, and L. A. Guillén, "Accuracy Assessment in Convolutional Neural Network-Based Deep Learning Remote Sensing Studies—Part 1: Literature Review," *Remote Sensing*, vol. 13, no. 13, p. 2450, Jun. 2021, doi: 10.3390/rs13132450.

[58] Z. He *et al.*, "Unified Convolutional Sparse Transformer for Disease Diagnosis, Monitoring, Drug Development, and Therapeutic Effect Prediction from EEG Raw Data," *Biology*, vol. 13, no. 4, pp. 203–203, Mar. 2024, doi: 10.3390/biology13040203.

[59] G. Brookshire *et al.*, "Data leakage in deep learning studies of translational EEG," *Frontiers in Neuroscience*, vol. 18, May 2024, doi: 10.3389/fnins.2024.1373515.

[60] M. Bouchane, W. Guo, and S. Yang, "Hybrid CNN-GRU Models for Improved EEG Motor Imagery Classification," *Sensors*, vol. 25, no. 5, pp. 1399–1399, Feb. 2025, doi: 10.3390/s25051399.

[61] J. Pachouly, S. Ahirrao, K. Kotecha, G. Selvachandran, and A. Abraham, "A systematic literature review on software defect prediction using artificial intelligence: Datasets, Data Validation Methods, Approaches, and Tools," *Engineering Applications of Artificial Intelligence*, vol. 111, no. 1, p. 104773, May 2022, doi: 10.1016/j.engappai.2022.104773.

[62] C. Aliferis and G. Simon, "Overfitting, Underfitting and General Model Overconfidence and Under-Performance Pitfalls and Best Practices in Machine Learning and AI," *Computers in health care (New York)*, pp. 477–524, Jan. 2024, doi: 10.1007/978-3-031-39355-6_10.

[63] Y. Ma *et al.*, "Emotion Recognition Model of EEG Signals Based on Double Attention Mechanism," *Brain Sciences*, vol. 14, no. 12, pp. 1289–1289, Dec. 2024, doi: 10.3390/brainsci14121289.

[64] T. Wang, X. Huang, Z. Xiao, W. Cai, and Y. Tai, "EEG emotion recognition based on differential entropy feature matrix through 2D-CNN-LSTM network," *EURASIP Journal on Advances in Signal Processing*, vol. 2024, no. 1, Apr. 2024, doi: 10.1186/s13634-024-01146-y.

[65] X. Wang, Y. Ren, Z. Luo, W. He, J. Hong, and Y. Huang, "Deep learning-based EEG emotion recognition: Current trends and future perspectives," *Frontiers in Psychology*, vol. 14, Feb. 2023, doi: 10.3389/fpsyg.2023.1126994.

[66] S. E. Güler and F. P. Akbulut, "Multimodal Emotion Recognition: Emotion Classification through the Integration of EEG and Facial Expressions," *IEEE Access*, vol. 13, pp. 24587-24603, Jan. 2025, doi: 10.1109/access.2025.3538642.

[67] H. Torkey, S. Hashish, S. Souissi, E. E.-D. Hemdan, and A. Sayed, "Seizure Detection in Medical IoT: Hybrid CNN-LSTM-GRU Model with Data Balancing and XAI Integration," *Algorithms*, vol. 18, no. 2, pp. 77–77, Feb. 2025, doi: 10.3390/a18020077.

[68] M. Khalid *et al.*, "Diagnosing epileptic seizures using combined features from independent components and prediction probability from EEG data," *Digital Health*, vol. 10, Jan. 2024, doi: 10.1177/20552076241277185.

[69] M. A. Afzal, Z. Gu, S. U. Bukhari, and B. Afzal, "Brainwaves in the Cloud: Cognitive Workload Monitoring Using Deep Gated Neural Network and Industrial Internet of Things," *Applied Sciences*, vol. 14, no. 13, pp. 5830–5830, Jul. 2024, doi: 10.3390/app14135830.

[70] M. Waqas and U. W. Humphries, "A Critical Review of RNN and LSTM Variants in Hydrological Time Series Predictions," *MethodsX*, vol. 13, pp. 102946–102946, Sep. 2024, doi: 10.1016/j.mex.2024.102946.

[71] I. H. Sarker, "Deep Learning: a Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions," *SN Computer Science*, vol. 2, no. 6, Aug. 2021, doi: 10.1007/s42979-021-00815-1.

[72] F. Rivas, J. E. S.-Garcia, and J. M. Camara, "Comparison of LSTM- and GRU-Type RNN Networks for Attention and Meditation Prediction on Raw EEG Data from Low-Cost Headsets," *Electronics*, vol. 14, no. 4, pp. 707–707, Feb. 2025, doi: 10.3390/electronics14040707.

[73] A. Glenn, P. LaCasse, and B. Cox, "Emotion classification of Indonesian Tweets using Bidirectional LSTM," *Neural Computing and Applications*, vol. 35, no. 13, pp. 9567-9578, Feb. 2023, doi: 10.1007/s00521-022-08186-1.

[74] M. K. Chowdary, J. Anitha, and D. J. Hemanth, "Emotion Recognition from EEG Signals Using Recurrent Neural Networks," *Electronics*, vol. 11, no. 15, p. 2387, Jul. 2022, doi: 10.3390/electronics11152387.

[75] A. Abgeena and S. Garg, "S-LSTM-ATT: a hybrid deep learning approach with optimized features for emotion recognition in electroencephalogram," *Health Information Science and Systems*, vol. 11, no. 1, Aug. 2023, doi: 10.1007/s13755-023-00242-x.

[76] J. Z. Wang *et al.*, "Unlocking the Emotional World of Visual Media: An Overview of the Science, Research, and Impact of Understanding Emotion," *Proceedings of the IEEE*, vol. 111, no. 10, pp. 1–51, 2023, doi: 10.1109/JPROC.2023.3273517.

[77] K. L. Lew, K. S. Sim, and Z. Ting, "Deep Learning Approach EEG Signal Classification," *JOIV: International Journal on Informatics*

Visualization, vol. 8, no. 3–2, p. 1693, Nov. 2024, doi: 10.62527/joiv.8.3-2.2959.

[78] A. D. Wibawa, B. S. Y. Mohammad, M. A. K. Fata, F. A. Nuraini, A. Prasetyo, and Y. Pamungkas, "Comparison of EEG-Based Biometrics System Using Naive Bayes, Neural Network, and Support Vector Machine," *2022 International Conference on Electrical and Information Technology (IEIT)*, pp. 408-413, Sep. 2022, doi: 10.1109/ieit56384.2022.9967861.

[79] I. H. Sarker, "Machine Learning: Algorithms, Real-World Applications and Research Directions," *SN Computer Science*, vol. 2, no. 3, pp. 1–21, 2021, doi: 10.1007/s42979-021-00592-x.

[80] S. S. Prasetiyowati and Y. Sibaroni, "Unlocking the potential of Naive Bayes for spatio temporal classification: a novel approach to feature expansion," *Journal of Big Data*, vol. 11, no. 1, Aug. 2024, doi: 10.1186/s40537-024-00958-x.

[81] M.-A. Schulz *et al.*, "Different scaling of linear models and deep learning in UKBiobank brain images versus machine-learning datasets," *Nature Communications*, vol. 11, no. 1, 2020, doi: 10.1038/s41467-020-18037-z.