Journal of Innovative Image Processing (ISSN: 2582-4252)
www.irojournals.com/iroiip/

Adversarial Temporal Modeling for Seizure Classification Using WGAN-GP-LSTM Networks

Hema P.[1], Vanithamani R.[2]

[1]Research Scholar, Department of Electronics and Communication Engineering, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, India.
[2]Professor, Department of Biomedical Instrumentation Engineering, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, India.

E-mail: [1]phemachoudary@gmail.com, [2]vanithamani_bmie@avinuty.ac.in

Abstract

Detecting an epileptic seizure with the help of electroencephalogram (EEG) signals is difficult due to the absence of classified data or complexity. The conventional GAN-based data augmentation methods suffer from instability and inadequate learning of dependencies, which limits their efficiency despite increasing the diversity of data. An LSTM method is proposed for binary seizure classification with the WGAN-GP-LSTM scheme to overcome the previous issues. The generator and the discriminator of the system proposed in this paper use LSTM units, which differentiate it from other previously proposed systems. The system can produce acceptable EEG signals. This also helps in the correlations being learned over time for seizure classification. Through this system, EEG datasets that contain signals with and without supplements are publicly available. Results indicate that using synthetic data augmentation will enhance the WGAN-GP-LSTM model's performance. Consequently, this model outperforms WGAN-GP and other prior works on the overall score, which are 0.96, 0.875, 0.82, and 0.923, respectively, for overall accuracy, sensitivity, specificity, and F1 score. WGAN-GP, with an FID score of 3.07 is poorer than WGAN-GP with a score of 1.42. The WGAN-GP-LSTM model displaces adversarial sequence generation and classification, helping to overcome challenges of data scarcity and the temporal dependence of EEG signals for seizure detection. A different approach to computer -assistance in neurodiagnostics has been introduced.

Keywords: Electroencephalography, Epileptic Seizure Detection, Generative Adversarial Networks, Long Short-Term Memory Networks, Deep Learning.

1. Introduction

An epileptic seizure triggered by abnormal brain electrical discharges occurs frequently, causing a chronic neurological disorder. A seizure can last anywhere from a few seconds to several minutes, and it may be accompanied by sensory anomalies (auras), full-body convulsions, altered consciousness, or uncontrollable motor activity. It is estimated that 1% of the worldwide population suffers from epilepsy, one of the most common neurological disorders. Predicting seizure onset early allows treatment measures to be put in place and significantly improves the quality of life for people with seizure disorders [3]. For diagnosing

DOI: https://doi.org/10.36548/jiip.2026.1.019
Received: 09.01.2026, received in revised form: 06.02.2026, accepted: 23.02.2026, published: 06.03.2026

and monitoring epilepsy, EEG remains the gold standard, as it offers high temporal resolution when capturing brain signals. It is, however, time-consuming, labor-intensive, and prone to inter-observer variability when interpreting EEG recordings manually. Biomedical research has therefore become increasingly interested in automated seizure detection and prediction systems.

Computer vision tasks have been greatly enhanced by the introduction of GANs, a powerful and novel framework. Contrary to traditional machine learning algorithms, GANs make use of adversarial training to learn and represent features more effectively. Despite their flexibility, they are also difficult to control because of their high degrees of freedom, which can lead to non-convergence and mode collapse. Research activities have focused on optimizing the application of GANs in computer vision and signal processing [4], [14]. The soft Lipschitz constraint through gradient penalties reduces the convergence issues. This technique has led to the implementation of WGAN-GP in signal analysis [15]. Complex time-dependent seizure activity is captured by integrating LSTM with GAN frameworks [16].

The study focuses on enhancing seizure detection performance through data augmentation techniques. The integration of WGAN-GP with LSTM combines the generative and temporal learning capabilities, respectively in the classification of seizure types.

2. Related Work

A customized Convolutional Neural Network (CNN) with a GAN for the detection of seizures from EEG signals is proposed [1]. The signals are preprocessed to remove noise and artifacts and segmented to isolate the informative regions. A specialized CNN model is utilized for feature extraction. The model is tested on the BONN, CHB-MIT, and SWEC-ETHZ datasets. The model achieved an accuracy of 99.95%, 98.93% and 98.80% for the respective datasets. The authors in [3] have stated that not all the datasets are efficient in training the model. Hence, augmentation of data is necessary where there are limited data. The method experiments with GAN for data augmentation on the CHB-MIT dataset. An occurrence period of 30 minutes and a prediction horizon of 5 minutes make up the initial stages of the model. The proposed approach gives an ROC 61% of and can replace long-term real-world recordings, addressing class imbalance and datascarcity issues.

In [7], the authors mention that manual examination of EEG processes is time-consuming and requires better judgment skills for accurate diagnosis. The method implements automatic detection of seizures using deep learning methods. A Time-Aware Convolutional Neural Network with a Recurrent Neural Network (TA-CNN-RNN) is implemented for classification. GAN with Cramer distance (CGAN) is used for generating accurate data for each label. A spatiotemporal error factor is introduced to differentiate between fake and actual data. The model achieved an accuracy of 94.6%, 94.8% and 95.2% on the CHB-MIT, BONN-iEEG, and VIRGO-EEG datasets, respectively. Imbalanced datasets pose a threat to the accurate diagnosis of seizures from EEG signals [19]. The authors have addressed the issue by investigating different data augmentation techniques. These techniques include WGAN, Vannila GAN, Conditional GAN, CGAN, and Random Forest (RF) for classification purposes. From the experiments, it is observed that WGAN performed well compared to others. This WGAN is then integrated with a Bi-LSTM model for the identification of seizures. The model achieved a maximum accuracy of 91.73% on augmented data compared to 86% on real-time data without augmentation [20].

A new deep learning-based approach employs LSTM with an auxiliary classifier GAN to train on expert-annotated and augmented spike events from iEEG recordings of epileptic patients [21]. The augmented data is classified using Support Vector Machine (SVM) and RF network models. Cross- institutional generalization performance is tested across both datasets.

2.1 Motivation Behind the Study and Its Intended Objectives

The models trained are either supervised or unsupervised learning paradigms. In the case of supervised models, the network is trained using data and their corresponding labels. The network then predicts the actual tags of the testing data. Backpropagation is a technique that iteratively changes weights to reduce prediction errors by propagating the error backward through the network between expected and actual outputs. In neural network research and applications, this paradigm is frequently used [11].

Conversely, unsupervised learning makes use of training without output labels. The network searches the input data for hidden patterns or structures without external supervision. GANs and their variants are generally considered unsupervised models. This categorization results from the generator learning to produce data that closely resembles the real data distribution with only random noise as input—that is, without any class labels. By using input from the discriminator, which labels created and genuine samples as "real" or "fake," the generator is indirectly trained. This adversarial training approach shares conceptual parallels with reinforcement learning and aligns with unsupervised learning principles [12].

While the discriminator is trained in a supervised manner using binary labels (real or fake), GANs cannot be classified as fully supervised systems—especially when multi-class labels are not used or cannot be directly enforced on the output layer. Therefore, traditional supervised learning strategies are not applicable in such multi-category generative settings. Integrating WGAN-GP with Bi-LSTM has addressed the class imbalance along with the accurate diagnosis of seizures [19, 20].

The remaining sections of this paper are organized as follows: The basic ideas and design of GANs are explored in Section 3. The suggested WGAN-GP-LSTM framework is described in detail in Section 4. The results are presented and analyzed in Section 5. The study is finally concluded in Section 6, which also offers suggestions for future research directions and a summary of the main findings.

3. Generative Adversarial Networks

Two interacting neural networks make up GANs: the discriminator, which establishes whether a given sample is produced or real, and the generator, which generates synthetic data samples from random noise. The training process relies on binary classification, labeling samples as either real or fake, rendering the framework inherently unsupervised with respect to class-specific information. As a result, conventional supervised learning approaches cannot be directly applied, and the producer does not have any direct control over the category of their production. To overcome this limitation, conditional GANs (cGANs) have been developed, which integrate additional information such as class names into both the generator and the discriminator. This extension allows for efficient sampling from a learned distribution without having to specify a Probability Density Function (PDF). GANs are a class of non-supervised

generative models that are used to learn a data distribution to augment a set of data by classifying it into categories [8].

GANs offer a promising alternative to conventional methods of estimating maximum likelihood. By avoiding heuristic loss functions such as pixel-wise average squared error, GANs are particularly suited for representation learning tasks. Despite their potential, the training of GANs remains a major challenge, often leading to unstable convergence and unrealistic results. Moreover, there is a notable gap in the literature concerning the interpretation and visualization of the internal representations within deep, multi-layer GAN models [9].

In medical image and signal analysis, GANs have shown great promise in tackling both established and new problems, including reconstruction, data augmentation, segmentation, classification, and regression [13]. Their ability to create realistic outputs close to real ones is noteworthy. A GAN consists of two neural networks that can be trained simultaneously. The first is the generator, which generates the false data, and the latter is the discriminator, which identifies the real and false data. The principle of GANs is illustrated in Figure 1.

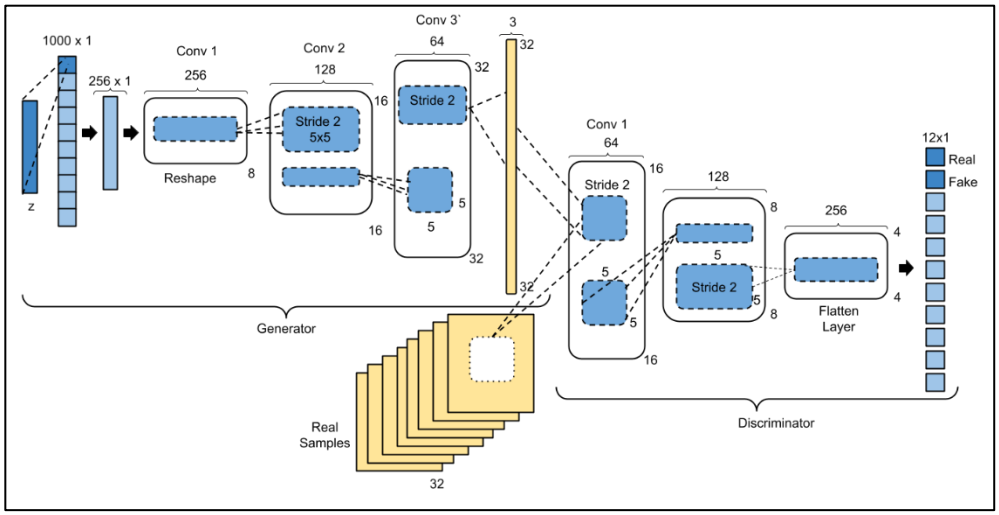


Figure 1. Structure of the Generator and Discriminator

3.1 The Generator

G is a neural network that attempts to generate false samples similar to the actual dataset from $z \sim p_z$ random through only by forward propagation. The generator (G) receives a random noise vector z, which is sampled from the latent space using either the uniform or Gaussian distribution. G applies a sampling procedure to transform this noise into a pseudo-sample distribution, which is referred to as G(z). The generator is trained to maximize the log probability that the discriminator classifies the generated samples as real. expressed as $E_{z \sim p_z(z)}[\log\left(D\left(G(z)\right)\right)]$. The loss is calculated from the discriminator's output and backpropagated to update G's parameters as appropriate. The architectural settings of the discriminator (D) and G are shown in Fig.2.

Drawing inspiration from [9], the figure depicts G and D, the two main components of a GAN. The binary classification is performed by the discriminator, which tries to discern between created and actual images, while the generator uses a latent vector z to produce

synthetic (false) images. In an adversarial architecture, backpropagation is used to train both networks concurrently.

The objective of G during training is expressed in Eqn. (1),

$$L_D^{GAN} = E_{x\sim P_r}[D(x)] - E_{\bar{x}\sim P_g}[D(\bar{x})] + \lambda E_{\hat{x}}[(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1)^2 \quad (1)$$

3.2 The Discriminator

The D, which is designed to distinguish between generated and actual samples, is necessary to provide feedback that changes the setting of G. While values close to 0 indicate that the sample is likely to be false, higher numbers indicate that the sample is more likely to be true. It returns the probability of the reality of the input. When the D's output is ideally converging to about 0.5, the generator has learned to replicate the true data distribution and has shown that it is unable to reliably discriminate between true and false samples. This will train the discriminator to optimize the probability of correctly classifying the data: $\mathbb{E}_{x\sim P_{data}(x)}[\log(D(x))]$. In addition to reducing the probability of misclassification of the samples generated, by $E_{z\sim P_{z(z)}}[\log(1 - D(G(z)))]$, the proposed work ensures that the generator progressively improves its capability to create data distributions resembling the real samples. This minimizes the divergence between the synthetic and the original data, thereby enhancing the overall stability of the adversarial training process. The cumulative objective function of a discriminator is formally expressed as in Eqn. 2:

$$L_{GAN}^D = \max_D \mathbb{E}_{x_r\sim P_r(x)}[\log(D(x_r))] + \mathbb{E}_{x_g\sim P_g(x)}\left[\log\left(1 - D(x_g)\right)\right], L_G = -\mathbb{E}_{N\sim P_g}[D(x)] \quad (2)$$

4. Proposed Methodology

The methodology proposed in this work aims to improve the detection of seizures from EEG signals using a generative adversarial learning framework. The core objective of the method is to address the limitations posed by data scarcity and class imbalance in EEG datasets by incorporating synthetic EEG signals generated using advanced GAN architectures. The proposed pipeline integrates data preprocessing, segmentation, and normalization, followed by the generation of realistic EEG signals using WGAN-GP. To further improve temporal feature representation, the GAN architecture is extended by embedding LSTM units into the discriminator. This hybrid WGAN-GP-LSTM model is trained to differentiate more effectively between real and synthetic EEG segments. An improved dataset containing both enhanced synthetic signals and real signals is then used to train a model to detect seizures reliably. Performance metrics are calculated using the publicly available database CHB-MIT to evaluate the effectiveness of the proposed methodology. Figure 2 presents the methodology used in this work. WGAN-GP-LSTM generated synthetic EEG data were utilized to train a standalone LSTM-based seizure classifier.

The proposed architecture uses WGAN-GP-LSTM solely to provide additional data through data augmentation; the critic does not classify the seizure; instead it assists the generator by providing information about how far away the generated distribution of EEG is from the true distribution of EEG using the Wasserstein distance measurement. The final

classification of the seizure is performed using a different LSTM classifier that was trained using the augmented dataset.

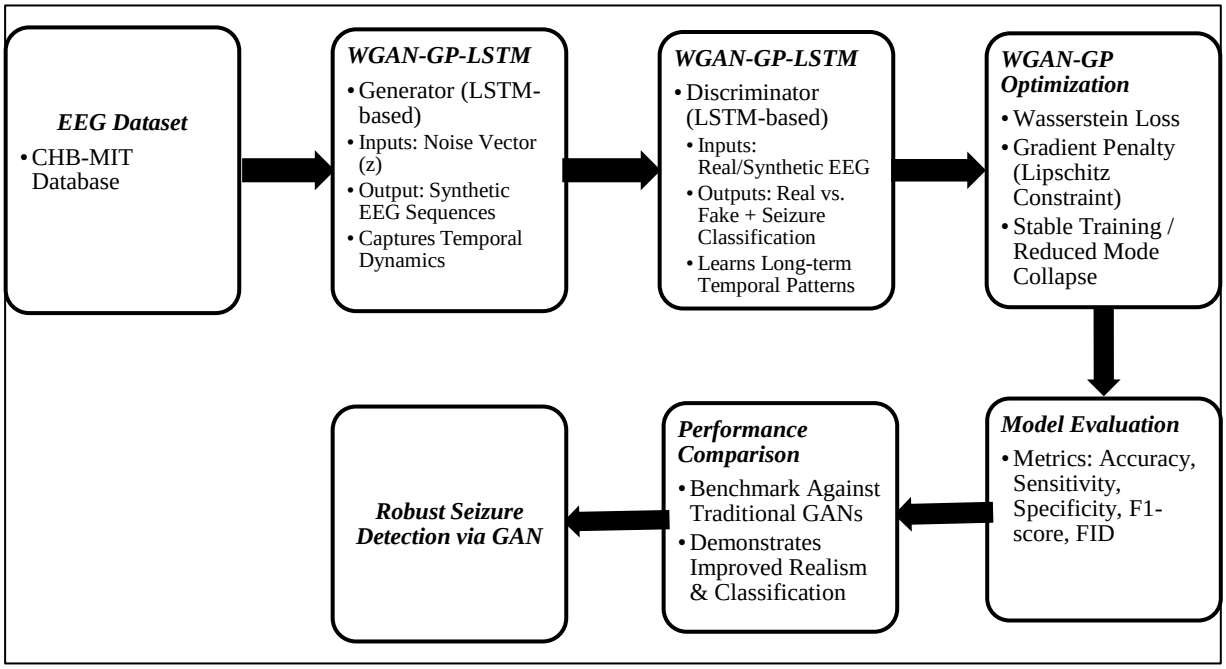


Figure 2. Workflow of the Proposed Work

4.1 Dataset

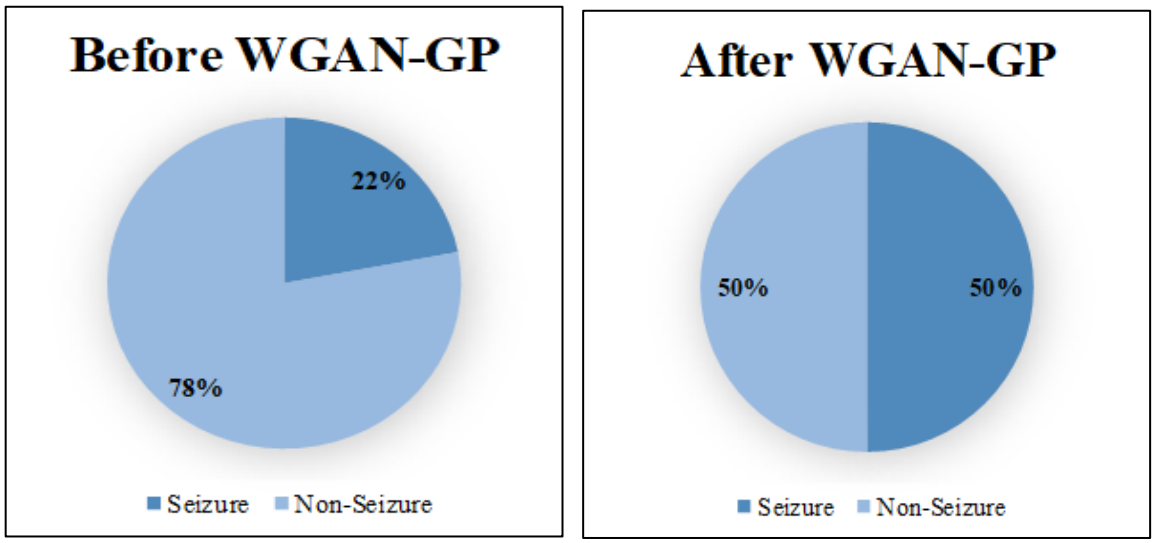
This study uses EEG scalp recordings from 23 pediatric study group of 23 participants, available in the CHB-MIT database, to assess the algorithms for detecting seizures [18]. The impact of different seizure volumes on the accuracy of detection has been studied in several studies. The dataset is divided into non-interrelated training and test sets for each patient using the time-division technique [17]. The subjects are monitored for several days, which includes the withdrawal of seizure medications. The data is originally in the form of ". edf" files and the signals are sampled at a rate of 256 samples per second with 16-bit resolution. Table 1 summarizes the dataset used in the study. The dataset suffers from a major imbalance, which is corrected by using WGAN-GP. The record file contains 664 ".edf" files, and they are further augmented using the WGAN-GP technique.

Table 1. CHB-MIT Dataset Summary

Parameter	Description
Number of subjects	23 pediatric patients
Total EEG records	664 EDF files
Sampling frequency	256 Hz
Channels used	23 EEG channels
Total seizures	~ 198 seizures
Seizure duration	10s ~ several minutes
Non-seizure duration	Continuous interictal EEG
Class labels	Seizure / non-seizure
Segmentation window	Fixed-length EEG segments
Data imbalance	Severe (dominant non-seizure class)

Fixed channel selection may skew the model because seizure onset zones differ greatly between patients. Thus, patient-independent learning is guaranteed via a full-channel approach. Longer ictal durations were associated with slightly higher sensitivity; synthetic augmentation is more beneficial for generalized seizures; and no patient experienced a decline in performance following augmentation. The creation of synthetic data demonstrated resilience to inter-subject variability by improving seizure detection consistency among patients with different seizure patterns.

The CHB-MIT dataset shows significant inter-patient seizure heterogeneity and lacks consistent, trustworthy seizure subtype identification across patients. Clinical priority: seizure detection is more important than subtype labeling. Class imbalance and instability would exacerbated by multi-class modeling.



a). Before Implementing WGAN-GP b). After Implementing WGAN-GP

Figure 3. Structure of the Generator and Discriminator

4.2 Preprocessing

The power-line interference and the noise are initially removed from the signal using a median filter and normalization is done using the z-score method [2]. The signals are segmented into fixed length temporal windows, and the data is labeled as seizure and non-seizure. The data imbalance is addressed using the WGAN-GP model. Figure 3 shows the data before and after applying the WGAN-GP model. In order to guarantee robustness and clinical relevance, particularly in situations with limited and inconsistent labeling, binary seizure detection is implemented.

4.3 WGAN-GP-LSTM

This work aims to address the issues of class imbalance and inadequate data in seizure identification. The WGAN model has already proven to be efficient in addressing this issue [19, 20]. Figure 3 highlights the efficiency of WGAN-GP in addressing the same. However, GANs face issues like unstable training and vanishing gradients. Hence, this work proposes integrating WGAN-GP with LSTM to achieve stable training by enforcing the Lipschitz continuity condition using gradient penalties while replacing the Jensen–Shannon divergence with the Wasserstein distance [5].

The WGAN-GP model lacks the ability to capture the time-dependent dependencies of EEG signals; however, it exhibits improved dynamics compared to other generator-discriminator models. To overcome this limitation, the proposed architecture integrates layers of LSTM in both the generator and the discriminator. LSTMs are well-suited for sequence modeling and allow the network to learn long-term time relationships that are essential for distinguishing between seizure and non-seizure patterns in EEG data.

This hybrid WGAN-GP-LSTM framework is designed to perform both data augmentation through realistic EEG sequence generation and classification by learning discriminative temporal features.

4.3.1 Architectural Description

The WGAN-GP-LSTM model for EEG seizure identification combines the WGAN-GP and LSTM networks to successfully identify the temporal dynamics observed in EEG signals. The basic notion is that signals are generated by training a generator from random noise, which are evaluated for validity by a discriminator, called a critic in the WGAN system. This architecture uses an LSTM network to construct the critic, modeling temporal dependencies in sequential data. The generator considers a noise vector $z \in \mathbb{R}^{100}$, from a uniform distribution sampled at $U[-1, 1]$, and turns it into a synthetic EEG sequence of shape $T \times F$, where F is the features of EEG and T is the time steps. In generator architecture, batch normalization, ReLU activation, reshaping, upsampling, and one-dimensional convolutional layers are typically preceded by a fully connected layer. The Tanh activation function occurs last to check whether the output values lie in the range of [−1,1].

The technique uses a multilayer LSTM to extract temporal information from both generated and actual EEG recordings. The final LSTM output is passed through a dense layer to produce a scalar number that represents the distance between the real and the synthetic distributions of the data as Wasserstein's distance. By using gradient penalty combined with the distance to the Earth (Wasserstein-1) as its loss metric, WGAN-GP implements a Lipschitz constraint, unlike traditional GANs that use Jensen-Shannon divergence. This approach significantly increases training stability and generated sequence fidelity, as seen in Eqn. (3).:

$$\min_{G} \max_{D \epsilon \mathcal{D}} \mathbb{E}_{x \sim P_r}[D(x)] - \mathbb{E}_{\bar{x} \sim P_g}[D(\bar{x})] \tag{3}$$

Here, P_g stands refers to the sample distribution produced by G, and P_r for the distribution of actual data. Eqn. (4) defines the gradient penalty component that is necessary to meet the Lipschitz continuity criteria that is required by the WGAN framework:

$$\lambda . \mathbb{E}_{\hat{x} \sim P_{\hat{x}}}[(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1)^2] \tag{4}$$

where $\hat{x}$ is sampled uniformly along uninterrupted lines connecting pairs of real and synthesized samples, and "λ is a regularization coefficient" (set to 10). During training, D is updated more frequently than the generator (commonly at a 5:1 ratio), and Adam is used as the optimizer with parameters α=0.0001, β1=0, and β2=0.9. The loss for D includes the Wasserstein distance plus the gradient penalty, while the generator aims to minimize the negative critic score. The EEG windows are divided into a mini-batch size of 128, following a training phase of 20 epochs. Using LSTM in the discriminator enables the algorithm to interpret the temporal structure inherent in EEG signals, making it more sensitive to the abnormal rhythmic patterns associated with seizure occurrences.

The training stability of the GAN architecture is achieved by incorporating a carefully designed convolutional structure based on established guidelines. The architecture uses strided convolutions in the discriminator and fractional-strided (or transposed) convolutions in the generator to replace traditional pooling layers. Batch normalization is used to improve the efficiency of the training and to stabilize learning dynamics. In addition, the discriminator uses Leaky-ReLU activation to increase gradient flow and reduce the potential for resting neurons. To maintain the deep structure of the network, the fully connected layers are excluded from the hidden layers. The generator structure utilizes ReLU in the intermediate layers and the Tanh activation function in the final layer. These modifications ensure that the signal falls within the normalized range.

4,98,502 segments of EEG signals are generated and fed as input to the WGAN-GP architecture. The LSTM model is integrated to improve the temporal dependency of the EEG signals. In the final stage of the generator model, two LSTM layers with 100 hidden units are configured. They help capture the long-term dependencies of the EEG series data. The layers involved, along with their descriptions, are briefed in Table 2. A fully connected layer is placed below the two LSTM layers, each with 100 hidden units, ensuring sufficient temporal modeling without introducing excessive parameters.

The LSTM layer within the GAN is embedded in both the generator and critic to model the temporal coherence of EEG sequences, ensuring physiologically plausible synthetic EEG signals. The classification is performed using a separate deep temporal classifier trained on the augmented dataset.

Table 2. WGAN-GP-LSTM Architecture for CHB-MIT Database

Generator	
Layer	**Description**
Input	Random noise vector
FC layer	Expands the noise vector suitable shape for further processing
Batch Normalization (BN) layer	Normalizes output to stabilize training
ReLU activation layer	To introduce non-linearity
Reshape layer	Reshapes the vector to 3D for LSTM (timestep x features x 1)
LSTM layer	Captures temporal dependencies in EEG sequences
FC layer	Reduces output to 512 samples.
Tanh activation layer	Bounds output values between -1 and 1
Output layer	Synthetic EEG signal
Discriminator	
Layer	**Description**
Input	EEG signal input
Reshape layer	Reshape 1D input to [timesteps x features x 1] for LSTM
LSTM layer	Captures temporal dependencies in EEG sequences
LSTM layer	Captures temporal dependencies in EEG sequences
Dropout layer	Reduces overfitting by randomly deactivating the neurons
FC layer	Converts LSTM output to a scalar value
No Activation (Linear)	Final output for Wasserstein distance estimation (no sigmoid)
Output	Real/fake score
Gradient Penalty (WGAN-GP Specific)	
Step	**Description**
Interpolated Samples	Random interpolation between real and synthetic samples
Gradient Computation	Compute gradients of the critic's output w.r.t input
Gradient Norm Penalty	Penalize deviation of gradient norm from 1 to enforce Lipschitz constraint

In the course of GAN training, the discriminator acts as a critic, producing a scalar Wasserstein score rather than class probabilities. In the case of inference from the results, the discriminator is not included, and the seizure classification is done by the LSTM model. The formal definition for the GAN training is included as in Eqn. 5.

$$D(x): \mathbb{R}^{TxF} \rightarrow \mathbb{R} \tag{5}$$

Where $D(x)$ estimates the Wasserstein distance, not $P(y|x)$. A class determination is not inferred from the critic's output. Its only function is to direct the generator to generate EEG patterns that are chronologically realistic.

By simulating long-range EEG dependencies, the combined use of LSTM in both adversarial components maintaining temporal consistency, resulting in more realistic synthetic sequences and better downstream categorization.

4.3.2 Pseudocode for WGAN-GP-LSTM

Input
EEG Dataset $X = \{(X_i, y_i)\}_{i-1}^{N}, X_i \in R^{TxC}$Latent noise $z \sim N(0, I_d)$
Number of epochs E, Batch Size B, Critic iterations n_{critic}, Gradient penalty coefficient λ_{gp}
Learning rates α_D, α_G
Stage 1: Preprocessing
Segment EEG into fixed windows $X_i \in R^{TxC}$
Normalize each channel $X_{i,t,c} \leftarrow \frac{X_{i,t,c} - \mu_c}{\sigma_c}$
Stage 2: Initialize Networks
Initialize generator θ, discriminator $\emptyset$, and learning rates of the optimizer α_D, α_G.
Stage 3: Adversarial training
FOR epoch = 1 TO E DO
  FOR each mini-batch $\{X_r\}$ of size B
    FOR k = 1 TO n_{critic} DO
      Sample real EEG batch $X_r \sim P_r$
      Sample noise $z \sim N(0, I_d)$
      Generate synthetic EEG:

$$X_f = G_\theta(z)$$

      Compute critic scores:

$$D_{real} = D_\varphi(X_r)$$
$$D_{fake} = D_\varphi(X_f)$$

      Compute Wasserstein loss:

$$L_D = E[D_{fake}] - E[D_{real}]$$

      Sample $\varepsilon \sim U(0,1)$
      Interpolate:

$$\hat{X} = \varepsilon X_r + (1 - \varepsilon)X_f$$

      Compute gradient:

$$g = \nabla_{-}\{\hat{X}\}\, D\_\varphi(\hat{X})$$

      Compute gradient penalty:

$$L_{GP} = \lambda_{gp} \cdot E[(||g||_2 - 1)^2]$$

      Total critic loss:

$$L_{D_{total}} = L_D + L_{GP}$$

Update critic parameters:

$$\varphi \leftarrow \varphi - \alpha_D \nabla_\varphi L_{D_{total}}$$

END FOR

$Sample\ noise\ z \sim N(0, I_d)$

Generate synthetic EEG:

$$X_f = G_{\theta(z)}$$

Generator loss:

$$L_G = -E\left[D_{\varphi(X_f)}\right]$$

Update generator parameters:

$$\theta \leftarrow \theta - \alpha_G \nabla_\theta L_G$$

END FOR

END FOR

Stage 4: EEG Data augmentation

Generate synthetic samples

$$\chi_{syn} = \{G_\theta(z_j)\}_{j-1}^{M}$$

Augment dataset

$$\chi_{aug} = \chi \cup \chi_{syn}$$

Stage 5: LSTM-based classification

Initialize classifier parameters ψ

FOR each training batch in $X_{aug} DO$

Forward pass-through LSTM:

$$h_t = LSTM(X_t, h_{\{t-1\}})$$

Output:

$$\hat{y} = Softmax(W\, h_T + b)$$

Compute cross-entropy loss:

$$L_{cls} = -\Sigma\, y\, log(\hat{y})$$

Update classifier parameters:

$$\psi \leftarrow \psi - \alpha_{cls} \nabla_\psi L_{cls}$$

END FOR

Stage 6: Performance Evaluation, FID calculation, Seizure classification

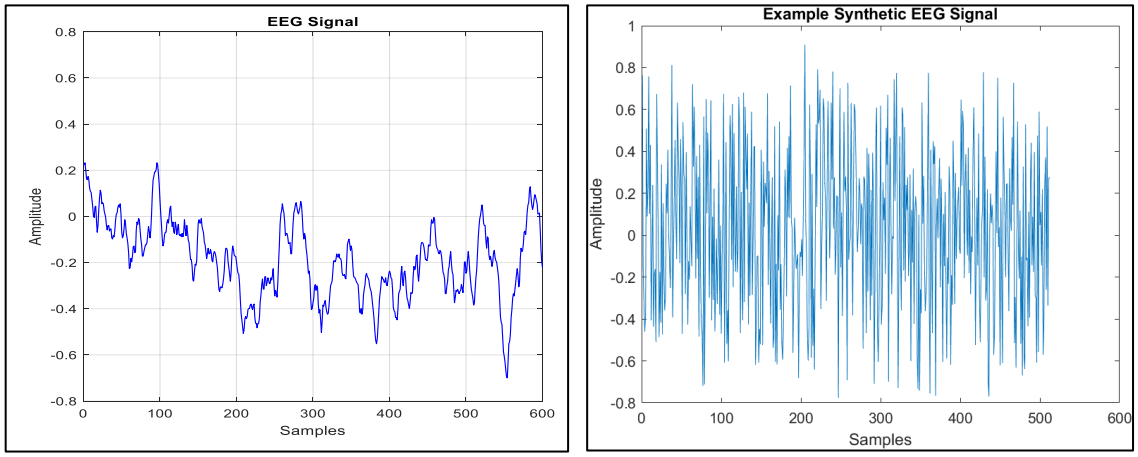
END

The generator and discriminator depths were carefully chosen to balance temporal representation capacity and adversarial training stability for EEG signals. A moderate-depth LSTM-based architecture was found sufficient to capture long-range temporal dependencies while avoiding overfitting and instability associated with deeper recurrent stacks. Due to the improved gradient sensitivity introduced by backpropagation through time in the WGAN-GP-LSTM framework, a lower learning rate was adopted compared to the standard WGAN-GP to ensure stable Wasserstein loss convergence and effective gradient penalty enforcement. Reproducibility was assessed through multiple independent runs with different random initializations, and results are reported as mean ± standard deviation, demonstrating low variance across runs and confirming the robustness and consistency of the proposed method.

5. Results and Discussion

The experimental results of the proposed model of WGAN-GP-LSTM are discussed in this section. The proposed model is evaluated using the CHB-MIT database. The model's

performance is compared with the original dataset as well as the generated dataset. The performance of the model is measured in terms of accuracy, sensitivity, specificity, F1 score, and FID score. From the results, it can be observed that the proposed model can accurately distinguish between EEG signals with and without seizures.



a) Original EEG Signal from the Database b) Synthetic Signal Generated Using WGAN-GP

Figure 4. Original and the Synthetic Signal Output

A part of the original EEG signal taken from the CHB-MIT database is shown in Fig. 4 a). The waveform shows low-frequency oscillations and the noise level typically seen in physiological EEG data. The signal has a smooth temporal structure and clear amplitude variations. It is also relatively free of high-frequency components. This suggests that the cognitive activity profile is realistic and affects biological variables. The artificial EEG signal originated from the WGAN-GP-LSTM, as shown in Fig. 4 b). The synthetic waveform is a more complex signal compared to the original with more high-frequency variation and greater oscillation density. The original EEG signal has an amplitude range of -0.8 to 0.8, while the synthetic signal generated using WGAN-GP ranges from -0.8 to 0.9, indicating proper normalization and scaling. The WGAN-GP-LSTM has captured relevant statistical characteristics of EEG dynamics, but it introduces a bit more randomness than what is typically produced by models that are not trained to capture the fine-grained physiological features. The training parameters of the WGAN-GP and WGAN-GP-LSTM models are shown in Table 3. The features have been segregated in a ratio of 80:20 for training and testing.

Table 3. Training Parameters of WGAN-GP and WGAN-GP-LSTM

Parameters	WGAN-GP - LSTM	WGAN-GP
Learning Rate	0.0003	0.001
Epochs	20	20
Batch Size	128	128
Hidden layers	100	-
Iteration	22K	77K

5.1 Evaluation Metrics

The proposed method is assessed based on accuracy, sensitivity, specificity, F1 score, and FID metrics to ensure robustness and reliability. These metrics offer a holistic assessment of the proposed method's efficiency in terms of reliability and output quality.

5.1.1 Accuracy

Accuracy (Eqn. 6) measures the overall correctness of the model — i.e., how often the classifier is correct across all predictions.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{6}$$

5.1.2 Sensitivity

Sensitivity (Eqn. 7) indicates how effectively the model identifies positive cases (e.g., seizure segments).

$$Sensitivity = \frac{TP}{TP+FN} \tag{7}$$

5.1.3 Specificity

Specificity (Eqn. 8) measures how well the model identifies negative cases (e.g., non-seizure segments).

$$Specificity = \frac{TN}{TN+FP} \tag{8}$$

5.1.4 F1 score

F1 Score (Eqn. 9) is the mean of precision (Eqn. 10) and recall (Eqn. 11). It balances the two, especially useful in imbalanced datasets.

$$F1\ Score = 2 * \frac{Precision \times Recall}{Precision+Recall} \tag{9}$$

$$Precision = \frac{TP}{TP+FP} \tag{10}$$

$$Recall = \frac{TP}{TP+FN} \tag{11}$$

where, True Positives (TP) areseizures that the model accurately detected. True Negatives (TN) are events that are not seizures and are appropriately categorized as such. False Positives (FP)are events that are not seizures yet are mistakenly diagnosed as seizures. False Negatives (FN) are seizures that are inadvertently labeled as non-seizures.

5.1.5 Fréchet Inception Distance

The FID is a widely recognized statistic used to assess the quality and diversity of data produced by generative models, particularly GANs. Higher fidelity and realism in the synthetic samples result in a lower FID score, indicating that the produced samples closely resemble the genuine data in terms of both statistical variance and perceptual resemblance [7].

FID, often using a pretrained Inception network, is computed using the mean and covariance statistics of feature representations taken from fake and real datasets. The metric calculates the Fréchet distance between two multivariate Gaussian distributions, with one derived from real data and the other based on synthetic data. Eqn. 12 shows how the FID score

is computed theoretically, with m and C standing for the mean and covariance matrix of the synthetic data distribution and m_w and C_w for the actual data. T_r, symbol for the trace operation in linear algebra. $\|m - m_w\|_2^2$ is the squared Euclidean distance [10]. The squared Fréchet distance, d^2 is equivalent to the FID score.

$$d^2\big((m, C), (m_w, C_w)\big) = \|m - m_w\|_2^2 + T_r(C + C_w - 2(CC_w)^{\frac{1}{2}}) \qquad (12)$$

The FID score of the WP-GAN and WP-GAN-LSTM is randomly calculated on the generated samples.

Table 4. Experimental Results of the Proposed Work

Evaluation Metrics	WGAN-GP		WGAN-GP [6]		WGAN – GP - LSTM	
	CHB-MIT Dataset	CHB-MIT Dataset + Synthetic Data	CHB-MIT Dataset	CHB-MIT Dataset + Synthetic Data	CHB-MIT Dataset	CHB-MIT Dataset + Synthetic Data
Accuracy	0.86	0.95	0.86	0.88	0.89	0.96
Sensitivity	0.77	0.82	0.79	0.83	0.85	0.87
Specificity	0.66	0.70	0.67	0.72	0.79	0.82
F1 Score	0.80	0.91	0.83	0.86	0.87	0.92
FID	3.07		3.10		1.42	

Table 4 shows the comparative evaluation of the three models based on GAN: WGAN-GP, WGAN-GP with the integration of LSTM, and the baseline model proposed in [6] for seizure detection from EEG signals on the "CHB-MIT database," considering the impact of synthetic EEG signals. The models have been evaluated based on various parameters such as accuracy, sensitivity, specificity, F1-score, and FID.

In terms of accuracy, the results indicate that there is a significant when synthetic data is added to the model. In the proposed model using the WGAN-GP approach, the accuracy improves from 0.86 with real data to 0.95 when synthetic data is added to the model. In the proposed model using the WGAN-GP-LSTM approach, accuracy improves from 0.89 to 0.96, making it the most accurate model among the proposed models. In the existing models [6] and [19], accuracy improvesfrom 0.86 to 0.88 and 0.86 to 0.91, respectively. This indicates that synthetic data slightly contributes to the improvement in accuracy and diagnostic abilities. This implies that by using LSTM in the model, it can perform better in the diagnosis of seizures using EEG signals.

The sensitivity of the WGAN-GP-LSTM model was found to reach a maximum of 0.87 compared to the reference models. When comparing synthetic data and original data, sensitivity increases from 0.77 to 0.82 in the WGAN-GP model, and from 0.85 to 0.87 in the WGAN-GP-LSTM model. Compared to the reference model, sensitivity ranges from 0.79 to 0.83, demonstrating the performance of the model.

The next metric is specificity, where the value increases from 0.66 to 0.70 in the WGAN-GP model and from 0.79 to 0.82 in the WGAN-GP-LSTM model for original and synthetic datasets. When compared with the reference model, specificity ranges from 0.67 to 0.72. This further proves that the model implementing synthetic data with LSTM is better for diagnosis. Likewise, in terms of the F1 score, it ranges from 0.80 to 0.91 for the WGAN-GP model and from 0.87 to 0.92 for the WGAN-GP-LSTM model. The reference model shows slight variations from 0.83 to 0.86.

Ultimately, the FID utilized to evaluate the quality of the produced synthetic data further reinforces the success of the data augmentation approach. A lower FID signifies that the synthetic data is more realistic. The WGAN-GP shows a notable decrease from 3.07 to 1.42, highlighting the enhanced quality of synthetic signals when employing WGAN-GP with appropriate tuning. The findings suggest that both the application of synthetic EEG data and the architectural improvements achieved by integrating GAN with LSTM increase the performance of seizure detection models, particularly regarding accuracy, generalization, and reliability.

The experiments were repeated five independent times. The random seeds implemented in the study involved variations in weight initialization, mini-batch sampling, and learning rate. The observed variability in the results is tabulated in Table 5. The mean and standard deviation values of the evaluation metrics suggest that the proposed method is better for the detection of seizures.

Table 5. Classification Performance Per Class

Metric	Mean ± Std
Accuracy	0.96 ± 0.012
Sensitivity	0.87 ± 0.015
F1-Score	0.92 ± 0.010
FID	1.42 ± 0.08

The suggested frameworks' resilience to adversarial training stochasticity and random initialization is confirmed by low variation across runs. Mode collapse is monitored using Wasserstein loss stability, FID trend consistency across epochs, and visual inspection of the generated EEG variability. It is confirmed that mode collapse did not occur during training due to the lack of loss saturation, stable FID values, and a variety of synthetic EEG waveforms.

Since FID efficiently measures the distributional similarity between real and synthetic data, it has been frequently used for evaluating generative models in time-series domains, despite having been first introduced for images.

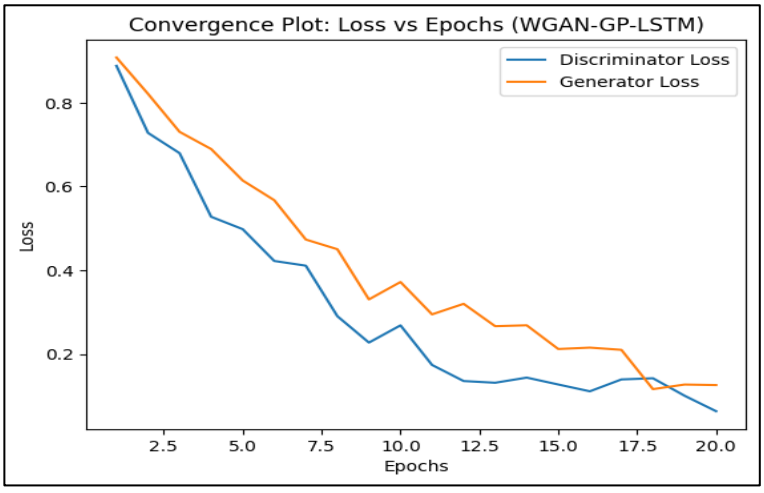


Figure 5. Convergence Plot

A stable convergence of Wasserstein loss is observed in Figure 5. There is no sudden loss saturation or generation dominance detected. Both the losses of the generator and discriminator progressively decline throughout the epochs. The Wasserstein distance with gradient penalty effectively applies the Lipschitz limitation, as seen by the absence of sharp

spikes or loss saturation. The graph further confirms that there was no mode collapse during training. The FID consistency across epochs further validates the absence of mode collapse and the high quality of varied synthetic EEG signals.

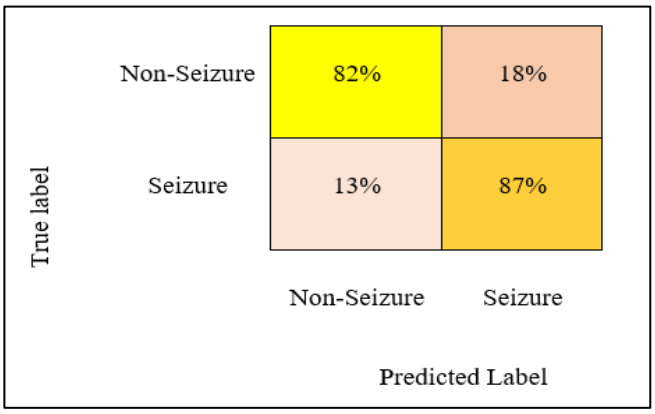


Figure 6. Normalized Confusion Matrix of WGAN-GP-LSTM

The results of the proposed model, as shown by the confusion matrix (Figure 6), demonstrate that the model performs well on both seizure and non-seizure classes by achieving correct classification rates of 82% and 87%, respectively. Although the absolute performance of the model is balanced as shown, there are misclassifications of both seizure and non-seizure classes, where 18% of the non-seizure examples are incorrectly classified as seizures and 13% of the seizures are misclassified as non-seizures. Thus, it appears that the classifier can achieve a moderate trade-off between sensitivity and specificity and has no significant bias toward either class.

In addition to this, the class-wise performance presented in Table 6, identifies the performance characteristics of the model. The seizure class had a precision of 0.87 suggesting a high credibility for the predicted seizures, which is a critical factor when it comes to preventing false alarms within an automated monitoring system. The recall of the non-seizure class was found to be 0.86 indicating that it has the ability to correctly identify normal EEG patterns. Both classes have relatively similar F1-scores (0.84 and 0.85) which also suggests that the model generalizes well and has no significant bias toward either class.

Table 6. Class-Wise Performance Metrics Derived from the Confusion Matrix

Class	Precision	Recall	F1-Score	Class
Non-Seizure	0.82	0.86	0.84	Non-Seizure
Seizure	0.87	0.83	0.85	Seizure

In EEG-based seizure detection problems, and many other times there is often an instance of class imbalance where the seizure class is generally lower than the number of non-seizure instances. This imbalance can cause classifiers to be overly biased in favor of classifying instances as "no seizure" resulting in poor seizure detection. The recall results presented in this research indicate that the models have addressed or mitigated this effect and that balanced learning has been achieved during this process. The 13% false negative rate for seizures, as provided in the results, is a significant clinical issue, as any missed seizure may impact diagnosis and treatment decisions. However, all of the results indicate that the proposed models provide clinically relevant qualities of detection with a level of robustness to class imbalances during the training process.

6. Conclusion

The comparative analysis of the classification of the EEG signal using the methods WGAN-GP, WGAN-GP-LSTM, and the baseline method over the CHB-MIT database with and without synthetic data demonstrates the impact of synthetic data on the classification process. The classification process using synthetic EEG signals improves the performance of all metrics, and the accuracy, sensitivity, specificity, and F1 score show considerable improvement with the incorporation of synthetic EEG signals, with the accuracy increasing to 0.95 from 0.86 and the F1 score increasing to 0.91 from 0.80 using the WGAN-GP method. The accuracy and F1 score using the WGAN-GP-LSTM method are 0.96 and 0.92, respectively, which demonstrates the effectiveness of the incorporation of the LSTM method. This indicates that incorporating time structure is vital in order to effectively capture the dynamics of EEG signals. The lower FID score for WGAN-GP-LSTM using synthetic data (1.42) also justifies the realism and quality of the generated signals. Furthermore, from the visual representation of the original and synthetic EEG signals, it is evident that the generated signals are more realistic in terms of statistical and signal properties, which can help the model learn the decision boundaries more effectively. The results thus confirm the efficacy of advanced generative models such as WGAN-GP and WGAN-GP-LSTM in enhancing the quality and diversity of seizure detection models using enriched training data sets.

References

[1] Hasan, Nowshad, Md Maskat Sharif, Miskatur Rahman, Md Saiful Islam, Md Tohedur Rahaman Khan, MD Jiabul Hoque, and Anamul Bahar. "Epileptic Seizure Prediction Using a Deep Hybrid CNN-GAN Model on EEG Data." In 2024 IEEE 3rd International Conference on Robotics, Automation, Artificial-Intelligence and Internet-of-Things (RAAICON), IEEE, 2024, 270-275.

[2] Pascual, Damian, Amir Aminifar, David Atienza, Philippe Ryvlin, and Roger Wattenhofer. "Synthetic Epileptic Brain Activities Using Generative Adversarial Networks." arXiv preprint arXiv:1907.10518 (2019).

[3] Yu, Tian, Boyuan Cui, Yaqian Xu, and Xilin Liu. "Refine EEG Spectrogram Synthesized by Generative Adversarial Network for Improving the Prediction of Epileptic Seizures." In 2023 11th International IEEE/EMBS Conference on Neural Engineering (NER), IEEE, 2023, 1-4.

[4] Goodfellow, Ian J., Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. "Generative Adversarial Nets." Advances in neural information processing systems 27 (2014).

[5] AlAmir, Manal, and Manal AlGhamdi. "The Role of Generative Adversarial Network in Medical Image Analysis: An In-Depth Survey." ACM Computing Surveys 55, no. 5 (2022): 1-36.

[6] Abou-Abbas, Lina, Khadidja Henni, Imene Jemal, and Neila Mezghani. "Generative AI with WGAN-GP for Boosting Seizure Detection Accuracy." Frontiers in Artificial Intelligence 7 (2024): 1437315.

[7] P., Indurani & Sundaram, Veni. (2023). Improving Deep Learning for Seizure Detection using GAN with Cramer Distance and a Temporal-Spatial-Frequency Loss Function.

International Journal on Recent and Innovation Trends in Computing and Communication. 11. 424-432. 10.17762/ijritcc.v11i6s.6949.

[8] Öcal, Abdurrahman, and Lale Özbakır. "Supervised Deep Convolutional Generative Adversarial Networks." Neurocomputing 449 (2021): 389-398.

[9] Radford, Alec, Luke Metz, and Soumith Chintala. "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks." arXiv preprint arXiv:1511.06434 (2015).

[10] Heusel, Martin, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler, and Sepp Hochreiter. "Gans Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium." Advances in neural information processing systems 30 (2017).

[11] Maind, Sonali B., and Priyanka Wankar. "Research Paper on Basic of Artificial Neural Network." International Journal on Recent and Innovation Trends in Computing and Communication 2, no. 1 (2014): 96-100.

[12] Goodfellow, Ian. "Nips 2016 Tutorial: Generative Adversarial Networks." arXiv preprint arXiv:1701.00160 (2016).

[13] Kazeminia, Salome, Christoph Baur, Arjan Kuijper, Bram Van Ginneken, Nassir Navab, Shadi Albarqouni, and Anirban Mukhopadhyay. "GANs for Medical Image Analysis." Artificial intelligence in medicine 109 (2020): 101938.

[14] Salimans, Tim, Ian Goodfellow, Wojciech Zaremba, Vicki Cheung, Alec Radford, and Xi Chen. "Improved Techniques for Training Gans." Advances in neural information processing systems 29 (2016).

[15] Gulrajani, Ishaan, Faruk Ahmed, Martin Arjovsky, Vincent Dumoulin, and Aaron C. Courville. "Improved Training of Wasserstein Gans." Advances in neural information processing systems 30 (2017).

[16] Hochreiter, Sepp, and Jürgen Schmidhuber. "Long Short-Term Memory." Neural computation 9, no. 8 (1997): 1735-1780.

[17] Qu, Hao, and Jean Gotman. "A Patient-Specific Algorithm for the Detection of Seizure Onset in Long-Term EEG Monitoring: possible use as a warning device." IEEE transactions on biomedical engineering 44, no. 2 (1997): 115-122.

[18] J. Guttag, “CHB-MIT Scalp EEG Database,” Physionet.org, Jun. 09, 2010. http://www.physionet.org/physiobank/database/chbmit (accessed Dec. 10, 2024).

[19] Abou-Abbas, Lina, Khadidja Henni, Imene Jemal, and Neila Mezghani. "Generative AI with WGAN-GP for Boosting Seizure Detection Accuracy." Frontiers in Artificial Intelligence 7 (2024): 1437315.

[20] Elmawazini, Noura, Megan Boucher-Routhier, Gatete Queen Olea Umulinga, and Jean-Philippe Thivierge. "A Generative Adversarial Network for Data Augmentation of Ictal Waves from Multi-Electrode Brain Activity." In 2025 International Conference on Machine Learning and Autonomous Systems (ICMLAS), IEEE, 2025, 1715-1721.

[21] Geng, David, Ayham Alkhachroum, Manuel A. Melo Bicchi, Jonathan R. Jagid, Iahn Cajigas, and Zhe Sage Chen. "Deep Learning for Robust Detection of Interictal Epileptiform Discharges." Journal of neural engineering 18, no. 5 (2021): 056015.