Journal of Computer Science

Original Research Paper

Novel Depression Classification Framework Using Optimal Feature Integration with Hybrid Convolution (1D/3D) Based Adaptive Residual DenseNet

[1,2]B. Manjulatha and [3]Suresh Pabboju

[1]*Osmania University, Hyderabad, Telangana, India*
[2]*Vignana Bharathi Institute of Technology, Aushapur, Ghatkesar, Hyderabad, Telangana, India*
[3]*Information Technology, Chaitanya Bharathi Institute of Technology, Gandipet, Hyderabad, Telangana, India*

Article history
Received: 13-09-2024
Revised: 14-11-2024
Accepted: 20-11-2024

Corresponding Author:
B. Manjulatha
Osmania University,
Hyderabad-500007,
Telangana, India.
Email: manjulathakbr@gmail.com

Abstract: This study introduces a novel depression classification method by incorporating multimodal data to reduce this limitation and improve the accuracy of depression diagnosis. In the beginning, the multimodal data such as speech signal, Electroencephalographic (EEG), and text data is obtained from the benchmark datasets. These acquired text data are subjected to the text pre-processing phase, where the stemming, character removal, punctuation and stop word removal operations are performed. After that, the resultant text is given to the Bidirectional Encoder Representations from Transformer (BERT), and the extracted features are considered Feature 1. From the EEG signals, feature 2 is attained from the wave features. Accordingly, feature 3 is attained from the linear and non-linear features. Finally, from the speech signals, the spectral feature is extracted and is considered Feature 4. Further, the extracted four features are optimally fused by using the proposed Modified Random Value of Osprey Optimization Algorithm (MRVOO). Subsequently, the optimally fused features and the video frames are subjected to the Hybrid 1D and 3D Convolution-based Adaptive Residual DenseNet (HCARDNet) for depression classification. Here, the network parameters are optimally determined by the same MRVOO. The performance is examined via distinct metrics and it outperforms with the better classification rather than baseline approaches.

Keywords: Depression Classification, Modified Random Value of Osprey Optimization Algorithm, Hybrid 1D and 3D Convolution-based Adaptive Residual DenseNet

Introduction

Countless people globally experience depressive disorders, a complex mental health condition characterized by a range of physical and mental symptoms, along with persistent feelings of sadness and disinterest in activities (Chen *et al*., 2022b). The rising prevalence of depression underscores the need for effective tools to identify, classify, and comprehend its multifaceted aspects (Li*et al*., 2023). The depression diagnosis framework incorporates a systematic approach to categorizing and evaluating depression, considering factors such as symptoms, severity, duration, and underlying causes (Xia*et al*., 2023). This framework is crucial for both medical practice and research, offering a structured method to assess, diagnose, and tailor interventions for individuals with depression. A well-constructed depression taxonomy framework is an invaluable resource for researchers, clinicians, and policymakers (Pérez-Toro*et al*., 2023). It enhances the quality of care for those struggling with mental illness and contributes to a better understanding of depressive disorders overall (Chen *et al*., 2019a).

Depression encompasses a diverse array of symptoms, with each individual's experience being highly unique (Khan *et al*., 2022). Existing categorization methods may struggle to capture this variability, potentially resulting in inaccurate diagnoses or overlooking critical information in symptom presentations (Shi *et al*., 2019). The intricate interplay between depression and other mental health conditions further complicates matters, as overlapping symptoms make it challenging to accurately distinguish between depressive disorders and concurrent issues (Fumarola*et al*., 2018). The cultural context significantly influences how depression is perceived and expressed. Current classification systems may not adequately consider this cultural nuance which leads to misunderstandings and inadequate treatment approaches for individuals from different cultural backgrounds (Kwon*et al*., 2019). Additionally, the static nature of categorization algorithms poses challenges, especially in cases of chronic or recurrent depression, where symptoms may evolve (Wagner, 2016). The inability of these algorithms to adapt as symptoms change could compromise the effectiveness of therapeutic planning and accurate diagnosis (Garg *et al*., 2023).

Incorporating Deep Learning Techniques into the development of classification systems for depression has resulted in significant advancements in the realm of mental well-being (Ganeshkumar *et al*.,2023). Deep learning, a subset of machine learning, employs neural networks to autonomously discern and extract intricate patterns from data (Zhao*et al*., 2020). This approach proves effective in identifying and classifying challenging-to-diagnose conditions, such as depression (Ahmad *et al*., 2020). Convolution Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) are the prominent types of deep neural networks utilized to analyze diverse data sources including text, Electroencephalogram (EEG) signals, and neuroimaging (Moussa *et al*., 2022). These networks capture intricate temporal and spatial patterns, enhancing our understanding of the neural correlates of depression (Villa-Pérez*et al*., 2023). They facilitate the identification of structural and functional abnormalities associated with depression by automatically extracting information from brain scans, aiding in the creation of imaging biomarkers for diagnostic purposes. While deep learning systems (Chen *et al*. 2022a) have made remarkable strides, they often face criticism for their lack of transparency. Efforts are underway to enhance the interpretability and understanding of depression categorization models. Techniques like saliency maps and attention processes are employed to reveal and interpret the decision-making process within the framework (Suen *et al*., 2021).

The objectives of the proposed depression classification method are depicted below:

- To create a deep learning-based depression classification model using advanced MRVOO and HCARDNetapproaches that aim to classify depression in early stages and enhance treatment outcomes through timely identification.
- To determine the weighted optimal fused features, where the weight and features are chosen accurately by MRVOO. After providing the diverse features, the average computation is done to get a better feature vector.
- To promote an MRVOO algorithm to tune the variables like the number of hidden neurons, epoch count, and steps per epoch in DensNet that is utilized for enhancing the precision and accuracy rate
- To adopt an effective model of HCARDNet for depression classification, this model influences the hybrid convolution mechanism in residual DenseNet. The parameter optimization is conducted using the MRVOO approach.
- To empirically validate the developed model using industry-standard benchmarks and appropriate evaluation criteria. This phase is crucial for confirming the developed approach's precision in classifying depression and assessing its performance in real-world scenarios.

Desk Review

Documentary Study

Rizwan *et al*. (2022) have suggested a comprehensive analysis that delved into the efficiency of four transformer-based small language algorithms in classifying the magnitude of depression using tweets. Notably, the focus was on models with fewer than 15 million adjustable variables: Electra Small Discriminator (ESD), Electra Small Generator (ESG), Albert Base V2 (ABV), and XtremeDistil-L6 (XDL). The models underwent optimization through the adjustment of various hyper-parameters to attain optimal results. Following this fine-tuning process, the models were systematically evaluated by categorizing the severity of depression in labeled tweets into three distinct groups: Severe, moderate, and mild.

Sam *et al*. (2023) have implemented a Spiking Neural Network (SNN) design and a Long Short-Term Memory (LSTM) architecture, marking the first-ever attempt to simulate the brain's fundamental structures during various stages of melancholy. The methods used in this study included activities such as categorizing and forecasting specific results, illustrating the anatomical changes in the brain associated with the expected outcomes, and providing analyses of the obtained results.

Jiang *et al*. (2021) have developed an efficient technique for classifying depression using geographical information based on EEG monitoring. The extraction and selection of features utilized an evolutionary algorithm and differential sensitivity, while classification was performed using a support vector machine. An intelligent method was suggested before feature extraction to address spatial disparities.

Seal *et al*. (2021) have suggested the electrical function of the brain is the EEG. It was utilized to generate accurate reports on the severity of depression. Previous research has suggested that deep learning models trained with EEG data can be employed for the diagnosis of mental illnesses. In this study, a CNN constructed with deep learning named DeprNetwas suggested to distinguish between EEG data from depressed and normal individuals. The degree of depression was measured using the patient health questionnaire's 9 rating.

Shen *et al*. (2023) have proposed the challenges associated with identifying melancholy using EEG technology involved in effectively maximizing the spatial data acquired from the multidimensional space of EEG signals. To address this difficulty, an adaptive channel fusion technique for melancholy recognition was proposed, that utilized EEG signals and improved Focal Loss (FL) functions. Two enhanced FL functions were presented in this method, aiming to increase their separation by weighting the losses of hard samples as optimization goals. Additionally, an adaptive channels fusion framework was suggested to optimize the channel strengths.

Kour and Gupta (2022) have implemented a combination of two deep learning structures, Bi-Directional LSTM (Bi-LSTM) and CNN, that was optimized to achieve a reliability of 94.28% on a benchmark melancholy dataset comprised of Twitter data. The CNN- Bi-LSTM model was compared with baseline methods, namely CNN and RNN. Our methodology contributed to improved prediction accuracy, as indicated by experimental data across various performance indicators. A comprehensive analysis of the issue was conducted using statistical and visualization methods, revealing a significant distinction between the language depiction of depressed and non-depressive content.

Swasthika Jain *et al*. (2023) have developed a hybrid Artificial Intelligence (AI) system, founded on machine learning that was developed for multi-modal depression analysis. This concept involved extracting written, audio, and video descriptions, along with other multi-modal data. Initially, the suggested method employed a combination framework of deep learning models to estimate the melancholy scale using the Patient Health Questionnaire (PHQ). Subsequently, the learning approach was designed to deduce people's emotional and physiological conditions concerning the Freudian characteristics of depression.

Uddin *et al*. (2022) have developed an effective method for finding texts that described self-perceived signs of depression, utilizing an RNN based on LSTM. The method was applied to a substantial dataset extracted from a Norwegian youth-oriented public internet stream that consists of inquiries written by the youth themselves. Following this, the robust features were extracted from the reflections of potential depressive symptoms, which had been identified by medical and psychology professionals. These features were then encoded using a one-hot method.

Recent References from the Depression Classification Model

Rejaibi *et al*. (2022) developed a deep Recurrent Neural Network-aided approach to detect depression and also its severity level from speech. The advanced features of the designed approach were non-invasiveness, non-intrusion, and fastness and also it applies to real-time applications.

Othmani*et al*. (2021) promoted anovel EmoAudioNet for emotion and depression recognition from speech. At last, the given developed model attained promising results and also it shows better classification results.

Othmani*et al*. (2022) introduced a deep structured-based approach for depression recognition using audiovisual data. It has shown a better accuracy of the F1-score rate using a standard dataset. The normality-based approach detects depression and in detecting depression relapses accurately. Here, a prospective monitoring system was designed to assist depressed patients.

Problem Statement

Millions of individuals globally suffer from depression, a complicated mental health illness that negatively affects their emotional health and general quality of life. Mental health practitioners used a system of categorization to better understand, determine, and treat different forms of sadness since depressive diseases are diverse. Using this classification made it easier to customize treatments according to the intensity, underlying reasons, and particular manifestations (Srujan *et al*., 2018). The upper comings and lower comings of the existing model are provided in Table (1):

- In the conventional depression classification model, the machine learning methods massively depend on manually annotated text, EEG, and speech signals and also that can be done only by experts. Therefore, hybridized techniques are suggested for a better depression classification model.
- The major challenges for extracting the features using conventional deep learning techniques are individual differences in the text, EEG, speech signals, and video frames of different subjects. It takes more time for processing and it reduces the accuracy rate. To rectify these issues, diverse feature extraction techniques can be suggested in the depression classification model.
- While choosing the accurate features, most of the conventional techniques use deep learning techniques that suffer from problems like scalability and stability. Hence, it is essential to use a hybridized algorithm that helps minimize computational time and achieves accurate features for classification.
- To perform depression classification, the baseline machine learning techniques face data availability and overfitting issues while dealing with complex models and small datasets. Therefore, diverse deep-learning techniques can be combined to promote a hybrid model for getting better classification outcomes.

Table 1: Merits and demerits of baseline depression classification system using deep structured network

Author [citation]	Technique	Merits	Demerits	Proposed Solutions
Rizwan*et al*. (2022)	Deep transfer learning	It allows pre-trained models on one task to be adapted for another task, saving computational resources and time	This may suffer if there is a significant dissimilarity between the target and source domains	The deep learning model used in the proposed work can address the degradation problem and extract more descriptions from the source and target domains
Sam *et al*. (2023)	LSTM	It is capable of learning and remembering patterns in sequential data	Training LSTMs can be computationally expensive	In this proposed model, the utilization of shared memory allocations efficiently reduces the training time
Jiang *et al*. (2021)	EEG	It provides an advanced technique for monitoring brain activity It offers high temporal resolution, capturing changes in brain activity over short time intervals	It can be sensitive to external noise and artifacts	The proposed model employs a wave-based features technique that efficiently removes the artifacts and noise from the EEG signal
Seal *et al*. (2021)	DL	It automatically learns hierarchical representations of features	It needs a large amount of data for effective training. Training deep CNNs can be computationally intensive	The proposed work supports feature reuse and thus it significantly deduces the count of variables while maintaining elevated performance, which allows a limited amount of data
Shen *et al*. (2023)	EEG signals	It is well-suited for capturing dynamic changes in the developed work	It can provide more information about the detection of diseases, but precise localization of activity within the disease is challenging	A deep learning-based depression model has the ability to provide more information about the depression model making the developed deep learning model analyze efficiently to provide accurate detection results
Kour and Gupta (2022)	CNN	It exploits the spatial hierarchies present in data, enabling them to capture patterns at diverse stages of abstraction	Training deep CNNs can be computationally expensive and time-consuming	In this proposed model, the utilization of shared memory allocations efficiently reduces the training time
Swasthika Jain *et al*. (2023)	SVM	Utilizing the operations of the kernel enables effective management of non-linear decision boundaries	Its sensitivity to noisy data may make it prone to overfitting	The deep learning model used for the proposed work supplies generalization techniques to resolve the overfitting issues
Uddin *et al*. (2022)	DL	Its flexibility allows for application across a diverse range of activities and domains	Efficient training necessitates the collection of a substantial amount of labeled data	The deep learning model used for the proposed work has efficient training resources with the collection of a substantial amount of labeled data

Materials and Methods

Description of Novel Framework of Depression Classification and Utilized Datasets

Recommended System of Depression Classification

Depression is a pervasive mental health condition that poses noteworthy issues for individuals, families, and society at large. Researchers and healthcare professionals utilize classification methods to comprehend and address the diverse manifestations of this illness (Muzammel *et al*., 2020). These systems provide a systematic framework for identifying and diagnosing depression-related conditions, aiding in more accurate patient diagnoses by establishing criteria and symptom clusters (Muzammel *et al*., 2020). The broad spectrum of depression encompassing various signs and severity levels, leads to misdiagnoses or oversimplified representations of the illness. The symptoms of depression often overlap with those of other mental health disorders, making it challenging to differentiate (Muzammel *et al*., 2021). Distinguishing between different conditions, such as bipolar disorder and anxiety disorders, may be intricate (Yasin *et al*., 2023; Othmani and Zeghina, 2022). Classification systems may fall short of acknowledging how cultural and societal factors influence the manifestation and understanding of depression. This limitation could impact the effectiveness of criteria across diverse demographics (Yasin *et al*., 2021). To mitigate these challenges, an advanced depression classification model is adopted using machine learning and machine learning approaches to enhance the efficiency of the categorization system and also help to reduce errors. This approach addresses the complexities inherent in diagnosing depression, accounting for the intricate interplay of symptoms, severity, and cultural influences. Figure (1) shows the heuristic approach and deep learning-based depression classification.

Recognizing and classifying depression is a challenging task due to its multifaceted nature. This study promotes a robust classification model named HCARDNet, which leverages diverse data sources, including text, EEG, and speech signals. A comprehensive dataset containing different kinds of data types like text, EEG, and speech signals is utilized for training and testing the developed text pre-processing model. Wave, non-linear, and linear features are derived from EEG signals and text features are obtained from text data. Additionally, speech signals are employed to extract spectral information. Incorporating multiple modalities ensures a comprehensive representation of the underlying characteristics associated with depression. The application of MRVOO ensures the proper selection of weights and features to generate optimal weighted fused features. After providing the diverse features, the average computation is done to get the better feature vector. The proposed HCARDNet model employs a combination of hybrid convolutions and residual DenseNet structure. Due to the architecture's capacity to capture temporal and geographical correlations in the data, the model is ensured to acquire the capability to learn intricate patterns associated with depression. The MRVOO method is used to optimize the variables of HCARDNet. The effectiveness of the developed model is empirically validated through appropriate evaluation criteria and industry-accepted benchmarks. This step is crucial for confirming the model's accuracy in detecting depression and evaluating its performance in real-world scenarios.

Input Data Collection

Collect diverse multimodal data from various sources to enrich the depression categorization model. This system specifically focuses on extracting text, EEG signals, and speech signals. Utilizing the provided URLs, three distinct sets of data were generated.

Text data:The text data is gathered from the link of https://www.kaggle.com/datasets/arashnic/the-depression-dataset accessed on 2024-01-06. This can establish a basis for evaluating diverse machine learning methods and methodologies that encompass oversampling strategies for addressing unbalanced class issues and cost-sensitive segmentation. The collected text data is represented as $Text^{input}$.

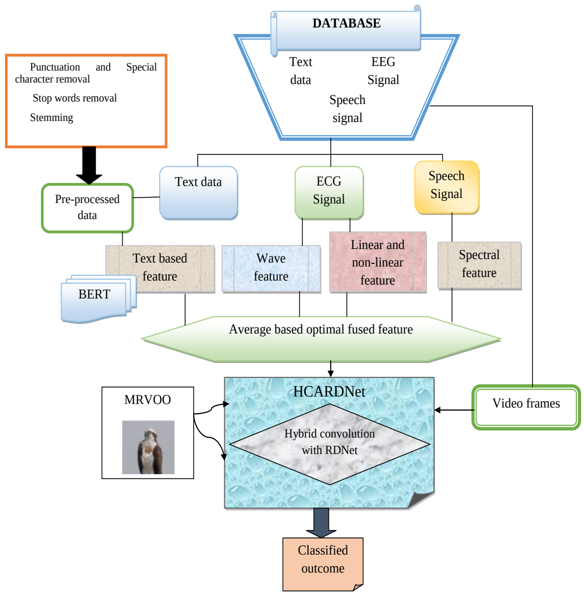


Fig. 1: Structural view of the deep learning-based depression classification

EEG signal:Using the link https://www.kaggle.com/datasets/tocodeforsoul/depression-rest-preprocessed accessed on 2024-01-06. The EEG signals encompass information from 122 individuals. Three subjects were omitted due to data invalidity and an additional three were excluded due to inaccurate results. The gathered EEG signal is represented as EEG^{input}.

Audio signal and video frames: The source is presented in the link https://zenodo.org/record/1188976, access date: 2024-01-06. This database consists of both audio signals and video frames. This consists of Twenty-four seasoned actors, comprising twelve men and twelve women who articulate two lexically matching phrases using a neutral North American accent. The vocalizations encompass emotions of anger, sadness, calmness, happiness, disgust, fear, and surprise both in speech and songs.

The target of the dataset is suicide or non-suicide. The gathered audio signal and video frames are specified as AU^{input} and VI^{input}.The sample images of the recommended system are shown in Fig. (2).

Text Pre-Processing

It is a critical phase in NLP, involving the preparation of raw text for analysis or machine learning applications by cleaning and formatting. It is crucial for developing reliable and accurate models for various applications. The objectives are to transform unstructured text into a structured format that is suitable for modeling and analysis. This involves three pre-processing steps, which are outlined below:

- Punctuation and special character removal
- Stop words removal
- Stemming

DOI: 10.3844/jcssp.2025.524.248

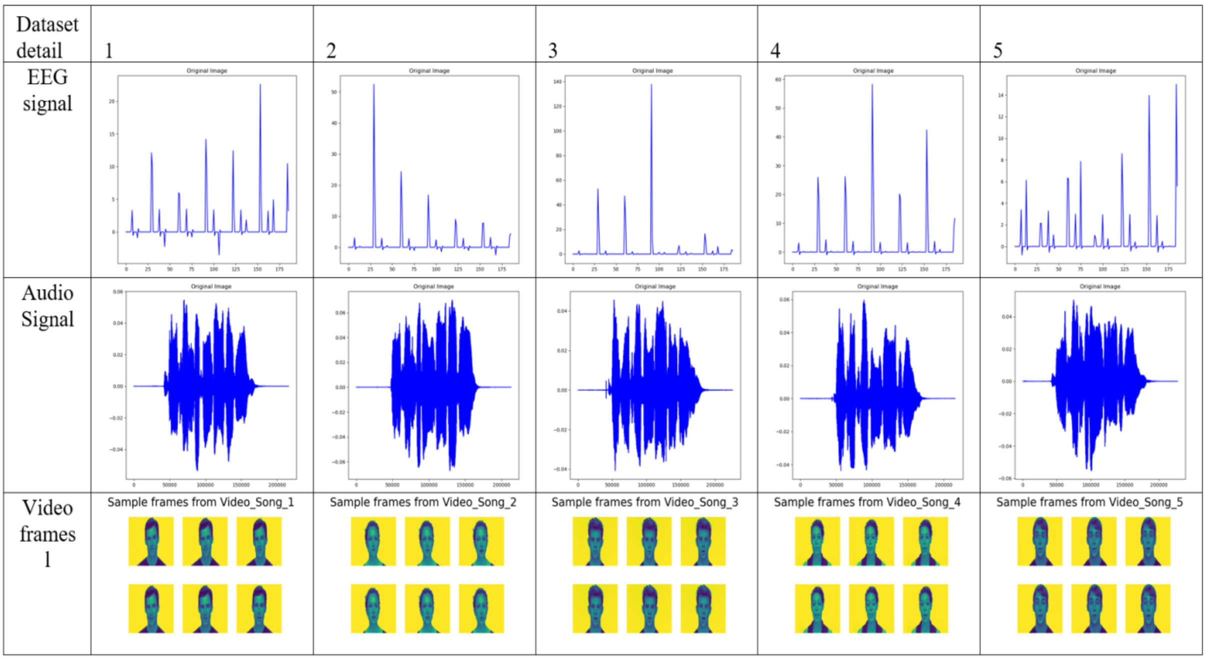


Fig. 2: Sample images of the recommended system

Punctuation and special character removal: The text data $Text^{input}$ is the input to this phase, the initial removal of special characters and punctuation marks (He *et al*., 2022) stands as a pivotal first step in the preliminary data processing phase. This procedure is designed to enhance the consistency and cleanliness of textual data, preparing it for tasks related to NLP and other forms of analysis. The systematic elimination of this serves to standardize the language, facilitating a more accessible comprehension and interpretation by subsequent models and computational processes. This foundational pre-processing stage significantly boosts the overall effectiveness and quality of text-based analyses and solutions. Finally got the special characters and punctuation marks removed data T^{pun}.

Stop words removal: The special characters and punctuation marks removed data T^{pun} is the input to this phase and addressing the challenge of extracting pertinent information from extensive text data, particularly in the context of NLP comprehension is a primary concern (He *et al*., 2022). Stop words are common terms with minimal semantic significance that pose a hindrance to the effectiveness and precision of text-based analytics. Therefore an indispensable pre-processing measure involves the removal of these stop words. This intentional removal enhances both the accuracy and comprehensibility of the processed text data, streamlining subsequent analyses for improved outcomes. Finally got the stop word removed data T^{stop}.

Stemming: The stop word removed data T^{stop} is the input to this phase, it involves systematically removing suffixes and prefixes from words to create a more standardized representation. This ensures consistent treatment of words with similar meanings and reduces redundancy in the process. By employing stemming (He *et al*., 2022), text analysis models and algorithms can identify and associate related terms that lead to enhanced accuracy and coherence in tasks such as sentiment analysis, document clustering, and data retrieval. The integration of stemming is indispensable for optimizing the processing capabilities of textual data, ultimately elevating the performance of various machine-learning applications. Finally, the pre-processed data$Text^{pre}$is obtained, which marks a significant milestone in text pre-processing. Having clean and well-organized data is crucial for subsequent analysis.

Feature Extraction of Input Source as Text and EEG Signal for Classifying the Depression

BERT-Based Text Features

The pre-processed text data $Text^{pre}$ is the input to this phase and the architecture of the multi-layer BERT (Ogunleye *et al*., 2024) model is rooted in its original design. Each encoder in this model comprises an array of m = 6 identical layers, with two sub-layers embedded within each of these layers. Following the normalization of each layer, an additional connection is established for both sub-layers. In other words, the input of each sub-layer is normalized using $Ln(y + Sb(y))$ where $Sb(y)$ represents the operation performed by the sub-layer.

Now, multi-head self-attention is denoted as $Att(t, y, u)$. The definition is as follows Eq. (1):

$$Att(t,y,u) = softmax\left(\frac{ty^T}{\sqrt{f_y}}\right)uText^{pre} \quad (1)$$

where, the term t represents a matrix of search queries, y is a vector of keys and u is a vector of corresponding answers. The variable f_y is denoted as the size of the matrices t and y. Analyze the attention from various perspectives by examining these matrices and their relationships are displayed in Eq. (2):

$$Mh\,(t,y,u) = con\,(h_1, h_2 .....h_h)\; p_o \quad (2)$$

Subsequently, perform the attention operation resulting in output values of dimension f_y.

Pretraining (Zeberga *et al*., 2022): There are two pre-training methods: Next Sentence Prediction (NSP) and Masked Language Model (MLM). In the MLM model, a certain proportion of input tokens is randomly masked using the [MASK] token and the task involves predicting the concealed tokens.

Fine-tuning (William *et al*., 2022): The self-attention mechanism in the transformer empowers BERT to effectively represent diverse upstream tasks, simplifying the fine-tuning process. All that is required is to input the specific inputs and outputs for a given task into BERT and adjust each value accordingly. Finally, the obtained text-based features are described as $Text^{features}$, these features serve as the foundation for various applications.

Wave-Based EEG Features

EEG is a powerful and non-invasive tool that allows researchers and clinicians to capture and analyze the electrical activity occurring in the brain. The wave-based feature$wave^{feature}$is extracted from the garnered EEG ^{input} signal. Within EEG analysis, wave-based features play a crucial role in characterizing the different frequency components of brain signals. The wave features (Lee *et al*., 2021) like the P wave, QRS complex wave, T wave, and U wave were explained as follows.

P Wave: The initial wave in an ECG, known as the P wave, represents the electrical activity of the upper chambers of the heart, known as atria.

QRS complex wave: The QRS complex in an ECG is formed by the combination of three waves: Q,R, and S. The Q wave represents the beginning of the negative displacement, the R wave denotes the positive displacement, and the S wave presents the subsequent negative displacement.

T wave: The T wave represents the electrical activity associated with ventricular depolarization, which is the recovery phase of the ventricles.

U wave: The U wave is a small, positive deflection that sometimes appears after the T wave.

Linear and Non-Linear Features

The EEG signal is a representation of the brain's electrical activity that serves as a rich source of information that can be harnessed to gain insights into cognitive processes, monitor neurological conditions, and develop innovative applications. From the EEG signal, both linear and non-linear features $Feature_{linear}^{Non\text{-}linear}$ are extracted. Linear features (Avots *et al*., 2022) such as mean, standard deviation, skewness, and kurtosis, along with non-linear features (Dehghani *et al*., 2023) like Hjorth parameters (activity, mobility, and complexity), collectively offer a comprehensive insight into the intricate dynamics within the brain. The formulas for skewness, mean, standard deviation, and kurtosis are as follows.

Mean: An essential measure reflecting the primary trend within a dataset is the mean, denoted byμ this is formulated in Eq. (3):

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i \quad (3)$$

Standard deviation: A metric indicating the extent of dispersion or spread among a set of data points is the standard deviation, denoted by the symbolσ. It quantifies the degree of variation from the mean, illustrating the extent to which data points deviate from the average. This is formulated in Eq. (4):

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(x_i - \mu\right)^2} \quad (4)$$

Skewness:Skewness serves as a measure for assessing the asymmetry of a probability distribution. Symmetry is indicated when skewness is zero, while positive skewness suggests a lengthening of the right tail and negative skewness implies a lengthening of the left tail. The equation employed to calculate skewness $\gamma$is in Eq. (5):

$$\gamma = \frac{\frac{1}{N}\sum_{i=1}^{N}\left(x_i - \mu\right)^2}{\left(\sum_{i=1}^{N}\left(x_i - \mu\right)^2\right)^{3/2}} \quad (5)$$

Kurtosis: Kurtosis is a measure that quantifies the tailedness or sharpness of a probability distribution, indicating its susceptibility to outliers compared to a normal distribution. The calculation of kurtosis $\kappa$is determined by the following Eq. (6):

$$\kappa = \frac{\frac{1}{N}\sum_{i=1}^{N}\left(x_i - \mu\right)^2}{\left(\sum_{i=1}^{N}\left(x_i - \mu\right)^2\right)^{3/2}} - 3 \quad (6)$$

The normal distribution exhibits a kurtosis of three and distributions with higher kurtosis typically feature heavier tails. From the above equations, the variable N is the amount of data points and x_i defines each data point.

Extracted Features of Speech Signal and Proposed MRVOO for Integrating the Resultant Features

Spectral Features from Speech Signal

The speech signal $speech^{input}$ is fed input to this phase. To derive resilient characteristics in the presence of noise, an array of features is extracted from the speech signals. These features encompass spectral centroid (Saga *et al*., 2022), spectral flux (Schubert *et al*., 2004), spectral density, spectral roll-off, MFCC (Chen *et al*., 2019b), peak amplitude, total harmonic distortion, zero-crossing rate, entropy, standard deviation and Root Mean Square of the Sum of Successive Differences (RMSSD) (Li *et al*., 2020). Finally, the spectral feature is represented as $^{the\,feature}$.

Proposed Meta-Heuristic Algorithm: MRVOO

The optimization method MRVOO utilizes the OOA (Berntson *et al*., 2005) for optimizing parameters. OOA is known for its efficient parallelizability, which leverages processing power effectively. It is crucial to acknowledge that robust optimization algorithms often derive their resilience from sound theoretical foundations. Users should exercise caution in evaluating the reliability of OOA, particularly if it lacks substantial scientific validation or a strong theoretical basis. To mitigate this limitation, the recommended approach is utilized to reduce the errors. Eq. (7) presents a mathematical representation of this proposed approach:

$$\gamma = \frac{cf}{bf + wf} \tag{7}$$

In the broader context of the proposed MRVOO, the variableγ is considered random. The factors *bf*, *wf*, and *cf* represent the respective current, worst, and best levels of fitness. Equation (12) articulates the update mechanism for introducing a random factor. The computational framework of the recommended MRVOO is outlined as follows.

The intelligent natural tendencies of ospreys in their pursuit and adept handling of fish for optimal feeding positions can be employed as an inspiration for devising an innovative optimization algorithm. Through a repetition-based methodology, the proposed OOA functions as a community-driven technique capable of providing effective solutions based on the collective search abilities of its solutions in problem-solving domains. The OOA community, comprising all ospreys, can be succinctly represented using a matrix corresponding to Eq. (8). The initiation of ospreys within the search space is determined arbitrarily at the outset of the OOA design, as indicated by Eq. (9):

$$E = \begin{bmatrix} E_1 \\ \vdots \\ E_2 \\ \vdots \\ E_n \end{bmatrix} \begin{bmatrix} E_{1,1} & \cdots & E_{1,d} & \cdots & E_{1,m} \\ \vdots & \ddots & \vdots & \cdot^{\cdot^{\cdot}} & \vdots \\ E_{i,1} & \cdots & E_{i,d} & \cdots & E_{i,m} \\ \vdots & \cdot^{\cdot^{\cdot}} & \vdots & \ddots & \vdots \\ E_{N,1} & \cdots & E_{N,d} & \cdots & T_{N,m} \end{bmatrix}_{N\times m} \tag{8}$$

$$E_{i,d} = \lambda + \gamma_{i,j}(\lambda - \chi) \tag{9}$$

Where E represents the locations in the osprey demographic matrix, E_i denotes the osprey, and $E_{i,m}$ describes the degree of the i^{th} problem variables. The quantity of ospreys is noted as m representing the random variable is noted as $\gamma_{i,j}$. Additionally, λ and x denote the lower and upper boundaries, for the j^{th} issue variable.

The objective function can be evaluated for each osprey since each one serves as a potential solution to the problem, as indicated by Eq. (10). Using vectors, the assessed solutions for the problem's objective function can be effectively represented:

$$S = \begin{bmatrix} S_1 \\ \vdots \\ S_2 \\ \vdots \\ S_n \end{bmatrix}_{N\times 1} = \begin{bmatrix} S(r_1) \\ \vdots \\ S(r_2) \\ \vdots \\ S(r_n) \end{bmatrix}_{N\times 1} \tag{10}$$

where, S_i represents the obtained objective function values for the i^{th} iteration and S is the vector containing the variables of the objective function.

Stage 1: Location Identification and Fish Hunting (Exploration)

Leveraging their sharp vision, ospreys emerge as formidable hunters capable of spotting fish beneath the surface. Their hunting process involves locating the fish, striking it, and subsequently submerging themselves to capture it.

The equation governing the catch for each osprey is presented in Eq. (11):

$$SL_i = \{E_S \quad L_S < L_i \ and \ E \in \{1,2...f\}\}U\{E_{best}\} \tag{11}$$

where, E_{best} represents the optimal solution (the best osprey) and SL_i denotes the set of fish placements for the i^{th} osprey.

In this process, the osprey randomly selects one of the previously described fish before striking it. Equation (12) is then employed to calculate the osprey's new position, utilizing a computer simulation that mimics the osprey's approach to the fish. Following Eq. (13), if the new osprey location enhances the significance of the objective operation, it replaces the old one:

$$h_{i,j}^{s_1} = h_{i,j} + \gamma_{i,j}^{s_1}(yy_{i,j} - K_{i,j}.\gamma_{i,j}) \quad (12)$$

$$W_i = \begin{cases} W_i^{s_1} & D_i^{s_1} < D_i \\ W_i & else \end{cases} \quad (13)$$

These variables $W_i^{s_1}$ are contingent on the initial phase of the OOA and denote the updated location of the i^{th} osprey. Specifically, $yy_{i,j}$ represents its j^{th} dimensions, and $K_{i,j}$ signifies the objective function value.

The variable γ serves as a random parameter, with a conventional random variable typically having a range between 0 and 1. This range could potentially lead to challenges in convergence and optimization. To address potential issues related to this range and minimize errors, the suggested formula in Eq. (7) is employed to compensate for inaccuracies.

Phase 2: Carrying the Fish to the Suitable Position (Exploitation)

Subsequently, as per Eqs. (14-15), if the value of the cost function increases at this updated location, the osprey's prior location is substituted with the new one:

$$h_{i,j}^{s_1} = h_{i,j} + (1 - 2\gamma_{i,j}).\frac{\lambda - \chi}{t} \quad (14)$$

$$W_i = \begin{cases} W_i^{s_2} & D_i^{s_2} < D_i \\ W_i & else \end{cases} \quad (15)$$

These variables $W_i^{s_2}$ are contingent on the second stage of the OOA and denote the updated location of the i^{th} osprey. Specifically, $yy_{i,j}$ represents its j^{th} dimensions, and $K_{i,j}$ signifies the function of objective value. Fig (3) Shows the flowchart of the offered MRVOO system.

Attainment of Optimal Feature Fusion

Feature fusion involves the amalgamation of numerous features that aim to enhance both the comprehensive representation of data and the efficiency of classification systems. Here, the extracted features from EEG and signals were considered as continuous signals. The objective is to refine the overall structure of the data, thus fostering improved performance in classification tasks.

From the extracted feature, optimally choose features $OText^{features}$, $Owave^{features}$, $OFeature_{linear}^{Non-linear}$ and $OSpectral^{features}$ from the extracted set by assigning optimal weights. This optimal selection of features and weight is done with the help of the recommended MRVOO approach. Without optimal selection, errors may occur, and using optimal selection is proposed as a way to minimize or lessen these errors. The mathematical representation of the acquired weighted feature is shown in Eq. (16):

$$\begin{aligned} &OW_1 \times OText^{features} = F_1 \\ &OW_2 \times Owave^{feature} = F_2 \\ &OW_3 \times OFeature_{linear}^{Non-linear} = F_3 \\ &OW_4 \times OSpectral^{feature} = F_4 \end{aligned} \quad (16)$$

The final step in the process entails concatenating the resulting features by computing their average value. This ultimately achieves optimal feature fusion $Optimal_{kl}^{fusion}$, a representation of which can be expressed mathematically in Eq. (17):

$$Optimal_{kl}^{fusion} = \frac{F_1 + F_2 + F_3 + F_4}{4} \quad (17)$$

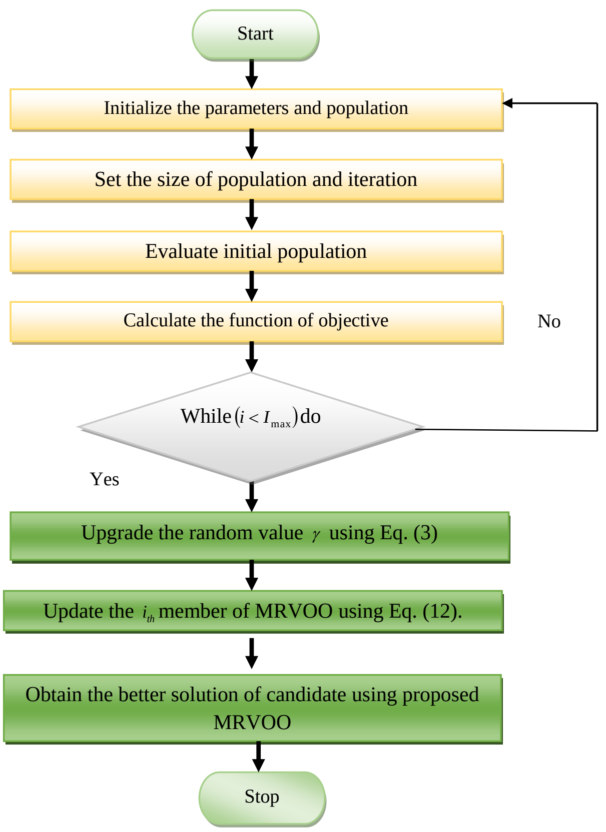


Fig. 3: Flowchart of recommended MRVOO

Optimizing weights in the feature extraction procedure results in a fixed configuration that may be insufficiently adaptable to the specific requirements of the task or the characteristics of the data. This lack of flexibility could potentially result in suboptimal performance across various scenarios. By optimizing the weights and features, the proposed technique aims to mitigate the issue and enhance the relief score. Equation (18) encapsulates the mathematical model of the objective parameter for optimal feature fusion:

$$UY_1 = \underset{\left\{\begin{array}{l} OW_1, OW_2, OW_3, OW_4, OText^{features}, \\ Owave^{feature}, OFeature_{linear}^{Non-linear}, \\ OSpectral^{feature} \end{array}\right\}}{argmax} (Rf) \tag{18}$$

Following the previously described computation, the parameters OW_1, OW_2, OW_3, OW_4, $Otext^{feature}$, $Owave^{features}$ $OFeature_{linear}^{Non-linear}$, $OSpectral^{feature}$signifies the optimal weights and features, also the range of optimal weights is denoted as [0.01-0.99] and the features are [1-No. of 4 feature sets], respectively. Additionally, the parameter value Rf serves as an indicator of the relief score. The mathematical representation of the relief score is articulated in Eq. (19).

Relief score: It assesses the significance of each feature in distinguishing among different classes within a dataset. This metric gauges the efficiency of a feature in separating instances belonging to a particular class from those belonging to other classes:

$$\Sigma Rf = k_i - (l_i - FM_i) + (l_i - FD_i) \tag{19}$$

The variable FM_i denotes the same class and FD_i denotes the various classes. Figure (4) shows the diagrammatic view of the recommended optimal weighted feature fusion.

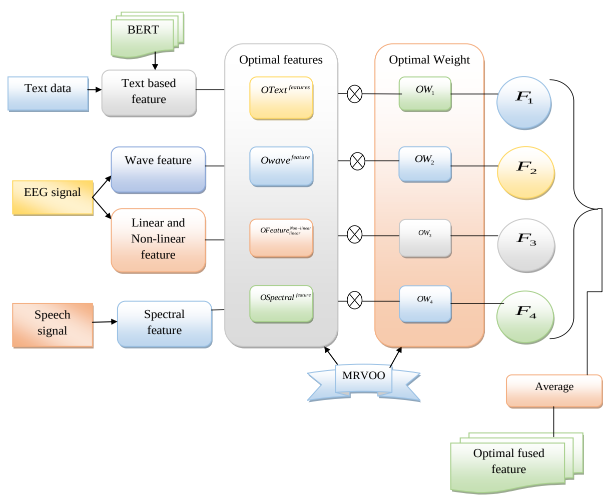


Fig. 4: Diagrammatic view of optimal weighted feature fusion process

Description of Optimal Weighted Feature Fusion Process

The optimal weighted feature fusion process is done with the support of the MRVOO algorithm to reduce the computational time and enhance the classification accuracy:

- Step 1: Diverse data such as text, EEG, and speech signals are collected
- Step 2: The feature extraction process is carried out

(i) Initially, the text features are extracted from the text data related to depression using the BERT model. It is utilized to extract features accurately without data loss from the text data. It is applicable to the image classification algorithm that has the ability to learn highly abstract features.
(ii) The features from EEG signals are extracted including wave, non-linear, and linear features.
(iii) The speech signals are used to extract spectral (features) information

- Step 3: In this step, the extracted features are fed into the MRVOO algorithm for optimal feature selection and weight optimization

(i) The optimally chosen features from the text, EEG signal, and speech signal such as$\{Otext^{feature}\}$,$\{Owave^{feature}\}$ $\{OFeature_{linear}^{Non-linear}\}$, and$OSpectral^{feature}\}$
(ii) The weights of each feature are optimized using the MRVOO algorithm
(iii) Accurately selected features are fused with their optimal weights to attain optimal weighted features
(iv) The concatenation process of optimal weighted features is carried out by computing their average value and the fused optimal weighted features are indicated by $Optimal_{kl}^{fusion}$

Hybrid Convolution-based Adaptive Residual DenseNet for Classifying the Depression

Residual DenseNet

By amalgamating crucial elements from two prominent architectures, residual DenseNet (Dehghani and Trojovský, 2023) emerges as a novel design in the realm of deep neural networks.

Backbone network: The output element of Residual-DenseNet employs a smoothing mechanism to integrate information from the Backbone Networks, aiming to generate a robust feature vector. Collaboratively, these modules give the formation of a characteristic vector comprising 64 dimensions. Notably, when establishing connections with DenseBlock4, the amalgamation of high-level characteristics is observed to be more dependable compared to the fusion of lower-level features. This calculated fusion significantly enhances the overall effectiveness of the model.

Feature outcome module: The outcome element of Residual-DenseNet employs a smoothing mechanism to integrate information from the Backbone Networks, aiming to generate a robust feature vector. Collaboratively, these modules present the formation of a characteristic vector comprising 64 dimensions. Notably, when establishing connections with DenseBlock4, the amalgamation of high-level characteristics is observed to be more dependable compared to the fusion of lower-level features. This calculated fusion significantly enhances the overall effectiveness of the model. This procedure improves the system's proficiency in associating image elements with the watermark. Equation (20) visually depicts this comparative procedure:

$$UI(y)=\begin{cases}1, & FDE(y)\geq\delta\\ 0, & otherwise\end{cases} \tag{20}$$

Here, the term$\delta$is displayed in Eq. (21):

$$\delta=\frac{1}{64}\sum_{k=0}^{63}FDE(y) \tag{21}$$

The comparison between the robust hashing vector $FDE(y)$ and the resilient vector of features generated by Residual-DenseNet is conducted using the mean binarization operationFDE. Figure (5) gives the architectural presentation of Residual DenseNet.

Hybrid Convolution Mechanism

The objective of a hybrid convolution system is to integrate multiple convolution techniques into a unified design. This integration may involve combining spatial convolutions, depth-wise convolutions, dilated convolutions, or other specialized convolution procedures. The aim is to harness the unique benefits of each process to optimize feature extraction, acquire contextual information, and enhance the network's ability to understand intricate patterns.

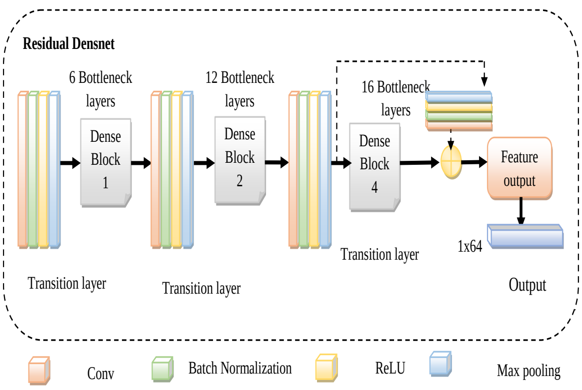


Fig. 5: Architectural view of residual DenseNet

Novel HCARDNet for Classification

HCARDNet emerges as a promising advancement in the field of depression classification models. The HCARDNet model is developed with the combination of 1D and 3D Convolution-based Adaptive Residual DenseNet networks for depression classification. Its unique combination of hybrid convolutions and residual DenseNet structure aims to overcome the challenges posed by diverse and complex patterns within depression-related data. 1D convolution excels in analyzing periodic or sequential data, including audio signals and time-series data, where its effectiveness shines. On the other hand, for analyzing multidimensional data, such as medical images (like MRI or CT scans), 3D models,or audio recordings where temporal dynamics are key, 3D convolution emerges as the preferred choice. This approach effectively captures the complexity inherent in such data types, balancing the need for spatial and temporal analysis. In this architecture, the optimally fused feature $Optimal_{kl}^{fusion}$ undergoes 1D convolution while the video frames VI^{input} are processed with 3D convolution. These two layers are then merged, forming a combined input that is subsequently fed into the next layer of a residual DenseNet and finally obtains the classified outcome. Enlarging the number of hidden neurons in DenseNet or any neural network contributes to higher model complexity. However, this heightened complexity comes with a risk of over fitting, where the model adapts too closely to the training data, capturing noise rather than the underlying patterns. Setting the epoch count too high exacerbates this risk by allowing the model to continuously learn from the training data, potentially hindering generalization to new data. To mitigate these issues and enhance accuracy and precision, the variables such as the number of steps per epoch, the number of epochs, and the number of suitably hidden neurons in DenseNet are optimally tuned. Equation (22) provides a mathematical formulation of the proposed system's objective function, aiming to strike a balance between model complexity, training duration, and the avoidance of overfitting, ultimately optimizing the accuracy and precision of the system:

$$UY_2=\underset{\left\{NH_{Dense},NE_{Dense},NS_{Dense}\right\}}{argmax}\left(\alpha+\lambda\right) \tag{22}$$

As per the given formula, the range of values for variablesNH_{Dense},NE_{Dense}, *and* NS_{Dense}spans [5-255], [5-50], and [100-500], representing the epochs, steps per epoch and number of hidden neurons in DenseNet. Figure (6) shows the proposed view of the novel HCARDNet for depression classification.

DOI: 10.3844/jcssp.2025.524.248

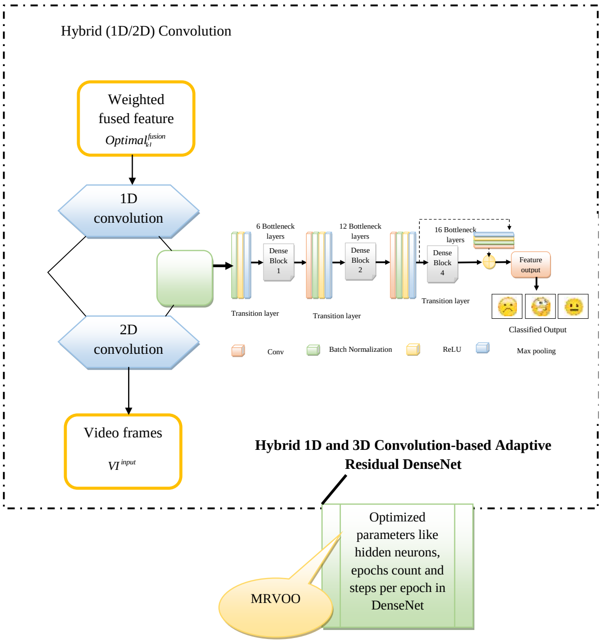


Fig. 6: Proposed view of novel HCARDNet for depression classification

Furthermore, the variable$\alpha$and λ defines precision and accuracy. Equations (23-24), which mathematically express the values of $\alpha$and λin the estimate, provide a comprehensive understanding of these parameters:

$$\alpha = \frac{ff + ll}{ff + ll + mm + uu} \tag{23}$$

$$\lambda = \frac{mm}{mm + ll} \tag{24}$$

Here, the words *ff* and *ll*, *mm*, and *uu* specify true, false positives and false, true negatives.

Results

Simulation Setup

Python facilitated the implementation of the depression categorization system that yielded significant results. The MRVOO method underwent 50 iterations, tailored for a population size of 10 individuals, and involved three chromosomes. The recommended model was leveraged in conjunction with optimization algorithms like Beluga Whale Optimization (BWO)-HCARDNet (Khened *et al.*, 2019), Cuttle Fish Optimization (CO)-HCARDNet (Zhang *et al.*, 2023), Mountaineering Team-Based Optimization (MTBO)-HCARDNet (Daweri *et al.*, 2020) and Osprey Optimization Algorithm (OOA)-HCARDNet (Berntson *et al.*, 2005). To provide a comprehensive comparison, the recommended model was assessed against well-established classifiers, including LSTM (Reddy and Ramanaiah, 2023), RNN (Wu *et al.*, 2021), TCN (Wang *et al.*, 2022) and HCRDNet (Dehghani and Trojovský, 2023).

Efficiency Metrics

The efficiency metrics for the offered approach are provided as follows.

Accuracy: It is validated in Eq. (25):

$$A = \frac{TP + TN}{TP + TN + FP + FN} \tag{25}$$

Precision: It is estimated in Eq. (26):

$$pre = \frac{TP}{TP + TN} \tag{26}$$

F1-Score: It is computed in Eq. (27):

$$F1\ S = 2 \times \frac{TN \times FP}{TN + FP} \tag{27}$$

Sensitivity: It is derived in Eq. (28):

$$Sen = \frac{TP}{TP + TN} \tag{28}$$

FPR: It is measured in Eq. (29):

$$FPR = \frac{FP}{FP + TN} \tag{29}$$

Specificity: It is shown in Eq. (30):

$$spec = \frac{FP}{FP + FN} \tag{30}$$

MCC: It is formulated in Eq. (31):

$$MCC = \frac{TP \times TN - FP \times TN}{\sqrt{(FN + FP)(FN + TN)(TN + FP)(TN + FN)}} \tag{31}$$

FDR: It is computed in Eq. (32):

$$FDR = \frac{FP}{TN + FP} \tag{32}$$

FNR: It is validated in Eq. (33):

$$FNR = \frac{FN}{FN + TN} \tag{33}$$

Recall: It is formatted in Eq. (34):

$$Re = \frac{TN}{TN + FN} \tag{34}$$

In this, the false positive and true positive values are described asTp and FP. The false negative and true negative values are depicted asTN and FN.

Statistical metrics:

- Best: The greatest rate is denoted as the best value
- Worst: The lowest rates are noted as the worst values
- Mean: It is the median rate of both worst and best rates
- Median: It is the central point of the worst and best rates
- Standard deviation: It is the grade of deviation among every implementation

Evaluation of Convergence Validation of the Offered MRVOOAlgorithm

The convergence performance of the recently developed MRVOO method was assessed by comparing its iteration values against those of traditional methods, as depicted in Fig. (7). The MRVOO-HCARDNet strategy exhibited superior performance, surpassing all other approaches. Specifically, it showcased a lead of 65.9% over BWO-HCARDNet, 34.5% over CO-HCARDNet, 89% over MTBO-HCARDNet, and 29% over OOA-HCARDNet at the 20th iteration, as shown in the image. This highlights the efficiency of the proposed MRVOO-HCARDNet technique compared to alternative optimization strategies.

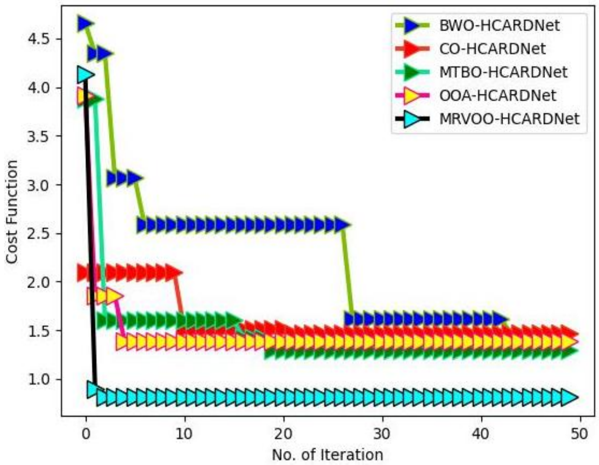


Fig. 7: Convergence validation over diverse optimization algorithms

Validation of Confusion Matrix of the Designed Depression Classification Model

Figure (8) displays a confusion matrix evaluation of the depression categorization paradigm, considering the recommended data sources. Upon calculating accuracy, it is evident that the dataset's accuracy was inferior to that of the current model in use. This validation underscores the effectiveness of the recommended depression categorization model, showcasing its ability to produce more accurate outcomes in the procedure.

Comparative Estimation of the Recommended Depression Classification Model

Examining Figs (9-10), the proposed depression classification system underwent evaluation against various traditional algorithms and classifiers at each level. The assessment of the latest model considered different activation function values. In Fig. (9a), where the accuracy value of the epoch was set to tanh, the offered approach explored superiority over BWO-HCARDNet, CO-HCARDNet, MTBO-HCARDNet, and OOA-HCARDNet by 58.9, 68.5, 54 and 22%.

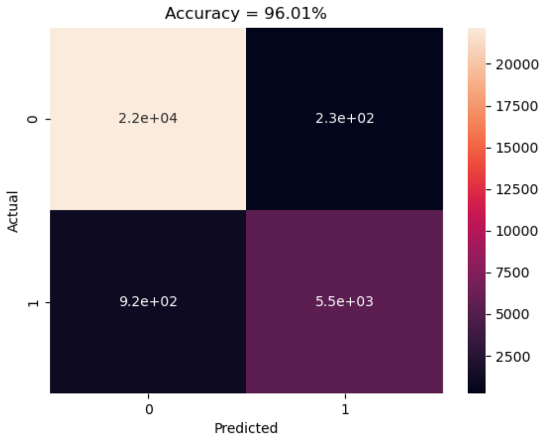
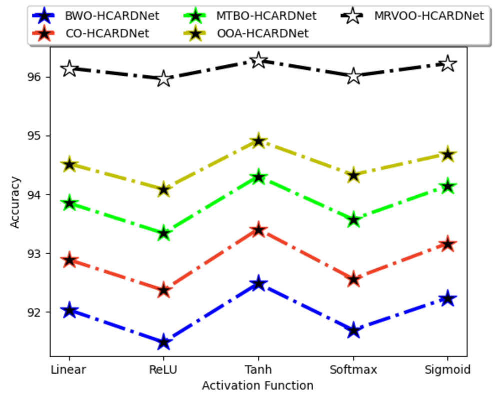
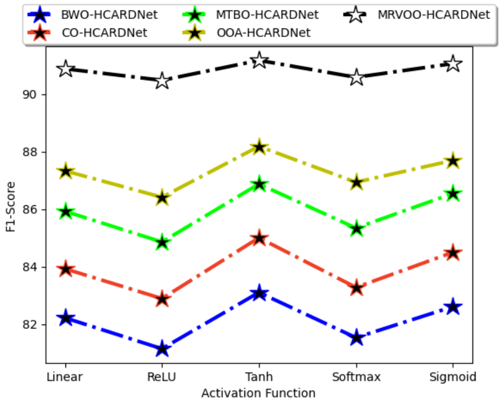


Fig. 8: Estimation of confusion matrix for the offered depression classification model

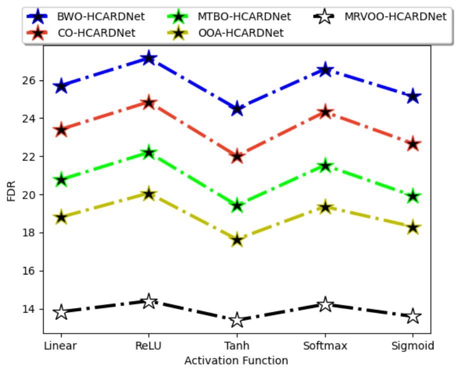


(a)

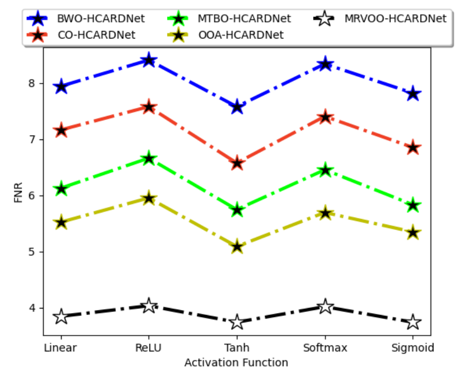
DOI: 10.3844/jcssp.2025.524.248



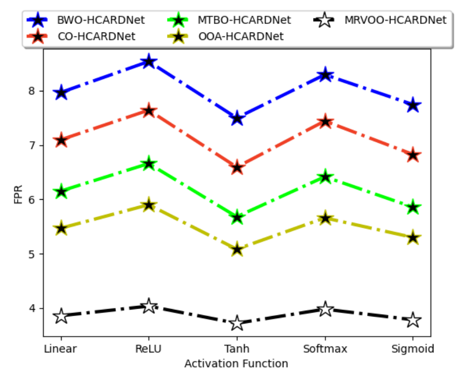
(b)



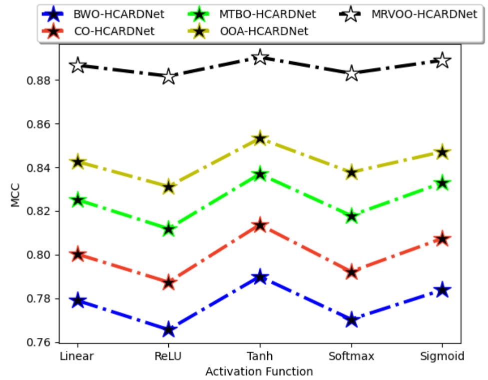
(c)



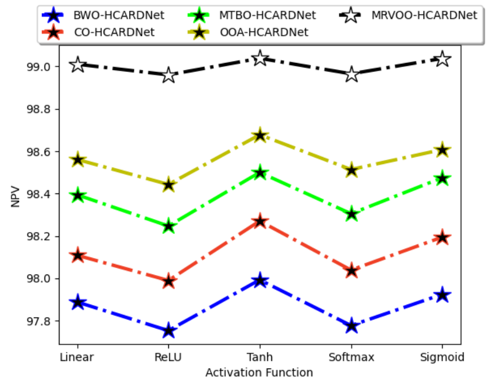
(d)



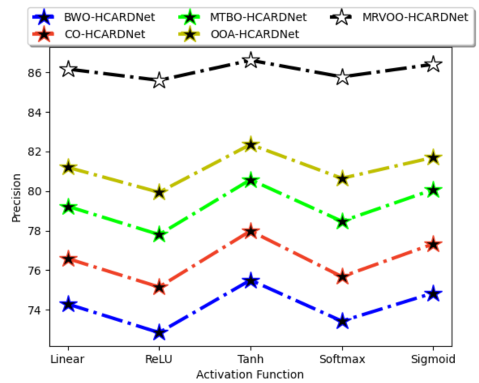
(e)



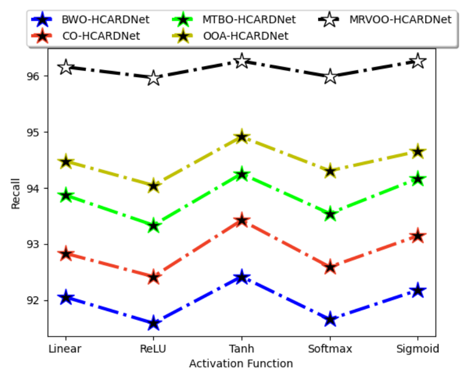
(f)



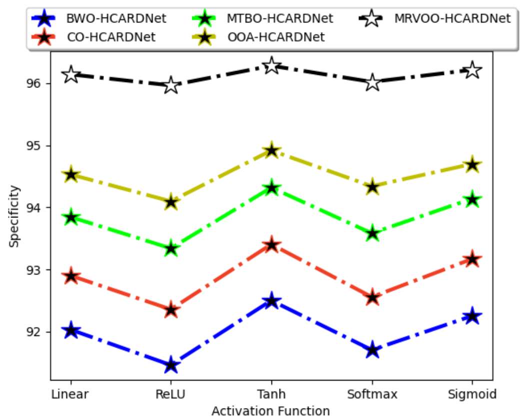
(g)



(h)

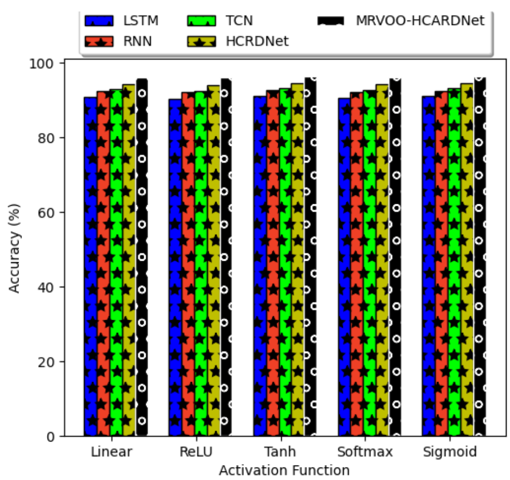


(i)

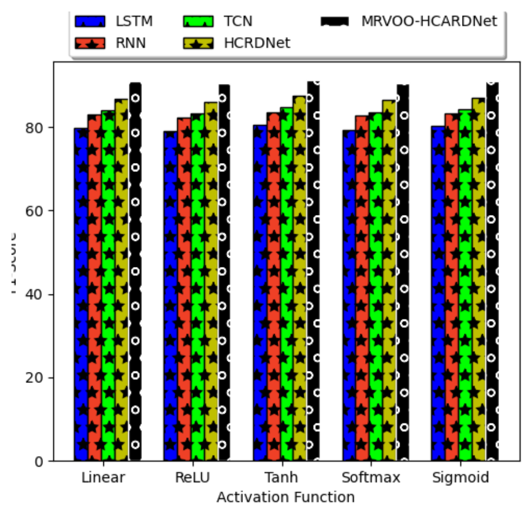


(i)

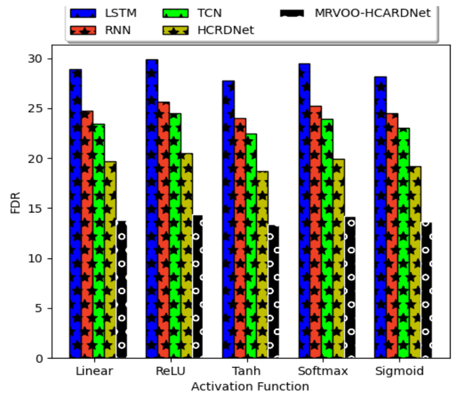
Fig. 9: Estimation of the designed depression classification model contrasts with traditional algorithms over (a) accuracy; (b) F1-score; (c) FDR; (d) FNR; (e) FPR; (f) MCC; (g) NPV; (h) Precision; (i) recall; and (j) Specificity



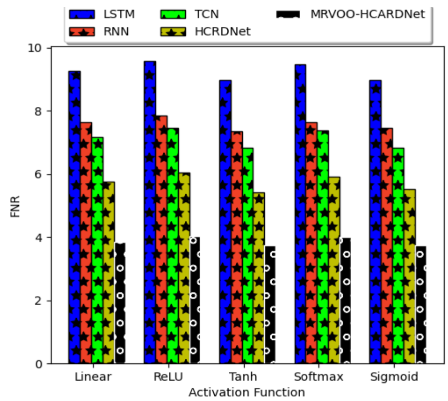
(a)



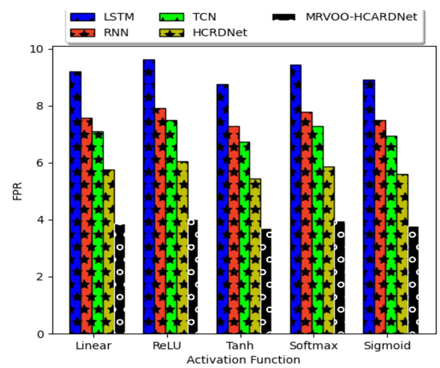
(b)



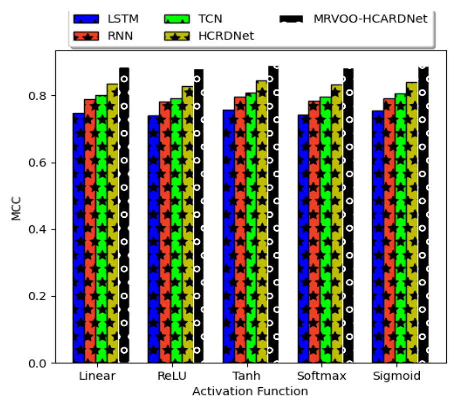
(c)



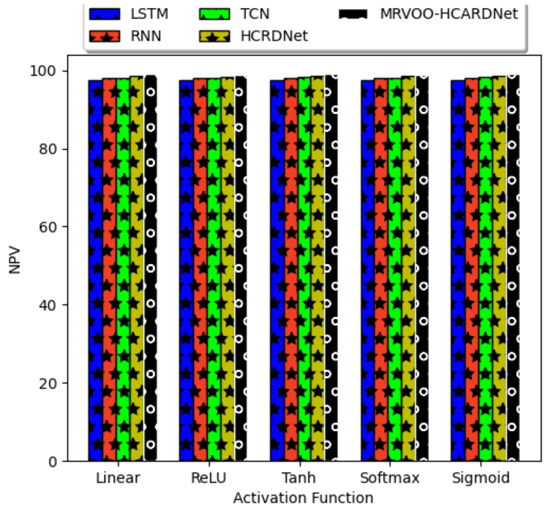
(d)



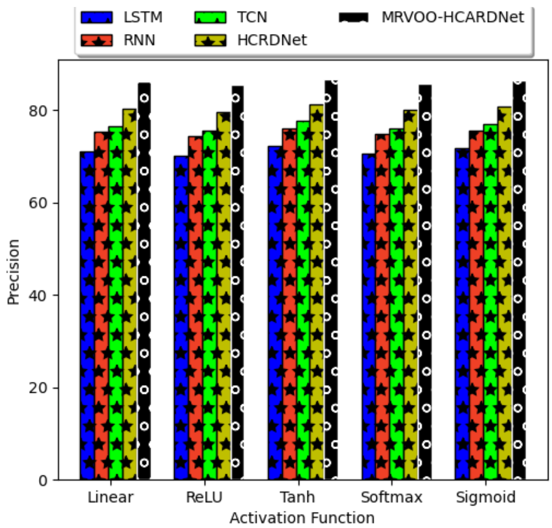
(e)



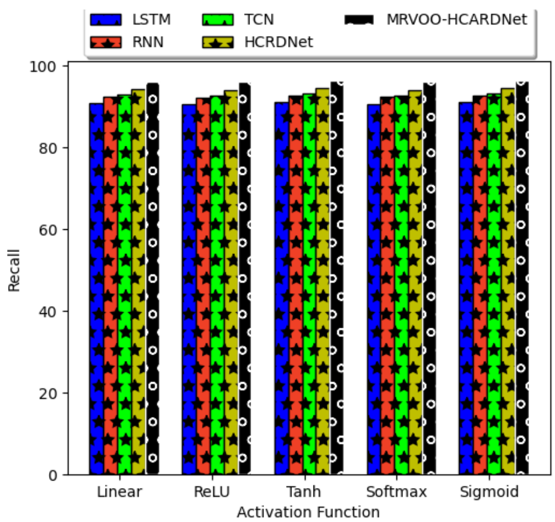
(f)



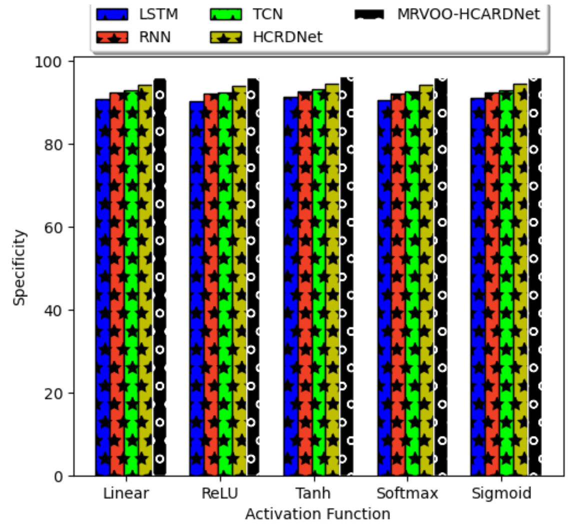
(g)



(h)



(i)



(j)

Fig. 10: Estimation of designed depression classification model contrasts with traditional classifiers over (a) accuracy; (b) F1-score; (c) FDR; (e) FNR; (f) MCC; (g) NPV; (h) Precision; (i) recall and (j) Specificity

Efficiency Validation of the Depression Classification Model

A thorough comparison of the efficiency of the offered method with various traditional methods and classifiers is given in Tables (2-3). The proposed depression categorization model explored precision rates surpassing 33.9% for LSTM, 45.5% for RNN, 67% for TCN, and 49% for HCRDNet. These outcomes emphasize the notable specificity rate and overall effectiveness of the suggested method.

Statistical Evaluation of the Developed Model

Table 4 presents a statistical analysis of the proposed depression classification framework, showcasing significant advancements. At its median value, the model has outperformed BWO-HCARDNet, CO-HCARDNet, MTBO-HCARDNet, and OOA-HCARDNet by 66.9, 33.5, 67 and 39%, respectively. These outcomes signify that, in the realm of depression categorization, the MRVOO-HCARDNet model has achieved substantially higher efficiency compared to baseline methodologies.

Validation of the Selection of Optimal Features

The validation of optimal feature selection on the suggested approach is shown in Fig. (11). The given graph findings visualize the enriched performance of the suggested approach using diverse feature extraction techniques. Here, the BWO algorithm attains the least performance and also the OOA algorithm attains second better performance when compared to the other algorithms.

DOI: 10.3844/jcssp.2025.524.248

Table 2: Overall efficiency validation of the designed model

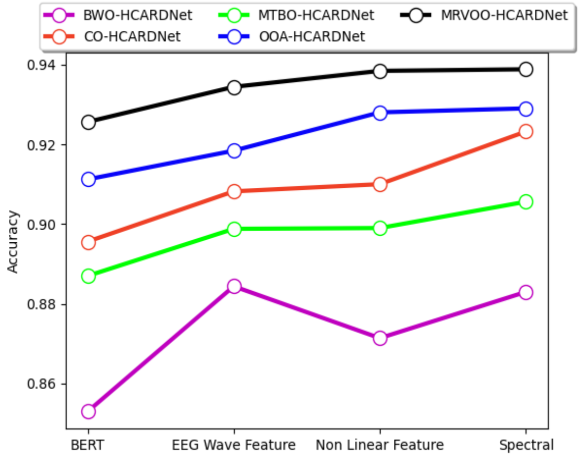
TERMS	BWO-HCARDNet Khened *et al.* (2019)	CO-HCARDNet Zhang *et al.* (2023)	MTBO-HCARDNet Daweri *et al.* (2020)	OOA-HCARDNet Berntson *et al.* (2005)	MRVOO-HCARDNet
Accuracy	92.23953	93.16719	94.14676	94.6902	96.22361
FPR	7.744894	6.827622	5.858429	5.300277	3.785912
FDR	25.15458	22.67241	19.92642	18.29997	13.59329
Specificity	92.25511	93.17238	94.14157	94.69972	96.21409
Precision	74.84542	77.32759	80.07358	81.70003	86.40671
F1-score	82.61207	84.50306	86.55054	87.70045	91.06836
MCC	0.783843	0.8074	0.832859	0.847093	0.888907
Sensitivity	92.17722	93.14642	94.16753	94.65213	96.26168
FNR	7.822776	6.853583	5.832468	5.347871	3.738318
NPV	97.92413	98.19425	98.47477	98.60786	99.03799

Table 3: Efficiency validation depression classification model

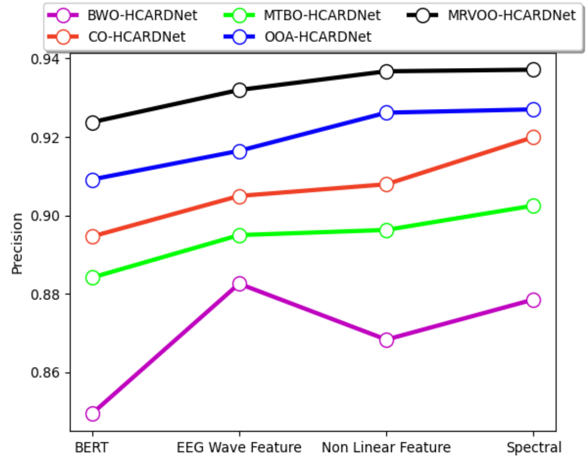
Classifier	LSTM Reddy and Ramanaiah (2023)	RNN Wu *et al.* (2021)	TCN Wang *et al.* (2022)	HCRDNet Dehghani and Trojovský (2023)	MRVOO-HCARDNet
Precision	71.81868	75.51193	76.999	80.82618	86.40671
Sensitivity	91.03496	92.54067	93.16372	94.47906	96.26168
F1-score	80.29309	83.16354	84.31357	87.12097	91.06836
FPR	8.930426	7.502596	6.957425	5.60315	3.785912
FNR	8.96504	7.459328	6.836276	5.520942	3.738318
Specificity	91.06957	92.4974	93.04258	94.39685	96.21409
NPV	97.59807	98.02375	98.19626	98.55891	99.03799
FDR	28.18132	24.48807	23.001	19.17382	13.59329
Accuracy	91.06265	92.50606	93.06681	94.41329	96.22361
MCC	0.754946	0.79078	0.805128	0.839966	0.888907

Table 4: Statistical validation of the offered depression classification method

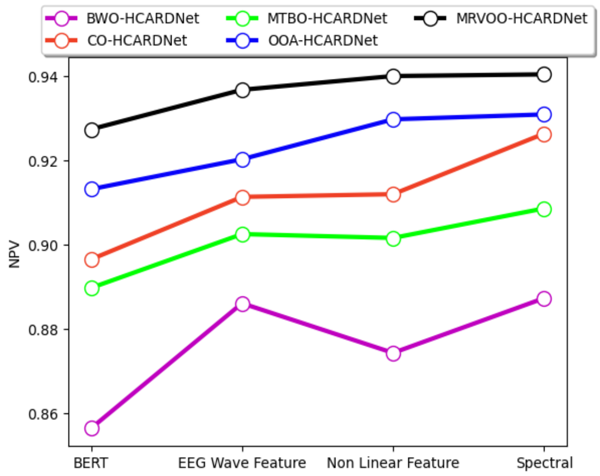
TERMS	BWO-HCARDNet Khened *et al.* (2019)	CO-HCARDNet Zhang *et al.* (2023)	MTBO-HCARDNetDaweri *et al.* (2020)	OOA-HCARDNet Berntson *et al.* (2005)	MRVOO-HCARDNet
Worst	1.330662	1.22442	1.249414	1.268384	1.222884
Median	1.444179	1.241699	1.297576	1.268791	1.285063
Mean	1.418235	1.31261	1.326565	1.333799	1.28624
Standard deviation	0.121119	0.170255	0.119037	0.137862	0.104575
Best	1.919915	1.933691	1.74814	1.828407	1.636788



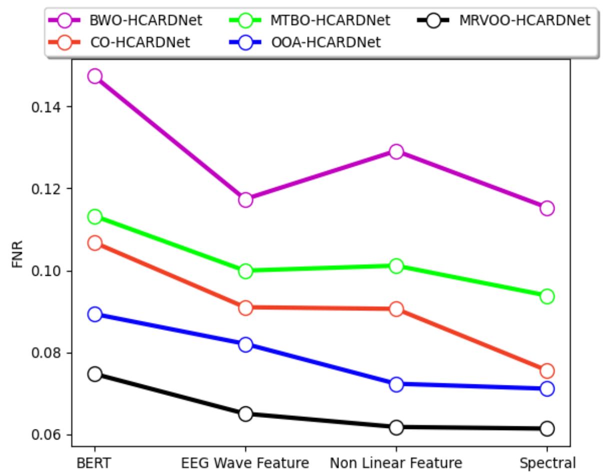
(a)



(b)



(c)



(d)

Fig. 11: Validation of selecting optimal features for the designed depression classification model regarding (a) Accuracy; (b) Precision; (c) NPV; and (d) FNR

Ablation Study

The ablation experiments of the developed model are shown in Table (5). It examines the efficiency of a developed system regarding diverse standard metrics.

Evaluation of Cross-Validation Results

The cross-validation results for the developed model are shown in Fig. (12). The graph findings show the effectiveness of the developed model using positive and negative metrics. The cross-validation results were used to deduce the overfitting issues.

Validation of the Offered Approach Using the New Depression Dataset

Explanation of new dataset: The depression dataset is taken from the link https://www.kaggle.com/datasets/arashnic/the-depression-dataset accessed on 2024-06-17. The danger of a depression is determined by its symptoms. It contains two portfolios one is data for the control and the other one is the condition group. Each patient has a CSV file that contains actigraph data. This dataset contains 23 condition files, 32 control files, and 2.54 kB scores.cv files.

Validation of the Suggested Approach Using Diverse Metrics

The evaluation of the suggested approach using the new depression dataset is shown in Fig. (13). Here, the accuracy and precision of the designed model are validated based on diverse activation functions like Linear, Relu, Tanh, Softmax, and Sigmoid. The given graph results proved that it attains better results than the other baseline approaches.

Table 5: Ablation study of the proposed depression classification model

Methods	ResNet Zhang *et al.*(2024)	Inception Morteza *et al.* (2024)	MobileNet Garg *et al.* (2024)	EfficiecntNet Sait (2024)	MRVOO-HCARDNet
Accuracy	90.20	92.66	90.28	93.80	96.22
Recall	90.31	92.67	90.39	93.77	96.26
Specificity	90.09	92.65	90.17	93.83	96.21
Precision	89.80	92.41	89.88	93.62	86.40
FPR	9.91	7.35	9.83	6.17	3.78
FNR	9.69	7.33	9.61	6.23	3.73
NPV	90.59	92.90	90.67	93.97	99.03
FDR	10.20	7.59	10.12	6.38	13.59
F1-Score	90.06	92.54	90.14	93.70	91.06
MCC	80.40	85.32	80.56	87.60	88.89

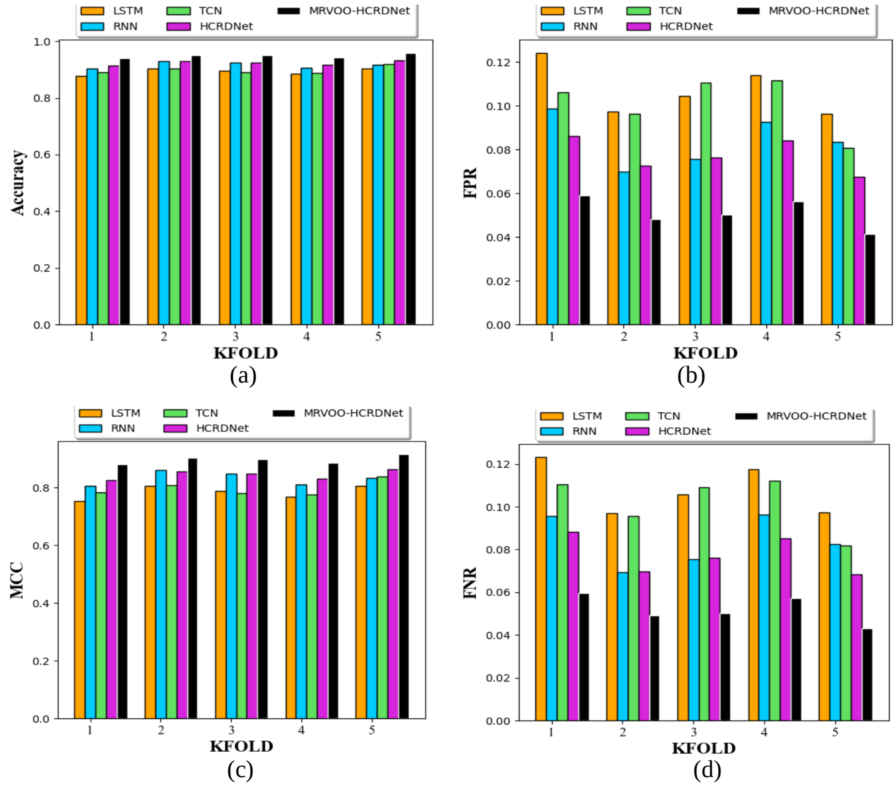


Fig. 12: Estimation of cross-validation results for the designed depression classification model regarding (a) accuracy; (b) FPR; (c) MCC; and (d) FNR

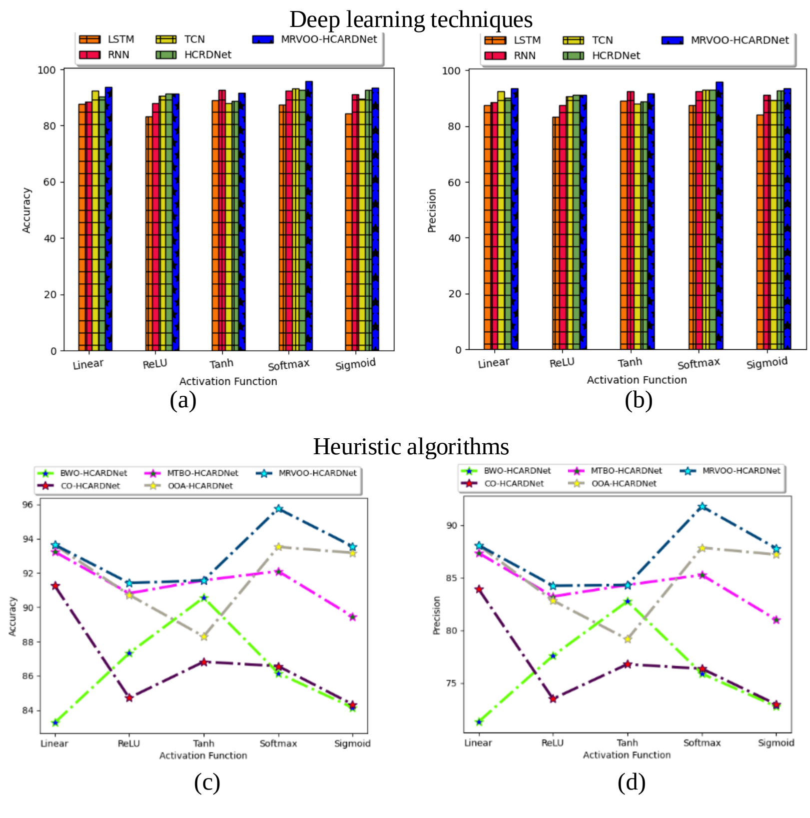


Fig. 13: Computation of the offered approach using the new depression dataset regarding (a) accuracy and (b) precision

Overall Efficiency Validation of the Developed Model Using the New Dataset

The evaluation of the overall efficiency of the suggested approach using the new dataset is given in Tables (6-7). The overall accuracy and specificity rate of the developed model is 96.22 and 96.26%. The validation of the developed model reveals its effective performance.

Comparison of the Suggested Approach with Baseline Models

The comparison of the suggested approach with the baseline models is given in Table (8). The effectiveness of the developed model is estimated according to the standard metrics. The efficiency of the suggested approach is higher than the other baseline methods. The CNN-BLSTM model has attained the worst performance since it does not have the capability to provide better detection results.

Discussion

Focusing on Contributions of the Proposed Work to Existing Literature

The discussion of the developed model is done regarding the attained outcomes. While analyzing Fig. (7), the cost function of the suggested model clearly shows that the given developed model attains a better convergence rate than the other traditional approaches. Section 7.3 visualizes the confusion matrix validation that helps to validate the model's error and calculate the accuracy rate with standard metrics. When taking Figs. (9-10), the validation of the developed model shows its better accuracy and precision rate. The conventional algorithms attain lower performance that suffers from issues like massive complexity and the speed of convergence gradually slows down in the late search period.

Table 6: Overall efficiency analysis of the depression classification models using heuristic algorithms

TERMS	BWO-HCARDNet Khened *et al*. (2019)	CO-HCARDNet Zhang *et al*. (2023)	MTBO-HCARDNet Daweri *et al*. (2020)	OOA-HCARDNet Berntson *et al*. (2005)	MRVOO-HCARDNet
Accuracy	84.15	84.34	89.47	93.18	96.22
Sensitivity	83.87	84.25	89.34	93.21	96.26
Specificity	84.29	84.39	89.53	93.16	96.21
Precision	72.75	72.96	81.01	87.20	86.40
FPR	15.71	15.61	10.47	6.84	3.78
FNR	16.13	15.75	10.66	6.79	3.73
NPV	84.29	84.39	89.53	93.16	99.03
FDR	27.25	27.04	18.99	12.80	13.59
F1-score	77.91	78.20	84.97	90.10	91.06
MCC	66.06	66.49	77.11	85.02	88.89

Table 7: Overall efficiency analysis of the depression classification models using deep learning strategies

Classifier	LSTM Reddy and Ramanaiah (2023)	RNN Wu *et al*. (2021)	TCN Wang *et al*. (2022)	HCRDNet Dehghani and Trojovský (2023)	MRVOO-HCARDNet
Accuracy	84.18	91.19	89.47	92.67	96.22
Sensitivity	84.25	91.23	89.43	92.74	96.26
Specificity	84.15	91.18	89.48	92.64	96.21
Precision	72.66	83.80	80.96	86.30	86.40
FPR	15.85	8.82	10.52	7.36	3.78
FNR	15.75	8.77	10.57	7.26	3.73
NPV	84.15	91.18	89.48	92.64	99.03
FDR	27.34	16.20	19.04	13.70	13.59
F1-score	78.03	87.35	84.98	89.40	91.06
MCC	66.21	80.79	77.13	83.94	88.89

Table 8: Comparison of the depression classification models using baseline models

Classifier	CNN-BLSTM Lilhore *et al*. (2024a)	SVM-RBF Rehmani *et al*. (2024)	IBi-LSTM Lilhore *et al*. (2024b)	LR-FNN Samsel *et al*. (2024)	MRVOO-HCARDNet
Accuracy	92.86	93.62	93.79	94.74	96.22
Recall	92.86	93.62	93.77	94.74	96.26
Specificity	92.86	93.62	93.80	94.74	96.21
Precision	92.35	93.16	93.35	94.35	86.40
FPR	7.14	6.38	6.20	5.26	3.78
FNR	7.14	6.38	6.23	5.26	3.73
NPV	93.35	94.06	94.20	95.11	99.03
FDR	7.65	6.84	6.65	5.65	13.59
F1-score	92.60	93.39	93.56	94.55	91.06
MCC	85.71	87.23	87.56	89.47	88.89

While analyzing Fig. (10), the efficiency of the existing deep learning models attains lower performance which generates exploding gradient and class imbalance issues. The LSTM model attains the least performance that is not applicable for solving the overfitting issues. The elevated outcomes of the developed model prove the enhanced performance of the depression classification model. The overall accuracy and precision rate of the designed model are 96.22 and 86.40%. The statistical analysis of the developed model provides meaningful interpretation and precise results.

Utilization of the proposed model for the clinicians or mental health professionals: The depression classification model helps clinicians or mental health professionals to perform the diagnosis procedures that can provide efficient treatment for depression patients. It is utilized to identify negative thought patterns and depressive symptoms. It prescribes medication in the case of a psychiatrist.

Diagnosis and assessment: It assigns clinical interviews and standardized depression scales to diagnose depression and determine its severity for mental health professionals.

Psychotherapy delivery: It provides diverse kinds of therapy like Cognitive Behavioral Therapy (CBT) and Interpersonal Therapy (IPT) based on the patient's needs.

Treatment planning: It supplies diverse treatment plans, including setting goals and monitoring progress. It helps to train educated patients about depression and its symptoms which can be helpful for earlier diagnosis.

Collaboration with other professionals: The collaboration of primary care physicians helps to manage medication and address urgent issues of patients.

Implications of the depression classification model: Interpersonal losses and stressful events can enlarge the risk of depression. The implemented model suggests a Cognitive Behavioral Theory (CBT) that can suggest that negative thought patterns can influence depression. Markov Decision Process (MDP) helps to optimize sequential treatment to facilitate personalized treatment decisions. Depression diseases mainly affect adult people. CBT is focused on changing negative patterns of behavior that lead to reducing the difficulties in functioning.

Conclusion

The research work aimed to develop a robust depression classification model that leveraged diverse data sources, including text, EEG, and speech signals. This multi-modal strategy aimed to provide a holistic representation of the underlying features associated with depression. After providing the diverse features, the average computation is done to get the better feature vector. The HCARDNet model utilized a combination of hybrid convolutions and residual DenseNet structure to capture spatial and temporal dependencies within the data, facilitating effective learning of intricate depression-related patterns. The parameters of HCARDNet were optimized using the MRVOO approach. The efficiency of the developed model was experimentally estimated by appropriate assessment criteria and industry-standard benchmarks in a crucial phase for evaluating its real-world performance. The precision value of the activation function was set to ReLu, the offered approach surpassed BWO-HCARDNet, CO-HCARDNet, MTBO-HCARDNet, and OOA-HCARDNet by 58.9, 68.5, 54, and 22%. This underscored the better performance of the offered approach compared to other traditional categorization techniques.

Limitations of the Proposed Work

However, the model's performance may be affected by external factors such as cultural nuances, variations in language, or changes in speech patterns, which were not explicitly addressed in the study. The main concern of this study is the potential biases there in the datasets utilized for validation and training. It contains high computational sources that are needed to deploy and train the model. Deep learning algorithms utilized in this study face issues such as complex architectures that demand substantial processing power and memory.

Future Work of this Research Work

Addressing these downsides and exploring the implemented future research directions could contribute to the refinement and applicability of the HCARDNet model for depression classification. In addition, the open-source implementation of the HCARDNet framework will be considered in the upcoming works. We will focus on correlation studies that can help to achieve more detailed outcomes. We will investigate more about the clinical applications of depression models and depressive and non-depressive contents.

Acknowledgment

Thank you to the publisher for their support in the publication of this research article. We are grateful for the resources and platform provided by the publisher, which have enabled us to share our findings with a wider audience. We appreciate the efforts of the editorial team in reviewing and editing our work, and we are thankful for the opportunity to contribute to the field of research through this publication.

Funding Information

This research did not receive any specific funding.

Author's Contributions

Both the authors have equally contributed to this manuscript.

Ethics

This article is original and contains unpublished material. The corresponding author confirms that all of the other authors have read and approved the manuscript and no ethical issues involved.

Reference

Ahmad, B., Usama, M., Huang, C.-M., Hwang, K., Hossain, M. S., & Muhammad, G. (2020). Discriminative Feature Learning for Skin Disease Classification Using Deep Convolutional Neural Network. *IEEE Access*, *8*, 39025–39033. https://doi.org/10.1109/access.2020.2975198

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. https://doi.org/10.3390/e24020211

Berntson, G. G., Lozano, D. L., & Chen, Y. (2005). Filter properties of root mean square successive difference (RMSSD) for heart rate. *Psychophysiology*, *42*(2), 246–252. https://doi.org/10.1111/j.1469-8986.2005.00277.x

Chen, G. K., Kumar, R., Sumbul, H. E., Knag, P. C., &Krishnamurthy, R. K. (2019a). A 4096-Neuron 1M-Synapse 3.8-pJ/SOP Spiking Neural Network With On-Chip STDP Learning and Sparse Weights in 10-nm FinFET CMOS. *IEEE Journal of Solid-State Circuits*, *54*(4), 992–1002. https://doi.org/10.1109/jssc.2018.2884901

Chen, W., Fan, J., Qian, C., Pu, B., Fan, X., & Zhang, G. (2019b). Reliability Assessment of Light-Emitting Diode Packages With Both Luminous Flux Response Surface Model and Spectral Power Distribution Method. *IEEE Access*, *7*, 68495–68502. https://doi.org/10.1109/access.2019.2916878

Chen, Q., Liu, Y.-B., Ge, M.-F., Liu, J., & Wang, L. (2022a). A Novel Bayesian-Optimization-Based Adversarial TCN for RUL Prediction of Bearings. *IEEE Sensors Journal*, *22*(21), 20968–20977. https://doi.org/10.1109/jsen.2022.3209894

Chen, T., Guo, Y., Hao, S., & Hong, R. (2022b). Exploring Self-Attention Graph Pooling With EEG-Based Topological Structure and Soft Label for Depression Detection. *IEEE Transactions on Affective Computing*, *13*(4), 2106–2118. https://doi.org/10.1109/taffc.2022.3210958

Daweri, M. S. A., Abdullah, S., & Ariffin, K. A. Z. (2020). A Migration-Based Cuttlefish Algorithm with Short-Term Memory for Optimization Problems. *IEEE Access*, *8*, 70270–70292. https://doi.org/10.1109/access.2020.2986509

Dehghani, M., & Trojovský, P. (2023). Osprey optimization algorithm: A new bio-inspired metaheuristic algorithm for solving engineering optimization problems. *Frontiers in Mechanical Engineering*, *8*, 1126450. https://doi.org/10.3389/fmech.2022.1126450

Dehghani, M., Torabi1, M., Asayesh, V., & Akhtari-Khosroshahi, S. (2023). Investigation Capability of EEG-based Non-linear Features in Depression Detection. *Research Gate*. https://doi.org/10.13140/RG.2.2.27997.59369

Fumarola, A., Sidler, S., Moon, K., Jang, J., Shelby, R. M., Narayanan, P., Leblebici, Y., Hwang, H., & Burr, G. W. (2018). Bidirectional Non-Filamentary RRAM as an Analog Neuromorphic Synapse, Part II: Impact of Al/Mo/Pr0.7Ca0.3MnO3 Device Characteristics on Neural Network Training Accuracy. *IEEE Journal of the Electron Devices Society*, *6*, 169–178. https://doi.org/10.1109/jeds.2017.2782184

Ganeshkumar, M., Vinayakumar, R., Sowmya, V., Gopalakrishnan, E. A., & Soman, K. P. (2023). Explainable Deep Learning-Based Approach for Multilabel Classification of Electrocardiogram. *IEEE Transactions on Engineering Management*, *70*(8), 2787–2799. https://doi.org/10.1109/tem.2021.3104751

Garg, D., Verma, G. K., & Singh, A. K. (2024). EEG-based emotion recognition using MobileNet Recurrent Neural Network with time-frequency features. *Applied Soft Computing*, *154*, 111338. https://doi.org/10.1016/j.asoc.2024.111338

Garg, S., Shukla, U. P., & Cenkeramaddi, L. R. (2023). Detection of Depression Using Weighted Spectral Graph Clustering With EEG Biomarkers. *IEEE Access*, *11*, 57880–57894. https://doi.org/10.1109/access.2023.3281453

He, Q., Zhao, L., Ji, K., & Kuang, G. (2022). SAR Target Recognition Based on Task-Driven Domain Adaptation Using Simulated Data. *IEEE Geoscience and Remote Sensing Letters*, *19*, 1–5. https://doi.org/10.1109/lgrs.2021.3116707

Jiang, C., Li, Y., Tang, Y., & Guan, C. (2021). Enhancing EEG-Based Classification of Depression Patients Using Spatial Information. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *29*, 566–575. https://doi.org/10.1109/tnsre.2021.3059429

Khan, D. M., Masroor, K., Jailani, M. F. M., Yahya, N., Yusoff, M. Z., & Khan, S. M. (2022). Development of Wavelet Coherence EEG as a Biomarker for Diagnosis of Major Depressive Disorder. *IEEE Sensors Journal*, *22*(5), 4315–4325. https://doi.org/10.1109/jsen.2022.3143176

Khened, M., Kollerathu, V. A., & Krishnamurthi, G. (2019). Fully convolutional multi-scale residual DenseNets for cardiac segmentation and automated cardiac diagnosis using an ensemble of classifiers. *Medical Image Analysis*, *51*, 21–45. https://doi.org/10.1016/j.media.2018.10.004

Kour, H., & Gupta, M. K. (2022). An hybrid deep learning approach for depression prediction from user tweets using feature-rich CNN and bi-directional LSTM. *Multimedia Tools and Applications*, *81*(17), 23649–23685. https://doi.org/10.1007/s11042-022-12648-y

Kwon, D., Lim, S., Bae, J.-H., Lee, S.-T., Kim, H., Kim, C.-H., Park, B.-G., & Lee, J.-H. (2019). Adaptive Weight Quantization Method for Nonlinear Synaptic Devices. *IEEE Transactions on Electron Devices*, *66*(1), 395–401. https://doi.org/10.1109/ted.2018.2879821

Lee, C.-C., Chiang, H.-S., & Hsiao, M.-H. (2021). Effects of screen size and visual presentation on visual fatigue based on regional brain wave activity. *The Journal of Supercomputing*, *77*(5), 4831–4851. https://doi.org/10.1007/s11227-020-03458-w

Li, D., Liu, J., Yang, Y., Hou, F., Song, H., Song, Y., Gao, Q., & Mao, Z. (2023). Emotion Recognition of Subjects With Hearing Impairment Based on Fusion of Facial Expression and EEG Topographic Map. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *31*, 437–445. https://doi.org/10.1109/tnsre.2022.3225948

Li, Q., Yang, Y., Lan, T., Zhu, H., Wei, Q., Qiao, F., Liu, X., & Yang, H. (2020). MSP-MFCC: Energy-Efficient MFCC Feature Extraction Method With Mixed-Signal Processing Architecture for Wearable Speech Recognition Applications. *IEEE Access*, *8*, 48720–48730. https://doi.org/10.1109/access.2020.2979799

Lilhore, U. K., Dalal, S., Faujdar, N., Simaiya, S., Dahiya, M., Tomar, S., & Hashmi, A. (2024a). Unveiling the prevalence and risk factors of early stage postpartum depression: a hybrid deep learning approach. *Multimedia Tools and Applications*, *83*(26), 68281–68315. https://doi.org/10.1007/s11042-024-18182-3

Lilhore, U. K., Dalal, S., Varshney, N., Sharma, Y. K., Rao, K. B. V. B., Rao, V. V. R. M., Alroobaea, R., Simaiya, S., Margala, M., & Chakrabarti, P. (2024b). Prevalence and risk factors analysis of postpartum depression at early stage using hybrid deep learning model. *Scientific Reports*, *14*(1), 4533. https://doi.org/10.1038/s41598-024-54927-8

Morteza, S. M., Shalbaf, R., & Shalbaf, A. (2024). Prediction of treatment response in major depressive disorder using a hybrid of convolutional recurrent deep neural networks and effective connectivity based on EEG signal. *Physical and Engineering Sciences in Medicine*, *47*(2), 633–642. https://doi.org/10.1007/s13246-024-01392-2

Moussa, M. M., Alzaabi, Y., & Khandoker, A. H. (2022). Explainable Computer-Aided Detection of Obstructive Sleep Apnea and Depression. *IEEE Access*, *10*, 110916–110933. https://doi.org/10.1109/access.2022.3215632

Muzammel, M., Salam, H., & Othmani, A. (2021). End-to-end multimodal clinical depression recognition using deep neural networks: A comparative analysis. *Computer Methods and Programs in Biomedicine*, *211*, 106433. https://doi.org/10.1016/j.cmpb.2021.106433

Muzammel, M., Salam, H., Hoffmann, Y., Chetouani, M., & Othmani, A. (2020). AudVowelConsNet: A phoneme-level based deep CNN architecture for clinical depression diagnosis. *Machine Learning with Applications*, *2*, 100005. https://doi.org/10.1016/j.mlwa.2020.100005

Ogunleye, B., Sharma, H., & Shobayo, O. (2024). Sentiment Informed Sentence BERT-Ensemble Algorithm for Depression Detection. *Big Data and Cognitive Computing*, *8*(9), 112. https://doi.org/10.3390/bdcc8090112

Othmani, A., & Zeghina, A. O. (2022). A multimodal computer-aided diagnostic system for depression relapse prediction using audiovisual cues: A proof of concept. *Healthcare Analytics*, *2*, 100090. https://doi.org/10.1016/j.health.2022.100090

Othmani, A., Kadoch, D., Bentounes, K., Rejaibi, E., Alfred, R., & Hadid, A. (2021). Towards Robust Deep Neural Networks for Affect and Depression Recognition from Speech. *Pattern Recognition. ICPR International Workshops and Challenges*, 5–19. https://doi.org/10.1007/978-3-030-68790-8_1

Othmani, A., Zeghina, A.-O., & Muzammel, M. (2022). A Model of Normality Inspired Deep Learning Framework for Depression Relapse Prediction Using Audiovisual Data. *Computer Methods and Programs in Biomedicine*, *226*, 107132. https://doi.org/10.1016/j.cmpb.2022.107132

Pérez-Toro, P. A., Vásquez-Correa, J. C., Bocklet, T., Nöth, E., & Orozco-Arroyave, J. R. (2023). User State Modeling Based on the Arousal-Valence Plane: Applications in Customer Satisfaction and Health-Care. *IEEE Transactions on Affective Computing*, *14*(2), 1533–1546. https://doi.org/10.1109/taffc.2021.3112543

Reddy, Y. V. K., & Ramanaiah, M. L. (2023). Solving Economic Load Dispatch Problem with Mountaineering Team-Based Optimization Technique. *2023 Second International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT)*, 1–6. https://doi.org/10.1109/iceeict56924.2023.10157715

Rehmani, F., Shaheen, Q., Anwar, M., Faheem, M., & Bhatti, S. S. (2024). Depression detection with machine learning of structural and non-structural dual languages. *Healthcare Technology Letters*, *11*(4), 218–226. https://doi.org/10.1049/htl2.12088

Rejaibi, E., Komaty, A., Meriaudeau, F., Agrebi, S., & Othmani, A. (2022). MFCC-based Recurrent Neural Network for automatic clinical depression recognition and assessment from speech. *Biomedical Signal Processing and Control*, *71*, 103107. https://doi.org/10.1016/j.bspc.2021.103107

Rizwan, M., Mushtaq, M. F., Akram, U., Mehmood, A., Ashraf, I., & Sahelices, B. (2022). Depression Classification from Tweets Using Small Deep Transfer Learning Language Models. *IEEE Access*, *10*, 129176–129189. https://doi.org/10.1109/access.2022.3223049

Saga, T., Tanaka, H., Iwasaka, H., & Nakamura, S. (2022). Multimodal Prediction of Social Responsiveness Score with BERT-Based Text Features. *IEICE Transactions on Information and Systems*, *E105. D*(3), 578–586. https://doi.org/10.1587/transinf.2021hcp0009

Sait, A. R. W. (2024). A LeViT–EfficientNet-Based Feature Fusion Technique for Alzheimer's Disease Diagnosis. *Applied Sciences*, *14*(9), 3879. https://doi.org/10.3390/app14093879

Sam, A., Boostani, R., Hashempour, S., Taghavi, M., & Sanei, S. (2023). Depression Identification Using EEG Signals via a Hybrid of LSTM and Spiking Neural Networks. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *31*, 4725–4737. https://doi.org/10.1109/tnsre.2023.3336467

Samsel, K., Tiwana, A., Ali, S., Sadeghi, A., Guergachi, A., Keshavjee, K., Noaeen, M., & Shakeri, Z. (2024). Predicting Depression among Canadians At-Risk or Living with Diabetes Using Machine Learning. *Artificial Intelligence in Medicine*. https://doi.org/10.1101/2024.02.03.24302303

Schubert, E., Wolfe, J., & Tarnopolsky, A. (2004). *Spectral centroid and timbre in complex, multiple instrumental textures*. Proceedings of the 8th International Conference on Music Perception and Cognition, North Western University, Illinois.

Seal, A., Bajpai, R., Agnihotri, J., Yazidi, A., Herrera-Viedma, E., & Krejcar, O. (2021). DeprNet: A Deep Convolution Neural Network Framework for Detecting Depression Using EEG. *IEEE Transactions on Instrumentation and Measurement*, *70*, 1–13. https://doi.org/10.1109/tim.2021.3053999

Shen, J., Zhang, Y., Liang, H., Zhao, Z., Zhu, K., Qian, K., Dong, Q., Zhang, X., & Hu, B. (2023). Depression Recognition From EEG Signals Using an Adaptive Channel Fusion Method via Improved Focal Loss. *IEEE Journal of Biomedical and Health Informatics*, *27*(7), 3234–3245. https://doi.org/10.1109/jbhi.2023.3265805

Shi, C., Miao, F., Jin, Z., & Xia, Y. (2019). Target Recognition of Synthetic Aperture Radar Images Based on Matching and Similarity Evaluation Between Binary Regions. *IEEE Access*, *7*, 154398–154413. https://doi.org/10.1109/access.2019.2948839

Srujan, K. S., Nikhil, S. S., Raghav Rao, H., Karthik, K., Harish, B. S., & Keerthi Kumar, H. M. (2018). Classification of Amazon Book Reviews Based on Sentiment Analysis. *Information Systems Design and Intelligent Applications*, 401–411. https://doi.org/10.1007/978-981-10-7512-4_40

Suen, P. J. C., Goerigk, S., Razza, L. B., Padberg, F., Passos, I. C., & Brunoni, A. R. (2021). Classification of unipolar and bipolar depression using machine learning techniques. *Psychiatry Research*, *295*, 113624. https://doi.org/10.1016/j.psychres.2020.113624

Swasthika Jain, T. J., Jacob, I. J., & Mandava, A. K. (2023). D-ResNet-PVKELM: Deep neural network and paragraph vector based kernel extreme machine learning model for multimodal depression analysis. *Multimedia Tools and Applications*, *82*(17), 25973–26004. https://doi.org/10.1007/s11042-023-14351-y

Uddin, M. Z., Dysthe, K. K., Følstad, A., & Brandtzaeg, P. B. (2022). Deep learning for prediction of depressive symptoms in a large textual dataset. *Neural Computing and Applications*, *34*(1), 721–744. https://doi.org/10.1007/s00521-021-06426-4

Villa-Pérez, M. E., Trejo, L. A., Moin, M. B., & Stroulia, E. (2023). Extracting Mental Health Indicators from English and Spanish Social Media: A Machine Learning Approach. *IEEE Access*, *11*, 128135–128152. https://doi.org/10.1109/access.2023.3332289

Wagner, S. A. (2016). SAR ATR by a combination of convolutional neural network and support vector machines. *IEEE Transactions on Aerospace and Electronic Systems*, *52*(6), 2861–2872. https://doi.org/10.1109/taes.2016.160061

Wang, J., Zhang, W., Yang, H., Yeh, C.-C. M., & Wang, L. (2022). Visual Analytics for RNN-Based Deep Reinforcement Learning. *IEEE Transactions on Visualization and Computer Graphics*, *28*(12), 4141–4155. https://doi.org/10.1109/tvcg.2021.3076749

William, D., Achmad, S., Suhartono, D., & Gema, A. P. (2022). Leveraging BERT with Extractive Summarization for Depression Detection on Social Media. *2022 International Seminar on Intelligent Technology and Its Applications (ISITIA)*, 63–68. https://doi.org/10.1109/isitia56226.2022.9855370

Wu, H., Ma, X., Yang, C.-H. H., & Liu, S. (2021). Attention Based Bidirectional Convolutional LSTM for High-Resolution Radio Tomographic Imaging. *IEEE Transactions on Circuits and Systems II: Express Briefs*, *68*(4), 1482–1486. https://doi.org/10.1109/tcsii.2020.3039526

Xia, L., Feng, Y., Guo, Z., Ding, J., Li, Y., Li, Y., Ma, M., Gan, G., Xu, Y., Luo, J., Shi, Z., & Guan, Y. (2023). MuLHiTA: A Novel Multiclass Classification Framework with Multibranch LSTM and Hierarchical Temporal Attention for Early Detection of Mental Stress. *IEEE Transactions on Neural Networks and Learning Systems*, *34*(12), 9657–9670. https://doi.org/10.1109/tnnls.2022.3159573

Yasin, S., Hussain, S. A., Aslan, S., Raza, I., Muzammel, M., & Othmani, A. (2021). EEG based Major Depressive disorder and Bipolar disorder detection using Neural Networks:A review. *Computer Methods and Programs in Biomedicine*, *202*, 106007. https://doi.org/10.1016/j.cmpb.2021.106007

Yasin, S., Othmani, A., Raza, I., & Hussain, S. A. (2023). Machine learning based approaches for clinical and non-clinical depression recognition and depression relapse prediction using audiovisual and EEG modalities: A comprehensive review. *Computers in Biology and Medicine*, *159*, 106741. https://doi.org/10.1016/j.compbiomed.2023.106741

Zeberga, K., Attique, M., Shah, B., Ali, F., Jembre, Y. Z., & Chung, T.-S. (2022). A Novel Text Mining Approach for Mental Health Prediction Using Bi-LSTM and BERT Model. *Computational Intelligence and Neuroscience*, *2022*, 1–18. https://doi.org/10.1155/2022/7893775

Zhang, L., Wang, L., Yu, M., Wu, R., Steffens, D. C., Potter, G. G., & Liu, M. (2024). Hybrid representation learning for cognitive diagnosis in late-life depression over 5 years with structural MRI. *Medical Image Analysis*, *94*, 103135. https://doi.org/10.1016/j.media.2024.103135

Zhang, R., Li, M., Chen, X., & Ding, Y. (2023). UAV Path Planning Method Based on Modified Beluga Whale Optimization. *2023 IEEE International Conference on Unmanned Systems (ICUS)*, 792–797. https://doi.org/10.1109/icus58632.2023.10318254

Zhao, Z., Bao, Z., Zhang, Z., Deng, J., Cummins, N., Wang, H., Tao, J., & Schuller, B. (2020). Automatic Assessment of Depression from Speech via a Hierarchical Attention Transfer Network and Attention Autoencoders. *IEEE Journal of Selected Topics in Signal Processing*, *14*(2), 423–434. https://doi.org/10.1109/jstsp.2019.2955012