Title: Major Depressive Disorder Diagnostics: Early Detection Through Advance Deep Learning Analysis of Physiological Signal

Abstract — Major Depressive Disorder (MDD) is a prevalent mental illness that poses significant challenges in diagnosis, traditionally relying on subjective assessments like personal questionnaires. These methods often result in inconsistent and inaccurate diagnoses. To address this issue, we propose an advanced diagnostic approach utilizing a one-dimensional convolutional neural network and long short-term memory (1DCNN-LSTM) model, designed to differentiate between MDD patients and healthy controls using electroencephalogram (EEG) data. Our model was trained using EEG data from multiple conditions (Eyes Closed, Eyes Open, Task, and their combinations) and extracted features from the time domain, frequency domain, and discrete wavelet transform (DWT). The model achieved outstanding results, particularly when analyzing eyes open data using time-domain features, with an accuracy of 99.67%, precision of 99.66%, recall of 99.70%, F1-score of 99.68%, and ROC-AUC of 99.70% and Combining all conditions the model achieved 99.52% accuracy. These results underscore the model's potential to enhance the early detection and diagnosis of MDD, offering a reliable tool for clinical application.

Key Terms — Deep Learning, Major Depressive Disorder or MDD, Depression, EEG, Time Domain, 1DCNN-LSTM, Discrete Wavelet Transform (DWT), Brainwaves, and Frequency Domain.

1. INTRODUCTION:

Major depressive disorder (MDD), a leading cause of disability, is characterized by a emotional instability, loss of interest and pleasure in previously enjoyable activities, recurrent thought of death, as well as physiological symptoms such as fatigue, weight loss, lack of motivation, sleep issues, and appetite disturbances [1], [2], [3]. According to the World Health Organization (WHO), MDD is the single largest contributor to global disability with an annual growth of 50 millions people-year, impacting 300 millions people of all ages worldwide [4], [5], [6]. Additionally, MDD also contributes to the economic burden, with overall direct and indirect cost (such as lost productivity) in the United States (USA) alone rising to 333.7 billion dollars in 2019 [7]. To decrease the impact of MDD on global health and productivity, early detection is essential for preventive strategies and effective treatment [8]. Traditional clinical diagnosis of MDD utilizes interviews and standard questionnaires such as Hamilton Depression Rating Scale (HDRS) and the Beck Depression Inventory (BDI) [9], but reliance on the patient's self report can be biased, resulting in diagnostic inconsistency and delayed treatment [10], [11]. Incorporating neurological biomarkers such as Electroencephalogram (EEG), a promising solution for improving the precision and validity of MDD diagnosis approach [12], [13], [14].

EEG captures the neurological activity of the patient, enabling real-time assessment of the individual's neural activity [15]. By analyzing these neural activities, cognitive and emotional disturbance can be identified [16]. Incorporating EEG signals with the clinical approaches, disturbance of neural activities can be identified, leading to an unbiased, and more improved disease management of MDD [10], [17]. With the advancements of Machine Learning (ML) and Deep Learning (DL), enabled extracting complex patterns from the EEG signal, further enhancing the detection precision of MDD. Compared to traditional techniques, ML, and DL exhibit superior recognition of neurophysiological changes, facilitating robust disease control and improved diagnostic outcomes [18], [19].

Several studies have employed ML and DL technologies for EEG-based MDD diagnosis. A ML and DL framework for MDD detection was proposed by Bashir et al.[20], emphasizing on the time, frequency, and non-linear domain features. They achieved the accuracy of 87.5% and 83.3% using

KNN and LSTM, respectively. However, Rafiei et al. [21] surpassed this performance with a customized InceptionTime model with the accuracy of 91.67%. After reducing the channel they achieved 87.5% accuracy. Though further improvement can be made with optimizing the channels. Ning et al. [22] developed a consumer-oriented depression detection system that is introduced by combining the EEG and speech signals. They achieved 86.11% and 87.44% accuracy on MDD and Healthy control recognition using a multi-model approach, though this performance can be significantly improved. Zhang et al. [23] proposed a secondary subject partitioning and attention mechanism (AM) based on a graph convolution network (GCN) to detect MDD, achieving 92.87% and 83.17% accuracy on two public datasets. Ksibi et al.[24] outperformed this study using multiple Machine Learning techniques. In their study, they marged demographic data with the extracted linear and non-learning features from the EEG signal, achieving 97% accuracy using the Convolutional Neural Network (CNN) model. However, they worked with 128 electrodes but using fewer electrodes can save resources and perform better. Wang et al. [25] introduced Graph convolutional Transformer Network (GCTNet) for detecting MDD using temporal and spatial features extracted from the EEG signal. In the study, they achieved the AUC scores of 0.7693% and 0.9755% on two datasets. This performance can be enhanced by better channel and feature selection.

Anik et al. [26] proposed 15s gamma band based MDD detention method, where they achieved 99.6% accuracy using optimized CNN model. In Spite of the outstanding performance, a more resource efficient approach can benefit efficient interpretation of MDD. Mahato and paul [27] divided the brain into six regions and extracted non-linear features to detect MDD, achieving 95.23% accuracy on utilizing SVM along with RelieF [28], While this region based method performed well, Choudhary et al. [29] surpassed it by utilizing fisher score-based method to reduce EEG channels, yielding 98.73% accuracy on dataset 1 and 95.48% accuracy on dataset 2. This highlights the effectiveness of optimized channel reduction to enhance both accuracy and resource efficiency in MDD detection.

Ying et al. [30] introduced EEG-based Depression Transformer (EDT) by employing an attention mechanism for feature learning and DL to MDD detection, achieving 92.25 ± 4.83% accuracy. Although this extensive sophisticated feature extraction, further improvements in performance can be achieved through more efficient feature selection. Wang et al. [13] proposed DiffMDD, a diffusion-based CNN framework along with a Reverse Diffusion Data Augmentation Module, acquiring 94% and 80% accuracy on two different MDD signal data. While effective, there is room for improvement in feature selection and performance. In the study by Dang et al. [31], Multivariate Pseudo Wigner Distribution is utilized to extract the time-frequency characteristics from the EEG signal, combining a frequency-dependent multilayer brain network with CNN. they achieved 97.27% accuracy, but a significant improvement can be done using optimal Channels and features. Xia et al. [32] provided a solution of the multimodal fusion problem using a GNN-BiLSTM fusion model, achieving 86.1% accuracy in detecting depression. Although the approach is effective, performance can be refined by feature selection. Chen et al. [33] proposed a hybrid CNN-LSTM model named DCLNet for depression classification, focusing on capturing temporal information the model acquired 99.15% accuracy, 99.01% specificity, and 99.30% sensitivity respectively. Despite this remarkable performance there is a room for improvement in effective feature extraction, channel reduction, and building a more effective and resource efficient model.

In this study, We aim to improve and introduce more reliable architecture to address these challenges by employing advanced features and DL based channel selection with a relatively smaller time segment to detect MDD more efficiently from EEG signals. Our contributions are as follows:

1) We proposed OptMDD architecture using optimal 1D-CNN and Bi-LSTM, validating it on a public dataset with 10-fold cross-validation.
2) We compared and selected feature extraction techniques that effectively contribute to MDD detection, further refining our proposed framework.
3) We performed an DL based evaluation approach for channel selection to identify the channels more responsible in MDD detection improving the resource effectiveness of the detection.

The remaining paper is organized as follows: Section 1) Materials and Methods: discusses the workflow and dataset used in this study. Section 2) Result: provides every result of our experiments. Section 3) Discussion: describes the findings and compares with the existing studies and Section 4) Conclusion: concludes this paper.

2.MATERIALS AND METHODS

3.1 Dataset Description

In this study, we utilized a publicly available "MDD Patients and Healthy Controls EEG Data (New)" dataset by Wajid Mumtaz to detect MDD based on the EEG signals. This dataset consists of a total 64 participants EEG data. Among those participants 34 individuals (17 males and 17 females) aged 40.3 years (±12.9) were diagnosed with MDD and 30 individuals (21 males and 9 females) aged 38.3 years (±15.6) were healthy controls (HC) participants. For acquisition, 19 electro-gel sensor caps were installed for both MDD and HC subjects during Eye open (EO), Eye close (EC), and Task states. The EEG recording adhered to the standard 10-20 montage system, with a linked ear reference. For data sampling 256 hz frequency was utilized [21]. All participants were conscious about the data acquisition process and the experimental setup. The Ethics committee of Hospital Universiti Sains Malaysia(HUSM) reviewed and approved the data acquisition process.

3.2 Experimental methods:

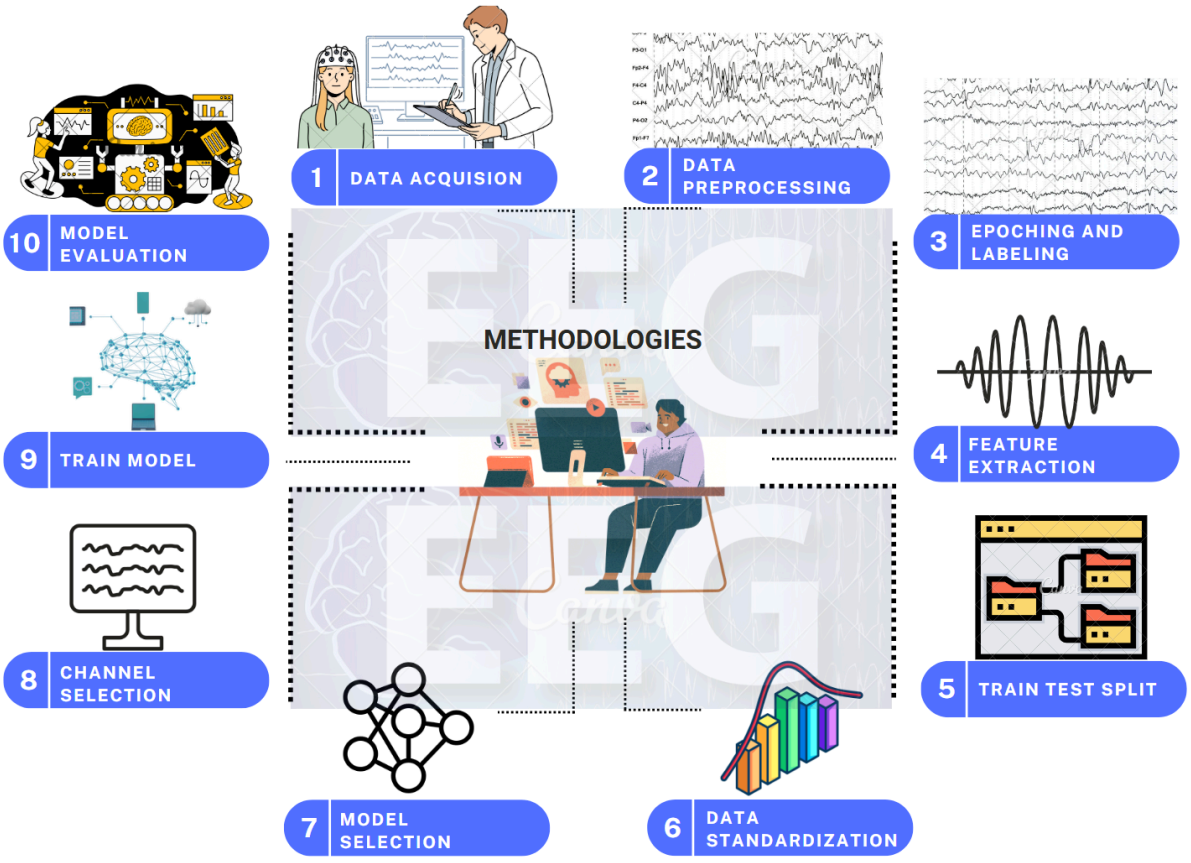


Figure 1: Workflow for detecting MDD.

In this study, we proposed a methodology to enhance the performance of the deep learning (DL) model for detecting Major Depressive Disorder (MDD) using a reduced set of EEG channels. The overall workflow of this study is illustrated in Figure 1. Initially, the EEG signals are denoised and preprocessed. To preserve temporal resolution while managing the computational complexity, 1s window size is applied. Various features are extracted and analyzed for feature selection. In this study, K-fold cross validation is performed to enhance the stability of the model. Before training, the data is standardized for consistency across all the features. Best performing layers are selected from 1D CNN and Bi-LSTM to optimize the proposed model for MDD detection. To reduce the experiment cost, single channel detection tests are implemented to the system. From that, the best channels are selected for better disease detection and management. Ultimately, Reduced channels and optimized model ensured the resource usage, cost effectiveness, and superior performance for MDD detection.

3.3 EEG Signal pre-preprocessing

To prepare the EEG data for analysis, several preprocessing steps were performed to enhance the quality and reliability of the signals. Initially, the raw data was loaded which was in European Data Format (.edf). Then the signal was interpolated to correct any electrode issues. Then, the data was filtered with FIR filtering between 0.5 Hz and 50Hz to remove noises and it's an effective range for filtering data suggested by previous studies on EEG data for MDD diagnosis. Also, we employed a notch filter of 50 Hz for noise reducing purposes. An average reference was set for all EEG channels to standardize the data. Furthermore, the standard 10-20 system montage was applied to ensure consistent brain region mapping across all participants and conditions and the final set of selected channels included: Fp1, F3, C3, P3, O1, F7, T3, T5, Fz, Fp2, F4, C4, P4, O2, F8, T4, T6, Cz, and Pz. Finally, Independent Component Analysis was then employed to underestimate different artifacts as there relies eye blink and heart artifact in the EEG data. With n-components of 19, ICA were applied and isolated the possible artifacts from the data. The cleaned signal then segmented into 1 second epochs corresponding to different experimental conditions (eyes closed, eyes open and task). Each epoch was labeled based on whether the participant was in the MDD group or healthy control group. These preprocessing steps ensured that the EEG data was of high quality and suitable for our further analysis of brain activity related to Major Depressive Disorder.

3.5 Feature Extraction

Extracting features from EEG signals is a crucial part in signal processing. As the raw data possess complexity, noise, and high dimensionality that affect the efficiency of the Deep learning process. By implementing feature extraction, relevant features can be acquired for detecting MDD with more efficiency. To select the best performing features for MDD detection, time domain, frequency domain and discrete wavelet Transformation (DWT) features are extracted. In this section, the feature extraction method applied in the study is discussed. In table 1 the extracted features are illustrated.

Table 1: Features considered in developing the MDD detection.

Features	Equation		Description
Time Domain			
Mean (μ)	$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$	(1)	x_i is the value, and N is the total number of value
Median			
Variance (var)	$var = \frac{1}{N}\sum_{i=1}^{N} (x_i - \mu)^2$	(2)	μ is the mean value.
Standard Deviation (σ)	$\sigma = \sqrt{var}$	(3)	

Features	Equation		Description
Time Domain			
Mean (μ)	$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$	(1)	x_i is the value, and N is the total number of value
Skewness (Skew)	$skew = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i-\mu}{\sigma}\right)$	(4)	
Kurtosis (Kurt)	$Kurt = \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i-\mu)^4}{(var)^2} - 3$		
Zero Crossing Rate (ZCR)	$ZCR = \frac{Number\ of\ zero\ crossings}{N}$		Number of times the signal crosses the zero line divided by total N number of points.
Wave Duration	$Wave\ Duration = \frac{Epoch\ Length}{Number\ of\ Peaks}$		
Peak Amplitude			
Hjorth Parameters	Activity, Mobility, Complexity		
Frequency Domain			
Mean PSD			
Variance PSD			
Standard Deviation PSD			The standard deviation is the square root of the variance of the PSD
Skewness			Skewness measures the asymmetry of the PSD distribution around its mean
PSD Kurtosis			Kurtosis quantifies how sharply peaked or flat the PSD distribution is compared to a normal distribution:
Band Power	$\text{Band Power}_\Theta = \int_4^7 PSD(f)\,df$		Delta (1-3 Hz), Theta (4-7 Hz), Alpha (8-12 Hz), Beta (13-30 Hz), Gamma (31-50 Hz)
Band Ratio			
DWT features			
Approximation Coefficients			
Detail Coefficients			
Entropy	$\text{Entropy} = -\sum_i p_i \log(p_i)$		pi is the probability associated with the occurrence of DWT coefficient..
percentiles			
mean values			
median			
Standard deviation			
variance			
root mean square	$RMS = \sqrt{\frac{1}{N}\sum_{i=1}^{N} C_i^2}$		Ci are the DWT coefficients.

Time Domain Feature:

Time-domain features capture the temporal characteristics of the signals by inspecting the signal's amplitude fluctuations over time. These features help to understand the dynamic changes of signals that occur in the brain, which helps to identify the crucial insights for temporal evolution of neural activity. Extracting the time-domain features is a computational efficient process which is suitable for a resource constrained environment. Table 1 summarizes the time domain features that were employed in this study. These features present an interportable understanding of the EEG signals, facilitating the effectiveness of MDD detection.

Frequency Domain feature:

Frequency-Domain features are essential in analyzing the power distribution and characteristics of the signal frequency. These features provide information about the frequency components of a signal by analyzing the amount of signal power distributed across different frequency bands. In case of MDD detection, extracted frequency domains are illustrated in table 1. Among those features, power spectral density (PSD) is used to identify patterns from the distribution of power across different frequency components of the EEG signals. The psd is particularly utilized for interpreting spectral and dominant frequencies that correspond in specific cognitive processes. PSD can be defined from equation (1).

$$S_{xx}(f) = \frac{1}{N}\left|\sum_{k=0}^{N-1} x(k)e^{-j2\pi fk}\right|^2 \quad (1)$$

Here, N is the number of data points,$x(k)$) is the signal value at time k,f is the frequency

Discrete wavelet Transformation (DWT):

DWT decomposes the signals into localized time-frequency coefficients. It captures both temporal and spectral information , which makes it suitable for detecting EEG signals. DWT applies wavelet functions to break down the signal into multilevel detail coefficients and approximation coefficients. In this experiment, we used level 5 decomposition of the wavelets. Here, The DWT of a discrete signal is described by the equation 2.

$$x[n] = \sum_{k=-\infty}^{\infty} c_{j_0}[k]\phi_{j_0}, k[n] + \sum_{j=j_0}^{\infty}\sum_{k=-\infty}^{\infty} d_j[k]\psi_j, k[n] \quad (2)$$

$$coefficients = \{cA_5, cD_5, cD_4, cD_3, cD_2, cD_1\}$$

where cA_5 represents the approximation coefficients at level 5, and cD_i represents the detail coefficients at level i. In table 1 extracted features from the DWT are given.

3.6 Deep Learning Model

3.6.1 OptMDD Architecture

In this study, the OptMDD model is proposed, which is a hybrid deep learning based architecture to detect MDD from EEG signals. The OptMDD fuses the 1D-Convolution Neural Network(1D CNN) and Bidirectional Long Short-Term Memory (Bi-LSTM) layers for diagnosing MDD in individuals. From the processed EEG signals the 1D CNN layers proficient spatial feature extraction and with the help of Bi-LSTM it can capture the long-term dependencies within the signal. This fusion of different layers is crucial for understanding the pattern from the MDD signals. The architecture of the optMDD model is elaborated in figure 2 and Table 2.

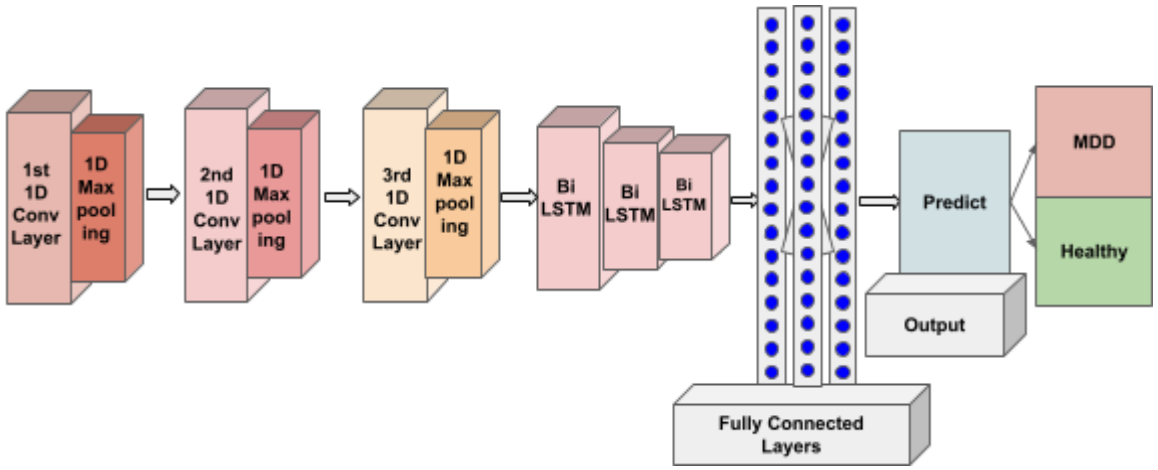


Figure 2: OptMDD Model

1D Convolutional Neural Network (1D CNN):

Convolving the input with a set of trainable filters 1D CNN processes one dimensional sequential data. The convolution layer facilitates the network to extract and learn from the local spatial features from the input signals. In equation (3) 1D CNN model is defined. Here, x_{i+k} is the input sequence, w_k is the filter weight for kernel K, b is the bias term, $K$is the kernel size (length of the filter), y_i is the output feature map at position i .

$$y_i = \sum_{k=0}^{K-1} w_k \cdot x_{i+k} + b \tag{3}$$

In the model architecture three 1D CNN layers are employed, where the layers consist of 64, 128, and 256 filters, respectively with a kernel size of 5. For all the layers, a rectified linear unit (Relu) activation function is used, followed by a Batch Normalization layer and Max Pooling layer. To justify the performance of the model and reduce the overfitting 30 % of neurons are randomly deactivated using a dropout layer.

Bidirectional Long Short-Term Memory (Bi-LSTM)

Bi-LSTM is an advanced Recurrent Neural Network (RNN) that extends the traditional LSTM model by processing the input data both forward and backward. It allows the model to learn the past and future subsequent time steps, thereby providing a more constrictive understanding of the temporal dynamics of data.This characteristics enhances the effectiveness of the model to detect EEG signals. In Bi-LSTM it runs two parallel LSTM layers, where one layer runs forward and the other runs backward.From each time step the output is generated by concating the hidden states of both directions. This process is depicted in figure 3. The LSTM layers are designed to handle long-term dependency which is best suited for retrieving information over long sequences of data. The LSTM units regulate the flow of data through the network, selectively remembering and forgetting particular information over time. The key operations in an LSTM layer are described as follows: here, forget gate f_t , input gate i_t, Cell State Update $\widehat{C}_t$, and output gate o_t. w is the weight matrix, C_t is the cell state, and h_t is the hidden state. σ represents the sigmoid activation function.

$$f_t = \sigma(w_f . x_t + U_f . h_{t-1} + b_f)$$

$$i_t = \sigma(w_i \cdot x_t + U_i \cdot h_{t-1} + b_i)$$

$$C_t = tanh(w_C \cdot x_t + U_C \cdot h_{t-1} + b_C)$$

$$\widehat{C_t} = f_t \circ \widehat{C_{t-1}} + i_t \circ C_t$$

$$o_t = \sigma(w_o \cdot x_t + U_o \cdot h_{t-1} + b_o)$$

$$h_t = o_t \circ tanh(C_t)$$

In our OptMDD architecture, after the 1D CNN three Bi-LSTM layers of 256 filter size are used for enabling the model to capture comprehensive information from the EEG signal to detect MDD.

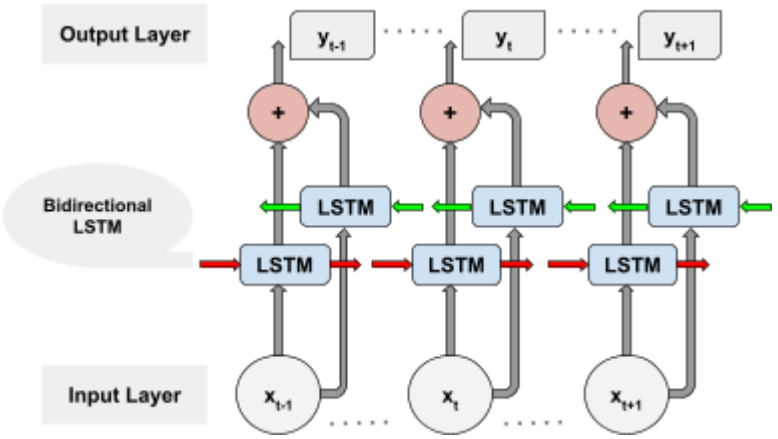


Figure 3: Structural representation of Bidirectional Long-Short term Memory

Dense Layers

The output of 1D-CNN-Bi-LSTM model is further refined by passing through Dense layers with 256,128, 64 neurons, each utilizing the Relu activation function. Finally, the sigmoid activation function is used for the single neuron in the output dense layer, which outputs the probability score for the MDD detection.

Table 2: Model architecture of OptMDD

Layer	Parameters
Conv1D	Filter = 64, kernel size = 5, Activation = ReLU, padding='same', input_shape = input_shape
BatchNormalization	-
MaxPooling1D	Pool size = 2
Dropout	Rate = 0.3
Conv1D	Filter = 128, kernel size = 5, Activation = ReLU, padding='same'
BatchNormalization	-
MaxPooling1D	Pool size = 2

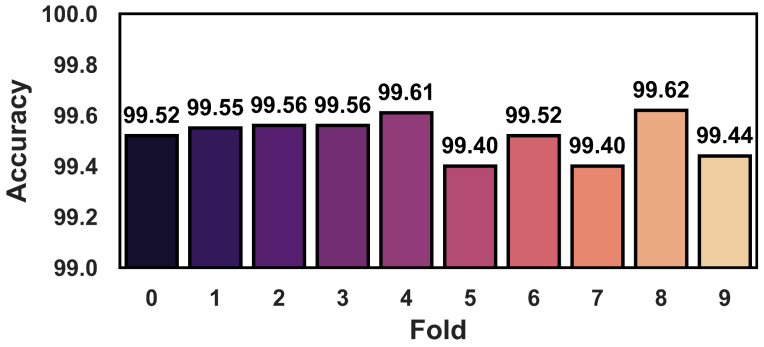
Dropout	Rate = 0.3
Conv1D	Filter = 256, kernel size = 5, Activation = ReLU, padding='same'
BatchNormalization	-
MaxPooling1D	Pool size = 2
Bidirectional LSTM	Units = 256, return_sequences=True
Bidirectional LSTM	Units = 256, return_sequences=True
Bidirectional LSTM	Units = 256
Dense	Units = 256, Activation = ReLU
Dense	Units = 128, Activation = ReLU
Dense	Units = 64, Activation = ReLU
Output	Units = 1, Activation = Sigmoid

Table 3: Configuration of the Hyperperameters

Configuration	**Value**
Learning Rate	0.001
Batch size	256
epoch	100
Loss Function	Binary Crossentropy
Optimizer	Adam

3.7 Cross Validation

For training the model we have employed a 10-fold StratifiedKfold cross validation method where the data was splitted 80-20 among which 80% data was used for training and the rest 20% was used to validate the model. This method ensures that all groups of data are evenly distributed.



3.8 Evaluation Metrics

Evaluating the model is a crucial step in detecting the MDD. From the metrics important insights can be achieved about the model's performance. Evaluation metrics that have been used are Accuracy, Area Under the Curve - Receiver Operating Characteristic (AUC-ROC), Precision, F1-score, Specificity, Sensitivity, and Cohen's Kappa. Here, TP = True Positives, TN = True Negatives, FP = False Positives, FN = False Negatives.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

$$Precision = \frac{TP}{TP+FP}$$

$$F1\ Score = \frac{2\times Tp}{2\times TP+Fp+FN}$$

$$Specificity = \frac{TN}{TN+FP}$$

$$Sensitivity = \frac{TP}{TP+FN}$$

$$Cohen's\ Kappa = \frac{p_o - p_e}{1 - p_e}$$

Here, p_o is the observed agreement, calculated as $p_o = \frac{TP+TN}{TP+TN+FP+FN}$

And $p_e = \left(\frac{TP+FP}{TP+TN+FP+FN} \times \frac{TP+FN}{TP+TN+FP+FN}\right) + \left(\frac{TN+FN}{TP+TN+FP+FN} \times \frac{TN+FP}{TP+TN+FP+FN}\right)$

$$\text{MCC} = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

Result:

Model performance on different features:

To develop a robust and effective method to classify MDD, this study implemented a comprehensive evaluation for the proposed OptMDD model. We have tested our model on DWT, Frequency Domain, Time Domain and combined (Time+Frequency Domain) feature extraction methods. 10 fold cross validation was performed for more accurate analysis of the extracted features performance using OptMDD. As different activities (Eye open, Eye Close, and Task) were involved while acquiring the dataset, features were extracted from them individually, and all of the activity combined. The averaged 10 fold result of our proposed model on features extracted from different activities is shown in table 4.1. Focusing on the MDD classification task, our proposed model performed outstanding using the time domain features. It achieved the accuracy of 99.32%, 99.22%, 99.34% and 99.14% using eye closed, eye open, task and combined (eye closed, eye open, and task) dataset respectively.

Table 4.1:OptMDD model performance comparison for different features and data conditions.

Features	**Conditions**	**Acc**	**Prec**	**F1**	**Auc Roc**	**Spec**	**Sens**	**Kap**	**Mcc**
DWT	Eyes Close	93.3	90.57	93.63	93.25	89.57	96.93	86.59	86.84
	Eyes Open	97.45	97.32	97.56	97.43	97.05	97.81	94.89	94.89
	Task	97.77	97.12	97.99	97.65	96.41	98.88	95.48	95.51
	Combined	95.69	94.59	96.02	95.56	93.63	97.5	91.33	91.38
Frequency Domain	Eyes Close	90.87	91.53	90.94	90.88	91.37	90.4	81.75	81.78
	Eyes Open	94.63	95.00	94.85	94.63	94.55	94.71	89.24	89.26
	Task	96.82	96.75	97.12	96.75	96.00	97.5	93.58	93.58
	Combined	95.71	95.91	95.97	95.68	95.31	96.05	91.37	91.38
Time Domain	**Eyes Close**	**98.32**	**98.47**	**98.34**	**98.32**	**98.43**	**98.21**	**96.64**	**96.64**
	Eyes Open	**99.22**	**99.38**	**99.25**	**99.22**	**99.32**	**99.12**	**98.43**	**98.43**
	Task	**99.34**	**99.45**	**99.4**	**99.33**	**99.33**	**99.34**	**98.66**	**98.66**
	Combined	**99.14**	**99.29**	**99.2**	**99.15**	**99.19**	**99.11**	**98.28**	**98.28**
Combined	Eyes Close	96.82	96.63	96.88	96.82	96.5	97.13	93.64	93.65
	Eyes Open	98.52	98.3	98.59	98.51	98.12	98.89	97.04	97.05
	Task	98.84	98.83	98.95	98.82	98.57	99.07	97.66	97.66
	Combined	98.69	98.74	98.77	98.68	98.56	98.81	97.37	97.38

Model optimization and layer selection:

Using the time domain feature on the combined data, our proposed model achieved an outstanding performance. To ensure the best performance of our model, we compared different layer based models performance that enabled us to analyze the optimal layer to use in our model. The comparison of the different layers is given in figure 3 and table 4.3 . We have analyzed a total 5 Conv1D layers, but 99.46% highest accuracy is achieved by utilizing 3 layers of the 1D Conv layer. Among different bi-LSTM layers, using 3 bi-Lstm layers gives the highest accuracy of 99.51%.

Fig 3: Different number of layer performance comparison, where A is 1D Conv layers and B is the LSTM layer comparison.

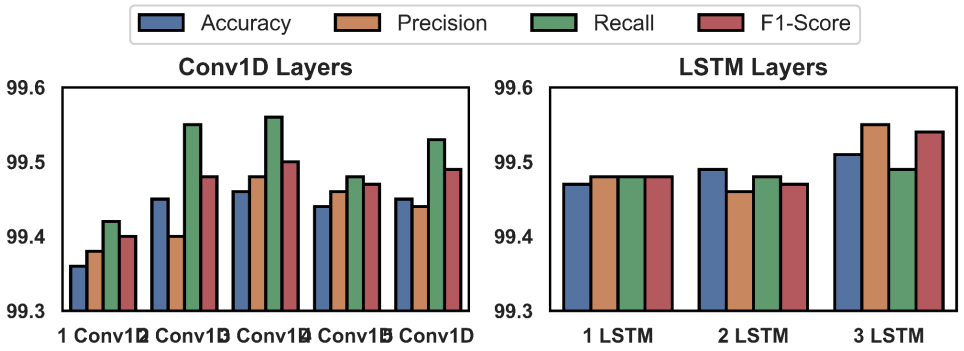


Table 4.2: performance metrics of 1D Conv and LSTM layer. For better identification of the MDD

Layer	Number of layers	Acc	Prec	Auc Roc	F1	Spec	Sens	Kap	Mcc
1D Conv	1	99.36	99.38	99.42	99.4	99.29	99.36	98.72	98.72
	2	99.45	99.40	99.56	99.48	99.32	99.44	98.89	98.89
	3	**99.46**	**99.48**	99.55	**99.50**	**99.40**	**99.46**	**98.92**	**98.92**
	4	99.44	99.46	99.48	99.47	99.39	99.43	98.87	98.87
	5	99.45	99.44	99.53	99.49	99.39	99.44	98.91	98.91
LSTM	1	99.47	99.48	99.45	99.48	99.41	99.45	98.92	98.92
	2	99.49	99.46	99.48	99.47	99.39	99.43	98.87	98.87
	3	**99.51**	**99.55**	99.39	**99.54**	**99.45**	**98.91**	**98.72**	**98.72**

pattern, the 1D Conv layers and the Bi-LSTM layers are combined. Different combinations of the layer are performed to find the optimal layer combination. However, after comparing different layers combinations for our OptMDD model in figure 4, we can observe that three 1D Conv layers with three Bi-LSTM layers outperformed the other fusion of layers. From Table 4.3 it can be concluded that the three layers combination of 1D Conv and Bi-LSTM more suitable choice for our proposed model. As we applied 10 fold cross validation to validate the evaluations of our model. In figure 5, the accuracy of the 10 folds for our optimal layers is given. The average accuracy of the fold is 99.52%. It solidifies the stability of the layer selection for our proposed model.

Figure 4: F1 Score comparison of different 1D conventional and LSTM layers

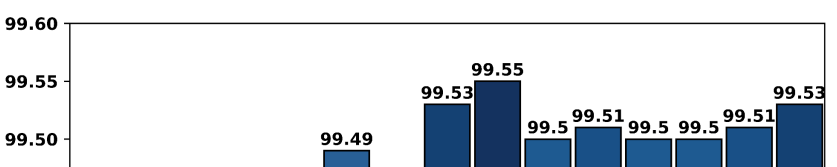
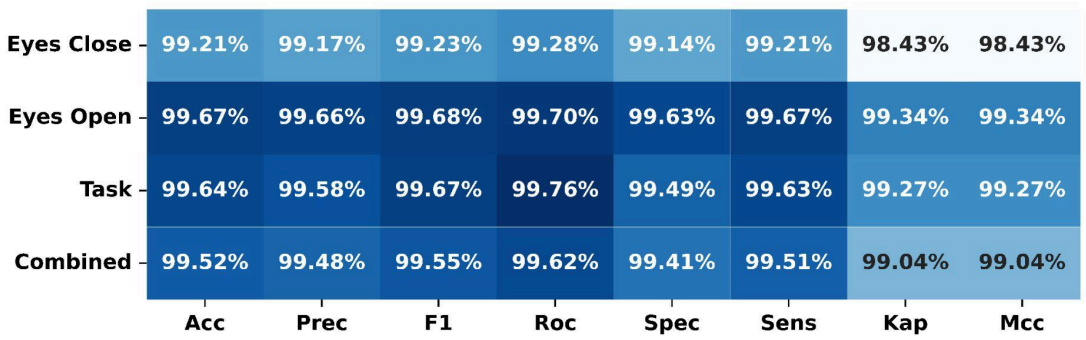


Table 4.3: Combination of layers performance comparison using combined data for OptMDD model.

Layers	Acc	Prec	Auc Roc	F1	Spec	Sens	Kap	Mcc
1 Conv1D + 1 LSTM	99.26	99.22	99.39	99.30	99.11	99.25	98.51	98.51
2 Conv1D + 1 LSTM	99.40	99.35	99.52	99.43	99.26	99.39	98.79	98.79
3 Conv1D + 1 LSTM	99.42	99.44	99.47	99.46	99.37	99.42	98.84	98.84
4 Conv1D + 1 LSTM	99.47	99.50	99.47	99.50	99.43	99.47	98.94	98.94
5 Conv1D + 1 LSTM	99.47	99.46	99.54	99.50	99.39	99.46	98.93	98.93
1 Conv1D + 2 LSTM	99.27	99.23	99.4	99.31	99.12	99.26	98.52	98.52
2 Conv1D + 2 LSTM	99.41	99.36	99.53	99.44	99.27	99.4	98.8	98.8
3 Conv1D + 2 LSTM	99.49	99.51	99.54	99.53	99.44	99.49	98.91	98.91
4 Conv1D + 2 LSTM	99.48	99.51	99.48	99.51	99.44	99.48	98.95	98.95
5 Conv1D + 2 LSTM	99.48	99.47	99.55	99.51	99.4	99.47	98.94	98.94
1 Conv1D + 3 LSTM	99.37	99.33	99.5	99.41	99.22	99.36	98.62	98.62
2 Conv1D + 3 LSTM	99.46	99.41	99.58	99.49	99.32	99.45	98.85	98.85
3 Conv1D + 3 LSTM	**99.52**	**99.48**	**99.62**	**99.55**	**99.41**	**99.51**	**99.04**	**99.04**
4 Conv1D + 3 LSTM	99.47	99.5	99.47	99.5	99.43	99.47	98.94	98.94
5 Conv1D + 3 LSTM	99.50	99.47	99.61	99.54	99.4	99.5	99.03	99.03

Table 4.4: 3*3 1D-CNN-BiLSTM based OptMDD model performance on different test conditions using Time Domain Features.

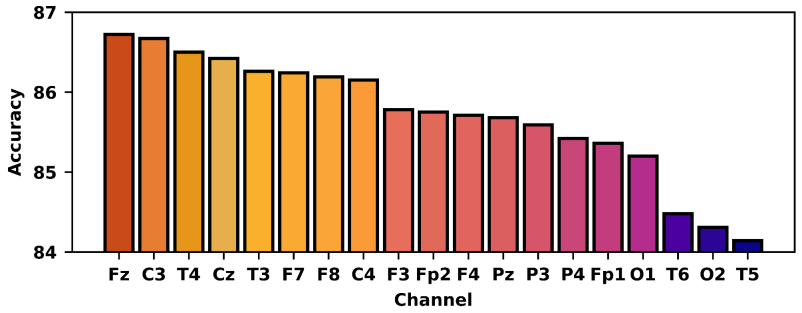


EEG Channel Optimization

In this study, the total number of the EEG channel is 19. But to analyze and identify the most responsible channels for MDD detection, we implemented a single channel testing method. Data was extracted for each of the channels. Then the OptMDD model is tested on channel wise extracted data. Our model's channel wise accuracy in figure 6, different in the accuracy of the channel can be observed from it. This is a crucial part where the important channels are identified. Channel wise

performance metrics are given in table 4.4, where the Fz channel achieved the highest accuracy of 86.72%, and F1 Score of 88.24%. Based on the performance in table 4.6. four, and six channels are selected for optimal channel performance. The 10 features are selected considering the high activity in the left frontal part and back of the brain in MDD.

Figure 6: Channel-wise Accuracy performance of the OptMDD model for detecting MDD.



In this study, the OptMDD model is trained on different channels (four, six, ten and all channels) to identify MDD at a lower data acquisition cost using EEG signal. From table 4.7 and figure 7 the comparison of different channel sets can be observed, where our proposed model achieved the accuracy of 99.52 %, Auc roc of 99.62%, 99.48% precision, and F1 score 99.55% using the combined dataset on time domain features. Utilizing the 10 channels this model achieved the accuracy of 98.58% ,Auc roc 98.81%, precision 99.53%, and F1 score 98.67%. As we reduced the channels using our channel wise performance based selection method, the accuracy slightly decreased. Using the 6 channel set our model achieved 97.54% accuracy, 97.99% Auc roc, 97.41% precision, and 97.7% F1 score. Finally, using only the most responsible 4 channels the our model achieved the accuracy of 95.73%, Auc roc 96.74%, precision of 95.45, and F1 Score of 96.08%. As the channels reduced the performance dropped but still a robust performance has been achieved by our model using redacted channels.

Table: 4.5: EEG channel wise performance metrics of the OptMDD Model.

Channels	**Acc**	**Prec**	Roc	**F1**	**Spec**	**Sens**	**Kap**	**Mcc**
C3	86.67	84.45	91.44	87.81	80.79	86.12	72.66	72.95
C4	86.15	83.32	92.52	87.68	78.87	85.7	71.95	72.47
Cz	86.42	83.42	93.01	87.95	78.90	85.95	72.50	73.07
F3	85.78	84.74	89.43	87.01	81.61	85.52	71.32	71.46
F4	85.71	83.53	91.17	87.18	79.49	85.33	71.12	71.47
F7	86.24	84.03	91.59	87.64	80.14	85.86	72.18	72.53
F8	86.19	84.65	90.50	87.48	81.27	85.89	71.12	72.33
Fp1	85.36	83.64	90.18	86.78	79.85	85.02	70.42	70.68
Fp2	85.75	83.74	90.95	87.19	79.83	85.39	71.21	71.54
Fz	**86.72**	83.50	93.56	88.24	78.91	86.23	73.09	73.72
O1	85.20	85.51	92.55	85.20	76.30	84.42	69.49	70.25
O2	84.31	80.82	92.54	86.28	74.93	83.74	68.17	69.03

P3	85.59	82.34	92.90	87.30	77.26	85.08	70.80	71.49
P4	85.42	82.57	92.09	87.06	77.80	84.95	70.46	71.03
Pz	85.68	82.35	93.09	87.39	77.23	85.16	70.97	71.69
T3	86.26	82.71	93.85	87.93	77.60	85.72	72.14	72.92
T4	86.50	84.41	91.59	87.85	87.85	91.59	80.69	86.14
T5	84.14	80.55	92.59	86.15	74.48	83.54	67.79	68.71
T6	84.48	80.56	93.42	86.51	74.27	83.84	68.46	69.51

Table 4.6: Optimal channel selection for detecting MDD

Channel	**Channel Names**
4 Channels	C3, Cz, Fz, T4
6 Channels	C3, Cz, Fz, T4, F7, T3
10 Channels	Fz, T3, F7, F8, C4, F3, Fp1, O2, T6, Pz

Table 4.7: Multiple Channels Performance on Combined Dataset

Channel	Channel Names	Acc	Prec	F1	Auc Roc	Spec	Sens	Kap	Mcc
4 Channels	(C3, Cz, Fz, T4)	95.79	95.43	96.08	96.74	94.71	95.73	91.54	91.55
6 Channels	(C3, Cz, Fz, T4, F7, T3)	97.54	97.41	97.70	97.99	97.02	97.51	95.05	95.05
10 Channels	(Fp1, F3, C4, F7, T3, Fz, O2, F8, T6, Pz)	98.58	99.53	98.67	98.81	98.32	98.56	97.15	97.15
All Channels	All	99.52	99.48	99.55	99.62	99.41	99.51	99.04	99.04

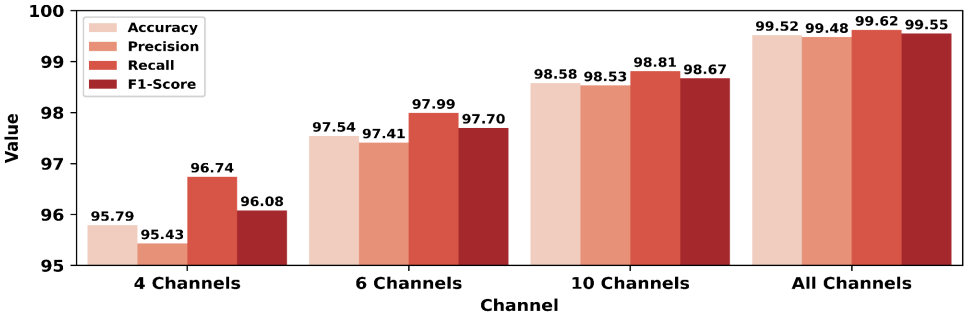


Figure 7: Evaluation metrics for different sets of EEG channels.

Discussion:

In this study we proposed OptMDD architecture combining 1D-CNN and Bi-LSTM models for effective and more efficient MDD detection, utilizing a robust feature extraction method with DL based Channel selection strategies. Although there are several studies using deep learning techniques implementing different features, this is the work that presents a hybrid deep learning model with promising efficiency in diagnosing MDD using EEG signals in all conditions of the patient. For better performance, we filtered the data including notch filtering and performed ICA on the data to remove artifacts and segment the data with 1 second epoch duration and standardized the data after labeling. Then, we extracted time-domain features, frequency-domain features and discrete wavelet transform (DWT) features from the data. First of all we directly developed a hybrid CNN-LSTM model to check the efficiency of the hybrid model and find out which features we would need to focus on. For the validation process we used 10-fold stratified cross validation with 80-20 data partitioning. Following table-4.1, it is clear that our model achieved in comparison with time-domain features poor results in dwt features with accuracy 95.69% in combined all conditions, 93.3% in eyes-close, 97.45% in eyes open and 97.77% in task state. Similarly in the frequency domain our model achieved 95.71% accuracy in combined dataset, 90.87% in eyes-close state, 94.63% in eyes-open state and 95.71% accuracy in task state. But we got our acceptable performance in time-domain features which were 99.14% accuracy in the dataset combining all conditions, 98.32% in eyes-close data, 99.22% in eyes-open condition and 99.34% in task condition. As the model showed off its effectiveness in time domain features so we just focused on that feature and worked on it only. As our first model performed over 99% accuracy in detecting depression, we next moved forward to optimize the model with a layer and model selection process to enhance the performance of this model. Therefore, We tried several convolutional and Bi-LSTM layers to find out the best layers for building a robust hybrid model and from figure-3 we found that 3-Convo layers achieved 99.46% accuracy and 3-LSTM layers got 99.51% accuracy which were the best layers so far among all other convo and Bi-LSTM layers. Although these individual models performed better, we were after haunting a robust model that would perform best in all conditions and . So, after selecting layers, we built several hybrid models with different numbers of CNN-LSTM layers combination and we compared the classification performances between each hybrid model we built to identify the outstanding performing one for our proposed framework. From table-4.3, it is evident that our final model achieved 99.52% accuracy along with 99.41% specificity and 99.51% sensitivity in combined conditions. The Figure-4 shows the comparison between the supreme model and other models we tried. Although the model performed comparatively lower in eyes close state acquiring 99.21% accuracy, according to the Table 4.4, the model performed best in data of eyes open state achieving accuracy 99.67% and 99.64% in task condition indicating the model's detectability in various situations of the patient. Furthermore we have employed an evaluation based channel selection procedure using combined dataset to minimize channels and the best performing channels were selected through this approach presented in the table-4.5. 4 Channels C3, Cz, Fz, T4 and 6 channels C3, Cz, Fz, T4, F7, T3 were selected through our evaluation based selection from figure-6. But, 10 channels we selected in our study were used by Rafiei et al.

[21] and they were Fz, T3, F7, F8, C4, F3, Fp1, O2, T6, Pz. The table-4.7 shows the model yielded 95.79% accuracy using combined data from 4 channels, 97.54% accuracy from 6 channels and 98.58% accuracy from10 channels. The figure-7 illustrates the comparison of the performance improvement throughout different sets of channels.

In table-5.1 we compared the performance of our proposed framework to the existing methods in previous studies for major depressive disorder classification using EEG signals.

Table-5.1: Comparison with Previous studies worked with different features

Study	Model	Features	Accuracy	Sensitivity	Specificity
Wang et al. [13]	1D-CNN Transformer	Spatial-Temporal Features	96.7%	92.4%	95.0%
Garg and Shukla et al. [34]	Weighted Spectral Clustering	Band Power, DFA, HFD	98%	-	-
Khan et al. [35]	2D-CNN	Wavelet	98.1%	98.0%	98.2%
Movahed et al. [36]	RBFSVM	Statistical, Spectral, Wavelet, Nonlinear, Functional	99%	98.4%	99.6%
Raghavendra et al. [37]	SVM	Wavelet (CWT, DWT, EWT)	99.33%	98.95%	99.70%
Our Study	1DCNN-LSTM	Time Domain, Frequency Domain, DWT	99.67%	99.67%	99.63%

REFERENCES

[1] L. Cui *et al.*, “Major depressive disorder: hypothesis, mechanism, prevention and treatment,” *Signal Transduct. Target. Ther.*, vol. 9, no. 1, p. 30, Feb. 2024, doi: 10.1038/s41392-024-01738-y.

[2] B. Wang *et al.*, “Characterizing Major Depressive Disorder (MDD) using alpha-band activity in resting-state electroencephalogram (EEG) combined with MATRICS Consensus Cognitive Battery (MCCB),” *J. Affect. Disord.*, vol. 355, pp. 254–264, Jun.

2024, doi: 10.1016/j.jad.2024.03.145.
[3] W. Marx *et al.*, “Major depressive disorder,” *Nat. Rev. Dis. Primer*, vol. 9, no. 1, p. 44, Aug. 2023, doi: 10.1038/s41572-023-00454-1.
[4] L. Cosgrove *et al.*, “Global Burden Disease Estimates for Major Depressive Disorders (MDD): A review of diagnostic instruments used in studies of prevalence,” *Community Ment. Health J.*, Jul. 2024, doi: 10.1007/s10597-024-01302-6.
[5] H. Qi *et al.*, “Identifying influencing factors of metabolic syndrome in patients with major depressive disorder: A real-world study with Bayesian network modeling,” *J. Affect. Disord.*, vol. 362, pp. 308–316, Oct. 2024, doi: 10.1016/j.jad.2024.07.004.
[6] C. S. H. Ho *et al.*, “Interpretable deep learning model for major depressive disorder assessment based on functional near-infrared spectroscopy,” *Asian J. Psychiatry*, vol. 92, p. 103901, Feb. 2024, doi: 10.1016/j.ajp.2023.103901.
[7] P. Greenberg *et al.*, “The Economic Burden of Adults with Major Depressive Disorder in the United States (2019),” *Adv. Ther.*, vol. 40, no. 10, pp. 4460–4479, Oct. 2023, doi: 10.1007/s12325-023-02622-x.
[8] Q. Xiao, F. Li, F. Jiang, Z. Zhang, and B. Xu, “The prospects for early detection with optical coherence tomography (OCT) and OCT angiography in major depressive disorder,” *J. Affect. Disord.*, vol. 347, pp. 8–14, Feb. 2024, doi: 10.1016/j.jad.2023.11.031.
[9] J. Sarlon *et al.*, “Adjunctive use of mindfulness-based mobile application in depression: randomized controlled study,” *Eur. Arch. Psychiatry Clin. Neurosci.*, Sep. 2024, doi: 10.1007/s00406-024-01884-y.
[10] X. Liu *et al.*, “EEG-based major depressive disorder recognition by neural oscillation and asymmetry,” *Front. Neurosci.*, vol. 18, p. 1362111, Feb. 2024, doi: 10.3389/fnins.2024.1362111.
[11] S. Yasin, S. A. Hussain, S. Aslan, I. Raza, M. Muzammel, and A. Othmani, “EEG based Major Depressive disorder and Bipolar disorder detection using Neural Networks:A review,” *Comput. Methods Programs Biomed.*, vol. 202, p. 106007, Apr. 2021, doi: 10.1016/j.cmpb.2021.106007.
[12] N. Van Der Vinne, M. A. Vollebregt, A. J. Rush, M. Eebes, M. J. A. M. Van Putten, and M. Arns, “EEG biomarker informed prescription of antidepressants in MDD: a feasibility trial,” *Eur. Neuropsychopharmacol.*, vol. 44, pp. 14–22, Mar. 2021, doi: 10.1016/j.euroneuro.2020.12.005.
[13] Y. Wang *et al.*, “DiffMDD: A Diffusion-Based Deep Learning Framework for MDD Diagnosis Using EEG,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 32, pp. 728–738, 2024, doi: 10.1109/TNSRE.2024.3360465.
[14] M. Ravan *et al.*, “Diagnostic deep learning algorithms that use resting EEG to distinguish major depressive disorder, bipolar disorder, and schizophrenia from each other and from healthy volunteers,” *J. Affect. Disord.*, vol. 346, pp. 285–298, Feb. 2024, doi: 10.1016/j.jad.2023.11.017.
[15] A. Habib, S. N. Vaniya, A. Khandoker, and C. Karmakar, “MDDBranchNet: A Deep Learning Model for Detecting Major Depressive Disorder Using ECG Signal,” *IEEE J. Biomed. Health Inform.*, vol. 28, no. 7, pp. 3798–3809, Jul. 2024, doi: 10.1109/JBHI.2024.3390847.
[16] C. Gu, T. Chou, A. S. Widge, and D. D. Dougherty, “EEG complexity in emotion conflict task in individuals with psychiatric disorders,” *Behav. Brain Res.*, vol. 467, p. 114997, Jun. 2024, doi: 10.1016/j.bbr.2024.114997.
[17] Y. Li *et al.*, “A novel EEG-based major depressive disorder detection framework with two-stage feature selection,” *BMC Med. Inform. Decis. Mak.*, vol. 22, no. 1, p. 209, Aug. 2022, doi: 10.1186/s12911-022-01956-w.
[18] L. Duan *et al.*, “Machine Learning Approaches for MDD Detection and Emotion Decoding Using EEG Signals,” *Front. Hum. Neurosci.*, vol. 14, p. 284, Sep. 2020, doi: 10.3389/fnhum.2020.00284.
[19] G. Sharma, A. M. Joshi, and E. S. Pilli, “An Automated MDD Detection System based on Machine Learning Methods in Smart Connected Healthcare,” in *2021 IEEE International*

Symposium on Smart Electronic Systems (iSES), Jaipur, India: IEEE, Dec. 2021, pp. 27–32. doi: 10.1109/iSES52644.2021.00019.
[20] N. Bashir *et al.*, “A Machine Learning Framework for Major Depressive Disorder (MDD) Detection Using Non-invasive EEG Signals,” *Wirel. Pers. Commun.*, May 2023, doi: 10.1007/s11277-023-10445-w.
[21] A. Rafiei, R. Zahedifar, C. Sitaula, and F. Marzbanrad, “Automated Detection of Major Depressive Disorder With EEG Signals: A Time Series Classification Using Deep Learning,” *IEEE Access*, vol. 10, pp. 73804–73817, 2022, doi: 10.1109/ACCESS.2022.3190502.
[22] Z. Ning, H. Hu, L. Yi, Z. Qie, A. Tolba, and X. Wang, “A Depression Detection Auxiliary Decision System Based on Multi-Modal Feature-Level Fusion of EEG and Speech,” *IEEE Trans. Consum. Electron.*, vol. 70, no. 1, pp. 3392–3402, Feb. 2024, doi: 10.1109/TCE.2024.3370310.
[23] Z. Zhang, Q. Meng, L. Jin, H. Wang, and H. Hou, “A novel EEG-based graph convolution network for depression detection: Incorporating secondary subject partitioning and attention mechanism,” *Expert Syst. Appl.*, vol. 239, p. 122356, Apr. 2024, doi: 10.1016/j.eswa.2023.122356.
[24] A. Ksibi, M. Zakariah, L. J. Menzli, O. Saidani, L. Almuqren, and R. A. M. Hanafieh, “Electroencephalography-Based Depression Detection Using Multiple Machine Learning Techniques,” *Diagnostics*, vol. 13, no. 10, p. 1779, May 2023, doi: 10.3390/diagnostics13101779.
[25] Y. Wang *et al.*, “GCTNet: a graph convolutional transformer network for major depressive disorder detection based on EEG signals,” *J. Neural Eng.*, vol. 21, no. 3, p. 036042, Jun. 2024, doi: 10.1088/1741-2552/ad5048.
[26] I. A. Anik, A. H. M. Kamal, M. A. Kabir, S. Uddin, and M. A. Moni, “A Robust Deep-Learning Model to Detect Major Depressive Disorder Utilising EEG Signals,” *IEEE Trans. Artif. Intell.*, pp. 1–10, 2024, doi: 10.1109/TAI.2024.3394792.
[27] S. Mahato and S. Paul, “Analysis of region of interest (RoI) of brain for detection of depression using EEG signal,” *Multimed. Tools Appl.*, vol. 83, no. 1, pp. 763–786, Jan. 2024, doi: 10.1007/s11042-023-15827-7.
[28] M. Robnik-Šikonja and I. Kononenko, “Theoretical and Empirical Analysis of ReliefF and RReliefF,” *Mach. Learn.*, vol. 53, no. 1/2, pp. 23–69, 2003, doi: 10.1023/A:1025667309714.
[29] S. Choudhary, M. K. Bajpai, and K. K. Bharti, “Electrode subset selection to lessen the complexity of brain activity measurement using EEG for depression detection,” *Trans. Inst. Meas. Control*, p. 01423312241263140, Jul. 2024, doi: 10.1177/01423312241263140.
[30] M. Ying, X. Shao, J. Zhu, Q. Zhao, X. Li, and B. Hu, “EDT: An EEG-based attention model for feature learning and depression recognition,” *Biomed. Signal Process. Control*, vol. 93, p. 106182, Jul. 2024, doi: 10.1016/j.bspc.2024.106182.
[31] W. Dang, Z. Gao, X. Sun, R. Li, Q. Cai, and C. Grebogi, “Multilayer brain network combined with deep convolutional neural network for detecting major depressive disorder,” *Nonlinear Dyn.*, vol. 102, no. 2, pp. 667–677, Oct. 2020, doi: 10.1007/s11071-020-05665-9.
[32] Y. Xia, L. Liu, T. Dong, J. Chen, Y. Cheng, and L. Tang, “A depression detection model based on multimodal graph neural network,” *Multimed. Tools Appl.*, vol. 83, no. 23, pp. 63379–63395, Jan. 2024, doi: 10.1007/s11042-023-18079-7.
[33] Y. Chen, S. Wang, and J. Guo, “DCTNet: hybrid deep neural network-based EEG signal for detecting depression,” *Multimed. Tools Appl.*, vol. 82, no. 26, pp. 41307–41321, Nov. 2023, doi: 10.1007/s11042-023-14799-y.
[34] S. Garg and U. P. Shukla, “A Multi-channel EEG Biomarker Weighted Spectral Clustering Model for MDD Prediction,” Apr. 27, 2022. doi: 10.36227/techrxiv.19654026.v1.
[35] D. M. Khan, K. Masroor, M. F. M. Jailani, N. Yahya, M. Z. Yusoff, and S. M. Khan, “Development of Wavelet Coherence EEG as a Biomarker for Diagnosis of Major

Depressive Disorder," *IEEE Sens. J.*, vol. 22, no. 5, pp. 4315–4325, Mar. 2022, doi: 10.1109/JSEN.2022.3143176.
[36] R. A. Movahed, G. P. Jahromi, S. Shahyad, and G. H. Meftahi, "A major depressive disorder classification framework based on EEG signals using statistical, spectral, wavelet, functional connectivity, and nonlinear analysis," *J. Neurosci. Methods*, vol. 358, p. 109209, Jul. 2021, doi: 10.1016/j.jneumeth.2021.109209.
[37] U. Raghavendra *et al.*, "Automated detection and screening of depression using continuous wavelet transform with electroencephalogram signals," *Expert Syst.*, vol. 40, no. 4, p. e12803, May 2023, doi: 10.1111/exsy.12803.

[1] Wang, Y., Zhao, S., Jiang, H., Li, S., Luo, B., Li, T., & Pan, G. (2024). DiffMDD: A Diffusion-Based Deep Learning Framework for MDD Diagnosis Using EEG. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 32, 728-738.(1)

[2] Anik, I. A., Kamal, A. H. M., Kabir, M. A., Uddin, S., & Moni, M. A. (Accepted/In press). A Robust Deep-Learning Model to Detect Major Depressive Disorder Utilizing EEG Signals. IEEE Transactions on Artificial Intelligence, 1-10. https://doi.org/10.1109/TAI.2024.3394792

[3] Sun X, Xu Y, Zhao Y, Zheng X, Zheng Y, Cui L. Multi-Granularity Graph Convolution Network for Major Depressive Disorder Recognition. IEEE Trans Neural Syst Rehabil Eng. 2024;32:559-569. doi: 10.1109/TNSRE.2023.3311458. Epub 2024 Feb 2. PMID: 37665697.

[4] Xuexiao Shao, Ming Ying, Jing Zhu, Xiaowei Li, Bin Hu,"Achieving EEG-based depression recognition using Decentralized-Centralized structure",Biomedical Signal Processing and Control, Volume 95, Part A, 2024, 106402, ISSN 1746-8094, https://doi.org/10.1016/j.bspc.2024.106402.

[5] Chen, Y., Wang, S. & Guo, J. DCTNet: hybrid deep neural network-based EEG signal for detecting depression. Multimed Tools Appl 82, 41307–41321 (2023). https://doi.org/10.1007/s11042-023-14799-y

[6] Shreeya Garg , Urvashi Prakash Shukla . A Multi-channel EEG Biomarker Weighted Spectral Clustering Model for MDD Prediction. TechRxiv. April 27, 2022. DOI: 10.36227/techrxiv.19654026.v1

[7] S. Garg, U. P. Shukla and L. R. Cenkeramaddi, "Detection of Depression Using Weighted Spectral Graph Clustering With EEG Biomarkers," in IEEE Access, vol. 11, pp. 57880-57894, 2023, doi: 10.1109/ACCESS.2023.3281453.

[8] Zhang, X., Zhang, Z., Diao, W., Zhou, C., Song, Y., Wang, R., ... & Liu, G. (2023). Early-diagnosis of major depressive disorder: From biomarkers to point-of-care testing. TrAC Trends in Analytical Chemistry, 159, 116904.

[9] Ksibi, A., Zakariah, M., Menzli, L.J., Saidani, O., Almuqren, L., & Hanafieh, R.A. (2023). Electroencephalography-Based Depression Detection Using Multiple Machine Learning Techniques. Diagnostics, 13.

[10] Sharma, G., Joshi, A.M., Gupta, R., & Cenkeramaddi, L. (2023). DepCap: A Smart Healthcare Framework for EEG Based Depression Detection Using Time-Frequency Response and Deep Neural Network. IEEE Access, 11, 52327-52338.

[11] Rafiei, A., Zahedifar, R., Sitaula, C., & Marzbanrad, F. (2022). Automated detection of major depressive disorder with EEG signals: a time series classification using deep learning. *IEEE Access*, *10*, 73804-73817. https://doi.org/10.1109/ACCESS.2022.3190502

[12] Movahed, R.A., Jahromi, G.P., Shahyad, S. et al. A major depressive disorder diagnosis approach based on EEG signals using dictionary learning and functional connectivity features. Phys Eng Sci Med 45, 705–719 (2022). https://doi.org/10.1007/s13246-022-01135-1

[13] H. Wu and J. Liu, "A Multi-stream Deep Learning Model for EEG-based Depression Identification," 2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Las Vegas, NV, USA, 2022, pp. 2029-2034, doi: 10.1109/BIBM55620.2022.9995246.

[14] Yan, D., Zhao, L., Song, X., Zang, X., & Yang, L. (2022). Automated detection of clinical depression based on convolution neural network model. Biomedical Engineering / Biomedizinische Technik, 67, 131 - 142.

[15] Rafiei, A., & Wang, Y. (2022). Automated Major Depressive Disorder Classification using Deep Convolutional Neural Networks and Choquet Fuzzy Integral Fusion. 2022 IEEE Symposium Series on Computational Intelligence (SSCI), 186-192.

[16] Loh, H. W., Ooi, C. P., Aydemir, E., Tuncer, T., Doğan, Ş., & Acharya, U. R. (2021). Decision support system for major depression detection using spectrogram and convolution neural network with eeg signals. Expert Systems, 39(3). https://doi.org/10.1111/exsy.12773

[17] Movahed, R. A., Jahromi, G. P., Shahyad, S., & Meftahi, G. H. (2021). A major depressive disorder classification framework based on EEG signals using statistical, spectral, wavelet, functional connectivity, and nonlinear analysis. Journal of neuroscience methods, 358, 109209. https://doi.org/10.1016/j.jneumeth.2021.109209

[18] Saeedi, A., Saeedi, M., Maghsoudi, A., & Shalbaf, A. (2020). Major depressive disorder diagnosis based on effective connectivity in EEG signals: a convolutional neural network and long short-term memory approach. Cognitive Neurodynamics, 15, 239 - 252.

[19] Dang, W., Gao, Z., Sun, X. et al. Multilayer brain network combined with deep convolutional neural network for detecting major depressive disorder. Nonlinear Dyn 102, 667–677 (2020). https://doi.org/10.1007/s11071-020-05665-9

[20] Uyulan, Ç., Ergüzel, T.T., Unubol, H., Çebi, M., Sayar, G.H., Nezhad Asad, M., & Tarhan, N. (2020). Major Depressive Disorder Classification Based on Different Convolutional Neural Network Models: Deep Learning Approach. Clinical EEG and Neuroscience, 52, 38 - 51

[21] Mumtaz, W. (2016). MDD Patients and Healthy Controls EEG Data (New) (Version 2). figshare. https://doi.org/10.6084/m9.figshare.4244171.v2

[22] D. M. Khan, K. Masroor, M. F. M. Jailani, N. Yahya, M. Z. Yusoff and S. M. Khan, "Development of Wavelet Coherence EEG as a Biomarker for Diagnosis of Major Depressive Disorder," in IEEE Sensors Journal, vol. 22, no. 5, pp. 4315-4325, 1 March1, 2022, doi: 10.1109/JSEN.2022.3143176.

[23] Raghavendra, U., Gudigar, A., Chakole, Y., Kasula, P., Subha, D. P., Kadri, N. A., … Acharya, U. R. (2021). Automated detection and screening of depression using continuous wavelet transform with electroencephalogram signals. Expert Systems. doi:10.1111/exsy.12803