Bioscene

Bioscene
Volume- 21 Number- 02
ISSN: 1539-2422 (P) 2055-1583 (O)
www.explorebioscene.com

Wearable IOT-Based Early Detection of Mental Disorders using Fast EXP-GRU from EEG Signals

Neetu Surendran & Vanshika Gupta (Assistant Professor)
Department of Computer Science and engineering, Inderprastha Engineering College, Ghaziabad

Abstract: Mental health issues have become a significant concern with the increasing competitiveness and pressure in today's society. To monitor mental health, smart wearable devices have emerged as valuable tools.However, there is insufficient comprehensive research on the classification of different mental health disorders. For addressing this gap, a Wearable IoT-based approach for Early and real-time Detection of Mental Disorders using FastExpGRU from EEG Signals (WIEDMD-FES) is proposed in this work.From publicly available sources, the approach gathers the Electroencephalogram (EEG) signals and pre-processes them. For analyzing the signals, it applies Blind Source Separation (BSS), bi-spectrum analysis, and 2D-Graph Fourier Transform (2D-GFT). Through Wigner-Ville Intrinsic Time-Scale Decomposition (WiVi-ITSD), the source signal acquired from BSS undergoes signal decomposition. Scalp map generation, band separation, Isoelectric Level Detection (ILD), and Weighted Directed Functional Brain Network (WDFBN) construction are done for the decomposed signal. Next, from the 2D-GFT signal, scalp map, ILD, and WDFBN, the features are extracted. Optimal features are selected by the Adaptively Mutated Bald Eagle Optimization(ADM-BEO) algorithm, and different types of mental disorders are classified by the Fast Exponential Linear Unit-Gated Recurrent Unit(FastExpGRU) model. The proposed approach's efficiency is demonstrated by the experimental evaluation.

Keywords: Ward Method-based K-Nearest Neighbor(Ward-KNN); Wigner-Ville Intrinsic Time-Scale Decomposition (WiVi-ITSD); Adaptively Mutated Bald Eagle Optimization (ADM-BEO); Fast Exponential Linear Unit -Gated Recurrent Unit (FastExp-GRU); 2D- Graph Fourier Transform (2D-GFT).

1. Introduction

Globally, mental health disorders have become a significant concern in recent years.Mental health disorders, namely schizophrenia, bipolar disorder, anxiety, and depression, which affect both individuals' lives and society as a whole, are prevalent globally(Khiani et al., 2022; Pacheco-Lorenzo et al., 2021). The high levels of competition, stress, and pressure in daily life have contributed to this increase in mental health issues(Singh et al., 2022). Thus, identifying mental health issues at

their earliest stages is essential. Timely intervention, appropriate treatment, and support are enabled by early detection. This could prevent the progression of mental disorders and reduce their negative consequences (D'Alfonso, 2020)(Ahmed et al., 2022).Early signs of mental disorders could be subtle and easily overlooked, which makes it necessaryfor developingefficient strategies and tools for their detection (Juneja et al., 2021).

It is challenging to develop efficient approaches for the early detection and intervention of mental disorders, which involves several challenges and considerations (Li, 2023).For addressing this challenge, one promising approach is the use of wearable Internet of Things (IoT) devices (Hussien & Mohialden, 2021).In recent years, these smart wearable devices, namely smartwatches, fitness trackers, and headsets have gained popularity to monitor several aspects of health and well-being (Gutierrez et al., 2021). The advantage of continuous and real-time data collection is offered, which provides valuable insights into an individual's physiological and behavioral patterns (Uban et al., 2021). Great potential for the early detection of mental disorders is held by the integration of wearable IoT devices with mental health monitoring (Arji et al., 2023). Such devices could capture and investigate a range of physiological and behavioral parameters, which are indicative of mental health status (Gomes et al., 2023). For example, valuable information about brain activity and patternsthat can be correlated with specific mental disorders can be provided by EEG signals gathered from wearable EEG headsets(Saito et al., 2022).

But, there is still a lack of comprehensive research on the classification of different types of mental disorders, despite the growing interest in using wearable IoT devices for mental health monitoring (Ji et al., 2022). In existing works, specific disorders and limited sets of features are concentrated.The understanding of the broader spectrum of mental health conditions may be limited by only considering specific disorders(Alwakeel et al., 2023).Thus, for addressing these problems, this paper proposes a novel framework called wearable IoT-based early detection of mental disorders using FastExp-GRU from EEG signals.

1.1. Problem statement

A few common drawbacks related to prevailing approaches are considered below:

- The issue of missing values was not adequately addressed by a few existing approaches, which results in reduced signal quality and potentially inaccurate classification results.

- In traditional techniques, only time-domain, frequency-domain, time-frequency domain, and spatial domain features were concentrated and the significant salient information was discarded.
- For effectively handling the computational complexity related to multi-channel EEG signals, prevailing approaches struggle, which causes longer processing times or resource-intensive computations.
- Prevailing works often concentrated on particular disorders.This limitation can hinder the comprehensive understanding and diagnosis of individuals who experience multiple mental illnesses.

The research approach's major objectives are described below:

- The common problem of missing values in EEG signals is addressed by this approach. It usesthe Ward-KNN algorithm for missing value imputations, thus ensuring a more complete dataset for analysis.
- BSS, bi-spectrum analysis, 2D-GFT analysis, and scalp map generation are included in the proposed approach for capturing relevant features.This technique aims to enhance the model's performance by considering a broader range of information from EEG signals.
- Using WiVi-ITSD, band separation, ILD, and WDFBN construction, the proposed framework recognizes the heterogeneity and complex associations among multi-channel EEG signals. For enhancing the system's generalization capabilities, the framework aims to capture complex associations.
- For classifying different types of mental health disorders, the proposed framework utilizes FastExp-GRU. The framework aims to minimize the chances of misdiagnosis by clinicians by leveraging real-time and accurate detection approaches.

The work is arranged as follows: the associated work concerning the proposed work is analyzed in section 2, the proposed approach is shown in section 3, and the proposed model's performance is analyzed in section 4. Lastly, the paper is concluded with future work in section 5.

2. Literature Survey

(Kim et al., 2022)established an unobtrusive dementia prediction modelto monitor elderly persons' physical activities via passive infrared motion sensors merged with data processing. Deep Neural Networks (DNNs), which predicted the risk of

dementia in a sensor-enabled home, were utilized. The system was non-invasive and cost-effective.However, the model failed for capturing cognitive function, memory, and emotional well-being that were necessary for an accurate dementia prediction.

(Kumar et al., 2021) propoundedmental stress state detectionsystemutilizing sensor-centric bio-signals. It utilized a multi-level DNN with hierarchical learning capabilities of a Convolution Neural Network (CNN). This classified the stress into 3 groups, namely baseline, stress, and amusement.The data quality problems, noise, and sensor variability could affect the model's performance even though the system achieved superior performance.

(Ahmed et al., 2023)explored Deep Hierarchical Attention Active Learning (DHAAL) for Mental Disorder prediction.The depression symptoms in mental health interventions were effectively extracted and classified into 9 diverseclasses by attention-centric active learning with deep entropy. As per the experimental outcomes, a superior F1 score was attained by the hierarchical attention approach. However, this system's generalizability was limited.

(Hassantabar et al., 2022)established anapproachnamedMental HealthDisorder Detection System (MHDDS)grounded on Wearable Sensors along with Artificial Neural Networks (ANN) and diagnosed3significant mental health disorders, namely bipolar, schizoaffective, and major depressive. MHDDS utilized 8 diverse categories of data acquired from sensors incorporated into a smartwatch and smartphone and achieved average accuracy. But, when applied to conditions that have overlapping symptoms, the approach's effectiveness might change.

(Thakur & Roy, 2021) explored mental health centered on the feature variables associatedwith dailyliving behavior utilizing smartphone usage and sensor data. It utilized an Independent-samples t-test, which analogized the variation in means betwixt the healthy group and the group with mental illness. For stress prediction and depression, the approach achieved a higher Area Under the Curve (AUC). Nevertheless, the approach was ineffective towards complex signals.

(Guo et al., 2022) presenteda data fusion for mental health detection. This work used 3 parts, such as physical appearance, academic performance, and representation learning. Lastly, for the final detection, a DNN was used. The approaches' promising performance in comparison to prevailing studies was demonstrated by the extensive results. But, only limited biological markers were considered by the scheme.

(Vaishnavi et al., 2022) propounded a Machine Learning (ML) technique, which identified mental health problems. (i) Logistic Regression, (ii) K-NN Classifier, (iii) Decision Tree Classifier, (iv) Random Forest, and (v) Stacking were the 5 ML

approaches. As per the experimental results, the approach achieved an overall accuracy of prediction of 79%. However, a very minimal dataset was utilized in the research.

(Gouda et al., 2023) utilized ML technique for psychological instability detection. It utilized 4 AI algorithms, such as Decision Tree, Logistic Regression, XGBoost, and KNN. The work's efficiency regarding dysfunctional behavior arrangement was exhibited by the results.Here, centered on actions and thoughts, mental disorders were established. However, the particular features or else variables utilized for prediction weren't completely captured.

(Jin et al., 2020) exploredan attention-centric block deep learning architecture within the device for multi-feature classification and fusion analysis. Different domain features' optimum fusion was autonomously acquired. According to the results, mental states were efficiently classified with superior performance by the architecture. But, the mental disorders' complexities weren't sufficiently captured.

(Dai & Ding, 2021) established multi-perception intelligent wearable devices, which monitored mental health. Here,as per the boosting decision fusion approach and ensemble learning idea, a strong classifier was designed. A superior classification effect was attained by the classifiers' superposition fusion outcomes. However, when applied to diverse mental health disorders, the approach's applicability and performance might vary.

3. Proposed Methodology for Early Detection of Mental Disorders Using Wearable IOT Devices

An approach for the early and real-time detection of mental disorders using wearable IoT devices is suggested in the proposed paper. Figure 1 illustrates the proposed work's architecture.

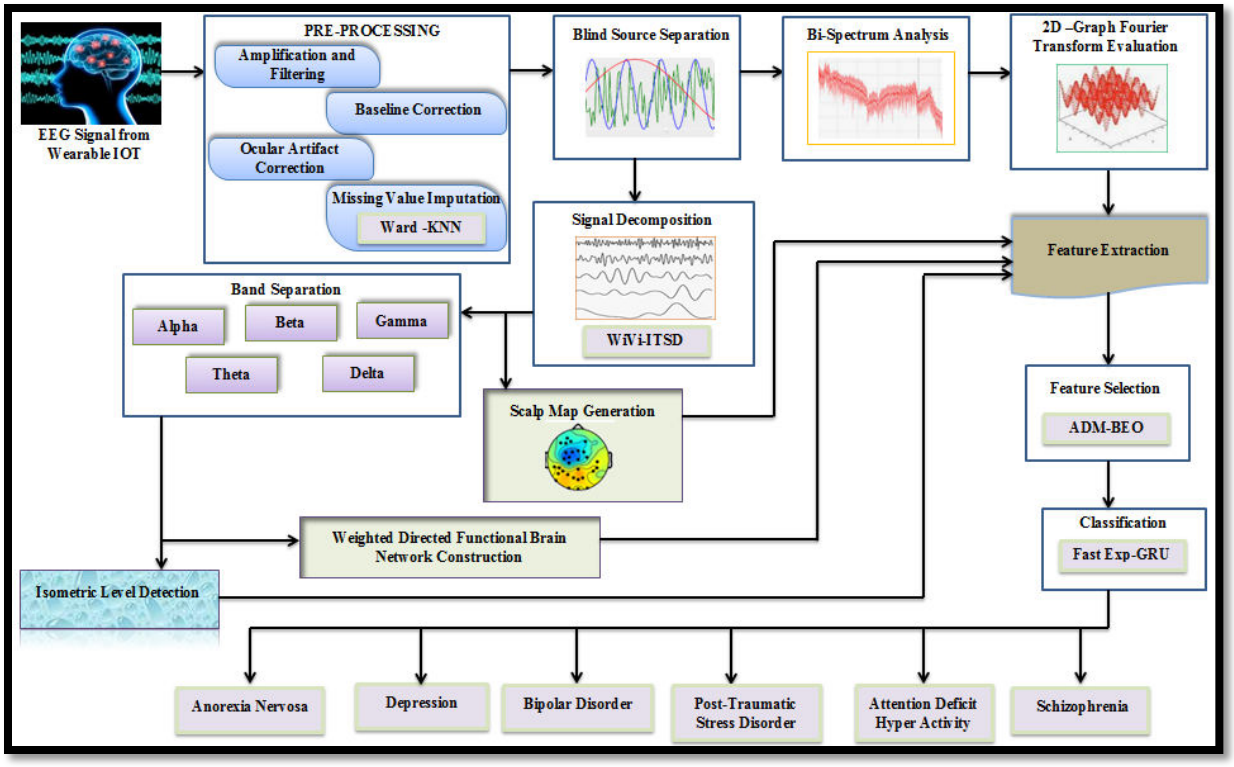


Figure 1: Proposed architecture

3.1. Input Source

EEG signals gathered from wearable devices are considered as the input source. Valuable information about brainwave patterns, which provide the emotional states of an individual, isincluded in the EEG signals recorded by these wearable devices. The EEG signals (E_a) are depicted as,

$$E_a = \{E_1, E_2, E_3, ..., E_A\} \quad a = (1,2,3,.....A) \qquad (1)$$

3.2. Preprocessing

(E_a) undergoes pre-processing; pre-processing function improves the signal quality, eliminates noise, artifacts, and unwanted components, and makes the signal more reliable for further analysis. The proposed approach performs the following pre-processing steps:

Amplification filtering: EEG signals captured by wearable IoT devices are often weak and prone to noise interference. Amplification filtering involves amplifying the signals for increasing their strength when simultaneously applying filtering techniques for removing unwanted noise. The amplified output (α) is,

$$\alpha = AF * E_a \quad (2)$$

Here, the input EEG signal is depicted as E_a, and the amplification factor is signified as AF.

Baseline correction:A process of removing baseline drift or slow variations in the signal is termed baseline correction. Itaidstoremove the effects of gradual changes in the electrode-skin interface or other slow variations unrelated to brain activity. The corrected signal (Cr) is,

$$Cr = E_a - Mean(E_a) \quad (3)$$

Here, the mean value of (E_a) over the desired baseline period is signified as $Mean(E_a)$.

Ocular artifact correction: From eye blinks, eye movements, or other external sources, artifacts in EEG recordings can arise. These artifacts could affect the interpretation of EEG signals. Thus, those ocular artifacts are corrected.

Missing Value Imputation (MVI): Due to electrode disconnections, EEG recordings have missing data points.For estimating or filling in the missing values based on surrounding data points, MVI techniques are applied. Here, for MVI, Ward-KNN is utilized. A versatile algorithm, which can handle various types of data, including continuous, discrete, ordinal, and categorical variables, is named KNN.This makes it well-suited to handle different kinds of missing data. But, when dealing with a large amount of data, especially for MVI tasks, the conventional KNN becomes slower. For overcoming this limitation, this work deploys the Ward Method, which considers the minimum variance distance betwixt the data points. This assists in estimating the missing value centered on the surrounding data points in a way that minimizes the overall variance of the dataset. Hence, the Ward-KNN ensures that the imputed values are consistent with the underlying patterns in the data, which enhances the reliability of the imputation process. The steps of Ward-KNN are,

Step 1: Initialize the number of data points (d_I). In this, the signal's voltage value is depicted as (d_I), and it is given as,

$$d_I = \{d_1, d_2, d_3, ..., d_J\} \quad (4)$$

Step 2: The optimal (k) value is determined; the number of nearest neighbors to consider for imputation is depicted as k. The k value is determined by,

$$k = \sqrt{\rho} \tag{5}$$

Here, the total number of data points in the EEG recordings is signified as ρ .

Step 3: Choose k data points with the shortest distances to the missing values. Such data points become the k nearest neighbors. Utilizing the ward approach, the distance is estimated, which is given by,

$$Dt(d_I, d_J) = \frac{|d_I|}{(|d_I| + |d_J|) * \ell} + \frac{|d_J|}{(|d_I| + |d_J|) * \ell} \tag{6}$$

Here, the Ward method distance between data points d_I and d_J is depicted as $Dt(d_I, d_J)$, the absolute value operator is signified as $|\ |$, and the data points' mean vectors are depicted as ℓ .

Step 4: Next, using the average of the corresponding variable from the K nearest neighbors, the imputing of the missing value is done, which is,

$$Iv = \frac{\sum V(k)}{k} \tag{7}$$

Here, imputing values are signified as Iv , and the sum of variable values from k nearest neighbors is depicted as $V(k)$.

Hence, the preprocessed EEG signals are depicted as,

$$PE_a = \{PE_1, PE_2, PE_3, ..., PE_A\} \tag{8}$$

3.3. Blind source separation

On the preprocessed signals (PE_a), the BSS is done.A signal processing approach utilized for separating original source signals from mixed signals devoid of any former knowledge about the sources or their mixing process is called BSS. The mixed EEG signals can be separated into their underlying source components, which represent the individual brainwave activities originating from different brain regions, by applying BSS.This separation permits for deeper analysis and understanding of the particular brain activities relatedto different mental disorders. The formula for BSS is given as,

$$B_M = Mt * S_M \tag{9}$$

Here, the observed mixed signal matrix is signified as B_M, the mixing matrix is depicted as Mt, and the source signals matrix is indicated as S_M. The S_M is,

$$S_M = (Mt)^{-1} * B_M \quad (10)$$

Here, each row and column represents a different source signal (S) and specific time sample (t), respectively. Once (S) is separated, it undergoes Bi-spectrum analysis and signal decomposition process,correspondingly.

3.3.1. Bi-spectrum analysis

To gain insights into their nonlinear interactions and phase relationships between different frequency components in the brainwaves, theBi-spectrum analysis is done on the separated source signals (S). In exploring the higher-order statistics and phase coupling betwixt different frequency components of the signals, Bi-spectrum analysis helps. It reveals nonlinear interactions and provides a more comprehensive understanding of signal dynamics. The bi-spectrum is computed as,

$$\beta(f_1, f_2) = \left|\varsigma(f_1, f_2)\right|^2 \quad (11)$$

Here, the bi-spectrum (β) at frequencies f_1 and f_2 is depicted as $\beta(f_1, f_2)$, cross power spectrum between f_1 and f_2 is signified as $\varsigma(f_1, f_2)$, and the modulus operation is indicated as $|\ |$.

3.3.1.1. 2D-Graph Fourier Transform

The 2D-GFT is applied to the bi-spectrum values (β) obtained from bi-spectrum analysis. It evaluates the bi-spectrum values' frequency characteristics and transforms them into the 2D frequency domain. It helps to identify the dominant frequency components and their spatial distribution within the bi-spectrum. The 2D-GFT representation is visualized as a graph, where the amplitude or intensity at each point corresponds to the strength of the corresponding frequency component in the bi-spectrum. This representation renders insights into the frequency content and spatial distribution of the nonlinear interactions present in the signal. The 2D-GFT is described as,

$$F(u,v) = \sum\sum \beta(f_1, f_2) * Exp\left(-j(2\pi(u * f_1 + v * f_2)/N)\right) \quad (12)$$

Here, 2D-GFT output at coordinates (u,v) in the frequency domain is signified as $F(u,v)$, bi-spectrum value is depicted as $\beta(f_1, f_2)$, the complex exponential term is

given as $Exp\left(-j(2\pi(u*f_1+v*f_2)/N)\right)$, where, the imaginary unit is indicated as j, and the number of frequency bins is signified as N.

3.3.2. Signal decomposition

Using WiVi-ITSD, the source signals (S) are decomposed. This decomposition helps to separate the source signals into constituent components, allowing for a more detailed analysis of individual components and rendering insights into the time-varying frequency content of the signals. ITSD conserves accurate temporal information about signal critical points and riding waves. However, the ITSD has the limitations of choosing extrema points while there are various successive data points with the same extrema value. For overcoming this issue, the WiVi function is utilized.The WiVi represents the signal using a bilinear time-frequency representation known as Cohen's class of time-frequency distributions and calculates the distribution measure, which identifies and distinguishes the different frequency components present in the signal. Hence, the WiVi-ITSD allows for automated and objective selection of extrema points, which minimizes the reliance on manual perception. The steps involved in WiVi-ITSD are illustrated as,

Extrema point calculation: Extrema points in a signal refer to the locations in which the signal reaches its maximum or minimum values. The extrema points of (S) is estimated utilizing WiVi $WV(T_m, f)$ at time (T_m) and frequency (f) is,

$$WV(T_m, f) = \int \left[S(T_m + \chi/2) * S*(T_m - \chi/2)\right] * Exp^{(-q2\pi f\chi)} d\chi \qquad (13)$$

Here, the input source signal at (T_m) is depicted as $S(T_m)$, the complex conjugate of $S(T_m)$ is indicated as $S*(T_m)$, the time-lag parameter is signified as χ, and the frequency content of the distribution is determined as $Ex^{(-q2\pi f\chi)}$.

The continuous extreme points $[h = 1,2,3,.....H]$ are defined by the interval $[\tau_h, \tau_{h+1}]$.

Baseline Extraction: This aims to separate (S) into 2 components, namely (L) and (H). The baseline component (low-frequency behavior of the signal) is depicted as L. Piecewise linear Residual Component (PRC) (high-frequency in the signal) is signified as H. This isolation process is given as,

$$S = L + H \qquad (14)$$

Piecewise Linear Baseline Extraction Operator: This operator is utilizedfor estimating the baseline signal within each interval between consecutive extreme

points and their time intervals. The σ controls the PRC's amplitude and can be adjusted for balancing the contribution of baseline and PRC.

$$LS = L_h + L_{h+1} - L_h S_{h+1} - S_h(S - S_h) \quad (15)$$

$$L_{h+1} = \sigma\left[S_h + (\tau_{h+1} - \tau_h \tau_{h+2} - \tau_h)(S_{h+2} - S_h)\right] + (1-\sigma)S_{h+1} \quad (16)$$

Here, the baseline extraction signal within the interval $[\tau_h, \tau_{h+1}]$ of continuous extreme points is signified as LS, and coefficients that control the amplitude of the baseline extraction signalare depicted as L_h and L_{h+1}.

Recursive Decomposition: It is the iterative process of decomposing the signal (S) into a series of PRCs and a residual component (H_{rc}), it is expressedas,

$$S = [H_1 + H_2 + H_3 + H_4 + ... + H_{rc}] \quad (17)$$

Hence, the decomposed signal is defined by D_{Sig}.

3.3.2.1. Scalp Map Generation

A process in which visual representations known as scalp maps are generated based on the decomposed signals D_{Sig} is named scalp map generation.These scalp maps give spatial information about the distribution of brain activity across the scalp. When EEG signals are recorded, electrodes are placed on the scalp at specific locations for measuring the electrical activity of the brain.The voltage fluctuations generated by the underlying brain regions are captured by each electrode.The scalp map visualizes the distribution of these voltage fluctuations across different electrode locations on the scalp. For generating scalp maps, D_{Sig} are typically plotted as color-coded topographical maps. A dot represents each electrode's location on the scalp, and the color intensity or contour lines on the map indicate the magnitude or power of the decomposed signal at each electrode location. The scalp map is denoted as SC_{Map}, and it is illustrated in Figure 2.

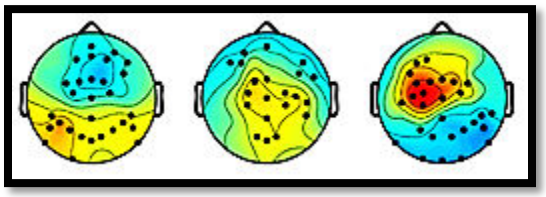


Figure 2: Scalp map

3.3.3. Band separation

The decomposed signals (D_{Sig}) are further separated into diverse frequency bands. These frequency bands, such as alpha, beta, gamma, theta, and delta depict specific ranges of frequencies that are relatedto different brainwave activities. Hence, the brainwave activities within each frequency range can be analyzed.

The frequency ranges relatedtoeach band are mentioned as,

Alpha Band (8-13 Hz): It is often associated with relaxed and calm states and is observed in the posterior regions of the brain.

Beta Band (13-30 Hz): It is relatedto active and alert mental states. It is observed in the frontal and central regions of the brain.

Gamma Band (30-100 Hz): It is associatedwith high-frequency brainwave activity and is associated with complex cognitive processes, including perception, memory, and consciousness. It is observed in various brain regions.

Theta Band (4-8 Hz): It is relatedto relaxation, daydreaming, and creativity. Theta activity is observed in the brain's frontal and temporal regions.

Delta Band (0.5-4 Hz): It is relatedto deep sleep, unconsciousness, and restorative processes. It is observed in the central and posterior regions of the brain.

Mathematically, the band separation is described as,

$$D_{Sig} \rightarrow \{G_{ab}, G_{bb}, G_{gb}, G_{tb}, G_{db}\} \quad (18)$$

After separating (D_{Sig}) into different frequency bands, the Weighted Directed Functional Brain Network Construction (WDFBNC) and the ILD are executed.

3.3.3.1. Weighted Directed Functional Brain Network Construction

Capturing the functional connectivity patterns between different brain regions based on the frequency information from each band $(G_{ab}, G_{bb}, G_{gb}, G_{tb}, G_{db})$ is the aim of WDFBN.It provides a comprehensive view of the functional connectivity patterns between different brain regions across specific frequency bands;thus, important insights into the interactions of brain regions can be examined. The construction of WDFBN involves the succeeding steps:

Define brain regions: First, the specific region of interest within the brain is identified, which is represented as nodes in the network.

Calculate connectivity measures: Based on the frequency-specific information extracted from the EEG signals, the functional connectivity between pairs of brain regions is calculated. This is expressedas,

$$C(f)=\frac{\left|\gamma_{xy}(f)\right|^2}{\gamma_{xx}(f)*\gamma_{yy}(f)} \quad (19)$$

Here, the coherence measure between two signals at frequency (f) is depicted as $C(f)$, the relationship between the signals in the frequency domain is signified as $\gamma_{xy}(f)$, the distribution of signal power across different frequencies is given as $\gamma_{xx}(f)$, and the distribution of signal power across different frequencies is depicted as $\gamma_{yy}(f)$.

Assign weights: The weights are assigned to the edges connecting the nodes in the network. The strength of the functional connectivity between the corresponding brain regions is depicted as the weights (Wt). The weight assignment process is given as,

$$Wt=\frac{C(f)}{Max(C)} \quad (20)$$

Here, the maximum coherence value across all connections is signified as $Max(C)$.

Directionality: The directionality of the edges in the network is determined. Identifying the direction of information flow between brain regions is included in this step. This is described by,

$$\varphi(f)=\varphi_{pq}(f)-\varphi_{pq}(f) \quad (21)$$

Here, the average phase difference at frequency (f) is indicated as $\varphi(f)$.

The edge'sdirectionality is determined by:

If $\varphi(f)>0$, then the edge from node p to node q is considered as directed.

If $\varphi(f)<0$, then the edge from node q to node p is considered as directed.

Network representation: Represent the WDFBN as a graph, which is displayed in Figure 3, where nodes and edges represent brain regions and functional connections betwixt them, respectively. The weights and directionality of the edges are integrated into the network representation.

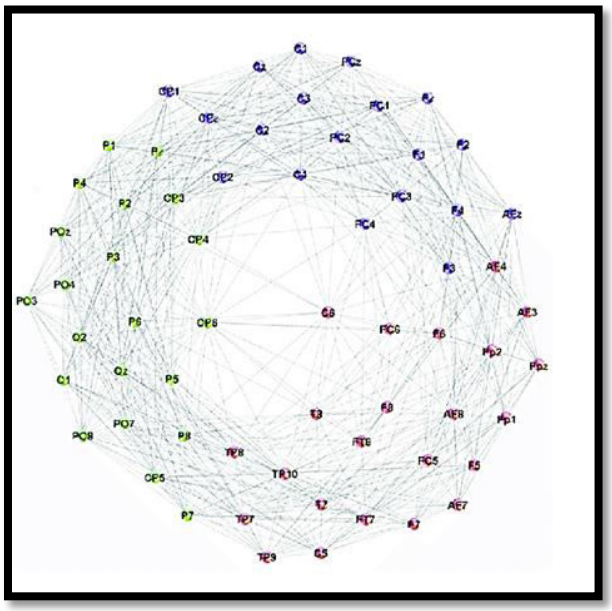


Figure 3: Illustration of WDFBN

3.3.3.2. Isoelectric Level Detection

For identifying the baseline activity in each band $(G_{ab}, G_{bb}, G_{gb}, G_{tb}, G_{db})$, the ILD is done. This baseline represents the absence of significant brainwave activity. By detecting baseline, it becomes possible to differentiate between periods of brainwave activity and without significant activity. This information is useful to analyze the EEG signals and extract relevant features for further analysis. The baseline activity is estimated as,

The time window $T(w)$ is defined, and the average value (ILD_{Avg}) within this time window is calculatedby,

$$ILD_{Avg} = \frac{1}{n(s)} * \sum [T(w)] \tag{22}$$

Here, the number of samples within the time window is signified as $n(s)$, $\sum [T(w)]$ represents the sum of all samples in the window, and the estimated baseline in each band is depicted as ILD_{Avg}. Any deviations or abnormalities from this baseline can

provide important insights into the brainwave activity associated with different mental disorders.

3.4. Feature extraction

From different sources, namely 2D-GFT, Scalp map, ILD, and WDFBN, the significant features are extracted. The extracted features are detailed as,

Sub-band energy, Sub-band fractional energy, Spectral flatness, Spectral brightness, Spectral roll-off, Total EEG energy, Spectral centroid, Spectral spread, Short-time energy, Spectral entropy, Maximal peaks of each sub-band, Spectral flux, and Zero cross rate are extracted from 2D-GFT. These features render the frequency components, energy distribution, and temporal variations relatedtodiverse mental disorders.

The activity power spectrum, Mobility, Fluctuations, Higuchi fractal, Power spectrum density, Lyapunov exponent, Lempel-Ziv complexity, and Teager energy operator are extracted from the Scalp map. Collectively, these features give valuable insights into functional characteristics, irregularities, and abnormalities relatedto mental disorders.

Interhemispheric asymmetry, Positive area, Absolute amplitude, Negative area, Total area, Absolute total area, Cross-correlation, Peak to Peak, RMS amplitude, energy, entropy, energy and variance of auto-covariance, variance, median, mean, skewness, standard deviation, and kurtosis are extracted from ILD. These features render information about the interhemispheric differences, functional connectivity, amplitude characteristics, energy distribution, temporal structure, and statistical properties of brain activity.

Node in-degree, Node out-degree, Node in-strength, Node out-strength, Global efficiency, Local efficiency, Betweenness centrality, Phase Lag Index, Magnitude square coherence, and Modularity are extracted from WDFBN. These features specify the connectivity patterns, information flow, integration, efficiency, centrality, synchronization, and modular organization of the functional brain network.

Hence, the extracted features (Z_F) are depicted as,

$$Z_F = \{Z_1, Z_2, Z_3, ..., Z_z\} \quad (23)$$

3.5. Feature selection

For minimizing the data's dimensionality and enhancing the classification system's efficiency, the selection of the most required features from (Z_F) is essential. For

optimal feature selection, the ADM-BEO algorithm is utilized. The conventional BEO attains superior optimal solutions with a minimal number of iterations. However, it has the drawback of premature convergence towards local optima and limited exploration of the search space. For addressing this problem, the Adaptive Mutation (ADM) factor is incorporated with BEO. The ADM operation disturbs the position of the population, which allows individuals for escaping from local optima and exploring the search space's new regions. So,the search space's more extensive exploration is allowed, which enables the algorithm to discover better optimal solutions that may be located in unexplored regions.The steps of ADM-BEO are,

Selecting stage

Firstly, the bald eagle population is initialized. In this, the extracted features (Z_F) are considered as the Bald eagle. This is given as,

$$Z_F = \{Z_1, Z_2, Z_3, ..., Z_z\} \quad (24)$$

Next, the fitness function (Ft) is defined. In this, the maximum classification accuracy $Max(Cl_{Acc})$ is the fitness function. This is described by,

$$Ft = Max(Cl_{Acc}) \quad (25)$$

The bald eagle selects the best location within the selected search space for the hunting process, and it is mathematically given as,

$$Q = Q_{Best} + g * R(Q_{Mean} - Q_i) \quad (26)$$

Here, a random number from 0 to 1 is indicated as R, g is the parameter that controls the position changes, Q refers to the new position, Q_{Best} denotes the best location, Q_{Mean} represents the mean position of all eagles, and the bald eagle's current position is signified as Q_i.

Searching stage

In this stage, to search for prey within the selected search space, bald eagles use a spiral movement pattern. Furthermore, for enhancing the exploration ability of the bald eagle, the ADM function $(\Im_{ADM})$ is used. The best position during this stage $(Q_{i,New})$ is written as,

$$Q_{i,New} = Q_i + Y(i) * (Q_i - Q_{i+1}) + X(i) * (Q_i - Q_{Mean}) * \Im_{ADM} \quad (27)$$

$$\Im_{ADM} = \hbar_1 * \left(\frac{itr_i}{Max_{itr}} \right) + \hbar_2 \quad (28)$$

$$X(i) = \frac{X\delta(i)}{Max(|X\delta|)}, \qquad Y(i) = \frac{Y\delta(i)}{Max(|Y\delta|)} \quad (29)$$

$$X\delta(i) = \delta(i) * Sin(\theta(i)), \qquad Y\delta(i) = \delta(i) * Cos(\theta(i)) \quad (30)$$

$$\theta(i) = g * \pi * Rand \quad (31)$$

$$\delta(i) = \theta i + \aleph * Rand \quad (32)$$

Here, the next position of the bald eagle is depicted as Q_{i+1}, $X(i)$ and $Y(i)$ represents scaling factors for controlling the movement, $X\delta(i)$ and $Y\delta(i)$ are the values between -1 and 1, $Max(|X\delta|)$ and $Max(|Y\delta|)$ are maximum absolute values of $X\delta(i)$ and $Y\delta(i)$, correspondingly, $\delta(i)$ and $\theta(i)$ are random values between 0 and 1, the parameters for the change in the spiral shape is depicted as $\aleph$, itr_i represents current iteration, Max_{itr} indicates maximum iterations, and the rate at which $(\Im_{ADM})$ increases with each iteration is determined as $\hbar_1$ and $\hbar_2$.

Swooping stage

The bald eagle moves from the selected optimal position in the search space toward its target prey. All other points in the search space also adjust their positions towards the best point, which is,

$$Q_{i,New} = Rand * Q_{Best} + X_1(i) * (Q_i - e_1 * Q_{Mean}) + Y_1(i) * (Q_i - e_2 * Q_{Best}) \quad (33)$$

Where, the parameters that control the movement of the points toward the best pointare signified as e_1 and e_2. Hence, the selected features (λ_F) are described as,

$$\lambda_F = \{\lambda_1, \lambda_2, \lambda_3, ..., \lambda_z\} \quad (34)$$

The Pseudo code for ADM-BEO is given below:

Input: Extracted features (Z_F)

Output: Selection of optimal features (λ_F)

Begin

Initialize the optimization parameters Q_i, Max_{itr}

Calculate the fitness function Ft

For $i = 1$ to Max_{itr} **do**

Select search space by using

$$Q = Q_{Best} + g * R(Q_{Mean} - Q_i)$$

Search the prey in search space by using

$$Q_{i,New} = Q_i + Y(i) * (Q_i - Q_{i+1}) + X(i) * (Q_i - Q_{Mean})$$

Swoop prey by using

$$Q_{i,New} = Rand * Q_{Best} + X_1(i) * (Q_i - e_1 * Q_{Mean}) + Y_1(i) * (Q_i - e_2 * Q_{Best})$$

If $Ft == Satisfied$

Return the optimal features

Else

$$itr = itr + 1$$

End if

End for

End

3.6. Classification

The selected features (λ_F) are inputted into FastExp-GRU.It classifies different types of mental disorders, including (a) Anorexia Nervosa, (b) Depression, (c) Bipolar Disorder, (d) Post Traumatic Stress Disorder, (e) Attention Deficit Hyperactivity Disorder, and (f) Schizophrenia.

GRU has a smaller number of gates, which makes them computationally effectivefor training and use in real-time applications.However, the GRU suffers from an output offset problem that refers to the mismatch between the predicted and actual output, which can affect the accuracy of the system's predictions. The proposed work alleviates this issue by utilizing FELU, which is a non-saturating activation function.FELUenhances the learning capacity of the network and mitigates the output offset problem. Furthermore, the FELU accelerates the network's computations leading to faster training times and potentially improved overall efficiency.Figure 4 shows the architecture of FastExp-GRU.

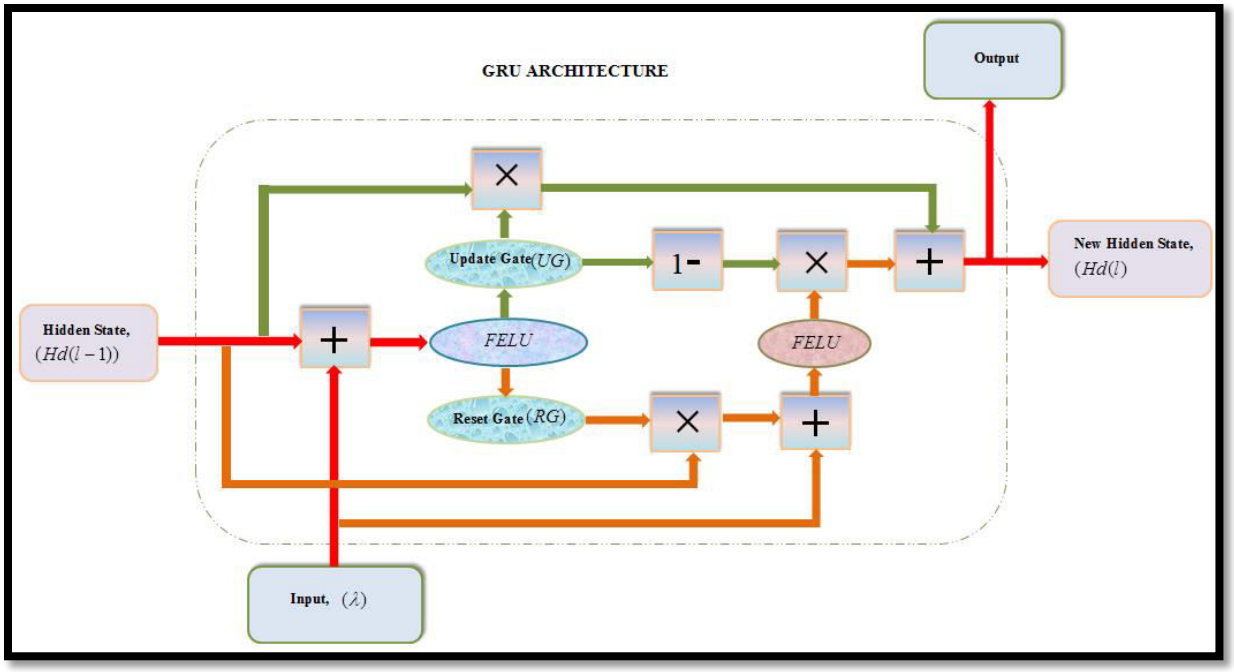


Figure 4: Fast Exp-GRU architecture

The selected features (λ_F) are given as the inputs, which are declared as,

$$\lambda_F = \{\lambda_1, \lambda_2, \lambda_3, ..., \lambda_z\} \quad (35)$$

Update Gate (UG) **:** UG determines the amount of the previous hidden state $(Hd(l-1))$ forward to the current time step (l) .

$$UG(l) = (wg_{UG} * [\lambda(l), Hd(l-1)] + Bias_{UG}) \quad (36)$$

Using the FELU activation function, the UG is computed, and is described as,

$$Exp^{UG(l)} = 2^{\vartheta}\left(1 + UG(l)/Log(2) - \vartheta\right) \quad (37)$$

$$\vartheta = \left[UG(l)/Log(2)\right] \quad (38)$$

Here, wg_{UG} and $Bias_{UG}$ are weight and bias of UG, correspondingly, and scaling factor is depicted as ϑ .

Reset Gate (RG) **:** RG determines how much of the previous hidden state $(Hd(l-1))$ should be ignored while computing the new hidden state candidate (Hd) . It is described as,

$$RG(l) = FELU(wg_{RG} * [\lambda(l), Hd(l-1)] + Bias_{RG}) \quad (39)$$

Here, weight and bias of RG are depicted as wg_{RG} and $Bias_{RG}$, correspondingly.

Hidden State Candidate $(\overline{Hd})$ **:** $(\overline{Hd})$ is the tentative value for the new hidden state at a time step (l). It is calculated by combining the current input and $(Hd(l-1))$, modulated by RG, which is given as,

$$\overline{Hd}(l) = FELU(wg_{\overline{Hd}} * [\lambda(l), RG(l) * Hd(l-1)] + Bias_{\overline{Hd}} \quad (40)$$

Here, $wg_{\overline{Hd}}$ and $Bias_{\overline{Hd}}$ are weight and bias of $(\overline{Hd})$,correspondingly.

Hidden State Update (Hd) **:**The (Hd) is a combination of $(Hd(l-1))$ and the $(\overline{Hd})$, weighted by (UG), which is given as,

$$Hd(l) = (1 - UG(l)) * Hd(l-1) + UG(l) * \overline{Hd}(l) \quad (41)$$

GRU's final output is derived from the updated hidden state, and then the error value (Er) is calculated by,

$$Er = \Pr - Ac \quad (42)$$

If $Error = 0$, then a back-propagation process is not needed. If $Error \neq 0$, then the back-propagation is performed by optimizing the weight values.

The pseudo-code for FastExp-GRU is given as:

Input: Selected features (λ_F)

Output: Classification of mental disorder

Begin

Initialize parameters wg **,** $Bias$,time step (l) , iteration z

Set $z = 1$

While $z \leq z_{\max}$

For $1 to\, z_{Max}$ **do**

Compute update gate UG

$$UG(l) = (wg_{UG} * [\lambda(l), Hd(l-1)] + Bias_{UG})$$

EvaluateFELU activation function using

$$Exp^{UG(l)} = 2^{\vartheta}\left(1 + UG(l)/Log(2) - \vartheta\right)$$

Compute reset gate RG

$$RG(l) = FELU(wg_{RG} * [\lambda(l), Hd(l-1)] + Bias_{RG})$$

Update hidden state Hd

$$Hd(l) = (1 - UG(l)) * Hd(l-1) + UG(l) * \overline{Hd}(l)$$

If $\Pr == Ac$

Terminate

Else

$z = z + 1$

End if

End for

End while

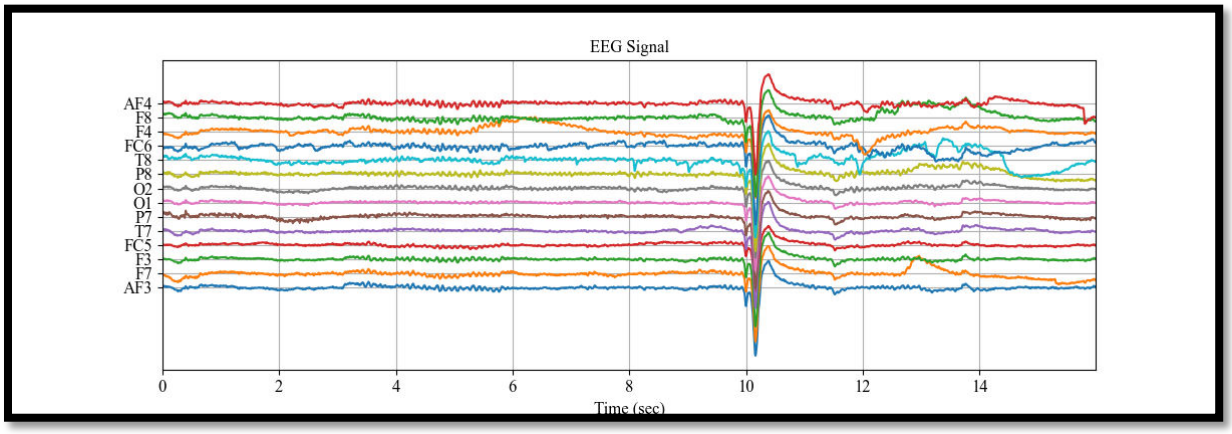
Return mental disorder class

End

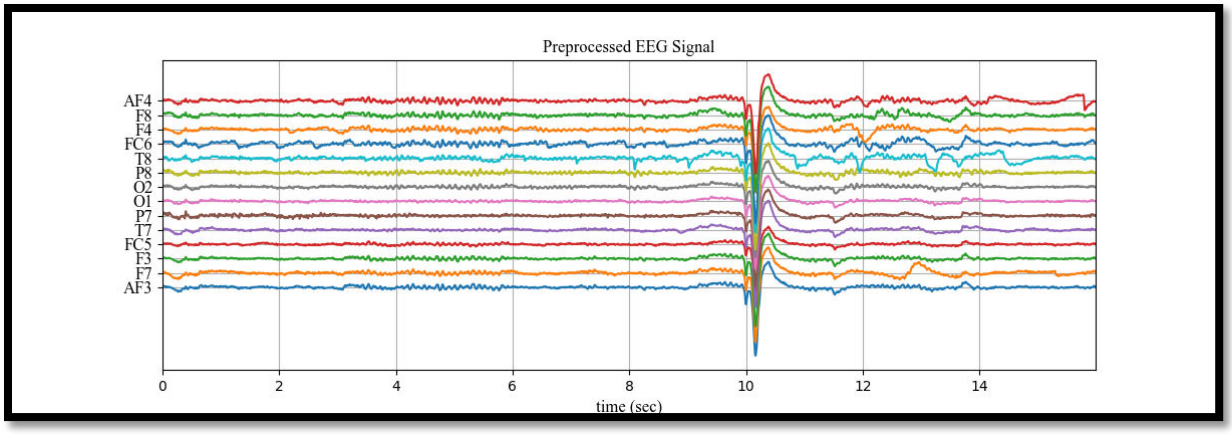
Hence, the FastExp-GRU effectivelycategorizes mental disorders into 6 different classes, namely (i) Anorexia Nervosa, (ii) Depression, (iii) Bipolar Disorder, (iv) Post Traumatic Stress Disorder, (v) Attention Deficit Hyperactivity Disorder, and (vi) Schizophrenia.

4. Results and Discussion

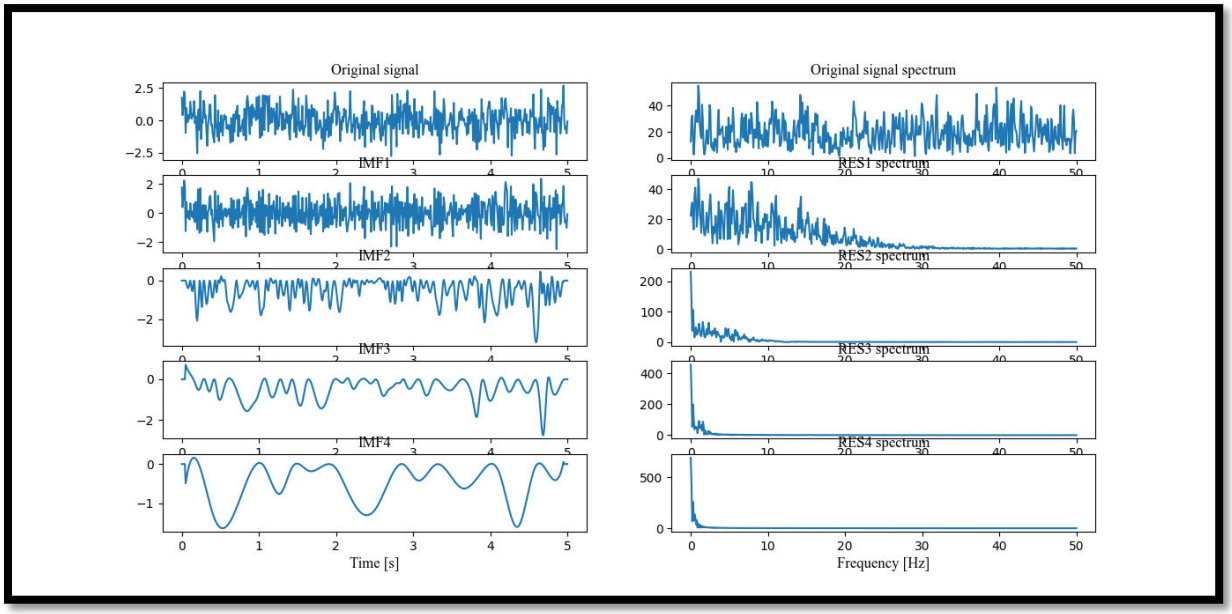
Utilizing publically available data, the proposed approach's performance is evaluated and in the working platform of PYTHON, the experiments were done.



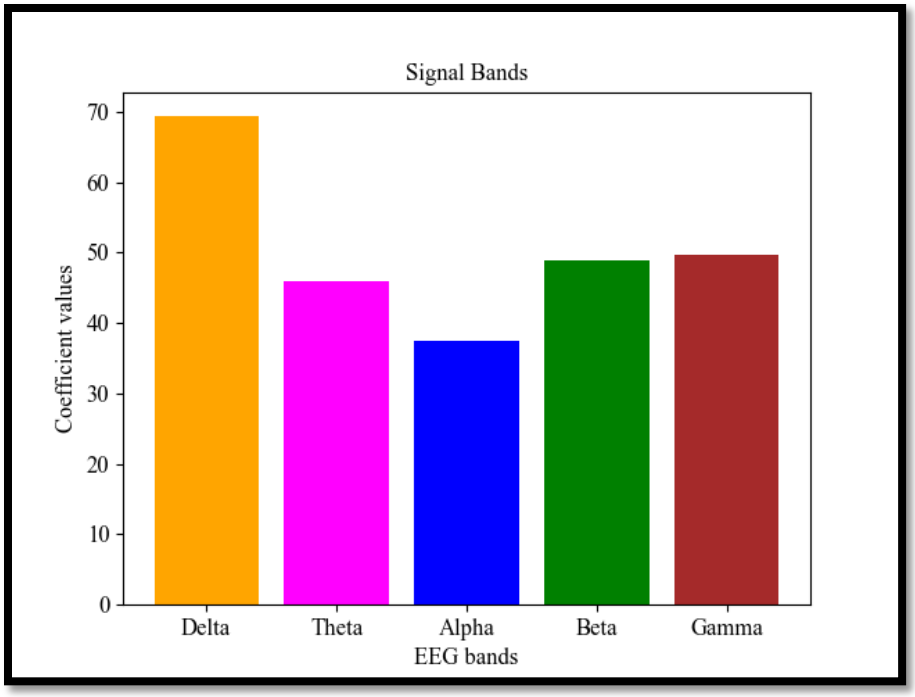
(a)


Preprocessed EEG Signal
AF4
F8
F4
FC6
T8
P8
O2
O1
P7
T7
FC5
F3
F7
AF3
0
2
4
6
8
10
12
14
time (sec)

(b)


Original signal
IMF1
IMF2
IMF3
IMF4
Time [s]
Original signal spectrum
RES1 spectrum
RES2 spectrum
RES3 spectrum
RES4 spectrum
Frequency [Hz]

(c)



(d)

Figure 5: EEGSignal analysis of the proposed WIEDMD-FES

The outcome of the respective phases of the proposed WIEDMD-FES is exhibited in Figure 5. The input EEG signal is depicted in Figure 5(a), a pre-processed signal is represented in Figure 5(b), the decomposed signal is displayed in Figure 5(c), and lastly, the signal is demonstrated with respect to several frequency bands in figure 5 (d).

4.1. Performance Analysis

Here, the proposed WIEDMD-FES'sperformanceis validated.In this, for proving the proposed approach's effectiveness,the performance and comparative analysis areperformed.

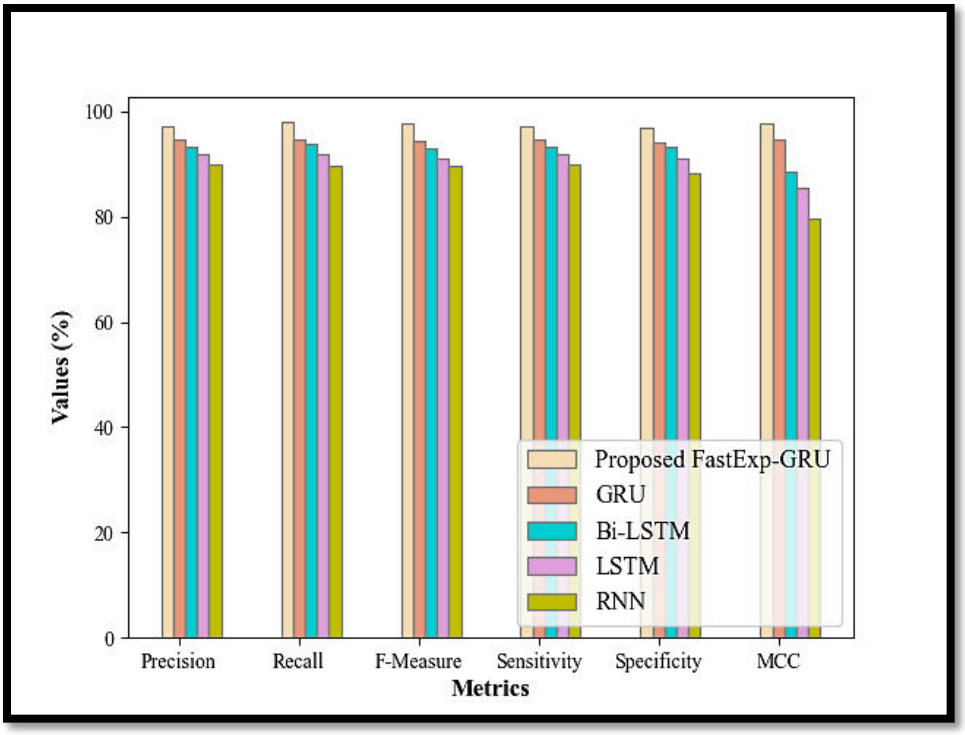


Figure 6: Comparative analysis of the proposed FastExp-GRU

The proposed FastExp-GRU's performance is analogized with prevailing approaches, namely GRU, Bidirectional Long Short-Term Memory (Bi-LSTM), LSTM, and Recurrent Neural Network (RNN) in Figure 6.The proposed FastExp-GRU is incorporated with FELU. The FELU's property includes non-linearity and avoidance of saturation contributing to more efficient learning. Hence, the FastExp-GRU attained precision, recall, F-Measure, Sensitivity, Specificity, and MCC at the rate of 97%, 97.92%, 97.81%, 97%, 96.98%, and 97.71%, correspondingly.However, the prevailing approaches acquire lower performance rates.

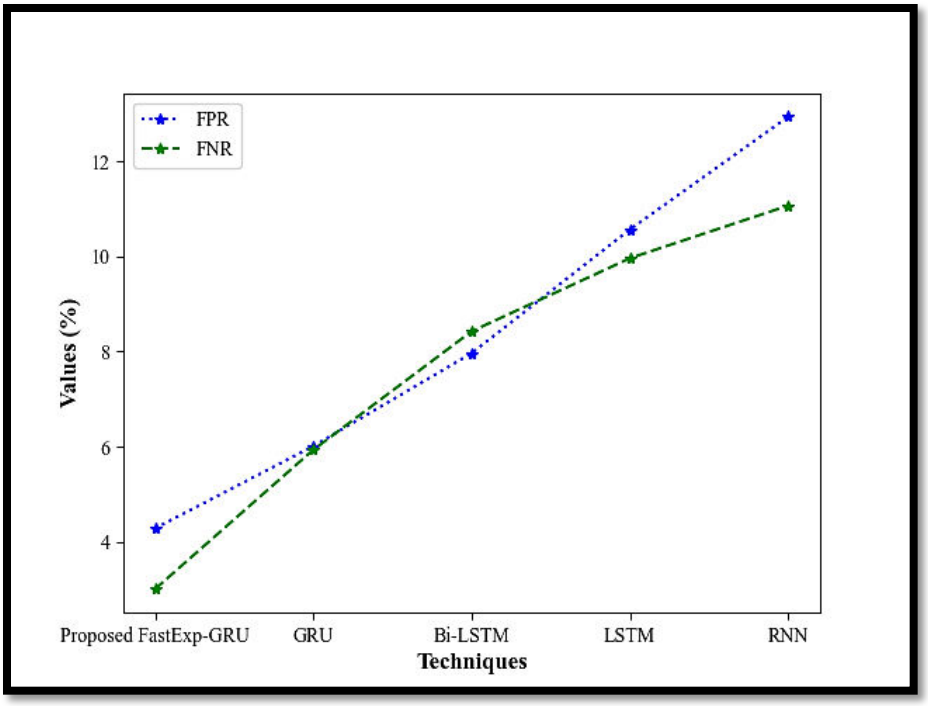


Figure 7: FPR and FNR comparison of the proposed FastExp-GRU

The False Positive Rate (FPR) and False Negative Rate (FNR) of the proposed FastExp-GRU and the other prevailingstudies are contrasted in Figure 7. FELU enables the system to learn more intricate decision boundaries that can help reduce error by enhancing the model's ability for capturing and representing non-linear patterns in the data. Hence, the proposed FastExp-GRU withstands 4.28% of FPR and 3% of FNR.Therefore, the proposed FastExp-GRU is an error-prone model.

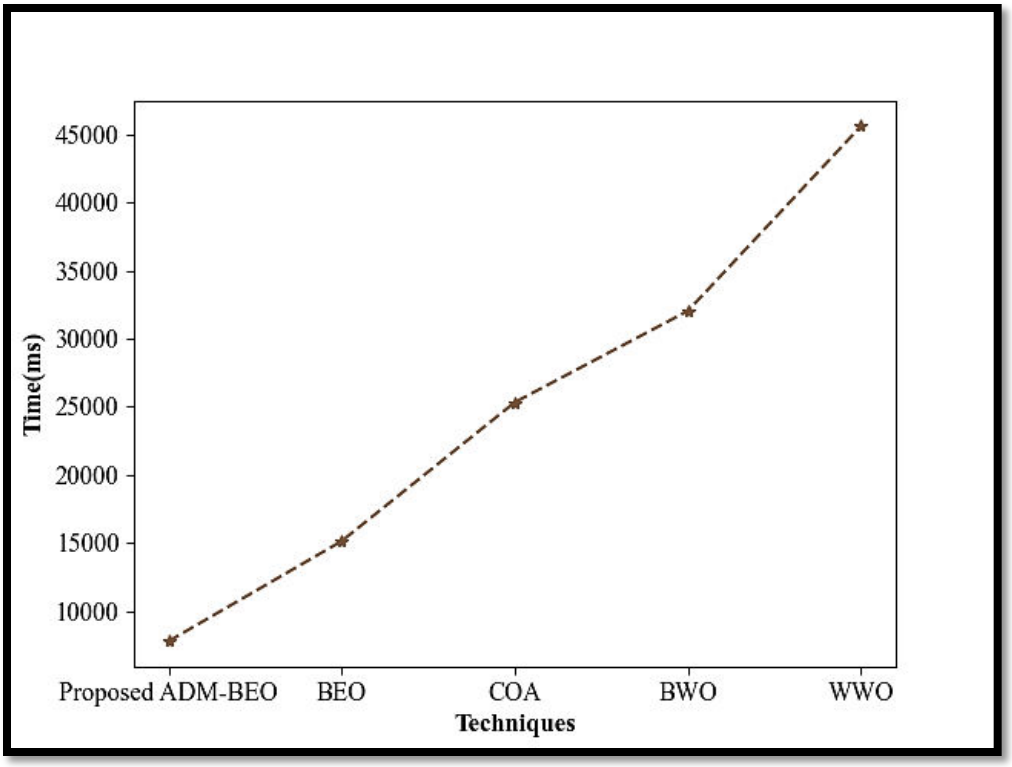


Figure 8: Training time evaluation

The training time of the proposed and existing techniques are analogized in Figure 8. The network is allowed to learn faster and more efficiently by FELU's linear approximation for positive inputs. The proposed FastExp-GRU completes the training process with 42173.97 ms. However, for training the data, the existent techniques require an average of 22047.04 ms. The analysis clearly shows that the proposed system has low time complexity.

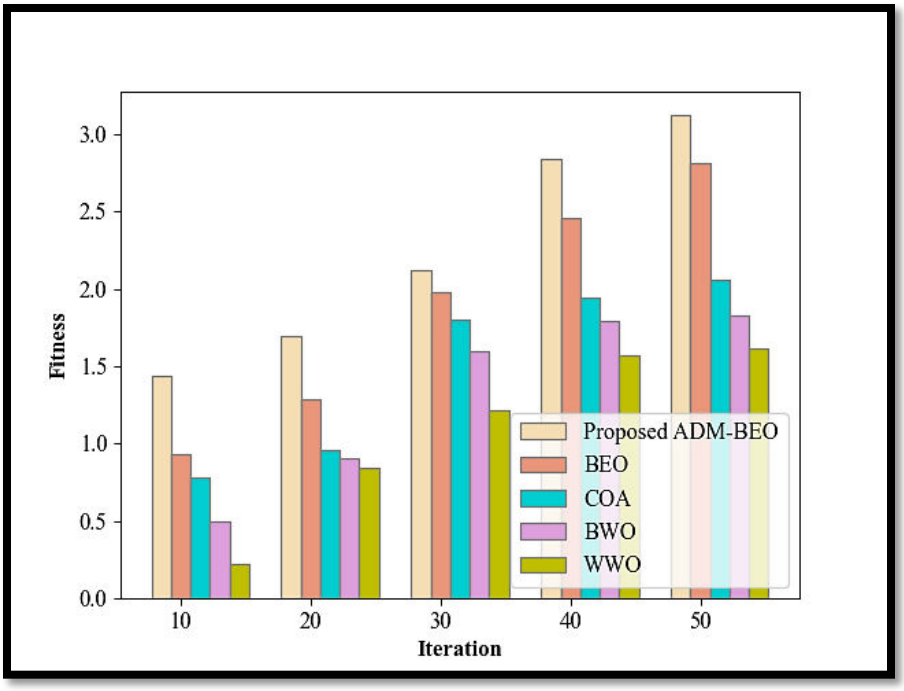


Figure 9: Fitness vs. Iteration Comparison

The fitness vs. iteration of the proposed ADM-BEO and existing algorithms, namely BEO, Coyote optimization algorithm (COA), Black Widow Optimization (BWO), and Water Wave Optimization (WWO) are evaluated in Figure 9. The ADM accelerates convergence by allowing individuals to quickly escape local optima.As a result, the proposed ADM-BEO converges to a better solution faster. Nevertheless,for attaining convergence, the existing techniques require more iteration.

Table 1: Feature selection time validation

Techniques	**Feature selection time (ms)**
Proposed ADM-BEO	7769.836
BEO	15127.26
COA	25324.59
BWO	32040.93
WWO	45617.03

The feature selection time of the proposed and existing approaches are validated in Table 1.The inclusion of ADM potentially leads to better solutions with limited time.

Hence, the proposedADM-BEO selects the features with 7769.83 ms. However, for choosing the features, the existent method requires an average of 29527.45 ms. Hence, the proposed approach's overall execution time is reduced.

Table 2: Performance validation of the clustering efficiency

Techniques	**Performance metrics**		
	Clustering time (ms)	**Accuracy (%)**	**NMSE**
Proposed Ward-KNN	12568	98.92	0.2387
KNN	19253	95.23	0.4581
FCM	26498	93.61	0.5984
KMA	32167	91.03	0.7138
GMM	43981	90.97	0.8934

The clustering efficiency of the proposed Ward-KNN and prevailing KNN, Fuzzy C-Means (FCM), K-Means Algorithm (KMA), and Gaussian Mixture Model (GMM) is exhibited in Table 2. Ward Method considers the minimum variance distance, which improves the time efficiency of the imputation process (12568 ms). Furthermore, the proposed Ward-KNN potentially enhances clustering accuracy (98.92%) and reduces the normalized Mean Squared Error (MSE) (0.2387) by preserving the underlying patterns in the data.

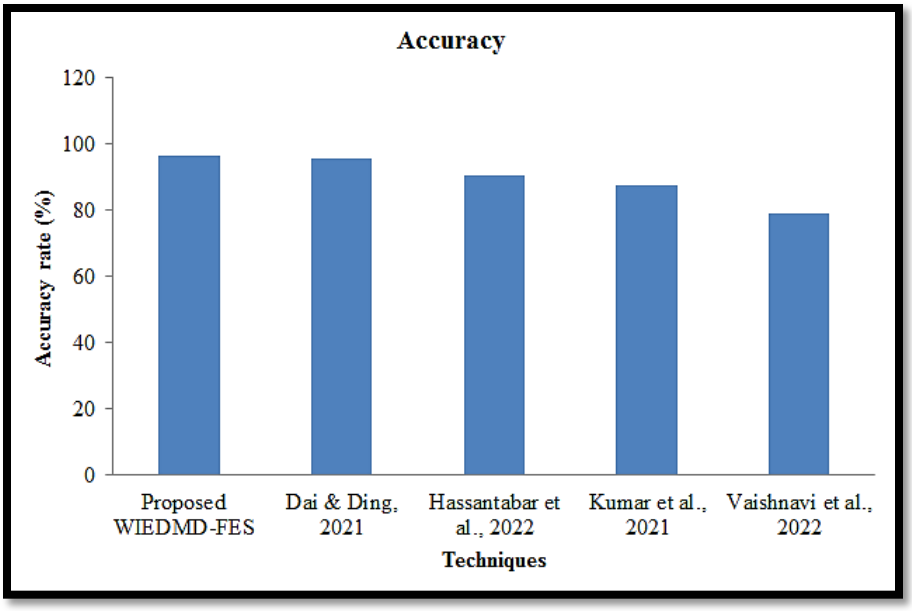


Figure 10: Accuracy comparison with related work

The detection accuracy of the proposed and existing approaches is demonstrated in Figure 10.The proposed approacheffectively addresses the missing values, assists in preserving the integrity and quality of the data, and understands the underlying patterns and characteristics in the EEG signals. Furthermore, the proposed approach potentially improves the classification accuracy (96.59%) by considering both temporal and spatial aspects of brain activity.

5. Conclusion

In this paper, a wearable IoT-based early and real-time detection framework for mental disorders using the FastExp-GRU algorithm is proposed. This work's significance lies in addressing the lack of exhaustive research on the type classification of diverse mental health disorders utilizing wearable IoT devices. The proposed approach offers a comprehensive technique for mental disorder detection by merging signal processing approaches, feature extraction, and Deep learning classification. The proposed approach's efficiency to detect and classify several mental disorders is demonstrated by its experimental evaluation. Utilizing publicly available datasets, the proposed WIEDMD-FES was evaluated. As per the analysis, WIEDMD-FES attained an impressive accuracy (96.59%). Moreover, for the framework, the training time was measured to be 42173.97 ms. Different types of mental disorders are effectively classified by the proposed WIEDMD-FES. Currently, it lacks a particular concentration on severity level estimation. In the future study, expanding the proposed approach's capabilities to include the estimation of severity levels of mental disorders should be concentrated.

References

1. Ahmed, U., Srivastava, G., Yun, U., & Lin, J. C. W. (2022). EANDC: An explainable attention network based deep adaptive clustering model for mental health treatment. Future Generation Computer Systems, 130, 106–113.
2. Alwakeel, A., Alwakeel, M., Zahra, S. R., Saleem, T. J., Hijji, M., Alwakeel, S. S., Alwakeel, A. M., & Alzorgi, S. (2023). Common Mental Disorders in Smart City Settings and Use of Multimodal Medical Sensor Fusion to Detect Them. Diagnostics, 13(6), 1–17.
3. Ahmed, U., Lin, J. C.-W., & Srivastava, G. (2023). Deep Hierarchical Attention Active Learning for Mental Disorder Unlabeled Data in AIoMT. ACM Transactions on Sensor Networks, 19(3), 1–18.
4. Arji, G., Erfannia, L., & Hemmat, M. (2023). Informatics in Medicine Unlocked A systematic literature review and analysis of deep learning algorithms in mental disorders. Informatics in Medicine Unlocked, 40, 1–21.
5. D'Alfonso, S. (2020). AI in mental health. Current Opinion in Psychology, 36,

112–117.

6. Dai, X., & Ding, Y. (2021). Mental Health Monitoring Based on Multiperception Intelligent Wearable Devices. Contrast Media and Molecular Imaging, 2021, 1–7.
7. Gouda, V., Ramya, N. J., Sathvika, T. V, L, V., & Y, V. (2023). Detecting psychological instability using machine learning. International Research Journal of Engineering and Technology (IRJET), 10(5), 76–80.
8. Guo, T., Zhao, W., Alrashoud, M., Tolba, A., Firmin, S., & Xia, F. (2022). Multimodal Educational Data Fusion for Students' Mental Health Detection. IEEE Access, 10, 70370–70382.
9. Gutierrez, L. J., Rabbani, K., Ajayi, O. J., Gebresilassie, S. K., Rafferty, J., Castro, L. A., & Banos, O. (2021). Internet of things for mental health: Open issues in data acquisition, self-organization, service level agreement, and identity management. International Journal of Environmental Research and Public Health, 18(3), 1–19.
10. Gomes, N., Pato, M., Lourenço, A. R., & Datia, N. (2023). A Survey on Wearable Sensors for Mental Health Monitoring. Sensors, 23(3), 1–16.
11. Hassantabar, S., Zhang, J., Yin, H., & Jha, N. K. (2022). MHDeep: Mental Health Disorder Detection System Based on Wearable Sensors and Artificial Neural Networks. ACM Transactions on Embedded Computing Systems, 21(6), 1–22.
12. Hussien, N. M., & Mohialden, Y. M. (2021). An IoT-based knowledge base system for the early detection of Autism diesis. The 2nd Specialized Scientific International Conference on Austim, 216–222.
13. Ji, S., Li, X., Huang, Z., & Cambria, E. (2022). Suicidal ideation and mental disorder detection with attentive relation networks. Neural Computing and Applications, 34(13), 10309–10319.
14. Jin, J., Gao, B., Yang, S., Zhao, B., Luo, L., & Woo, W. L. (2020). Attention-Block Deep Learning Based Features Fusion in Wearable Social Sensor for Mental Wellbeing Evaluations. IEEE Access, 8, 89258–89268.
15. Juneja, S., Dhiman, G., Kautish, S., Viriyasitavat, W., & Yadav, K. (2021). A Perspective Roadmap for IoMT-Based Early Detection and Care of the Neural Disorder, Dementia. Journal of Healthcare Engineering, 2021, 1–11.
16. Khiani, S., Mohamed Iqbal, M., Dhakne, A., Sai Thrinath, B. V., Gayathri, P. G., & Thiagarajan, R. (2022). An effectual IOT coupled EEG analysing model for continuous patient monitoring. Measurement: Sensors, 24, 1–8.
17. Kim, J., Cheon, S., & Lim, J. (2022). IoT-Based Unobtrusive Physical Activity Monitoring System for Predicting Dementia. IEEE Access, 10, 26078–26089.
18. Kumar, A., Sharma, K., & Sharma, A. (2021). Hierarchical deep neural network for mental stress state detection using IoT based biomarkers. Pattern Recognition Letters, 145, 81–87.

19. Li, B. (2023). Design of Early Warning System for Mental Health Problems Based on Data Mining and Database. Revista Brasileira de Medicina Do Esporte, 29, 1–6.
20. Pacheco-Lorenzo, M. R., Valladares-Rodríguez, S. M., Anido-Rifón, L. E., & Fernández-Iglesias, M. J. (2021). Smart conversational agents for the detection of neuropsychiatric disorders: A systematic review. Journal of Biomedical Informatics, 113, 1–14.
21. Saito, T., Suzuki, H., & Kishi, A. (2022). Predictive Modeling of Mental Illness Onset Using Wearable Devices and Medical Examination Data: Machine Learning Approach. Frontiers in Digital Health, 4, 1–16.
22. Singh, J., Wazid, M., Singh, D. P., & Pundir, S. (2022). An embedded LSTM based scheme for depression detection and analysis. Procedia Computer Science, 215, 166–175.
23. Thakur, S. S., & Roy, R. B. (2021). Predicting mental health using smart-phone usage and sensor data. Journal of Ambient Intelligence and Humanized Computing, 12(10), 9145–9161.
24. Uban, A. S., Chulvi, B., & Rosso, P. (2021). An emotion and cognitive based analysis of mental health disorders from social media data. Future Generation Computer Systems, 124, 480–494.
25. Vaishnavi, K., Kamath, U. N., Rao, B. A., & Reddy, N. V. S. (2022). Predicting Mental Health Illness using Machine Learning Algorithms. Journal of Physics: Conference Series, 2161(1), 1–7.