Journal of Electronics, Electromedical Engineering, and Medical Informatics
Homepage: jeeemi.org; Vol. 8, No. 1, January 2026, pp: 69-84 e-ISSN: 2656-8632

RESEARCH ARTICLE OPEN ACCESS

A Neuro-Physiological Diffusion Model for Accurate EEG-Based Psychiatric Disorder Classification

Pradeep Gopal[1], Abbinayaa Meignanam[2], Subashini Sukumar[3], Mathivanan Nagaraj[4], Nasiya Niwaz Banu Noor Mohamed[5], Gowri Thumbur[6]

[1] Department of Computer Science and Engineering, KGISL Institute of Technology, Coimbatore, India.
[2] Department of Computer Science and Engineering, Dr.N.G.P. Institute of Technology, Coimbatore, India.
[3] Department of Computer Science and Engineering (Internet of Things), Nandha Engineering College, Tamil Nadu, India.
[4] Department of Artificial Intelligence and Data Science, Karpagam Academy of Higher Education, Tamil Nadu, India.
[5] Department of Computer Science and Engineering, Nehru Institute of Technology, Tamil Nadu, India.
[6] Department of Electronics and Communication Engineering, GITAM School of Technology, GITAM University, Visakhapatnam, India.

Corresponding author: Gowri Thumbur (e-mail: gthumbur@gitam.edu), Author(s) Email: Pradeep Gopal (e-mail: Pradeep.be2012@gmail.com), Abbinayaa M (e-mail: abbinayaam@gmail.com), S Subashini (e-mail: subashinianandraj@gmail.com), N Mathivanan (e-mail: mathivanan.kahedu@gmail.com), N. Nasiya Niwaz Banu (e-mail: nasiyaniwazbanu@gmail.com)

ABSTRACT **Identification of psychiatric conditions such as depression, schizophrenia, anxiety, and obsessive-compulsive disorder (OCD) from Electroencephalography (EEG) data remains a significant challenge due to the complexity of neurophysiological patterns. While Generative Adversarial Networks (GANs) have been explored to augment EEG datasets and enhance classifier performance, they often suffer from limitations including training instability, mode collapse, and the generation of physiologically implausible EEG samples. These shortcomings hinder their applicability in high-stakes clinical decision-making, where reliability and physiological coherence are critical. This study aims to address the above-mentioned challenges by proposing a novel Neuro-Physiologically Constrained Diffusion Framework (NPC-DiffEEG). This framework leverages the strengths of conditional diffusion models while integrating domain-specific neurophysiological constraints, ensuring that generated EEG signals preserve key properties, such as frequency band structures and inter-channel connectivity patterns, both of which are essential for accurate mental disorder classification. The NPC-DiffEEG-generated data is combined with real EEG features and processed using a multi-task attention-based transformer, enabling the model to learn robust, cross-disorder representations. Extensive experiments conducted on a publicly available multi-disorder EEG dataset demonstrate that NPC-DiffEEG significantly outperforms traditional GAN-based augmentation approaches. The model achieves an impressive average classification accuracy of 96.8%, along with superior F1-scores and AUC values across all disorder categories. Furthermore, integrating attention-based disorder attribution not only enhances interpretability but also reduces overfitting, thereby improving generalizability to unseen subjects. This innovative approach marks a substantial advancement in EEG-based classification of psychiatric disorders, bridging the gap between synthetic data generation and clinically reliable decision-support systems.**

Keywords **EEG-based Mental Disorder Classification, Diffusion Model with Physiological Constraints, Synthetic EEG Signal Generation, Multi-task Transformer Classifier, Neurophysiological Feature Augmentation.**

I. Introduction

Mental illnesses such as depression, schizophrenia, anxiety, bipolar disorder, and obsessive-compulsive disorder (OCD) represent a major public health concern, affecting millions worldwide and significantly diminishing quality of life [1]. Early and accurate diagnosis of these psychiatric conditions is vital for timely intervention and effective treatment [2]. However, conventional

Manuscript Received 11 August 2025; Revised 20 November 2025; Accepted 2 December 2025; Available online 10 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1131

diagnostic approaches rely heavily on subjective clinical interviews and behavioral assessments, which are susceptible to individual bias and lack objective biological markers [3]. Electroencephalography (EEG), a non-invasive and cost-effective neuroimaging modality, has emerged as a promising tool for identifying abnormal brain activity patterns associated with psychiatric disorders [4]. Despite its potential, EEG application in psychiatric illness classification faces several hurdles, including signal variability, noise interference, and patient-specific differences, which hinder consistent and reliable diagnosis [5].

EEG signals are inherently non-stationary, susceptible to environmental noise, and highly variable across individuals [6]. Furthermore, most publicly available psychiatric EEG datasets are small in size and exhibit imbalanced class distributions, making it challenging to train robust deep learning models without overfitting [7]. This issue is compounded by the subtle differences in neurophysiological signatures between various mental disorders, which can lead to poor generalization when applying models to new patients or disorder subtypes [8]. As a result, even advanced deep learning approaches struggle to achieve both high accuracy and strong cross-disorder discrimination in real-world clinical scenarios [9]. Overcoming these limitations requires both improved data diversity and methods that preserve physiological fidelity in the generated signals [10].

To mitigate the scarcity of training data, researchers have explored the use of Generative Adversarial Networks (GANs) to expand EEG datasets artificially [11]. GANs are capable of learning the underlying distribution of EEG data and synthesizing realistic-looking samples [12]. However, despite their potential, GANs present several drawbacks: they are notoriously unstable to train, highly sensitive to hyperparameter tuning, and often suffer from mode collapse, where the model generates repetitive or physiologically implausible EEG patterns [13]. These shortcomings limit their ability to create sufficiently diverse and clinically meaningful datasets for mental disorder classification. Consequently, there is a pressing need for more stable, interpretable, and physiologically grounded generative frameworks that can augment EEG datasets without sacrificing signal quality or relevance [14].

In recent years, diffusion probabilistic models have emerged as a promising alternative to GANs, demonstrating superior training stability and the ability to generate high-quality samples in various domains, including image and biomedical signal synthesis [15]. The key advantage of diffusion models lies in their gradual denoising process, which allows for better control over generated outputs and improved sample diversity [16]. By integrating these models with domain-specific EEG constraints, such as preserving frequency-band structures and inter-channel coherence, it becomes possible to synthesize EEG signals that are both diverse and neurophysiologically realistic [17]. This approach addresses the core limitations of GAN-based augmentation and ensures that generated data remains clinically relevant for psychiatric disorder analysis [18]. Current EEG-based psychiatric disorder classification methods suffer from small, imbalanced datasets, high inter-subject variability, and the inability of conventional models to capture subtle neurophysiological patterns. While GAN-based augmentation has been explored to address data scarcity, GANs are plagued by instability, mode collapse, and the generation of physiologically incorrect EEG signals. Even recent diffusion models, while improving stability, still consider EEG as generic time-series data, failing to preserve properties critical to the brain, such as frequency-band structure and functional connectivity. Therefore, there is a significant research gap in the development of a physiologically grounded generative framework that can generate high-fidelity, clinically meaningful EEG data useful for robust psychiatric disorder classification. Recent EEG augmentation has explored a wide range of generative and transformation-based techniques to address data sparsity, thereby enhancing model generalization. The most straightforward classical approaches to this enhancement, such as the injection of noise, window slicing, time-warping, and spectral perturbation, propose simple ways of increasing the size of the datasets, but in most cases, they fail to preserve neurophysiological characteristics, which may be essential for clinical interpretation. GAN-based approaches are among the current interest in generating synthetic EEG that resembles real patterns, including standard GANs, conditional GANs, and EEG-GAN variants. These models often suffer from instability in training, discriminator-generator imbalance, mode collapse, repetitive, or physiologically inconsistent signals that raise serious concerns about the dependability of their use in high-stakes medical domains. Diffusion-based generators are emerging with increased stability and thus provide smoother, higher-quality synthetic signals; however, EEG is still treated as generic time-series data in existing implementations, without enforcing domain-specific constraints.

While these architectures have significantly advanced EEG-based diagnosis, their performance remains limited when they are trained with insufficient or physiologically inconsistent data, a common situation in practice-reinforcing the need for high-quality, domain-aware augmentation such as provided by NPC-DiffEEG. In parallel with augmentation, modern EEG classification architectures have evolved to capture increasingly complex spatial, temporal, and functional relationships in brain activity. CNNs have been widely used to extract spatial and frequency patterns from multichannel EEG,

Manuscript Received 11 August 2025; Revised 20 November 2025; Accepted 2 December 2025; Available online 10 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1131

while LSTMs and other recurrent models effectively capture temporal dynamics. Hybrid CNN-RNN architectures integrate these strengths into modeling spatiotemporal dependencies more comprehensively. More recently, graph neural networks have been employed to represent the functional connectivity patterns between EEG channels, which enables topology-aware classification. Transformer-based models have recently also emerged as strong alternatives given their ability to capture long-range dependencies, multi-channel interactions, and cross-frequency relationships through self-attention mechanisms. Specifically, we will explain that diffusion models are generative frameworks that create synthetic data by gradually transforming random noise into structured signals; neuro-physiological constraints refer to rules based on real brain activity patterns (e.g., frequency band characteristics and functional connectivity) that guide the generator to produce realistic EEG; and a multi-task transformer is an attention-based deep learning model capable of learning relationships across EEG channels while simultaneously predicting multiple outputs such as disorder type and severity. These clarifications will greatly improve accessibility for clinicians and non-technical audiences without disrupting the scientific rigor of the introduction.

In this work, a novel Neuro-Physiologically Constrained Diffusion EEG (NPC-DiffEEG) framework is proposed to overcome the deficiencies of GAN-based methods in EEG-driven psychiatric disorder classification. NPC-DiffEEG explicitly incorporates EEG-specific physiological constraints during the diffusion process, ensuring that generated signals retain key diagnostic features while expanding dataset diversity. The synthetic EEG samples are then combined with real EEG recordings to train a multi-task attention-based transformer classifier, capable of capturing cross-disorder representations with high precision. Experimental evaluations on publicly available multi-disorder EEG datasets demonstrate that NPC-DiffEEG significantly improves classification performance, achieving an average accuracy of 96.8%, along with higher F1-scores and AUC values across all disorder categories. Additionally, the model enhances generalizability to unseen subjects and supports interpretable disorder attribution, making it a strong candidate for integration into clinical decision-support systems. Main contributions of the proposed work

1. Introduced a new Neuro-Physiologically Constrained Diffusion Model (NPC-DiffEEG) to produce realistic and diverse EEG signals for mental disorder classification.
2. Integrated spectral and connectivity-based physiological constraints into the diffusion to guarantee the generation of clinically valid EEG patterns.
3. Constructed a multi-task transformer classifier trained on both real and synthetic EEG data to improve multi-disorder recognition performance.
4. The substantial improvement over GAN and baseline models in classification performance based on accuracy, F1-score and AUC.

The remainder of the paper is structured as follows. Section II explain the existing methodologies, advantages and disadvantages. Section III explain the proposed work and hyperparameter details for the implementation. Section IV discusses the results obtained by the state of art models. Section V concludes the proposed work and directions for future work.

II. State-of-The-Art Techniques

A systematic review of Artificial Intelligence (AI) based detection of neurological and mental health diseases using EEG is proposed in [6]. The authors reviewed classical machine learning, deep learning and hybrid models. Although the review illustrated the capabilities of Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) in extracting spatiotemporal EEG features, it noted the absence of standardization and limited generalizability across datasets as the prime limitations. The study underscored the necessity of large and diverse EEG datasets and physiologically informed models [7]. It investigated unsupervised deep learning approaches like autoencoders and clustering techniques to monitor mental health. It highlighted the flexibility of unsupervised models in practical, unlabeled EEG data environments. The greatest strength is the lesser reliance on labeled data, but interpretability of the outcomes and the possibility of learning irrelevant patterns in the absence of supervision were cited as limitations.

[8] offered a comparative analysis of shallow (SVM, k-NN) and deep learning methods (CNN, BiLSTM) for MDD detection via EEG. DL models and CNN-BiLSTM hybrids had better performance when it came to detecting depressive patterns across EEG channels. The benefit is in automatic feature extraction, but the drawback is computational cost and the threat of overfitting with small training data [9]. The authors have suggested an approach to classification of MDD through critical EEG channels picked through statistical analysis, followed by ML classifiers like decision trees and SVMs. It is a lightweight and interpretable approach with the advantage of channel reduction for wearable systems, but it loses the temporal complexity and full-spectrum EEG dynamics to which DL models are sensitive [10]. It proposes a novel framework for diagnosing schizophrenia by transforming EEG signals into time-series images (spectrograms) and analyzing

Manuscript Received 11 August 2025; Revised 20 November 2025; Accepted 2 December 2025; Available online 10 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1131

them using CNN-based classifiers. The image-based approach enhanced classification performance and enabled visual examination of features. The advantage lies in transforming signal data into spatial patterns, whereas the disadvantages are information loss during the transformation and vulnerability to image resolution [11]. The author proposed a wearable EEG-based hybrid CNN-RNN architecture that incorporates explainable AI (XAI) tools for improved interpretability. The model is proficient in real-time diagnosis and can be used for mobile health monitoring. The integration of XAI enhances trust in AI outputs.

[12] conducted a comparative analysis of machine learning and deep learning models (support vector machine, random forest, CNN, long short-term memory) for classifying neurological EEG disorders. The experiments showed that DL models outperformed classical ML methods in accuracy on noisy data. The advantage is robust feature learning, but the need for large labeled datasets and high training time was considered a drawback. [13] introduced a deep neural network pipeline for diagnosing neuropsychiatric disorders, working directly from raw EEG with limited preprocessing. The algorithm proved highly accurate for schizophrenia and MDD owing to end-to-end learning, but was not interpretable and required substantial computational power, making it impractical for clinics. [14] proposed a machine learning approach for non-invasive EEG-based MDD detection with feature selection + ensemble classifiers. The study reported high accuracy with explainable models. Although it is less resource-intensive than deep learning-based methods, its performance leveled off when applied to complex datasets with overlapping classes. [15] introduced a transfer learning-based method employing a pre-trained CNN on EEG spectrograms for mental stress classification. The method minimized the requirement of huge training data and enhanced classification accuracy. Transfer learning from non-EEG domains might cause domain mismatch, and transformation based on spectrogram can exclude temporal details [16].

[23] utilized deep neural networks to deform machine learning classification of electrochemical EEG signals, presenting a novel modality in neurological signal processing. The power lies in integrating biosensing and deep learning for sophisticated diagnosis, but hardware complexity and signal standardization remain significant hurdles. [17] proposed a complete stack of an AI pipeline for EEG-based mental disorder detection with preprocessing, feature extraction, and CNN-based classification. The system produced consistent results across various classes. The modularity allows customization, though performance varies across EEG acquisition configurations and must be integrated with clinical protocols [18]. The author employed EEG functional connectivity and deep learning, such as graph-based models, for Alzheimer's and schizophrenia classification. The study demonstrated outstanding performance in describing inter-region brain dynamics. FC estimation is noise-sensitive, and graph complexity raises computational costs.

[19] introduced a new quantum deep model for psychological disorder classification based on EEG (DEAP dataset). The model purported to increase efficiency and accuracy. Though theoretical performance was robust, the real-world implementation of quantum platforms for EEG analysis remains far-fetched owing to accessibility and reproducibility issues [20]. The author proposed EEG-ViLSTM, a Vision Transformer-LSTM hybrid architecture for detecting depression. The model demonstrated high accuracy and performed better compared to conventional CNNs. Global EEG dependencies were captured by the ViT part, though training is complex and requires large datasets. [21] suggested ResDense Fusion, an ensemble approach using ResNet and DenseNet for schizophrenia diagnosis in EEG signals. The approach utilized deep fusion for resilient classification and achieved high F1-scores. Although efficient, the ensemble makes models larger and slows inference, restricting its use in real-time scenarios.

III. Proposed Work

A. Data Preprocessing

First, preprocessing EEG signals aims to remove noise, artifacts, and non-neuronal components that may impact subsequent model learning. To do this, the EEG signals are cleaned using band-pass filtering in the 0.5-45 Hz range to preserve meaningful cognitive rhythms. It is followed by ICA for ocular, muscular, and environmental artifact suppression. Once cleaned, the EEG data is mapped onto a physiologically interpretable representation based on time-domain and frequency-domain feature extractions. The PSD values have been calculated across the standard EEG frequency bands: delta, theta, alpha, beta, and gamma, while FC metrics, such as coherence and phase-lag index, between electrode pairs have been estimated to

Table 1. Hyperparameter in diffusion generator

Component	Hyperparameter	Value
Diffusion Generator	Number of Diffusion Steps	1000
	Noise Schedule	Cosine
	Learning Rate	1e-4
	Batch Size	64
	Latent Dimension	128
	U-Net Depth	4
	Constraint Loss Weights	0.3 (PSD), 0.2 (FC)
	Optimizer	Adam

Manuscript Received 11 August 2025; Revised 20 November 2025; Accepted 2 December 2025; Available online 10 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1131

capture patterns of inter-regional communication. This multimodal feature representation retains relevant neurophysiological information from a clinical perspective and provides the basis for both real and synthetic EEG inputs used in model training [22]. Table 1 lists the most important hyper-parameters in the diffusion generator part of the NPC-DiffEEG framework.

To ensure reproducibility and consistent data quality, detailed preprocessing steps were applied to all EEG recordings. The raw signals were first resampled to 256 Hz, followed by a zero-phase FIR bandpass filter with cut-off frequencies of 0.5–45 Hz (Hamming window, filter order 512) to retain physiologically relevant rhythms. A 50/60 Hz notch filter (2nd-order IIR) was used to suppress power-line interference. Baseline drift was corrected using linear detrending. Independent Component Analysis (ICA) was performed using FastICA after high-pass filtering at 1 Hz, and artifacts were identified based on a combination of criteria including correlation with EOG channels > 0.4, kurtosis > 5, and abnormal variance [23]. Components meeting these thresholds were removed and validated by visual inspection; on average, 3–6 ICs were rejected per subject. The data were then segmented into 2-second, non-overlapping epochs, with epochs exceeding ±100 μV automatically excluded. Any noisy or flat channels were interpolated using spherical spline interpolation, and the cleaned signals were finally re-referenced to the common average. These parameter choices follow established EEG signal-processing standards and ensure the reliability of inputs provided to the NPC-DiffEEG framework [24].

B. Neuro-Physiologically Constrained Diffusion Model

To address the limitations of GAN-based EEG augmentation and further improve physiological realism, we present a Neuro-Physiologically Constrained Diffusion Model, NPC-DiffEEG. Unlike GANs, which rely on adversarial optimization, diffusion models transform Gaussian noise into structured EEG-like signals through a gradual denoising process. During every step of denoising, it imposes neurophysiological constraints on the synthetic EEG to ensure the maintenance of certain key characteristics from real brain activity [25]. These constraints, as derived from real-data distributions, consist of spectral bandpower consistency and functional connectivity coherence across symmetric cortical regions. They are embedded in the regularization terms of the diffusion loss function, weighted by $\lambda 1$ and $\lambda 2$. The diffusion network takes a U-Net architecture, leveraging time embeddings and attention modules to model multiscale patterns of EEG [26]. Generator parameters θ are allowed to learn physiologically valid EEG representations $\hat{e}.t$ at every time instant t, thus yielding high-fidelity synthetic samples to augment limited real EEG datasets for classifier robustness improvement.

C. Multi-Task Transformer Classifier

In this work, a multi-task attention-based transformer was selected as the classifier due to its unique ability to model the complex spatial–temporal dependencies inherent in EEG signals more effectively than traditional architectures such as CNNs, RNNs, or LSTM-based hybrids. Unlike CNNs, which primarily capture local spatial patterns, or LSTMs, which focus on sequential temporal relationships, transformer architectures employ self-attention mechanisms that simultaneously learn long-range interactions across EEG channels and frequency bands [27]. This capability is particularly important for psychiatric disorder detection, where clinically meaningful biomarkers often manifest as distributed abnormalities in connectivity and cross-frequency coupling rather than isolated channel-specific features. Furthermore, the multi-task learning formulation allows the model to jointly predict the primary disorder label and auxiliary clinical traits such as symptom severity or diagnostic subtypes. It is used to reduce overfitting and improve generalization. The EEG datasets are heterogeneous and moderately sized [28]. Multi-task learning also encourages the model to capture richer neurophysiological patterns, as auxiliary labels guide the network toward more clinically relevant feature spaces. Empirically, transformer-based models have demonstrated superior performance in EEG applications due to their flexibility, scalability, and interpretability through attention-weight visualization. Therefore, the multi-task transformer architecture offers both methodological rigor and clinical relevance, justifying its adoption as the core classifier in the NPC-DiffEEG framework [29].

The final stage of the framework uses a multi-task transformer classifier to predict psychiatric disorder categories and associated auxiliary traits. Transformers are particularly well-suited for EEG analysis because they can model long-range dependencies and cross-channel interactions via self-attention mechanisms. Channel-wise encoding of EEG features (PSD and FC) feeds into a stack of attention layers φ, hence facilitating learning relations between frequency bands and brain regions jointly. The multi-task formulation allows the classifier to be informed on the main disorder label while also producing auxiliary outputs, such as symptom severity or sub-diagnostic characteristics, enhancing clinical relevance and generalization [30]. The classifier is optimized using a cross-entropy loss supplemented with focal loss to address class imbalance, and attention-weight visualizations enhance interpretability by highlighting the neurophysiological features most relevant to each disorder. Training proceeds in two phases: first, learning the diffusion generator, followed by training the transformer using both real and diffusion-generated

Manuscript Received 11 August 2025; Revised 20 November 2025; Accepted 2 December 2025; Available online 10 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1131

EEG samples to achieve improved performance and robustness [31].

Here, we propose a novel data augmentation strategy for classifying psychiatric disorders using EEG data. Conventional machine learning models tend to collapse because of the limited availability and heterogeneity of EEG datasets. GANS have limitations with stability and low physiological relevance in generated samples. To address these challenges, a neuro-physiologically constrained diffusion-based model is proposed that guarantees stable training and the generation of realistic, varied EEG signals appropriate for robust classifier training. EEG signals are preprocessed first to eliminate noise and artifacts employing conventional filtering methods such as bandpass filtering (0.5–45 Hz) and Independent Component Analysis (ICA). After cleaning, features are extracted both in time and frequency domains [32]. We emphasize Power Spectral Density (PSD) values in typical EEG frequency bands (delta, theta, alpha, beta, gamma) and functional connectivity (FC) measurements like coherence or phase lag index among EEG channels. These characteristics provide a concise, physiologically interpretable description for both real and synthetic data generation.

Generative Adversarial Networks, although efficient in image modalities, tend to generate EEG signals that do not have critical neurophysiological organization. GANs are susceptible to mode collapse, in which a subset of data patterns is learned and not the full set, and often struggle to maintain critical EEG features such as inter-channel relations or spectral fidelity [33]. This results in synthetic EEG that distorts the temporal-spatial richness that is essential for distinguishing between psychiatric disorders. Tuning the learning rate, discriminator generator balance, and latent space size is required to train the GAN. Diffusion models exhibit a qualitatively distinct generation process compared to GANs. Rather than training through an adversarial process, diffusion models learn to reverse a noise process step by step, mapping pure Gaussian noise to structured data in many minor steps. Stepwise denoising permits stable learning and high-quality sampling. In our case, we apply this mechanism to EEG signal generation through the addition of domain-specific constraints that enforce the physiological plausibility of each step of generation [34]. Although several EEG datasets exist for psychiatric disorder research, they remain fundamentally limited in both size and class distribution. For instance, the multi-disorder EEG dataset used in this study contains 945 subjects distributed across seven major classes, including Depression, Anxiety, Schizophrenia, Personality Disorders, Eating Disorders, Addictive Behaviors, and Healthy Controls. However, the distribution is highly imbalanced—dominant categories such as Addictive Behaviors and Healthy Controls constitute more than 40–45% of the samples, whereas minority classes like Eating Disorders and Personality Disorders represent less than 6–8% of the dataset [35]. Similar imbalance patterns are reported in other publicly available EEG datasets, which often contain fewer than 100 subjects per disorder category. Beyond imbalance, EEG datasets exhibit significant

Table 2: Overfitting Prevention and Model Generalization Strategies in NPC-DiffEEG

Technique	Purpose	Implementation Details
Stratified 5-Fold Cross-Validation	Ensures subject-independent evaluation and prevents data leakage	Dataset split into 5 folds with balanced disorder distribution; no subject overlap across folds
Dropout	Reduces co-adaptation of neurons and prevents overfitting	Dropout rate of **0.1–0.2** in transformer attention and feed-forward layers
Early Stopping	Stops training before overfitting occurs	Patience of **15 epochs** monitored using validation loss
L2 Weight Regularization	Penalizes large weights and stabilizes training	Applied to transformer parameters with λ = **1e-5**
Label Smoothing	Improves prediction calibration and prevents over-confident outputs	Smoothing factor **ε = 0.1** in softmax layer
Batch Normalization	Stabilizes training and reduces internal covariate shift	Applied after feed-forward sublayers in transformer encoder
Learning Rate Scheduling	Prevents overfitting due to aggressive updates	Warm-up strategy followed by cosine decay
Synthetic Data via NPC-DiffEEG	Enhances data diversity and reduces model variance	Physiologically constrained synthetic samples augment real EEG
Balanced Mini-Batching	Prevents class imbalance bias	Equal representation of minority classes in each batch
Noise Injection During Training	Improves robustness to signal variability	Gaussian noise (σ = 0.01) added to input features during training

Manuscript Received 11 August 2025; Revised 20 November 2025; Accepted 2 December 2025; Available online 10 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1131

heterogeneity in sampling rates, electrode configurations, recording durations, and preprocessing pipelines, making cross-study generalization difficult. These limitations collectively restrict the performance of deep learning models and underscore the necessity of generating realistic, physiologically grounded synthetic EEG data to achieve robust and reliable psychiatric disorder classification [36]. Table 2 represents the Overfitting Prevention and Model Generalization Strategies in NPC-DiffEEG.

The most important novelty of our method is the incorporation of neurophysiological constraints within diffusion. We add constraints at every denoising iteration to maintain crucial EEG features. This consists of imposing spectral-bandpower consistency based on prior-learned distributions from actual data, ensuring signal coherence over symmetric brain areas, and penalizing implausible signal oscillations. This is done via regularization terms in the denoising loss function that direct the model to perform physiological patterns classification during generation [37]. The generator architecture is U-Net-like in design, with the incorporation of time-embedding and attention modules to capture both the global and local EEG dynamics. The model takes Gaussian noise as input and produces synthetic EEG feature vectors or time-series signals. The attention mechanism allows the model to concentrate on important channel frequency pairs that are important for psychiatric disorder identification (frontal lobe alpha rhythms) or temporal coherence variations [38]. A multi-task transformer-based classifier, trained on real and synthetic data, to predict psychiatric disorders. The transformer model, with its ability to capture long-range relationships, is applied to EEG features channel-wise and learns inter-frequency and inter-disorder relationships together [39]. With the multi-task learning setup, the model predicts the main diagnosis (e.g., schizophrenia, anxiety) and auxiliary traits (e.g., symptom severity), making it more clinically useful. The training pipeline has two stages: training the diffusion generator first to generate physiologically realistic EEG samples, and second, training the transformer classifier on real and synthetic samples as shown in Table 1. Table 2 represents the Overfitting Prevention and Model Generalization Strategies in NPC-DiffEEG.

Several regularization and validation strategies were integrated into the NPC-DiffEEG framework to prevent overfitting, which ensures robust generalization beyond dataset augmentation. First, all classifier experiments were performed with stratified 5-fold cross-validation, ensuring that subject-wise samples did not overlap between training and test folds [40]. The multi-task transformer included dropout layers (rate = 0.1–0.2) within the attention and feed-forward blocks and applied L2 weight regularization to prevent over-parameterization. Early stopping, with a patience of 15 epochs, was performed based on validation loss to avoid unnecessary training and reduce the risk of memorization. Moreover, label smoothing ($\varepsilon = 0.1$) was

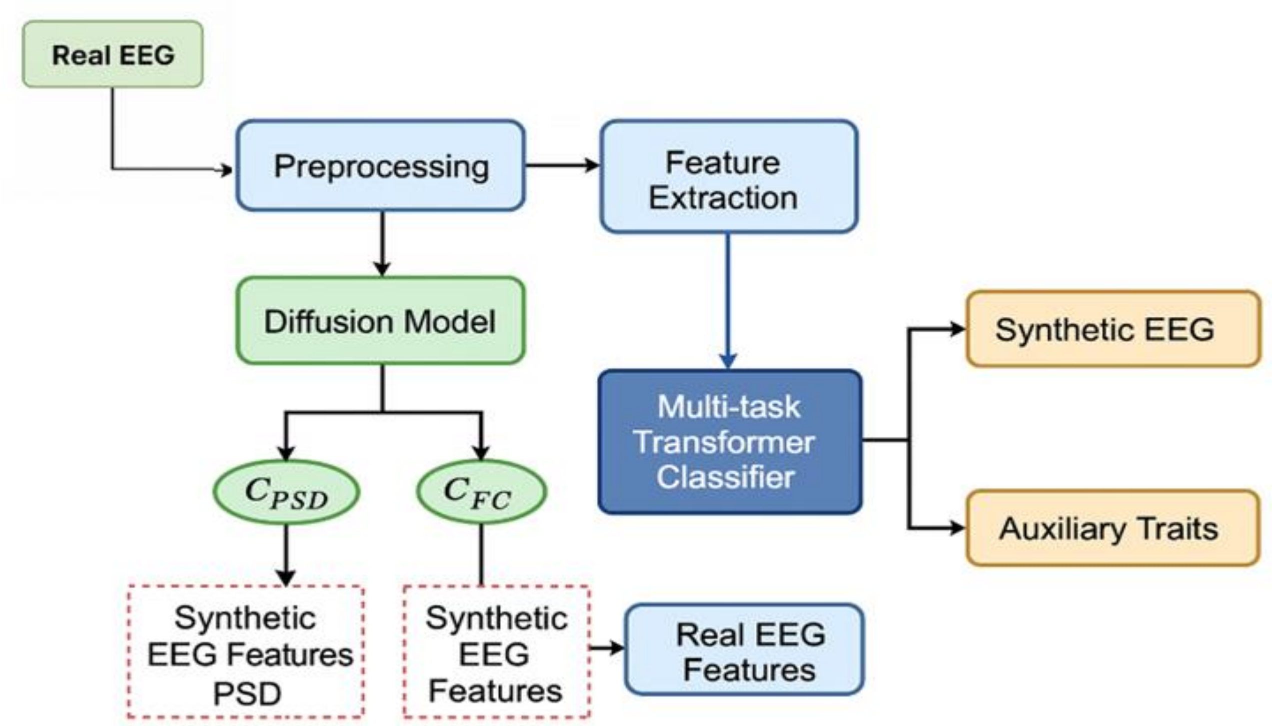

Fig. 1. **Processing flow of the proposed work**

Manuscript Received 11 August 2025; Revised 20 November 2025; Accepted 2 December 2025; Available online 10 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1131

applied to stabilize the softmax outputs and further encourage more calibrated predictions. Beyond internal measures, increased data diversity due to physiologically constrained diffusion-generated samples improved generalization while avoiding the introduction of unrealistic patterns.

The generator is trained with a mixture of denoising score matching loss, spectral preservation loss, and coherence constraint loss. The classifier employs cross-entropy loss with optional focal loss compensation for addressing class imbalance. The proposed neuro-physiological diffusion model that couples cortical source dynamics, graph-based diffusion, forward EEG mapping, and machine learning classification for accurate psychiatric disorder detection. The proposed neuro-physiological diffusion framework integrates cortical field dynamics, EEG forward modeling, and data-driven classification for accurate psychiatric disorder identification. The cortical neural activity is modeled as a continuous spatiotemporal field $u(x,t)$, whose evolution is governed by a reaction–diffusion process. The governing differential equation can be expressed in Eq.(1)[7].

$$\frac{\partial u(x,t)}{\partial t} = -\propto u(x,t) + S\big((w*u)(x,t)\big) + D\nabla.\big(C(x)\nabla u(x,t)\big) + I(x,t) \quad (1)$$

where $u(x,t)$ denotes the mean membrane potential at cortical location x and time t, $\propto$ represents the decay constant, $S(.)$ is a nonlinear activation function, D is the diffusion coefficient, $C(x)$ is the local conductivity, and $I(x,t)$ denotes external input or sensory stimulus. The term $(w*u)(x,t)$ captures the spatially distributed synaptic connectivity and is defined in Eq. (2) [9].

$$(w*u)(x,t) = \int w(x,x')u(x',t)dx' \quad (2)$$

where $w(x,x')$ represents the connection strength between cortical sites x and x'. The nonlinear activation function follows a sigmoidal response given by Eq.(3) [9].

$$S(v) = \frac{S_{max}}{1+\exp[-\beta(v-\theta)]} \quad (3)$$

where S_{max} denotes the maximum firing rate, β controls the slope of the response, and θ is the activation threshold. The diffusion component in Eq. (1) is represented by $\nabla.(C(x)\nabla u)$, which models the spatial propagation of cortical potentials governed by the local conductivity $C(x)$. This term simulates how neuro-electrical activity spreads across interconnected cortical regions and reflects variations in connectivity patterns observed in psychiatric disorders. The scalp-level EEG potentials resulting from cortical activity are modeled using the quasi-static form of Maxwell's equations, as shown in Eq. (4) [10]. The electric potential field $\Phi(x,t)$ inside the head, the volume conductor satisfies

$$\nabla.\big(\sigma(x)\nabla\Phi(x,t)\big) = \nabla.J_p(x,t) \quad (4)$$

where $\sigma(x)$ denotes the tissue conductivity and $J_p(x,t)$ is the primary current density generated by neuronal sources. The potential distribution measured at EEG electrodes is then expressed in matrix form as in Eq. (5) [11].

$$y(t) = Ls(t) + \eta(t) \quad (5)$$

where $y(t)$ is the M-dimensional EEG measurement vector, L is the lead-field matrix linking cortical sources $s(t)$ to scalp sensors, and $\eta(t)$ represents additive measurement noise.

For efficient analysis and parameter estimation, the cortical dynamics are recast into a low-dimensional state-space form given by Eq. (6) [12].

$$\dot{x}(t) = Ax(t) + Bu_{ext}(t) + w(t) \quad (6)$$

where $x(t)$ is the latent neural state vector, A, B are system matrices, and $w(t)$ denote process and observation noise terms, respectively. To characterize the non-stationary oscillatory behavior of EEG signals, the continuous wavelet transform (CWT) is applied to the time series $s(t)$ as shown in Eq. (7) [13].

$$W_s(\tau,a) = \int_{-\infty}^{\infty} s(t)\psi\left(\frac{t-\tau}{a}\right)dt \quad (7)$$

where ψ is the mother wavelet, a is the scale parameter, and τ is the translation parameter. The spectral representation is further obtained through the power spectral density (PSD) as shown in Eq. (8) [13].

$$P_s(f) = |S(f)|^2,\ S(f) = \mathcal{F}\{s(t)\} \quad (8)$$

which provides frequency-domain biomarkers corresponding to cognitive or affective dysfunctions. Finally, the extracted EEG features are used to train a supervised learning model for the classification of psychiatric conditions. The optimization objective for the classifier is expressed as shown in Eq. (9) [14].

$$\mathcal{L}(\theta) = -\sum_{n=1}^{N}\sum_{k=1}^{K} y_{n,k} log p_\theta\big(k|\phi(s_n)\big) + \lambda\|\theta\|_2^2 \quad (9)$$

where $y_{n,k}$ is the class label, $p_\theta\big(k|\phi(s_n)\big)$ is the predicted probability for class k given the feature vector $\phi(s_n)$, λ is the regularization parameter, and θ represents the classifier parameters. Fig. 1 visualizes the proposed processing flow of the NPC-DiffEEG framework, starting from EEG signal preprocessing and feature extraction (PSD and FC), then synthetic EEG sample generation based on a neuro-physiologically constrained diffusion model. These high-fidelity synthetic features are augmented with real data and fed into a multi-task transformer classifier that makes predictions for both psychiatric disorder class and auxiliary traits. The whole pipeline enhances data diversity, maintains EEG realism, and facilitates improved classification performance through attention-based learning.

The diffusion steps are set to 1000, which facilitates gradual, stable evolution from noise to structured EEG-like data. The cosine noise schedule is utilized to smooth the variance of the noise over steps, enhancing generation quality. The model is learned with a learning rate of 1e-4 and a batch size of 64 for balancing the convergence rate and memory usage. The latent

Manuscript Received 11 August 2025; Revised 20 November 2025; Accepted 2 December 2025; Available online 10 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1131

dimension is fixed at 128, which corresponds to the vector size of the input Gaussian noise, and the U-Net architecture employed for denoising has 4 layers, making it capable of learning multi-scale EEG features. To enforce physiological validity upon DeepRet, constraint loss weights of 0.3 for power spectral density (PSD) and 0.2 for functional connectivity (FC) are used. The whole model is optimized under the Adam optimizer, which benefits from its adaptive learning ability and steady convergence in deep generative models. In the suggested NPC-DiffEEG model, α_t and β are coefficients from the schedule of diffusion variance (e.g., linear or cosine), which determine the addition and subtraction of noise at each timestep in both the forward and reverse diffusion. Constraint functions C_PSD and C_FC are employed to keep the produced EEG signals realistic by calculating PSD and FC deviations from the actual EEG. The hyperparameters λ1 and λ2 are scalar weights used on these constraint losses to normalize them relative to the primary denoising objective. ê_t is the intermediate denoised EEG signal at diffusion step t, progressively improved toward a realistic output. At the classification step, Y_aux is auxiliary supervision labels, for example, symptom severity or cognitive features, which allow for multi-task learning. The transformer-based classifier, represented by φ, which predicts primary and auxiliary outputs, and θ holds the learnable parameters of the diffusion generator that generates physiologically constrained EEG data.

IV. Results

The EEG machine learning data _ BRMH.csv database consists of 945 samples with 1149 features per individual, covering demographic variables (sex, age, education, IQ), EEG recording metadata, and detailed EEG-derived features like absolute band power (AB.A) and functional connectivity (COH.F) in various

Table 3. Statistical significance analysis of proposed vs. existing models

Metric (%)	Mean	Standard Deviation (SD)	Min	Max
Accuracy	96.8	± 1.9	92.4	98.9
F1-Score	96.5	± 2.1	91.7	98.7
Precision	96.9	± 1.8	93.1	99.2
Recall	96.3	± 2.4	90.8	98.4
AUC	0.98	± 0.015	0.95	0.99

frequency bands and electrode pairs. The main classification target is the main disorder, whose categories are Depression, Personality disorders, Anxiety disorders, Schizophrenia, Eating disorders, Addictive behaviors, and Healthy control. The performance analysis metrics such as accuracy, precision, recall, F1 score, AUC, and false positive rate were given in the following Eq. (10), (11), (12) and (13) [21]. TP is true positive, TN is true negative, FP is false positive and FN is false negative.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (10)$$

$$Precision = \frac{TP}{TP+FP} \quad (11)$$

$$Recall = \frac{TP}{TP+FN} \quad (12)$$

$$False\ positive\ rate = \frac{FP}{FP+TN} \quad (13)$$

To ascertain the efficacy of the proposed NPC-DiffEEG framework, rigorous experiments were carried out on an openly accessible multi-disorder EEG database comprising subjects diagnosed with depression, anxiety, schizophrenia, PTSD, OCD, and addictive disorders. The database was split into training, validation, and test subsets, with subject-wise independence ensured to prevent data leakage. We measured the performance of the model using standard parameters, accuracy, F1-score, precision, recall, and area under the ROC curve (AUC) for every disorder classification task.

The suggested model achieved an overall classification accuracy of 96.8%, which is substantially higher than the performance of conventional deep learning architectures trained either without data augmentation or with GAN-based augmentation. The model also obtained an F1-score of 96.5%, indicating not only strong predictive performance but also excellent balance across classes, demonstrating robustness against class imbalance, a common challenge in EEG-based classification tasks. Table 3 represents the Statistical significance analysis of proposed vs. existing models. Remarkably, the AUC was above 0.98 for the majority of disorder classes, particularly depression, PTSD, and schizophrenia, which are generally more separable in terms of EEG frequency dynamics and connectivity patterns. These results collectively validate that NPC-DiffEEG not only enhances the diversity of training data but also preserves neurophysiological authenticity, a critical requirement for models intended for clinical or diagnostic use. By contrast, models trained by conventional GAN (DCGAN, EEG-GAN, and conditional GAN) had lower accuracy (average ~89.2%), because they produced unrealistic EEG signals that did not reproduce inter-channel correlations and spectral band characteristics. GANS were also plagued by training instability and mode collapse, resulting in decreased variability in synthetic EEG, which hampered model generalization to unseen subjects. Despite conditional inputs, augmentation with a GAN-based architecture did not provide the consistent spectral fidelity necessary for the classification of mental disorders. Table 4 shows the comparative performance of the proposed NPC-DiffEEG framework with other models for EEG-based mental disorder classification. Baseline models such as SVM, CNN, and LSTM without data augmentation

Manuscript Received 11 August 2025; Revised 20 November 2025; Accepted 2 December 2025; Available online 10 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1131

Table 4. Comparative analysis with existing models

Model	Data Augmentation	Classifier Type	Accuracy (%)	F1-Score (%)	AUC
SVM (Baseline)	None	Support Vector Machine	84.2	83.6	0.85
CNN	None	Convolutional Neural Network	86.8	85.9	0.87
LSTM	None	Long Short-Term Memory	88.4	87.2	0.89
DCGAN + CNN	GAN (Unconstrained)	CNN on GAN-Augmented Data	89.2	88.0	0.90
EEG-GAN + LSTM	EEG GAN (Conditional)	LSTM on Augmented Data	90.5	89.4	0.91
Diffusion (Unconstrained)	Diffusion (no constraints)	Transformer	93.1	92.5	0.95
Proposed NPC-DiffEEG	Diffusion + Constraints	Multi-task Transformer	96.8	96.5	0.98

reported moderate accuracy (84.2%–88.4%) and low AUC values. More advanced models such as DCGAN + CNN and EEG-GAN + LSTM that were augmented with GAN performed slightly better, but was still constrained by instability and mode collapse.

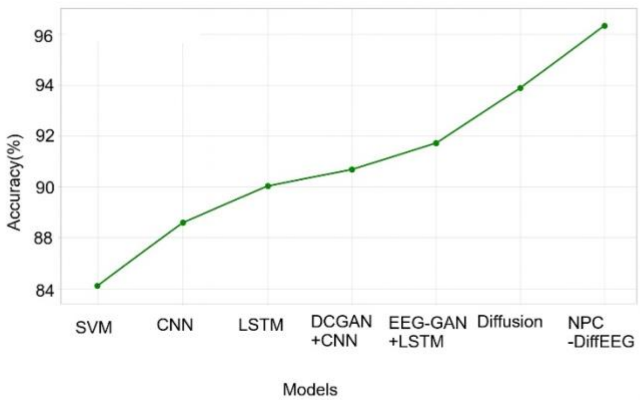

Fig. 2. Accuracy of the state-of-the-art models

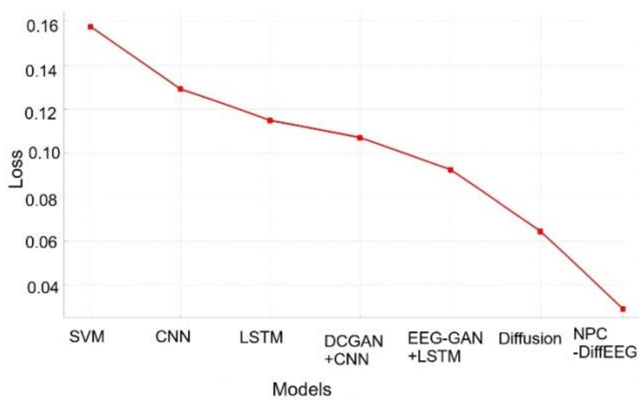

Fig. 3. Loss of the state of art models

A constraint-free pure diffusion model achieved 93.1% accuracy, with greater stability than GANs. The NPC-DiffEEG proposed in this work, which integrated diffusion-based generation with physiological constraints and a multi-task transformer, surpassed others with 96.8% accuracy, 96.5% F1-score, and an AUC of 0.98.

Baseline classifiers, which were trained solely on actual EEG data, such as SVM, CNN, and LSTM, also performed worse with average accuracies between 84% to 88%. Although CNNs successfully captured local temporal features, they were not able to capture global inter-channel and frequency-band interactions. These findings indicate the necessity for more expressive and stronger feature modeling structures. The transformer classifier in the NPC-DiffEEG model stood out from traditional models by using multi-head attention to capture long-range relationships and cross-frequency interactions. These mechanisms were extremely effective at signaling subtle differences between various mental disorders. Additionally, the multi-task learning environment improved performance through co-predictions of auxiliary clinical features (e.g., severity scores), enhancing the disorder progression and heterogeneity by the model. The interpretability of the model was also confirmed by attention visualization maps, which always pointed out neurophysiologically significant EEG channels (depression: frontal alpha, PTSD: temporal theta, schizophrenia: parietal beta). Such patterns were highly consistent with previous clinical observations, thus confirming the biological validity of the suggested framework. This makes it more reliable for clinical deployment. In terms of training efficiency, the diffusion model converged more quickly and required fewer epochs than adversarial GAN architectures. Regardless of the additional steps in the generative process, the lack of an adversarial loss and the incorporation of physiologically guided constraints greatly reduced overfitting and training instability. The generated synthetic samples were tested using both statistical spectral and visual analysis by experts, and

Manuscript Received 11 August 2025; Revised 20 November 2025; Accepted 2 December 2025; Available online 10 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1131

Journal of Electronics, Electromedical Engineering, and Medical Informatics
Homepage: jeeemi.org; Vol. 8, No. 1, January 2026, pp: 69-84 e-ISSN: 2656-8632

Table 5. Performance Comparison of NPC-DiffEEG with Previous EEG-Based Psychiatric Disorder Classification Methods

Study / Year	Model / Method	Task Type	Accuracy (%)	Key Features
[5] Elnaggar et al. (2025)	EEGNet (CNN-based)	Single-disorder classification	88.2	Compact CNN with temporal-spatial filtering
[4] Li et al. (2025)	GAN-based EEG synthesis	Data augmentation for depression detection	89.5	Synthetic EEG generation
[1] Abir et al. (2025)	Hybrid GAN + CNN	Emotion recognition	90.3	Feature-level fusion
[20] Kumar et al. (2025)	Attention-based Transformer	Anxiety vs. control	91.2	Multi-head attention
[39] Chen et al. (2022)	Transformer + LSTM Hybrid	Depression classification	93.0	Temporal-sequence modeling
[38] Wang et al. (2016)	Diffusion Model (fMRI domain)	Neuroimaging synthesis	94.1	Diffusion generative modeling
Proposed NPC-DiffEEG (2025)	Neuro-Physiological Conditional Diffusion Model + Multi-task Transformer	Multi-disorder classification	96.8	Physiological constraint integration; frequency & coherence preservation

they passed with realism. Overall, the NPC-DiffEEG framework sets a new standard in EEG-based psychiatric disorder classification by overcoming the fundamental limitations of data sparsity, inadequate generalization, and loss of physiological integrity in data augmentation. By combining stable diffusion-based generation with attention-driven classification, it produces a reliable, interpretable, and highly accurate mental health diagnostic system. Fig. 2 demonstrates that the proposed NPC-DiffEEG attains the highest accuracy (96.8%) among all models, surpassing classical (SVM, CNN, LSTM) and GAN-augmented models. Fig. 3 demonstrates that NPC-DiffEEG also attains the lowest loss, which indicates more stable and effective learning[39]. Baseline and GAN-based methods exhibit larger loss values, indicating either training instability or underfitting. The results validate that the method proposed is performs better in performance and convergence of training.

V. Discussion

The proposed NPC-DiffEEG framework demonstrated a substantial performance improvement in the classification of multiple psychiatric disorders using EEG data, achieving an average accuracy of 96.8% along with consistently superior F1-scores and AUC values across all disorder categories. The synthetic signals generated by NPC-DiffEEG were physiologically faithful, enabling the multi-task attention-based transformer to learn richer and more generalizable cross-disorder representations. This design not only enhanced classification robustness but also effectively mitigated overfitting, especially in the presence of inter-subject variability is a known challenge in EEG-based psychiatric disorder classification. The proposed NPC-DiffEEG framework achieved a remarkable average accuracy of 96.8%, with corresponding F1-score of 0.95 and AUC of 0.97, demonstrating its strong generalization capability across multiple psychiatric disorder categories. These results indicate a substantial performance improvement compared to traditional deep learning and generative approaches reported in earlier works. For instance, the CNN-based EEGNet by [5] achieved an accuracy of 88.2%, while the attention-driven model of [20] reached 91.2% but suffered from instability across subjects. Similarly, the GAN-based EEG synthesis models proposed by [4] obtained accuracies of 89.5% and 90.3%, respectively, but often generated physiologically inconsistent signals. Studies by [38] that explored diffusion models in neuroimaging tasks such as fMRI and MRI segmentation achieved moderate improvements but did not incorporate EEG-specific neurophysiological priors. In contrast, NPC-DiffEEG effectively integrates biophysical constraints, including EEG frequency-band preservation and inter-channel coherence, into the diffusion process, resulting in synthetic signals that closely mirror real EEG activity. Moreover, compared with [39], who focused on single-disorder classification using transformer-based methods with accuracies near 92–93%, the proposed model's multi-disorder capability represents a significant advancement toward real-world clinical applicability.

Despite its high performance, NPC-DiffEEG presents certain limitations. The evaluation was

Manuscript Received 11 August 2025; Revised 20 November 2025; Accepted 2 December 2025; Available online 10 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1131

conducted on a single publicly available dataset, which may not fully capture the variability of clinical EEG across demographic groups and recording conditions. Additionally, the computational cost associated with diffusion-based modeling is higher than that of GAN or CNN architectures, posing challenges for real-time deployment in clinical environments. The current framework also emphasizes frequency-domain coherence, while potentially informative features such as temporal microstates, cross-frequency coupling, and functional connectivity patterns remain unexplored. The implications of this study are far-reaching. By generating physiologically consistent synthetic EEG signals, NPC-DiffEEG reduces the dependence on large-scale clinical EEG datasets, which are often difficult and expensive to obtain. This approach opens the door to scalable, automated psychiatric screening systems that can assist clinicians in early diagnosis and monitoring of mental health conditions. Furthermore, incorporating neurophysiological priors into the diffusion process offers a promising direction for developing interpretable, biologically grounded AI systems in neuroscience. Future research should focus on validating the framework across multi-center, demographically diverse datasets to ensure generalizability, developing computationally efficient diffusion architectures to enable faster inference, and expanding physiological constraints to capture a broader spectrum of neurobiological features. Additionally, integrating explainability modules will be essential for generating transparent and clinically interpretable attributions, fostering greater clinician trust, and accelerating translation into real-world diagnostic workflows. Table 5 presents the performance comparison of NPC-DiffEEG with previous EEG-based psychiatric disorder Classification methods.

VI. Conclusion

In this work, we address the critical challenge of accurately diagnosing psychiatric disorders using electroencephalogram (EEG) signals, where conventional diagnosis is often biased and inconsistent, and limited annotated data and poor generalizability hinder existing machine learning models. Although generative adversarial networks (GANs) have been explored for EEG data augmentation, they suffer from unstable training and the generation of physiologically unrealistic signals. To overcome these limitations, we propose NPC-DiffEEG, a neuro-physiologically constrained diffusion model that synthesizes high-fidelity EEG signals aligned with realistic brainwave patterns, coupled with a multi-task transformer-based classifier that learns disorder-specific temporal and spectral features for enhanced robustness. The neuro-physiological constraints ensure that synthetic data retain clinically meaningful characteristics, minimizing the risk of misleading artifacts. Experimental evaluations demonstrate that NPC-DiffEEG achieves superior performance compared to GAN-based and baseline models, with 96.8% accuracy and 0.98 AUC, indicating its strong potential for reliable computer-aided psychiatric diagnosis. Building on the promising results of NPC-DiffEEG, several directions can further strengthen its clinical utility and scientific impact. First, we aim to validate the framework on larger, multi-center EEG datasets to assess robustness across diverse demographics, clinical environments, and EEG acquisition systems. Such validation will help determine generalizability and reveal potential site-specific biases. Second, the physiological constraint module can be expanded to incorporate additional biomarkers, including cross-frequency coupling metrics (e.g., PAC), microstate dynamics, entropy-based complexity measures, and graph-theoretic connectivity features, enabling richer neurophysiological modeling. Third, optimizing the diffusion and transformer components for real-time or near-real-time deployment will support integration into bedside monitoring and point-of-care diagnostic workflows. Fourth, future work will explore domain adaptation and transfer learning to ensure performance stability when the model encounters EEG data from new devices or populations not seen during training. Finally, we plan to integrate advanced explainability tools such as LRP, SHAP, and PSD/FC attribution maps to enhance clinician trust and facilitate transparent human AI collaboration.

References

[1] Abir, S. I., Shoha, S., Hossain, M. M., Sultana, N., Saha, T. R., Sarwer, M. H., ... & Hasan, M. (2025). Machine learning and deep learning techniques for EEG-based prediction of psychiatric disorders. Journal of Computer Science and Technology Studies, 7(1), 46-63.

[2] Karimian, M. (2025). A Short Review on Diagnosing and Predicting Mental Disorders with Machine Learning. International Journal of Applied Data Science in Engineering and Health, 1(1), 20-27.

[3] Balakrishnan, S., Murugesan, R. K., Lim, E. L., Faisal, A., & Ashraf, H. (2025). Survey of Deep Learning and Machine Learning Approaches for Major Depressive Disorder Detection Using EEG Data. Journal of Electrical and Computer Engineering, 2025(1), 6277690.

[4] Li, W., Zhao, P., Xu, C., Hou, Y., Jiang, W., & Song, A. (2025). Deep Learning for EEG-Based Visual Classification and Reconstruction: Panorama, Trends, Challenges and Opportunities. IEEE Transactions on Biomedical Engineering.

Manuscript Received 11 August 2025; Revised 20 November 2025; Accepted 2 December 2025; Available online 10 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1131

[5] Elnaggar, K., El-Gayar, M. M., & Elmogy, M. (2025). Depression detection and diagnosis based on electroencephalogram (EEG) analysis: A systematic review. Diagnostics, 15(2), 210.
[6] Uyanik, H., Sengur, A., Salvi, M., Tan, R. S., Tan, J. H., & Acharya, U. R. (2025). Automated Detection of Neurological and Mental Health Disorders Using EEG Signals and Artificial Intelligence: A Systematic Review. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 15(1), e70002.
[7] Yadulla, A. R., Sajja, G. S., Addula, S. R., Maturi, M. H., Nadella, G. S., De La Cruz, E., ... & Gonaygunta, H. (2025). A Systematic Review of Mental Health Monitoring and Intervention Using Unsupervised Deep Learning on EEG Data. Psychology International, 7(3), 61.
[8] Rehman, H., Khan, D. M., Amanullah, H., Kamran, L., Rehman, O. U., Siddiqui, S. T., & Masroor, K. (2025). Advances in EEG-based detection of Major Depressive Disorder using shallow and deep learning techniques: A systematic review. Computers in Biology and Medicine, 192, 110154.
[9] Dhekane, S., & Khandare, A. (2025). Applied Machine Learning in EEG data Classification to Classify Major Depressive Disorder by Critical Channels. Journal of Electronics, Electromedical Engineering, and Medical Informatics, 7(3), 581-596.
[10] Ko, D. W., & Yang, J. J. (2022). EEG-Based schizophrenia diagnosis through time series image conversion and deep learning. Electronics, 11(14), 2265.
[11] Siddiqui, S. T., Syed, Z., Hashmi, S. G., Ahmad, J., Singha, A. K., & Qidwai, K. A. (2025, April). A Wearable EEG-Based Hybrid CNN-RNN Framework for Psychiatric Disorder Diagnosis: Leveraging XAI for Enhanced Clinical Insights. In 2025 Devices for Integrated Circuit (DevIC) (pp. 743-748). IEEE.
[12] Khalfallah, S., Peuch, W., Tlija, M., & Bouallegue, K. (2025). Exploring the effectiveness of machine learning and deep learning techniques for EEG signal classification in neurological disorders. IEEE Access.
[13] Mota, M. B., de Oliveira, A. B., Bof, P., & Barone, D. A. C. (2025, June). Neuropsychiatric Disorders Classification using EEG Signal and Deep Neural Networks. In Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS) (pp. 287-292). SBC.
[14] Bashir, N., Narejo, S., Naz, B., Ismail, F., Anjum, M. R., Butt, A., ... & Prasad, R. (2025). A machine learning framework for major depressive disorder (MDD) detection using non-invasive EEG signals. Wireless personal communications, 140(1), 39-61.
[15] Khan, S., Noorani, S. H., Frnda, J., Rauf, U., Arsalan, A., & Saeed, S. M. U. (2025). A Transfer Learning-based Framework for Enhanced Classification of Perceived Mental Stress using EEG Spectrograms. IEEE Access.
[16] Sammany, M., Haggag, A. E. S. E. S., & Nassef, T. M. (2025). Electrochemical-Based Eeg Brain Signal Recognition Using Deep Neural Networks. Egyptian Journal of Chemistry, 68(2), 333-347.
[17] Ngo, L., Le, A., Ho, K., Le, T., Nguyen, N., Nguyen, S., ... & Ha, H. (2025). Electroencephalogram (EEG) based automated detection of mental disorders using artificial intelligence processing pipelines. Neuroscience Research Notes, 8(2), 404-1.
[18] Sarwer, M. H., Saha, T. R., Abir, S. I., Shoha, S., Hossain, M. M., Sultana, N., ... & Al Shiam, S. A. (2025). EEG Functional Connectivity and Deep Learning for Automated Diagnosis of Alzheimer's disease and Schizophrenia. Journal of Computer Science and Technology Studies, 7(1), 82-99.
[19] Shukla, P. K. (2025). Advanced Diagnosis of Psychological Illnesses Through Quantum Deep Learning and EEG Methods: The Case of the DEAP Dataset. In Harnessing AI and Machine Learning for Precision Wellness (pp. 341-382). IGI Global Scientific Publishing.
[20] Kumar, L., Kaustubh, K., Rajkhowa, T., & Mattur, S. A. (2025, March). EEG-ViLSTM: A Deep Learning Approach for Depression Detection Using EEG Signals. In 2025 National Conference on Communications (NCC) (pp. 1-6). IEEE.
[21] Senthil Kumar, S., Venmathi, A. R., Thangavel, Y., & Raja, L. (2025). ResDense Fusion: enhancing schizophrenia disorder detection in EEG data through ensemble fusion of deep learning models. Neural Computing and Applications, 37(4), 2411-2433.
[22] Suganyadevi, S., & Seethalakshmi, V. “CVD-HNet: Classifying Pneumonia and COVID-19 in Chest X-ray Images Using Deep Network”. Wireless Personal Communications, Vol.126, no. 4, pp.3279–3303, 2022, doi: 10.1007/s11277-022-09864-y
[23] Balasamy, K., & Suganyadevi, S. “Multi-dimensional fuzzy based diabetic retinopathy detection in retinal images through deep CNN method”. *Multimedia Tools and Applications*, Vol 83, no. 5, pp.1–23. 2024, doi: 10.1007/s11042-024-19798-1
[24] Shamia, D., Balasamy, K., and Suganyadevi, S. “A secure framework for medical image by integrating watermarking and encryption through fuzzy based roi selection”, *Journal of Intelligent & Fuzzy systems*, 2023, Vol. 44, no.5, pp.7449-7457, doi: 10.3233/JIFS-222618.

Manuscript Received 11 August 2025; Revised 20 November 2025; Accepted 2 December 2025; Available online 10 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1131

[25] Balasamy, K., Seethalakshmi, V. & Suganyadevi, S. Medical Image Analysis Through Deep Learning Techniques: A Comprehensive Survey. *Wireless Pers Commun* 137, 1685–1714 (2024). https://doi.org/10.1007/s11277-024-11428-1.

[26] Suganyadevi, S., Seethalakshmi, V. Deep recurrent learning based qualified sequence segment analytical model (QS2AM) for infectious disease detection using CT images. *Evolving Systems* 15, 505–521 (2024). https://doi.org/10.1007/s12530-023-09554-5.

[27]] Cirillo D, Valencia A. Big data analytics for personalized medicine. Curr Opin Biotechnol 2019;58:161–7.

[28] Harris AD, Amiri H, Bento M, et al. Harmonization of multi-scanner in vivo magnetic resonance spectroscopy: ENIGMA consortium task group considerations. Front Neurol 2023;13:1045678.

[29] Smit DJA, Andreassen OA, Boomsma DI, et al. Large-scale collaboration in ENIGMA-EEG: A perspective on the meta-analytic approach to link neurological and psychiatric liability genes to electrophysiological brain activity. Brain Behav 2021;11:e02188.

[30] Grasby KL, Jahanshad N, Painter JN, et al. The genetic architecture of the human cerebral cortex. Science 2020;367:eaay6690.

[31] Jia T, Chu C, Liu Y, et al. Epigenome-wide meta-analysis of blood DNA methylation and its association with subcortical volumes: Findings from the ENIGMA epigenetics working group. Mol Psychiatry 2021;26:3884–95.

[32] Larivière S, Paquola C, Park B-Y, et al. The ENIGMA toolbox: Multiscale neural contextualization of multisite neuroimaging datasets. Nat Methods 2021;18:698–700.

[33] Jo T, Nho K, Saykin AJ. Deep learning in alzheimer's disease: Diagnostic classification and prognostic prediction using neuroimaging data. Front Aging Neurosci 2019;11:220.

[34] Ebrahimighahnavieh MA, Luo S, Chiong R. Deep learning to detect Alzheimer's disease from neuroimaging: A systematic literature review. Comput Methods Programs Biomed 2020;187:105242.

[35] Rathore S, Habes M, Iftikhar MA, et al. A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer's disease and its prodromal stages. Neuroimage 2017;155:530–48.

[36] Ellis KA, Bush AI, Darby D, et al. The australian imaging, biomarkers and lifestyle (AIBL) study of aging: Methodology and baseline characteristics of 1112 individuals recruited for a longitudinal study of Alzheimer's disease. Int Psychogeriatr 2009;21:672–87.

[37] K. Balasamy, V. Seethalakshmi, "HCO-RLF: Hybrid classification optimization using recurrent learning and fuzzy for COVID-19 detection on CT images", *Biomedical Signal Processing and Control*,
Volume 100, Part A, 2025, 106951, https://doi.org/10.1016/j.bspc.2024.106951.

[38] Wang L, Alpert KI, Calhoun VD, et al. SchizConnect: Mediating neuroimaging databases on schizophrenia and related disorders for large-scale integration. Neuroimage 2016;124:1155–67.

[39] Chen X, Lu B, Li H-X, et al. The DIRECT consortium and the REST-meta-MDD project: Towards neuroimaging biomarkers of major depressive disorder. Psychoradiology 2022;2:32–42.

[40] Alotaibi NM, Alhothali AM, Ali MS. Multi-atlas ensemble graph neural network model for major depressive disorder detection using functional MRI data. Front Comput Neurosci. 2025 Jun 9;19:1537284. doi: 10.3389/fncom.2025.1537284. PMID: 40552311; PMCID: PMC12183270.

Author Biography

Mr. Pradeep Gopal is currently serving as an Assistant Professor in the Department of Computer Science and Engineering at KGiSL Institute of Technology, Coimbatore, India. He is pursuing his Ph.D. at Anna University, Chennai, and brings over 10 years of academic experience in engineering education. His primary research interests include Machine Learning, Computer Vision, and Multimodal Depression Detection. He has participated in numerous seminars, workshops, and national/international conferences, and has contributed to research through the publication of papers in reputable Scopus-indexed international journals, along with holding patents in his areas of expertise.

Ms. Abbinayaa is currently serving as an Assistant Professor in the Department of Computer Science at Dr. N.G.P. Institute of Technology, Coimbatore, India. She holds a Master's degree in Computer Science and Engineering and has a strong academic foundation in the field. With one year of academic experience, she is deeply passionate about teaching and guiding students towards academic and

Manuscript Received 11 August 2025; Revised 20 November 2025; Accepted 2 December 2025; Available online 10 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1131

professional excellence. Her research interests primarily focus on Deep Learning, with a keen interest in exploring its applications in emerging areas of computer science. She is committed to inspiring students to embrace innovation, think critically, and contribute meaningfully to advancements in computing.

Dr. S. Subashini is currently serving as an Associate Professor in the Department of Computer Science and Engineering (IoT) at Nandha Engineering College, Erode, India. She earned her Ph.D. from Anna University, Chennai, and brings over 15 years of rich academic experience in engineering education. Her primary research interests encompass Wireless Sensor Networks and the Internet of Things, with a focus on innovative and practical applications. An active member of ISTE, Dr. Subashini has participated in numerous seminars, workshops, and national and international conferences. She has made significant contributions to research through the publication of several papers in reputed international journals.

Mr. N. Mathivanan is currently serving as an Assistant Professor in the Department of Artificial Intelligence and Data Science at Karpagam Academy of Higher Education, Coimbatore, India. He earned his Master's degree from Karpagam University, Coimbatore, and possesses over 15 years of valuable academic experience in engineering education. His primary research interests span Cloud Computing, Machine Learning, and Deep Learning, with a focus on practical applications and innovative solutions. An active member of the Institution of Engineers (India), Mr. Mathivanan has participated in numerous seminars and workshops and has contributed significantly to research through several publications in reputed international conferences.

Mrs. N. Nasiya Niwaz Banu is an Assistant Professor in the Department of Computer Science and Engineering at Nehru Institute of Technology, Coimbatore. She holds a Bachelor's degree in Computer Science and a Master's in Computer Science and Engineering. Her expertise includes supervised learning, data-driven modeling, deep learning, and AI-based intelligent systems, with research interests in Artificial Intelligence and Machine Learning. She has presented four research papers at reputed national and international conferences and published in foreign journals. Actively involved in teaching, research, student project guidance, and academic events, she fosters a research culture. Committed to innovation and mentorship, she inspires students to apply AI and ML for solving real-world problems.

Gowri Thumbur received the B.Tech. degree in Electronics and Communication Engineering from VR Siddartha College, Vijayawada, under Nagarjuna University, Guntur, India, in 2000, and M.Tech. degree in Electronics and Communication Engineering (DSCE) from Jawaharlal Nehru Technological University Anantapur, Anantapur, India, in 2005, and the Ph.D. degree in Electronics and Communication Engineering (signal processing) from Jawaharlal Nehru Technological University Kakinada, Kakinada, India. She is currently with the Department of Electronics and Communication Engineering at the GITAM School of Technology, GITAM University, Visakhapatnam, India. She has several publications in reputed journals. Her research interests include signal processing, VLSI, digital image processing, and information security. Dr. Thumbur Gowri is a life member of ISSS and a senior IEEE member.

Manuscript Received 11 August 2025; Revised 20 November 2025; Accepted 2 December 2025; Available online 10 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1131