PAPER

A hybrid human–machine framework integrating DCGAN and CNN for depression detection under incomplete eeg data conditions

To cite this article: Na Cheng *et al* 2026 *Eng. Res. Express* **8** 025217

Engineering Research Express

PAPER

RECEIVED
24 July 2025

REVISED
25 December 2025

ACCEPTED FOR PUBLICATION
8 January 2026

PUBLISHED
20 January 2026

A hybrid human–machine framework integrating DCGAN and CNN for depression detection under incomplete eeg data conditions

Na Cheng[1], **Yongheng Li**[1], **Qianqian Zhang**[1,*], **Yun-Bo Zhao**[2], **Yu Kang**[3], **Qinggaoyuan Zhou**[1] and **Zihui Wu**[1]

[1] School of Artiffcial Intelligence, Anhui University, Hefei, People's Republic of China
[2] Department of Automation, University of Science and Technology of China, Hefei, People's Republic of China
[3] School of Electrical Engineering and Automation, Hefei University of Technology, Hefei, People's Republic of China
[*] Author to whom any correspondence should be addressed.

E-mail: qqzhang@ahu.edu.cn

Keywords: depression diagnosis, incomplete data, generative adversarial networks, convolutional neural networks, human-machine hybrid, adaptive diagnosis

Abstract

In clinical EEG-based diagnosis, the acquisition of sufficient and well-labeled data is often hindered by high data collection costs, limited accessibility to subjects, and the inherent difficulty of expert labeling. These constraints result in insufficient data availability, significantly limiting the performance and reliability of traditional machine learning models. To address this challenge, we propose a hybrid human–machine diagnostic framework that integrates deep convolutional generative adversarial networks (DCGANs) and convolutional neural networks (CNNs) for depression detection under incomplete data conditions. The DCGAN module synthesizes realistic EEG samples to augment scarce datasets, while CNNs are employed for feature extraction and classification. A confidence-aware fusion strategy dynamically integrates expert assessments with model predictions, effectively improving diagnostic accuracy in scenarios with limited labeled data. Experimental results on real EEG datasets demonstrate that the proposed approach achieves superior performance, offering a practical solution for intelligent diagnosis in resource-constrained settings.

1. Introduction

The human brain is an extraordinarily intricate organ that governs perception, cognition, and behavior, forming the foundation of all higher-level mental functions[1]. Ensuring the stability and health of this system is therefore essential for maintaining effective cognitive performance and sound decision-making[2]. However, a variety of neurological and psychiatric conditions can disrupt normal brain functioning. Among them, Major Depressive Disorder (MDD) has emerged as one of the most prevalent and debilitating mental health disorders [3], posing substantial clinical and societal challenges.

An epidemiological data reveal its alarming prevalence, as demonstrated by Huang *et al.*'s nationwide survey indicating a 6.8% lifetime depression incidence in China [4]. The disorder's socioeconomic impact proves equally concerning, accounting for 6.2% of China's total disease-related economic burden and ranking as the second most costly mental health condition [5]. Globally, the World Health Organization (WHO) classifies major depressive disorder (MDD) as the fourth leading cause of disability worldwide [6], further underscoring its clinical significance [7]. Characterized by persistent emotional distress, cognitive impairment, and functional decline, depression not only disrupts daily activities but also severely compromises social engagement capabilities [8]. Particularly vulnerable populations like the elderly face exacerbated challenges, exhibiting poorer recovery trajectories compared to younger cohorts [9]. These multifaceted consequences highlight the critical need for precise severity assessment and dynamic monitoring [10], which remain fundamental to developing targeted interventions and alleviating this escalating mental health epidemic.

However, the diagnosis and treatment of depression remain highly challenging. Social stigma often causes individuals to underreport symptoms out of fear of judgment [11], reinforcing social desirability bias that leads patients to deliberately minimize their emotional difficulties. Meanwhile, current diagnostic procedures still rely heavily on subjective clinical evaluations and self-reported symptoms, making them vulnerable to human error, inconsistent interpretation, and limited reliability [2, 3]. Collectively, these limitations highlight an urgent need for objective, reliable, and supportive diagnostic tools that can assist clinicians in making accurate and timely decisions.

Electroencephalography (EEG) is a technique that records the brain's spontaneous electrical activity by placing electrodes on the scalp in a non-invasive manner and plays a crucial role in diagnosing neurological disorders, including depression [12, 13]. By capturing the weak electrical signals generated by cortical neurons, EEG reflects the brain's bioelectrical activity during cognitive processes such as thinking and perception. EEG signals can reveal alterations in brainwave patterns and are extensively utilized in medical settings to assist in diagnosing neurological disorders, including motor neuron disease, Parkinson's disease, epilepsy, sleep disorders, coma, encephalopathy, and brain death [14–17].

Accurate analysis of EEG data provides doctors with deeper insights into a patient's neurological condition, enabling the development of personalized treatment plans. The integration of artificial intelligence with clinical expertise allows machine learning algorithms to efficiently analyze vast amounts of clinical data and imaging information. Recently, deep learning has gained popularity in EEG signal processing. For example, models predicting Beck Depression Inventory (BDI) scores have shown promising results [18]. Techniques like Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks have achieved exceeding 92% classification accuracy by analyzing features from preprocessed EEG signals [19]. Studies indicate that depressed patients often exhibit reduced alpha wave activity, a key diagnostic feature. Additionally, a machine learning method for screening depression in young adults using wireless EEG has been proposed [20]. This method filters EEG data into six frequency bands, extracts features such as Hjorth parameters, Shannon entropy, and log energy entropy, and uses a Cubic SVM classifier with 5-fold cross-validation. The model achieves 97.22% accuracy in the Beta band (12–30Hz), with 97.2% precision and 95.8% specificity, outperforming others in distinguishing depressed individuals.

However, EEG-based intelligent diagnosis of depression still faces three major challenges: (1) the high cost of EEG data collection limits large-scale dataset construction; (2) EEG data annotation relies heavily on clinical experts, which is time-consuming and lacks standardization; and (3) existing datasets are often small, incomplete, or lack labeling. These data-related constraints significantly hinder the generalization and real-world deployment of machine learning models in clinical settings. In this study, the term 'incomplete data conditions' reflects several inherent limitations of the Mumtaz HUSM EEG dataset, including limited clinical metadata, the absence of channel-level quality-control information, and variability in the amount of usable EEG across subjects. Together with the relatively small sample size, these factors restrict the completeness of the data representation and pose challenges for developing reliable automated depression classifiers.

Recent studies have explored data-efficient modeling strategies, such as employing generative models for data augmentation and using deep learning techniques to extract robust features. Building upon our previous human–machine collaborative diagnosis framework for depression [21], this study proposes a hybrid human–machine diagnostic framework that differs substantially from prior DCGAN-based EEG studies. While existing research typically combines GAN-generated signals with CNN classifiers, our method introduces a decision-level fusion mechanism that explicitly incorporates human clinical judgment into the computational pipeline. Rather than treating clinicians and machine learning models as independent diagnostic entities, the proposed framework integrates them through a confidence-aware adaptive weighting strategy, enabling dynamic adjustment between machine predictions and expert assessment. This hybrid design enhances robustness, and reduces model-driven bias that have not been addressed in previous GAN–EEG approaches.

The major contributions of this work are summarized as follows:

- **DCGAN-based EEG Data Augmentation**: We employ a deep generative model (DCGAN) to synthesize realistic EEG signals and expand the limited original dataset. By training the classifier on a combination of real and synthetic samples, the model captures richer spatial–temporal EEG representations associated with depressive symptoms, effectively alleviating the small-sample problem common in clinical EEG studies.
- **Adaptive Human-Machine Hybrid Intelligence**: Unlike prior works that independently combine DCGAN and CNN, this study implements an adaptive human-machine hybrid intelligence model that integrates expert knowledge with machine learning techniques. In this model, adaptive physician diagnosis is combined with automated analysis, enhancing the reliability and accuracy of diagnostic results. This hybrid strategy reduces over-reliance on machine outputs, mitigates data-driven bias, and enhances practical applicability in real-world medical scenarios-an aspect unexplored in existing GAN-based EEG literature.

- **High-Accuracy Diagnosis under Incomplete Data Conditions**: By combining data augmentation with human–machine fusion, the proposed framework maintains strong diagnostic performance even under limited or incomplete EEG data conditions. This approach offers a practical and clinically meaningful solution for intelligent depression diagnosis, providing new insights into AI-assisted neuropsychiatric assessment.

The paper is organized as follows: section 2 reviews related work, and section 3 outlines the methodology. Section 4 presents experiments and results, and section 5 concludes the paper and discusses future work.

2. Related research

Electroencephalogram (EEG)-based emotion recognition has emerged as a pivotal direction in affective computing, providing valuable insights into the neurological underpinnings of mental health disorders such as depression. In recent years, the rapid advancement of artificial intelligence (AI) techniques has accelerated the development of data-driven diagnostic tools in both neuroscience and clinical applications. Among these, EEG analysis stands out due to its non-invasive nature and ability to capture subtle cognitive and affective patterns.

In particular, the integration of machine learning algorithms with clinical expertise has shown great promise in enhancing the accuracy and interpretability of depression diagnosis. This hybrid approach not only addresses the limitations of conventional subjective assessments but also offers practical solutions to data scarcity and annotation challenges commonly encountered in EEG-based research. This section reviews prior studies in depression recognition across multiple modalities—including EEG, speech, and facial expression—and highlights their methodological innovations and remaining limitations.

2.1. Depression diagnosis

Proper diagnosis and treatment are essential to prevent depression. Screening tools such as the Beck Depression Inventory-II (BDI-II) [22], Center for Epidemiologic Studies Depression (CES-D) [23], Hamilton Depression Rating Scale (HDRS) [24], and the three-page Patient Health Questionnaire (PHQ-9) [25] are commonly used to detect depression. Early automatic depression detection primarily relied on handcrafted acoustic or prosodic features extracted from speech [26–28]. While significant progress has been made in speech-based depression recognition, there is still room for improvement, especially in temporal modeling.

Beyond speech, medical imaging–based studies have explored multimodal fusion approaches. For example, the Local-Global Multimodal Fusion Graph Neural Network (LGMF-GNN) model [29] integrates functional MRI, structural MRI, and EHR data, revealing clinically meaningful connectivity abnormalities in MDD patients and enhancing diagnostic accuracy.

With the rise of artificial intelligence, facial-expression–based recognition has also been investigated, where CNNs utilize regions of interest such as the eyes and mouth to distinguish depressive states [30]. It further enhances classification by identifying key regions of interest (ROI), including the facial, eye, and mouth regions, and using them to train a pre-trained 2D CNN model for improved accuracy.

More recently, transformer-based and multimodal architectures have further advanced depression detection. Representative examples include MTNet, which fuses EEG and eye-tracking signals, achieving 91.79% accuracy and demonstrating the advantages of multimodal feature alignment. [31]. In text-based assessment, transformer ensembles such as vanilla BERT, BERTweet, and ALBERT have been employed to estimate depression severity from social media posts [32], while other studies leverage BERT and MentalBERT with additional extra-linguistic cues for detecting depression and stress [33]. Moreover, multimodal transformer frameworks such as TensorFormer [34] further demonstrate the benefits of combining heterogeneous signals. Collectively, these studies highlight a growing trend toward deep multimodal representation learning for depression assessment.

2.2. Human-machine hybrid intelligence in the medical field

As medical decision-making often involves complex, uncertain, and high-dimensional information, integrating human expertise with computational intelligence has become an increasingly valuable strategy in clinical applications. Human–machine hybrid intelligence leverages the strengths of both sides: machine learning models excel at uncovering latent patterns in physiological and behavioral data, while clinicians provide contextual understanding and high-level reasoning that cannot be fully captured by algorithms alone. In depression diagnosis, such synergy enables more accurate and personalized assessments by compensating for the limitations inherent in both human judgment and automated inference.

Within this paradigm, machine learning (ML) and deep learning (DL) have demonstrated strong potential for analyzing EEG signals in neurological and psychiatric evaluation [35]. DL architectures, in particular, are

capable of automatically extracting discriminative neural representations that distinguish major depressive disorder (MDD) patients from healthy individuals [36]. Despite these advantages, conventional ML/DL systems typically rely on centralized training frameworks, which raise practical limitations—most notably significant privacy concerns due to the sensitive nature of medical data, and the need for substantial computational resources for data storage and model optimization [35, 37].

Beyond EEG-only modeling, recent ML research has expanded toward learning from multimodal physiological and behavioral data. For example, ML frameworks have been developed to detect clinically meaningful patterns in imaging, biomedical signals, and electronic health records [38]. By analyzing EEG along with physiological and biochemical markers of depression, Mumtaz *et al* proposed a machine learning approach [39] using synchronization likelihood (SL) features for automatic MDD diagnosis, demonstrating notable potential for early screening. Concurrently, natural language processing (NLP) techniques have enabled innovative approaches to psychological assessment, including computational analyses of counseling dialogues to uncover linguistic markers associated with depressive tendencies [40]. By integrating sequence modeling, message clustering, and psycholinguistic metrics, these methods offer quantitative and interpretable insights for intelligent mental health assessment.

Traditional depression diagnosis methods, such as interviews and questionnaires, are often time-consuming and costly. In addition, some individuals may struggle to verbally express their depressive symptoms, making diagnosis more challenging. To address this issue, a study developed a Persian-language chatbot [41] based on deep learning to assist in diagnosing depression. The chatbot was trained using textual data from both individuals with depression and healthy individuals, including question-and-answer exchanges. Experimental results showed that the chatbot achieved an accuracy of over 85% and an F1 score of 80.5%, outperforming similar studies. These findings highlight the chatbot's potential as a valuable tool for supporting depression diagnosis and treatment.

3. Diagnostic method

This paper proposes an innovative depression diagnosis method based on human-machine hybrid intelligence, aiming to achieve robust depression detection under incomplete data scenarios. The approach consists of three core interconnected steps: data processing, model architecture, and adaptive diagnosis. Each stage builds upon the previous one, collaboratively enhancing the accuracy, robustness, and reliability of the diagnosis, effectively overcoming challenges posed by incomplete or imperfect data.

The method starts by acquiring EEG signals and using independent component analysis (ICA) to remove noise, ensuring clean data for further analysis. Following this, we introduce a data augmentation strategy using DCGAN to generate synthetic data and expand the dataset, facilitating better model training. Then we outline the design of Convolutional Neural Networks (CNN), which are employed to extract key features from the augmented dataset and perform classification. Finally, we discuss an adaptive diagnosis approach, combining the machine-generated predictions with expert clinical assessments to refine the final diagnosis. figure 1 illustrates the overall depression diagnosis system structure.

In this study, we employed the DCGAN architecture to generate synthetic EEG data, thereby expanding the training set and improving model generalization. The DCGAN consists of a Generator and a Discriminator: the Generator produces realistic EEG signals, while the Discriminator learns to distinguish between real and synthetic data. As shown in figure 2, these components are designed to learn EEG data distributions effectively. To prevent data leakage and ensure unbiased evaluation, we first partitioned the original dataset into training and testing subsets, and then applied data augmentation only to the training set.

A convolutional neural network (CNN) is further employed as the classifier to extract both spatial and temporal representations from the preprocessed EEG segments. The overall architecture of the proposed model is shown in figure 3. The network begins with a spatial convolutional layer, which applies 8 filters with a (19×1) kernel, allowing the model to learn inter-channel spatial relationships across the 19 EEG electrodes. Next, a depthwise temporal convolution is performed using a (1×32) kernel with depth multiplier $d = 2$, enabling multi-scale temporal feature extraction from each spatially filtered component. This is followed by a ReLU activation and batch normalization to enhance nonlinearity and stabilize training. A large AveragePooling2D layer with pool size $(1 \times (\text{samples -kernelLength} + 1))$ is then used to collapse the temporal dimension, generating a compact representation of each feature map. Finally, the flattened feature vector is fed into a fully connected dense layer with 2 output units, followed by a softmax activation to classify EEG segments into Healthy or MDD. This architecture is lightweight, stable, and specifically designed for small-sample EEG scenarios with incomplete data, making it suitable for depression detection tasks.

The diagnosis of depression is a complex process that typically requires clinicians to perform multi-dimensional assessments rather than relying on a single test or analysis. Therefore, our diagnostic system

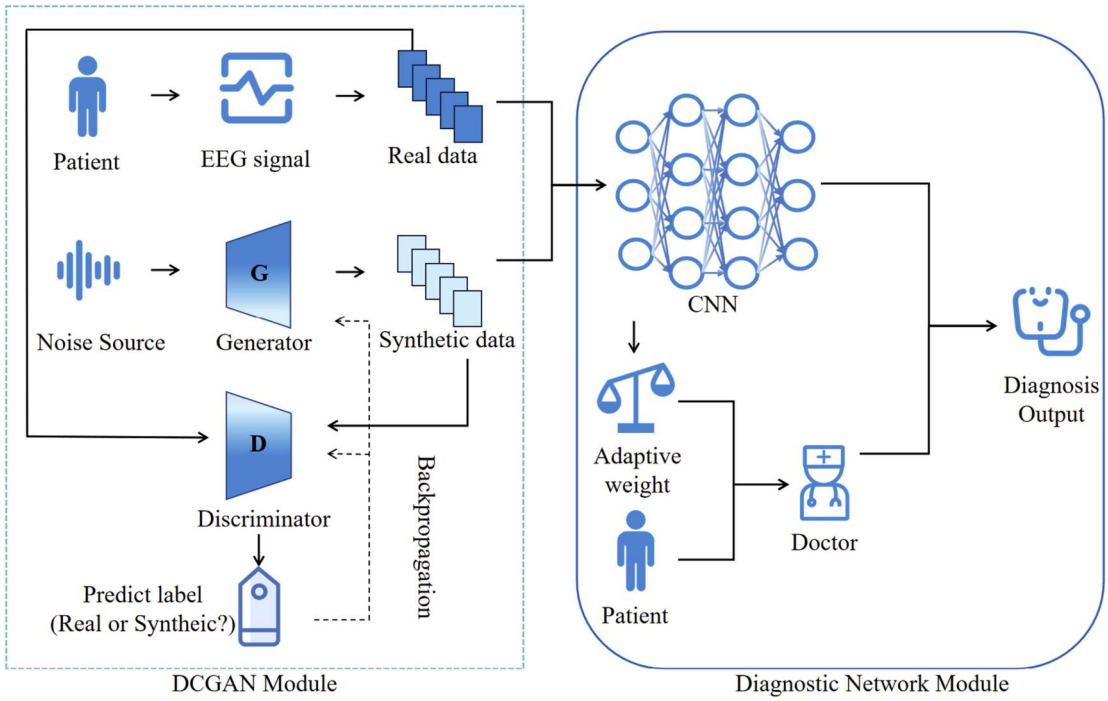


Figure 1. Architecture of the proposed human-machine hybrid depression diagnosis system, which combines EEG data augmentation via DCGAN Module and classification via Diagnostic Network Module, with low-confidence predictions refined by expert.

integrates the expertise of clinicians as a critical criterion for the final diagnosis, while maintaining the integrity of EEG features. The specific steps are as follows:

1. EEG Signal Acquisition: EEG signals are obtained through electrodes placed on the patient's scalp. To avoid interference from electrooculographic (EOG) signals, this study employs Independent Component Analysis (ICA) to isolate and remove artifacts related to eye activity, such as blinking and eye movements.
2. Feature Extraction and Classification: The purified EEG signals are then passed through a trained Convolutional Neural Networks (CNN), which automatically extracts signal features and classifies the signals as either healthy or depressed.
3. Final Diagnosis Integration: The classification results from the CNN are combined with expert clinical assessments using a weighted approach to derive the final depression diagnosis, where adaptive weight adjustment plays a key role.

To better reflect real-world diagnostic scenarios, we implement a dynamic weight adjustment strategy based on specific evaluation metrics, where prediction confidence determines the weight distribution. In this adaptive framework, confidence scores indicate the model's certainty in its predictions and dynamically influence weight allocation. Higher confidence increases the reliance on model-generated results, whereas lower confidence shifts greater weight to expert evaluations. This adaptive mechanism simulates real-world decision-making, ensuring that when the model exhibits high certainty, its influence is prioritized. By effectively integrating expert insights, this method enhances diagnostic precision and reliability, as illustrated in figure 4(a). The main algorithmic concept of the entire article is described in algorithm 1.

The algorithm achieves dynamic adaptive diagnosis through four stages:

In the first stage, ICA is used to denoise raw EEG signals, while DCGAN generates synthetic data to expand the training set, ensuring that generated data is used exclusively for training.

In the second stage, a CNN model is constructed for feature extraction and classification, obtaining initial prediction probabilities for distinguishing patients from healthy individuals through end-to-end training.

In the third stage, an innovative dynamic weight fusion mechanism is introduced—expert evaluation weight $\omega_e^{(i)}$ is dynamically adjusted based on the model's confidence score ρ_i, and the final comprehensive diagnosis is derived using the weighted fusion formula $D_i = \frac{(p_i + s_e^{(i)})}{2}$.

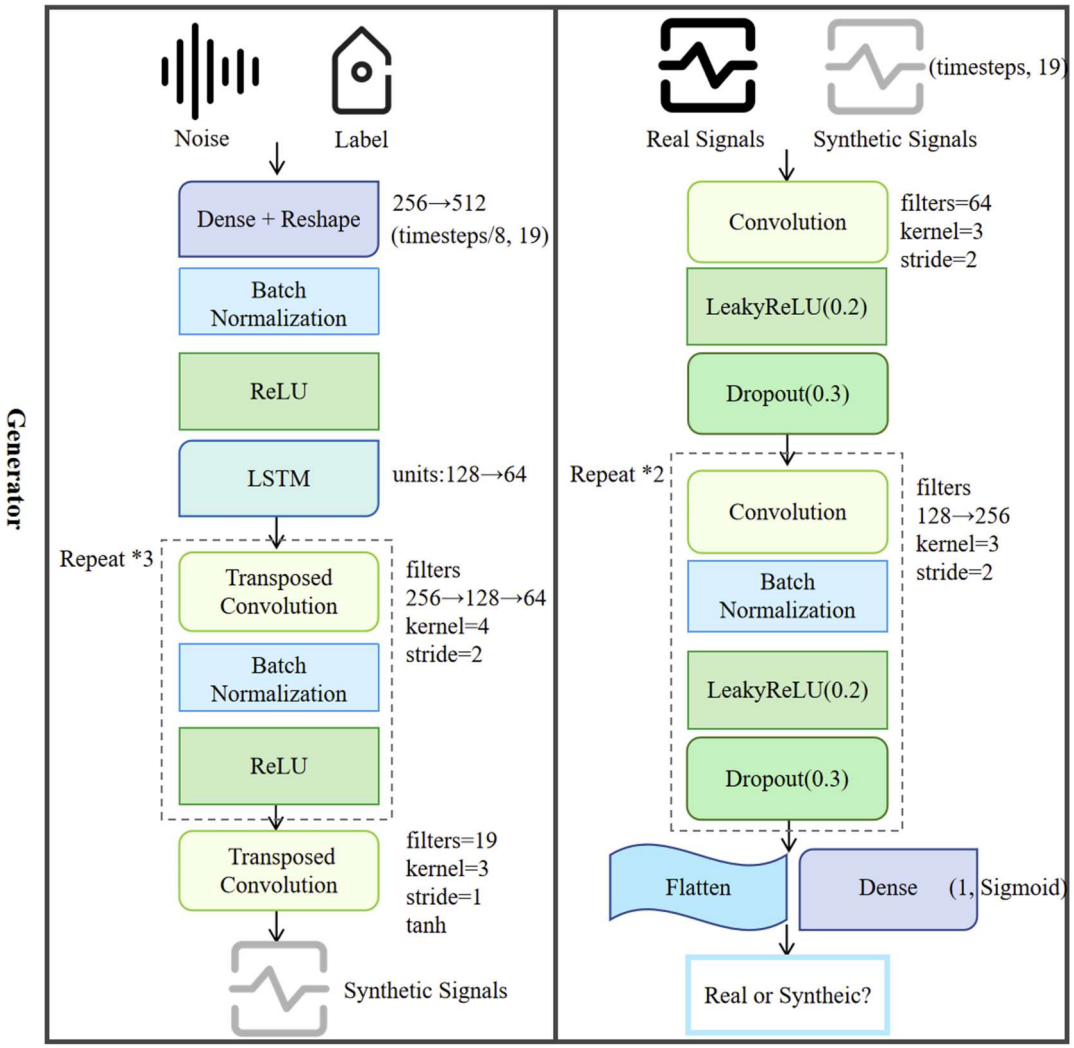


Figure 2. Detailed architecture of the DCGAN model, illustrating the progressive upsampling operations in the generator and the hierarchical downsampling structure of the discriminator.

In the fourth stage, system performance is validated using a confusion matrix and multidimensional evaluation metrics, with model selection based on accuracy, precision, and F1-score.

By leveraging a confidence-driven elastic weight allocation mechanism, this process retains the efficiency of machine learning when the model exhibits high confidence, while enhancing the interpretability of expert knowledge when confidence is low, achieving an optimized synergy between human and machine intelligence.

4. Experiments and results

4.1. Data collection

This study utilized a publicly available EEG dataset collected by Mumtaz *et al* at the Universiti Sains Malaysia Hospital (HUSM). The dataset includes EEG signals from 34 patients with major depressive disorder (mean age 40.33 years, standard deviation ± 12.861) and 30 healthy controls (mean age 38.227 years, standard deviation ± 15.64). The study design was approved by the ethics committee. EEG data were acquired based on the international 10-20 system, with linked ears as the reference. Nineteen electrodes were placed on the scalp, covering the frontal (Fp1, Fp2, F3, F4, F7, F8, Fpz), temporal (T3, T4, T5, T6), parietal (P3, P4, P7, P8), occipital (O1, O2), and central regions (C3, C4), as shown in figure 4(b). The data were high-pass filtered at 0.70 Hz and processed with a 50 Hz notch filter to reduce power line noise. EEG signals were recorded at a sampling rate of 256 Hz. This study focuses on analyzing resting-state EEG signals.

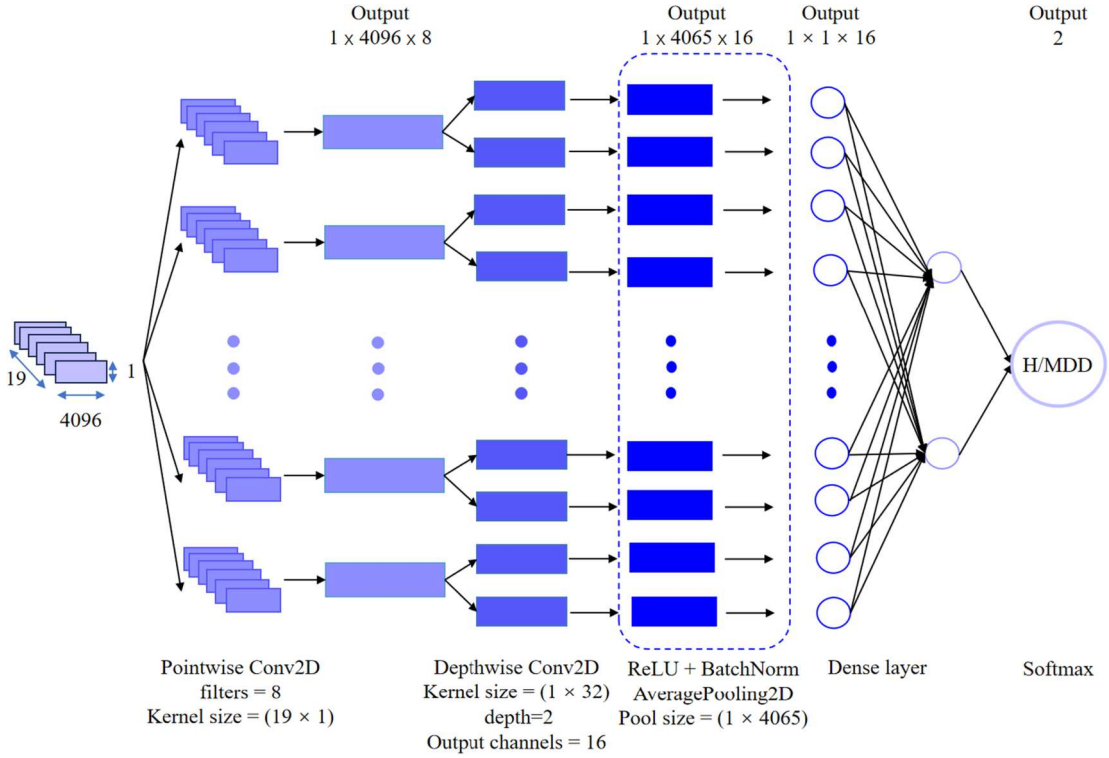


Figure 3. Detailed CNN architecture for diagnosing and classifying EEG signals. The model includes a spatial convolutional layer (8 filters, 19 × 1 kernel), a depthwise temporal convolution layer (1 × 32 kernel, depth multiplier = 2), ReLU activation, batch normalization, and a global temporal average-pooling layer. The extracted features are flattened and fed into a fully connected layer followed by softmax for two-class classification (Healthy versus MDD).

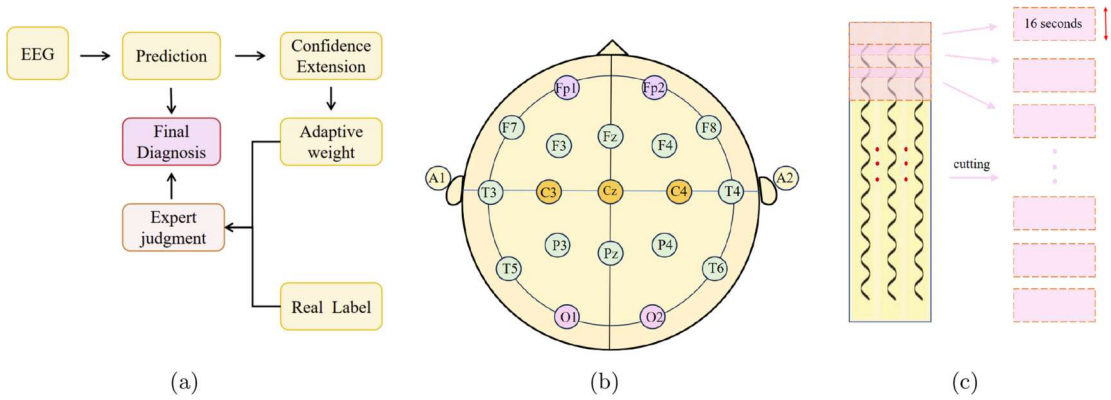


Figure 4. (a). Confidence-based adaptive diagnosis flowchart. (b). EEG Electrode Placement Diagram. (c). Segmented EEG signal processing.

Algorithm 1. Confidence-Adaptive Depression Diagnosis Algorithm

Require:

1: Raw EEG signal $\mathcal{X} \in \mathbb{R}^{n \times c}$ (e.g., $n = 64$ samples(34 depression samples, 30 healthy control samples), $c = 19$ channels)

2: True labels $y \in \{0, 1\}^n$ (0 = healthy, 1 = depression)

3: Confidence score $\rho_i \in (0, 1)$; confidence threshold $\tau = 0.7$

Ensure: Comprehensive diagnosis result $D \in [0, 1]^n$, evaluation metrics

4: **Stage 1: Data Preprocessing**

5: ICA denoising: $\tilde{\mathcal{X}} \leftarrow \text{ICA}(\mathcal{X})$

6: Split data: $\mathcal{X}_{\text{train}}, \mathcal{X}_{\text{test}}, y_{\text{train}}, y_{\text{test}} \leftarrow \text{split}(\tilde{\mathcal{X}}, y)$

7: DCGAN augmentation: $\mathcal{X}_{\text{aug}} \leftarrow \text{DCGAN}(\mathcal{X}_{\text{train}})$

(Continued.)

8: Expand training set: $\mathcal{X}_{\text{train}} \leftarrow \mathcal{X}_{\text{train}} + \mathcal{X}_{\text{aug}}$
9: **Stage 2: Model Training and Prediction**
10: Initialize CNN model $\mathcal{M}$ with parameters θ_m
11: Train model: $\theta_m^* \leftarrow \arg\min_{\theta_m} \mathcal{L}(\mathcal{M}(\mathcal{X}_{\text{train}}), y_{\text{train}})$
12: Predict probabilities: $p_i \leftarrow \text{softmax}(\mathcal{M}(\mathcal{X}_{\text{test}}[i]))$, for $i = 1, \ldots, n_{\text{test}}$
13: Predict labels: $\hat{y}_i \leftarrow \arg\max(p_i)$
14: **Stage 3: Dynamic Weight Fusion**
15: **for** each sample $i \in \mathcal{X}_{\text{test}}$ **do**
16: Prediction confidence: $\rho_i \leftarrow \max(p_i)$
17: Model weight: $\omega_m^{(i)} \leftarrow \rho_i$
18: Expert weight:

$$\omega_e^{(i)} \leftarrow \alpha(1 - \rho_i), \quad \alpha = \begin{cases} 1, & \rho_i < \tau \\ 0.4, & \rho_i \geqslant \tau \end{cases}$$

19: Expert judgment: $s_e^{(i)} \leftarrow \omega_e^{(i)} \cdot y_i$
20: Model judgment: $s_m^{(i)} \leftarrow \omega_m^{(i)} \cdot p_i$
21: Final diagnosis: $D_i \leftarrow \frac{s_e^{(i)} + s_m^{(i)}}{2}$
22: **end for**
23: **Stage 4: Performance Evaluation**
24: Confusion matrix: $CM \leftarrow \text{ConfusionMatrix}(y_{\text{test}}, \hat{y})$
25 Compute metrics: {Accuracy, Precision, Recall, F1, ...} ← Metrics(CM)
26:**return** D, CM, {Accuracy, Precision, F1}

4.2. Preprocessing of data

4.2.1. Removal of artifacts

When processing EEG data, identifying and removing interference signals from other physiological activities is crucial to ensure data quality. These interference signals can include artifacts such as eye blinks, eye movements, and heartbeat-related noise, which may obscure the true brain activity signals in the EEG recordings. To address this issue, the study employed Independent Component Analysis (ICA), specifically the efficient FastICA algorithm, to separate and remove these artifact signals. This approach helps retain cleaner EEG signals, which are essential for subsequent analysis.

4.2.2. Data segmentation, augmentation, and partitioning

The original dataset consists of EEG recordings from 64 subjects, including 30 healthy individuals and 34 patients diagnosed with major depressive disorder (MDD). Using DCGAN generated 225 synthetic healthy samples and 230 synthetic depressive samples.

Then, the original 5-minute EEG signals were segmented into shorter 16-second intervals to further enhance the diversity and effectiveness of model training, resulting in a total of 1200 segments. To maintain signal continuity and preserve key features, each segment (except the first and last) partially overlapped with the preceding and subsequent segments. This approach prevents signal discontinuity while retaining the core characteristics of the EEG data, as shown in figure 4(c). With a sampling rate of 256 Hz, each segment contained 4096 data points per channel.

To ensure that model evaluation reflects performance on real EEG signals, the train–test partition was conducted strictly on the real dataset. Specifically, 90% of the real samples were assigned to the training set, while the remaining 10% were held out as the final test set. Within the 90% training pool, 10-fold cross-validation was applied for model training and validation (figure 5). After obtaining the train–validation split in each fold, DCGAN-generated samples were added exclusively to the training subset. Both the validation subset and the independent test set contained only real EEG data, ensuring an unbiased estimate of generalization performance.

4.2.3. Feature extraction

The EEG signals in this study were recorded from 19 channels, each capturing electrophysiological activity from different brain regions. Seven features were extracted from each channel, including relative power in four frequency bands (delta wave: 0.5–4 Hz; theta wave: 4–8 Hz; alpha wave: 8–12 Hz; beta wave: 12–30 Hz) and three nonlinear features (sample entropy, Higuchi's fractal dimension, and the Hurst exponent), which are key indicators of brain functional states. These features were then used to analyze differences in EEG signals for depression diagnosis, as they are closely related to the brain functional differences observed in depressive patients. Furthermore, these features were utilized to train deep learning classification models, enabling efficient and accurate diagnosis of depression.

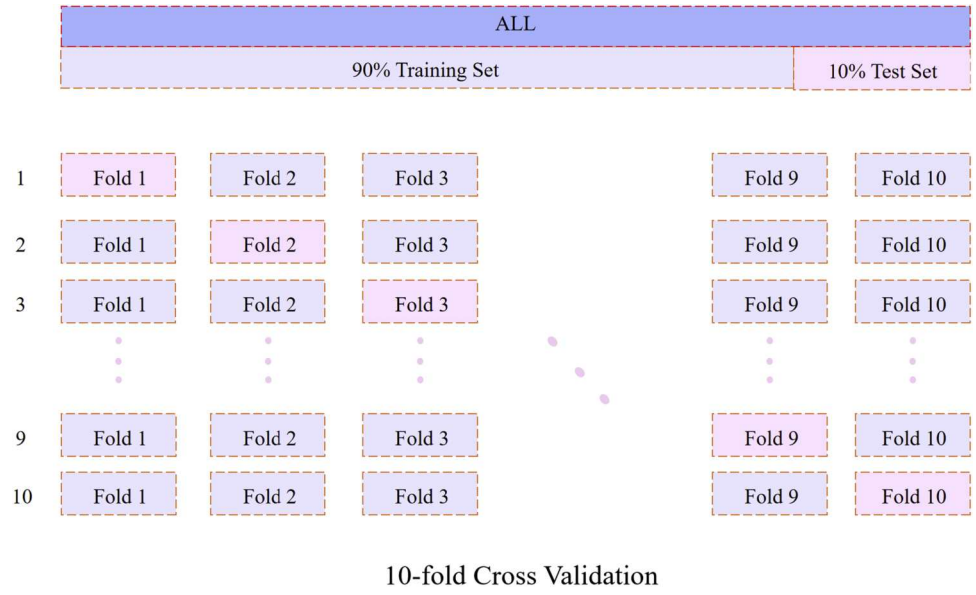


Figure 5. 10-fold cross-validation.

4.3. Diagnostic network

In the statistical analysis phase, we extracted features from both real and DCGAN-generated EEG signals of healthy and depressive individuals to provide a comprehensive comparison. As illustrated in figure 6, two sets of comparisons are presented: the delta wave (0.5–4 Hz) characteristics of the original EEG signals and those of the generated EEG signals. Both comparisons consistently reveal clear differences between healthy and depressive groups, indicating that delta-band activity is strongly associated with mental health status. The similarity between real and synthetic patterns further demonstrates that the generated EEG preserves meaningful physiological characteristics.

To evaluate the effectiveness of the CNN diagnostic network in our experiments, we trained the raw EEG data using CNN, KNN, and SVM models separately and compared their performance. The classification performance of each model is illustrated in figure 7, with detailed numerical results summarized in table 1. As shown in the figure, the CNN model consistently achieves higher classification accuracy than both KNN and SVM across all evaluation steps. Furthermore, the shaded error bands in the figure represent the standard deviation across multiple experimental runs, reflecting the robustness and stability of the CNN model. The CNN's higher average accuracy clearly demonstrates its superior generalization performance for EEG-based depression diagnosis.

For a comprehensive comparison, we adopted standard evaluation metrics as defined in equation 1:

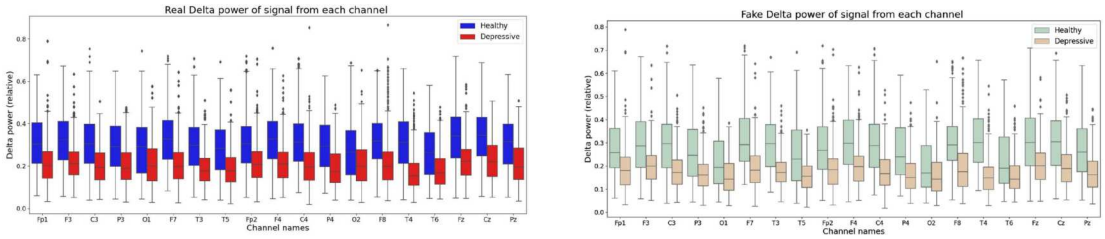
$$\begin{cases} Accuracy = \dfrac{TP + TN}{TP + FN + FP + TN} \\ Precision = \dfrac{TP}{TP + FP} \\ F\text{-}score = \dfrac{2 \times Recall \times Precision}{Recall + Precision} \end{cases} \tag{1}$$

where TP represents True Positive, FN represents False Negative, FP represents False Positive and TN represents True Negative.

Accuracy refers to the proportion of correctly classified samples relative to the total number of samples. Precision is the proportion of correctly predicted positive samples among all samples predicted as positive. The F1 score, which is the harmonic mean of precision and recall, provides a balanced measure that accounts for both precision and recall. Recall, or sensitivity, indicates the proportion of correctly predicted positive samples out of all actual positive samples.

4.4. Data generation

In this study, we used DCGAN to generate 225 healthy EEG samples and 230 depressive EEG samples. To verify the similarity in features between the synthetic data and the real data, and to ensure that the synthetic data could be used for CNN model training, we extracted spectral features (delta, theta, alpha, and beta wave power) and nonlinear features (sample entropy, Higuchi's fractal dimension, and the Hurst exponent) from both the



(a) Original EEG: delta wave (0.5–4 Hz) (b) Generated EEG: delta wave (0.5–4 Hz)

Figure 6. Comparison of delta wave (0.5–4 Hz) between healthy and depressed subjects: (a) original EEG; (b) DCGAN-generated EEG. The generated signals preserve key spectral characteristics of real EEG, demonstrating their physiological plausibility.

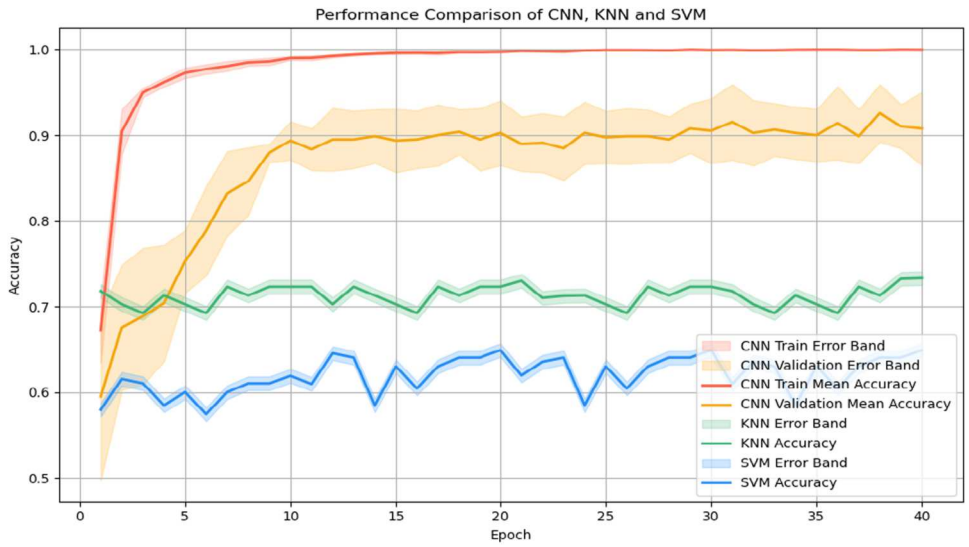


Figure 7. Performance comparison of CNN, KNN, and SVM on raw data classification.

synthetic and real data. Statistical analyses, including t-tests, ANOVA, and boxplots, were conducted to compare feature distributions across channels. The results showed that the synthetic data closely resembled real data, especially in terms of spectral features, confirming the reliability of the synthetic samples.

We divided the data into healthy and depressive groups and compared the distributions of the Hurst exponent feature between the synthetic and real data. The x-axis represents the long-term memory of the time series, while the y-axis represents the kernel density estimate, which shows the probability density of data points at each x-axis value. A higher y-axis value indicates a greater probability of a data point appearing at the corresponding x-axis position. The purple curve represents real data, and the green curve represents synthetic data, with the healthy group shown in figure 8. It can be observed that the distribution curves of the two datasets have a large overlap in the central region, providing visual evidence that not only confirms the statistical similarity between the synthetic and real samples, but also emphasizes the considerable potential of GAN technology in the field of medical data augmentation.

To further validate the effectiveness of the generated data, both real and generated EEG samples were input into a pre-trained CNN model for depression diagnosis. The results, shown in figure 9, demonstrated that the model's performance using generated data was comparable to that using real data. This confirms that the DCGAN model effectively generates data that closely resembles real EEG signals, thereby contributing to a more diverse dataset and enhancing the model's generalization ability for depression diagnosis.

Overall, our findings not only provide a powerful tool for depression diagnosis but also lay a solid foundation for the future application of GAN technology in the diagnosis and research of neuropsychiatric disorders. With ongoing advancements and optimizations in technology, we hope that the generated EEG data can be progressively improved and contribute to societal benefits, offering valuable insights for clinical diagnosis and scientific research.

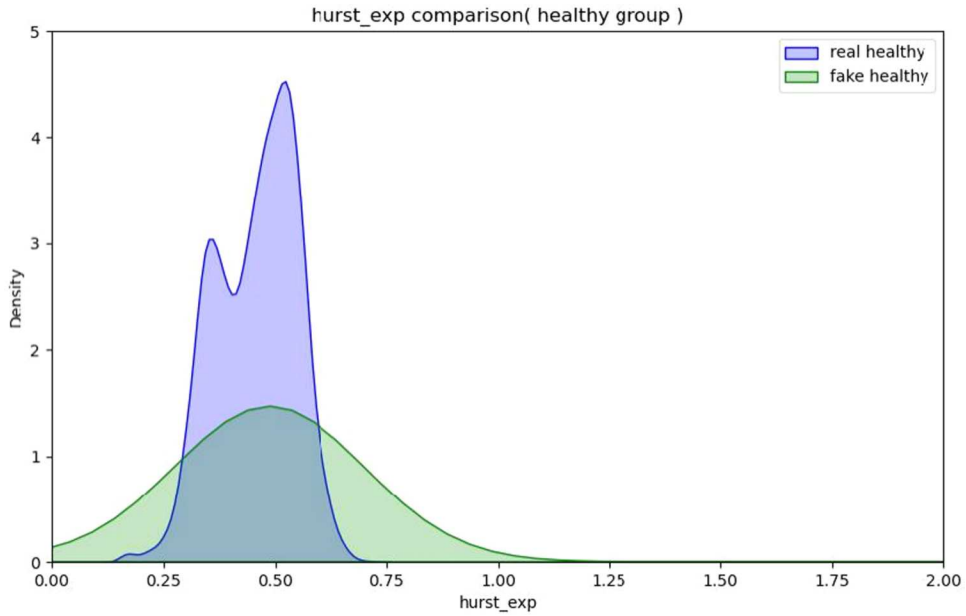


Figure 8. Comparison of Hurst exponent distribution between synthetic and real healthy data.

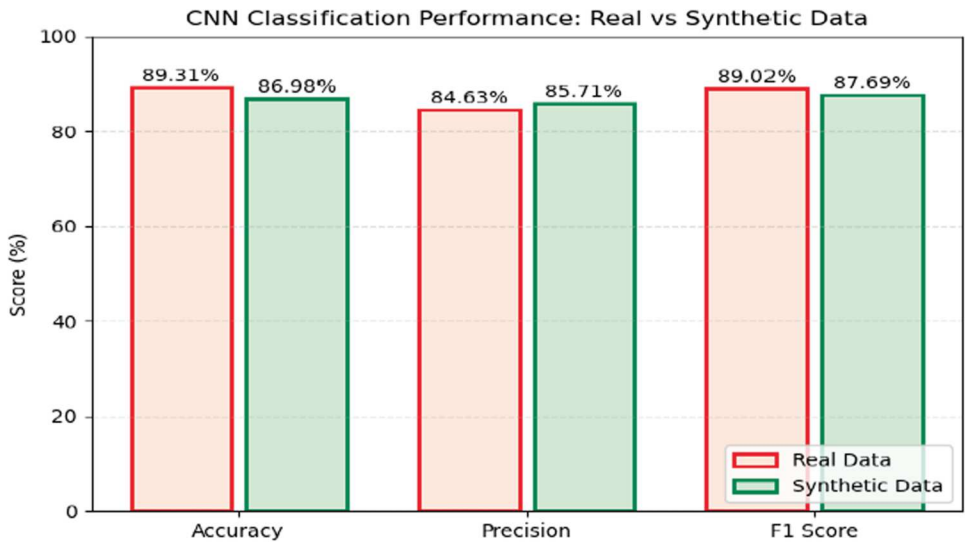


Figure 9. The computed classification results of CNN, trained separately on real and synthetic data.

Table 1. Classifying raw data using three models: CNN, KNN, and SVM.

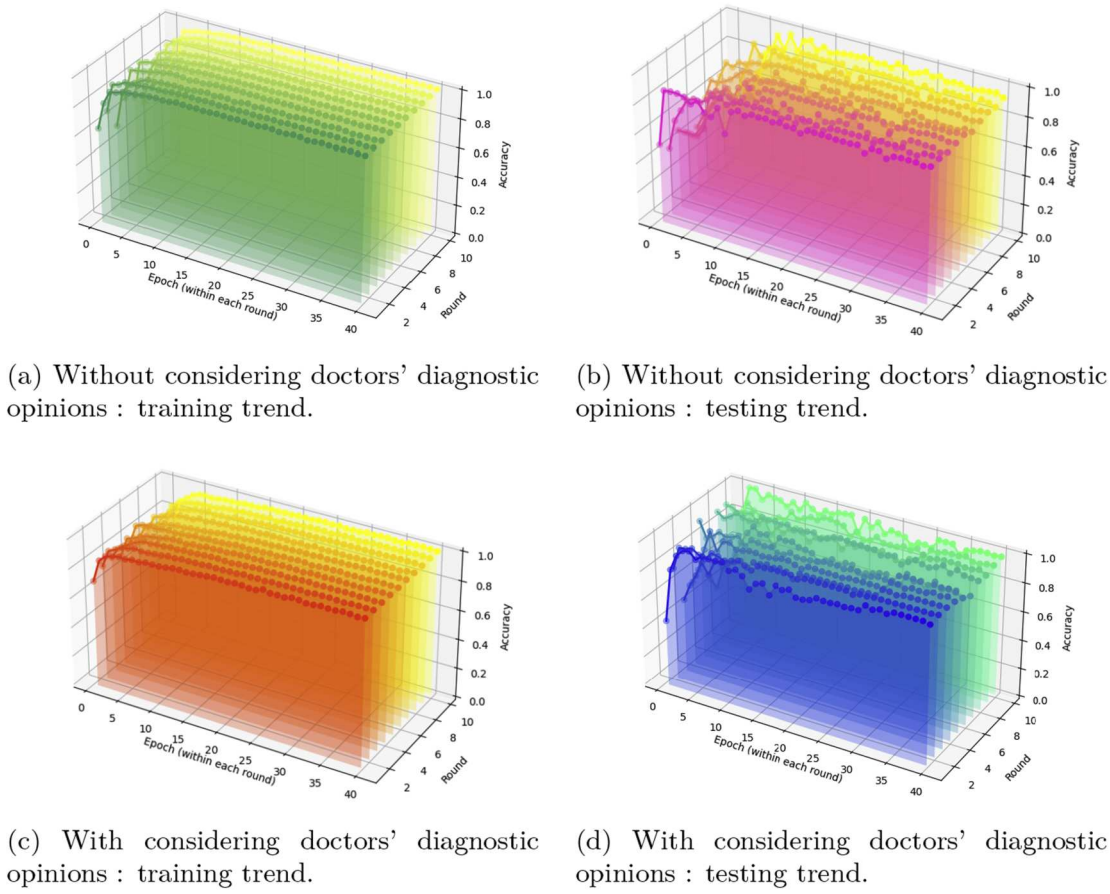
	Aver-accuracy	Aver-precision	Aver-F1 score
CNN	89.31%	84.63%	89.02%
KNN	71.01%	67.28%	69.80%
SVM	59.47%	55.70%	57.78%

4.5. Depression diagnosis

To situate our work within the broader research landscape and to provide a clear overview of the application of Generative Adversarial Networks (GANs) for EEG data augmentation in various classification tasks, we have compiled a comparative summary of relevant studies [42–44] in table 2.

Table 2. Comparison of GAN-based EEG Data Augmentation Methods for Classification Tasks.

Study	Method	Application	Dataset	Performance
Zhang *et al* (2018)	Conditional DCGAN (cDCGAN)	Motor Imagery Classification	BCI Competition II Dataset III	82.86% accuracy (raw), 82.86% (synthetic), about 84% with augmentation
Ling *et al* (2022)	Residual Dense Block (RDB)-DCGAN	Sleep Stage Classification	Sleep-EDF Database	+6% overall accuracy, +19% for N1 stage
Carrie *et al.* (2023)	Conditional Wasserstein GAN (WGAN) with CNN	Major Depressive Disorder (MDD) Diagnosis	Cai *et al.* EEG Dataset; Mumtaz *et al* EEG Dataset	+10% accuracy on the first dataset, no significant improvement on the second dataset
Proposed method	DCGAN + Human-Machine Hybrid CNN	Depression Diagnosis	Mumtaz *et al* EEG Dataset	97.20% accuracy, 96.36% precision, 97.38% F1-score



(a) Without considering doctors' diagnostic opinions : training trend.

(b) Without considering doctors' diagnostic opinions : testing trend.

(c) With considering doctors' diagnostic opinions : training trend.

(d) With considering doctors' diagnostic opinions : testing trend.

Figure 10. Training and testing accuracy curves of the CNN model under two conditions: with and without consideration of the doctor's diagnosis.

This table synthesizes key methodologies, target applications, datasets, and reported performance gains across different studies, including our own. The comparison reveals a consistent trend: GAN-based data augmentation is a potent strategy for mitigating data scarcity and class imbalance in EEG analysis, leading to performance improvements across diverse domains such as brain-computer interfaces, sleep staging, and mental disorder diagnosis. It also highlights the methodological evolution from standard DCGANs to more stabilized variants like WGAN and architecturally enhanced models like RDB-DCGAN. Our proposed method, which integrates DCGAN-based augmentation with an adaptive human-machine hybrid model, achieves state-of-the-art performance on the depression diagnosis task, as evidenced by the highest accuracy and F1-score among the compared studies. This structured comparison underscores the effectiveness of our approach and facilitates a direct understanding of its contributions relative to existing literature.

To validate the depression diagnosis system proposed in this paper, we conducted two experiments to assess its performance in diagnosing depression. The dataset was divided into two groups: depressed patients and healthy individuals.

In both experiments, after preprocessing the data and splitting it into training and testing sets, we applied data augmentation to the training set and assessed the model's performance using 10-fold cross-validation to ensure robustness and reliability.

In the first experiment, we evaluated the performance of the CNN model without incorporating doctors' diagnostic opinions. The classification accuracy on the training and testing sets is shown in figures 10(a) and (b), respectively. Quantitative results in table 3 further confirm the model's strong generalization capability in distinguishing between healthy and depressed states, highlighting the potential of machine learning in assisting diagnostic tasks. The confusion matrices are presented in table 4.

In the second experiment, we incorporated the doctor's diagnosis, which was based on patient interactions, questionnaires, and observations. Each data segment was labeled as 0 for healthy or 1 for depressed, and these labels were used as the doctor's diagnostic results. During network training, one-hot encoding was applied to facilitate subsequent hybrid AI diagnosis. Additionally, a dynamic weight adjustment method based on the model's confidence score was implemented to better simulate real-world diagnostic scenarios. In this experiment, the data labels were used as the doctor's diagnosis results.

Table 3. Results of Experiments 1 and 2.

Experiment	Accuracy	Precision	F1-score
Experiment 1	91.07%	86.21%	90.90%
Experiment 2	97.20%	96.36%	97.38%

Table 4. Confusion matrices on the test set for Experiments 1 and 2.

Experiment	Experiment 1		Experiment 2	
	Pred. Depressed	Pred. Healthy	Pred. Depressed	Pred. Healthy
Actual Depressed	58	2	59	1
Actual Healthy	9	51	2	58

The model predicted a probability distribution for each sample, with the maximum predicted probability serving as the confidence score. These confidence values were processed into a two-dimensional array consistent with the label shape to generate adaptive expert judgments. Finally, by combining and averaging the adaptive expert judgments with the model's predictions, a comprehensive diagnostic result was obtained. This dynamic weight adjustment method effectively balances the contributions of the model's predictions and the doctor's diagnosis, improving diagnostic accuracy and reliability. Experimental results are presented in figures 10(c), (d), table 3 and the confusion matrices in table 4.

5. Conclusion

In this study, we propose a diagnostic framework tailored for depression detection under conditions of incomplete and limited EEG data. The approach integrates human–machine collaborative intelligence with a deep convolutional generative adversarial network (DCGAN) to address the challenges posed by data scarcity. By generating synthetic EEG signals, the DCGAN module effectively augments the training dataset, while convolutional neural networks (CNNs) are employed to extract meaningful features. Additionally, a confidence-based fusion mechanism dynamically adjusts the weight between model predictions and expert evaluations, leading to more accurate final diagnostic outcomes. Experimental results demonstrate that the proposed method achieves superior performance in terms of accuracy, precision, and F1 score compared to conventional models.

This study has several limitations. First, the size of the available EEG dataset is relatively small, which may restrict the diversity of both real and generated samples. Second, we adopted a baseline conditional DCGAN architecture to assess the feasibility of EEG data augmentation. More advanced and stable GAN variants—such as Wasserstein GANs, GANs with gradient penalty, or models incorporating spectral normalization—were not explored. Third, the current framework focuses exclusively on EEG signals, and its applicability to other physiological modalities remains unexamined.

Although the proposed CNN classifier effectively captures spatial–temporal EEG representations, the current work does not incorporate post-hoc interpretability methods such as Grad-CAM, saliency maps, or EEG-specific attribution techniques. Such visualization tools are valuable for enhancing clinical interpretability and understanding model decision patterns. In this study, however, we focus primarily on evaluating the feasibility of GAN-based EEG augmentation and CNN-based classification. As a direction for future work, we will integrate explainable AI (XAI) techniques tailored for multichannel EEG to facilitate more transparent and clinically meaningful interpretations.

Another important limitation is that cross-subject EEG generalization was not evaluated. Cross-subject EEG analysis constitutes an independent and complex research direction, which typically requires systematic methods and sufficiently large-scale datasets. Similarly, cross-dataset validation is beyond the scope of this study due to the limited availability and substantial heterogeneity of public depression-related EEG datasets.

Beyond methodological constraints, potential demographic or clinical imbalances in the datasets may introduce bias and affect model performance when applied to other populations. Regulatory compliance and comprehensive clinical validation will be required before any deployment in real-world healthcare settings. Ethical considerations concerning data usage, potential biases, and clinical implications have been addressed;

all data were de-identified and used in accordance with their respective data usage agreements, and the framework is intended solely for research and proof-of-concept purposes.

For future work, we plan to improve the GAN architecture to generate more diverse and higher-quality EEG samples, collect larger-scale EEG datasets, and extend the framework toward multimodal physiological monitoring, incorporating signals such as heart rate variability (HRV) and electrodermal activity (EDA). In addition, we aim to further explore cross-subject and cross-dataset generalization, for example by leveraging transfer learning and multi-domain adversarial training methods, to enhance the model's adaptability across different subjects or datasets. We also plan to integrate explainable AI (XAI) techniques tailored for multi-channel EEG to facilitate more transparent and clinically meaningful interpretations. Through these directions, we strive to provide more accurate diagnoses and more effective treatment options for individuals with depression, ultimately improving their quality of life.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant 62203006, 62173317.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

References

[1] Vohryzek J, Sanz-Perl Y, Kringelbach M L and Deco G 2025 Human brain dynamics are shaped by rare long-range connections over and above cortical geometry *Proc. Natl Acad. Sci.* **122** e2415102122
[2] Hagan A T, Xu L, Klugah-Brown B, Li J, Jiang X and Kendrick K M 2025 The pharmacodynamic modulation effect of oxytocin on resting state functional connectivity network topology *Frontiers in Pharmacology* **15** 1460513
[3] Kreivinienė B, Šaltytė-Vaisiauskė L and Mačiulskytė S 2025 Therapeutic effect of proprioceptive dolphin assisted activities on health-related quality of life and muscle tension, biomechanical and viscoelastic properties in major depressive disorder adults: case analysis *Front. Hum. Neurosci.* **18** 1487293
[4] Huang Y *et al* 2019 Prevalence of mental disorders in china: a cross-sectional epidemiological study *The Lancet Psychiatry* **6** 211–24
[5] Hu T-w 2004 The economic burden of depression and reimbursement policy in the asia pacific region *Australasian Psychiatry* **12** s11–5
[6] Otte C, Gold S M, Penninx B W, Pariante C M, Etkin A, Fava M, Mohr D C and Schatzberg A F 2016 Major depressive disorder *Nature Reviews Disease Primers* **2** 1–20
[7] Rehm J and Shield K D 2019 Global burden of disease and the impact of mental and addictive disorders *Current Psychiatry Reports* **21** 1–7
[8] Pampouchidou A, Simos P G, Marias K, Meriaudeau F, Yang F, Pediaditis M and Tsiknakis M 2017 Automatic assessment of depression based on visual cues: A systematic review *IEEE Trans. Affective Comput.* **10** 445–70
[9] Liao P 2020 Analysis of depression for aging people in china *In 2020 International Conference on Public Health and Data Science (ICPHDS* 8–12
[10] Oluboka O J, Katzman M A, Habert J, McIntosh D, MacQueen G M, Milev R V, McIntyre R S and Blier P 2018 Functional recovery in major depressive disorder: providing early optimal treatment for the individual patient *Int. J. Neuropsychopharmacolog.* **21** 128–44
[11] Balakrishna N, Mukesh Krishnan M B and Ganesh D 2024 Handling depression: An in-depth study on its causes and treatment through novel machine learning approaches *In 2024 8th International Conference on Electronics, Communication and Aerospace Technology (ICECA* 830–5
[12] Teplan M *et al* 2002 Fundamentals of eeg measurement *Measurement Science Review* **2** 1–11
[13] Cai H *et al* 2018 A pervasive approach to eeg-based depression detection *Complexity* **2018** 5238028
[14] Derry C P, Duncan J S and Berkovic S F 2006 Paroxysmal motor disorders of sleep: the clinical spectrum and differentiation from epilepsy *Epilepsia* **47** 1775–91
[15] Xu S, Wang Z, Sun J, Zhang Z, Wu Z, Yang T, Xue G and Cheng C 2020 Using a deep recurrent neural network with eeg signal to detect parkinson's disease *Annals of Translational Medicine* **8** 1–9
[16] Lee S, Hussein R, Ward R, Jane Wang Z and McKeown M J 2021 A convolutional-recurrent neural network approach to resting-state eeg classification in parkinson's disease *J. Neurosci. Methods* **361** 109282
[17] Sakkalis V 2011 Review of advanced techniques for the estimation of brain connectivity measured with eeg/meg *Comput. Biol. Med.* **41** 1110–7 Special Issue on Techniques for Measuring Brain Connectivity
[18] Kang M, Oh S H, Oh K S, Kang S and Lee Y 2021 The deep learning method for predict beck's depression inventory score using eeg *In 2021 International Conference on Information and Communication Technology Convergence (ICTC* 490–3
[19] Pamungkas Y, Wibawa A D and Rais Y 2022 Classification of emotions (positive-negative) based on eeg statistical features using rnn, lstm, and bi-lstm algorithms *In 2022 2nd International Seminar on Machine Learning, Optimization, and Data Science (ISMODE* 275–80
[20] Sakib N, Islam K and Faruk T 2023 Machine learning based depression screening among young adults using wireless eeg *In 2023 International Conference on Artificial Intelligence Innovation (ICAII)* 110–5
[21] Li Y, Cheng N, Gu J, Zhang Q, Song J, Ding Z and He S 2024 A human-machine hybrid approach for diagnosing depression based on convolutional neural networks *In 2024 14th Asian Control Conference (ASCC)* 129–34
[22] BECK DEPRESSION Inventory-Ii 2010 Beck depression inventory-ii *Corsini Encycl. Psychol* **1** 210

[23] Radloff L S 1991 The use of the center for epidemiologic studies depression scale in adolescents and young adults *Journal of Youth and Adolescence* **20** 149–66
[24] Williams J B W 1988 A structured interview guide for the hamilton depression rating scale *Archives of General Psychiatry* **45** 742–7
[25] Kroenke K, Spitzer R L and Williams J B W 2001 The phq-9: validity of a brief depression severity measure *Journal of General Internal Medicine* **16** 606–13
[26] France D J, Shiavi R G, Silverman S, Silverman M and Wilkes M 2000 Acoustical properties of speech as indicators of depression and suicidal risk *IEEE Trans. Biomed. Eng.* **47** 829–37
[27] Cohn J F, Kruez T S, Matthews I, Yang Y, Nguyen M H, Padilla M T, Zhou F and De la Torre F 2009 Detecting depression from facial actions and vocal prosody *In 2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops* 1–7
[28] Yang Y, Fairbairn C and Cohn J F 2013 Detecting depression severity from vocal prosody *IEEE Trans. Affective Comput.* **4** 142–50
[29] Liu S *et al* 2024 An objective quantitative diagnosis of depression using a local-to-global multimodal fusion graph neural network *Patterns* **5** 1–23
[30] Acharya R and Dash S P 2022 Automatic depression detection based on merged convolutional neural networks using facial features *In 2022 IEEE International Conference on Signal Processing and Communications (SPCOM* 1–5
[31] Zhu F, Zhang J, Dang R, Hu B and Wang Q 2025 Mtnet: Multimodal transformer network for mild depression detection through fusion of eeg and eye tracking *Biomed. Signal Process. Control* **100** 106996
[32] Ahmed T, Ivan S, Munir A and Ahmed S 2024 Decoding depression: Analyzing social network insights for depression severity assessment with transformers and explainable ai *Natural Language Processing Journal* **7** 100079
[33] Ilias L, Mouzakitis S and Askounis D 2023 Calibration of transformer-based models for identifying stress and depression in social media *IEEE Transactions on Computational Social Systems* **11** 1979–90
[34] Sun H, Chen Y-W and Lin L 2022 Tensorformer: A tensor-based multimodal transformer for multimodal sentiment analysis and depression detection *IEEE Trans. Affective Comput.* **14** 2776–86
[35] Rahul J, Sharma D, Sharma L D, Nanda U and Sarkar A K 2024 A systematic review of eeg based automated schizophrenia classification through machine learning and deep learning *Front. Hum. Neurosci.* **18** 1347082
[36] Subhani A R, Mumtaz W, Saad M N B M, Kamel N and Malik A S 2017 Machine learning framework for the detection of mental stress at multiple levels *IEEE Access* **5** 13545–56
[37] Umair M, Tan W-H and Foo Y-L 2024 Dynamic federated learning aggregation for enhanced intrusion detection in iot attacks *2024 International Conference on Artificial Intelligence in Information and Communication (ICAIIC* 524–9(IEEE)
[38] Ed-Driouch C, Mars F, Gourraud P-A and Dumas C 2022 Addressing the challenges and barriers to the integration of machine learning into clinical practice: An innovative method to hybrid human–machine intelligence *Sensors* **22** 8313
[39] Mumtaz W, Saad Azhar Ali S, Yasin M A M and Malik A S 2018 A machine learning framework involving eeg-based functional connectivity to diagnose major depressive disorder (mdd) *Med. Biol. Eng. Comput.* **56** 233–46
[40] Althoff T, Clark K and Leskovec J 2016 Large-scale analysis of counseling conversations: An application of natural language processing to mental health *Transactions of the Association for Computational Linguistics* **4** 463–76
[41] Sohofi A, Lesani F S and Jamalou Z 2024 Persian depression detection chatbot using deep learning *In 2024 10th International Conference on Web Research (ICWR)* 370–4
[42] Zhang Q and Liu Y 2018 *Improving Brain Computer Interface Performance by Data Augmentation with Conditional Deep Convolutional Generative Adversarial Networks* arXiv:1806.07108
[43] Ling H, Luyuan Y, Xinxin L and Bingliang D 2022 Staging study of single-channel sleep eeg signals based on data augmentation *Frontiers in Public Health* **10** 1038742
[44] Carrle F P, Hollenbenders Y and Reichenbach A 2023 Generation of synthetic eeg data for training algorithms supporting the diagnosis of major depressive disorder *Frontiers in Neuroscience* **17** 1219133