Machine Learning for Depression Diagnosis based on Resting-state Electroencephalography (EEG)

I. Introduction

Major depressive disorder (MDD), commonly termed depression, is a significant public health issue affecting people all over the world. It is characterized by at least one discrete depressive episode lasting at least two weeks and involving clear-cut changes in mood, interests and pleasure, cognition, and vegetative symptoms [1]. MDD was the second leading contributor to the global disease burden, as expressed in disability-adjusted life years, in both developed and developing countries [2]. Currently, the most common way to diagnose depression is an interview conducted by a medical expert. Clinical professionals commonly use the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) [3], or the International Statistical Classification of Diseases and Related Health Problems—11th revision (ICD-11) [4]. Diagnosis is often based on the subjective assessment and clinical experience of the clinicians and is further complicated by the fact that many psychiatric symptoms are polymorphous and may coexist in different psychiatric diseases [5].

Apart from subjective clinical assessment, brain activities of patients can be monitored objectively by applying various imaging modalities like Computed Tomography (CT), functional Magnetic Resonance Imaging (fMRI), and electroencephalography (EEG). EEG is a medical test used to measure the electrical activities of the brain and evaluate brain states and disorders. Among all methods used for diagnosing neurological disorders and studying brain functions, EEG stands out as the most simple and cost-effective device, making the detecting mental states and abnormalities using EEG an actively researched field showing promising results [7, 8].

Resting-state EEG (rsEEG) refers to endogenous or intrinsic neural activity without a specific stimulus or task imposed. Generally, rsEEG is recorded in the eyes-closed and/or eyes-open condition for a few minutes [15]. Although participants are simply doing nothing during rsEEG signal recording, it can provide plenty of meaningful information for diagnosing neurological disorders, including MDD. Previously, many studies have tried to classify rsEEG signals between MDD and healthy control (HC) groups by employing various feature extraction and data analysis methods. These include analysis of spectral features [16, 17], nonlinear neuronal dynamics [18, 19], functional connectivity [20, 21], microstate [22], or a combination of many features [23-25], and in many studies, conventional machine learning techniques [16, 18-20, 23-25] and deep learning technique [9, 10] are also applied.

Although machine learning methods have been widely employed, there may be a limitation due to a lack of specific domain knowledge. Each relevant study utilized specific features based on the assumption that these features could be useful for diagnosing depression. Therefore, this project aims to explore various linear and nonlinear features that have been previously used for depression classification, utilizing feature selection and dimensionality reduction methods to select meaningful features for machine learning. In addition, because of the need of feature extraction for conventional machine learning, prior knowledge is required to extract the truly meaningful features. Thus, I also aim to develop a novel neural network model for automatic depression classification that can work with raw EEG signals without the need of feature extraction procedure.

II. Method

a) Data Collection

This study used the EEG data set [11], a public dataset collected by Mumtaz et al. at the Hospital Universiti Sains Malaysia (HUSM). The dataset included EEG signals from 34 MDD patients (mean age = 40.33, SD = ±12.861) and 30 healthy controls (mean age = 38.227, SD = ±15.64). The experiment design for the HUSM study was approved by the ethics committee. The participants agreed to sign the consent forms and were fully aware of the experimental procedure adopted for experimental data acquisition. In addition, the MDD participants have been confirmed diagnosis based on the symptoms of depression as mentioned in the Diagnostic and Statistical Manual for Depression (DSM-IV) and underwent a two-week medication washing period before collecting EEG signals. MDD participants with psychotic symptoms, pregnant patients, alcoholics, smokers, and patients with epileptic problems were excluded. The healthy participants were examined for clinical symptoms to exclude the possibility of any physical and mental disability and were found normal.

The EEG data collection was based on the international 10–20 system [12] with linked ear (LE) as a reference. The 19 electrodes covering the scalp include frontal (Fp1, Fp2, F3, F4, F7, F8, Fpz), temporal (T3, T4, T5, T6), parietal (P3, P4, P7, P8), occipital (O1, O2), central (C3, C4). The data were filtered between 0 and 70.0 Hz, then a 50 Hz notch filter was used to reduce the effect of power line noise. The EEG data were recorded with the eyes closed for five minutes, the eyes open for five minutes, and with a specific visual stimulus (10 min). The data were gathered at a rate of 256 samples per second. The resting-state EEG signals, which are eye-closed (EC) and eye-opened (EO) EEG data from the HUSM data collection, were used in this study.

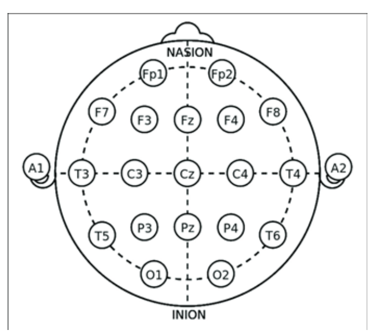

Fig. 1: Electrode locations of International 10-20 system for EEG recording[1]

[1] "The 10-20 International system of EEG electrode placement.." https://www.researchgate.net/figure/The-10-20-International-system-of-EEG-electrode-placement_fig1_324361441. Accessed 15 May. 2023.

b) EEG signal pre-processing

The recorded EEG data were confounded with composite signals generated by sources other than neuronal activity and were classified as artifacts or noise. Those include eye blinks, eye movements, and muscular activities (e.g., heartbeats); therefore, the reduction of noise from the recorded EEG data was a required pre-processing step before continuing further data analysis. In this study, an Independent Component Analysis (ICA) conducted based on the FastICA algorithm [13] was employed to decompose EEG signals into the weighted sum of multiple independent components (ICs), which was implemented with MNE-Python. Then, the ICLabel plugin [14] was used to automatically classify IC into brain signals and noise, such as eye movements and muscle activities. This enabled us to identify the artifact components and potentially remove them from the EEG signals.

c) Feature extraction

Seven features were extracted from EEG signals of all 19 channels. These include four EEG band power (delta, theta, alpha, and beta), sample entropy, Higuchi fractal dimension (HFD), and Hurst exponent. The total number of resulting features are 133 features which were then used as inputs of machine learning algorithms.

1. EEG band power

The EEG signals are filtered with band-pass butterworth filters to extract four common frequency bands, delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–12 Hz) and beta (12–30 Hz). For each band of each channel, the Welch method was applied to calculate the power spectrum of EEG bands [26]. In the Welch method, time series is divided into overlapping segments and then modified periodograms of all segments are averaged.

2. Sample Entropy

Sample Entropy (SampEn) was computed according to Richman and Moorman [27]. SampEn estimates signal complexity by computing the conditional probability that two sequences of a given length, m, similar for m points, remain similar within tolerance r at the next data point (when self-matches are not included). Mathematically, SampEn is the negative natural logarithm of the conditional probability that two sequences similar for m points remain similar at the next point. Thus, SampEn measures the irregularity of the data (the higher values, the less regular signal) that is related to signal complexity. For calculating SampEn, an MNE-features library was used.

3. Higuchi Fractal Dimension (HFD)

The fractal dimension of EEG was calculated using Higuchi's algorithm [28] demonstrated to be the most appropriate for electrophysiological data [29]. This method works directly in the time domain, gives a reasonable estimate of the fractal dimension even in the case of short signal segments and is computationally fast. The Higuchi algorithm was computed with the maximal scale kmax = 10. Fractal dimensions were calculated for each electrode for the same duration of signal (the epoch of recorded EEG) for all the participants. For calculating HFD, an MNE-features library was used.

4. Hurst exponent from Detrended fluctuation analysis (DFA)

The Hurst exponent was calculated based on DFA methods. DFA is applied to evaluate the presence and persistence of long range correlations in time in EEG signals [30]. Scaling exponent from DFA (aka. Hurst exponent) represent the correlation properties of the signals. For calculating Hurst exponent, an MNE-features library was used.

d) Statistical analysis

Student's t-tests were performed for 7 extracted features (delta power, theta power, alpha power, beta power, sample entropy, HFD and hurst exponent), calculated from EEG signal from each subject both eye-closed and eye-opened conditions to determine if there are any significant differences between MDD and HC groups.

e) Data Preparation

Before applying machine learning classification model, EEG data were prepared according to these following steps

i) Epoching

The artifact-free signals were chopped to increase the number of samples (Fig. 2).

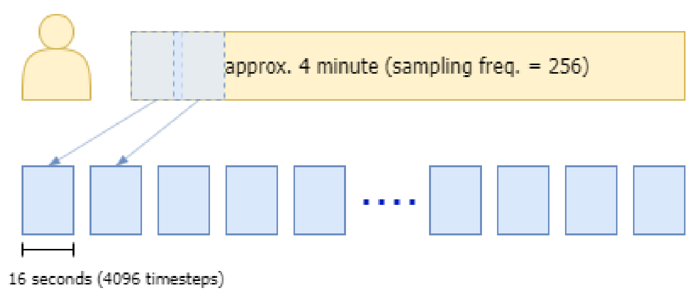

Fig. 2: Epoching procedure

ii) Train test split

The dataset has been splitted into training set and test set by subjects which can assure that EEG epochs from the same subject would be placed in the same set.

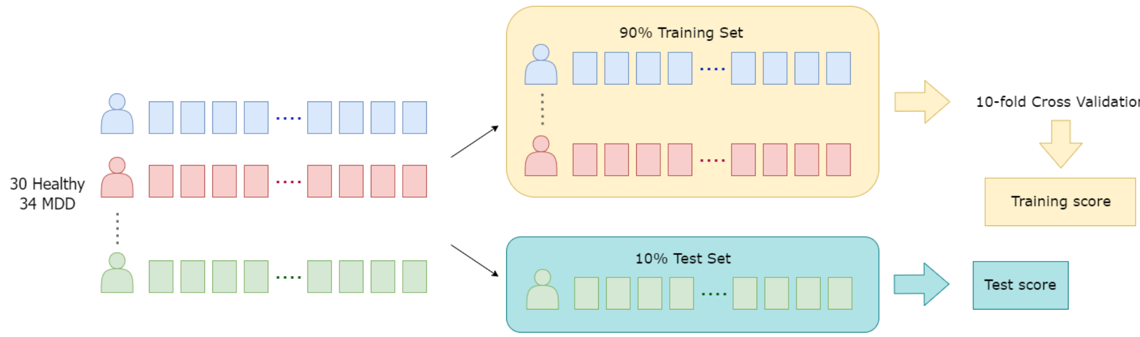

Fig. 3: Train test split procedure

f) Classification methods

i) Baseline method : State-of-the-art EEGNet-8,2

EEGNet-8,2 was proposed as a compact CNN model designed to capture discriminative EEG features [31]. Despite its compactness, EEGNet-8,2 demonstrated high robustness by achieving outstanding performance in various EEG-based brain–computer interface paradigms. It was widely accepted as the state-of-the-art method in various EEG classification tasks. However, EEGNet-8,2's performance on depression classification was not explored as of now. In this article, EEGNet-8,2 is reproduced for the depression detection task, keeping the optimal set of hyperparameters as recommended in the original publication [31].

ii) Conventional machine learning methods

ii-1) Feature selection + SVM

The univariate feature ranking algorithm helps to understand the significance of each feature by examining the importance of each predictor individually using an F-test. 30 features (cover >80% variation) out of 133 features were selected to be inputs of SVM classifier.

ii-2) Principal component analysis + SVM

The principal component analysis (PCA) is used to reduce the dimensionality of data by projecting them to a lower dimensional space which still contains most of the information of the data. First 10 principal components (cover >85% variation) were used as inputs of the SVM classifier.

iii) Extracted features + Fully-connected network with convolutional layer

All 133 features were used as inputs of the neural network model. The first layer is a convolutional layer which is designed to be a feature extraction layer. The subsequent layers are used to learn more complex signal characteristics.

iv) Novel neural network architecture : DepressioNet

1) Network design

According to the assumption that EEG signals can be viewed as a mixture of cortical source signals generated from different areas of the brain, the proposed network architecture has been designed to learn these spatial and temporal features of signals (Fig. 4). The first pointwise convolution layer is designed to function similar to a well-known independent component analysis (ICA) [32] or common spatial pattern (CSP) [33] which demix the signal into separate independent sources. Next, the depthwise layer is designed to learn temporal features from each demixed signal independently. Lastly, a ReLU activation layer, a batch normalization layer, a pooling layer, a dense layer, and a softmax activation layer were used for classification.

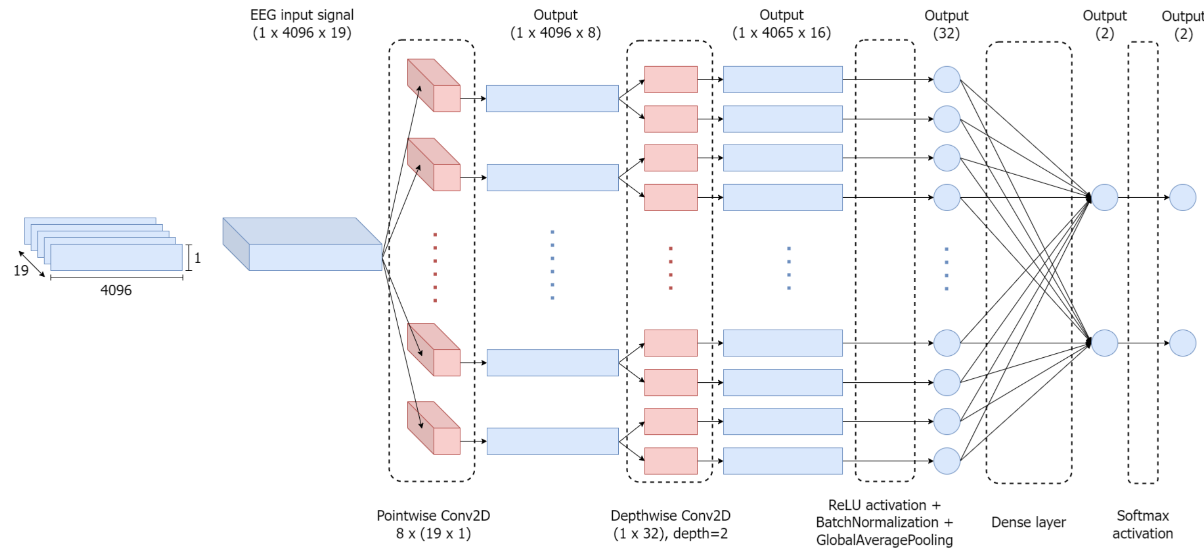

Fig. 4: DepressioNet architecture

g) Experimental Evaluation

The performance of each model was evaluated with accuracy, sensitivity and specificity. Accuracy computes the percentage of EEG segments that are computed correctly as in

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

where TP: true positives, TN: true negatives, FP: false positives, FN: false negatives.

Sensitivity provides the true positive rate of the relevance class (MDD) as demonstrated in (2), while specificity provides the true negative rate of the nonrelevance class (HC) as shown in (3).

$$\text{Sensitivity} = \frac{TP}{TP + FN} \tag{2}$$

$$\text{Specificity} = \frac{TN}{TN + FP} \tag{3}$$

All models have been evaluated using a 10-fold cross validation scheme to ensure reliability of results. Finally, the proposed DepressioNet model has been evaluated on an independent test set.

III. Results

a) Statistical analysis

Results from a Student's t-test demonstrated that there were significant differences between some extracted feature values, especially beta power, sample entropy and hurst exponent (**Table 1**).

Table 1. Features extracted from EEG signal of MDD patients and healthy controls

Features	EC				EO			
	Mean $\pm$ SD		t-stats	p-value	Mean $\pm$ SD		t-stats	p-value
	HC	MDD			HC	MDD		
Delta power	0.263 $\pm$ 0.103	0.160 $\pm$ 0.047	4.929	<0.001	0.339 $\pm$ 0.094	0.231 0.077	4.897	<0.001
Theta power	0.188 $\pm$ 0.042	0.182 $\pm$ 0.069	0.364	0.717	0.187 $\pm$ 0.048	0.195 $\pm$ 0.086	-0.435	0.665
Alpha power	0.276 $\pm$ 0.073	0.246 $\pm$ 0.062	1.699	0.095	0.209 $\pm$ 0.048	0.182 $\pm$ 0.043	2.245	0.028
Beta power	0.259 $\pm$ 0.033	0.343 $\pm$ 0.084	-4.919	<0.001	0.278 $\pm$ 0.323	0.358 $\pm$ 0.086	-4.737	<0.001
SampEnt	0.748 $\pm$ 0.091	0.676 $\pm$ 0.077	3.260	0.002	0.843 $\pm$ 0.113	0.709 $\pm$ 0.107	4.716	<0.001
HFD	1.341 $\pm$ 0.096	1.338 $\pm$ 0.079	0.142	0.887	1.445 $\pm$ 0.081	1.419 $\pm$ 0.078	1.261	0.212
Hurst exponent	0.337$\pm$ 0.065	0.135 $\pm$ 0.267	3.870	0.003	0.377 $\pm$ 0.055	0.221 $\pm$ 0.247	3.324	0.001

b) Classification results

Table 2 shows the classification performance of each model employed in this study. The best results are achieved by the PCA +SVM model. Results demonstrated that our proposed network, DepressioNet, provides classification accuracy, sensitivity and specificity scores of 0.911 ± 0.057, 0.922 ± 0.096 and 0.918 ± 0.116 respectively, which are greater than the scores from the state-of-the-art EEGNet-8,2. It also requires less training parameters than EEGNet-8,2 and less training time. Without the need of feature extraction, DepressioNet can provide comparable results to the conventional SVM model based on extracted features.

Table 2. Classification performance (mean ± sd) from 10-fold cross validation

Model	Accuracy	Sensitivity	Specitivity	Trainable parameters
EEGNet-8,2	0.793 ± 0.023	0.917 ± 0.043	0.669 ± 0.007	5,586
F-test + SVM	0.870 ± 0.095	0.865 ± 0.136	0.896 ± 0.121	NA
PCA + SVM	**0.940 ± 0.053**	**0.934 ± 0.089**	**0.960 ± 0.063**	NA
Neural network	0.935 ± 0.063	0.927 ± 0.127	0.952 ± 0.053	1,065
DepresioNet	0.911 ± 0.057	0.922 ± 0.096	0.918 ± 0.116	3,522

c) Results from independent test dataset

Table 3 and Fig.5 show DepressioNet performance on independent test dataset.

Table 3. Classification performance on test dataset

Model	Accuracy	Sensitivity	Specitivity
DepresioNet	0.889	0.933	0.844

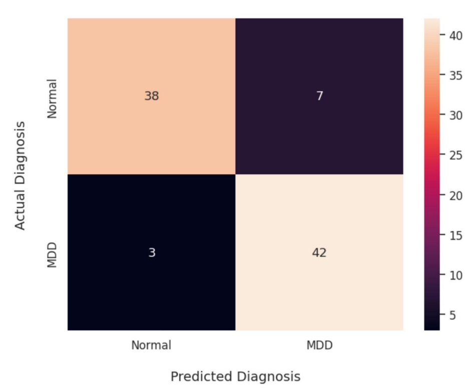

Fig. 5: Confusion matrix of predictions from DepressioNet on test dataset

IV. Conclusion and Discussion

In this study, I analyzed resting EEG of 34 MDD patients and 30 healthy controls by power of EEG bands (delta, theta, alpha and beta power) and three non-linear features (SampEnt, HFD, Hurst exponent). Results from statistical analysis indicate differences between brain signals of MDD patients and healthy controls. There are very prominent differences in beta power, sample entropy and hurst exponent. Elevated beta power is associated with stress, anxiety and brain desynchronization [34] found in MDD patients' brain oscillation. Significantly lower sample entropy in MDD patients indicates some underlying predictable patterns associated with neural processes which are absent in healthy people [35]. Hurst exponents of MDD patients' EEG signals are less than 0.5, which indicate anticorrelated LRTC behavior of signals that is absent in healthy controls. This behavior may be associated with brain physiology alteration related to depressive symptoms which is currently still unknown.

In classifying MDD patients and normal healthy subjects based on extracted features, the results demonstrated that conventional machine learning can perform better than state-of-the-art EEGNet-8,2 and achieve the maximum accuracy, sensitivity and specificity of 0.940 ± 0.053, 0.934 ± 0.089 and 0.960 ± 0.063 respectively (PCA + SVM method), suggesting that these features can potentially be used to differentiate between MDD patients and healthy people.

A novel neural network architecture for depression classification, DepressioNet, has also been designed and developed in this study. DepressioNet is an end-to-end model that works directly with EEG signals, meaning that feature extraction is not required. The model achieved 0.911 ± 0.057 accuracy, 0.922 ± 0.096 sensitivity and 0.918 ± 0.116 specificity, which are higher than EEGNet-8,2's classification scores. Moreover, DepressioNet can achieve pretty high scores on independent test dataset, suggesting that our model is generalized and well-performed on unseen data.

In conclusion, EEG signals can be a useful tool in studying depression. Our proposed end-to-end neural network has the capability of extracting important features from the EEG signals automatically and learning to distinguish between MDD and healthy subjects effectively. The results suggest the potential of using this framework as a depression screening or diagnostic tool in clinical practice.

V. References

[1] Otte, C. et al. (2016). Major depressive disorder. *Nature reviews Disease primers*, *2*(1), 1-20.

[2] Vos, T. et al (2015). Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *The lancet*, *386*(9995), 743-800.

[3] American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 5th Edition: DSM-5 (American Psychiatric Association, 2013).

[4] International Classification of Diseases, Eleventh Revision (ICD-11), World Health Organization (WHO) 2019/2021 https://icd.who.int/browse11.

[5] Ho, C. S. et al. (2020). Diagnostic and predictive applications of functional near-infrared spectroscopy for major depressive disorder: a systematic review. *Frontiers in psychiatry*, *11*, 378.

[6] Strawbridge, R., Young, A. H., & Cleare, A. J. (2017). Biomarkers for depression: recent insights, current challenges and future prospects. *Neuropsychiatric disease and treatment*, 1245-1262.

[7] de Aguiar Neto, F. S., & Rosa, J. L. G. (2019). Depression biomarkers using non-invasive EEG: A review. *Neuroscience & Biobehavioral Reviews*, *105*, 83-93.

[8] Mahato, S., & Paul, S. (2019). Electroencephalogram (EEG) signal analysis for diagnosis of major depressive disorder (MDD): a review. *Nanoelectronics, Circuits and Communication Systems: Proceeding of NCCS 2017*, 323-335.

[9] Safayari, A., & Bolhasani, H. (2021). Depression diagnosis by deep learning using EEG signals: A systematic review. *Medicine in Novel Technology and Devices*, *12*, 100102.

[10] Sarkar, A., Singh, A., & Chakraborty, R. (2022). A deep learning-based comparative study to track mental depression from EEG data. *Neuroscience Informatics*, 100039.

[11] Mumtaz, Wajid (2016): MDD Patients and Healthy Controls EEG Data (New). figshare. Dataset. https://doi.org/10.6084/m9.figshare.4244171.v2

[12] Jasper, H. H. (1958). Ten-twenty electrode system of the international federation. *Electroencephalogr. Clin. Neurophysiol.*, *10*, 371-375.

[13] Hyvarinen, A. (1999). Fast and robust fixed-point algorithms for independent component analysis. *IEEE transactions on Neural Networks*, *10*(3), 626-634.

[14] Pion-Tonachini, L., Kreutz-Delgado, K., & Makeig, S. (2019). ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. *NeuroImage*, *198*, 181-197.

[15] Snyder, A. Z., & Raichle, M. E. (2012). A brief history of the resting state: the Washington University perspective. *Neuroimage*, *62*(2), 902-910.

[16] Mumtaz, W., Xia, L., Ali, S. S. A., Yasin, M. A. M., Hussain, M., & Malik, A. S. (2017). Electroencephalogram (EEG)-based computer-aided technique to diagnose major depressive disorder (MDD). *Biomedical Signal Processing and Control*, *31*, 108-115.

[17] Grin-Yatsenko, V. A., Baas, I., Ponomarev, V. A., & Kropotov, J. D. (2009). EEG power spectra at early stages of depressive disorders. *Journal of clinical neurophysiology*, *26*(6), 401-406.

[18] Hosseinifard, B., Moradi, M. H., & Rostami, R. (2013). Classifying depression patients and normal subjects using machine learning techniques and nonlinear features from EEG signal. *Computer methods and programs in biomedicine*, *109*(3), 339-345.

[19] Acharya, U. R., Sudarshan, V. K., Adeli, H., Santhosh, J., Koh, J. E., Puthankatti, S. D., & Adeli, A. (2015). A novel depression diagnosis index using nonlinear features in EEG signals. *European neurology*, *74*(1-2), 79-83.

[20] Mumtaz, W., Ali, S. S. A., Yasin, M. A. M., & Malik, A. S. (2018). A machine learning framework involving EEG-based functional connectivity to diagnose major depressive disorder (MDD). *Medical & biological engineering & computing*, *56*, 233-246.

[21] Sun, S. et al. (2019). Graph theory analysis of functional connectivity in major depression disorder with high-density resting state EEG data. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *27*(3), 429-439.

[22] Murphy, M. et al. (2020). Abnormalities in electroencephalographic microstates are state and trait markers of major depressive disorder. *Neuropsychopharmacology*, *45*(12), 2030-2037.

[23] Liu, S., et al. (2022). Alterations in patients with first-episode depression in the eyes-open and eyes-closed conditions: A resting-state EEG study. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *30*, 1019-1029.

[24] Avots, E., Jermakovs, K., Bachmann, M., Päeske, L., Ozcinar, C., & Anbarjafari, G. (2022). Ensemble approach for detection of depression using EEG features. *Entropy*, *24*(2), 211.

[25] Cai, H., Qu, Z., Li, Z., Zhang, Y., Hu, X., & Hu, B. (2020). Feature-level fusion approaches based on multimodal EEG data for depression recognition. *Information Fusion*, *59*, 127-138.

[26] Welch, P. (1967). The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. *IEEE Transactions on audio and electroacoustics*, *15*(2), 70-73.

[27] Richman, J. S., & Moorman, J. R. (2000). Physiological time-series analysis using approximate entropy and sample entropy. *American journal of physiology-heart and circulatory physiology*.

[28] Higuchi, T. (1988). Approach to an irregular time series on the basis of the fractal theory. *Physica D: Nonlinear Phenomena*, *31*(2), 277-283.

[29] Esteller, R., Vachtsevanos, G., Echauz, J., & Litt, B. (2001). A comparison of waveform fractal dimension algorithms. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, *48*(2), 177-183.

[30] Lee, J. S., Yang, B. H., Lee, J. H., Choi, J. H., Choi, I. G., & Kim, S. B. (2007). Detrended fluctuation analysis of resting EEG in depressed outpatients and healthy controls. *Clinical Neurophysiology*, *118*(11), 2489-2496.

[31] Lawhern, Vernon J., et al. "EEGNet: a compact convolutional neural network for EEG-based brain–computer interfaces." *Journal of neural engineering* 15.5 (2018): 056013.

[32] Zhukov, L., Weinstein, D., & Johnson, C. (2000). Independent component analysis for EEG source localization. *IEEE Engineering in Medicine and Biology Magazine*, *19*(3), 87-96.

[33] Ang, K. K., Chin, Z. Y., Zhang, H., & Guan, C. (2008, June). Filter bank common spatial pattern (FBCSP) in brain-computer interface. In *2008 IEEE international joint conference on neural networks (IEEE world congress on computational intelligence)* (pp. 2390-2397). IEEE.

[34] Díaz, H., Cid, F. M., Otárola, J., Rojas, R., Alarcón, O., & Cañete, L. (2019). EEG Beta band frequency domain evaluation for assessing stress and anxiety in resting, eyes closed, basal conditions. *Procedia Computer Science*, *162*, 974-981.

[35] Faust, O., Ang, P. C. A., Puthankattil, S. D., & Joseph, P. K. (2014). Depression diagnosis support system based on EEG signal entropies. *Journal of mechanics in medicine and biology*, *14*(03), 1450035.