EEG Anomalies Detection and Removal Using Generative Adversarial Networks

Amine Ahardane[1], Silverio Manganaro[1] and Roberta Avanzato[1]

[1]Department of Electrical, Electronics and Computer Engineering, University of Catania, Catania, Italy

Abstract

The human brain generates electrical activities measurable through electroencephalography (EEG). These signals are often contaminated by noise, and so its not easy to do an accurate analysis and interpretation which is essential for clinical applications such as epilepsy diagnosis, cognitive neuroscience, and brain-computer interfaces. Traditional denoising techniques frequently fall short in effectively distinguishing between signal and noise, especially when the noise sources exhibit complex and nonlinear characteristics. This paper explores the application of Generative Adversarial Networks (GANs) in denoising EEG signals, offering a data-driven approach to learn the complex structures of both clean and noisy EEG data. We detail the training of a classifier to distinguish between normal and abnormal EEG signals, the development of an AutoEncoder to compress and reconstruct signals, and the use of a Wasserstein GAN (WGAN) to manipulate abnormal signals towards normality in the latent space. Our results demonstrate the potential of GAN-based methods in enhancing EEG signal quality, paving the way for more accurate clinical analyses.

Keywords

EEG Denoising, Generative Adversarial Networks, AutoEncoders, Deep Learning

1. Introduction

The human brain is a complex network of connected neurons that produces electrical activity; this can be measured through EEG. Although EEG signals are quite useful to understand brain functioning, they are heavily contaminated by numerous sources of noise of very different kinds: environmental interference, muscle activity, and electrical artifacts[1]. For instance, myogenic artifacts are created in muscle movements around the head and face, while ocular artifacts are generated by eye blinks and movements. Moreover, brain signals are complex and non-stationary; this makes the separation of neural activity from noise a challenging task. Denoising EEG signals is important because it allows us to better analyze and interpret signals. This is important in various settings, such as clinical applications for the diagnosis of epilepsy, cognitive neuroscience, and brain-computer interfaces[2, 3].

To meet these challenges, traditional denoising techniques, such as band-pass filtering, independent component analysis (ICA) [4, 5, 6], and wavelet transforms and other domain transforms [7, 2, 8], have been put into practice. For example, linear filtering methods may remove significant portions of the EEG signal along with the noise, leading to a loss of critical information. There is an assumption at the root of ICA that the sources are statistically independent, which might not hold in practical scenarios. Wavelet transformation does provide a multi-resolution approach but can be computation-intensive, and also this technique cannot deal with noise that is highly non-stationary in nature. These techniques often tend to delete the main important features of the original EEG signal, especially in cases when noise sources become complex and nonlinear in nature[9, 10].

In recent years, machine learning techniques like Generative Adversarial Networks have shown great potential in denoising EEG signals. GANs were first introduced by Goodfellow et al. in 2014 [11, 12], where we have two neural networks, named the generator and the discriminator. The generator tries to generate data samples that can come from the real data distribution, and the discriminator tries to differentiate between the real data examples and those created by the generator. In the end, the GAN updates its weights so that the generator can fool the discriminator. This adversarial training process makes GANs capable of learning complex structures of the data, which subsequently becomes useful for problems like denoising where typical methods fall short.

Traditional GANs have several training problems, such as mode collapse and instability during training. To handle these problems, (Arjovsky et al. 2017 [13]) introduced Wasserstein Generative Adversarial Networks, or WGANs[14]. The methodology presented in this paper is comprised of three principal stages:

1. **Classifier Training:** Initially, a classifier is developed to accurately discern between normal and abnormal EEG signals using the provided dataset. This classifier forms the basis of our subsequent denoising process by ensuring the precise identification of signal types.

SYSTEM 2025: 11th Sapienza Yearly Symposium of Technology, Engineering and Mathematics. Rome, June 4-6, 2025
roberta.avanzato@unict.it (R. Avanzato)

2. **AutoEncoder Development:** Subsequently, an AutoEncoder is constructed and trained exclusively on normal EEG data. This network is tasked with compressing and reconstructing the EEG signals, thereby establishing a representative latent space that captures the salient features of normal brain activity. This latent space serves as a critical reference for subsequent processing.
3. **WGAN Training:** In the final stage, a Wasserstein Generative Adversarial Network (WGAN) is employed to refine the latent representations of abnormal signals. The generator within the WGAN is designed to transform these abnormal latent representations so that they more closely resemble those of normal signals, effectively deceiving the discriminator. This transformation facilitates an improved reconstruction of the original signal, ultimately converting unhealthy signals into representations that are consistent with healthy EEG characteristics—a process that holds significant potential for clinical applications.

By leveraging the advanced capabilities of GAN-based architectures, the proposed approach addresses several limitations associated with traditional denoising techniques. The integration of classifier training, AutoEncoder development, and WGAN-based latent space manipulation provides a comprehensive and robust solution for EEG signal denoising. This integrated methodology not only enhances the quality of EEG signals but also increases their utility in clinical diagnostics and cognitive neuroscience research, paving the way for the development of more effective brain-computer interface technologies.

2. Related Works

There are several recent studies that deals with denoising EEG signals using various methodologies. Among these, (Peng Yi et al. 2021 [15]) using transformer, proposes a novel approach integrating non-local and local self-similarity of EEG signals through a 1-D EEG signal denoising network with a 2-D transformer.Using self-similarity characteristics, EEGDnet has a improvements in removing ocular and muscle artifacts compared to other state-of-the-art models.

In (Eoin Brophy et al. 2022 [16]), the focus is on utilizing Generative Adversarial Networks (GANs) to denoise EEG time series data. We have to deal with artifacts induced in real-world Brain-Computer Interface (BCI) applications[17, 18], which degrade performance.Using GANs the model is able to, given noisy EEG signals transforming it into clean ones,demonstrating promising results in quantitative metrics such as power spectral density and signal-to-noise ratio. This study has the capability to handle multiple artifcat types, having the potential of being used in applications in portable EEG devices.

Similarly we have (Yang An et al. 2022 [19]) approach that proposes a new loss function to retain original information and energy in the filtered signals, demonstrating that the performances are comparable to manual denoising methods while at the same time, we have a significant reduction of the processing time.In this method, we have also an incorporation of a new normalization method ensures stable generation of EEG signals by the GAN model, allowing for automatic denoising across different subjects' data[20].

Another notable contribution in this domain is presented by (Wang et al. 2022 [21]) in their work titled "An improved Generative Adversarial Network for Denoising EEG signals of brain-computer interface systems[22, 23].

The authors propose a novel GAN-based framework that includes a generator with BiLSTM and LSTM layers and a discriminator composed of multiple CNN layers. This architecture aims to address the limitations of previous models by reducing mode collapse and improving the convergence stability during training. The study demonstrates that their improved GAN significantly outperforms traditional methods and other deep learning approaches, particularly in scenarios with high noise levels. Their results show enhanced performance in terms of root mean square error (RMSE) and Pearson correlation coefficient (CC), making it a robust solution for real-time EEG denoising in brain-computer interface (BCI) systems.

The improved GAN framework not only excels in denoising accuracy but also enhances the robustness of EEG signal processing by effectively handling artifacts such as eye blinks and muscle movements. The authors utilized the EEGdenoiseNet (Zhang et al. 2022 [24]) dataset to benchmark their model, which includes a diverse range of artifact-contaminated EEG signals. The proposed model's ability to maintain high performance across varying signal-to-noise ratios (SNR) underscores its potential for practical applications in BCI systems, where real-time processing and reliability are crucial [25, 26].

Moreover, this study show the importance of incorporating BiLSTM and LSTM layers in the generator network to capture temporal dependencies in EEG signals. The discriminator network, consisting of multiple CNN layers, ensures that the generated clean EEG signals closely resemble the true signals, thus enhancing the overall quality of the denoised output.

In addition to the structural improvements, the training strategy employed by Wang et al. achieves superior performance. By carefully balancing the learning rates and optimizing the loss functions for both the generator and discriminator, the model is able to avoid the common problems of the GAN, overfitting and non-convergence.

Other interesting works that apply GAN-based models for anomaly detection (Zenati et al. 2018 [27]). In this works, state-of-the-art performances are shown on image and network intrusion datasets. The approach consists of using a GAN that learns a latent representation of normal samples in such a way that the GAN is able to detect anomalies by measuring reconstruction and discriminator-based loss. High performance is shown on the MNIST and KDD99 datasets in the method proposed.

The proposed method by (Niu et al. 2020 [28]), is called LSTM-based VAE-GAN. This solves the inefficiency of real-time space-to-latent space mapping when doing anomaly detection. Here we exploit the temporal dependencies that an LSTM Network can capture and the reconstruction and discrimination abilities of VAE and GAN respectively.

(Zhang et al. 2023 [29]) is a novel GAN-based model for unsupervised anomaly detection in multivariate time series (MTS). In this work, the authors use a self-training framework wherein they have a teacher model generating high-quality pseudo-labels iteratively training a student model. STAD-GAN involves a generator-discriminator structure with a neural network classifier. The generator maps the normal data distribution. The discriminator amplifies the reconstruction error of abnormal data to enhance recognition performance. It means that the performance of the anomaly classifier will be improved through self-training by iteratively refining the dataset[30].

Our approach offers a different contribution. Unlike existing methodologies that often rely on direct correspondences between healthy and unhealthy signals for denoising, our study uses a comprehensive data-driven approach without such direct correspondences, by transforming various types of unhealthy signals into healthy ones within the same architectural framework.

Many methods proposed in the literature typically address only one type of artifact or abnormality, these methods limits versatility and applicability across different scenarios and they often require a specialized design or customization for each specific type of artifact or abnormality, which can be both time-consuming and resource-intensive.

Specifically, our model is designed to handle potentially a diverse range of signal abnormalities and artifacts, ensuring that it can be retrained and adapted to improve any given signal's quality without the need for significant modifications.

3. Proposed Model

Our approach to enhance abnormal data consists of three main phases. The goal of our approach is to transform the latent representation of an abnormal signal to resemble those of normal signals. By doing so, the reconstructed signal from this transformed latent representation will be an enhanced version of the original abnormal signal. This enhancement process exploit the combination of classification, compression, and adversarial training to improve the quality and normalcy of abnormal data.

3.1. Phase 1: Classifier

In the first step, we train a very simple binary classifier. The classifier learns the difference between normal and abnormal data, in fact it's role is to accurately identify abnormal signals, which will later be processed to enhance their quality.

3.2. Phase 2: AutoEncoder

We now train an AutoEncoder (as shown in Figure 2) using only the normal data. The job of this AutoEncoder is to compress and reconstruct each given signal. Once it learns what the typical data distribution is, the AutoEncoder can then be able to reconstruct signals that reflect normal behavior.

3.3. Phase 3: WGAN

For the final phase, we train a Wasserstein Generative Adversarial Network (WGAN). The Generator in the WGAN works on the encoded anomalous latent-space representation in such a way that the latter becomes more like the latent representations of normal signals, those obtained by pre-trained AutoEncoders. The Discriminator's role is to distinguish between the true normal latent representations and the manipulated ones produced by the Generator. The objective is to train the Generator to produce enhanced versions of the abnormal signals that fool the Discriminator into classifying them as normal.

The objective function for Wasserstein Generative Adversarial Networks (WGAN) is given by:

$$\min_{G} \max_{D \in \mathcal{D}} \mathbb{E}_{\mathbf{x} \sim \mathbb{P}_r}[D(\mathbf{x})] - \mathbb{E}_{\mathbf{z} \sim \mathbb{P}_z}[D(G(\mathbf{z}))] \quad (1)$$

Here, $\mathbb{P}_r$ denotes the real data distribution, $\mathbb{P}_z$ denotes the prior distribution of the input noise $\mathbf{z}$ to the Generator G, and $\mathcal{D}$ is the set of 1-Lipschitz functions which the Discriminator D is a part of.

In particular, we will enforce the 1-Lipschitz constraint, a gradient penalty term is added to the loss function. The modified objective with gradient penalty (WGAN-GP) is:

$$\begin{aligned}\mathcal{L} = \mathbb{E}_{\mathbf{x} \sim \mathbb{P}_r}[D(\mathbf{x})] - \mathbb{E}_{\mathbf{z} \sim \mathbb{P}_z}[D(G(\mathbf{z}))] \\ + \lambda \mathbb{E}_{\hat{\mathbf{x}} \sim \mathbb{P}_{\hat{\mathbf{x}}}}\left[(\|\nabla_{\hat{\mathbf{x}}} D(\hat{\mathbf{x}})\|_2 - 1)^2\right]\end{aligned} \quad (2)$$

where $\mathbb{P}_{\hat{\mathbf{x}}}$ is the distribution of the interpolated samples between the real and generated data, and λ is the penalty coefficient.

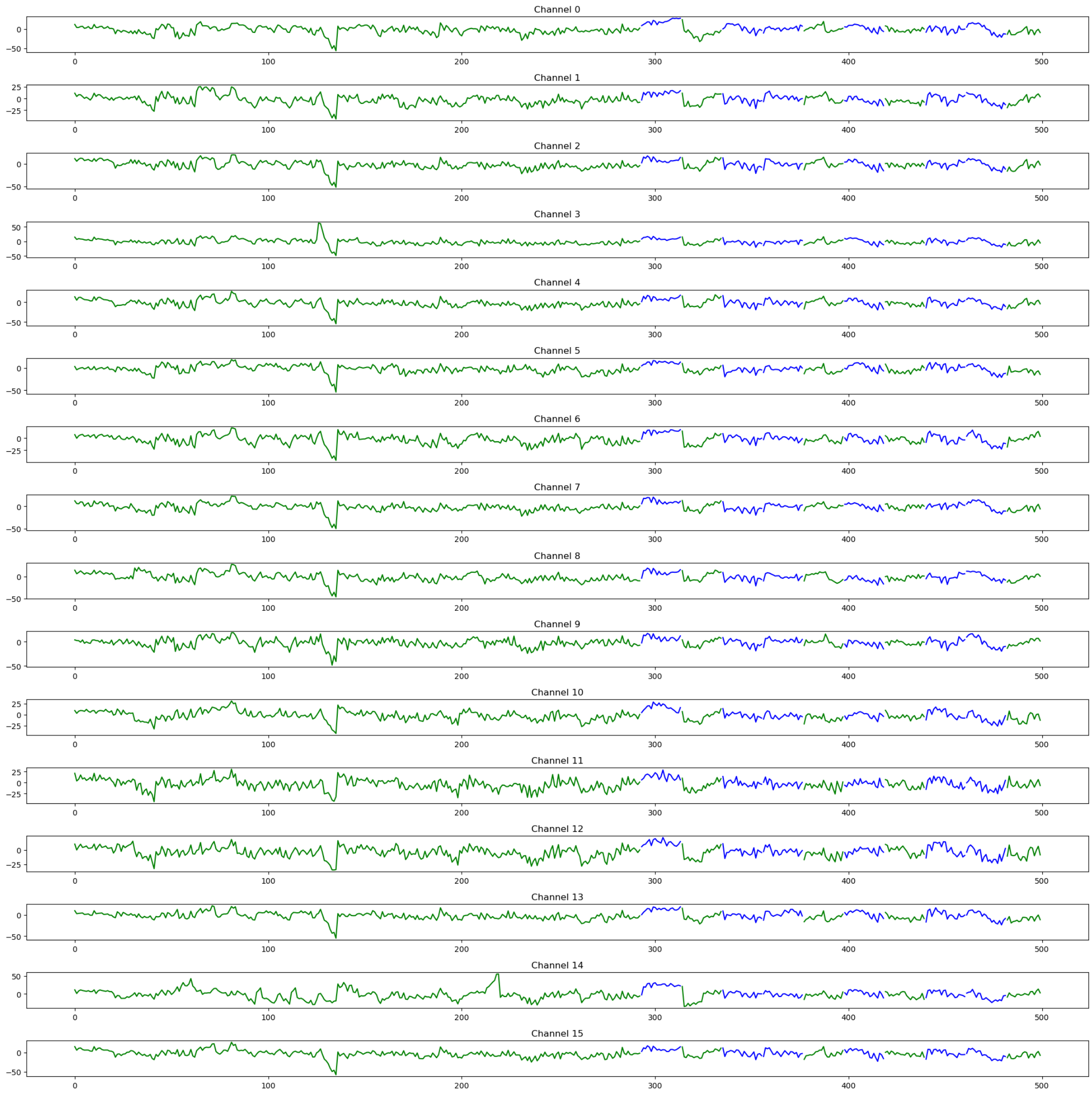

Figure 1: example of EEG signals taken from the dataset.

4. Implementation

More specifically, our method is developed in three principal phases: training of the classifier to differentiate between normal and abnormal signals, training of the AutoEncoder signal compression and its reconstruction, training of Wasserstein Generative Adversarial Network (WGAN) that transform unhealthy into healthy signals. Then each phase is described in detail with extensive descriptions for data preparation, architectural choices, training processes, and integration steps. Note that in our experiments, we used Weights & Biases (wandb)—a popular tool used for tracking machine learning experiments. It's a strong platform for logging metrics, visualizing results, and collaborating on project implementations. In our implementation, we have integrated wandb to monitor training and the evaluation processes of our models. In Table 1, we depict the various hyperparameters of the models.

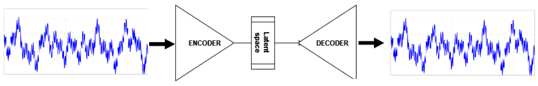

Figure 2: Architecture of the AutoeEncoder. Reconstruction of EEG signal .

Parameter	Classifier	WGAN	Autoenc.
Batch Size	32	128	128
Epochs	1000	1000	1,000
Sequence Length	50	50	50
Hidden Size	-	256	256
Discrimin. Size	-	256	-
Generator Size	-	1,024	-
Learning Rate	-	0.0001	-
Learning Rate	-	0.0005	-
Learning Rate	0.001	-	0.0005
Critic Iterations	-	5	-

Table 1
Training Parameters for Classifier, WGAN, and Autoencoder

4.1. Phase 1: Classifier

4.1.1. Data Preparation

The main dataset (shown in Figure 1) used in our experiments consists of EEG signals recorded under various conditions. The Healthy dataset consists of over 1,500 one- and two-minute EEG recordings obtained from 109 volunteers. The EEG recordings were collected using a 64-channel setup with the BCI2000 system. Each subject performed 14 experimental runs, which included:

1. Baseline, eyes open
2. Baseline, eyes closed
3. Task 1: Open and close left or right fist
4. Task 2: Imagine opening and closing left or right fist
5. Task 3: Open and close both fists or both feet
6. Task 4: Imagine opening and closing both fists or both feet
7. Task 1 (repeat)
8. Task 2 (repeat)
9. Task 3 (repeat)
10. Task 4 (repeat)
11. Task 1 (repeat)
12. Task 2 (repeat)
13. Task 3 (repeat)
14. Task 4 (repeat)

During these tasks, subjects performed or imagined performing the movement corresponding to the stimuli presented on the screen. All these tasks are designed to evoke different types of motor and imagery behaviors and have the event type tag-based indicators:

- **T0**: Rest
- **T1**: Onset of motion (real or imagined) of the left fist (Tasks 1 and 2) or both fists (Tasks 3 and 4)
- **T2**: Onset of motion (real or imagined) of the right fist (Tasks 1 and 2) or both feet (Tasks 3 and 4)

4.1.2. Unhealthy Dataset

The Unhealthy dataset comprises raw 18-channel EEG recordings from 7 human participants with orthopedic impairment during motor imagery (MI) tasks. Due to component removal, some subjects have slightly fewer channels (14 or 15) to eliminate noisy channels and improve data quality.

Participants performed a series of MI-related trials across three sessions, each consisting of 40 trials with four different MI tasks presented in random order. Each trial included:

- 3 seconds of fixation cross
- 4 seconds of visual cue
- 3 seconds of letters indicating the ready state
- 5 seconds of imaginary movement

The dataset includes filtered EEG data (8-30 Hz with a notch filter) and labels for 10 movement types and a rest state. The corresponding electrode names are provided in TSV files, ensuring a 1:1 mapping with the CSV data.

Note that we have healthy and unhealthy signals, and no correspondences between them (so given an unhealthy signal, we don't have the corresponding healthy one) .

4.1.3. Supplementary Dataset

After training the classifier for WGAN, we tested it using the Epilepsy2 dataset, which contains single-channel EEG measurements from 500 subjects. [31]

Each subject's brain activity was recorded for 23.6 seconds, resulting in a comprehensive collection of EEG data. To facilitate detailed analysis and model training, the dataset was divided and shuffled into 11,500 samples, each representing a 1-second segment of EEG data sampled at 178 Hz. This shuffling mitigates sample-subject association and ensures a robust evaluation environment.

The dataset is divided into three groups: 60 samples for training, 20 samples for validation, and 11,420 samples for the test. The validation set was appended directly to the end of the training file to allow easier reproducibility in case a validation set is required. The small size of the training set will actually assist in testing the transfer learning capabilities, and the test set retains the original distribution for a fair evaluation.

The dataset contains five unique classification labels related to different conditions or measurement locations:

1. Eyes open
2. Eyes closed
3. EEG measured in a healthy brain region
4. EEG measured in the region of a tumor
5. Subject experiencing a seizure episode

Our classifier achieved an impressive accuracy of 95% on this dataset, demonstrating its effectiveness in distinguishing between seizure and non-seizure states. For reference, the Mini Rocket classifier, a well-known model in the field, achieves a test accuracy of 96.25%, highlighting the competitive performance of our approach.

The actual classifier, autoencoder and wgan they will be trained using the main dataset.

4.1.4. Classifier Architecture

We implemented a binary classifier using a neural network architecture. The classifier consists of multiple fully connected (dense) layers, each followed by a Rectified Linear Unit (ReLU) activation function, and a final Sigmoid activation function to output the probability of the signal being abnormal. The detailed architecture is as follows:

- **Input Layer**: Accepts the preprocessed EEG signal.
- **Hidden Layers**:
  - First layer: 512 neurons, ReLU activation.
  - Second layer: 256 neurons, ReLU activation.
  - Third layer: 64 neurons, ReLU activation.
- **Output Layer**: 1 neuron, Sigmoid activation.

This architecture was chosen for its balance between complexity and performance that allows the model to classify effectively between unhealthy and healthy signal.

4.1.5. Training the Classifier

The classifier was trained using the Adam optimizer with a binary cross-entropy loss function. The training process involved multiple epochs, where the model parameters were updated iteratively to minimize the loss function. During training, the following steps were performed:

- **Forward Pass**: The input data was passed through the network to obtain the output probabilities.
- **Loss Calculation**: The binary cross-entropy loss between the predicted probabilities and the true labels was calculated.
- **Backward Pass**: Gradients were computed using backpropagation, and the network weights were updated using the Adam optimization algorithm.

The training loss is described in Figure 4 and its correspondig loss of evaluation dataset Figure 3

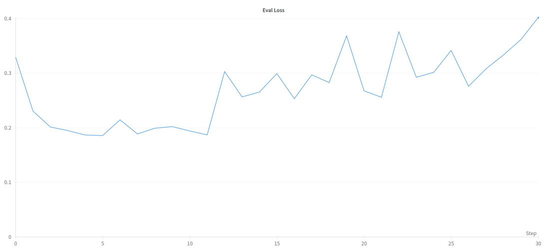

Figure 3: Classifier eval loss while training

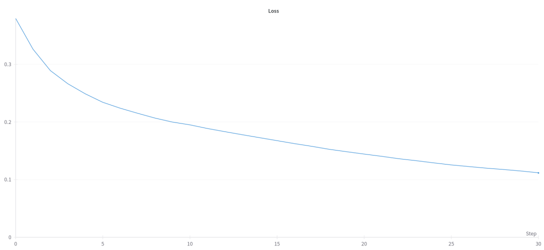

Figure 4: Classifier loss while training

4.2. Phase 2: AutoEncoder

4.2.1. AutoEncoder Architecture

The AutoEncoder was designed to compress and reconstruct the EEG signals. The AutoEncoder consists of two main parts: the encoder and the decoder. The encoder reduces the dimensionality of the input data to a latent space, capturing the essential features, while the decoder reconstructs the data back to its original form from this compressed representation.

- **Encoder**:
  - First layer: 128 neurons, ReLU activation.
  - Second layer: 64 neurons, ReLU activation.
  - Third layer: 32 neurons (latent space), ReLU activation.

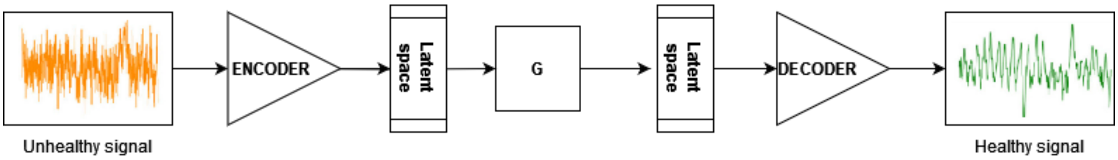

Figure 5: WGAN in the inference stage, the latent space of the generator output will be converted by the decoder of the AutoEncoder (previously trained) to the corresponding Healthy signal.

- **Latent Space**: Represents the compressed form of the input data.
- **Decoder**:
  - First layer: 64 neurons, ReLU activation.
  - Second layer: 128 neurons, ReLU activation.
  - Output layer: Same size as input, linear activation.

This architecture allows the AutoEncoder to learn a compressed representation of normal EEG signals, needed for the training of WGAN.

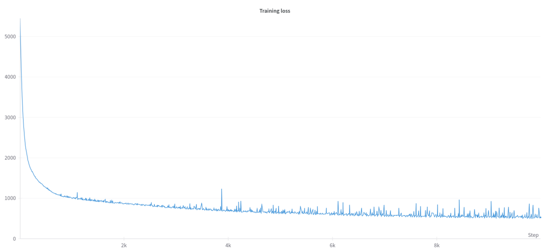

Figure 6: AutoEncoder loss while training

4.2.2. Training the AutoEncoder

The AutoEncoder was trained using the mean squared error (MSE) loss function, which measures the difference between the input and the reconstructed output. The training involved the following steps:

- **Forward Pass**: The normal EEG signals were passed through the encoder to obtain the latent representation, and then through the decoder to reconstruct the signals.
- **Loss Calculation**: The MSE loss was computed between the original and reconstructed signals.
- **Backward Pass**: Gradients were computed, and the network weights were updated using the Adam optimizer.

The training was done over several epochs while checking the performance of the model on the validation set in terms of the reconstruction error. The objective was to minimize the reconstruction error, which should allow the AutoEncoder to reconstruct normal EEG signals. The training loss is described in Figure 6

4.3. Phase 3: WGAN

4.3.1. WGAN Architecture

The WGAN consists of a generator and a discriminator. The generator aims to transform the encoded latent space of abnormal signals to resemble the latent space of normal signals, while the discriminator evaluates the authenticity of the generated signals. The detailed architectures are as follows:

- **Generator**:
  - First layer: 64 neurons, ReLU activation.
  - Second layer: 128 neurons, ReLU activation.
  - Output layer: Same size as latent space, linear activation.
- **Discriminator**:
  - First layer: 128 neurons, ReLU activation.
  - Second layer: 64 neurons, ReLU activation.
  - Output layer: 1 neuron, Sigmoid activation.

4.3.2. Training the WGAN

In order to provide stable training dynamics and avoid problems like mode collapse, training of the WGAN was performed with Wasserstein loss and gradient penalty (WGAN-GP). The training process involved alternating between optimizing the discriminator and the generator. The following steps were performed during training:

- **Discriminator Training**:
  - Real latent representations (from the encoder) and fake latent representations (from the generator) were fed into the discriminator.
  - The discriminator's loss was calculated based on its ability to discern real from fake representations.

    - Gradients were computed, and the discriminator's weights were updated.
- **Generator Training**:
    - The generator produced fake latent representations from abnormal signals.
    - These representations were fed into the discriminator.
    - The generator's loss was calculated based on its ability to fool the discriminator.
    - Gradients were computed, and the generator's weights were updated.

The discriminator was trained more frequently than the generator to maintain a balance between the two networks. Thanks to this adversarial training process, the generator learned how to convert signals from unhealthy to healthy. The training loss is described in Figure 7

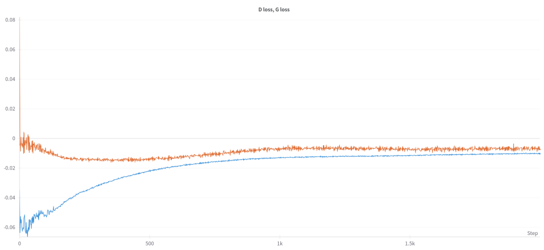

Figure 7: WGAN loss while training

4.4. Integration and Testing

After training each component, we integrated the models (shown in Figure 8) to form a cohesive system. So, each sub-model work will be:

- **Classification**: The classifier identified abnormal signals from the EEG data.
- **Encoding**: The identified abnormal signals were encoded into latent representations using the autoencoder's encoder.
- **Transformation**: The WGAN generator transformed these latent representations to match the distribution of normal latent representations.
- **Reconstruction**: The autoencoder's decoder reconstructed the enhanced signals from these transformed latent representations.

At inference time (shown in Figure 5) the output of the generator will be the input of the decoder to have the corresponding Healthy signal. The combined system was additionally analyzed with other performance measurements such as: accuracy, RMSE, MPC. mean entropy. Additionally, qualitative assessments of the reconstructed signals were performed to make sure that we remove abnormalities but also preserve essential features of the EEG signals.

With this method, we will effectively enhance the quality of EEG signals and transform unhealthy signals to healthy ones. The deployment of our model provides a scalable solution for EEG signal denoising applicable in many clinical and research applications.

5. Results

The evaluation of our model's performance included several key metrics: Root Mean Square Error (RMSE) calculated in the latent space, Pearson correlation, classifier accuracy, and entropy changes of the samples after manipulation. The results are summarized in Table 2.

Table 2
Performance Evaluation Metrics

Metric	Value
RMSE in Latent Space (GAN)	0.004166
Mean Pearson Correlation (GAN)	-0.116
Classifier Accuracy (Healthy Samples)	0.9047
Samples with Increased Mean Entropy	585
Samples with Decreased Mean Entropy	1728

The RMSE in the latent space for GAN architecture is very low (0.004166), meaning that the signals reconstructed with the use of GAN are close to the expected latent representations. This further means that the RMSE value is quite low and close to zero, thus suggesting the fact that GAN can successfully capture the structure of signals.

The mean Pearson correlation for the GAN is -0.116, which indicates a slight negative correlation between the generated signals and the true signals. While it is not a strong negative correlation, it still denotes that there is more room of improvements for the model to generate signals that closely resemble the true signal patterns.

The classifier, which identify the healthy samples for the unhealthy one, achieved an accuracy of 0.9047. This high accuracy shows that the classifier is effective in differentiating between the EEG recordings of normal and abnormal cases, giving a good basis for further processing by the GAN. In an entropy perspective, after manipulation was performed on the 2334 samples, 585 samples showed increased mean entropy while 1728 samples showed decreased mean entropy. Entropy represents a measure of randomness or disorder in the signals. If the entropy has increased, this can be taken as an indication that the signals grew to be more complex, possibly indicating the introduction of noise or artifacts. On the other hand, a decrease in entropy means a decrease in signal complexity, which may be explained by effective signal

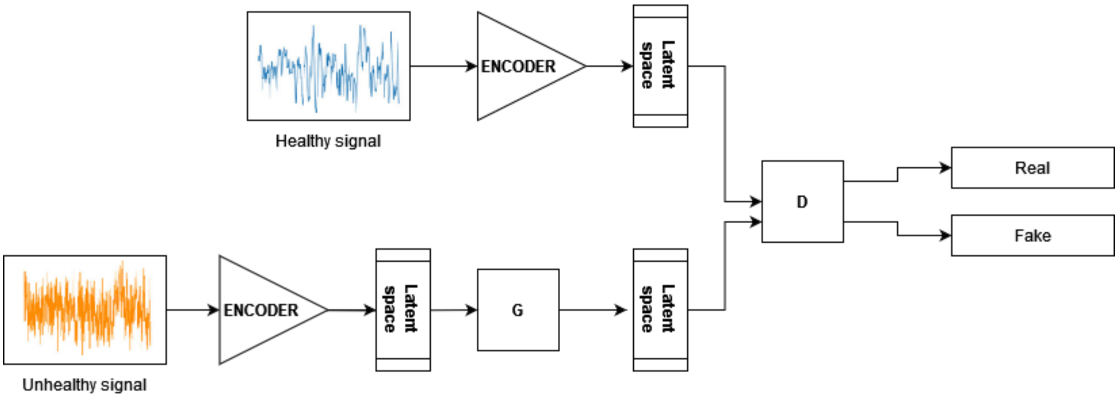

Figure 8: WGAN architecture. Unhealthy signal EEG to Healthy one.

denoising and enhancement by the GAN. In Figure 9, is shown an example of transforming an unhealthy signal to healthy one, its clear that probably by adding to the loss also a constraint to the amplitude, we will obtain a better signal.

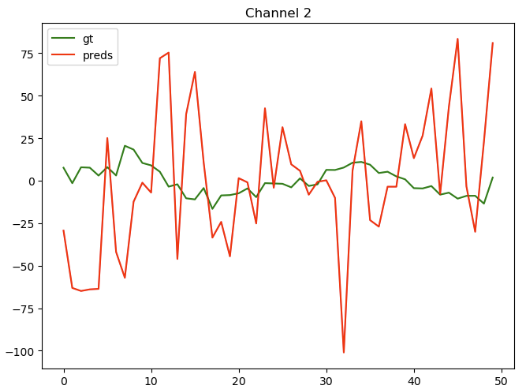

Figure 9: Example of denoising an unhealthy signal

These results show promise for enhancing abnormal EEG signals using a GAN-based approach. Particularly encouraging is the very low RMSE and high classifier accuracy, while the entropy results provide further insight into the nature of signal transformations. Further improvements in the model could potentially address the negative correlation and refine the signal manipulation process to achieve even better outcomes.

6. Conclusions and Future Works

This study has presented a novel framework for enhancing abnormal EEG signals using a GAN-based approach that synergistically integrates a classifier, an AutoEncoder, and a WGAN. Using the unique capabilities of each component, the proposed method effectively distinguishes between normal and abnormal signals, learns a compact latent representation of healthy EEG patterns, and transforms noisy or abnormal latent representations towards normality. The experimental results, characterized by a low RMSE in the latent space and robust classifier performance, underscore the potential of the framework to achieve meaningful denoising while preserving critical signal features.

Our integrated approach addresses several of the limitations inherent in traditional signal processing techniques, offering a data-driven alternative that accommodates the complex, non-stationary, and nonlinear nature of EEG noise. Despite demonstrating promising results, particularly in terms of classifier accuracy and latent space convergence, the study also reveals areas where further refinement is warranted, such as improving the Pearson correlation between reconstructed and true signals and optimizing entropy measures in the enhanced output.

Future work should explore the scalability of the framework across broader and more diverse datasets, including multi-channel recordings and clinical datasets encompassing a wider range of neurological conditions. Moreover, the real-time implementation of the denoising framework could significantly enhance its applicability in BCI systems. Integrative research efforts that combine EEG with other neuroimaging modalities, such as fMRI or MEG, may further enrich the diagnostic precision and clinical relevance of the proposed methodology. In general, the promising results of this study pave the way for advanced EEG signal processing paradigms, offering valuable insights into both cognitive neuroscience research and practical clinical diagnostics.

Declaration on Generative AI

During the preparation of this work, the authors used ChatGPT, Grammarly in order to: Grammar and spelling check, Paraphrase and reword. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] N. Boutarfaia, S. Russo, A. Tibermacine, I. E. Tibermacine, Deep learning for eeg-based motor imagery classification: Towards enhanced human-machine interaction and assistive robotics, life 2 (2023) 4.

[2] V. Ponzi, S. Russo, A. Wajda, R. Brociek, C. Napoli, Analysis pre and post covid-19 pandemic rorschach test data of using em algorithms and gmm models, in: CEUR Workshop Proceedings, volume 3360, 2022, p. 55 – 63.

[3] N. N. Dat, V. Ponzi, S. Russo, F. Vincelli, et al., Supporting impaired people with a following robotic assistant by means of end-to-end visual target navigation and reinforcement learning approaches, in: CEUR Workshop Proceedings, volume 3118, CEUR-WS, 2021, pp. 51–63.

[4] S. Makeig, A. J. Bell, T.-P. Jung, T. J. Sejnowski, Independent Component Analysis of Electroencephalographic Data, Advances in Neural Information Processing Systems (1996). URL: https://www.researchgate.net/publication/2242002_Independent_Component_Analysis_of_Electroencephalographic_Data.

[5] I. E. Tibermacine, A. Tibermacine, W. Guettala, C. Napoli, S. Russo, Enhancing sentiment analysis on seed-iv dataset with vision transformers: A comparative study, in: Proceedings of the 2023 11th international conference on information technology: IoT and smart city, 2023, pp. 238–246.

[6] S. I. Illari, S. Russo, R. Avanzato, C. Napoli, A cloud-oriented architecture for the remote assessment and follow-up of hospitalized patients, in: CEUR Workshop Proceedings, volume 2694, 2020, p. 29 – 35.

[7] Z. Weidong, L. Yingyuan, EEG Multiresolution Analysis Using Wavelet Transform, Technical Report, Defense Technical Information Center, 2001. URL: https://apps.dtic.mil/sti/tr/pdf/ADA411450.pdf.

[8] N. Brandizzi, A. Fanti, R. Gallotta, S. Russo, L. Iocchi, D. Nardi, C. Napoli, Unsupervised pose estimation by means of an innovative vision transformer, in: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), volume 13589 LNAI, 2023, p. 3 – 20. doi:10.1007/978-3-031-23480-4_1.

[9] B. Ladjal, I. E. Tibermacine, M. Bechouat, M. Sedraoui, C. Napoli, A. Rabehi, D. Lalmi, Hybrid models for direct normal irradiance forecasting: A case study of ghardaia zone (algeria), Natural Hazards 120 (2024) 14703–14725.

[10] S. Russo, I. E. Tibermacine, A. Tibermacine, D. Chebana, A. Nahili, J. Starczewscki, C. Napoli, Analyzing eeg patterns in young adults exposed to different acrophobia levels: a vr study, Frontiers in Human Neuroscience 18 (2024). doi:10.3389/fnhum.2024.1348154.

[11] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative Adversarial Networks, arXiv preprint arXiv:1406.2661 (2014). URL: https://arxiv.org/abs/1406.2661.

[12] A. Tibermacine, W. Guettala, I. E. Tibermacine, Efficient one-stage deep learning for text detection in scene images., Electrotehnica, Electronica, Automatica 72 (2024).

[13] L. B. Martin Arjovsky, Soumith Chintala, GWasserstein GAN, arXiv preprint arXiv:1701.07875 (2017). URL: https://arxiv.org/abs/1701.07875.

[14] C. Napoli, C. Napoli, V. Ponzi, A. Puglisi, S. Russo, I. E. Tibermacine, Exploiting robots as healthcare resources for epidemics management and support caregivers, in: CEUR Workshop Proceedings, volume 3686, 2024, p. 1 – 10.

[15] Y. Peng, C. Kecheng, M. Zhaoqi, Z. Di, P. Xiaorong, R. Yazhou, EEGDnet: Fusing Non-Local and Local Self-Similarity for 1-D EEG Signal Denoising with 2-D Transformer, arXiv preprint arXiv:2109.04235 (2021). URL: https://arxiv.org/pdf/2109.04235.

[16] B. Eoin, R. Peter, F. Andrew, D. Maarten, B. Geraldine, W. Tomás, Denoising EEG Signals for Real-World BCI Applications Using GANs, Frontiers in Neuroscience (2022). URL: https://www.frontiersin.org/journals/neuroergonomics/articles/10.3389/fnrgo.2021.805573/full.

[17] S. Russo, S. Ahmed, I. E. Tibermacine, C. Napoli, Enhancing eeg signal reconstruction in cross-domain adaptation using cyclegan, in: 2024 International Conference on Telecommunications and Intelligent Systems (ICTIS), IEEE, 2024, pp. 1–8.

[18] A. Tibermacine, I. E. Tibermacine, M. Zouai, A. Rabehi, Eeg classification using contrastive learning and riemannian tangent space representations, in: 2024 International Conference on Telecommunications and Intelligent Systems (ICTIS), IEEE, 2024, pp. 1–7.

[19] Y. An, H. K. Lam, S. H. Ling, Auto-Denoising for EEG Signals Using Generative Adversarial Net-

work, Sensors (2022). URL: https://www.mdpi.com/1424-8220/22/5/1750.
[20] I. Naidji, A. Tibermacine, W. Guettala, I. E. Tibermacine, et al., Semi-mind controlled robots based on reinforcement learning for indoor application., in: ICYRIME, 2023, pp. 51–59.
[21] S. Wang, Y. Luo, H. Shen, An improved generative adversarial network for denoising eeg signals of brain-computer interface systems, Proceedings of the 2022 China Automation Congress (CAC) (2022). URL: https://ieeexplore.ieee.org/document/10055145.
[22] B. Nail, B. Djaidir, I. E. Tibermacine, C. Napoli, N. Haidour, R. Abdelaziz, Gas turbine vibration monitoring based on real data and neuro-fuzzy system, Diagnostyka 25 (2024).
[23] S. Bouchelaghem, I. E. Tibermacine, M. Balsi, M. Moroni, C. Napoli, Cross-domain machine learning approaches using hyperspectral imaging for plastics litter detection, in: 2024 IEEE Mediterranean and Middle-East Geoscience and Remote Sensing Symposium (M2GARSS), IEEE, 2024, pp. 36–40.
[24] H. Zhang, M. Zhao, C. Wei, Eegdenoisenet: A benchmark dataset for deep learning solutions of eeg denoising, arXiv preprint arXiv:2009.11662 (2021). URL: (https://arxiv.org/abs/2009.11662.
[25] B. Nail, M. A. Atoussi, S. Saadi, I. E. Tibermacine, C. Napoli, Real-time synchronisation of multiple fractional-order chaotic systems: an application study in secure communication, Fractal and Fractional 8 (2024) 104.
[26] C. Napoli, C. Napoli, V. Ponzi, A. Puglisi, S. Russo, I. E. Tibermacine, Exploiting robots as healthcare resources for epidemics management and support caregivers, in: CEUR Workshop Proceedings, volume 3686, 2024, p. 1 – 10.
[27] Z. Houssam, F. Chuan-Sheng, L. Bruno, M. Gaurav, R. C. Vijay, Efficient GAN-based anomaly detection, arXiv (2018). URL: https://arxiv.org/abs/1802.06222.
[28] Z. Niu, K. Yu, X. Wu, LSTM-Based VAE-GAN for Time-Series Anomaly Detection, Sensors (2020). URL: https://www.mdpi.com/1424-8220/20/13/3738.
[29] Z. Zhijie, L. Wenzhong, D. Wangxiang, Z. Linming, L. Qingning, H. Peng, G. Tong, L. Sanglu, STAD-GAN: Unsupervised Anomaly Detection on Multivariate Time Series with Self-training Generative Adversarial Networks, ACM Transactions on Intelligent Systems and Technology (2023). URL: https://dl.acm.org/doi/10.1145/3572780.
[30] A. Tibermacine, D. Akrour, R. Khamar, I. E. Tibermacine, A. Rabehi, Comparative analysis of svm and cnn classifiers for eeg signal classification in response to different auditory stimuli, in: 2024 International Conference on Telecommunications and Intelligent Systems (ICTIS), IEEE, 2024, pp. 1–8.
[31] P. Rabich, Dataset: Epilepsy2, 2023. URL: (https://www.timeseriesclassification.com/description.php?Dataset=Epilepsy2.