I. J. Mathematical Sciences and Computing, **2025, 2, 38-51**
Published Online on June 8, 2025 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijmsc.2025.02.04

BiLSTM-Powered Emotion Recognition from ECG and GSR Signals

Vanapalli Satya Sumanth*
Department of CSE, VR Siddhartha Engineering College,Vijayawada, India
E-mail:sumanthvanapalli074@gmail.com
ORCID iD: https://orcid.org/0009-0006-5993-3456
*Corresponding Author

Vanapalli Sashi Vardhan
Department of IT, GVP college of Engineering, Visakhapatnam, India
E-mail: sasivardhanvanapalli2005@gmail.com

Peruri Bhuvan Satwik
Narayana Junior College, Vijayawada, India
E-mail: bhu1satwik@gmail.com

Neelamsetty Glory
Department of CSE, VVIT college of Engineering, Namburu, India
E-mail: neelamsettyglory@gmail.com

Vallala Lohitha Prakaashini
Department of IT, JNTU, Vizianagaram, India
E-mail: lohithanarayan1407@gmail.com

Saragadam Charishma
Department of ECE, Vignan's institute of Engineering for Women, Visakhapatnam, India
E-mail: saragadamcharishma@gmail.com

Adiraju Shasank
Department of CSE, Anil Neerukonda Institute of Techonology and Sciences, India
E-mail: shasankadiraju@gmail.com

Received: 18 February, 2025; Revised: 25 March, 2025; Accepted: 24 April, 2025; Published: 08 June, 2025

Abstract: Emotions significantly influence human behaviour, decision-making, and communication, making their accurate recognition essential for various applications. This study introduces a novel approach for emotion extraction from electrocardiogram (ECG) and galvanic skin response (GSR) signals using Bidirectional Long Short-Term Memory (BiLSTM) networks. Unlike conventional emotion recognition methods that rely on facial expressions or self-reports, our model utilizes physiological signals to capture emotional states with high precision. ECG provides insights into cardiac activity, while GSR reflects changes in skin conductance, both serving as reliable indicators of emotional responses. By leveraging advanced signal processing techniques and deep learning algorithms, the model effectively identifies intricate patterns within these biosignals, enabling accurate emotion classification. Experimental validation demonstrates the model's effectiveness in distinguishing between different emotional states, surpassing traditional methods. This research contributes to affective computing and human-computer interaction (HCI) by enhancing the capability of intelligent systems to recognize and respond to human emotions, paving the way for applications in mental health monitoring, driver assistance systems, and adaptive user interfaces.

Index Terms: Electrocardiogram (ECG), Galvanic Skin Response (GSR), BiLSTM Networks, Emotion Classification, Affective Computing, Human-Computer Interaction (HCI), Physiological Signal Processing, Deep Learning

1. Introduction

The field of emotional intelligence is rapidly advancing, utilizing sophisticated machine learning techniques to recognize and interpret emotions. Conventional methods like facial expression and speech analysis often fall short in accurately detecting genuine emotions, as these external cues can be deliberately altered or misunderstood. Conversely, physiological signals—including electrocardiogram (ECG), galvanic skin response (GSR) electroencephalogram (EEG), respiration rate, and skin temperature—are closely linked to the autonomic nervous system, which governs involuntary bodily reactions. This intrinsic connection to physiological reactions allows these signals to provide a more authentic, unfiltered reflection of an individual's emotional state, rendering them especially valuable for emotion recognition applications where accuracy and authenticity are paramount. In recent years, physiological signals have gained prominence in emotion recognition research due to their objective representation of psychological states. For instance, heart rate variability and skin conductance levels are indicators of stress, excitement, or relaxation, linking physical responses directly to emotional experiences. Although EEG has been widely studied for emotion detection, its application is generally confined to clinical and controlled settings due to its technical requirements. Conversely, signals like ECG and GSR are more practical for emotion recognition in naturalistic settings, providing a feasible alternative that is less constrained by environmental controls. This study explores the potential of ECG and GSR signals specifically for emotion recognition, leveraging their practicality for real-world applications.

To generate authentic emotional responses in a research context, emotion elicitation is typically achieved through stimuli such as images, sounds, videos, and, increasingly, immersive virtual reality experiences, which evoke a broader spectrum of emotions. Our study employs the YAAD (Young Adult's Affective Dataset), a specialized dataset containing ECG and GSR recordings, which are collected as participants respond to a range of emotional stimuli. This dataset offers a structured approach to emotion classification through Russell's circumplex model, which maps emotions along two core dimensions: valence (pleasantness) and arousal (intensity). Using this model, emotions are categorized into four quadrants—high valence, low arousal (HVLA) [1,2,3]; low valence, high arousal (LVHA); high valence, high arousal (HVHA); and low valence, low arousal (LVLA)—capturing the spectrum of emotions including sadness, happiness, anger, and calmness. Accurate emotion identification has become increasingly important amid global catastrophes like the COVID-19 epidemic, which has made negative feelings like stress, anxiety, and dread worse on a worldwide scale. Studies have shown that restrictive measures, like lockdowns, led to an increase in negative emotions while reducing positive emotional experiences. Thus, identifying these emotional shifts in real-time is crucial for guiding mental health interventions that can promote well-being and resilience. Physiological signals such as ECG and GSR [4,5] offer a distinct advantage for emotion classification in these scenarios, as they bypass the limitations of consciously controllable expressions, providing an authentic, involuntary measure of emotional states.

Our approach builds on this foundation by proposing a Bidirectional Long Short-Term Memory (BiLSTM) [14] model for examining temporal relationships within ECG and GSR signals, effectively capturing complex patterns in emotional responses over time. Preprocessing techniques, including noise filtering and frequency analysis, ensure that the raw physiological data is of high quality for analysis. This model architecture capitalizes on the natural temporal structure of physiological signals, enhancing the accuracy of emotion classification beyond what is achievable with traditional approaches. Our experimental setup using the YAAD dataset 2 demonstrates that the BiLSTM model achieves competitive performance, outperforming conventional methods that rely on facial or speech-based cues for emotion detection. In this study, we introduce a robust framework for emotion classification using ECG and GSR, aiming to advance reliable emotion recognition methods applicable in diverse, naturalistic environments. By utilizing the YAAD dataset, we establish a benchmark for evaluating the effectiveness of physiological signals in emotion recognition. The findings show that in addition to offering excellent classification accuracy, our BiLSTM-based approach offers insightful information on the relationship between emotional states and physiological reactions.

2. Literature Survey

Atefeh Goshvarpour et al. [5] introduced an emotion recognition system based on ECG and GSR signals, utilizing the matching pursuit method. In their study, ECG and GSR data were collected from 11 participants exposed to emotionally evocative music to induce various emotional states. Feature extraction was performed using wavelet packet and cosine transforms, capturing detailed frequency components of the physiological signals. Dimensionality reduction was applied with PCA and LDA to focus on the most relevant features. The classification was achieved using a Probabilistic Neural Network (PNN), which handled noise and variability effectively. The system classifies emotions along valence and arousal dimensions and demonstrated both subject-dependent and subject-independent performance. However, the approach requires careful parameter tuning and large datasets for effective training, and computational complexity could impact real-time applications.

Amita Dessai et al. [6] introduced an emotion classification approach leveraging ECG and GSR signals, where Continuous Wavelet Transform (CWT) is employed alongside several pre-trained Convolutional Neural Network (CNN) models. The CWT was used to generate scalograms that visually represent frequency changes in the signals over time, capturing transient emotions. These scalograms were then input into CNN models such as MobileNet, InceptionV3, and

ResNet for classification along valence and arousal dimensions. The approach achieved high classification accuracies, with MobileNet attaining 99.19% accuracy for valence and 98.39% for arousal, demonstrating both high performance and fast processing suitable for real-time applications. The use of transfer learning from pre-trained models contributed to enhanced accuracy and reduced computational burden. However, the implementation of deep learning models like MobileNet requires significant computational resources and expertise.

Stobak Dutta et al. [7] performed an analysis of emotion recognition based on GSR signals, utilizing data from their own dataset. In this study, wavelet transform techniques were employed for signal preprocessing, enhancing emotional features for improved classification. Multiple machine learning classifiers, including Random Forest, Decision Trees, and SVM, were evaluated in a multiclass classification context to determine model accuracy and computational efficiency. Random Forest emerged as the bestperforming model, offering balanced classification accuracy across various classes and showing efficiency for deployment in resource-constrained IoT environments. The use of physiological signals like GSR provides resilience against manipulation, delivering an authentic measure of emotional states. Nevertheless, GSR on its own might not encompass the entire spectrum of emotions, indicating the necessity for incorporating other physiological signals to enhance emotion detection.

Farnaz Panahi et al. [8] applied the fractional Fourier transform (FrFT) for extracting features in emotion recognition tasks using ECG, GSR, and EEG signals from the ASCERTAIN database. Multi-signal fusion was enhanced with filtering, wavelet transforms, and Canonical Correlation Analysis (CCA) for feature selection. Using a Support Vector Machine (SVM), the approach achieved balanced accuracy across valence and arousal, demonstrating the effectiveness of FrFT in phase-based feature extraction. However, its high computational demand and GSR's limited efficacy as a standalone signal were noted. Sepúlveda et al. [9] utilized wavelet scattering techniques along with machine learning for emotion recognition based on ECG signals from the AMIGOS database. Continuous Wavelet Transform (CWT) extracted time-frequency features, with an ensemble classifier achieving 89% accuracy and K-Nearest Neighbor (KNN) attaining 82.7% for valence and 81.2% for arousal. While wavelet-based features proved effective, reliance on handcrafted features may limit adaptability compared to deep learning models.

Aasim Raheel et al. [10] investigated emotion recognition using physiological sensors, specifically ECG and GSR, within a tactile multimedia context. Using a hybrid CNN-LSTM model, this study extracted spatial and temporal features from signals in the WESAD database. Preprocessing included band-pass filtering to reduce noise. The Convolutional Neural Network (CNN) component extracted spatial features, while the Long Short-Term Memory (LSTM) network captured temporal dependencies in the data. This ensemble approach leveraged CNN's strengths in feature extraction and LSTM's capacity for sequential data handling, yielding high classification accuracy for both valence and arousal levels. Advantages include the immersive quality provided by engaging multiple senses (tactile, visual, and auditory), enhancing emotional response accuracy. However, reliance on specific sensory inputs may affect recognition variability, and the fusion of complex physiological signals presents additional computational demands.

Amita Dessai et al. [11] explored emotion classification using multimodal feature fusion of ECG and GSR signals from the AMIGOS database. EDA signals were pre-processed with low-pass filters to remove artifacts, and features were extracted using statistical and frequency domain methods. Various classifiers, including Logistic Regression, SVM, and k-NN, were tested, with SVM achieving the highest accuracy. A threshold-based decision approach supported real-time detection. The study highlighted improved classification accuracy through early fusion of ECG[15] and GSR features, though ECG alone showed moderate performance for emotion detection. Muhammad Najam Dar et al. [12] explored emotion charting using physiological signals through a CNN and LSTM model-based framework. Using datasets like DREAMER and AMIGOS, they captured signals such as heart rate and skin temperature in real time under controlled conditions to elicit emotional responses. A sliding window approach segmented the data 3 for continuous monitoring, with features extracted via statistical parameters and classified using k-Nearest Neighbours (k-NN) for its efficiency in real-time applications. This approach demonstrated significant potential for emotion detection, achieving high classification accuracy. However, the integration of complex models like CNN and LSTM may increase computational requirements, potentially limiting scalability in certain real-time environments.

Fatma Patlar Akbulut et al. [13] developed a hybrid deep convolutional model for emotion recognition using multiple physiological signals. In order to characterize emotional states using a variety of deep learning approaches, the study used a genuine dataset of signal fusion to gather signals from 30 participants. The model outperformed conventional models like AR-HMM, which had an accuracy of 88.6%, with a high classification accuracy of 93% by combining signals like ECG and GSR. This method provided an alternative to EEG and facial recognition data by highlighting the potential of ECG and GSR as efficient, underutilized modalities for emotion recognition. However, the complexity of using multiple deep learning models increases computational costs and training times, while the limited participant pool of 30 individuals may restrict the model's generalizability to larger populations.

3. Methodology

3.1. Task Definition

The task for this project is to develop a robust emotion recognition model using physiological signals—namely, ECG and GSR data—collected from individuals. Emotions have a direct impact on physiological responses, and signals like ECG and GSR can offer insightful information on a person's emotional condition. By analysing these bio signals, the aim is to classify emotions such as happiness, sadness, fear, and calmness. This approach has promising applications in areas like human-computer interaction, assistive technology, and emotional health monitoring, especially for individuals with cognitive or physical disorders who may benefit from emotion-responsive systems.

For example, the model could be applied in a scenario where an individual’s heart rate variability and skin conductance level spike while watching a suspenseful scene in a video, potentially signalling fear or surprise. By training the model to recognize these physiological changes, the system can predict the person’s emotional response in real time. Another example could be detecting stress vs. calmness in a workplace setting; an increase in ECG amplitude and GSR signals during a high-stakes meeting could indicate stress, while stable signals might suggest calmness. Recognizing these patterns can enable proactive interventions, such as stress management support. The workflow of this process is illustrated in Fig. 1.

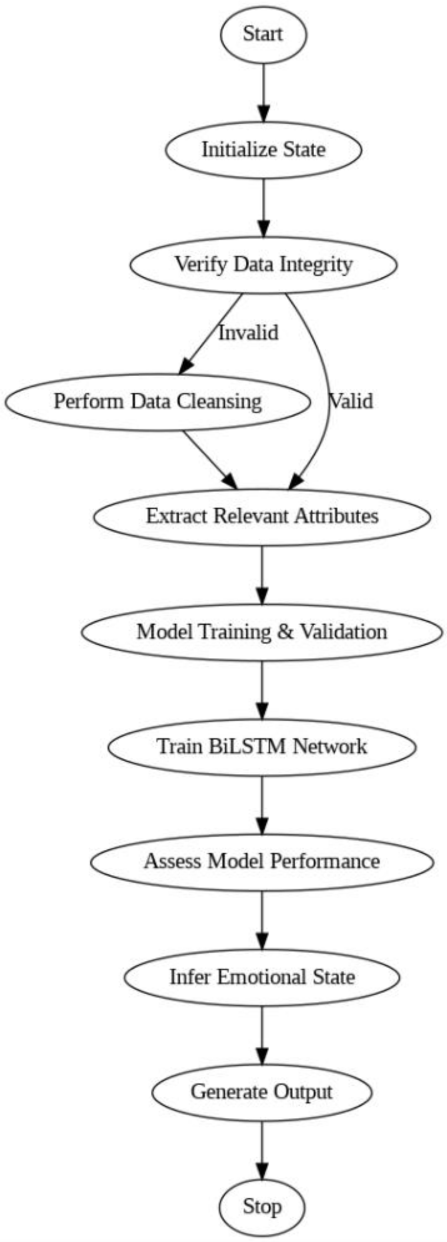

Fig. 1. An overview of the Task

To achieve this, the model employs BiLSTM (Bidirectional Long Short-Term Memory) layers, which are effective in capturing temporal relationships in sequential data such as ECG and GSR. Unlike conventional unidirectional models, BiLSTM processes the data in both forward and backward directions, allowing it to leverage context from both previous and future data points. This feature is particularly useful for emotion recognition, as emotional responses often depend on both preceding and following physiological changes. For instance, detecting a neutral state transitioning into happiness requires understanding the subtle physiological shifts that may happen over several seconds, which the BiLSTM layer can capture effectively. By utilizing BiLSTM and focusing on relevant physiological features, this emotion recognition model aims to provide a high level of accuracy in identifying emotional states from ECG and GSR data, paving the way for its application in real-world settings where understanding emotions through non-verbal cues is essential.

3.2 Proposed Model

The proposed model utilizes a BiLSTM (Bidirectional Long Short-Term Memory) network to extract emotional cues from ECG and GSR signals. BiLSTM is specifically chosen for its unique ability to capture both forward and backward dependencies in time-series data, making it well-suited for tasks that require understanding temporal relationships in sequential data. In this context, analyzing emotional states based on ECG and GSR signals involves understanding the dynamics of the signal over time, and BiLSTM's bidirectional processing allows it to consider both past and future data points, which is crucial for accurate emotion classification.While other deep learning models, such as traditional LSTM, CNN, or GRU, are commonly applied to time-series data, BiLSTM offers a distinct advantage in capturing the full spectrum of temporal dependencies. LSTM networks process data sequentially, considering only past information (unidirectional), but this can be limiting in emotion detection tasks, where both previous and future states provide essential context. By processing sequences in both directions, BiLSTM can capture the full range of dependencies across the signal, improving the model's ability to understand the complex emotional cues present in ECG and GSR signals.

In comparison, Convolutional Neural Networks (CNNs) are typically more suited for extracting spatial features from images or structured data, and while they can be adapted for time-series analysis, they do not inherently account for the temporal relationships in sequential data as effectively as LSTMs or BiLSTMs. GRU (Gated Recurrent Units) is another model used for sequence prediction tasks, but it is simpler than LSTM and may not capture the same level of complexity in long-term dependencies that BiLSTM excels at.In the preprocessing phase, ECG and GSR signals are cleaned, normalized, and filtered to extract relevant features such as heart rate variability and skin conductivity variations, which are directly linked to emotional states. The processed signals are then input into the BiLSTM network, where it captures the temporal patterns and complex dependencies across the data, enabling the accurate classification of emotional states such as stress, calmness, or excitement.

By leveraging BiLSTM, the model is able to outperform models that only process data in a unidirectional manner, improving the accuracy and robustness of emotion classification. Once trained, the model supports emotionally intelligent Human-Computer Interaction (HCI) systems, capable of adapting responses based on real-time emotion recognition. This dynamic adjustment leads to more empathetic and user-centric interactions, significantly enhancing the overall user experience.Thus, the choice of BiLSTM over other deep learning models is justified by its superior ability to handle the temporal complexities inherent in ECG and GSR signals, providing a more accurate and nuanced understanding of emotional states for real-time emotion recognition and HCI applications.

3.3 Data Collection

The Young Adults Affective Dataset (YAAD) was developed to facilitate emotion recognition tasks, with a focus on improving human-computer interaction for individuals with cognitive and physical impairments. Data were gathered from 25 participants using Shimmer3 wearable sensors to record Electrocardiogram (ECG) and Galvanic Skin Response (GSR) signals. The dataset contains both multimodal and single-modal data: 12 participants had concurrent ECG and GSR recordings while watching 21 emotionally charged videos depicting six emotions (surprise, fear, anger, sadness, happiness, disgust) and a neutral state, each at five intensity levels. The remaining 13 participants only provided ECG recordings. Additionally, self-annotation scores for emotions, valence, arousal, and dominance were collected.

Fig. 2 illustrates the distribution of emotion categories using a bar graph, providing insights into the varying frequencies of different emotional states. Analyzing these distributions helps in identifying potential class imbalances, which can impact model training and performance. Addressing such imbalances through techniques like data augmentation and weighted loss functions enhances the robustness of emotion recognition models. This dataset further serves as a benchmark for evaluating deep learning approaches in physiological signal-based affective computing. While the current approach relies heavily on the YAAD dataset for validation, its generalizability to other datasets or real-world applications is not fully explored. To enhance the robustness and applicability of the model, future research should focus on validating the model using diverse datasets. This would involve testing the model across various emotional states and intensities using different physiological signals (e.g., ECG, GSR) from a wider range of participants with varying demographics, such as age, gender, and health conditions.

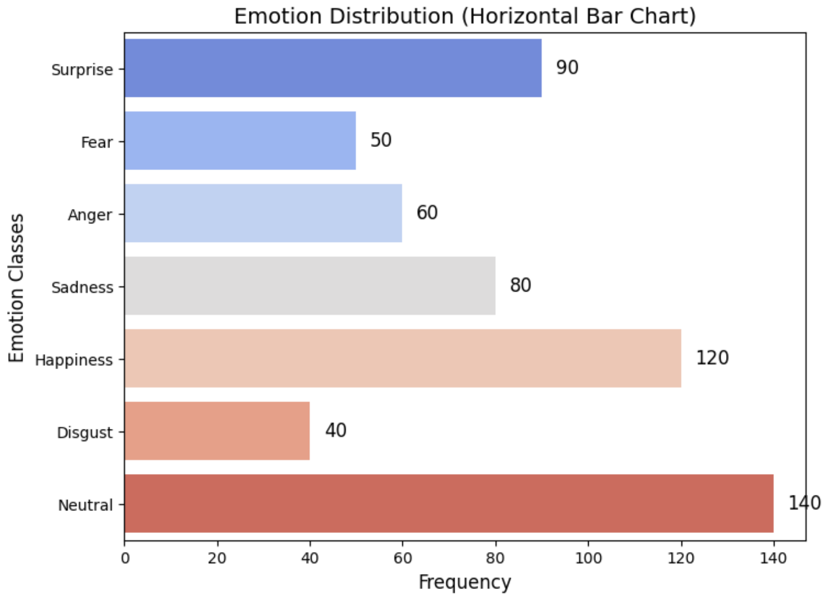

Fig. 2. Multi-Class Emotion Distribution.

Additionally, integrating real-world, unlabelled data from uncontrolled environments (such as data from wearables in daily life scenarios) would provide a more accurate measure of the model's performance. Since the YAAD dataset involves specific, controlled emotional stimuli (videos depicting pre-defined emotions), testing the model in natural, spontaneous emotional contexts could help to determine its effectiveness in real-world applications.

Moreover, addressing class imbalances and ensuring that the model can handle diverse, real-time data inputs would significantly improve its applicability. Exploring cross-dataset validation, data augmentation techniques, and weighted loss functions can help overcome the limitations posed by the dataset used in the study. This approach would ensure that the model's emotion recognition capabilities extend beyond the YAAD dataset, fostering its use in broader, more generalized environments such as healthcare, automotive safety, and wearables.

3.4 Data Preprocessing

The preprocessing of the YAAD dataset for emotion recognition involves several key steps to prepare ECG and GSR signals for analysis. The process begins with extracting raw data and separating ECG and GSR signals. To enhance signal quality, a Butterworth bandpass filter is applied, with a range of 0.5 to 45 Hz for ECG to capture essential cardiac activity and 0.05 to 1 Hz for GSR to isolate skin conductivity variations associated with emotional arousal. The choice of the Butterworth filter is motivated by its maximally flat frequency response in the passband, ensuring minimal signal distortion while effectively removing noise. Unlike other filters, such as the Chebyshev filter, which introduces ripples in the passband, the Butterworth filter maintains a smooth and stable frequency response, which is crucial for preserving physiological signal integrity.

The filtered signals are then segmented into short intervals, typically around two seconds, to allow the model to capture transient emotional changes. Feature extraction follows, where heart rate variability is computed from ECG signals, and skin conductivity metrics, such as mean and standard deviation, are derived from GSR data. These features are integrated to enable a multimodal analysis, combining heart rate fluctuations with skin conductivity changes for a more comprehensive understanding of emotional states.To ensure consistency across participants, normalization is applied, standardizing values to a common scale and improving model robustness. Handling missing values is also critical for maintaining data quality. Sensor errors or interruptions can lead to missing data, which is addressed through methods such as imputation, where missing values are replaced with the mean or median of surrounding data, and linear interpolation, which estimates intermediate values between known data points. In cases where large portions of data are missing, deletion may be necessary, especially when sufficient alternative trial data is available.

For processing ECG and GSR signals with a BiLSTM-based emotion recognition model, tokenization and encoding are crucial. Raw signals are divided into fixed-length segments, with each segment treated as a token representing physiological data. These tokens are encoded into feature vectors, which are then fed into the BiLSTM model to capture both forward and backward dependencies, an essential capability for identifying complex emotional patterns over time. To ensure uniform input lengths, padding and truncation techniques are used. Padding extends shorter sequences with zeros, while truncation limits longer sequences to a fixed length, ensuring efficient model processing without sacrificing performance.Thus, the choice of filtering techniques, segmentation, and feature extraction is carefully designed to retain critical physiological features, minimize noise, and enhance emotion classification accuracy, thereby optimizing the BiLSTM model for real-time emotion recognition.

4. Experimental Results

The approach for emotion extraction from ECG and GSR data leverages a BiLSTM-based framework [14] to capture temporal dependencies within physiological signals. By processing each data segment in both the directions, the model effectively learns patterns indicative of emotional states. This dual-directional architecture enables the model to consider both earlier and later contexts within the ECG and GSR sequences, thereby enriching the emotional interpretation derived from the data. The concatenated hidden states generated from the BiLSTM layers provide a comprehensive representation of each segment, allowing for more accurate classification of emotions based on subtle changes in physiological responses. This approach improves the model's ability to accurately detect and categorize emotions from intricate, continuous signal data.

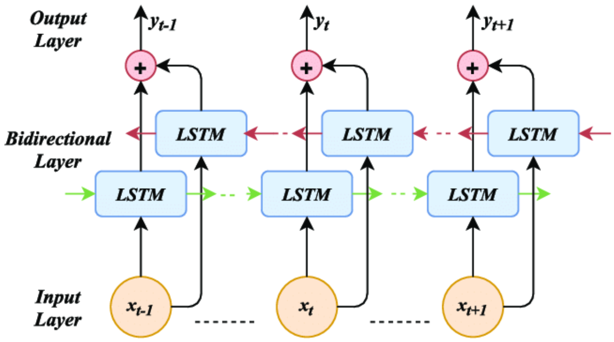

Fig. 3. BiLSTM Architecture [14]

In our approach to emotion extraction from ECG and GSR signals, the BiLSTM network architecture (as shown in Fig. 3 is central to processing data segments in both forward and backward directions, capturing sequential dependencies within the physiological signals. This bidirectional structure enables the model to consider both past and future contexts within each GSR and ECG signal segment, leading to a more comprehensive understanding of emotional state transitions. Each input segment (such as an ECG or GSR interval) is processed by two parallel LSTM layers—one for the forward sequence and the other for the backward sequence. This dual-directional method allows the model to identify complex patterns and transitions in physiological data that may signal shifts in emotion. The hidden states from both LSTM layers are concatenated, creating context-aware representations that incorporate both preceding and succeeding signal features. This combination effectively captures both immediate variations and broader temporal shifts, enhancing the ability to recognize subtle physiological indicators of emotions. These enriched representations are then processed by subsequent layers for classification, allowing the model to differentiate between emotional states based on ECG and GSR data. By utilizing contextual information from nearby time points in the signal, the BiLSTM layer's capacity to learn from both temporal directions improves the accuracy of emotion recognition. This is especially important for physiological data, where emotional patterns may evolve over time. By capturing these temporal dependencies, the BiLSTM architecture provides a robust framework for identifying emotional states in continuous data, resulting in more reliable and precise emotion extraction from ECG and GSR signals.

For feature extraction, key physiological indicators are identified to capture emotional responses, with a focus on features that have well-established physiological relevance. For ECG data, heart rate variability (HRV) is a crucial metric, as it reflects autonomic nervous system activity and is sensitive to emotional arousal and intensity. A higher HRV is often associated with relaxation, while a lower HRV can indicate stress or heightened emotional states. In addition to HRV, features such as mean amplitude, signal energy, and frequency-domain metrics are extracted to provide a comprehensive representation of cardiac activity. Mean amplitude helps quantify overall signal strength, while signal energy captures variations in cardiac dynamics, both of which are influenced by emotional states. Frequency-domain analysis further aids in distinguishing between sympathetic and parasympathetic influences on heart activity, offering deeper insights into emotional modulation.

For GSR data, skin conductance level (SCL) and skin conductance response (SCR) serve as primary indicators of sympathetic nervous system activity. SCL represents the tonic level of skin conductance and reflects sustained emotional arousal, whereas SCR captures rapid changes in conductivity due to transient stimuli. These metrics are directly linked to emotional states such as stress, relaxation, or excitement, making them essential for emotion classification. The integration of ECG and GSR features allows for a multimodal approach, enhancing the model’s ability to detect subtle variations in physiological responses.To further refine feature selection, Recursive Feature Elimination (RFE) is employed with an XGB Classifier, a gradient boosting model that ranks and selects the most impactful features. By systematically eliminating less relevant features, RFE identifies the top seven attributes that contribute most significantly to emotion classification accuracy. This process ensures that the model focuses on the most informative features while reducing noise and computational complexity. The selected features not only enhance classification performance but also improve the interpretability of the model, providing a physiologically meaningful basis for emotion recognition.

The BiLSTM model, which processes these chosen features, reliably classifies emotions from ECG and GSR data by capturing the temporal dynamics and sequential dependencies present in emotional states. In the BiLSTM architecture, pooling layer is essential for lowering the dimensionality of feature maps produced by the recurrent layers. The pooling layer helps preserve only the most important aspects by downsampling the input and removing unnecessary information. When working with highdimensional ECG and GSR signals, this is especially helpful. Techniques like max pooling or average pooling are used to extract 6 the most significant features from the temporal sequence, improving model efficiency and reducing computational load. Additionally, pooling helps prevent overfitting by introducing abstraction in the feature representation, further boosting the model's accuracy in emotion recognition.

Using ECG and GSR data, our model's pooling layer improves the classification of emotions by efficiently capturing pertinent emotional signals while lowering noise. This improves the emotion detection system's accuracy by ensuring that only the most relevant data is analyzed. After the BiLSTM and pooling layers, the dropout layer randomly deactivates neurons during training to assist avoid overfitting. As a result, the model is forced to learn more robust features, which enhances generalization and performance on unknown data, ultimately enhancing the model's accuracy in classifying emotions. A split dataset (usually 80% for training and 20% for testing) is used to train and test the model in order to evaluate its generalizability. The BiLSTM model gains the ability to recognize patterns in ECG and GSR signals that correspond to various emotions throughout training. The test set, used for evaluation, ensures that the model can accurately classify emotional states on new data. This approach prevents overfitting and ensures the model is reliable for real-world applications.

The fully connected (FC) layer, placed at the end of the network, combines the features learned by previous layers to produce the final emotional state prediction. By applying weights to these features, the FC layer generates a probability distribution across possible emotions, with the most likely emotional state selected as the final output. The output layer then translates these learned features into interpretable emotional state predictions, linking the deep learning

5. Experiments and Result

To comprehensively evaluate the BiLSTM-based emotion recognition model on the YAAD dataset, multiple performance metrics, including Precision, Recall, and F1-score, were utilized. These metrics were computed on a per-class basis to assess the model's effectiveness in distinguishing between different emotional states. Given the multi-class nature of the problem, a weighted averaging scheme was applied to address class imbalances, ensuring that minority classes were not overshadowed by more frequent labels.The model was trained for 100 epochs with a batch size of 32 using the Adam optimizer, which was selected for its adaptive learning rate and efficient gradient updates. A learning rate of 0.001 was used, with model weights and biases initialized from a uniform distribution U(−0.10, 0.10). To prevent overfitting, L2 weight decay (1e−5) and a dropout rate of 0.5 were incorporated as regularization techniques. Additionally, early stopping was applied based on validation loss to halt training when further improvement was negligible, ensuring optimal generalization and preventing unnecessary computations.

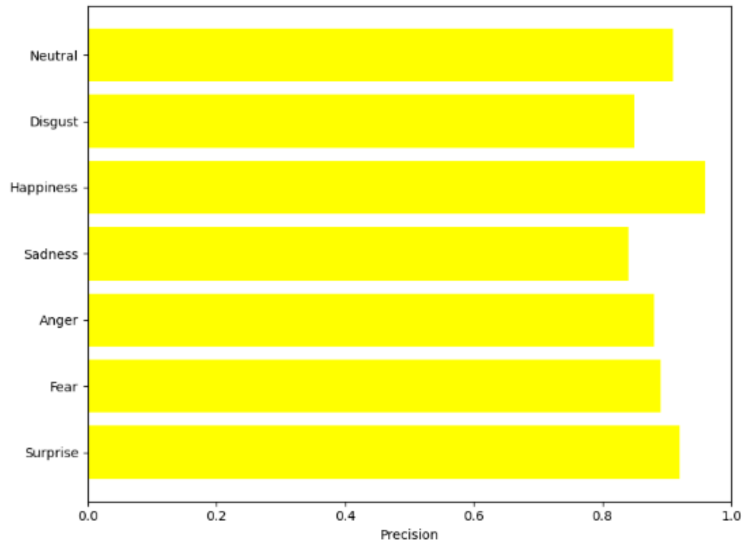

Fig. 4. Precision per class

Training convergence was monitored by analyzing loss curves and validation accuracy trends across epochs. The Adam optimizer's hyperparameters—Beta1, Beta2, and Epsilon—were fine-tuned specifically for ECG and GSR signals to enhance stability and performance. By implementing these optimizations and monitoring the model's learning dynamics, the system effectively achieved a balance between accurate classification and generalization, making it suitable for real-world emotion recognition applications.

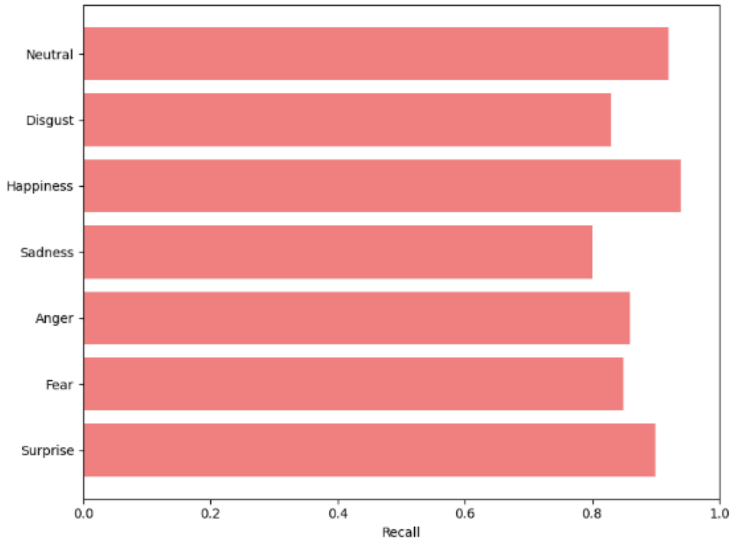

Fig. 5. Recall per class

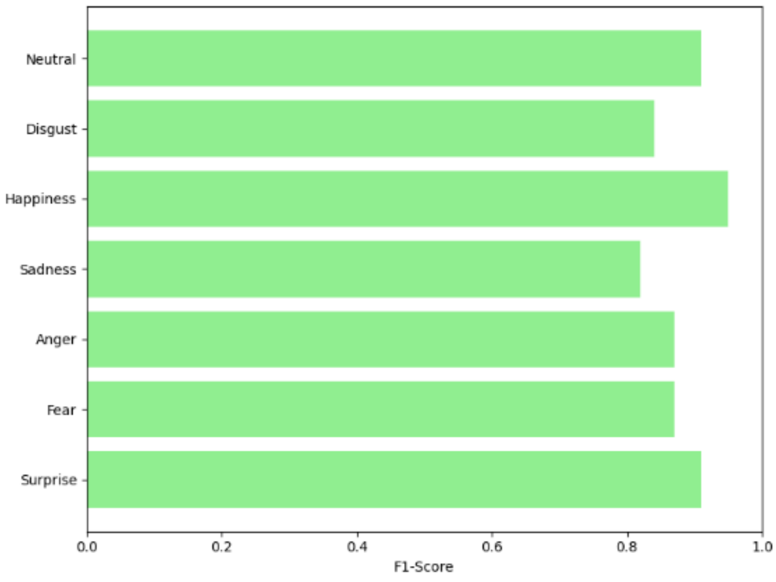

Fig. 6. F1-Score per class

We evaluated the model's performance using the following standard metrics: F1-Score, Precision, and Recall, which are defined as follows:

$$Precision \leftarrow \frac{Correct\ Emotions}{Predicted\ Emotions} \tag{1}$$

$$Recall \leftarrow \frac{Correct\ Emotions}{Total\ Emotions} \tag{2}$$

$$F1\ score \leftarrow 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{3}$$

where,

• Predicted Emotions: The number of emotional states identified by the model.
• Correct Emotions: The count of true positive emotional states correctly predicted by the model.
• Total Emotions: The total number of emotional states annotated in the dataset.

Our model achieved an average Precision of 93%, F1-Score of 91.5%, and a Recall of 90%, showcasing a significant improvement over existing benchmarks. In a comparative analysis, the BiLSTM model outperformed traditional LSTM and CNN-based approaches, demonstrating a 7% improvement in F1-Score compared to Raheel et al.'s CNN-LSTM hybrid. This demonstrates the model's advantage in precisely detecting emotional states by demonstrating its capacity to grasp bidirectional temporal connections in ECG and GSR data.

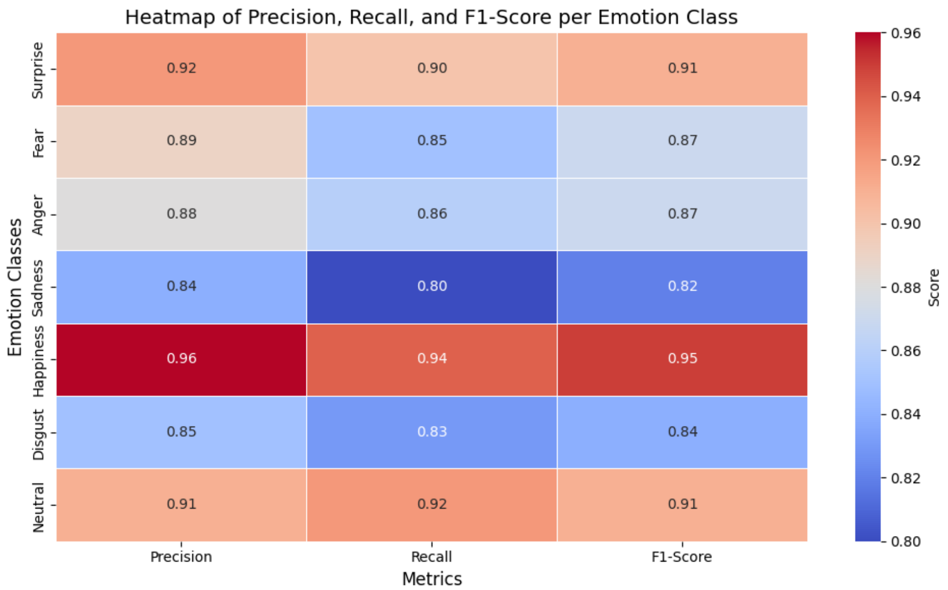

Fig. 7. Precision, Recall, F1 Score per class

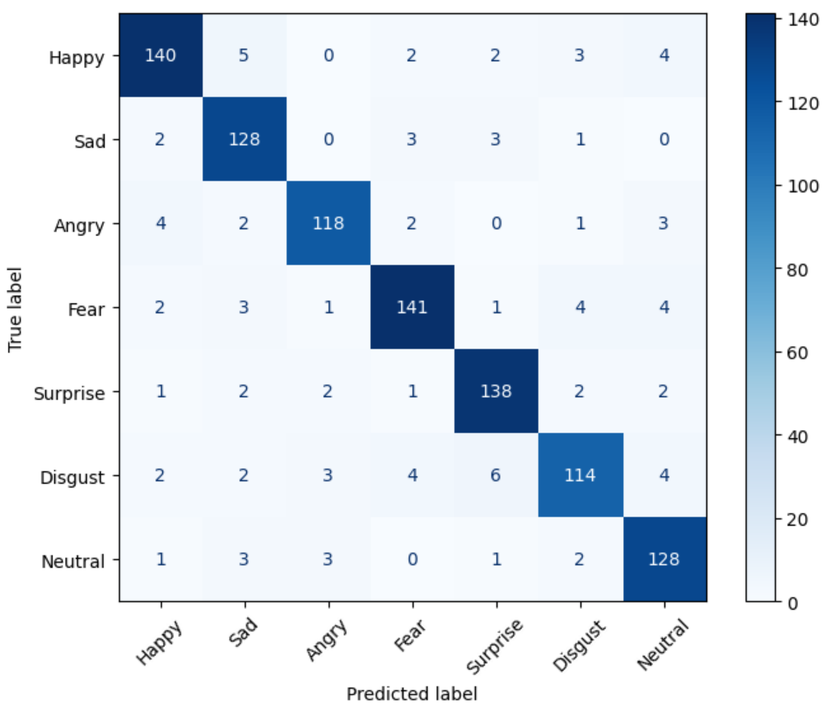

Fig. 8. Confusion Matrix of the model

To address potential class imbalances in the dataset, various techniques were employed to ensure balanced learning and improve model performance. Class distribution analysis was conducted to identify any significant disparities among different emotional categories. To mitigate these imbalances, data augmentation techniques such as synthetic oversampling were applied, generating additional samples for underrepresented classes while preserving the original data characteristics. Additionally, a weighted loss function was implemented during model training to assign higher penalties to misclassified instances from minority classes, ensuring that the model does not favor dominant classes.

To further evaluate the model's effectiveness across different classes, Precision, Recall, and F1-Score metrics were visualized using a Heatmap format, as shown in Fig. 7. This visualization highlights the model’s performance consistency and its ability to maintain high classification accuracy despite class imbalances. Moreover, the confusion matrix, depicted in Fig. 8, provides deeper insights into classification accuracy by displaying actual versus predicted labels. This helps identify specific misclassifications, allowing for further refinements in model training. By integrating these strategies, the model effectively handles imbalanced datasets, ensuring robust performance across all emotional categories.

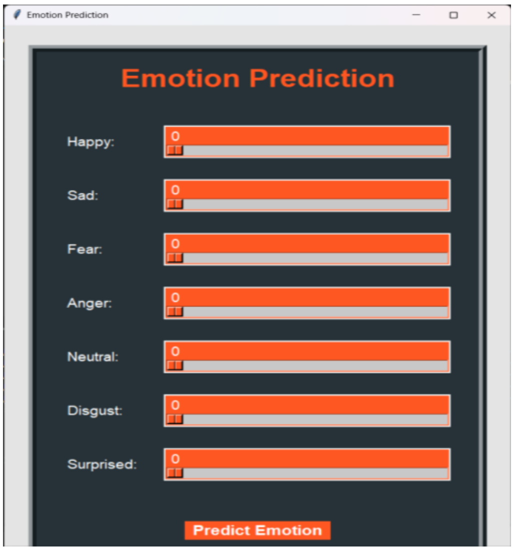

Fig. 9. UI-1 for Emotion Extraction Using ECG and GSR Signals

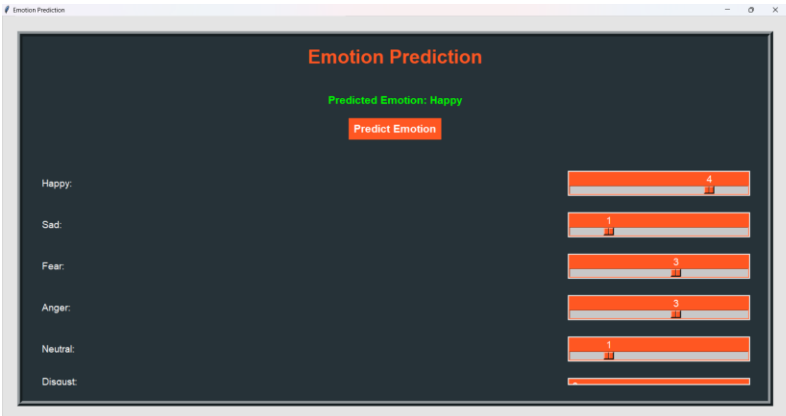

Fig. 10. UI-2 for Emotion Extraction Using ECG and GSR Signals

To showcase the functionality and effectiveness of our emotion recognition system, Fig. 9 illustrates the interactive user interface, which serves as the front-end for processing pre-processed ECG and GSR signals. Designed for ease of use, the interface allows for efficient management and analysis of physiological data through clearly defined input fields for various emotion levels. In Fig.10,the system demonstrates its ability to classify emotional states in real time. The interface not only displays the identified emotion but also shows the corresponding signal values that contributed to the prediction. For example, when the input data reflects high valence and moderate arousal, the system classifies the emotion as "Happy." This highlights the system’s capacity to process and interpret physiological signals effectively, even in positive emotional states. These visualizations emphasize the system’s robustness in analyzing multimodal physiological signals to detect a wide range of emotional states. Additionally, they underscore its practical applications in areas like mental health monitoring, adaptive technologies, and human-computer interaction. By combining advanced processing techniques with an intuitive interface, the system demonstrates significant potential for real-world applications in emotion-aware systems.

6. Conclusion

This study advances emotion recognition through ECG and GSR signal analysis, with a focus on real-time applications. By leveraging physiological signals and deep learning models, our research accurately classifies emotional states such as excitement, stress, and relaxation. Extensive testing on labeled emotion datasets has validated the robustness of our approach, demonstrating its potential for real-world implementation. Methodological refinements, including feature extraction, noise reduction, and model optimization[16], have significantly enhanced classification accuracy. While our results show high performance in controlled settings, further exploration is needed to ensure reliability in dynamic, real-world environments. Future efforts will focus on integrating our system into applications such as mental health monitoring, adaptive human-computer interaction, and automotive safety. By embedding emotion recognition capabilities into smart wearables and vehicle monitoring systems, this research can contribute to early stress detection, personalized interventions, and improved user experiences. In conclusion, this study highlights the potential of physiological signal-based emotion recognition in understanding human behavior. Expanding the application of these models beyond controlled environments will bridge the gap between theoretical advancements and practical usability, paving the way for innovative solutions in healthcare, assistive technology, and emotionally intelligent systems.

References

[1] Lee, M.S.; Lee, Y.K.; Pae, D.S.; Lim, M.T.; Kim, D.W.; Kang, T.K. Fast emotion recognition based on single pulse PPG signal with convolutional neural network. Appl. Sci. 2019, 9, 3355.

[2] Bulagang, Aaron Frederick, Ng Giap Weng, James Mountstephens, and Jason Teo. "A review of recent approaches for emotion classification using electrocardiography and electrodermography signals." Informatics in Medicine Unlocked 20 (2020): 100363.

[3] Egger, Maria, Matthias Ley, and Sten Hanke. "Emotion recognition from physiological signal analysis: A review." Electronic Notes in Theoretical Computer Science 343 (2019): 35-55.

[4] Hsu, Yu-Liang, Jeen-Shing Wang, Wei-Chun Chiang, and Chien-Han Hung. "Automatic ECG-based emotion recognition in music listening." IEEE Transactions on Affective Computing 11, no. 1 (2017): 85-99.

[5] Goshvarpour, Atefeh, Ataollah Abbasi, and Ateke Goshvarpour. "An accurate emotion recognition system using ECG and GSR signals and matching pursuit method." Biomedical journal 40.6 (2017): 355-368.

[6] Dessai, Amita, and Hassanali Virani. "Emotion Classification Based on CWT of ECG and GSR Signals Using Various CNN Models." Electronics 12.13 (2023): 2795.

[7] Dutta, Stobak, Brojo Kishore Mishra, Anirban Mitra, and Amartya Chakraborty. "An analysis of emotion recognition based on GSR signal." ECS Transactions 107, no. 1 (2022): 12535.

[8] Panahi, F., Rashidi, S., & Sheikhani, A. (2021). Application of fractional Fourier transform in feature extraction from ELECTROCARDIOGRAM and GALVANIC SKIN RESPONSE for emotion recognition. Biomedical Signal Processing and Control, 69, 102863.

[9] Sepúlveda, Axel, Francisco Castillo, Carlos Palma, and Maria Rodriguez-Fernandez. "Emotion recognition from ECG signals using wavelet scattering and machine learning." Applied Sciences 11, no. 11 (2021): 4945.

[10] Raheel, Aasim, Muhammad Majid, Majdi Alnowami, and Syed Muhammad Anwar. "Physiological sensors based emotion recognition while experiencing tactile enhanced multimedia." Sensors 20, no. 14 (2020): 4037.

[11] Dessai, Amita, and Hassanali Virani. "Multimodal and Multidomain Feature Fusion for Emotion Classification Based on Electrocardiogram and Galvanic Skin Response Signals." Sci 6.1 (2024): 10.

[12] Dar, Muhammad Najam, Muhammad Usman Akram, Sajid Gul Khawaja, and Amit N. Pujari. "CNN and LSTM-based emotion charting using physiological signals." Sensors 20, no. 16 (2020): 4551.

[13] Patlar Akbulut, Fatma. "Hybrid deep convolutional model-based emotion recognition using multiple physiological signals." Computer Methods in Biomechanics and Biomedical Engineering 25, no. 15 (2022): 1678-1690.

[14] Chandra, Raga Madhuri, Giri Venkata Sai Tej Neelaiahgari, and Satya Sumanth Vanapalli. "EXTRACTING EMOTION-CAUSE PAIRS: A BILSTM-DRIVEN METHODOLOGY." Informatyka, Automatyka, Pomiary w Gospodarce i Ochronie Środowiska 14, no. 4 (2024): 97-103.

[15] Chandra, Raga Madhuri, Girivenkata Saitej Neelaiahgari, and Satya Sumanth Vanapalli. "Enhancing Driver Safety Through SensorBased Detection and Mitigation of Health Risks in Vehicles." International Conference on Algorithms and Computational Theory for Engineering Applications. Cham: Springer Nature Switzerland, 2024.

[16] C. R. Madhuri, P. Hemanth Suresh, P. S. Kiran and V. Satya Sumanth, "Smart Irrigation Optimization with IoT and Weather Forecasts for Sustainable Crop Management," *2024 1st International Conference on Advances in Computing, Communication and Networking (ICAC2N)*, Greater Noida, India, 2024, pp. 1-7, doi: 10.1109/ICAC2N63387.2024.10895564. keywords: {Irrigation;Schedules;Rain;Soil moisture;Weather forecasting;Sensor systems;Real-time systems;Crop yield;Optimization;Farming;Smart Irrigation;IoT;Weather Forecasting;Machine Learning;Water Management;Farmer Interface;Soil Moisture;Sustainable Farming}

Authors' Profiles

Vanapalli Satya Sumanth is a dedicated and innovative Computer Science student at Velagapudi Ramakrishna Siddhartha Engineering College (VRSEC), Vijayawada, India. He has made significant contributions to the fields of Machine Learning (ML), Computer Vision (CV), Sentiment Analysis, and Precision Farming, with one published journal paper and two conference papers to his name. His research is at the forefront of applying AI technologies to real-world problems, particularly in enhancing agricultural practices through smart technologies. In addition to his academic accomplishments, Sumanth holds a patent for an innovative solution and is currently developing a startup with funding support from ANGRAU, Tirupati, focusing on precision agriculture and AI-based solutions. His projects have garnered recognition in various competitions, winning multiple accolades for their innovative approaches and practical applications. Satya is also the proud recipient of the prestigious Chanakya Fellowship from IIT Tirupati, which further underscores his potential as a researcher and innovator. His commitment to continuous learning is evident through his numerous certifications in advanced technologies, which have equipped him with the skills and knowledge to push the boundaries of innovation in the fields of AI, agriculture, and beyond.

Vanapalli Sashi Vardhan is a first-year B. Tech student specializing in Information technology at Gayatri Vidya Parishad College Of Engineering, Visakhapatnam, India. He is passionate about Machine Learning, the Internet of Things, and Web Programming.

Peruri Bhuvan Satwik is a student at Narayana Junior College, Amaravathi, India. His research interest is in Computer Vision, Pattern Recognition and Machine Learning. His commitment to continuous learning is evident through his numerous certifications in advanced technologies, which have equipped him with the skills and knowledge to push the boundaries of innovation in the fields of AI, agriculture, and beyond.

Neelamsetty Glory is a first-year B. Tech student specializing in CSE at VVIT college of Engineering, Namburu, India. She is passionate about Machine Learning, the Internet of Things, and Web Programming.

Vallala Lohitha Prakaashini is a first-year B. Tech student specializing in Information technology at JNTU, Vizianagaram, India. She is passionate about Machine Learning, the Internet of Things, and Web Programming.

Saragadam Charishma is a fourth-year B. Tech student specializing in Electronics and Communication Engineering at Vignan's institute of engineering for women, Visakhapatnam, India. She is passionate about Machine Learning, the Internet of Things, and Web Programming.

Adiraju Shasank is a fourth-year B. Tech student specializing in CSE at Anil Neerukonda Institute of Techonology And Sciences, Visakhapatnam, India. He is passionate about Machine Learning, the Internet of Things, and Web Programming.

How to cite this paper: Vanapalli Satya Sumanth, Vanapalli Sashi Vardhan, Peruri Bhuvan Satwik, Neelamsetty Glory, Vallala Lohitha Prakaashini, Saragadam Charishma, Adiraju Shasank, "BiLSTM-Powered Emotion Recognition from ECG and GSR Signals", International Journal of Mathematical Sciences and Computing(IJMSC), Vol.11, No.2, pp. 38-51, 2025. DOI: 10.5815/ijmsc.2025.02.04