

Enhanced Depression Detection from EEG Using Deep Learning and Channel Selection on Synthetic-Balanced Data

Sudhir Dhekane¹ and Anand Khandare²

¹Thakur College of Engineering and Technology, Mumbai, India
Email: sudhir.dhekane@tcetmumbai.in

²Thakur College of Engineering and Technology, Mumbai, India
Email: sudhir.dhekane@djsce.ac.in

Abstract— EEG, or electroencephalogram, has emerged as a valuable and non-invasive method for evaluating depression. This study introducing a deep learning-based channel identification model for detecting depression using EEG signals. The dataset, sourced from AHEPA General Hospital, consists of EEG recordings from 293 patients, including 95 healthy individuals and 198 diagnosed with depression. To address the data imbalance, we applied Generative Adversarial Networks (GANs), increasing the healthy patient count to 184 and expanding the dataset to 382 subjects. EEG signals were recorded using 19 electrodes, following the 10–20 system across various frequency bands (Alpha, Beta, Delta, Gamma, High Beta, and High Gamma). ANOVA and Tukey’s test was used to filter the Significant Frequency Band (Alpha) for detailed research. Feature selection using Recursive Feature Elimination (RFE) reduced the model’s input to 11 key channels: Fp1, Fp2, F7, F4, F8, T3, C3, Cz, T4, T5, and P3. A Multi-Layer Perceptron (MLP) model was employed, achieving an accuracy of 98.70%, with a Precision Score of 1.00, an F1-Score of 0.983, and a Recall Score of 0.967. These results highlight the efficacy of using deep learning in identifying key brain channels for the detection of depression.

Keywords— Multi-Layer Perceptron, Electroencephalograms, Recursive Feature Elimination, Channels, Depression.

I. INTRODUCTION

Over 264 million people globally suffer from depression, one of the most prevalent and difficult mental health conditions that has an enormous impact on their everyday lives and overall well-being [1]. The emotional and physical symptoms significantly hinder an individual's ability to perform at work or in social settings. Depression frequently presents as chronic sorrow, a lack of interest in once-enjoyable activities of interest, and more [2]. Traditional methods for diagnosis which rely on the symptoms reported by people, can make accurate diagnosis challenging. It can result in false diagnosis or delayed treatment [3]. The timely detection of depression can greatly improve treatment plans and help prevent serious consequences like the alarmingly high risk of suicide among young adults [4]. It is imperative to create unbiased, data driven decisions to diagnose depression as the stigma related to mental health issues discourage people from pursuing treatment [5].

Advancements in technology have led to more precise and uniform diagnoses. They ensure that those affected receive the support and care required at a suitable time. It has also made it possible to use physiological signals such as electroencephalogram (EEG) data to evaluate mental health conditions. Neural patterns associated with depressive mood can be uncovered through EEG data. This provides an important method to study brain activity [6]. Deep learning algorithms can help us to better understand the underlying causes of depression by analysing EEG data and identifying key channels linked to it [7].

This study uses deep learning models trained on EEG data to understand the neural channels which correlate strongly with the signs of depression. This research attempts to develop a reliable framework for detecting symptoms of depression based on patterns of mental activity by utilizing neural networks, Gradient Boosting, Random Forest, and XGBoost, and other machine learning models. In order to improve patient outcomes and minimize the general negative impacts of depression on society, we hope to help create objective diagnostic tests for that can be easily integrated into clinical practice.

II. RELATED WORKS

Depression is a constant mental health illness which has badly affected the intellectual capacity and emotional well-being. Continuous sorrow disinterest in activities and irregularity in daily functioning are among its characteristics. According to Corrigan, P. W. et al. [5] early and accurate identification of depression is essential in delivering effective therapy nevertheless existing approaches to diagnosis mostly rely on subjective and variance with self-reports and clinical observations. Olbrich, S. et al. [6] stated that the potential of electroencephalogram (EEG) data to monitor real-time neuronal activity and said neural mechanisms that correspond to low moods has led to a boost in interest in using this information as a means of diagnosis for depression.

Deep learning models have shown their incredible capacity to interpret difficult EEG signals for the aim of spotting signals associated with depression. Huang, G. et al. [8] demonstrated that the Multilayer Perceptron (MLP) model is increasingly popular among these owing to its robust creation, which successfully represents non-linear relationships in data. The knack of MLP to handle multidimensional patterns makes it an ideal instrument for biological applications; this is especially the case when analyzing EEG data to establish gloom. Using EEG data, the MLP model was effectively trained to distinguish between depressed and non-depressive states.

Zhang, X. et al. [9] proposed that the model's multilayer structure allowed it to process to identify minor changes suggestive of depressed symptoms by processing the complex data patterns typical of brain activity. Such an ability is necessary for medical diagnostics, as advanced methods of analysis are often required due to the complexity of the data.

In machine learning, feature selection is important because it finds the most useful data points, improving model accuracy and reducing computational complexity. Guyon, I. et al. [10] used Recursive Feature Elimination (RFE) to detect essential EEG channels that are crucial for identifying depressed states, including F7, T5, Fp1, and F8. Riera-Serra, P. et al. [16] recommended that these channels are commonly associated with cognitive and emotional processes, which are often disrupted in individuals experiencing depression.

Focusing on specific EEG features not only streamlines data analysis but also enhances the accuracy of depression detection. Ros, F. et al. [11] proposed that this approach aligns with the broader trend in machine learning where reducing the data dimensionality helps in avoiding overfitting and improves the overall performance of diagnostic models. This scenario's implementation of the MLP model demonstrated how choosing features might result in improved precision and effectiveness of depression diagnostic results.

Comparisons were conducted with other sophisticated technologies such as CatBoost, LightGBM, and XGBoost in order to evaluate the performance of the MLP. Zhang, Y. et al. [12] proposed that these ensemble models excel at handling high-dimensional data and uncovering complex interactions, thus, the MLP Classifier was particularly effective in capturing non-linear patterns inherent in EEG signals. Yasin, S. et al. [15] claimed that its ability to represent data at multiple layers enabled a deeper analysis of neural activity, which is crucial in distinguishing subtle signs of depression from other mental states.

The comparative analysis verified MLP's scalability to a broad spectrum of biomedical applications, highlighting that it would be suitable for EEG data alongside neurophysiological datasets of various varieties. Acharya, U. R. et al. [14] claimed that this adaptability is key in developing scalable diagnostic tools that can be integrated into clinical practice for mental health assessment.

Although considering the MLP model's perks, there are still difficulties with deploying it globally to an assortment of patient populations. Due to age, gender, and context-dependent influences, EEG data demonstrates

significant differences between people in patterns. The study by He, H. et al. [13], integrated approaches such as data augmentation and transfer learning have been suggested proposed to conquer these issues and improve the model's resiliency and pliability.

By permitting models to apply previously acquired data to new datasets, transfer learning improves training on less amounts of data while preserving high precision. Still, Supratak, A. et al. [18] proposed that data augmentation produces synthetic data that enriches the dataset and strengthens the framework's capabilities to make assumptions in a variety of scenarios. In practical healthcare settings, these approaches have promise to enhance the efficacy existing EEG-based diagnostic models.

Integrating machine learning models like MLP with EEG data might revolutionize the recognition of depression. Riera-Serra, P. et al. [16] concluded that swift action dramatically improves the effects of therapy. It also mitigates the risk of suicide. Models enable objective and data-driven decisions and these lessen the shortcomings of conventional diagnostic approaches which depend on subjective evaluations.

In real-world applications, these diagnostic techniques help design and curate customized therapies to match a patient's neurological profile. Kaur, C. et al. [17] believe that healthcare providers can generate healthier outcomes by designing treatments which target the physiological causes of depression. It is crucial to understand the specific EEG patterns linked to depressive disorders to achieve this. The gaps between effective therapies and psychological diagnosis can be lessened by implementing these models in routine clinical practice. The analysis of EEG data by MLP framework is crucial in the identification of psychological illnesses. It is a proven model for diagnosing depressive states and takes advantages of advanced feature selection methods. It also focuses on non-stationary data patterns. Techniques like data augmentation and transfer learning enhance the adaptability of the model which makes it a powerful tool for clinical use. These data-driven methods offer more accurate, customized treatment to improve patient outcomes and hence could potentially be the way of the future for identifying depression.

III. METHODOLOGY

The proposed methodology shown in Figure. 2 explore a structured approach to classify MDD based on EEG data. The process starts with collecting EEG recordings from a standardized dataset that is AHEPA dataset. To handle class imbalance, the data is subjected to augmentation using Generative Adversarial Networks (GANs) and validated with Chi-square tests and p value analysis, ensuring a statistical validation of the augmented data. Data pre-processing techniques standardization, minimax scaling, ANOVA, and Tukey tests-are used for normalizing the data. Heuristic measures, like Asymmetric Variance Ratio (AVR) and Amplitude Asymmetry Ratio (AAR), techniques are combined with channel selection techniques such as Recursive Feature Elimination (RFE) and entropy based analysis for the selection of the most critical EEG channels. The features from important frequency bands, like Alpha, Beta, Gamma, Delta, High Beta, and High Gamma, are also extracted. A comparative heuristic approach is used to fine-tune channel selection by integrating AVR, AAR, and entropy measures. Further methodology focuses on, designing and training classifiers. The machine learning algorithms such as, Random Forest, Gradient Boosting, Support Vector Machines (SVM), Logistic Regression, AdaBoost, and Naïve Bayes implemented for feature extraction and classifications. In addition to that, deep learning models such as, Long Short-Term Memory (LSTM), Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and EEGNet are implemented and evaluated. Then the models use different performance metrics such as accuracy, precision, recall, F1-score, and ROC-AUC curve that helps to ensure, the evaluation of the implemented models and reliable classification results. This integrated pipeline allows the development of an optimized approach for stages detection in MDD in the proposed feature work, through a combination of advanced data pre-processing and feature selection along with robust classification techniques.

A. Dataset

A standard EEG dataset serves as the original source of the input. This dataset comprises EEG signal observations made in the second department of neurology at the AHEPA General Hospital in Thessaloniki [19]. This dataset consists of electroencephalograms of 293 patients divided into two groups: Healthy (95 subjects) and Depressive (198 subjects). For recording EEG, the dataset maker used a Nihon Kohen device (EEG-2100) with 19 electrodes situated on the head as per the 10–20 scheme: {Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, O2}. With sampling rate of 500 Hz for each signal. The duration of EEG recordings ranged between 5 min to 21 min. An inactive EEG was noted with the eyes closed. Here's a breakdown of what the labels mean:

- F: Frontal lobe

- P: Parietal lobe
- T: Temporal lobe
- C: Central region (between frontal and parietal lobes)
- O: Occipital lobe
- z: Midline (center of the head)
- Numbers indicate the lateralization:
 - Odd numbers (e.g., F3, C3) indicate the left hemisphere.
 - Even numbers (e.g., F4, C4) indicate the right hemisphere.
 - "z" (e.g., Fz, Cz, Pz) refers to the midline (along the center of the head).

Diagnosis is the target column with

- '0' indicating the patient is healthy
- '1' indicating the patient is depressed.

Generative Adversarial Networks (GANs) was used to increase the number of healthy patients from 95 to 184. This increased the dataset size to 382. The initial dataset was set with 19 channels records for each Frequency Band: Alpha, Beta, Delta, Gamma, High Beta and High Gamma. ANOVA and Tukey's test was used to filter the Significant Frequency Band (Alpha) for further research.

B. Feature Selection

Feature 1 selection is a machine learning process that involves selecting the most relevant features from a dataset to use in a model. In this study, feature selection was done using Recursive Feature Elimination. Top-ranked feature subsets determined via Recursive Feature Elimination (RFE) have been employed for assessing the performance of numerous machine-learning models. Based on the model's empirical feedback, RFE, a wrapper procedure, incrementally minimizes less essential features, remaining only the most significant ones. To identify interactions between traits and the target variable, the dataset first served to train the Random Forest Classifier. Then, attributes were classified by slowly eliminating the least noteworthy ones through Recursive Feature Elimination (RFE). RFE iteratively clipped the feature set until solely the top 15 features were retained, using the Random Forest Classifier as the foundational model. The importance of these traits was classified, with lower numbers implying more significance. Considering most models supplied good accuracy, feature counts from ten to fifteen were implemented for comparing the models.. The same top-ranked features were applied uniformly across all models without recalculating rankings for each, ensuring a consistent basis for comparison.

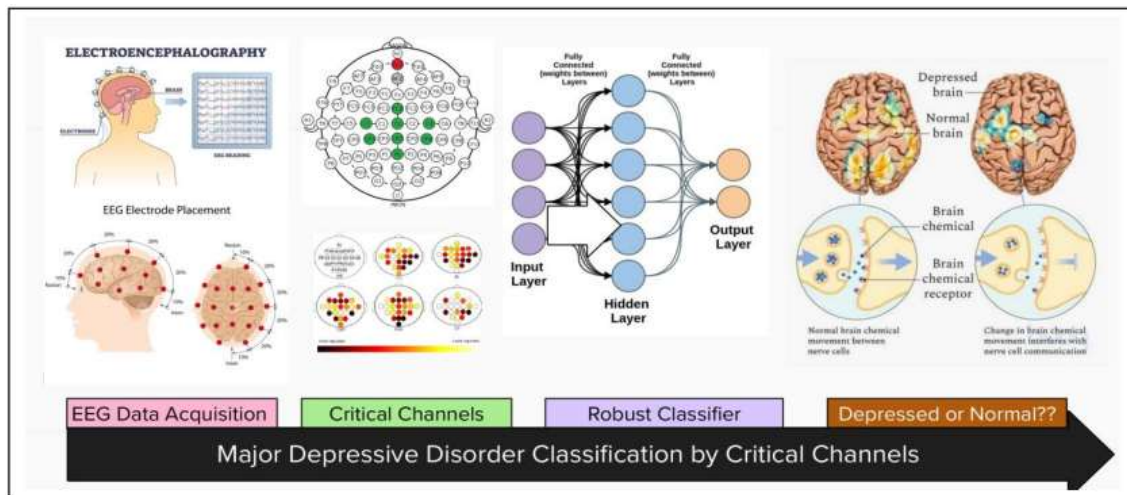


Figure 1: Process for Major Depressive Disorder Classification by Critical Channels

C. Model Selection

The features in the dataset were scaled using Min-Max Scaler. The data was also standardized. Nine distinct machine and deep learning models were fit to the data-Random Forest, Gradient Boosting, XGBoost, Logistic Regression, Support Vector Classifier (SVC), Multi-Layer Perceptron (MLP), TabNet, CatBoost and LGBM. The accuracies and performance achieved by these models after training on the entire dataset are summarized in Table I.

TABLE I: PERFORMANCE OF DIFFERENT ALGORITHMS ON ENTIRE DATASET

Sr No.	Model Name	Accuracy (in %)	Precision	F1-Score	Recall-Score
1)	Random Forest	88.31	0.8000	0.8615	0.9333
2)	Gradient Boosting	92.21	0.8529	0.9063	0.9667
3)	XGBoost	93.51	0.9032	0.9180	0.9333
4)	Logistic Regression	92.21	0.8333	0.9091	1.0000
5)	SVC	72.73	0.5918	0.7342	0.9667
6)	MLP Classifier	98.70	1.0000	0.9831	0.9667
7)	TabNet Classifier	50.65	0.4259	0.5476	0.7667
8)	CatBoost Classifier	92.21	0.8750	0.9032	0.9333
9)	LGBM Classifier	94.81	0.9063	0.9355	0.9667

Feature Selection was then applied to identify the channels in the human brain which can help detect depression. These classifiers underwent training on the dataset with reduced features with MLP demonstrating the highest accuracy amongst them with a commendable accuracy of 98.70%. It also displayed strong performance metrics with a Precision Score of 1.00, F1-Score of 0.983051 and Recall Score of 0.966667. Table II summarizes the accuracies of the algorithms with different feature counts.

TABLE II: ACCURACIES OF DIFFERENT ALGORITHMS AFTER FEATURE SELECTION

Sr No.	Model Name	Accuracy % (10 features)	Accuracy % (11 features)	Accuracy % (12 features)	Accuracy % (13 features)	Accuracy % (14 features)	Accuracy % (15 features)
1)	Random Forest	87.01	90.91	90.91	92.21	89.61	89.61
2)	Gradient Boosting	90.91	92.21	92.21	92.21	90.91	93.51
3)	XGBoost	92.21	96.10	94.81	94.81	93.51	93.51
4)	Logistic Regression	89.61	90.91	89.61	89.61	89.61	90.91
5)	SVC	76.62	74.03	74.03	74.03	74.03	72.73
6)	MLP Classifier	96.10	98.70	98.70	98.70	98.70	98.70
7)	TabNet Classifier	68.83	61.04	72.73	61.04	41.56	38.96
8)	CatBoost Classifier	88.31	94.81	92.21	94.81	93.51	93.51
9)	LGBM Classifier	92.21	92.21	94.81	97.40	96.10	93.51

The table above clearly shows that the MLP Classifier is the best model among all the algorithms, achieving a consistent accuracy of 98.70% with a feature selection of 11 or more. The 11 most important channels were identified and are displayed in Table III.

TABLE III: CRITICAL CHANNELS FOR DEPRESSION DETECTION

Sr. No	1	2	3	4	5	6	7	8	9	10	11
EEG Channel	Fp1	Fp2	F7	F4	F8	T3	C3	Cz	T4	T5	P3

These findings demonstrate that the identified channels can be effectively used for depression detection. Given the serious nature of depression, early detection through these channels can play a crucial role in timely intervention and treatment. By utilizing these methods, we can contribute to improving mental health outcomes and potentially mitigate the impact of depression through early-stage identification and support. Figure 2 compares the best accuracies of the nine models along with the number of features required to achieve it.

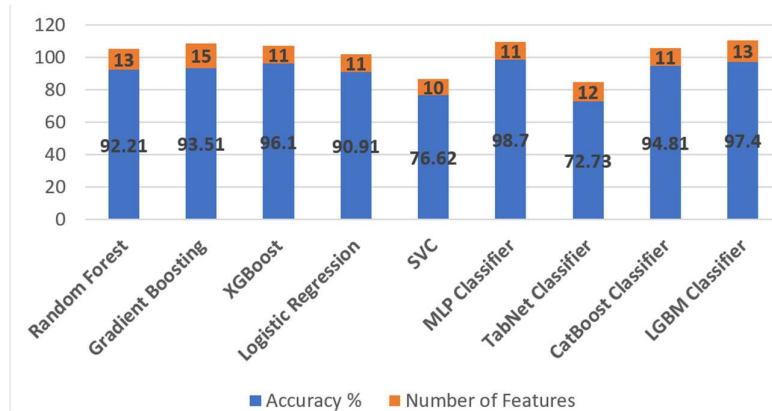


Figure 2: Accuracy of the learning with features

D. Flowchart of Methodology

Figure 3 shows a flowchart that summarizes the complete process in order to provide an in-depth understanding of the steps that are required. It makes the flow of the process easier to comprehend by providing a graphical representation of how it works from data collection to the final evaluation.

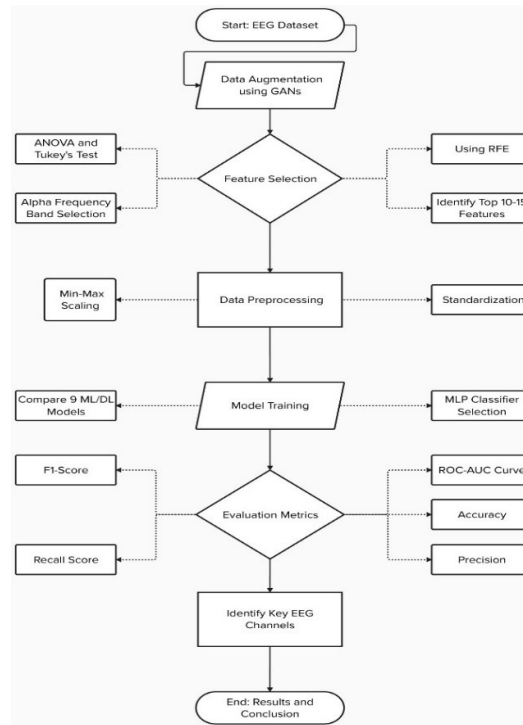


Figure 3: Flowchart depicting the workflow of the study

IV. RESULTS

The classification power of the MLP model is demonstrated with the help of the confusion matrix in Figure 4. Achieving a commendable accuracy of 98.70%, the model also achieved strong performance metrics with a precision of 1.00, recall score of 0.966667 and F1-Score of 0.983051. Confusion matrix is a crucial method to check how classification models work. The predictions are classified into four groups - True Positives (TP), False Positives (FP), True Negatives (TN) and False Negatives (FN). This division is crucial in the calculation of accuracy, precision, F1-score and recall score which give us better insights into the model's performance. The examination of the confusion matrix gives us a measure of the model's overall accuracy in addition to the specific regions where errors may arise resulting in greater emphasis on improvements in efficiency.

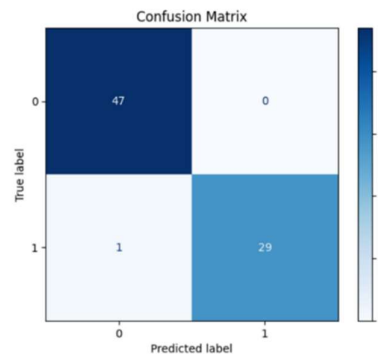


Figure 4: Confusion Matrix of MLP Model

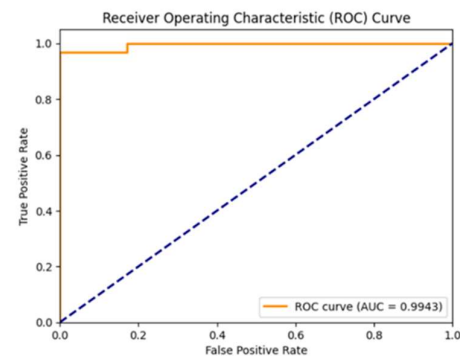


Figure 5: ROC-AUC Curve of MLP Model

The ROC-AUC curve (Receiver Operating Characteristic - Area Under the Curve) is important to evaluate the performance of binary classifiers. It plots the True Positive Rate (TPR) against the False Positive Rate (FPR). This visualises the relationship between recall (TPR) and fall-out (FPR) across different thresholds. The AUC (Area Under the Curve) measures the ability of the model to classify between classes. A higher AUC score indicates better model performance. This curve is particularly useful for comparing models and assessing their effectiveness. Figure 5 shows the ROC-AUC curve of the MLP model. It is calculated by the equation shown in equation 1.

$$\int_0^1 TPR(x) dFPR(x) \quad (1)$$

The high AUC score of 0.9943 of the MLP model demonstrates its powerful classifying power. The effectiveness of the chosen channels in detecting depression is indicated by the strong performance of the model. Therefore, these channels can help in the effective and quick identification of depression. This facilitates quick action and therapy.

V. CONCLUSION AND FUTURE WORK

The MLP model proposed by this study has been trained on EEG signal data from 382 patients to diagnose their depression. Through careful data analysis and training, the following EEG channels were identified as primary features for depression detection: Fp1, Fp2, F7, F4, F8, T3, C3, Cz, T4, T5, and P3. The high accuracy of the model of 98.70% showcases its ability in differentiating between depressed and non-depressed subjects. The ideal precision score of the model of 1.00 along with Recall-Score of 0.966 and F1-Score of 0.983 demonstrate its powerful and reliable performance. This helps us to conclude that the proposed model can be used for the early detection of depression demonstrating significant potential for clinical applications. To discuss about future work, though even we worked with small and dataset and enhanced accuracy by using stratified k-fold cross-validation to maintain class balance and by data augmentation, using GAN as per required class data to make it nearly balanced. In the further study, we have target, to scale dataset up to the mark by ensuring diversity in demographics information present in the current dataset. This we must have to explore with the objective of identification of MDD stages (Mild, Moderate, Server). In addition to that, to find efficiency in computational time of training period over entire dataset vs. training period with critical channels which will be ensuring the computational optimization in classification. Finally, explainable AI models such as LIME, SHAP, etc. will support our classification of MDD stages.

REFERENCES

- [1] Fields, S.A., Schueler, J., Arthur, K.M. et al. The Role of Impulsivity in Major Depression: A Systematic Review. *Curr Behav Neurosci Rep* 8, 38–50 (2021).
- [2] Li, Z., Ruan, M., Chen, J. et al. Major Depressive Disorder: Advances in Neuroscience Research and Translational Applications. *Neurosci. Bull.* 37, 863–880 (2021).
- [3] K. Kroenke and R. L. Spitzer, "The PHQ-9: A New Depression Diagnostic and Severity Measure," *Psychiatric Annals*, vol. 32, no. 9, pp. 509–515, Sep. 2002.
- [4] Riera-Serra, P., Navarra-Ventura, G., Castro, A. et al. Clinical predictors of suicidal ideation, suicide attempts and suicide death in depressive disorder: a systematic review and meta-analysis. *Eur Arch Psychiatry Clin Neurosci* 274, 1543–1563 (2024).
- [5] Corrigan, P. W., Druss, B. G., & Perlick, D. A. (2014). The Impact of Mental Illness Stigma on Seeking and Participating in Mental Health Care. *Psychological Science in the Public Interest*, 15(2), 37-70.
- [6] Olbrich, S., & Arns, M. (2013). EEG biomarkers in major depressive disorder: Discriminative power and prediction of treatment response. *International Review of Psychiatry*, 25(5), 604–618.
- [7] Ay, B., Yildirim, O., Talo, M. et al. Automated Depression Detection Using Deep Representation and Sequence Learning with EEG Signals. *J Med Syst* 43, 205 (2019).
- [8] G. Huang, Z. Liu, L. Van Der Maaten and K. Q. Weinberger, "Densely Connected Convolutional Networks," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017, pp. 2261-2269, doi: 10.1109/CVPR.2017.243
- [9] X. Zhang et al., "EEG-based Depression Detection Using Convolutional Neural Network with Demographic Attention Mechanism," 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Montreal, QC, Canada, 2020, pp. 128-133, doi: 10.1109/EMBC44109.2020.9175956.
- [10] Guyon, I., Weston, J., Barnhill, S. et al. Gene Selection for Cancer Classification using Support Vector Machines. *Machine Learning* 46, 389–422 (2002).
- [11] Ros, F., Riad, R. (2024). Dimensionality reduction. In: *Feature and Dimensionality Reduction for Clustering with Deep Learning. Unsupervised and Semi-Supervised Learning*. Springer, Cham.

- [12] Zhang, Y., Chen, J., Tan, J. H., Chen, Y., Chen, Y., Li, D., Yang, L., Su, J., Huang, X., & Che, W. (2020). An investigation of deep learning models for EEG-based emotion recognition. *Frontiers in Neuroscience*. Frontiers.
- [13] H. He and D. Wu, "Transfer Learning for Brain–Computer Interfaces: A Euclidean Space Data Alignment Approach," in *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 2, pp. 399–410, Feb. 2020, doi: 10.1109/TBME.2019.2913914.
- [14] U. Rajendra Acharya, Shu Lih Oh, Yuki Hagiwara, Jen Hong Tan, Hojjat Adeli, D. P Subha, Automated EEG-based screening of depression using deep convolutional neural network, *Computer Methods and Programs in Biomedicine*, Volume 161, 2018, Pages 103–113, ISSN 0169-2607
- [15] Sana Yasin, Syed Asad Hussain, Sinem Aslan, Imran Raza, Muhammad Muzammel, Alice Othmani, EEG based Major Depressive disorder and Bipolar disorder detection using Neural Networks: A review, *Computer Methods and Programs in Biomedicine*, Volume 202, 2021, 106007, ISSN 0169-2607
- [16] Riera-Serra, P., Navarra-Ventura, G., Castro, A. et al. Clinical predictors of suicidal ideation, suicide attempts and suicide death in depressive disorder: a systematic review and meta-analysis. *Eur Arch Psychiatry Clin Neurosci* 274, 1543–1563 (2024).
- [17] Chamandeep Kaur, Amandeep Bisht, Preeti Singh, Garima Joshi, EEG Signal denoising using hybrid approach of Variational Mode Decomposition and wavelets for depression, *Biomedical Signal Processing and Control*, Volume 65, 2021, 102337, ISSN 1746-8094
- [18] A. Supratak, H. Dong, C. Wu and Y. Guo, "DeepSleepNet: A Model for Automatic Sleep Stage Scoring Based on Raw Single-Channel EEG," in *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 11, pp. 1998–2008, Nov. 2017, doi: 10.1109/TNSRE.2017.2721116.
- [19] Velichko, A.; Belyaev, M.; Izotov, Y.; Murugappan, M.; Heidari, H. Neural Network Entropy (NNetEn): Entropy-Based EEG Signal and Chaotic Time Series Classification, Python Package for NNetEn Calculation. *Algorithms* 2023, 16, 255.