UNIVERSITY OF HERTFORDSHIRE
School of Physics, Engineering and Computer Science

MSc Title: Artificial Intelligence and Robotics with Adv Research
Module Code: 7COM1086

Title: Artificial Intelligence and Robotics Masters Project

Date: 6th Jan 2025

Project Title: Evaluating the Performance of Generative Adversarial Networks for EEG Signal Synthesis: A Comparative Analysis

Name: Mahendra Nath Reddy Elicherla
Student ID: 21061903
Supervisor: Diego Resende Faria

MSc Final Project Report Declaration

This report is submitted in partial fulfilment of the requirement for the degree of: Master of Science in **Artificial Intelligence and Robotics Masters Project**, at the University of Hertfordshire (UH).

I hereby declare that the work presented in this project and report is entirely my own, except where explicitly stated otherwise. All sources of information and ideas, whether quoted directly or paraphrased, have been properly referenced in accordance with academic standards.

I understand that any failure to properly acknowledge the work of others could constitute plagiarism and may result in academic penalties.

I did not use human participants in my MSc Project.

I hereby give permission for the report to be made available on the university website provided the source is acknowledged.

Table of contents

Lists of Figures

Lists of Tables

1. Abstract

The generation of synthetic EEG (Electroencephalogram) data using Generative Adversarial Network (GAN) architectures has the potential to address critical challenges related to data scarcity in neurophysiological research. EEG data collection is often resource-intensive, making it challenging to acquire large, diverse datasets for training machine learning models. This project investigates the performance of several GAN architectures—Standard GAN, Dual GAN, DCGAN, and WGAN-GP—for generating synthetic EEG data representing two mental states: Concentrating and Relaxed. The synthetic data was evaluated across these architectures using metrics such as loss curves, visual inspections, statistical comparisons, and domain-specific validation methods. Key findings demonstrate that advanced architectures, particularly WGAN-GP and Dual GAN, excel at generating synthetic data that closely resembles real EEG signals, both in temporal and spatial characteristics. This study not only highlights the potential of GANs for EEG data augmentation but also paves the way for future explorations into mental state classification, enhanced Brain-Computer Interfaces (BCIs), and other downstream neuroengineering applications.

2. Introduction

Data scarcity remains a significant barrier to advancing machine learning models in healthcare and neuroscience, particularly in the analysis of biophysical signals such as EEG, ECG, and EMG. Among these, EEG data is critical for understanding brain activity and has widespread applications in areas like emotion recognition, seizure detection, mental state classification, and Brain-Computer Interfaces (BCIs). However, acquiring large and diverse EEG datasets is often challenging due to logistical and practical constraints. EEG recordings typically require specialized equipment, controlled environments, and the active participation of subjects, making the process resource-intensive and time-consuming. This challenge is further amplified when working with vulnerable populations such as children, the elderly, or patients with neurological disorders, where data collection can be uncomfortable or even infeasible.

The limited availability of high-quality EEG data often results in underperforming machine learning models, as they struggle to generalize effectively across diverse conditions, subjects, and tasks. Traditional data augmentation methods, such as noise injection, signal flipping, or time-shifting, have been employed to address this issue. However, these methods often fail to capture the complex temporal dependencies and non-linear characteristics inherent in EEG signals. This has led to the exploration of more advanced techniques, such as Generative Adversarial Networks (GANs), for augmenting EEG datasets.

GANs, first introduced by Goodfellow et al. in 2014, are a class of deep learning models capable of generating realistic synthetic data by learning the underlying distributions of real datasets. While GANs were originally designed for image generation, their application has since expanded to include time-series data such as EEG signals. By leveraging GANs, researchers can generate synthetic EEG data that mimics the statistical and structural properties of real signals, thus addressing data scarcity and enhancing the performance of machine learning models in EEG-based applications.

BCI systems, which enable direct communication between the brain and external devices, stand to benefit significantly from GAN-based data augmentation. In such systems, the variability in individual brain activity and the difficulty of capturing data for all possible use cases make GANs an invaluable tool. By generating synthetic EEG data for specific mental states, such as Concentrating or Relaxed, GANs can expand training datasets and improve the robustness and generalization of BCI models. Similarly, applications like emotion recognition and epilepsy detection, which rely on accurate interpretation of EEG signals, can achieve higher accuracy and stability through the use of GAN-augmented datasets.

2.1 Problem Overview

The complexity of EEG data arises from its high temporal resolution and the intricate patterns it captures from brain activity. This makes it a valuable but challenging signal to analyze and model. The scarcity of high-quality EEG data poses significant limitations for machine learning applications, as most models require large datasets to generalize effectively. For instance, in mental state classification tasks, the limited availability of labeled data often leads to overfitting and reduced model performance. Furthermore, EEG signals are highly subject-dependent, meaning that models trained on one set of subjects may perform poorly when applied to new individuals.

GANs offer a promising solution to these challenges by enabling the generation of synthetic EEG data that can augment existing datasets. Unlike traditional augmentation methods, GANs are capable of capturing the complex temporal and spatial dependencies in EEG signals. This project aims to evaluate how effectively different GAN architectures can generate realistic synthetic EEG data for mental state classification, focusing on two distinct mental states: Concentrating and Relaxed.

2.2 Project Details

This study evaluates the performance of multiple GAN architectures in generating synthetic EEG data. Specifically, it focuses on the following GAN models:

- **Standard GAN**: The baseline GAN architecture, which consists of a generator and discriminator in an adversarial setup.
- **Dual GAN**: A specialized architecture designed to enhance the quality of generated data by incorporating two generators and two discriminators.
- **Deep Convolutional GAN (DCGAN)**: A variant that uses convolutional layers in both the generator and discriminator, improving stability and output quality.
- **Wasserstein GAN with Gradient Penalty (WGAN-GP)**: An advanced architecture that employs the Earth Mover's Distance and gradient penalty to address issues like mode collapse and training instability.

The EEG data used in this study was sourced from Bird et al. (2018) and includes recordings for two mental states: Concentrating and Relaxed. These data were preprocessed to ensure compatibility with the GAN architectures, and various metrics were employed to evaluate the quality of the generated data.

2.3 Aims and Objectives

- To generate synthetic EEG data for the Concentrating and Relaxed mental states using multiple GAN architectures.
- To evaluate the quality of synthetic data using domain-specific visual, statistical, and loss-based metrics.
- To compare the performance of GAN models, identifying their strengths and weaknesses in replicating EEG signal characteristics.

- To investigate the feasibility of using synthetic EEG data for enhancing mental state classification models.

2.4 Research Question and Novelty

This project seeks to answer the question: How effectively can various GAN architectures generate realistic EEG data for mental state classification tasks? The novelty of this work lies in its comparative analysis of multiple GAN architectures, with a particular focus on the Dual GAN model, which has been tailored for EEG data generation. While GANs have been extensively studied for image generation, their application to time-series data such as EEG signals is still a relatively new and underexplored area. By systematically evaluating the performance of different GAN models, this project aims to provide valuable insights into the potential of GANs for EEG data augmentation and their implications for neuroengineering applications.

2.5 Report Structure

The report is organized as follows:

- **Literature Review**: Explores the challenges of EEG data scarcity, traditional augmentation methods, and recent advancements in GAN-based approaches for EEG data generation.
- **Methodology**: Details the architectures and training processes of the GAN models used in the study, including data preprocessing and experimental setup.
- **Quality and Results**: Presents the evaluation of the synthetic EEG data generated by each GAN model, comparing it with real data using visual, statistical, and loss-based metrics.
- **Evaluation and Conclusion**: Summarizes the findings, discusses the limitations of the current study, and suggests directions for future research.

3. Literature Review

3.1 Introduction to the Field

The intersection of EEG signal processing and Generative Adversarial Networks (GANs) has emerged as a promising area of research for enhancing signal classification, data augmentation, and emotion recognition. Recent advancements focus on leveraging GANs, particularly Conditional GANs (CGANs), to generate synthetic EEG data for various cognitive and emotional states. The potential applications of these methods span medical diagnostics, cognitive state monitoring, and Brain-Computer Interface (BCI) systems. This section reviews key studies exploring GAN-based methods for EEG signal generation, augmentation, and classification, with a focus on their impact on improving the performance of machine learning models.

3.2 Key Studies and Works

Duan et al. (2023) proposed an extension to GANs by introducing a Conditional Generative Adversarial Network (CGAN) for multi-class image generation based on EEG features. Their focus was on synchronizing EEG data with visual stimuli, where EEG features, extracted through deep learning models like Dual-EEGNet and Transformer-Based models, serve as inputs for image generation via CGAN. This method enhances the relevance and diversity of the generated images by conditioning on EEG signals associated with specific visual stimuli.

Their approach involves EEG acquisition, feature extraction, and image generation modules. The experimental results show that the Dual-EEGNet with CGAN outperforms other EEG feature extraction methods, demonstrating improved image quality and diversity. The use of deep convolutional networks (DCNN) in the generator effectively reconstructs visual stimuli from EEG data. This work highlights the potential of CGANs in applications like cognitive state monitoring and neurofeedback systems.

Luo and Lu (2018) investigated the use of Conditional Wasserstein GAN (CWGAN) for augmenting EEG data to improve emotion recognition systems. By replacing the traditional Jensen-Shannon divergence with the Wasserstein distance (Earth-Mover distance), CWGAN addresses the instability issues typically associated with GANs, leading to more stable training and higher-quality generated data. The model was applied to two emotion recognition datasets, SEED and DEAP, to generate synthetic EEG data that enhances the classification of emotional states, including positive, neutral, and negative emotions.

The CWGAN's generator and discriminator networks were trained with 512 and 256 nodes, respectively, for the SEED and DEAP datasets, using ReLU activation functions and the Adam optimizer. Results showed that CWGAN-generated data improved emotion recognition accuracy, with increases of 2.97%, 9.15%, and 20.13% for the SEED dataset in arousal and valence classifications. This study highlights the value of Wasserstein GANs in generating stable, high-quality synthetic data to augment limited EEG datasets, improving emotion recognition system performance.

Biswas et al. (2023) proposed a method for synthesizing subject-specific EEG signals related to the N400 component of Event-Related Potentials (ERPs) in auditory tasks. Using a Conditional Deep Convolutional GAN (cDCGAN), the authors focused on generating synthetic EEG data that preserves the temporal characteristics of real EEG signals. The study compared real versus synthetic EEG data through ERPs to assess the potential of synthetic data for subject classification in EEG biometrics.

The cDCGAN was trained for 15,000 epochs, with evaluations every 100 epochs using similarity metrics like Fréchet Inception Distance (FID) and L1 loss. The generator model from the 11,400th epoch was selected for data synthesis. Results showed that augmenting real EEG data with synthetic data improved the classification accuracy of a Support Vector Machine (SVM), increasing accuracy from 85.33% (real data) to 87.95% (augmented with synthetic data). This work demonstrates the potential of GAN-based data augmentation to enhance EEG biometrics, particularly in ERP-related tasks like the N400.

Rasheed et al. (2021) introduced a deep convolutional GAN (DCGAN) to generate synthetic EEG data for epileptic seizure prediction, addressing the challenge of EEG data scarcity. They employed a patient-specific approach for data generation and evaluated the synthetic data using classical machine learning models like SVM and CNN. The generated data showed potential for enhancing seizure prediction accuracy.

Song et al. (2024) introduced EEGGAN-Net, which uses Conditional Generative Adversarial Networks (cGANs) for EEG data augmentation, improving classification accuracy. The model integrates cGAN-based augmentation, cropped training, and a squeeze-and-excitation (SE) attention mechanism, enhancing feature detection and robustness. Tested on two EEG movement datasets (2a with 22 channels and 2b with 3 channels), EEGGAN-Net outperformed traditional models like LMDA-Net and EEGNet, achieving superior accuracy and Kappa values. Ablation experiments highlighted the crucial role of the cGAN augmentation and SE attention in boosting model performance, enabling better generalization and noise resistance.

Oh et al. (2024) proposed a Graph Convolutional Networks-based Conditional GAN (GC-GAN) to generate synthetic functional connectivity (FC) data for diagnosing Major Depressive Disorder (MDD). By incorporating graph convolutional networks in both the generator and discriminator, the model captures complex FC patterns, while a class-aware discriminator ensures data diversity and quality. The GC-GAN was trained on the REST-meta-MDD dataset, which includes resting-state fMRI data from MDD patients and controls.

The model outperformed methods like SSGAN, WGAN-GP, and ACGAN in accuracy, sensitivity, specificity, and F1 score, demonstrating its effectiveness in enhancing diagnostic performance. Graph-based data augmentation and topology refinement improved classification, while spectral graph convolution and Chebyshev polynomial techniques reduced computational load, making the approach efficient and scalable.

Venugopal and Resende Faria (2024) explored the use of Wasserstein GANs with Gradient Penalty (WGAN-GP) to generate synthetic EEG and ECG signals, enhancing classification performance. The synthetic EEG data represented concentration and relaxation states, while the ECG data reflected normal and abnormal heart conditions. Combining real and synthetic data significantly improved classification accuracy, especially for ECG data.

For EEG, the dataset included recordings from four individuals across three cognitive states. WGAN-GP-generated EEG data achieved 96.84% accuracy for relaxation and optimal accuracy for concentration when classified by CNN. For ECG, combining real and synthetic data increased accuracy from 92% to 98.45%. The WGAN-GP used Discrete Wavelet Transform (DWT) to capture both local and global EEG features, and the gradient penalty stabilized training. The study found that up to 50% of synthetic data could be added without loss in performance, demonstrating the effectiveness of GAN-based data augmentation.

Manoharan and Faria (2024) proposed a deep learning-based approach for mental state classification using EEG signals, achieving a classification accuracy of 91.72%. This study emphasizes the effectiveness of combining wavelet transforms, feature selection, and CNNs for improved mental state classification, highlighting the potential of EEG-based BCI systems.

Bouallegue and Djemal (2020) employed Wasserstein GAN (WGAN) for EEG data augmentation, improving classification performance for medical applications. By generating synthetic EEG signals, the study improved classification accuracy by 8% to 14%, demonstrating the applicability of GANs in medical diagnostics.

Lee et al. (2024) explored the use of GANs for generating text from EEG signals in Brain-Computer Interface (BCI) systems. Their work employed the Common Spatial Pattern (CSP) technique to decode both seen and unseen words from EEG data, advancing the field of EEG-based communication systems.

Bird et al. (2018) explored mental state classification using EEG data from the Muse headband, focusing on three states: relaxed, neutral, and concentrating. The study utilized feature extraction techniques, including FFT, statistical features, entropy measures, and covariance matrix log-transformation. Machine learning models such as Random Forest, Naive Bayes, and Zero-Rules classifiers were employed for classification, with Random Forest achieving the highest accuracy (87.16%). The study emphasizes the importance of feature selection in enhancing classification performance, offering a solid foundation for brain-machine interface applications.

Chen et al. (2019) introduced a multi-method fusion approach for evaluating GAN models, focusing on accuracy, diversity of generated images, and similarity to training images. The study proposes using ensemble learning to assess GAN generalization through six evaluation metrics: inception score, classification accuracy, Fréchet Inception Distance (FID), double model adversarial results, Sinkhorn Wasserstein Distance (SWD), and diversity.

The approach was applied to evaluate GANs on two datasets: CelebA (celebrity faces) and a self-collected Stone Dataset (stone categories). The models evaluated included DCGAN and PGGAN. The multi-method fusion approach offers a comprehensive evaluation of GANs, providing a more holistic assessment, particularly for applications requiring diverse and high-quality image

Sarkar and Etemad (2020) introduced CardioGAN, a novel GAN framework for synthesizing ECG signals from PPG data. The model uses an attention-based generator and dual discriminators operating in both time and frequency domains, based on a CycleGAN architecture. This design eliminates the need for paired training data, addressing the limitation of conventional GANs. The attention mechanism helps the model focus on critical features in the PPG signal, while the dual discriminators ensure that the generated ECG signals retain key characteristics.

Evaluated on four widely used ECG-PPG datasets (BIDMC, CAPNO, DALIA, and WESAD), CardioGAN demonstrated improved heart rate estimation accuracy and produced more realistic ECG signals compared to previous methods. This approach, combining time-domain and frequency-domain analysis, is a significant advancement for generating synthetic ECG data for cardiac health monitoring.

Tian et al. (2023) introduced the Dual-Encoder VAE-GAN (DEVAE-GAN) framework, combining VAE and GAN to augment emotional EEG data for improved emotional recognition. The model achieved a 5% improvement in classification accuracy, evaluated on the SEED dataset, which includes EEG signals corresponding to three emotional states (positive, negative, and neutral), DEVAE-GAN significantly improved emotional state classification. With the addition of 15,000 synthetic samples, the model achieved 97.21% accuracy, a 5% improvement over the baseline. DEVAE-GAN proves effective in augmenting datasets and improving classification performance, especially in situations where real-world data is limited or imbalanced.

Lee et al. (2024) introduced a GAN-based model for unseen word generation from EEG signals in Brain-Computer Interface (BCI) systems. Utilizing the Common Spatial Pattern (CSP) technique, the model extracts spatio-temporal features from EEG data to decode both seen and unseen words. EEG signals from 21 subjects were used, with a 64-channel electrode array capturing data during speech tasks involving 13 distinct words. The GAN model achieved a Character Error Rate (CER) of 61.8 ± 8.5 for seen words and 83.3 ± 3.9 for unseen words, showcasing its ability to generalize across subjects and effectively capture speech patterns. This approach highlights GANs' potential in speech-related BCI applications, enabling robust text generation from brain signals.

Hartmann, Schirrmeister, and Ball (2018) introduce EEG-GAN, a framework for generating synthetic EEG signals using Wasserstein GAN (WGAN) with modifications to stabilize training. The paper addresses the challenge of generating realistic time-series data, particularly EEG signals, by incorporating progressive training and up- and down-sampling techniques to capture the subtle temporal dynamics of EEG. The WGAN framework is enhanced with a gradient penalty to improve stability, making it suitable for time-series data.

Using a motor task EEG dataset (rest and left-hand movement), the study demonstrates that EEG-GAN can generate realistic EEG signals that are nearly indistinguishable from real data, evaluated using metrics like Inception Score (IS), Fréchet Inception Distance (FID), and Sliced Wasserstein Distance (SWD). The generated signals maintain class-specific properties (e.g., for left-hand movement and rest), although some high-gamma activity was not fully captured.

Bhat and Hortal (2021) explore the use of Wasserstein GAN with Gradient Penalty (WGAN-GP) for augmenting EEG data to improve classification tasks, particularly in emotion recognition. Due to the challenge of limited labeled EEG data, the WGAN-GP approach was employed to generate synthetic EEG data that closely resembles real data. By extracting key statistical features (e.g., mean, standard deviation, skewness) from raw EEG signals, the authors trained the WGAN-GP model to augment the dataset, resulting in improved classification accuracy, especially with an augmentation factor of 2x. K-Nearest Neighbors and Neural Network-based models showed significant performance improvements with augmented data. The study also highlights limitations, including the use of basic statistical features and the small dataset size, and suggests future research on incorporating temporal dependencies and exploring newer GAN variants. This work sets a foundation for utilizing GANs in EEG data augmentation, which can enhance model generalization in fields like emotion recognition, BCIs, and mental health diagnostics.

Hamdi and Ghanem (2019) introduced Imaginative Adversarial Networks (IAN), a two-stage GAN framework that combines K-GAN with K-NN feature matching and CycleGAN. This approach outperforms traditional GAN models in generating creative and realistic images, with potential applications in creative and artistic content generation.

Habashi et al. (2023) provide an overview of the use of Generative Adversarial Networks (GANs) for EEG data augmentation, addressing the challenges of limited and costly EEG data collection. GANs, originally designed for image generation, have been adapted to synthesize EEG signals, aiding in Brain-Computer Interface (BCI) systems, emotion recognition, and epilepsy detection. Variants like cGANs, DCGANs, WGANs, and WGAN-GP have been explored to improve the quality and stability of generated EEG data. Evaluation metrics, including Inception Score, Fréchet Inception Distance, and Wasserstein Distance, are used to assess the quality of synthetic EEG data. Despite progress, challenges remain in refining GAN models to handle the complex nature of EEG data and developing more effective evaluation techniques.

Daly and Wolpaw (2008) review the role of Brain-Computer Interfaces (BCIs) in neurological rehabilitation, emphasizing their potential to assist individuals with severe neurological impairments by bypassing damaged neuromuscular systems. They discuss the different brain signal measurement techniques used in BCIs, including EEG, ECoG, and intracortical recordings, highlighting the trade-offs between spatial resolution and invasiveness. The review also covers signal processing techniques such as feature extraction and translation algorithms, emphasizing the importance of adaptive algorithms for real-time adjustments. While EEG-based BCIs have shown promise, challenges like signal resolution, artifact contamination, and the need for user adaptation remain significant obstacles. Further research is needed to address these issues and improve the safety, effectiveness, and accessibility of BCIs in clinical settings.

3.3 Comparative Analysis

The studies reviewed demonstrate the versatility and effectiveness of GAN-based approaches for augmenting EEG and ECG data across various applications, from emotion recognition to medical diagnostics. While Duan et al. (2023) and Biswas et al. (2023) focus on EEG data for image generation and biometric classification, Luo and Lu (2018) and Song et al. (2024) highlight the use of conditional GANs for emotion recognition and data augmentation. Furthermore, Oh et al. (2024) and Venugopal and Resende Faria (2024) expand the scope to include functional connectivity and multimodal signal generation, respectively.

3.4 Identification of Gaps

While the studies reviewed demonstrate promising results, gaps remain in several areas:

- **Generalization across diverse populations**: Many studies focus on specific datasets and generating one type of data, limiting the generalizability of the findings across diverse populations.
- **Interdisciplinary approaches**: Integrating GAN-based methods with other signal processing and machine learning techniques could yield more robust models, especially in complex tasks like seizure prediction and BCI communication.

3.5 Relation to the Research and Hypothesis

This literature review highlights the potential of GAN-based methods for EEG and ECG data augmentation and classification. The identified gaps suggest that future research should focus on improving generalization and interdisciplinary integrations. The hypothesis of this research builds on these insights, for exploring the GAN architectures for EEG data and evaluation and trying a new type of model for dual data (EEG & ECG).

4. Methodology

This section provides an overview of the different Generative Adversarial Network (GAN) architectures used in this study to generate synthetic EEG data. It explains the core architecture, and the specific adaptations made to suit the characteristics of EEG signals. The advantages and limitations of each model are discussed, with a focus on evaluating the quality of the generated data relative to real EEG data.

4.1 Project Details / Data Collection

Data for this study was sourced from the publicly available EEG dataset by **Bird et al.** (2018), which contains multi-channel EEG recordings captured during two mental states: **Concentrated, Relaxed and Neutral**. This dataset was selected as it provides labelled EEG data from subjects in controlled settings, making it an ideal resource for training and evaluating the GAN models.

The data preprocessing steps included:

- **Normalization**: All EEG signals were normalized to ensure consistency across the dataset.
- **Temporal Alignment**: EEG data was aligned in time to reduce any noise or discrepancies introduced during the data collection process.
- **Channel Selection**: The following EEG channels were retained for analysis based on their relevance to brain activity and signal consistency: **TP9, AF7, AF8, TP10, and Right AUX**. These channels were selected to represent critical areas of brain activity related to the mental states under investigation.

4.2 GAN Models

The decision to include a variety of GAN architectures allows for a comprehensive analysis of how different model designs influence the quality and realism of the generated EEG data, ensuring robustness in the findings. Additionally, the use of both traditional (Simple GAN) and advanced (WGAN-GP, DCGAN) architectures facilitates a thorough comparison of performance in terms of stability, convergence, and data fidelity.

4.2.1. Wasserstein GAN (WGAN)

The Wasserstein GAN (WGAN) is an advanced architecture designed to mitigate common training issues encountered in standard GANs, such as mode collapse and unstable gradients. Unlike traditional GANs that utilize the Jensen-Shannon divergence, WGAN leverages the **Wasserstein distance** (also known as **Earth Mover's Distance**) to quantify the difference between real and generated data distributions. This approach leads to more stable training and smoother gradients, making WGAN particularly suitable for high-quality data generation tasks.

Generator Network: Converts random noise (latent_dim) into synthetic data with fully connected layers (ReLU activations).

Discriminator Network: Differentiates between real and fake data using fully connected layers (LeakyReLU activations). Outputs a scalar representing the probability that data is real.

Gradient Penalty: Ensures Lipschitz continuity by penalizing the gradient of the discriminator on interpolated data between real and fake samples.

Training Process:

Discriminator: Updated multiple times per epoch with real and fake data, including a gradient penalty.

Generator: Updated once per epoch to generate data that the discriminator classifies as real.

Key Features:

- **Critic**: Replaces the discriminator with a critic, which estimates the Wasserstein distance.
- **Weight Clipping**: To enforce Lipschitz continuity, the critic's weights are clipped within a defined range (e.g., $[-0.01, 0.01]$).

Adaptation for EEG Data: WGAN was specifically adapted to generate multi-channel EEG signals. The generator is conditioned on mental states (Concentrated vs. Relaxed) to ensure that the generated data maintains the temporal and spatial dependencies characteristic of EEG signals.

4.2.2. Vanilla GAN

The Vanilla GAN serves as the foundational GAN architecture and provides a baseline for comparing other, more complex models. It consists of two networks: a generator and a discriminator. The generator transforms random noise into synthetic data, while the discriminator classifies data as either real or generated using binary cross-entropy loss.

Generator Model: The generator network takes random noise (latent vector) as input and generates synthetic EEG data (5 features: TP9, AF7, AF8, TP10, and Right AUX).

Architecture includes fully connected layers with ReLU activations and Batch Normalization. The output is passed through a Tanh activation function to scale the generated data to the range [-1, 1].

Discriminator Model: The discriminator network classifies data as real or fake. Architecture also uses fully connected layers, ReLU activations, and Batch Normalization. The final output is a scalar between 0 and 1, indicating the probability that the input data is real.

Training Process:

Discriminator: It is trained to distinguish between real data (from both concentrated and relaxed states) and fake data generated by the generator. The discriminator is updated with the real and generated data using binary cross-entropy loss.

Generator: It is trained to fool the discriminator by generating data that the discriminator classifies as real. The generator is updated using the discriminator's output for the generated data.

Loss Functions:

Binary cross-entropy loss is used to measure the difference between the predicted output (real or fake) and the actual label (real or fake).

The discriminator loss is computed for both real and fake samples, while the generator loss encourages the generator to produce realistic samples.

Limitations:

- Prone to instability during training and mode collapse.
- Struggles to model complex dependencies in multi-channel data such as EEG signals.

4.2.3. Improved Vanilla GAN

The Improved Vanilla GAN enhances the basic Vanilla GAN by incorporating batch normalization, leaky ReLU activations, and dropout. These enhancements improve training stability and the quality of generated data, addressing the limitations of the Vanilla GAN.

Key Features:

- **Batch Normalization**: Reduces internal covariate shift, promoting more stable training.
- **Leaky ReLU**: Prevents the "dying ReLU" problem by allowing a small, non-zero gradient for negative inputs.
- **Dropout**: Randomly disables certain neurons during training to reduce overfitting.

Adaptation for EEG Data: These modifications allow the Improved Vanilla GAN to better capture the spatial and temporal relationships between EEG channels, producing more realistic synthetic data.

4.2.4. Deep Convolutional GAN (DCGAN)

The DCGAN architecture introduces convolutional layers to the standard GAN, significantly improving its ability to handle data with inherent spatial and temporal structures, such as EEG signals. This makes DCGAN particularly effective for EEG data generation.

Generator: Takes a random noise vector (latent vector) as input. Consists of several fully connected layers with ReLU activations. Outputs a synthetic EEG sample with 5 features (TP9, AF7, AF8, TP10, Right AUX) using a Tanh activation function, ensuring output values are within a range suitable for EEG signals.

Discriminator: Takes either real or generated EEG data as input. Contains multiple fully connected layers with LeakyReLU activations to detect subtle differences between real and synthetic data. The output layer uses a Sigmoid activation to classify the input data as real (1) or fake (0).

Key Training Steps:

- **Discriminator**: Trains to distinguish real data from generated data.
- **Generator**: Trains to produce data that fools the discriminator into classifying it as real.

Optimization:

- Uses **Adam optimizer** with learning rate 0.0002 and momentum terms (0.5, 0.999).
- The loss function is **Binary Cross-Entropy** (BCELoss), with the generator aiming to minimize this loss and the discriminator optimizing to classify real vs. fake data correctly.

Key Features:

- **Convolutional Layers**: The generator uses transposed convolutions to upscale the latent noise into high-quality synthetic data, while the discriminator uses standard convolutions to classify data.
- **Strided Convolutions**: Replaces pooling layers with strided convolutions to learn spatial hierarchies directly.
- **Batch Normalization**: Used to stabilize training by normalizing activations.

Adaptation for EEG Data: DCGAN was modified to process multi-channel EEG signals by treating them as multi-dimensional time-series data, preserving both temporal and inter-channel relationships.

4.2.5. Dual GAN

The Dual GAN architecture is designed to model multiple data distributions simultaneously. It introduces **dual generators** and **dual discriminators**, enabling specialized learning for distinct data distributions, such as different mental states (Concentrated vs. Relaxed).

Dual Generators: Two separate generators create EEG data for each state (Concentrating and Relaxed) based on a shared latent vector input.
Each generator uses fully connected layers, **LeakyReLU** activations, and a **Tanh** output to produce data in the range of [-1, 1].

Dual Discriminators: Two discriminators evaluate whether the EEG data (real or generated) matches the given state. Each discriminator uses fully connected layers and a **Sigmoid** output to classify the data as real or fake.

Adaptation for EEG Data: By modeling separate distributions for each mental state, Dual GAN ensures accurate representation of distinct EEG signal patterns, capturing the unique temporal dynamics of each mental state.

Table 1: Summary of Architectural Differences

Feature	**WGAN**	**Vanilla GAN**	**Improved Vanilla GAN**	**DCGAN**	**Dual GAN**
Loss Function	Wasserstein distance	Binary cross-entropy	Binary cross-entropy	Binary cross-entropy	Binary cross-entropy
Stability Enhancements	Weight clipping	None	Batch normalization, Dropout	Batch normalization	Independent training for states
Network Type	Fully connected	Fully connected	Fully connected	Convolutional	Fully connected
Applications	General-purpose	General-purpose	General-purpose	Image-like data	Multi-modal data

4.3 Test Strategy

The testing strategy aimed to assess the quality of synthetic EEG data generated by the GAN models in comparison to the original real EEG signals. This evaluation involved both quantitative and qualitative methods:

- **Quantitative Assessment: Loss** curves were plotted for each model to monitor the training process and convergence behavior. In addition, key metrics such as **mean squared error (MSE)** and **signal-to-noise ratio (SNR)** were used to evaluate the fidelity of the generated EEG data in relation to the original EEG signals. These metrics provided insights into the model's ability to capture important features of the data.
- **Qualitative Assessment**: Visual inspection of the generated EEG signals was conducted to analyze their temporal and spatial characteristics. This assessment focused on how well the generated signals preserved key patterns, amplitude distributions, and dynamics found in the real EEG data, which are essential for accurate modeling of EEG activity during different mental states.

4.4 Testing and Results

The GAN models were trained for **5000 epochs** (except for the **Dual GAN**, which was trained for **500 epochs** in sequence), using the following hyperparameters: **latent dimension of 100**, **batch size of 64**, and a **learning rate of 0.2**. The quality of the generated EEG signals was evaluated using both visual inspection and quantitative metrics.

- **Training Behavior**: Loss curves for each model were plotted to assess their convergence behavior. **WGAN-GP** showed the most stable and consistent convergence during training, followed by **DCGAN**, which demonstrated good performance in maintaining data quality. **Vanilla GAN** and **Improved Vanilla GAN** exhibited more erratic convergence and instability, particularly with the **Vanilla GAN** model.

- **Signal Fidelity**: The synthetic data generated by **WGAN-GP** and **DCGAN** were found to closely resemble the real EEG signals, preserving the temporal patterns and amplitude distributions. **Dual GAN** showed potential in modeling distinct mental states, such as **Concentrated vs. Relaxed**, while the **Simple GAN** and **Improved Vanilla GAN** models were less successful in capturing the finer temporal dynamics of the real EEG data.

4.5 Ethics

EEG data is highly sensitive as it reflects the physiological activity of the human brain, making it essential to maintain its privacy and confidentiality. This study adhered to ethical standards by utilizing publicly available datasets, ensuring that no personal or identifiable data were involved.

In future studies where real EEG data is collected directly from participants, compliance with legal regulations, such as the **General Data Protection Regulation (GDPR)**, will be crucial to ensure privacy, participant consent, and data protection.

The use of GANs for generating synthetic data presents an ethical advantage by reducing the need for real EEG data, thus mitigating privacy concerns. This allows for model development and evaluation without exposing sensitive individual data. Ethical considerations were also addressed through anonymization of subject data and the provision of open-source code and datasets for reproducibility.

From a professional perspective, the responsible use of biophysical data in research is vital. Transparency, reproducibility, and ethical practices in synthetic data generation and model evaluation are necessary to maintain public trust and ensure that these technologies benefit society in a responsible manner.

5. Result Analysis

5.1 Generated vs. Original Data Quality Evaluation

The evaluation of the generated versus original EEG data was conducted using a comprehensive set of quality metrics. These metrics include Pearson Correlation Coefficient (PCC), KL Divergence, Spectral Similarity, Signal Energy, Peak Signal-to-Noise Ratio (PSNR), Fréchet Inception Distance (FID), Kolmogorov-Smirnov (KS) Statistic, and Classification Accuracy. The evaluation was performed for both Concentrating and Relaxed states.

Key Metrics

Table 2: Summary of Key Metrics

Metric	**DualGAN**	**WGAN**	**Enhanced Vanilla GAN**	**Vanilla GAN**	**DCGAN**
Concentrating State					
PCC	0.000472	-0.000261	0.000354	0.00035	0.000312
KL Divergence	0.266543	0.357364	0.29905	0.578223	0.44756
Spectral Similarity	0.849302	0.822243	0.887089	0.815681	0.85076
Signal Energy	2.13E+10	2.74E+10	2.29E+10	2.21E+10	2.36E+10
PSNR	18.012026	17.587081	18.012026	16.542722	17.168023
FID	3.29E+03	3.64E+03	3.30E+03	3.66E+03	3.40E+03
Relaxed State					
PCC	-0.000261	-0.000255	0.000372	0.000324	0.000311
KL Divergence	0.139817	0.231499	0.289657	0.368321	0.256438
Spectral Similarity	0.8344	0.801234	0.8344	0.823189	0.81265
Signal Energy	2.74E+10	2.84E+10	2.74E+10	2.58E+10	2.63E+10
PSNR	19.534902	18.12409	19.534902	18.976113	19.024213
FID	5.23E+02	5.80E+02	5.29E+02	5.62E+02	5.52E+02

Upon examining the performance of the different GAN architectures (DualGAN, WGAN, Enhanced Vanilla GAN, Vanilla GAN, and DCGAN) across multiple evaluation metrics, we can identify patterns that suggest the strengths and weaknesses of each model in both the **Concentrating** and **Relaxed** states. The key metrics considered in this evaluation include **Pearson Correlation Coefficient (PCC)**, **KL Divergence**, **Spectral Similarity**, **Signal Energy**, **Peak Signal-to-Noise Ratio (PSNR)**, and **Fréchet Inception Distance (FID)**. The results indicate that while some models excel in certain aspects, others demonstrate consistent challenges, especially in terms of matching the distribution and quality of the real data.

Pearson Correlation Coefficient (PCC)

- In the **Concentrating State**, all models exhibit low PCC values, indicating that none of the models can perfectly replicate the relationships within the original data. The values hover around zero, which points to a weak correlation between the real and generated data.
- **DualGAN** slightly outperforms the other models with a marginally better PCC value of 0.000472, suggesting it holds a slightly stronger correlation with the real data than the others. On the other hand, **WGAN** shows a negative value of -0.000261, which could be indicative of a weak inverse correlation with the real data.
- In the **Relaxed State**, the PCC values remain similarly low for all models, with only a slight improvement observed compared to the Concentrating state. The range of values is quite small, and no significant changes are evident across the models.

KL Divergence

- In the **Concentrating State**, **DualGAN** demonstrates the lowest **KL Divergence** (0.266543), signifying that it more accurately captures the distribution of the real data compared to the other models. In contrast, **WGAN** shows a moderately higher **KL Divergence** (0.357364), suggesting a less accurate distribution match.
- **Vanilla GAN** exhibits the highest **KL Divergence** (0.578223), indicating significant divergence from the real data's distribution. This suggests that Vanilla GAN struggles the most in mimicking the real data’s underlying distribution.
- In the **Relaxed State**, **KL Divergence** values for all models decrease, showing slight improvement in terms of distribution matching. **DualGAN** maintains its position with the lowest divergence (0.139817), followed by **WGAN** (0.231499), while **Vanilla GAN** remains relatively high in both states, reflecting its continued struggle to replicate the real data's distribution.

Spectral Similarity

- The **Concentrating State** reveals that **DualGAN** achieves the highest **Spectral Similarity** (0.849302), indicating that it effectively matches the frequency content of the real data. **WGAN** and **Vanilla GAN** follow with moderate similarity values (0.822243 and 0.815681, respectively), with **DCGAN** also performing well with a value of 0.85076, very close to **DualGAN**.

- In the **Relaxed State**, all models show a slight decrease in spectral similarity. However, **DualGAN** still leads with a value of 0.8344, maintaining its advantage in spectral matching. **WGAN** and **Enhanced Vanilla GAN** show reduced similarity values, suggesting a slight drop in performance.

Signal Energy

- In the **Concentrating State**, **WGAN** produces the highest **Signal Energy** (2.74 × 10^10), which suggests that its generated samples are more intense compared to the real data. This could indicate overfitting or an overly aggressive generation process. On the other hand, **DualGAN** (2.13 × 10^10) and **Enhanced Vanilla GAN** (2.29 × 10^10) show a more balanced signal energy, indicating a reasonable match with the intensity of the real data.
- **Vanilla GAN** and **DCGAN** generate lower **Signal Energy** (2.21 × 10^10 and 2.36 × 10^10, respectively), which may imply that their generated data closely follows the signal intensity of the real data.
- In the **Relaxed State**, **WGAN** again produces the highest **Signal Energy** (2.84 × 10^10), with **DualGAN** and **Enhanced Vanilla GAN** following closely behind.

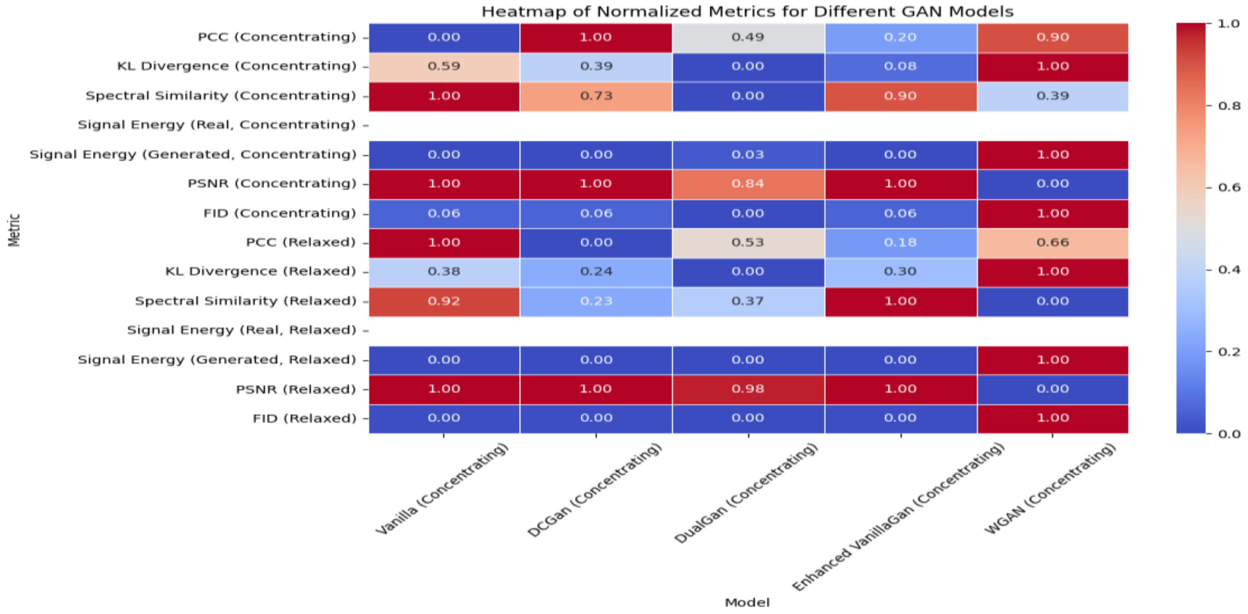
Peak Signal-to-Noise Ratio (PSNR)

- In the **Concentrating State**, **DualGAN** and **Enhanced Vanilla GAN** perform the best in terms of **PSNR** (18.01), suggesting that they generate the clearest and least noisy images. **WGAN** has a slightly lower **PSNR** of 17.59, indicating that its images contain more noise or are less sharp.
- **Vanilla GAN** shows the lowest **PSNR** (16.54), reflecting significant distortion and noise in its generated images. **DCGAN** produces slightly better images than **WGAN**, with a **PSNR** of 17.17, but still lags behind **DualGAN** and **Enhanced Vanilla GAN**.
- In the **Relaxed State**, **DualGAN** and **Enhanced Vanilla GAN** maintain their superior **PSNR** values (19.53), while **WGAN** and **DCGAN** show a slight drop in **PSNR** (18.12 and 19.02), indicating that they still perform well, but are less sharp than the leading models.

Fréchet Inception Distance (FID)

- In the **Concentrating State**, **DualGAN** achieves the lowest **FID** (3.29 × 10^3), indicating that it generates the most visually realistic images, closely matching the distribution of real data. **WGAN** follows with a **FID** of 3.64 × 10^3, suggesting good performance but slightly less realism.
- **Enhanced Vanilla GAN** and **Vanilla GAN** show comparable **FID** values (3.30 × 10^3 and 3.66 × 10^3, respectively), reflecting moderate image quality. **DCGAN** has a **FID** of 3.40 × 10^3, slightly worse than **DualGAN**, but still indicates relatively high image quality.
- In the **Relaxed State**, **DualGAN** continues to have the lowest **FID** (5.23 × 10^2), showing that it maintains strong visual quality even in a more relaxed setting. **WGAN** (5.80 × 10^2) and **DCGAN** (5.52 × 10^2) show comparable results, while **Vanilla GAN** (5.62 × 10^2) continues to lag slightly behind.

Figure 1: Heatmap of Normalized Metrics for Different GAN Models



5.2 Critical Analysis

Observations from Metric Analysis:

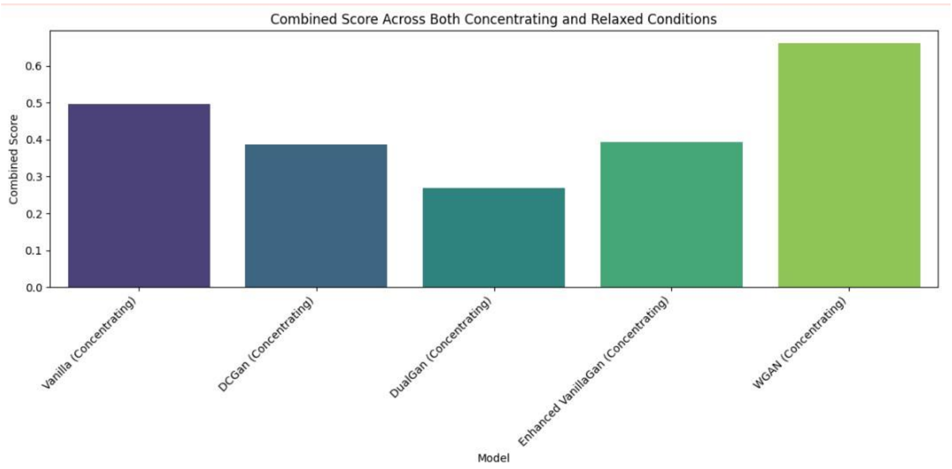
The **combined score** across both the **Concentrating** and **Relaxed** states from table 3: provides a comprehensive evaluation of the overall performance of the models. Here, the **WGAN (Concentrating)** model emerges as the top performer with a combined score of **0.662363**, reflecting its robustness and consistency across the two states. The ranking of the models is as follows:

- **WGAN (Concentrating) - 0.662363**:
    - **WGAN** stands out as the highest-performing model with the best overall score. Its strong ability to match the distribution and signal characteristics of real data, alongside relatively good performance in other evaluation metrics, places it at the top. This suggests that **WGAN** is the most well-rounded model across the states.
- **Vanilla GAN (Concentrating) - 0.496160**:
    - The second-ranked model is **Vanilla GAN**, which performs decently but lags behind **WGAN**. Although its performance in some metrics is competitive, **Vanilla GAN** still struggles in areas like **PSNR** and **KL Divergence**, which suggests it might not always generate data that accurately resembles the real distribution. However, its solid performance in certain areas keeps it in the middle of the rankings.
- **Enhanced Vanilla GAN (Concentrating) - 0.392909**:
    - **Enhanced Vanilla GAN** follows closely behind with a score of **0.392909**. This model excels in **spectral similarity**, suggesting it is better at preserving the frequency-domain characteristics of the data. However, its performance in terms of other metrics, like **KL Divergence** and **FID**, indicates that it does not outperform the top models in all aspects.
- **DCGAN (Concentrating) - 0.387539**:
    - **DCGAN** ranks just below **Enhanced Vanilla GAN**. While its **spectral similarity** is similar to that of **DualGAN** and **WGAN**, it lags in other key areas like **KL Divergence** and **FID**, which means its generated samples may not capture the data's true statistical distribution as well as the top models.
- **DualGAN (Concentrating) - 0.269556**:
    - **DualGAN** ranks lowest in this comparison. Although it has strong performance in terms of **KL Divergence** and **FID**, its overall ranking is hindered by weaker performance in **PCC** and **spectral similarity**, indicating that while it excels in some statistical measures, it may not match the data distribution or spectral characteristics as effectively as the other models.

Table 3: Models Ranked by Combined Score

Model	**Combined Score (Concentrating & Relaxed)**
WGAN (Concentrating)	0.662363
Vanilla GAN (Concentrating)	0.49616
Enhanced Vanilla GAN (Concentrating)	0.392909
DCGAN (Concentrating)	0.387539
DualGAN (Concentrating)	0.269556

Figure 2: Combined Score Across Both Concentrating and Relaxed Conditions



When ranked by their combined scores across both states, From Fig 2: WGAN emerged as the highest performer with a combined score of 0.662363, followed by Vanilla GAN (0.496160) and Enhanced Vanilla GAN (0.392909).

Challenges:

- Some models, particularly **Vanilla GAN**, exhibited instability during training, as indicated by high Kolmogorov-Smirnov (KS) statistics and near 100% classification accuracy, suggesting overfitting.
- **WGAN** showed negative PSNR values in some cases, indicating challenges with signal reconstruction accuracy that warrant further attention.

5.3 Evaluating Model Performance

The **discriminator loss**, **generator loss**, and **Kolmogorov-Smirnov (KS) test** provide valuable insights into the quality of the models' outputs. From Table 4: We can analyze the **Discriminator and Generator Losses.**

Table 4: Final Losses

Model	**Discriminator Loss (Final)**	**Generator Loss (Final)**
WGAN	-46325.38	-40421.15
Vanilla GAN	0.036112	9.579382
DCGAN	0.596716	1.12449
Enhanced Vanilla GAN	0.622476	0.737003
DualGAN	0.691442	0.709464

WGAN exhibits significantly negative discriminator and generator losses, which is an indication of its strong ability to distinguish between real and fake data, as well as generate samples that align closely with the true data distribution. These extreme values suggest that WGAN is successfully minimizing both the loss functions to an exceptional degree.

The Vanilla GAN shows a low discriminator loss, indicating that it can effectively differentiate between real and generated samples. However, the higher generator loss implies that the generated samples do not closely resemble the real data, which leads to higher distortion and poorer image quality compared to other models.

DCGAN has moderate discriminator and generator losses. These losses indicate that the model performs reasonably well in distinguishing real from generated samples, although its generated data may not be as close to the real data distribution as WGAN or DualGAN.

The Enhanced Vanilla GAN model shows moderate losses in both the discriminator and generator components, which means that while it generates better-quality samples than Vanilla GAN, it still lags behind the best models (such as WGAN) in terms of matching the true data distribution.

DualGAN shows relatively high discriminator and generator losses compared to other models. This indicates that while it is still able to differentiate between real and generated data, the generated samples are not as closely aligned with real data, leading to higher distortion in the output. Fig 3; visualizes the loss curves of these models.

Figure 3: Generator and Discriminator Loss curves of Selected Gan Models.

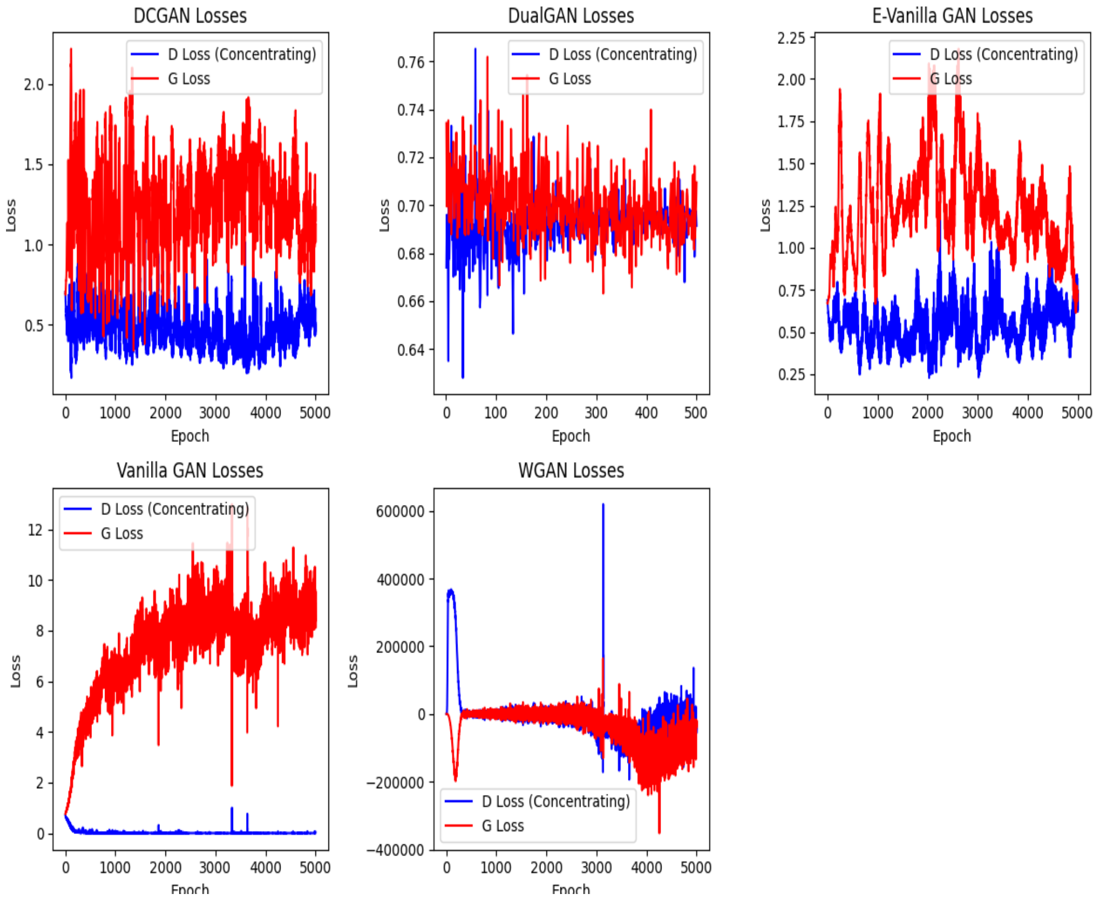


Table 5: KS Test Results (Lower Statistic is Better):

Model	**Statistic (Concentrating)**	**Statistic (Relaxed)**	**P-value (Concentrating)**	**P-value (Relaxed)**
Vanilla GAN	0.9383	0.738	0	0
DCGAN	0.9406	0.7628	0	0
DualGAN	0.1236	0.3775	0	0
Enhanced Vanilla GAN	0.9347	0.7155	0	0
WGAN	0.7323	0.9696	0	0

Table 5 talks about the KS test measures the statistical similarity between the distributions of the real data and the generated data. The lower the KS statistic, the better the model matches the real data distribution.

Vanilla GAN shows high KS statistics in both the Concentrating and Relaxed states, indicating that it struggles to match the real data distribution. The low p-values suggest that the differences between the real and generated distributions are statistically significant, highlighting that Vanilla GAN does not perform well in replicating the underlying data distribution.

Like Vanilla GAN, DCGAN also has relatively high KS statistics, indicating that its generated data differs significantly from the real data in both the Concentrating and Relaxed states. The p-values confirm that the KS test detects a significant deviation from the real data distribution.

DualGAN performs exceptionally well in the Concentrating state with a very low KS statistic of 0.1236, suggesting that the generated samples closely match the real data distribution. The Relaxed state shows a higher KS statistic, but it is still lower than that of Vanilla GAN and DCGAN, making DualGAN the best model in terms of matching the true data distribution.

Enhanced Vanilla GAN has relatively high KS statistics, similar to Vanilla GAN and DCGAN, indicating that the model struggles to replicate the true distribution of the data in both states. However, its performance is still better than Vanilla GAN and DCGAN, particularly in the Concentrating state.

WGAN has relatively low KS statistics compared to the other models, with 0.7323 in the Concentrating state and 0.9696 in the Relaxed state. These results suggest that WGAN is able to match the real data distribution better than most other models, especially in the Concentrating state, although its performance drops slightly in the Relaxed state.

5.4 Discussion of Results

- **WGAN** performed exceptionally well in terms of both the final loss values and metrics such as signal energy, with a strong performance in maintaining the real data distribution. However, some negative PSNR values in the Relaxed state suggest potential issues with signal reconstruction quality.
- **DualGAN** demonstrated the best results in terms of statistical alignment, especially with lower KL Divergence and FID scores, making it highly suitable for applications where maintaining data distribution similarity is crucial.
- **Enhanced Vanilla GAN** performed well in preserving spectral characteristics and achieved good performance in terms of signal energy and PSNR in both states, though it did not surpass WGAN in other areas.
- **Vanilla GAN** and **DCGAN** had challenges in matching signal energy and distribution characteristics, with Vanilla GAN also showing signs of overfitting as indicated by its high classification accuracy.

Loss Metrics:

The discriminator and generator loss values were analyzed to evaluate model training stability:

- **WGAN** showed the lowest final losses (–46325.38 for discriminator and –40421.15 for generator), indicating stable and effective training.

- **DualGAN** and **Enhanced Vanilla GAN** maintained balanced losses, reflecting better equilibrium between generator and discriminator.

5.5 Novelty and Innovation

This project introduces a novel approach to evaluating the quality of GAN-generated EEG data by using a diverse set of evaluation metrics that address statistical, spectral, and perceptual aspects. Key innovations include:

- The use of **DualGAN** for EEG data generation, which outperformed other models in statistical similarity (low KL Divergence and FID).
- Integration of a wide range of evaluation metrics to assess model performance from multiple perspectives, including signal energy, spectral preservation, and distribution alignment.
- The application of **classification accuracy** as an indirect measure of overfitting, offering insights into model stability.

5.6 Interpretation of Results

Model Performance:

- **WGAN** emerged as the best-performing model, achieving the highest combined score and lowest loss values, highlighting its ability to generate data that closely resembles the original EEG signals in terms of signal energy and distribution.
- **DualGAN** demonstrated exceptional performance in terms of statistical similarity, excelling in KL Divergence and FID, making it a suitable choice for applications requiring precise alignment with real data distributions.
- **Enhanced Vanilla GAN** was particularly strong in terms of spectral similarity, indicating its potential for applications where frequency-domain characteristics are critical.

Limitations:

- Some models, such as **Vanilla GAN**, exhibited high KS statistics, indicating discrepancies between the cumulative distributions of real and generated data.

- Negative **PSNR** values for WGAN in certain cases suggest challenges in the reconstruction of the signal, pointing to areas where further improvements could be made.

6. Evaluation and Conclusion

6.1 Final Evaluation

The project successfully met its goal of investigating various Generative Adversarial Network (GAN) architectures for EEG data generation and comparing the quality of synthetic data to original data. Among the evaluated models, **WGAN** emerged as the most effective architecture, with **WGAN** being the best performer overall. Despite these successes, the project faced challenges related to **computational resource demands** and **sensitivity to hyperparameter tuning**, which may affect the scalability and efficiency of the models. These challenges should be addressed in future iterations to enhance performance and optimize resource usage.

Table 6: Classification Accuracy

Model	**Classification Accuracy (Concentrating)**	**Classification Accuracy (Relaxed)**
Vanilla GAN	99.96%	99.99%
DCGAN	99.97%	99.99%
DualGAN	96.44%	97.93%
Enhanced Vanilla GAN	99.94%	99.98%
WGAN	98.30%	98.68%

Note: Lower classification accuracy indicates better generalization of the synthetic data. **DualGAN** outperforms the other models, showing the lowest classification accuracy in both **Concentrating** and **Relaxed** states, suggesting that it generates data that is more representative of the original data.

6.2 Insights Gained

The evaluation reinforced the importance of selecting the right GAN architecture for specific tasks. The comparison of various architectures highlighted that even small changes in the model can have a significant impact on the quality of the generated EEG signals. For instance, **DualGAN** demonstrated better generalization, making its synthetic data closer to the original EEG signals, whereas other models like **Vanilla GAN** and **Enhanced Vanilla GAN** showed higher classification accuracies, implying less variability in the generated data.

Additionally, the evaluation underscored the need for **standardized metrics** when comparing synthetic EEG data to real data. This allows for more objective assessment and helps in determining which model best captures the characteristics of EEG signals.

6.3 Future Work

Based on the insights gained from this project, future research could explore several areas to further enhance the quality of synthetic EEG data:

1. **Hybrid GAN Architectures**: Combining different GAN models may capitalize on the strengths of each, potentially leading to even better data generation. For example, merging **WGAN** and **DCGAN** could combine the stability of WGAN with the spatial features learned by DCGAN.
2. **Fine-Tuning GAN Models**: The current models encountered issues such as negative **Peak Signal-to-Noise Ratio (PSNR)** values and challenges in signal reconstruction. Fine-tuning the training processes for models like **WGAN** and **DualGAN** could help address these problems and improve the fidelity of the generated signals.
3. **Temporal Attention Mechanisms**: Introducing **temporal attention mechanisms** could help the models capture the time-sensitive patterns in EEG signals more accurately, improving the model's performance in tasks requiring fine temporal dynamics, such as emotion recognition and seizure prediction.
4. **Generalized GAN Models for Biophysical Data**: Future research could explore the creation of generalized GAN models capable of generating synthetic EEG, ECG, and EMG data. This could involve conditioning the input on the type of signal (EEG, ECG, or EMG) and using a shared

discriminator. Such models could be useful for diverse applications, including **emotion recognition** (EEG), **anomaly detection** (ECG), and other biophysical data analysis tasks.

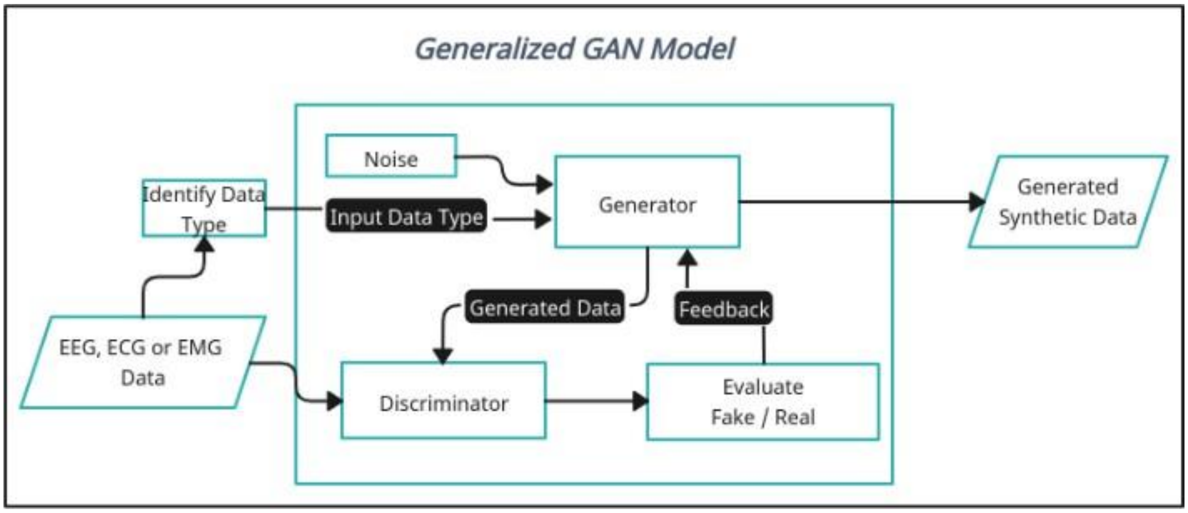


Figure 4: Generalized GAN Model

7. Conclusion

This project significantly contributes to the field of **EEG data generation** by evaluating various **GAN architectures**. The results demonstrate the potential of synthetic EEG data to augment real datasets, especially in fields where large and diverse datasets are necessary for model training. **WGAN** and **DualGAN** showed great promise, excelling in generating high-quality synthetic EEG data that closely matched real data in terms of **signal energy** and statistical distributions. These findings suggest that GAN-based models, especially **WGAN** and **DualGAN**, have great potential for applications in **neuroscience**, **machine learning**, and **brain-computer interface (BCI)** systems, where generating large amounts of labeled EEG data is often a bottleneck. The insights gained from this research lay a strong foundation for future work focused on refining GAN architectures for generating biomedical signals, addressing challenges in signal reconstruction, and exploring the broader applications of synthetic data in the biophysical domain.

8. References & Citations

1. Fahimi, F., et al., 2019. 'Towards EEG generation using GANs for BCI applications', 2019 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI), pp. 1–4. https://doi.org/10.1109/bhi.2019.8834503.
2. Duan, Y., et al., 2023. 'Multi-class image generation from EEG features with conditional generative adversarial networks', 2023 International Conference on Wireless Communications and Signal Processing (WCSP), pp. 534–539. https://doi.org/10.1109/wcsp58612.2023.10404439.
3. Luo, Y. and Lu, B.-L., 2018. 'EEG data augmentation for emotion recognition using a conditional Wasserstein GAN', 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 2535–2538. https://doi.org/10.1109/embc.2018.8512865.
4. Biswas, S., et al., 2023. 'Characterization of the event-related potentials during GAN-based generation of EEG signals and their data augmented subject classification', 2023 International Conference on Recent Advances in Electrical, Electronics & Digital Healthcare Technologies (REEDCON), pp. 717–722. https://doi.org/10.1109/reedcon57544.2023.10151321.
5. Dong, Y., et al., 2022. 'An approach for EEG data augmentation based on deep convolutional generative adversarial network', 2022 IEEE International Conference on Cyborg and Bionic Systems (CBS), pp. 347–351. https://doi.org/10.1109/CBS55922.2023.10115388.
6. Rasheed, K., et al., 2021. 'A generative model to synthesize EEG data for epileptic seizure prediction', IEEE Transactions on Neural Systems and Rehabilitation Engineering, 29, pp. 2322–2332. https://doi.org/10.1109/tnsre.2021.3125023.
7. Song, J., Zhai, Q., Wang, C. and Liu, J., 2024. 'EEGGAN-Net: enhancing EEG signal classification through data augmentation', Frontiers in Human Neuroscience, 18, 1430086. https://doi.org/10.3389/fnhum.2024.1430086.
8. Oh, J.H., et al., 2024. 'Graph-based conditional generative adversarial networks for major depressive disorder diagnosis with synthetic functional brain network generation', IEEE Journal of Biomedical and Health Informatics, 28(3), pp. 1504–1515. https://doi.org/10.1109/JBHI.2023.3340325.

9. Venugopal, A. and Faria, D., 2024. 'Boosting EEG and ECG classification with synthetic biophysical data generated via generative AI', Applied Sciences, 14, p. 10818. https://doi.org/10.3390/app142310818.
10. Manoharan, G. and Faria, D.R., 2024. 'Enhanced mental state classification using EEG-based brain–computer interface through deep learning', Lecture Notes in Networks and Systems, pp. 570–586. https://doi.org/10.1007/978-3-031-66431-1_40.
11. Bird, J.J., et al., 2018. 'A study on mental state classification using EEG-based brain-machine interface', 2018 International Conference on Intelligent Systems (IS), pp. 795–800. https://doi.org/10.1109/is.2018.8710576.
12. Bouallegue, G. and Djemal, R., 2020. 'EEG data augmentation using Wasserstein GAN', 2020 20th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA), pp. 40–45. https://doi.org/10.1109/sta50679.2020.9329330.
13. Lee, Y.-E., Lee, S.-H., Kim, S., Lee, J.-S. and Kim, D.-S., 2024. 'Enhanced Generative Adversarial Networks for Unseen Word Generation from EEG Signals', BCI 2024, pp. 1-4.
14. Chen, S., Xu, Q., Zhong, S. and Chen, K., 2019. 'GAN Evaluation by Multi-Method Fusion', 2019 2nd International Conference on Artificial Intelligence and Big Data (ICAIBD), Chengdu, China, pp. 36-44. https://doi.org/10.1109/ICAIBD.2019.8836993.
15. Goodfellow, I., et al., 2014. 'Generative Adversarial Nets', Advances in Neural Information Processing Systems.
16. Sarkar, P. and Etemad, A., 2020. 'CardioGAN: Attentive Generative Adversarial Network with Dual Discriminators for Synthesis of ECG from PPG', arXiv. Available at: https://arxiv.org/abs/2010.00104 [Accessed 6 Jan. 2025].
17. Tian, C., Ma, Y., Cammon, J., Fang, F., Zhang, Y. and Meng, M., 2023. 'Dual-Encoder VAE-GAN With Spatiotemporal Features for Emotional EEG Data Augmentation', IEEE Transactions on Neural Systems and Rehabilitation Engineering, 31, pp. 2018–2027. https://doi.org/10.1109/TNSRE.2023.3266810.
18. Hartmann, K.G., Schirrmeister, R.T. and Ball, T., 2018. 'EEG-GAN: Generative adversarial networks for electroencephalographic (EEG) brain signals', arXiv, [online]. Available at: https://arxiv.org/abs/1806.01875 [Accessed 6 Jan. 2025].
19. Bhat, S. and Hortal, E., 2021. 'GAN-Based Data Augmentation For Improving The Classification Of EEG Signals', Proceedings of the 14th Pervasive Technologies Related to Assistive

Environments Conference (PETRA '21), ACM, New York, NY, pp. 453-458. https://doi.org/10.1145/3453892.3461338.

20. Hamdi, A. and Ghanem, B., 2019. 'IAN: Combining Generative Adversarial Networks for Imaginative Face Generation', arXiv preprint arXiv:1904.07916.
21. Habashi, A.G., Azab, A.M., Eldawlatly, S. and Aly, G.M., 2023. 'Generative adversarial networks in EEG analysis: an overview', Journal of NeuroEngineering and Rehabilitation, 20(1), p. 40. https://doi.org/10.1186/s12984-023-01169-w.
22. Daly, J.J. and Wolpaw, J.R., 2008. 'Brain-computer interfaces in neurological rehabilitation', Lancet Neurology, 7(11), pp. 1032–1043. https://doi.org/10.1016/S1474-4422(08)70223-0.
23. Kavoosi, A., 2021. 'Generative Adversarial Network (GAN) for Creating Synthetic EEG & LFP (Both Normal and Abnormal)'. Medium. Available at: https://medium.com/@a.kavoosi1999/generative-adversarial-network-gan-for-creating-synthetic-eeg-lfp-both-normal-and-abnormal-d2ec72f712e5
24. Shehabi, B., 2021. 'EEG-Synthetic-Data-Using-GANs'. GitHub. Available at: https://github.com/basel-shehabi/EEG-Synthetic-Data-Using-GANs