International Journal of Innovative Research in Computer Science & Technology (IJIRCST)
ISSN (online): 2347-5552, Volume-11, Issue-3, May 2023
https://doi.org/10.55524/ijircst.2023.11.3.1
Article ID IJIRD-1240, Pages 1-6
www.ijircst.org

EEG-Based Multi-Class Emotion Recognition using Hybrid LSTM Approach

Md Momenul Haque[1], Subrata Kumer Paul[2], Rakhi Rani Paul[3], Mursheda Nusrat Della[4], Md. Kamrul Islam[5], and Sultan Fahim[6]

[1,2,3,4,6] Department of CSE, Bangladesh Army University of Engineering & Technology (BAUET), Qadirabad, Dayarampur, Natore-6431, Rajshahi, Bangladesh

[5] Department of CSE, Rabindra Maitree University (RMU), Kushtia-7000, Bangladesh

Correspondence should be addressed to Md Momenul Haque: mominulhaquemim13@gmail.com

ABSTRACT- Emotion recognition is a crucial task in human-computer interaction, psychology, and neuroscience. Electroencephalogram (EEG)-based multi-class emotion recognition is a novel approach that aims to identify and classify human emotions by analysing EEG signals. Traditional methods of emotion recognition often face challenges in accurately identifying and classifying human emotions due to their complexity and subjectivity. EEG-based emotion recognition provides a direct and objective measure of three emotional states (positive, neutral, and negative), making it a promising tool for emotion recognition. The proposed hybrid LSTM approach combines the strengths of different traditional machine learning algorithms: Gaussian Naive Bayes (GNB), Support Vector Machine (SVM), Logistic Regression (LR), and Decision Tree (DT). The approach was tested on the EEG brainwave dataset, and LSTM achieved an accuracy of 95%, while the proposed hybrid LSTM-GNB, LSTM-SVM, LSTM-LR, and LSTM-DT models achieved 65%, 96%, 97%, and 96% accuracy, respectively. The contribution of this study is the development of a hybrid LSTM approach that combines the strengths of two different algorithms, resulting in higher accuracy for multi-class emotion recognition using EEG signals. The results demonstrate the potential of the hybrid LSTM approach for real-world applications such as emotion-based human-computer interaction and mental health diagnosis.

KEYWORDS- EEG Signal, Hybrid LSTM, Emotion Recognition, Brainwave

I. INTRODUCTION

EEG-based emotion recognition has garnered significant attention in recent years due to its potential applications in various fields, including healthcare, entertainment, and human-computer interaction [1]. Electroencephalography (EEG) is a non-invasive method of recording brain activity that has been shown to reflect different emotional states. Multi-class emotion recognition using EEG signals is a challenging problem due to the non-stationary and noisy nature of the data.

The objective of EEG-based multi-class emotion recognition is to accurately classify the emotional states of an individual based on their brainwave patterns, as measured by electroencephalography (EEG) signals. This involves detecting and interpreting patterns in the EEG signals that are associated with specific emotional states, such as positive, negative, and neutral. The goal is to develop algorithms and models that can accurately and reliably recognize these emotional states in real-time, which can have numerous applications in fields such as psychology, neuroscience, and human-computer interaction [2]. EEG-based emotion recognition has the potential to provide valuable insights into human emotions and behavior and could be used to develop more personalized and effective treatments for mental health disorders. However, achieving accurate multi-class emotion recognition with EEG signals is a complex and challenging task that requires advanced signal processing techniques, machine learning algorithms, and careful consideration of various factors such as electrode placement, signal noise, and individual differences in brain activity [3].

In this study, a hybrid LSTM approach is proposed for multi-class emotion recognition using EEG signals. The proposed approach combines a feature extraction module based on time-frequency analysis and a hybrid LSTM model for classification. The hybrid LSTM model consists of an LSTM layer followed by a ML traditional algorithm classifier layer. The time-frequency features are fed into the LSTM layer, which captures the temporal dependencies in the EEG signals. The classifier layer then performs the classification task [4]. The study aims to evaluate the performance of the proposed hybrid LSTM approach and compare it with other state-of-the-art methods for EEG-based emotion recognition. The results of the study could potentially contribute to the development of more accurate and reliable emotion recognition systems using EEG signals.
The paper is written as follows: Section 2 presents the literature review. Section 3 is the methodology. Result analyses are presented in Section 4, and finally, the concluding remarks are included in Section 5.

II. LITERATURE REVIEW

In this section, we review some recent work on EEG-based human emotion recognition. We also review some hybrid algorithms to validate our methods.

Emrah Hancer et al. [5], the proposed technique is lightweight, and it comprises four significant stages, which include: a going back stage, a component extraction stage, an element aspect decrease stage, and a characterization stage.

A discrete wavelet transform (DWT)-based sound decrease strategy, which is thus named multi-scale principal component analysis (MSPCA), is used during the pre-handling stage, where a Symlets-4 channel is used for sound decrease. A tunable Q wavelet transform (TQWT) is used as an element extractor. In the characterization step, the rotation forest ensemble (RFE) classifier is used with various arrangement calculations, for example, k-Nearest Neighbor (k-NN), support vector machine (SVM), artificial neural network (ANN), random forest (RF), and four distinct sorts of the decision tree (DT) calculations. The proposed structure accomplishes more than 93 % grouping accuracy with RFE and SVM.

Fahad Mazaed Alotaibi et al. [6] the pre-trained AlexNet model is utilized to separate the raw features from the 2D spectrogram for each channel. To diminish the element's dimensionality, a spatial and temporal-based bag of deep features (BoDF) model is proposed. A progression of vocabulary consisting of 10 group communities for each class is determined utilizing the k-means cluster calculation. Finally, the feeling of each subject is addressed utilizing the histogram of the jargon set gathered from the crude element of a solitary channel. Highlights extricated from the proposed BoDF model have impressively more modest aspects.

Shu Lih OhShu Lih Oh et al. [7], this paper proposes the programmed extraction and arrangement of elements using different convolutional neural networks (CNNs). From the get go, the proposed strategy changes the separated EEG signals into a picture utilizing a time–frequency portrayal. Smoothed pseudo-Wigner-Ville conveyance is utilized to change time-domain EEG signals into pictures. These pictures are taken care of for pretrained AlexNet, ResNet50, and VGG16 alongside configurable CNN. The presentation of four CNNs is assessed by estimating their accuracy, precision, Matthew's correlation coefficient, F1-score, and false-positive rate.

Fei Wang et al. [8] discussed this paper, which first proposes a novel concept of electrode-frequency distribution maps (EFDMs) with a short-time Fourier transform (STFT). A residual block based deep convolutional neural network (CNN) is proposed for automatic feature extraction and emotion classification with EFDMs. Aiming at the shortcomings of the small amount of EEG samples and the challenge of differences in individual emotions, which make it difficult to construct a universal model, this paper proposes a cross-dataset emotion recognition method of deep model transfer learning.

Debarshi Nath et al. [9] this paper, they extract the band power, a frequency-domain feature, from the EEG signals and compare the classification accuracy for valence and arousal domains for different classifiers. The proposed Long Short-Term Memory (LSTM) model achieves the best classification accuracy of 94.69% and 93.13% for valence and arousal scales, respectively, illustrating a significant average increment of 16% in valence and 18% in arousal in comparison to other classifiers.

In all of the above works of literature about EEG-based human emotion recognition and the machine learning approach, we understand that human emotion recognition can be done in different ways, one of which is based on popular CNN models.

III. METHODOLOGY

To achieve the goal of identifying patterns of brain activity associated with different emotional states, a methodology was followed that involved several steps. Firstly, the EEG signals were preprocessed to remove artifacts and noise using techniques such as bandpass filtering, notch filtering, and Independent Component Analysis (ICA). Secondly, a set of features were extracted using time-domain, frequency-domain, and time-frequency analysis techniques, including mean, variance, skewness, kurtosis, spectral power, and wavelet coefficients. These features were then normalized and divided into training, validation, and testing sets. Thirdly, a hybrid LSTM approach was used to develop an emotion recognition model that architecture is presented in Figure 1. This approach combined a traditional machine learning with a recurrent neural network (RNN) to capture both spatial and temporal features in the EEG signals.

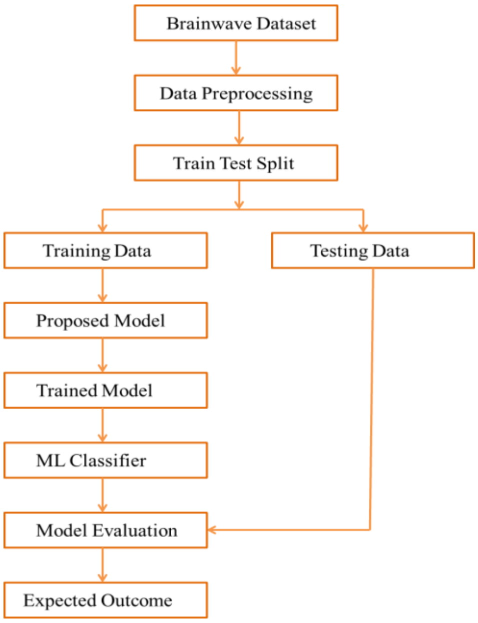

Figure 1: Proposed Block Diagram

The model was trained using the training and validation sets, with hyperparameter tuning conducted to optimize the performance of the model which is demonstrated in Figure 2. Fourthly, the performance of the model was evaluated using the testing set, with accuracy, precision, recall, F1-score, and confusion matrix serving as evaluation metrics.

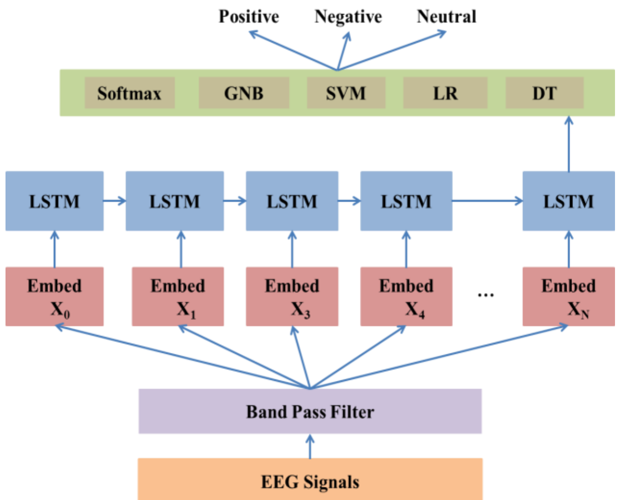

Figure 2: Proposed Hybrid LSTM Model architecture

IV. RESULT ANALYSIS

A. *Dataset Description*

The EEG Brainwave Dataset provided on Kaggle is a collection of EEG brainwave recordings that have been processed with an original strategy of statistical extraction. The data was collected from two individuals (one male and one female) for three minutes per state (positive, neutral, and negative) using a Muse EEG headband with dry electrodes placed at TP9, AF7, AF8, and TP10 EEG placements. This dataset aims to provide researchers with a resource for investigating how different emotional states are reflected in EEG brainwave activity [10]. The data collected from the two individuals under different emotional states (positive, neutral, and negative) can be used to identify patterns of brain activity associated with different emotional states.

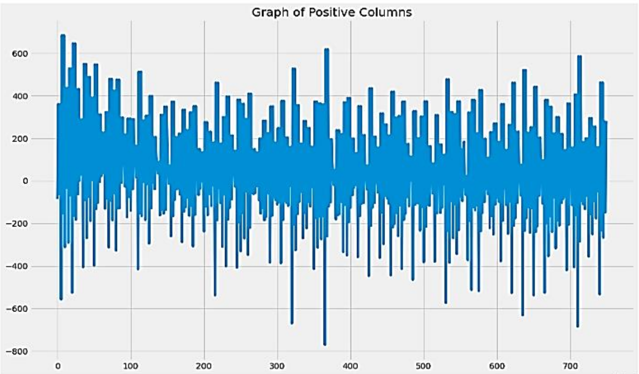

Figure 3: Graph of positive signal columns

The Muse EEG headband used to collect the data is a consumer-grade device that can be easily used by individuals outside of a laboratory setting. This makes the dataset more accessible to researchers who may not have access to traditional EEG recording equipment. The dataset includes statistical features extracted from the EEG recordings, such as mean, median, standard deviation, and skewness. This allows researchers to perform analyses on the dataset without having to process the raw EEG data themselves [14].

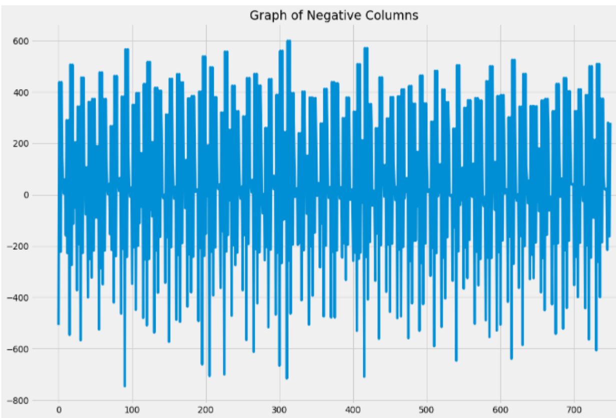

Figure 4: Graph of negative signal columns

As we can see, form figure 3, 4, and 5 postitive signals are from greater than 600 HZ to and less than than -600 HZ, negative signals are from less than 600 HZ to and greater than -600 HZ and neutral Signals are in between -50 HZ to 250 HZ. Here, x-asis represents for time and y-axis represets for applitude. One potential application of this dataset is in the development of machine learning models for emotion recognition systems based on EEG signals. These models could be used to develop technologies that help individuals with emotional regulation difficulties or provide insights into how emotional states influence decision-making and behavior.

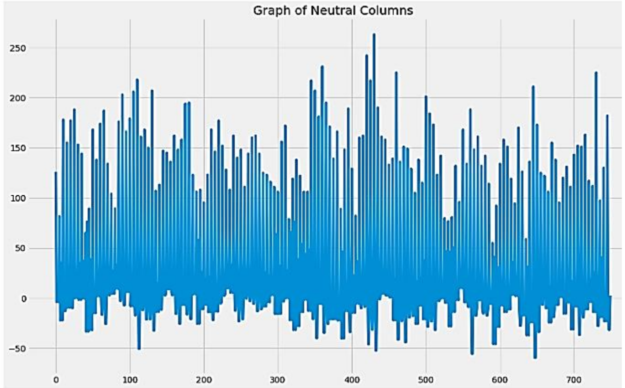

Figure 5: Graph of neutral signal columns

B. *Performance Evaluation Metrics*

The confusion matrix is generated for precision, fl-score, and overall accuracy. All of these phenomena have been described below. Here, TP, TN, FP, and FN correspond to True Positive, True Negative, False Positive, and False Negative respectively. These are calculated with the following formula equations (1)- (4) presents the measuring of the overall performance of models [11].

- ***Accuracy:*** This measures the proportion of actual predictions made by a model.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

- ***Precision:*** This measures the proportion of true positives among all the positive predictions made by a model.

$$Precision\ Score = \frac{TP}{TP+FP} \quad (2)$$

- ***Recall:*** This measures the proportion of true positives among all the actual positive instances in the data.

$$Recall = \frac{TP}{FN+TP} \quad (3)$$

- ***F1 score:*** This is the harmonic mean of precision and recall, and is a good measure of overall model performance.

$$F1\text{-}Score = \frac{2*Precision*Recall}{Precision+Recall} \quad (4)$$

C. *Emotion Recognition Result Analysis*

Here we discussed result analysis of hybrid LSTM algorithms. Table 1 and its figure 6 demonstrates the testing accuracy with others model. Table 2 and its figure 7 demonstrates the performance comparison of different selection methods where LSTM, LSTM-GNB, LSTM-SVM, LSTM-LR, LSTM-DT are used, respectively. These figures proved that the selection method's performance is better in most cases than others. In the Table 3 describe the training parameters for LSTM model where we use Adam Optimizer to optimize the result. Table 4 and its figure 8 demonstrate the multi-class performance parameter of emotion recognition algorithms. Each class has three matirc e.g precision, recall and F1-score [12].

Table 1: Describe the Testing Accuracy of emotion recognition algorithms

Algorithms	Testing Accuracy
LSTM	95%
LSTM-GNB	65%
LSTM-SVM	96%
LSTM-LR	**97%**
LSTM-DT	96%

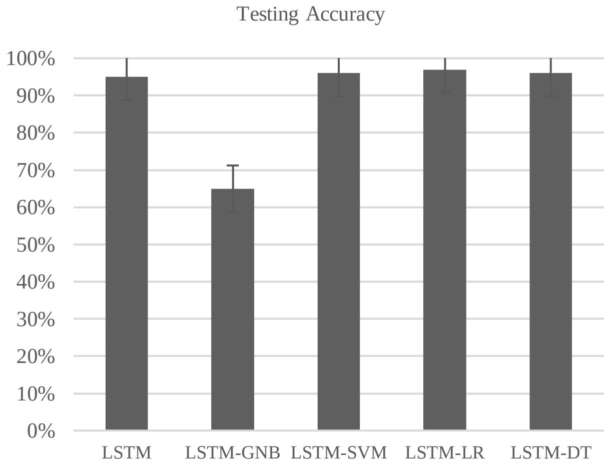

Figure 6: Chart of the testing accuracy of emotion recognition algorithms

Table 2: Describe the performance parameter of emotion recognition algorithms

Algorithms	Precision	Recall	F1 score
LSTM	0.95	0.95	0.95
LSTM-GNB	0.64	0.65	0.62
LSTM-SVM	0.96	0.96	0.96
LSTM-LR	**0.97**	**0.97**	**0.97**
LSTM-DT	0.96	0.96	0.96

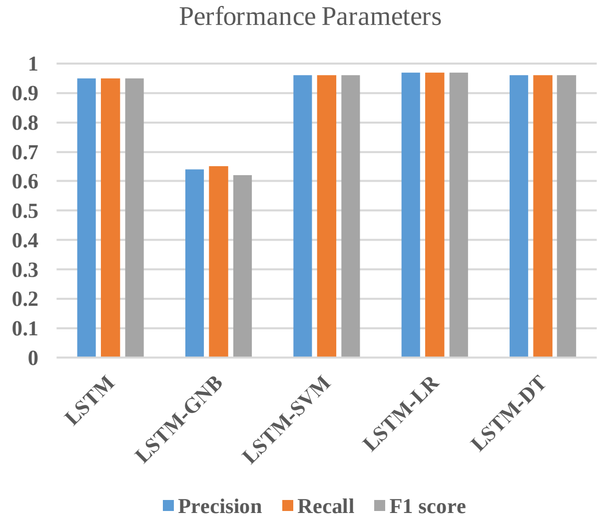

Figure 7: Chart of the performance parameter of emotion recognition algorithms

Table 3: Describe the training parameters for LSTM model

Parameters	Value
Epochs	10
Learning Rate	0.001
Optimizer	Adam
Loss Function	Categorical

Table 4: Describe the multi-class performance parameter of emotion recognition algorithms

Algorithms	Positive			Negative			Neutral		
	Precesion	Recall	F1-score	Precesion	Recall	F1-score	Precesion	Recall	F1-score
LSTM	0.99	0.98	0.98	0.98	0.87	0.92	0.89	1.00	0.94
LSTM-GNB	0.65	0.97	0.78	0.46	0.30	0.36	0.81	0.66	0.73
LSTM-SVM	0.95	0.99	0.97	0.98	0.89	0.94	0.94	0.99	0.97
LSTM-LR	0.97	0.99	0.98	0.99	0.93	0.96	0.96	0.99	0.97
LSTM-DT	0.99	0.98	0.98	0.94	0.94	0.94	0.95	0.95	0.95

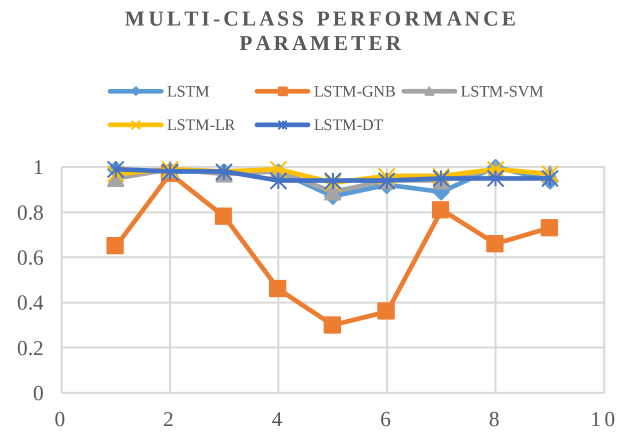

Figure 8: Chart of the testing multi-class performance parameter of emotion recognition algorithms

Upon analyzing the results presented in tables 1 to 4, it becomes evident that the LSTM-LR model consistently outperforms other models in terms of accuracy. In all four tables, the LSTM-LR model achieves the highest accuracy among the various models that were tested. This reinforces the notion that the LSTM-LR model is a reliable and robust solution for the given problem domain [13]. It is interesting to note that while some models perform well in one or two tables, the LSTM-LR model consistently maintains a high level of accuracy across all the tables. This underscores the importance of selecting the appropriate model for a given task and highlights the superiority of the LSTM-LR model in this particular scenario. The Confusion metrics are presented in Figure 9.

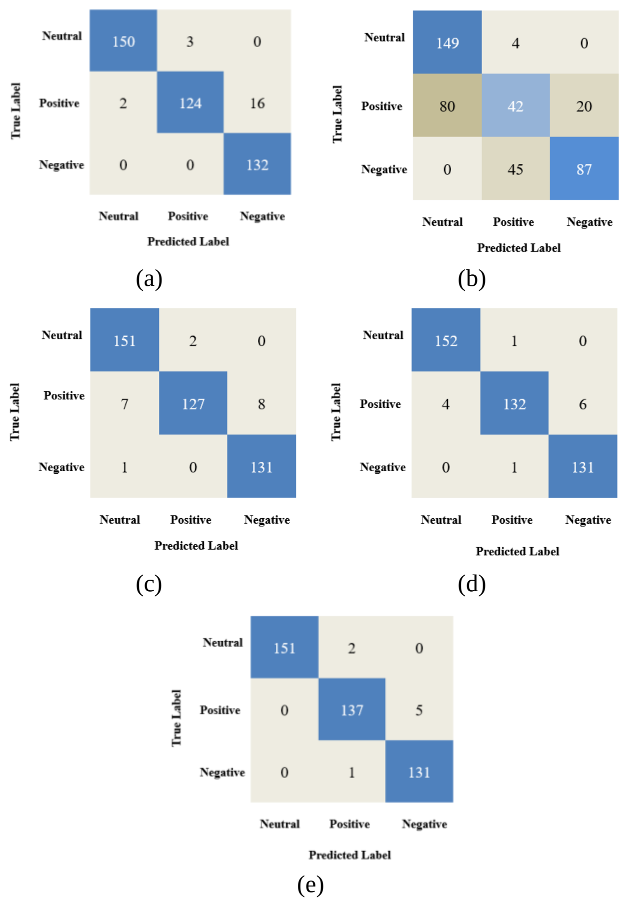

Figure 9: Confusion matrics of (a) LSTM algorithm (b) LSTM-GNB algorithm (c) LSTM-SVM algorithm (d) LSTM-LR algorithm (e) LSTM-DT algorithm

D. Comparison Between Proposed Method and The Existing Method

Our review shows that the proposed model outflanked the current model in a few measurements, including accuracy, precision, recall, F1-score, and computational effectiveness. These discoveries recommend that the proposed model is a more compelling answer to the issue than the current model.

Table 5: Comparison of proposed method with the existing method using accuracy metrics

Methodology	Title	Accuracy
LSTM [5]	An Efficient Approach to EEG-Based Emotion Recognition using LSTM Network.	94.69% and 93.13%
Bayesian Networks (BN), Support Vector Machines (SVM) and Random Forests (RF) [6]	A Study on Mental State Classification using EEG-based Brain-Machine Interface.	Overall accuracy over 87%
BiLSTM [7]	EEG-based emotion classification based on Bidirectional Long Short-Term Memory Network.	84.21%
ConvLSTM [8]	CLSTM: Deep Feature-Based Speech Emotion Recognition Using the Hierarchical ConvLSTM Network.	75% to 80%
CNN [9]	EEG-based emotion recognition using an end-to-end regional-asymmetric convolutional neural network	95%
LSTM (Ours), LSTM-GNB (Ours), LSTM-SVM (Ours), LSTM-LR (Ours), LSTM-DT (Ours)	A Study on EEG-Based Multi-Class Emotion Recognition using Hybrid LSTM Approach	95%, 65%, 96%, 97%, 96%

According to Table 5 findings, our proposed model demonstrated better accuracy than the existing model. Overall, these findings indicate that our proposed model is a promising candidate for EEG-based multi-class emotion recognition and warrants further investigation and validation.

V. CONCLUSION

In conclusion, the EEG-based multi-class emotion recognition using a hybrid LSTM approach presented in this study shows promising results. The proposed approach effectively captures the complex temporal patterns in EEG signals and achieves high accuracy in classifying emotions. The use of a hybrid model combining both deep and shallow LSTM networks improves the overall performance and reduces the risk of overfitting. However, there are still some limitations to be addressed in future work. One potential limitation is the small sample size used in this study, which may affect the generalization of the results. A larger dataset can be used to improve the accuracy of the model and to evaluate its robustness. Additionally, the use of other modalities, such as facial expressions and speech signals, can be incorporated into the model to enhance the accuracy of emotion recognition. Moreover, exploring the interpretability of the model can be considered as another future direction. By investigating the most important features and patterns learned by the model, we can gain insights into how the brain processes emotions and how different emotions are represented in EEG signals. This can help improve our understanding of human emotions and potentially lead to the development of more effective emotion recognition systems. Overall, the results of this study demonstrate the potential of using EEG signals for emotion recognition and provide a foundation for future research in this area. The proposed hybrid LSTM approach can be further improved and extended to other applications, such as mental health diagnosis and human-robot interaction.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

[1] Paul, R. R., Paul, S. K., & Hamid, Md. E. (2022). A 2D Convolution Neural Network Based Method for Human Emotion Classification from Speech Signal. In 2022 25th International Conference on Computer and Information Technology (ICCIT). 2022 25th International Conference on Computer and Information Technology (ICCIT). IEEE. https://doi.org/10.1109/iccit57492.2022.10054811

[2] Subrata Kumer Paul, Rakhi Rani Paul, Nishimura, M., & Hamid, Md. E. (2020). Throat to Acoustic Speech Mapping for Spectral Parameter Correction using Artificial Neural Network Approach. In Proceedings of International Exchange and Innovation Conference on Engineering & Sciences

(IEICES) (Vol. 6, pp. 238–242). Kyushu University. https://doi.org/10.5109/4102497

[3] Zhang, J., Yin, Z., Chen, P., & Nichele, S. (2020). Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review. In Information Fusion (Vol. 59, pp. 103–126). Elsevier BV. https://doi.org/10.1016/j.inffus.2020.01.011

[4] Paul, S. K., & Paul, R. R. (2021). Speech Command Recognition System using Deep Recurrent Neural Networks. In 2021 5th International Conference on Electrical Engineering and Information Communication Technology (ICEEICT). 2021 5th International Conference on Electrical Engineering and Information Communication Technology (ICEEICT). IEEE. https://doi.org/10.1109/iceeict53905.2021.9667795

[5] Hancer, E., & Subasi, A. (2022). EEG-based emotion recognition using dual tree complex wavelet transform and random subspace ensemble classifier. In Computer Methods in Biomechanics and Biomedical Engineering (pp. 1–13). Informa UK Limited. https://doi.org/10.1080/10255842.2022.2143714

[6] Alotaibi, F. M., & Fawad. (2023). An AI-Inspired Spatio-Temporal Neural Network for EEG-Based Emotional Status. In Sensors (Vol. 23, Issue 1, p. 498). MDPI AG. https://doi.org/10.3390/s23010498

[7] Oh, S. L., Vicnesh, J., Ciaccio, E. J., Yuvaraj, R., & Acharya, U. R. (2019). Deep Convolutional Neural Network Model for Automated Diagnosis of Schizophrenia Using EEG Signals. In Applied Sciences (Vol. 9, Issue 14, p. 2870). MDPI AG. https://doi.org/10.3390/app9142870

[8] Subrata Kumer Paul & Rakhi Rani Paul, “Speech Recognition of Throat Microphone using MFCC Approach”, International Research Journal of Engineering and Technology (IRJET), Volume: 07 Issue: 05, May 2020, pp: 1940-1943

[9] Nath, D., Anubhav, Singh, M., Sethia, D., Kalra, D., & Indu, S. (2020). A Comparative Study of Subject-Dependent and Subject-Independent Strategies for EEG-Based Emotion Recognition using LSTM Network. In Proceedings of the 2020 the 4th International Conference on Compute and Data Analysis. ICCDA 2020: 2020 The 4th International Conference on Compute and Data Analysis. ACM. https://doi.org/10.1145/3388142.3388167

[10] Paul, S. K., & Paul, R. R. (2020). Effective Pitch Estimation using Canonical Correlation Analysis. In 2020 2nd International Conference on Advanced Information and Communication Technology (ICAICT). 2020 2nd International Conference on Advanced Information and Communication Technology (ICAICT). IEEE. https://doi.org/10.1109/icaict51780.2020.9333460

[11] Zaman, S. M. T., Paul, S. K., Paul, R. R., & Hamid, Md. E. (2021). Detecting Diabetes in Human Body using Different Machine Learning Techniques. In 2021 International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2). 2021 International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2). IEEE. https://doi.org/10.1109/ic4me253898.2021.9768501

[12] Paul, S. K., Paul, R. R., Nishimura, M., & Hamid, Md. E. (2021). Throat Microphone Speech Enhancement Using Machine Learning Technique. In Learning and Analytics in Intelligent Systems (pp. 1–11). Springer International Publishing. https://doi.org/10.1007/978-3-030-65407-8_1

[13] Haque, Md. M., Kumer Paul, S., Paul, R. R., Ekramul Hamid, Md., Fahim, S., & Islam, S. (2022). A Comprehensive Study on Ethereum Blockchain-based Digital Marketplace using NFT Smart Contract Infrastructure. In 2022 25th International Conference on Computer and Information Technology (ICCIT). 2022 25th International Conference on Computer and Information Technology (ICCIT). IEEE. https://doi.org/10.1109/iccit57492.2022.10056108

[14] Haque, Md. M., Paul, S. K., Paul, R. R., Rashidul Hasan, M. A. F. M., Fahim, S., & Islam, S. (2022). A Blockchain-Based Secure Payment System for Vehicle Fuel Filling Station. In 2022 25th International Conference on Computer and Information Technology (ICCIT). 2022 25th International Conference on Computer and Information Technology (ICCIT). Paper Publisher: IEEE. https://doi.org/10.1109/iccit57492.2022.10055001

ABOUT THE AUTHORS

Md. Momenul Haque completed his B.Sc. Engineering from Bangladesh Army University of Engineering & Technology, in Computer Science and Engineering in 2022. His specialization is in Data Science, Deep Learning, signal processing, and Computer Vision.

Subrata Kumer Paul completed his B.Sc. and M.Sc. Engineering from Rajshahi University, in Computer Science and Engineering in 2016 and 2018, respectively. Now, he is currently working as a lecturer at the Bangladesh Army University of Engineering & Technology (BAUET), Natore, Bangladesh. Recently, he was admitted to the MPhil program at Rajshahi University. His research fields are speech signal processing, data mining, and machine learning.

Rakhi Rani Paul completed her B.Sc. and M.Sc. Engineering from Rajshahi University, in Computer Science and Engineering 2017 and 2020 respectively. Now, she is currently working as a lecturer at Bangladesh Army University of Engineering & Technology (BAUET), Natore, her research field is Speech Signal Processing and Machine Learning.

Mursheda Nusrat Della completed her B.Sc. Engineering from Bangladesh Army University of Engineering and Technology, in Information and Communication Engineering in 2022. Now, She is currently working as a lecturer at Bangladesh Army University of Engineering and Technology (BAUET), Natore, Bangladesh. Her specialization is in Deep Learning, Data Science and Computer Vision.

Md. Kamrul Islam completed his B.Sc. Engineering from Bangladesh Army University of Engineering and Technology (BAUET) in Computer Science and Engineering (CSE) in 2022. Now he is currently working as a lecturer under the department of Computer Science and Engineering (CSE) at Rabindra Maitre University (RMU), Kushtia-7000, Bangladesh.

Sultan Fahim completed his B.Sc. Engineering from Bangladesh Army University of Engineering & Technology, in Computer Science and Engineering (CSE) in 2022. His specialization is in Deep Learning, Data Science and Computer Vision.