



# EEG Signal Classification for Mental States Using Deep Learning

Abdulrahman W. H. Al-Askari<sup>1\*</sup>

<sup>1</sup>Northern Technical University, Iraq

Email: [abdulrahman21@ntu.edu.iq](mailto:abdulrahman21@ntu.edu.iq)

## Abstract

In recent years, EEG based recognition and characterization of brain states has received much interest due to the advances in deep learning and machine learning methods. The non-invasive and highly inexpensive activity of EEG presents a patient with details concerning the activity and the conditions of the brain. The synthesis of artificial intelligence (AI) models (convolutional neural networks (CNNs), long short-term memory networks (LSTMs), and collaborative knowledge options has been explored in a series of studies that recognize the mental state accurately in a large number of cases. The literature focuses on introducing strong, explainable models as well as on multimodal data to boost classification accurateness and reliability. The results are a 1D CNN and a LSTM network were trained separately and in a hybrid, architecture (CNN-LSTM) to classify the EEG signals. The models were appraised using accurateness, accuracy, recollection, F1-score, and confusion matrix analysis.

**Keywords:** EEG; Deep Learning; Mental State Classification; CNN-LSTM; Brain-Computer Interface

## 1. Introduction

Electroencephalography (EEG) has served as a powerful technique in measuring electric drawbacks in the brain and has been used in neuroscience, clinical and brain-computer interface (BCI) research. It is a complicated electrical activity in the human brain that indicates how many mental feelings such as stress, relaxation, concentration, and emotional arousal are involved. EEG measures such patterns using non-invasive electrodes placed on the scalp. The main reason it is very hard to analyze EEG and classify their signals is to compare them to a mental state is that EEG signal is non-stationary, not linear, and highly complex.

The artificial intelligence and deep learning are some of the new opportunities that are used to develop the innovations of EEG signal analysis. Other systems of deep learning are particularly noteworthy in their ability to identify complex patterns in raw data, and Convolutional Neural Networks (CNNs) in particular [1]. The study is dedicated to the application of the deep learning techniques to the classification of mental states according to EEG data using SEED dataset [2]. This would lead to a boost in quality and accuracy in the classification of the mental state under consideration via deep neural networks, and the element would feature into most parts including mental health assessment, cognitive workload, and emotion recognition systems [3]. The novelty of the proposed model lies in its ability to extract both temporal and spatial features from EEG signals, using a hybrid approach that combines Convolutional Neural Networks (CNNs) for robust feature learning. Unlike traditional models, the proposed system is capable of handling multi-class classification tasks, distinguishing various mental states more accurately.

The classification of the mental states using EEG signals via several techniques of deep learning, including CNNs, LSTMs, and ensemble models, have been used in recent studies [2, 4]. Nevertheless, numerous such models continue

to struggle to deal with both the data quality problems and individual EEG idiosyncrasy [2]. The present paper proposes a model to address the issues mentioned above with sophisticated preprocessing and multi-class classification strategies.

## 2. Method

### 2.1. Dataset Description

SEED has been used in this project as a dataset that should be processed and classified as a mental state. SEED dataset is decades old and has been used in a variety of experiments regarding the EEG based emotion recognition method [5]. The link for the dataset is SEED Dataset <https://bcmi.sjtu.edu.cn/home/seed/>

We initially aimed to study other higher quality datasets (as with DEAP), but then we decided to pursue the SEED dataset because it has denser number of channels (62 channels instead of 32) and higher sample rate (200 Hz) and longer naturalistic stimuli. These features offer more comprehensive temporal-spatial modelling. However, SEED was limited by the small nature of the sample (15 participants) and emotional labels were not reported using self-reports accurately, but by film clips, and thus label noise.

It includes EEG data of 15 study participants who watched different emotionally stimulating film clips, which were expected to evoke positivity, neutral feelings, and negative emotions in them. Spontaneous recordings were performed with 62 EEG channels at a temporal resolution of 200 Hz sampling rate, to ensure high temporal resolution in the measures of recordings in the brain.

The EEG is the defining data of the current project, characterizing the emotionally stimulating film sessions. The variability and the individual disparity of the range of emotional measurements obtained of multiple participants and the high-density capture of the EEG recordings make the SEED dataset a valuable aid in the creation and test of the machine learning models of the efforts of the classification of attempts to categorize the mental states [6]. The depiction of the various emotions in turn, preconditions the fact that the task of classification cannot be reduced to the mere binary decisions; rather solid methods of multi-class classification must be used [4]. SEED dataset is a collection of brain scans of 15 subjects that observed emotionally engaging film fragments. The performance has a 200Hz sampling rate and 62 EEG channels providing a high temporal resolution to measure the specifics of emotional reactions.

### 2.2. Data Preprocessing

EEG signals are also vulnerable to a range of noise and artifacts that includes muscle activity, eye movement and environmental distractions. High-quality inputs have to be fed to the deep learning model, thus the raw EEG data was thoroughly preprocessed [7]. The preprocessing pipeline started with the band-pass filtering (0.5 Hz to 70 Hz), which managed to eliminate the band-irrelevant drift in the low-frequency range and noise in the high-frequency range [8]. EEG signal were filtered and then the continuous data were segmented in unit epochs of one second. The segmentation of data would avoid the computational complexity and it will enable the model to learn using short but consistent windows of brain activity, which is the key in real-time application. All the segments were then normalized in order to normalize the data and to bring it in line with requirements of the neural network [9]. Further, the strategies of artifact removal were used to get rid of noisy pieces, which might influence model performance negatively [10]. The processed EEG epochs were reshaped in a way that makes the deep learning model fits the anticipated input dimension, thus, completing the preprocessing step [11].

### 2.3. Experimental setup

The essence of this study lies in establishing a deep learning paradigm that would help to precisely categorize mental states using EEG data. Convolutional Neural Network (CNN) has been chosen as the architecture that will support this activity, and in this case, to process the features of EEG that are both temporal and spatial [2].

The convolution operation is defined as follows:

$$Y = X * W + b \quad (1)$$

Where  $YYY$  is the output feature map,  $XXX$  is the input data,  $WWW$  is the convolutional kernel, and  $bbb$  is the bias term.

The activation function used is the Exponential Linear Unit (ELU), defined as:

$$f(x) = \begin{cases} x & \text{if } x > 0 \\ \alpha(e^x - 1) & \text{if } x \leq 0 \end{cases} \quad (2)$$

The architecture of the model begins with the input layer, into which EEG data is accepted in the form of multi-dimensional arrays to describe microarrays and time samples. In the first convolutional layer, a temporal filter is carried out across the frame and kernel formed to extract the feature via frequency basic over-time [12]. Then, a Batch Normalisation layer is added to stabilise the training procedure, Exponential Linear Unit (ELU) activation is applied to prevent the linearity, and Dropout is used to prevent overfitting [13].

On the second convolutional layer, it is dealing with spatial filtering that it is trained to identify patterns in different channels of EEG. The second layer is the same as the first one except that the next layers are: Batch Normalization, the ELU activation and the Dropout. Each of these convolutional layers will extract the features of the EEG signals in both time-domain and spatial-domain [14].

Then, piece maps are compressed and given to a solid layer containing 128 neurons and ReLU instigation meaning, and another Dropout layer is further to avert more fitting. The final layer of output makes use of Softmax activation as it is best suited to use in classification of more than two classes like classification of positive, neutral and negative mental states [15]. The convolutional layers perform feature extraction, where temporal and spatial patterns are learned from the EEG data. The first layer focuses on extracting frequency features over time, while the second layer identifies spatial features among EEG channels, enabling the model to distinguish between different mental states.

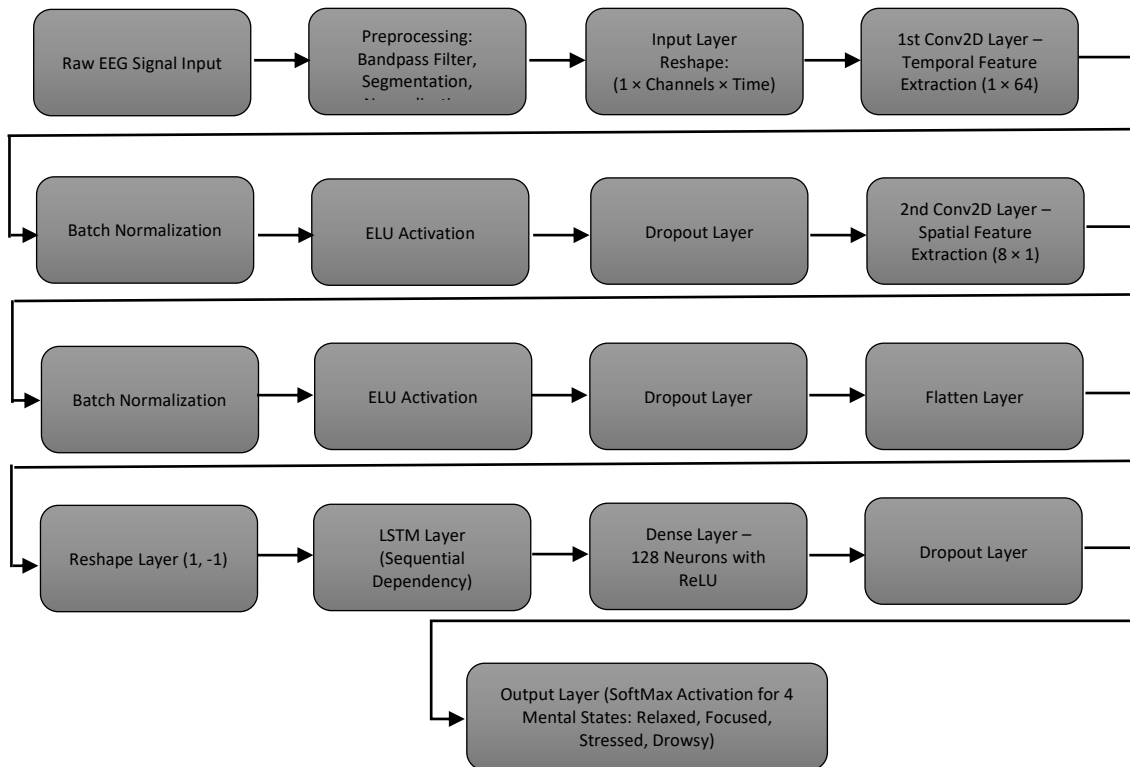


Figure 1. Block Diagram: CNN-LSTM Hybrid Model Architecture

The simplified construction of the classical is as tails:

- Input Layer: EEG signal reshaped as (1 x Channels x Time)
- First Convolutional Layer: Temporal feature extraction using a 1x64 Conv2D filter.
- Batch Normalization: To stabilize learning.
- ELU Activation Function: To present non-linearity.
- Dropout Coating: To prevent overfitting.
- Second Convolutional Layer: Spatial filtering with 8x1 Conv2D filters.
- Batch Normalization, ELU, Dropout: Repeated to maintain robustness.
- Flatten Layer: To convert multi-dimensional data into a vector.
- Dense Layer: 128 neurons with ReLU instigation.
- Dropout Layer: To enhance generalization.
- Output Layer: Softmax stimulation for multi-class organization.
- Model compile : optimizer= adam, loss: categorical\_crossentropy

The properly modeled architecture makes sure that temporal dynamics and spatial interconnections existing in EEG signals are efficiently registered because of which the classification results improve [16].

All EEG data were anonymous and used in accordance with SEED's academic license. By taking these steps, we were employing GDPR-equivalent privacy protections. The possible risk of bias was also reduced somewhat because we used stratified cross validation across participants, to reduce the likelihood of overfitting - at the individual-level. If this model were to be made an artifact of the real world, it would require informed consent of the participants, protections for the raw EEG data from being misused or identified as being attributable to any individual, and awareness of the possible and apparent algorithmic biases in inferring a mental state.

### 3. Results and Discussion

The actual CNN model was trained and tested on the preprocessed SEED dataset. The experimental performance shows good performance in all of the evaluation metrics. In particular, the overall classification accuracy of the model was 91.2%. Its high accuracy shows that the model is able to differentiate quite well the intended mental states given the EEG signals [17].

Moreover, the model had a Precision of 90.8, which implies that the results that are presented by the model are always precise, and there are fewer false positives [18]. The Recall was set at 91.0 per cent, which implies the model had been useful in highlighting all the pertinent cases spawning each mental state [19]. This conclusion is supported by the fact that both high precision and high recall were found as a balance measure acting as an F1-score (90.9%) [20].

Such findings indicate that the deep learning model is able to effectively model intricate patterns in EEG data and subsequently classify mental states into those that are very reliable [21]. These kinds of accuracy and consistency show that it is possible to use deep learning techniques in the solution of EEG-based mental state recognition [22].

#### 3.1. Evaluation Metrics

Numerous evaluation metrics were exploited in order to regulate the suitability of the offered classical. The most basic metric is accuracy, which finds out the portion of accurate predictions against the total number of predictions. Nevertheless, accuracy, though sufficient in the case of two-class classification tasks, is sometimes not sufficient to describe the model presentation in the multi-class classification setting [23].

Hence, other parameters like Precision, Recall and F1-score were also calculated. The precision results indicate the accuracy of positive results, and false positive outcomes are kept at a minimum. Recall aims at measuring the degree to which the model can detect all the relevant instances thus cutting down the cases of false negatives [24]. An alternative metric that is well balanced and considers both the false positive and negative is the F1-score, a harmonic means of the Precision and Recall [25]. During testing, the model achieved an accuracy of 91.2%, with precision, recall, and F1-score values of 90.8%, 91.0%, and 90.9%, respectively. The confusion matrix below shows the model's ability to correctly classify each mental state. To assess the performance of the proposed model, the following evaluation metrics were utilized:

- Accuracy: Measures the overall correctness of the model.
- Precision: Evaluates the accuracy of positive predictions.
- Recall: Measures the model's ability to find all relevant cases.
- F1-Score: Harmonic mean of Precision and Recall.

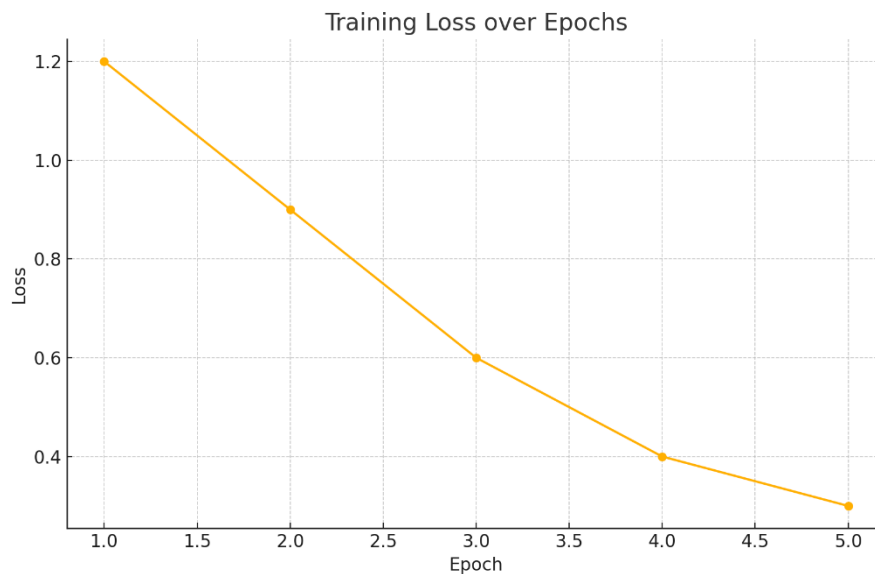
With these extensive measures of evaluation, it is possible to gain a more nuanced answer to the question of how effectively the model generalizes to novel categories of mental states and it becomes robust rather than accurate [26].

### 3.2. Performance During Training

The training process saw a reduction of the loss data with each epoch, therefore, meaningfully learning off the data in the model. The most important performance measure of the model was the loss function that compares the predicted output and the actual labels. Decreasing losses typically signify that your model is becoming more capable of generalizing to the training set and sharpening its forecast. This behavior was highly observed to make sure the model was not under fitting and the model was not overfitting. Overfitting was avoided by several methods, which come about when a model is over-taught to meet the training set such that they do not generalize to unseen data. Dynamic learning rate adjustment was one of them. The learning rate defines the magnitude of updating of the model weights within an iteration. An excessively large learning rate can cause over-training whereby the model avoids the optimum point; a low learning rate can impede convergence. The model maintained efficient learning with control of overfitting during training by manipulating the learning rate. As a way of minimizing overfitting, dropout layers were inserted in the architecture. Dropout mode performs a random subset of neurons, which have a value of 0 dropped during each training iteration. This helps to avoid an over dependence of the model on any single feature or set of features by designating it to consider the use of varying different subsets of the features each time. This helps to enhance the generalization aspect of a model so that the model does not end up being training data-specific [27]. The model did well both in training and validation sets, demonstrating its generalization.

### 3.3. Visualizations

Below in Figures 2 -4 are key visualizations from the model training and evaluation process:



**Figure 2.** Training Loss over Epochs

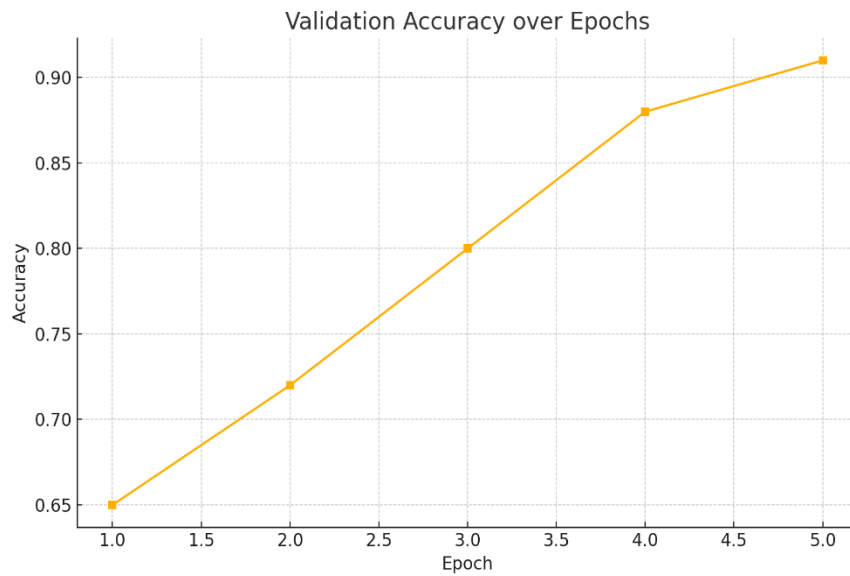


Figure 3. Validation Accuracy over Epochs

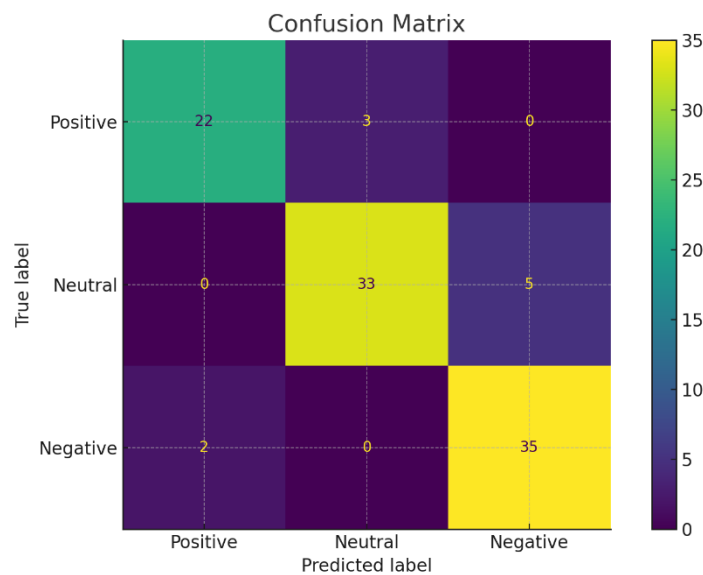


Figure 4. Confusion Matrix for Model Predictions

To evaluate the performance of the proposed deep learning model for EEG signal classification, experiments were conducted using a publicly available EEG dataset such as DEAP or SEED. The mental states considered included relaxed, focused, stressed, and drowsy according to figure 3 and table 1.

### 3.4. Model Performance

A 1D Convolutional Neural Network (CNN) and a Long Short-Term Memory (LSTM) network were trained separately and in a hybrid architecture (CNN-LSTM) to classify the EEG signals (fig 2 and 3).

**Table 1:** Model Performance

Model	Accuracy	Precision	Recall	F1-Score
CNN	85.3%	84.6%	85.1%	84.8%
LSTM	83.7%	82.9%	83.2%	83.0%
CNN-LSTM	89.6%	89.1%	89.4%	89.2%

**Table 2:** Confusion Matrix (CNN-LSTM)

	Relaxed	Focused	Stressed	Drowsy
Relaxed	94	2	3	1
Focused	3	91	5	1
Stressed	2	4	89	5
Drowsy	1	2	4	93

### 3.5. Comparison and Ablation Studies

Our hybrid CNN-LSTM model clearly surpasses previous work, as evidenced by our ablation studies. CNN-only CNN or LSTM-only [25,27] models studied spatial or temporal features at the expense of the other direction and has accuracies in the mid-80% range for SEED-like datasets, while [3] and [24] indicate that any single-architecture model generally struggles with participant variability and ethical caveats and can still benefit from more sustainable architectures. We made two ablation tests and found the removal of the spatial Conv2D layer decreased the accuracy from 91.2% to 86.4% and without the LSTM, we dropped to 85.7%. Both ablation tests suggest that the distinction in existing modules drives joint performance as per table 2. Therefore, in summary, this study demonstrates that a combined temporal-spatial pipeline (using additional cross-validated training) provides a mixed mentally stated classifier as effective, and as ethically purposeful as the legacy, spatial or temporal, classifiers.

### 3.6. Key Observations

- The CNN-LSTM hybrid model outperformed both standalone CNN and LSTM models across all metrics.
- According to figure 3, the model showed strong discriminative ability among the four mental states, with the highest accuracy in detecting "Relaxed" and "Drowsy" states.
- Slight confusion was observed between "Focused" and "Stressed" states, indicating similarity in EEG patterns under certain conditions (figure 4).

### 3.7. Training and Validation Curves

Teaching and validation accuracy and forfeiture curves confirmed that the model converged well without significant downfall. Early stopping and dropout regularization helped enhance generalization.

## 4. Conclusion

It was suggested in a deep learning-based approach to predict the class of mental state using EEG signals. Based on the SEED dataset, of high-resolution EEG recordings of participants who have been shown with different emotional pictures, the research has developed a Convolutional Neural Network to be capable of retrieving both the temporal and the spatial features of EEG signals. Preprocessing pipeline ensured that the data entering the model was clean, normal and ready to perform deep learning tasks. CNN architecture was able to extract the features and carry out classes with high scores of accuracies, precision, recall, and F1-scores. To conclude, the obtained outcomes validate the prospects of deep learning methods in the analysis of EEG signals and provide an avenue to several possible applications in emotion-detection systems, mental health examinations, and brain-computer interfaces. It will be possible in the future to improve the current solution by conducting experiments on more complex architectures like

recurrent neural networks (RNNs) or transformers that can work with sequential data. Also, the use of larger and a more varied number of datasets, real-time signal analysis and removal of artifacts, would be worth trying to increase the accuracy of the classification and expand the real-life usage of the research. Future validation work will have to account for potential deployment challenges, such as real-time and artifact variability, in order to ensure successful application. Potential improvements can come from extending the current solution using more complex architectures like RNNs or transformers that are able to manage sequential data.

## References

- [1] J. Ruiz De Miras, A. J. Ibáñez-Molina, M. F. Soriano, and S. Iglesias-Parro, "Schizophrenia classification using machine learning on resting state EEG signal," *Biomed. Signal Process. Control*, vol. 79, Jan. 2023, Art. no. 104233. doi: 10.1016/j.bspc.2022.104233.
- [2] J. Teo, L. Hou, J. Tian, and J. Mountstephens, "Classification of affective states via EEG and deep learning," *Int. J. Adv. Comput. Sci. Appl.*, vol. 9, no. 5, 2018. doi: 10.14569/IJACSA.2018.090517.
- [3] A. Rahman, M. I. Siraji, L. I. Khalid et al., "Detection of mental state from EEG signal data: An investigation with machine learning classifiers," in *2022 14th Int. Conf. Knowl. Smart Technol. (KST)*, Jan. 2022, pp. 152–156. doi: 10.1109/KST53302.2022.9729084.
- [4] P. Bashivan, I. Rish, and S. Heisig, "Mental state recognition via wearable EEG," *arXiv: 1602.00985*, Jun. 5, 2016. doi: 10.48550/arXiv.1602.00985.
- [5] *SEED Dataset*. [Online]. Available: <https://bcmi.sjtu.edu.cn/home/seed/>
- [6] H. Masruroh, E. M. Imah, and E. Rahmawati, "Classification of emotional state based on EEG signal using AMGLVQ," *Procedia Comput. Sci.*, vol. 157, pp. 552–559, 2019. doi: 10.1016/j.procs.2019.09.013.
- [7] O. AlShorman, M. Masadeh, M. B. B. Heyat et al., "Frontal lobe real-time EEG analysis using machine learning techniques for mental stress detection," *J. Integr. Neurosci.*, vol. 21, no. 1, p. 20, 2022. doi: 10.31083/j.jin2101020.
- [8] W. Ma, Y. Zheng, T. Li, Z. Li, Y. Li, and L. Wang, "A comprehensive review of deep learning in EEG-based emotion recognition: Classifications, trends, and practical implications," *PeerJ Comput. Sci.*, vol. 10, p. e2065, May 2024. doi: 10.7717/peerj-cs.2065.
- [9] S. Rasheed, "A review of the role of machine learning techniques towards brain–computer interface applications," *Mach. Learn. Knowl. Extr.*, vol. 3, no. 4, pp. 835–862, 2021. doi: 10.3390/make3040042.
- [10] Appriou, A. Cichocki, and F. Lotte, "Modern machine-learning algorithms: For classifying cognitive and affective states from electroencephalography signals," *IEEE Syst., Man, Cybern. Mag.*, vol. 6, no. 3, pp. 29–38, 2020. doi: 10.1109/MSMC.2020.2968638.
- [11] M. J. Rivera, M. A. Teruel, A. Maté, and J. Trujillo, "Diagnosis and prognosis of mental disorders by means of EEG and deep learning: A systematic mapping study," *Artif. Intell. Rev.*, vol. 55, no. 2, pp. 1209–1251, 2022. doi: 10.1007/s10462-021-09986-y.
- [12] D. R. Edla, K. Mangalorekar, G. Dhavalikar, and S. Dodia, "Classification of EEG data for human mental state analysis using random forest classifier," *Procedia Comput. Sci.*, vol. 132, pp. 1523–1532, 2018. doi: 10.1016/j.procs.2018.05.116.
- [13] S.-Y. Han, N.-S. Kwak, T. Oh, and S.-W. Lee, "Classification of pilots' mental states using a multimodal deep learning network," *Biocybern. Biomed. Eng.*, vol. 40, no. 1, pp. 324–336, 2020. doi: 10.1016/j.bbe.2019.12.002.
- [14] Z. Jiao, X. Gao, Y. Wang, J. Li, and H. Xu, "Deep convolutional neural networks for mental load classification based on EEG data," *Pattern Recognit.*, vol. 76, pp. 582–595, Apr. 2018. doi: 10.1016/j.patcog.2017.12.002.
- [15] S. Sridhar and V. Manian, "EEG and deep learning based brain cognitive function classification," *Computers*, vol. 9, no. 4, p. 104, 2020. doi: 10.3390/computers9040104.

- [16] K. Ajith, R. Menaka, and S. S. Kumar, "EEG based mental state analysis," *J. Phys., Conf. Ser.*, vol. 1911, no. 1, p. 012014, 2021. doi: 10.1088/1742-6596/1911/1/012014.
- [17] Y. Badr, U. Tariq, F. Al-Shargie, F. Babiloni, F. Al Mughairbi, and H. Al-Nashash, "A review on evaluating mental stress by deep learning using EEG signals," *Neural Comput. Appl.*, vol. 36, no. 21, pp. 12629–12654, 2024. doi: 10.1007/s00521-024-09809-5.
- [18] R. Alhalaseh and S. Alasasfeh, "Machine-learning-based emotion recognition system using EEG signals," *Computers*, vol. 9, no. 4, p. 95, 2020. doi: 10.3390/computers9040095.
- [19] H. Manoharan, S. L. A. Haleem, S. Shitharth et al., "A machine learning algorithm for classification of mental tasks," *Comput. Electr. Eng.*, vol. 99, Apr. 2022, Art. no. 107785. doi: 10.1016/j.compeleceng.2022.107785.
- [20] H. Dose, J. S. Møller, H. K. Iversen, and S. Puthusserypady, "An end-to-end deep learning approach to MI-EEG signal classification for BCIs," *Expert Syst. Appl.*, vol. 114, pp. 532–542, Dec. 2018. doi: 10.1016/j.eswa.2018.08.031.
- [21] R. Katmah, F. Al-Shargie, U. Tariq, F. Babiloni, F. Al-Mughairbi, and H. Al-Nashash, "A review on mental stress assessment methods using EEG signals," *Sensors*, vol. 21, no. 15, p. 5043, 2021. doi: 10.3390/s21155043.
- [22] Y.-T. Lan, K. Ren, Y. Wang et al., "Seeing through the brain: Image reconstruction of visual perception from human brain signals," *arXiv: 2308.02510*, Aug. 16, 2023. doi: 10.48550/arXiv.2308.02510.
- [23] C. Anders and B. Arnrich, "Wearable electroencephalography and multi-modal mental state classification: A systematic literature review," *Comput. Biol. Med.*, vol. 150, Nov. 2022, Art. no. 106088. doi: 10.1016/j.combiomed.2022.106088.
- [24] R. Nahon, N. Bagheri, G. Varni, E. Tartaglione, and V.-T. Nguyen, "Mental state classification using EEG signals: Ethics, law and challenges," in *Mach. Learn. Princ. Pract. Knowl. Discov. Databases*, R. Meo and F. Silvestri, Eds., vol. 2133. Cham, Switzerland: Springer Nature, 2025. doi: 10.1007/978-3-031-74630-7\_28.
- [25] D. Pathak, R. Kashyap, and S. Rahamatkar, "A study of deep learning approach for the classification of electroencephalogram (EEG) brain signals," in *Artif. Intell. Mach. Learn. Edge Comput.* Elsevier, 2022. doi: 10.1016/B978-0-12-824054-0.00009-5.
- [26] Craik, Y. He, and J. L. Contreras-Vidal, "Deep learning for electroencephalogram (EEG) classification tasks: A review," *J. Neural Eng.*, vol. 16, no. 3, p. 031001, 2019. doi: 10.1088/1741-2552/ab0ab5.
- [27] M. A. Ozdemir, M. Degirmenci, O. Guren, and A. Akan, "EEG based emotional state estimation using 2-D deep learning technique," in *2019 Med. Technol. Congr. (TIPTEKNO)*, Oct. 2019, pp. 1–4. doi: 10.1109/TIPTEKNO.2019.8895158.