Resting-State EEG Classification of Children and Adolescents Diagnosed With Major Depressive Disorder Using Convolutional Neural Networks

A. Jahanian Najafabadi[1,2] and K. Bagh[3]

1. Dept. Cognitive Neuroscience, Bielefeld University, Bielefeld, Germany
2. School of Business, Social and Decision Sciences, Constructor University Bremen, Bremen, Germany
3. Dept. of Technology, Bielefeld University, Bielefeld, Germany
amir.jahanian@uni-bielefeld.de, khaled.bagh@uni-bielefeld.de

***Abstract*— The aim of this research was to use Convolutional Neural Network (CNN) based on VGG16 model to identify biomarkers in Children and Adolescents with Major Depressive Disorder (MDD) from age-matched healthy young individuals. For this purpose, resting-state eyes-closed electroencephalography (EEG) was pre-processed, and analysis was performed based on frequency bands and Regions Of Interest (ROI). Overall, results achieved an accuracy of 0.875 and F1-Score of 0.638, and further revealed lower Delta activity (0.856 accuracy and 0.539 F1-Score), higher Theta activity (0.895 accuracy and 0.717 F1-Score), and higher Alpha activity (0.804 accuracy and 0.606 F1-Score) which were classified in MDD compared to healthy individuals. In this study, Beta and Gamma activities were not found as biomarkers for MMD with higher accuracy. Moreover, results showed that while in MDD group, Delta frequency bands were featured in left temporal, occipital, bilateral frontal, and central regions; Theta frequency bands were featured in left temporal and frontal, left occipital, and central regions. The Alpha frequency band was further featured in left frontal, central, left occipital, and left temporal regions.**

Keywords*— *Electroencephalography, Children, Adolescents, Major Depressive Disorder, Convolutional Neural Network

I. Introduction

Clinical depression, also known as major depressive disorder (MDD) is a more severe form of depression. Unlike normal depression which can occur due to loss, such as the death of a loved one, MDD is similar to a disease that causes depression without the existence of a catalyst. The current investigation methods used for detecting MDD is based mainly on medical specialist e.g., psychiatric evaluation of patients' state using reported symptoms, events, occurrences and psychological assessments [1]. The causes of MDD are not yet clear, however, a major contributing cause is thought to be a combination of several factors including genetics, injury, trauma, etc.

The timely identification of MDD plays a vital role in effective treatment and can contribute to reducing the intensity of symptoms and improving patient outcomes [2] [3]. Hence, comprehending the biomarkers associated with MDD holds great significance in the diagnosis and treatment of this condition and can be used by medical specialists as a complementary or confirmatory method for their research and diagnostic purposes.

Biomarkers serve as objective indicators of biological, pathogenic, or pharmacological processes and responses, playing a crucial role in assessing therapeutic interventions [4]. The quest for non-invasive biomarkers continues, particularly in the realm of electroencephalography (EEG) research, as it holds the potential to provide an objective diagnosis of the disorder, free from the subjectivity of clinicians or patients [5].

The EEG recording as a cost effective and non-invasive method captures the electrical signals naturally produced by the brain by means of an electrogram. These signals have been proven to reflect the postsynaptic potentials of pyramidal neurons in both the neocortex and allocortex. The recorded resting-state EEG data is capable of representing the neurological condition of a person [6]. Under specific conditions (at rest, eye closed, eye open, etc.) it is possible to distinguish between a healthy brain waveform and unhealthy ones [7].

Using EEG in order to detect and correctly diagnose MDD can help provide the correct treatment for patients and help us better evaluate the different methods used for treating depression or help clinicians to further propose individualised treatment. However, analysing EEG data using only eye examination, behavioral measurements, or subjective judgment of clinicians are insufficient for diagnosing psychiatric and psychological disorders such as MDD [7]. EEG data contain a lot of information to allow humans to detect any existing patterns. To that end, there exist several approaches to extract useful information, some use mathematical tools [8], and some use machine based strategies, with the most promising being machine learning using Principal Component Analysis, Genetic Algorithm, Support Vector Machine, Linear Discriminant Analysis [9], and deep learning [10][11][12].

II. Background

II-A. Deep learning network in EEG classification

There has been several studies categorizing and testing the efficiency and accuracy of using Neural Networks (NNs) in analysing EEG data. Roy et al [13], reviewed over 154 different studies and found that the biggest trend among researchers is to use NNs on EEG data for classification tasks, with the most common architectures

used being CNNs, Recurrent Neural Networks (RNNs) and Auto Encoders (AEs). The use of NNs also showed a 5.4% increase in performance over other methods and benchmarks. There has been several studies that aimed to either distinguish Healthy Brain Networks (HBN) from MDD or for time-series forecasting in order to predict EEG signals [14][15].

II-B. EEG finding

A study by Ayan Seal et al. [16] tackles the classification of MDD. Their study investigated 33 subjects (18-normal, 15-depressed), and Independent Component Analysis (ICA) was performed in order to exclude eye movement components as the data were recorded with open eye state. Their study introduces and tests CNNs network called DeprNet and used a method to generate a heat map of the brain built of the CNN model that represents the area with the most noticeable difference between HBN and MDD and the network resulted in a 0.914 accuracy.

According to an extensive analysis of resting-state EEG frequency bands in psychiatric disorders, it was observed that in 18 studies investigating depression, individuals diagnosed with MDD displayed elevated absolute power in the theta and beta bands when compared to control groups. For example, in a recent study conducted by Zhang et al. (2020) [17], the MDD group exhibited significantly higher absolute power in the theta and alpha2 (10-12 Hz) bands across all electrodes. Furthermore, Grin-yatsenko et al. (2010) [18] reported increased power in the theta, alpha, and beta bands specifically in the occipital and parietal regions. In this line another recent study by Wolff et al. (2019) [19] revealed participants with MDD exhibited increased theta peak frequency and a reduced, more consistent coefficient of variation, along with decreased alpha peak frequency and an elevated, less stable coefficient of variation.

To our knowledge, despite a series of studies in this line of research, there is still no clear and optimized model introduced by the scientific community to detect abnormal biomarkers to distinguish MDD from the healthy group with highest accuracy and especially in children and adolescents. Additionally, lack of consistency observed in literature, e.g. with regard to the age of participants and sample sizes, motivated us to test CNNs to classify both normal and abnormal biomarkers in psychiatric disorders particularly MDD amongst children and adolescents contributing to this line of research and further applications.

III. Metrial and methods

III-A. MDD dataset

This study involved 214 datasets of children and adolescents aged 5 to 21, with 44 diagnosed with MDD and 170 classified as healthy (labelled HBN). The data were obtained from the publicly available Healthy Brain Network (HBN) dataset [20]. Resting-State data under closed eyes conditions were chosen for this study, recorded at a sampling rate of 500 Hz and a band pass of 0.1 to 100 Hz using a 128 Channel EEG HydroCel Geodesic system by Electrical Geodesics Inc. However, after processing and excluding outer channels, only 107 channels were retained. The EEG electrode distribution on the scalp is depicted in figure 1.

III-B. Pre-processing

Data pre-processing involved three stages (Code available at https://github.com/AmirJahanian/MDD_CNN_pipeline.git). Initially, data scraping entailed downloading datasets in raw CSV format using Amazon AWS service. Subsequently, the datasets were pre-processed in the second stage utilizing Python's MNE library. The default MNE GSN-HydroCel-128 montage was applied, excluding outer channels. The pyprep library's pipeline [21] was utilized to identify bad channels, followed by resampling the datasets to 256 Hz and applying a band-pass filter of 1-70 Hz. A notch filter at 60 Hz eliminated power line artifacts, commonly encountered in USA-recorded EEG data. Interpolation resolved any remaining bad channels, and a channel average-based re-referencing was conducted. Muscle movement was annotated using MNE's filter [22] with a Z-score of 5 and a frequency filter of 110-140 Hz, along with the application of ICA to detect ECG and EOG artifacts.

In the final stage, data was prepared for input into the Neural Networks. Each EEG dataset underwent filtering into distinct frequency bands (**Delta: 1-4 Hz, Theta: 4-8 Hz, Alpha: 8-12, Beta: 12-30, and Gamma: 30-70 Hz**) and was segmented into chunks of 4000 samples (15.6 seconds per chunk). TensorFlow's image resizing function with bilinear interpolation was used to resize the datasets to the shape (107, 350). The CNN module was configured with an input of (107, 350, 3), where the 3 RGB channels were duplicates of a single sample.

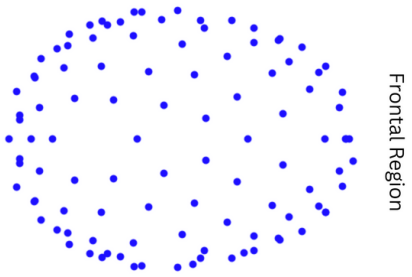

Figure 1. Channel location of 128 channels in the HBN dataset

III-C. Architecture of CNN

The CNN model's architecture is built upon the VGG16 network, with the top layers replaced by a global average pooling layer. This was succeeded by a flattening layer and two fully connected layers, each containing 512 neurons activated by ReLU. Subsequently, a dropout layer of 0.2 was inserted, followed by a batch normalization layer. Another dense layer comprising 62 neurons activated by ReLU was included, followed by a decision-making layer with 1 neuron activated by Sigmoid. The AdaBelief optimizer was employed [23], with binary cross-entropy serving as the loss function. The evaluation metrics employed were accuracy, F1 score, precision, sensitivity, and specificity.

While accuracy was measured, it was not utilized for analysis in this manuscript. Instead, emphasis was placed on specificity, sensitivity, and F1-Score. However, future work may incorporate different metrics.

III-D. Model training

During the CNN training implementation, various factors were taken into account, with particular emphasis on the data size. To overcome the limitations posed by the small dataset, we opted to segment the data, resulting in an expanded dataset of up to 5840 samples (MDD + HBN). While we did contemplate data augmentation, it was ultimately omitted from the paper due to its failure to significantly enhance the learning process, while the current method achieve the similar output without altering the data. Additionally, standard image manipulations such as rotation, cropping, and enlargement were not considered to avoid invalidating the ROI analysis discussed in the subsequent section. We also employed the SMOTE algorithm [24] to generate additional data, but this approach did not yield any improvements over the current implementation. Data augmentation was also considered to fix the data imbalance issue, but it was found that adjusting class weights was sufficient to resolve the issue without the need for artificial data generation.

To address overfitting, the model was trained using a subject-wise split to prevent the mixing of inherent subject features between the training and testing phases. Furthermore, we introduced an early stopping mechanism during the training process.

Several NN architectures have been explored, this included training Long Short-Term Memory (LSTM) network on the dataset as well as a Transformer network. However, our attempt to train the dataset on a Transformer network was unsuccessful due to limited processing power. Although this aspect, falling beyond the scope of this paper and it has been omitted from the discussion.

III-E. ROI analysis

Seven regions of interest (ROI) are identified, illustrated in figure 2, to discern discrepancies in MDD and HBN across different frequency bands. We aim to employ a generic detection method, different from more architecture specific method such as in DeprNet [16]. Our approach involves training and testing each network with the original pre-processed data containing the relevant features. Subsequently, specific channels related to each ROI are 'zeroed out' during testing, effectively removing all features associated with that ROI. This alteration is solely implemented in the test data, without affecting the training process.

The alteration in the test score following testing on each ROI may lead to two possible outcomes. First, it can create a bias towards HBN classification (high specificity and low sensitivity) or MDD classification (low specificity and high sensitivity), indicating the significance of features in that ROI for the corresponding frequency. Second, no bias might be observed, suggesting the lack of statistical importance in the ROI's features. i.e removing statistically insignificant features will not have any effect on the test score, as the network has already learned to ignore these features to some extent.

It is important to note that this logic remains valid only if the model is reliable, as indicated by the Area Under Curve (AUC) score and the F1-Score. Furthermore, it is crucial to ensure the independence of channel information during the resizing step, ensuring that zeroed channels do not affect features associated with other channels in different ROIs. This is achieved by maintaining the number of channels (107 in this case).

IV. Experimental results

IV-A. Setup

Training method was 5 K-fold cross-validation training instead of the usual 10 K-fold, the reason being is to maintain an average data split of 20 percent test data.

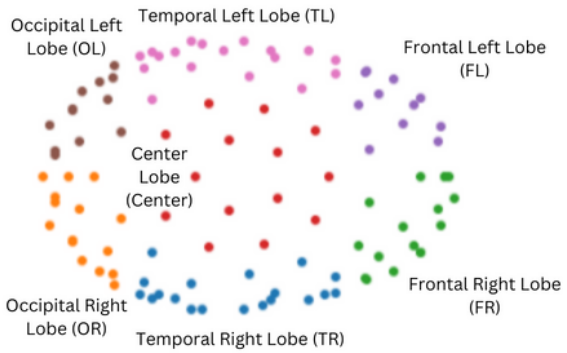

Figure 2. Region Of Interest (ROI)

The fold split was based on subjects. Five different models were trained for each fold. The training was set to 100 epoch with early stopping triggered when the F1-Score hits 0.8, batch size was set to 90 and learning rate set to 0.0001. Before initiating training, weights were added to each class to account for the classes imbalance. The weights were calculated in each fold according to equation [1].

$$W_j = \frac{n_{samples}}{n_{classes} * n_{jsamples}} \quad (1)$$

Where W_j is the weight of class j, $n_{samples}$ is the total number of samples used for training, $n_{classes}$ is the total number of classes, and $n_{jsamples}$ is the number of training samples belonging to class j. Additionaly, the training of the CNN was done in the cloud using Google Colab. With Python 3.10.12, MNE 1.4.2, Tensorflow 2.12.0, Keras 2.12.0, dask 2022.12.1, scipy 1.10.1, numpy 1.22.4. The system had a Intel(R) Xeon(R) CPU @ 2.00 GHz as CPU, 85 GB RAM, 12.0 CUDA version, A100 GPU and 40 GB GPU RAM.

IV-B. Results

Training results are shown in table 1. Training results over the entire frequency band (1-70 Hz) achieved an accuracy of 0.875 and F1-Score of 0.638. It is also noted that all frequency bands achieved slightly worse results except the Theta frequency band, which managed to achieve a better result of 0.895 accuracy and 0.717 F1-Score, indicating the clearest feature difference lies in the Theta band. We observed that all frequency bands

Table 1. CNN testing score in each frequency band

Frequency	Accuracy	F1 Scores	Precision	Sensitivity	Specificity
All	0.875	0.638	0.643	0.693	0.923
Delta	0.856	0.539	0.686	0.467	0.955
Theta	0.895	0.717	0.659	0.826	0.907
Alpha	0.804	0.606	0.524	0.761	0.815
Beta	0.820	0.526	0.525	0.570	0.877
Gamma	0.744	0.399	0.350	0.501	0.800

produced valid classifier in figure 3, but looking at table 2, we also observed that only Delta, Theta, and Alpha produces strong unbiased classifier. In figure 4 and table

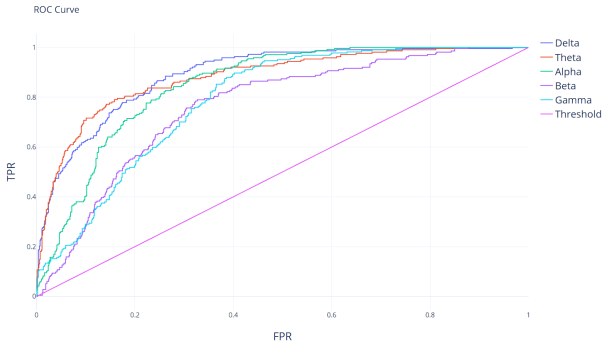

Figure 3. Receiver Operating Characteristic (ROC) Curve, where the threshold line represented a purely random classifier

Table 2. Area Under Curve (AUC), where score above 0.80 is an excellent discrimination score as specified by the rule of thumb in Hosmer and Lemeshow [25] (p. 177)

Frequency	Delta	Theta	Alpha	Beta	Gamma
AUC	0.890	0.881	0.848	0.763	0.788

3, interestingly, Delta shows the lowest change in F1-Score across all ROI (Specially in TR and OR). But, as reported in table 3 it is observed that zeroing any ROI causes a sharp increase in Sensitivity with a relative drop in Specificity. This would indicate some level of lower activities in MDD patients in this frequency band.

It was also observed that the highest drop in F1-Score is in Theta band, where FR, Center, FL, OL and TL showed the highest effect on classification, where F1-Score dropped to zero or near zero. These findings indicate that zeroing a ROI in this frequency band would make the classifier believe that a sample is healthy, which means that a higher level of activity is observed in Theta band for these ROIs for MDD patients compared to healthy controls. While TR and OR show a high level of similarity between MDD and healthy patients, destroying the features in that region did not allow the classifier to conclusively classify the test samples.

In Alpha, we can see that OL, TL, FL, and Center had the largest influence, indicating a high level of activity in these ROIs in Alpha bands for MDD patients compared to healthy controls.

As for Beta and Gamma frequency bands, we can't conclusively make a statement about the importance of these bands or ROIs. As seen in table 2 and 1, the classifiers in these cases had achieved a low AUC score or a low F1-Score.

Table 3. CNN testing scores on test data with ROI channels zeroed

Freq	ROI	F1 Scores	Precision	Sensitivity	Specificity
Delta	TR	0.418	0.278	0.900	0.335
	OR	0.476	0.339	0.856	0.528
	FR	0.352	0.219	**0.939**	**0.056**
	Center	0.347	0.216	**0.930**	**0.046**
	FL	0.342	0.212	**0.935**	**0.020**
	OL	0.349	0.216	**0.961**	**0.014**
	TL	0.350	0.216	**0.968**	**0.009**
Theta	TR	0.405	0.265	0.974	0.369
	OR	0.421	0.380	0.517	0.817
	FR	**0.045**	0.145	**0.028**	**0.999**
	Center	**0.001**	0.001	**0.001**	**1.000**
	FL	**0.001**	0.001	**0.001**	**1.000**
	OL	**0.001**	0.001	**0.001**	**0.997**
	TL	**0.001**	0.001	**0.001**	**0.999**
Alpha	TR	0.387	0.248	0.944	0.256
	OR	0.403	0.284	0.768	0.490
	FR	0.411	0.397	0.480	0.793
	Center	**0.065**	0.194	**0.040**	**0.981**
	FL	**0.001**	0.001	**0.001**	**1.000**
	OL	**0.001**	0.001	**0.001**	**1.000**
	TL	**0.001**	0.001	**0.001**	**1.000**

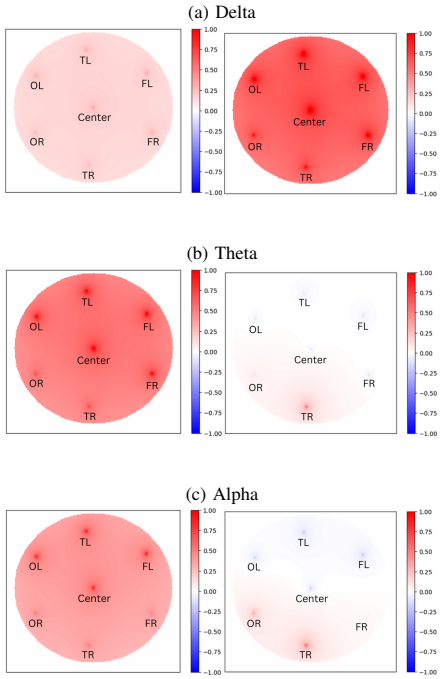

Figure 4. Heatmap for F1-Score(left) and Specificity(right). A red Z value (max score of 1) indicate a drop in test score after zeroing the ROI, a blue Z score (max score of -1) indicate an increase in test score after zeroing a ROI.

IV-C. Discussion

In this study, CNN was implemented to propose a diagnostic model for MDD using EEG as related biomarkers compared to the healthy children and adolescents. Our model managed to achieve a best score of 0.895 for accuracy and 0.717 F1-Score. Using a generic method to analyse our modules to extract useful information on ROI, our results revealed lower delta activity, higher theta activity and higher alpha activity in MDD group compared to healthy controls while we could not show any differences in beta and gamma frequency bands. Our results confirmed previous findings that delta [26], theta, and alpha are found to be associated with MDD symptoms [19]. The theta frequency band was featured in the left temporal, left occipital, central, and left frontal regions; the delta frequency band was featured in the left temporal, left occipital, bilateral frontal, and central regions; the alpha frequency band was featured in the left frontal, central, left occipital, and left temporal regions [27]. Increased theta activity is a biomarker for MDD patients compared to healthy controls.

Theta activity is commonly found in the temporal regions of the brain. The activity of the hippocampus has been associated with theta waves e.g., memory impairment and processes like working memory and episodic memory, as well as functions such as spatial navigation, attention, and learning. All of these functions can be disrupted in people with depressive disorders [7]. In addition, patients with MDD frequently tend to heavily introspect and reflect upon their self on a higher rate than the average healthy participant which highlight the association between delta activity and its role in MDD.

Alpha activities are related to state of relaxation and were also found in previous studies to be associated with MDD. Alpha activity is observed in different regions of the brain and predominantly in posterior regions and is involved in MDD leading to difficulties in memory tasks, concentration and attention [28]. However, differences in the methods and conditions of EEG datasets used in different studies could be responsible for some contradictions in reported findings which is very important for development of such diagnostic specific classification methods [2]. For example, beta band specifically was one of activities that was found to be associated with number of depressive periods in MDD patients and predicted quality of life after patients were successfully finished their treatment [29]. However, our classifier could not distinguish beta activities among both groups. Such discrepancy can be due to several factors, one being the average age difference, as the HBN dataset is for children and adolescents compared to the work Koshiyama et al. 2020 [29] which reported predictions in adults age participants. The other being the fact that the HBN dataset participants are yet to receive treatment.

Replication of this study using resting-state EEG in children and adolescents diagnosed with MDD would determine whether brain activities involved in MDD changes with age as previous studies e.g., [19] revealed lower theta activity in MDD adults (age 54±18 years) and healthy controls (age 46±16 years). We also suggest future studies to test the classifier on higher MDD sample size as we only had 46 MDD compared to 173 healthy controls, although we segmented our datasets and increased the samples up to 5840.

V. Conclusions

We conclude that MDD can be characterized by atypical delta, theta and alpha band activities. The potential clinical applications of these differences in the delta band requires further research. Therefore, we encourage future studies to closely look at delta, theta and alpha activities and to further investigate whether other frequency bands are involved in MDD before receiving any treatment (and co-existence of other disorders) amongst children and adolescents compared to healthy individuals.

Acknowledgments

This manuscript was prepared using limited access to the datasets obtained from the USA BioBank provided

by Child Mind Institute (CMI), Healthy Brain Network. This manuscript reflects the views of the authors and does not necessarily reflect the opinions or views of the CMI. EEG Datasets were credited by CMI to Dr. Amir Jahanian Najafabadi under the official agreement signed by Constructor University which is highly appreciated (formerly known as Jacobs University Bremen).

REFERENCES

[1] W. M. Reynolds, "A model for the screening and identification of depressed children and adolescents in school settings." *Professional school psychology*, vol. 1, no. 2, p. 117, 1986.

[2] P. C. Koo, J. Thome, C. Berger, P. Foley, and J. Hoeppner, "Current source density analysis of resting state eeg in depression: a review," *Journal of Neural Transmission*, vol. 124, pp. 109–118, 2017.

[3] T. Mendelson and S. D. Tandon, "Prevention of depression in childhood and adolescence," *Child and Adolescent Psychiatric Clinics*, vol. 25, no. 2, pp. 201–218, 2016.

[4] B. D. W. Group, A. J. Atkinson Jr, W. A. Colburn, V. G. DeGruttola, D. L. DeMets, G. J. Downing, D. F. Hoth, J. A. Oates, C. C. Peck, R. T. Schooley *et al.*, "Biomarkers and surrogate endpoints: preferred definitions and conceptual framework," *Clinical pharmacology & therapeutics*, vol. 69, no. 3, pp. 89–95, 2001.

[5] F. S. de Aguiar Neto and J. L. G. Rosa, "Depression biomarkers using non-invasive eeg: A review," *Neuroscience & Biobehavioral Reviews*, vol. 105, pp. 83–93, 2019.

[6] T. R. Goodwin and S. M. Harabagiu, "Multi-modal patient cohort identification from eeg report and signal data," *AMIA Annual Symposium Proceedings*, vol. 2016. American Medical Informatics Association, 2016, p. 1794.

[7] Y. Liu, C. Pu, S. Xia, D. Deng, X. Wang, and M. Li, "Machine learning approaches for diagnosing depression using eeg: A review," *Translational Neuroscience*, vol. 13, no. 1, pp. 224–235, 2022.

[8] S. Mahato and S. Paul, "Detection of major depressive disorder using linear and non-linear features from eeg signals," *Microsystem Technologies*, vol. 25, no. 3, pp. 1065–1076, 2019.

[9] S. Mahato and S. Paul, "Electroencephalogram (eeg) signal analysis for diagnosis of major depressive disorder (mdd): a review," *Nanoelectronics, Circuits and Communication Systems: Proceeding of NCCS 2017*, pp. 323–335, 2019.

[10] G. Li, C. H. Lee, J. J. Jung, Y. C. Youn, and D. Camacho, "Deep learning for eeg data analytics: A survey," *Concurrency and Computation: Practice and Experience*, vol. 32, no. 18, p. e5199, 2020.

[11] P. Nagabushanam, S. T. George, and S. Radha, "Eeg signal classification using lstm and improved neural network algorithms," *Soft Computing*, pp. 1–23, 2019.

[12] A. Craik, Y. He, and J. L. Contreras-Vidal, "Deep learning for electroencephalogram (eeg) classification tasks: a review," *Journal of neural engineering*, vol. 16, no. 3, p. 031001, 2019.

[13] Y. Roy, H. Banville, I. Albuquerque, A. Gramfort, T. H. Falk, and J. Faubert, "Deep learning-based electroencephalography analysis: a systematic review," *Journal of neural engineering*, vol. 16, no. 5, p. 051001, 2019.

[14] C. Bhaskarachary, A. J. Najafabadi, and B. Godde, "Machine learning supervised classification methodology for autism spectrum disorder based on resting-state electroencephalography (eeg) signals," *2020 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*. IEEE, 2020, pp. 1–4.

[15] G. Ciodaro, A. J. Najafabadi, and B. Godde, "Resting state eeg classification of children with adhd," *2020 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*. IEEE, 2020, pp. 1–6.

[16] A. Seal, R. Bajpai, J. Agnihotri, A. Yazidi, E. Herrera-Viedma, and O. Krejcar, "Deprnet: A deep convolution neural network framework for detecting depression using eeg," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–13, 2021.

[17] B. Zhang, G. Yan, Z. Yang, Y. Su, J. Wang, and T. Lei, "Brain functional networks based on resting-state eeg data for major depressive disorder analysis and classification," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 29, pp. 215–229, 2020.

[18] V. A. Grin-Yatsenko, I. Baas, V. A. Ponomarev, and J. D. Kropotov, "Independent component approach to the analysis of eeg recordings at early stages of depressive disorders," *Clinical Neurophysiology*, vol. 121, no. 3, pp. 281–289, 2010.

[19] A. Wolff, S. De la Salle, A. Sorgini, E. Lynn, P. Blier, V. Knott, and G. Northoff, "Atypical temporal dynamics of resting state shapes stimulus-evoked activity in depression—an eeg study on rest–stimulus interaction," *Frontiers in Psychiatry*, vol. 10, p. 719, 2019.

[20] L. M. Alexander, J. Escalera, L. Ai, C. Andreotti, K. Febre, A. Mangone, N. Vega-Potler, N. Langer, A. Alexander, M. Kovacs *et al.*, "An open resource for transdiagnostic research in pediatric mental health and learning disorders," *Scientific data*, vol. 4, no. 1, pp. 1–26, 2017.

[21] N. Bigdely-Shamlo, T. Mullen, C. Kothe, K.-M. Su, and K. A. Robbins, "The prep pipeline: standardized preprocessing for large-scale eeg analysis," *Frontiers in neuroinformatics*, vol. 9, p. 16, 2015.

[22] "Mne - annotate muscle artifacts," https://mne.tools/dev/auto_examples/preprocessing/muscle_detection.html.

[23] J. Zhuang, T. Tang, Y. Ding, S. C. Tatikonda, N. Dvornek, X. Papademetris, and J. Duncan, "Adabelief optimizer: Adapting stepsizes by the belief in observed gradients," *Advances in neural information processing systems*, vol. 33, pp. 18 795–18 806, 2020.

[24] J. Wang, M. Xu, H. Wang, and J. Zhang, "Classification of imbalanced data by using the smote algorithm and locally linear embedding," *2006 8th international Conference on Signal Processing*, vol. 3. IEEE, 2006.

[25] D. W. Hosmer Jr, S. Lemeshow, and R. X. Sturdivant, *Applied logistic regression*. John Wiley & Sons, 2013, vol. 398.

[26] C. Uyulan, T. T. Ergüzel, H. Unubol, M. Cebi, G. H. Sayar, M. Nezhad Asad, and N. Tarhan, "Major depressive disorder classification based on different convolutional neural network models: Deep learning approach," *Clinical EEG and neuroscience*, vol. 52, no. 1, pp. 38–51, 2021.

[27] S. Sun, X. Li, J. Zhu, Y. Wang, R. La, X. Zhang, L. Wei, and B. Hu, "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*, vol. 27, no. 3, pp. 429–439, 2019.

[28] P. Fernández-Palleiro, T. Rivera-Baltanás, D. Rodrigues-Amorim, S. Fernández-Gil, M. del Carmen Vallejo-Curto, M. Álvarez-Ariza, M. López, C. Rodriguez-Jamardo, J. Luis Benavente, E. de Las Heras *et al.*, "Brainwaves oscillations as a potential biomarker for major depression disorder risk," *Clinical EEG and neuroscience*, vol. 51, no. 1, pp. 3–9, 2020.

[29] D. Koshiyama, K. Kirihara, K. Usui, M. Tada, M. Fujioka, S. Morita, S. Kawakami, M. Yamagishi, H. Sakurada, E. Sakakibara *et al.*, "Resting-state eeg beta band power predicts quality of life outcomes in patients with depressive disorders: A longitudinal investigation," *Journal of Affective Disorders*, vol. 265, pp. 416–422, 2020.